Wear 426–427 (2019) 1761–1770

Contents lists available at ScienceDirect

Wear

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

Integrated model of BP neural network and CNN algorithm for automatic wear debris classification

S. Wang[a], T.H. Wu[a,*], T. Shao[a], Z.X. Peng[b]

[a] Key Laboratory of Education Ministry for Modern Design and Rotor-Bearing System, Xi'an Jiaotong University, China
[b] School of Mechanical and Manufacturing Engineering, the University of New South Wales, Sydney 2052, Australia

ARTICLE INFO

Keywords:
Wear debris
BP neural network
Deep learning
Wear debris classification

ABSTRACT

Mechanism-based wear debris classification (WDC) is important for root cause analysis and prediction of wear related faults. Compared to manual classifications, automatic WDC is more efficient and often more reliable for a wide range of industrial applications. However, existing methods unavoidably encounter some difficulties when dealing with those wear particles with highly geometric similarity, especially for fatigue particles and severe sliding particles. To meet the requirement for automatic WDC, an integrated, automated method for identifying typical wear debris is proposed with a two-level classification procedure. By referring to the traditional ferrography – a widely used wear particle imaging and analysis technique, the first-level classification is performed by a general back-propagation (BP) neural network with selected particle's morphological features. By doing this, three types of wear particles including rubbing, cutting, and spherical particles can be determined. In the second-level classification, a deep learning model of a 6-layer convolution neural network (CNN) is adopted to identify fatigue particles and severe sliding particles by analyzing their very slight surface details in pixel-level. The method is tested with over 100 images of real particles generated from an extruder machine in a petrochemical plant and identified by a ferrograph specialist. A high recognition rate of over 80% is achieved for the three types including rubbing, cutting, and spherical particles with the first procedure. Further, the identification rates are 85.7% and 80% for fatigue particles and severe sliding particles, respectively, which is distinctly improved from the reported values (they are 45.5% and 36.4%, respectively) of other intelligent methods.

1. Introduction

As the direct production of friction interactions between mechanical parts, wear particles carry valuable information for the identification of wear mechanisms and ongoing failures [1]. Therefore, wear debris analysis (WDA) has been the most efficient means for both diagnosis and prognosis of wear failures [2]. Among existing techniques for wear debris imaging and analysis, ferrography is one of the most recognized and effective WDA methods [3]. However, the traditional ferrography has inherent disadvantages, such as experience-dependence and time-consuming, which retard its applications. With the new requirement of condition-based maintenance and assistance of advanced information technologies, increasing interests are stimulated on developing automatic ferrography. Therefore, a comprehensive method for industrial application remains a promising prospect.

Morphological characteristics of wear particles are directly correlated to their generation mechanism and the operation of the equipment. Moreover, different types of wear particles can be identified using descriptors of the particle shape and surface morphology. By referring to analytical ferrography, intelligent algorithms [5–7] were introduced to extract particle morphological and color characteristics. Shape parameters are available to identify regular, irregular, elongated, and spherical particles by the classification and regression tree method [8]. Some texture parameters (energy, entropy, and correlation) are also extracted from the grayscale symbiotic matrix to establish a new classification algorithm by combining the principal component analysis (PCA) with the gray correlation [9]. With color features, the typical classification includes copper alloys, red oxidation particles and dark oxidation particles are distinguished by adopting the clustering algorithm [10]. Although current WDA methodologies have gained some successes in the automation of wear particle identification, their accuracy rates of wear particle recognition need to be improved. In addition, they have difficulty in identifying fatigue particles and severe sliding particles which exhibit similar shapes and edge features in 2D images.

Fatigue particles and severe sliding particles are produced in their corresponding wear processes, and are often treated as the precursor of

* Corresponding author.
E-mail address: wt-h@163.com (T.H. Wu).

https://doi.org/10.1016/j.wear.2018.12.087
Received 3 September 2018; Received in revised form 15 December 2018; Accepted 28 December 2018

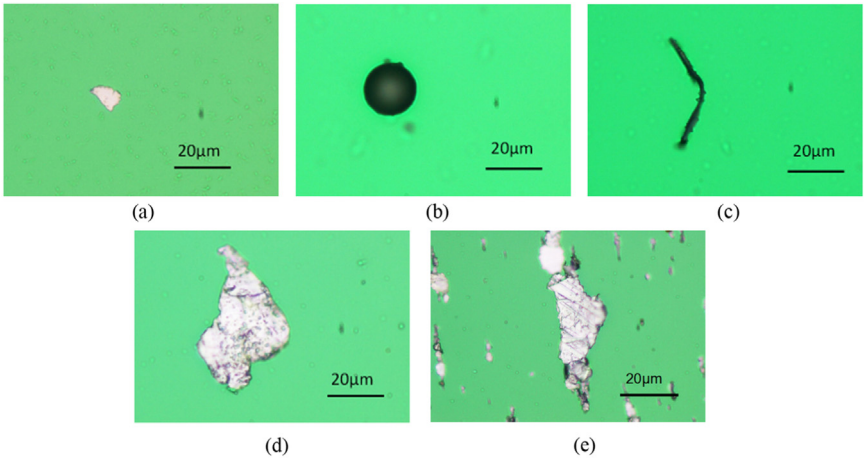

Fig. 1. Five typical wear particles: (a) a rubbing wear particle; (b) a spherical wear particle; (c) a cutting wear particle; (d) a fatigue wear particle; (e) a severe sliding wear particle.

ongoing failures. To identify the two typical wear particles, research efforts have directed towards extracting additional features. Among these attempts, stereo scanning electron microscopy (SEM) is adopted to acquire wear debris contour and surface information [11]. Then these features are inputted into a support vector machine (SVM) system [11] or a two neural networks system [12] to classify the two kinds of wear particles, demonstrating that surface-based classification system is more efficient and accurate than the classification system based on 2D images. However, these advanced techniques cannot be popularized for industrial applications due to the expensive equipment and complex operations.

As mentioned above, traditional particle recognition methods can achieve high accuracy for the particles which have typical characteristic parameters but have difficulty in recognizing the others, especially in fatigue particles and severe sliding particles. With the development of self-learning methods, target images can be directly analyzed without artificial parameters. These network models mainly include LeNet-5 [13], AlexNet [14], GoogleNet [15], VGG [16] and ResNet [17]. However, AlexNet, GoogleNet, VGG and ResNet require too much computation time and hardware configuration to train the network structure. The LeNet-5 model is a smaller Convolutional Neural Network (CNN) structure. The mapping relation between the input and the output is trained without any precise mathematical expressions between the input and the output. In addition, weight sharing can reduce training parameters to simplify the neural network structure and improve its applicability. The LeNet-5 model is the first CNN method that has been successfully applied to deal with digital recognitions. However, particle images are more complicated than digital images in textures. Therefore, further research is required on the CNN-based wear particle identification.

This paper introduces an integrated method with BP neural network and CNN algorithm to identify five typical groups wear particles, including: rubbing, cutting, spherical, fatigue, and severe sliding wear debris. With selected morphological features, BP neural network can easily identify four typical groups of particles, including rubbing, cutting, spherical particles and the unknowns (that is, fatigue particles and severe sliding particles). Fatigue particles and severe sliding particles are then identified by a Convolutional Neural Network (CNN) model, which can avoid the requirement for a large number of samples and long training time. The performance of this innovative methodology will be verified by comparing with results observed from traditional machine learning methods using wear particles that are produced from an extruder machine in a petrochemical plant.

This paper is further organized as follows: Section 2 contains the wear particle identification with BP neural network and CNN algorithm; the verification of the proposed method is given in Section 3; Discussions are presented in Section 4; and conclusions are given in Section 5.

2. Materials and methods

Based on distinctive shape and surface features of five types of wear particles, a hierarchical concept was introduced for wear particle identification. A 2-step approach was used for this development. Firstly, wear particle features were quantitatively characterized using numerical parameters and a BP neural network was developed to identify rubbing, cutting, and spherical particles based on their characterized shape features while non-identified particles were grouped separately. And then, fatigue particles and severe sliding wear particles were identified with a 6-layer convolution neural network. These 2 steps are described as follows.

2.1. Four typical wear particle identification based on BP neural network

2.1.1. Wear particle feature extraction

Five typical wear particles were generated from an extruder machine in a petrochemical plant, and their images were collected using the ferrography as shown in Fig. 1. The correlation between morphological characteristics of these five typical particles and the condition of friction pairs are described in Table 1[4,7,8,11,12]. As can be observed, each typical wear particle has unique features corresponding to its generating mechanism. For example, rubbing wear debris, which is generated from a normal sliding wear process, is a thin slice with a polishing surface and in small size. Therefore, the shape and/or surface features of wear particles can facilitate to identify the types of wear particles.

Numerical parameters were used to characterize the sizes, shapes and surface textures of the five types of wear particles. Rubbing, cutting, and spherical wear particles show obvious shape differences, thus, Area, Roundness and Aspect ratio can characterize these particles [18]. Fatigue particles and severe sliding particles exhibit high similar in their shape and size features, resulting in low recognition accuracy. Thus, the two typical wear particles are grouped into a “unknowns” group in step 1 and will be separated in step 2. Due to unique surface textures, the unknowns can be identified from the other three typical groups by texture features, including Energy, Entropy, Relevance and Inertia, which are calculated by gray level co-occurrence matrix [19]. The definitions of selected morphological parameters are given in Table 2.

A total of 215 wear particles, that is, 43 wear particles per type,

Table 1
Correlation between morphological characteristics and generation mechanisms of wear particles [4,7,8,11,12].

Particle types	Features	Generation mechanism	Wear level	Possible machine operating condition
Rubbing	Thin, plane and oblong shape Size < 50 μm	By rubbing, particles caused from the damaged frictional part in the shear mixing layer	Normal	Normal wear particles without dramatic increase
Cutting	Slender or curved particles	Caused by sharp edges or hard particles penetrating the soft surface of parts	Severe	Impending trouble of severe cutting wear
Spherical	Ball-like, hollow spheres	Generated by welding, grinding and erosion with the effect of temperatures in the fatigue cracks	Abnormal	An onset of early surface pitting or the severe wear under way
Fatigue	Smooth surfaces with pitting	Pitting after surface fatigue of friction pairs	Severe	Heavy load or over-speed
Severe sliding	Parallel scratches or grooves on its surface	Generated from severe sliding wear caused by excessive stress on the surface of friction pairs	Severe	The damage of oil film

Table 2
Definitions of seven morphological parameters [18–20].

Descriptor	Definition
Area	The coverage area of wear particles
Roundness	The proximity degree between the shape of wear particle and its minimum enclosing rectangle
Aspect ratio	The ratio of the major axis and the minor axis
Energy	The uniformity degree of the pixel distribution and texture in the grayscale image
Entropy	The complexity of textures on wear particle surfaces
Relevance	The similarity degree of elements in gray level co-occurrence matrix from row or column direction.
Inertia	The clarity degree of textures on wear particle surfaces

Table 3
The average value of seven parameters of four typical wear particles.

Parameter	Wear particle type			
	Rubbing	Cutting	Spherical	Unknowns
Area (μm^2)	52.1759	241.7771	127.2898	495.9439
Roundness	0.6874	0.0912	0.9165	0.6470
Aspect ratio	1.4341	10.2587	1.0601	1.6142
Energy	0.3798	0.5896	0.1762	0.2221
Entropy	1.6278	1.2297	2.2171	2.3378
Inertia	0.5245	0.9733	2.1836	0.5212
Relevance	0.8970	0.1044	0.0401	0.2347

were selected to develop the particle classification system. The average values of these wear particles are listed in Table 3. As can be observed, the selected parameters can characterize the differences in these wear particles. For example, the area of rubbing particles is smaller than the others and the aspect ratio of cutting particles is the largest.

2.1.2. BP neural network model

In BP neural network, the complex nonlinear relationship and the simple learning process can ensure that four typical groups of particles can be effectively identified with distinctive parameters. BP neural network mainly includes an input layer, a hidden layer and an output layer, which are connected by weight values, as shown in Fig. 2. And its basic principle is that the network is trained with labeled samples until the difference between the output value and the expected value reached the set threshold value.

According to the Kolmogorov theory [21], when the number of neurons in the hidden layer meets the condition: $D \geq 2M + 1$, where D is the number of neurons in the hidden layer and M is the number of input nerves, the three-layer neural network can accurately realize any

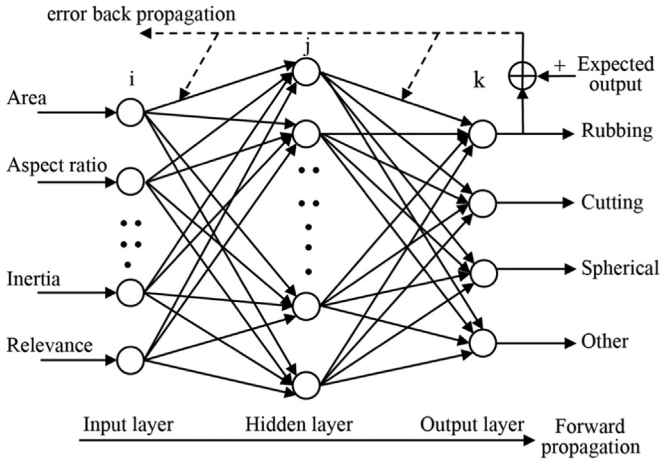

Fig. 2. Wear particle classification model based on BP neural network.

continuous mapping. Therefore, a nonlinear mapping relationship from the input (wear particle parameters) to the output (wear particle types) can be established through appropriately setting the number of hidden layer neurons [22]. The specific procedures are described as following:

(1) Initialize connection weight values and thresholds with random values.
(2) Calculate the output of each unit in the hidden layer and the output layer based on selected parameters in the input and the output mode.
(3) Calculate new connection weights and thresholds with equation (1)–(4).
The modification of the neuron threshold:

$$\theta_k(t+1) = \theta_k(t) + \eta_t \sigma_k, (t = 1, 2, ...,p; k = 1, 2, ...,l) \quad (1)$$

$$\beta_j(t+1) = \beta_j(t) + \eta_t \sigma_j, (t = 1, 2, ...,p; j = 1, 2, ...,m) \quad (2)$$

where, η_t is the learning rate of the t-th training iteration; σ_k is the error of the k-th node in the output layer; θ_k is the threshold value of the k-th neuron in the output layer, σ_j is the error of the j-th node in the hidden layer; β_j is the threshold value of the j-th neuron in the hidden layer.
The weight value updating formula:

$$w_{jk}(t+1) = w_{jk}(t) + \Delta w_{jk}(t), (t = 1, 2, ...,p) \quad (3)$$

$$v_{ij}(t+1) = v_{ij}(t) + \Delta v_{ij}(t), (t = 1, 2, ...,p) \quad (4)$$

where, t is the number of training iteration, w_{ij} is the weight value updating from the hidden layer to the output layer, v_{ij} is the weight value updating from the input layer to the hidden layer.
(4) Update the learning input mode and re-execute the second step to train the neural network until the final output error reached the set value.

Using the above procedures, the mean values of trained particles in Table 3 were input into the BP Neural network. And the number of nodes in the hidden layer was adjusted according to the selected input parameters. The output was divided into four types, including: rubbing, cutting, spherical wear particle and unknowns. The output vector was calculated with the BP neural network, and particle types were identified by referring to Table 4.

2.2. Fatigue particle and severe sliding particle recognition based on CNN

Severe sliding particles always have obvious scratches or grooves on the surface while fatigue particles have smooth surfaces with pitting. However, there is a low identification accuracy of fatigue particles and severe sliding particles due to the fact that traditional methods have difficulty in manually extracting the distinctive surface textures. To address this problem, a recognition model based on self-learning features was developed to identify the two types of wear particles. To achieve the objective, two steps were involved including (1) wear particle image preprocessing, and (2) development of a CNN-based wear particle recognition model.

Table 4
The output table of on BP Neural network.

Output	Output 01	Output 02	Output 03	Output 04
Rubbing	1	0	0	0
Spherical	0	1	0	0
Cutting	0	0	1	0
Unknowns	0	0	0	1

2.2.1. Wear particle image preprocessing

Various types of wear particles may co-exist in an image, which will cause a poor classification if the image is directly taken as the input. In addition, a large number of particle images in the same type are required as training samples for CNN. Therefore, it is necessary to preprocess original wear particle images to obtain enough typical particle images, and the procedure is described as follows.

Fatigue particles and severe sliding wear particles were firstly extracted from original images based on the maximum between-class variance method (OTSU) [23]. Examples of segmented image samples are shown in Fig. 3. It is noted that the excessive size of the original particle image (1600 × 1200 pixels) results in too much network structural parameters, but too small size is not conducive to extract sufficient features from wear particle images. Therefore, the size of the wear particle images was then collectively normalized to 240 × 240 pixels to ensure the consistent size of the input vector into CNN. Finally, due to a small amount of typical wear particles in actual enterprises, existing wear particle images were expanded from one to five by randomly flipping, shifting, cutting, and amplifying [24] to meet the requirements of CNN for a large amount of images. The expanded images of a single wear particle are shown in Fig. 4.

2.2.2. The CNN-based wear particle recognition model

Compared with the BP neural network, CNN can directly analyze target image without inputting artificial parameters [25]. Furthermore, the neurons between layers are partially connected instead of full connection, and the number of training parameters can be continuously optimized by sharing the weights and the sparse connection in the network [26,27]. The LeNet-5 model is the most widely adopted CNN model in the field of computer vision. The model has 7 layers and its framework is shown in Fig. 5.

Based on the initial framework in Fig. 5, improvements were made so the developed model can accommodate the size, color and complexity of wear particles as well as higher efficiency with less required sample quantity for training. Then, a CNN-based wear particle recognition network was constructed. These 2 steps are detailed below.

2.2.2.1. Improvement on the LeNet-5 model. The original LeNet-5 network is designed for identifying digit images (eg, the image of digit 0) which are relatively simple 32 × 32-pixel images. In this study, the size of a particle images is 240 × 240 pixels and the textures are more complicated than digit images. Meanwhile, the output type that needs to be identified is also different. Therefore, some improvements need to be made on the parameters of each layer in the original LeNet-5 network for identifying wear particles. The main improvements are listed as follows.

(1) Rectified Linear Unit (ReLU) activation function [28] was adopted to replace traditional Tanh and sigmoid activation functions. Compared with other activation functions, ReLU activation function has high efficiency gradient descent and reverse propagation, which can reduce the training time and avoid the escalator explosion.
(2) In order to improve training efficiency and reduce overfitting, dropout layer was added in the first fully connected layer for randomly abandoning some neurons.
(3) Fatigue particles and severe sliding wear particles need to be identified by CNN. Thus, the number of output was changed from 10 to 2. And, a full link layer F6 was removed to simplify the network.
(4) In view of the complexity of wear particle images, the numbers of convolution kernel in first convolution layer and second convolution layer were increased from 6 and 16 to 32 and 64, respectively, to extract more features from wear particle images.
(5) With the increase of the input image size and convolution kernel number, the number of neurons in CNN was increased correspondingly. The neuron node number of C5 layer was increased

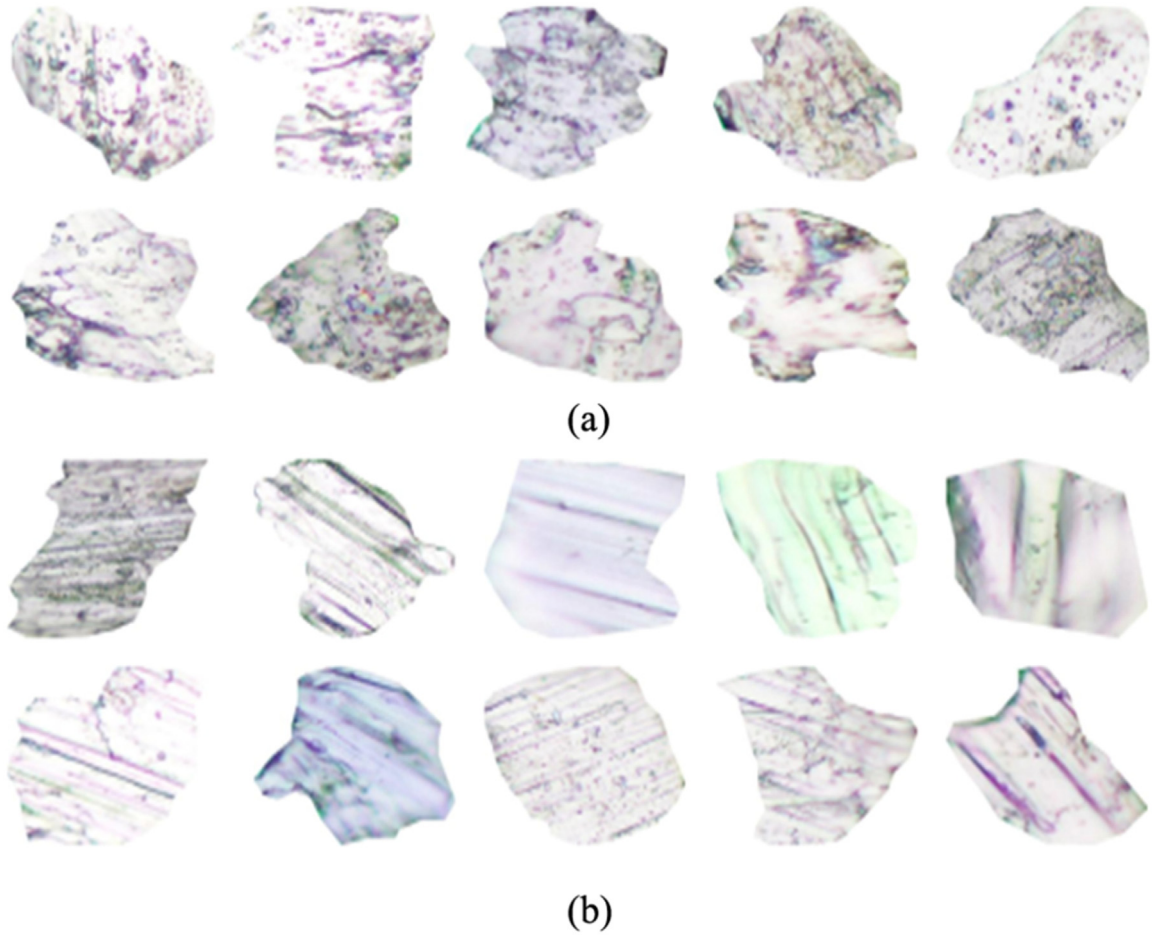

Fig. 3. Segmented image samples: (a) fatigue wear particles; (b) severe sliding wear particles.

from 120 to 1024 to ensure the recognition performance of CNN.

(6) Too many simultaneously trained images may be result in memory overflow, slow convergence and local optimum. Thus, the parameter Batch_size, representing the number of simultaneously trained images, was added to limit the number of simultaneously trained images.

2.2.2.2. Construction of CNN-based network.

(1) CNN-based network structure

Based on the improvements made on the existing LeNet-5 network, a CNN-based wear particle recognition network was constructed and is shown in Fig. 6. The input is single wear particle image with the size of 240 × 240 pixels. In the first convolution layer, 32 convolution kernels with a size of 5 × 5 pixels and 1-pixel step length were adopted to output 32 feature images with a size of 236 × 236 pixels by point multiplication processing. Then the dimension of 32 feature images was reduced to the size of 118 × 118 pixels by the maximum pool layer through a filter with the size of 2 × 2 pixels and 2-pixel step length. In the second convolution layer, 64 convolution kernels with a size of 5 × 5 pixels and 1-pixel step length were adopted to output 64 feature images with a size of 114 × 114 pixels from by point multiplication processing. And then the dimension of 64 feature images was reduced to a size of 57 × 57 pixels by the maximum pool layer through a filter with the size of 2 × 2 pixels and 2-pixel step length. Finally, in the first full connection layer, 1024 neurons were adopted to connect all pixels of 64 feature images in the second layer. The output of the second full connection layer is the types of wear particles - fatigue particles and severe sliding wear particles. In the CNN-based wear particle recognition model, the ReLU activation function was applied in the convolution layer and all connection layers.

Fig. 4. Expanded images of a single wear particle.

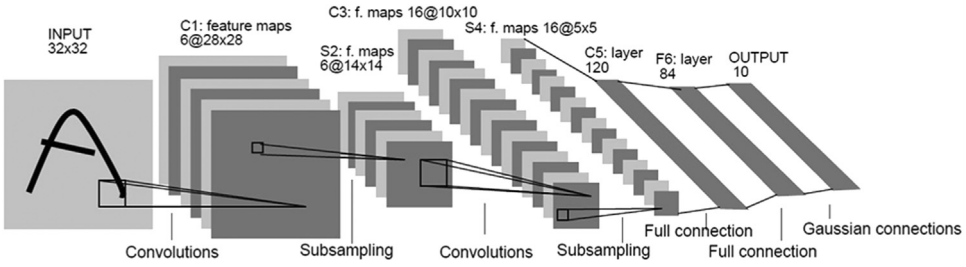

Fig. 5. The framework of the LeNet-5 model.

(2) Loss function selection

In the neural network, the multi-class problem is addressed by outputting an n-dimensional array (n is the class number), in which each output node corresponds to one type. Ideally, when a sample is identified to a certain type, the corresponding node should be assigned to be 1 and the other nodes are 0. However, the actual situation can not satisfy this assumption due to various noises on collected particle images. Thus a loss function was applied to evaluate the proximity between the output vector and the expected vector. The cross entropy evaluates the proximity by characterizing the distance between two probability distributions, which is one of the most commonly used methods in classifications [29]. The cross entropy was calculated using Eq. (5).

$$H(p, q) = -\sum_{x} p(x)\log q(x) \tag{5}$$

Where, $p(x)$ is the expected value; $q(x)$ represents the predicted value.

(3) The training process

In the CNN-based wear particle recognition network, the relationship between the input (wear particle images) and the output (wear particle types) was trained using a large amount of sample images. The training processes included forward propagation and backward propagation. In the forward propagation, wear particle samples were inputted into the neural network and the features of the fatigue particles and severe sliding wear particles were automatically calculated to obtain the output value. In the backward propagation, the cross entropy was applied to calculate the error between the output value and the expected value.

In addition, weight parameters in the network need to be adjusted by the error return algorithm. A number of error return algorithms, including the batch gradient descent (BGD) [30], the stochastic gradient descent (SGD) [31] and the adaptive gradient descent (AGD) [32], were considered. BGD requires too much learning time and hardware configuration to update model parameters in real time. SGD may not update parameters in the correct direction which will result in the optimization fluctuation. In contrast, AGD can establish an adaptive learning rate for each parameter to improve the robustness of CNN. Sparse features will get a large learning rate, and non-sparse features will get a small learning rate. Therefore, AGD was adopted as the error return algorithm in the CNN-based wear particle identification network. Based on the error return algorithm, the training process was repeated again and again until the number of iterations reached the set value.

3. Method verification and performance evaluation

As mentioned in Section 2, based on the hierarchical concept, five typical groups of wear particles can be identified by integrating BP neural network and CNN algorithm. In particular, the developed particle classification system is able to identify fatigue particles and severe sliding particles, which have similar shape features. In this section, a set of experiments were performed to demonstrate the performance of the proposed method.

3.1. Four types of particle classification based on BP neural network

As stated in Section 2, 215 wear particles were selected from the sample base as training samples. To verify the performance of BP neural network, the remaining 52 samples were used as tested samples, including 11 fatigue particles, 11 severe sliding particles, 10 rubbing particles, 10 cutting particles and 10 spherical particles. Part of tested particle images is shown in Fig. 7. Fatigue particles and severe sliding wear particles were grouped into the same class, named as “unknowns”. Extracted features were inputted into the BP-based classifier and the results of final classification are shown in Table 5. It can be seen that the identification of four typical particles has a high accuracy of over 80%, demonstrating that selected seven features can character four typical groups of wear particles.

In comparison, Table 6 shows the recognition results of wear particle through the seven selected features and same BP neural network except that the output was changed from four types to five types. Although the recognition accuracy of rubbing, spherical, and cutting particles is the same, the recognition of fatigue particles and severe sliding particles has a low accuracy (lower than 46%). The main reason is that fatigue particles and severe sliding wear particles have highly

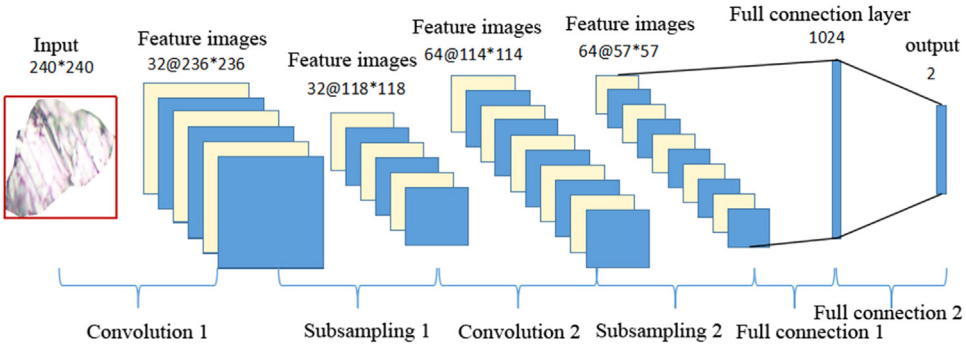

Fig. 6. The diagram of the developed CNN-based wear particle recognition model.

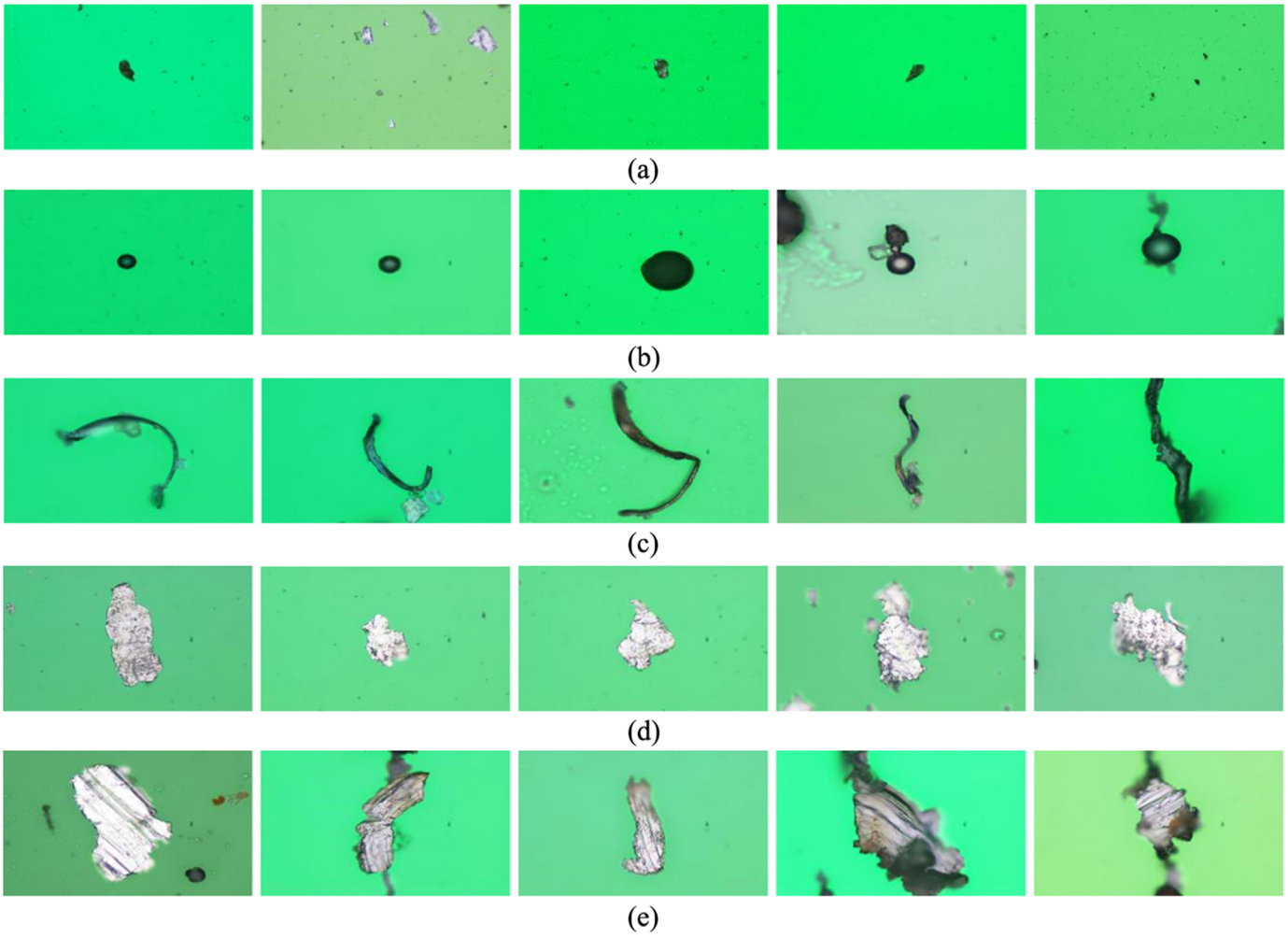

Fig. 7. Images of test samples to train the BP neural network: (a) rubbing wear particles; (b) spherical wear particles; (c) cutting wear particles; (d) fatigue wear particles; (e) severe sliding wear particles.

Table 5
The classification results of wear particles based on BP neural network.

Wear particle type	Correct recognition	Incorrect recognition	Recognition rate
Rubbing	8	2	80%
Spherical	9	1	90%
Cutting	10	0	100%
Unknowns	22	0	100%

Table 6
The classification result of wear particles based on BP neural network.

Wear particle type	Correct recognition	Incorrect recognition	Recognition rate
Rubbing	8	2	80%
Spherical	9	1	90%
Cutting	10	0	100%
Fatigue	5	6	45.5%
Severe sliding	4	7	36.4%

similar shape and surface texture features, and features extracted by artificial methods for have limitation.

3.2. Fatigue particle and severe sliding particle recognition based on CNN

In the experiment, 108 wear particles were selected from an extruder machine in the petrochemical plant, including 54 fatigue particles and 54 severe sliding particles. The size of segmented particle images was collectively normalized to 240 × 240 pixels. In addition, wear particle images were expanded from one image to five images by the data amplification method. Therefore, the original particle image database was expanded from 108 images to 540 images, including 270 fatigue particles and 270 severe sliding wear particles. 35 fatigue particles and 35 severe sliding particles were selected as the testing particles while the others were served as the training samples. Part of tested particle images is shown in Fig. 8.

The CNN-based wear particle identification model was trained using the AGD method. The training parameter (Batch_size) is 30, the learning rate is 0.0004, and the iteration number is 300. In order to verify the performance, the improved LeNet-5 network model was compared with the original networks in the training process of wear particle classification, and the results are shown in Fig. 9.

As can be observed from Fig. 9(a), the improved LeNet-5 network has achieved good results in the training process. The network starts to converge nearly at 75-th time, and the training accuracy tends to be nearly 100% when it is iterated to 200-th time.

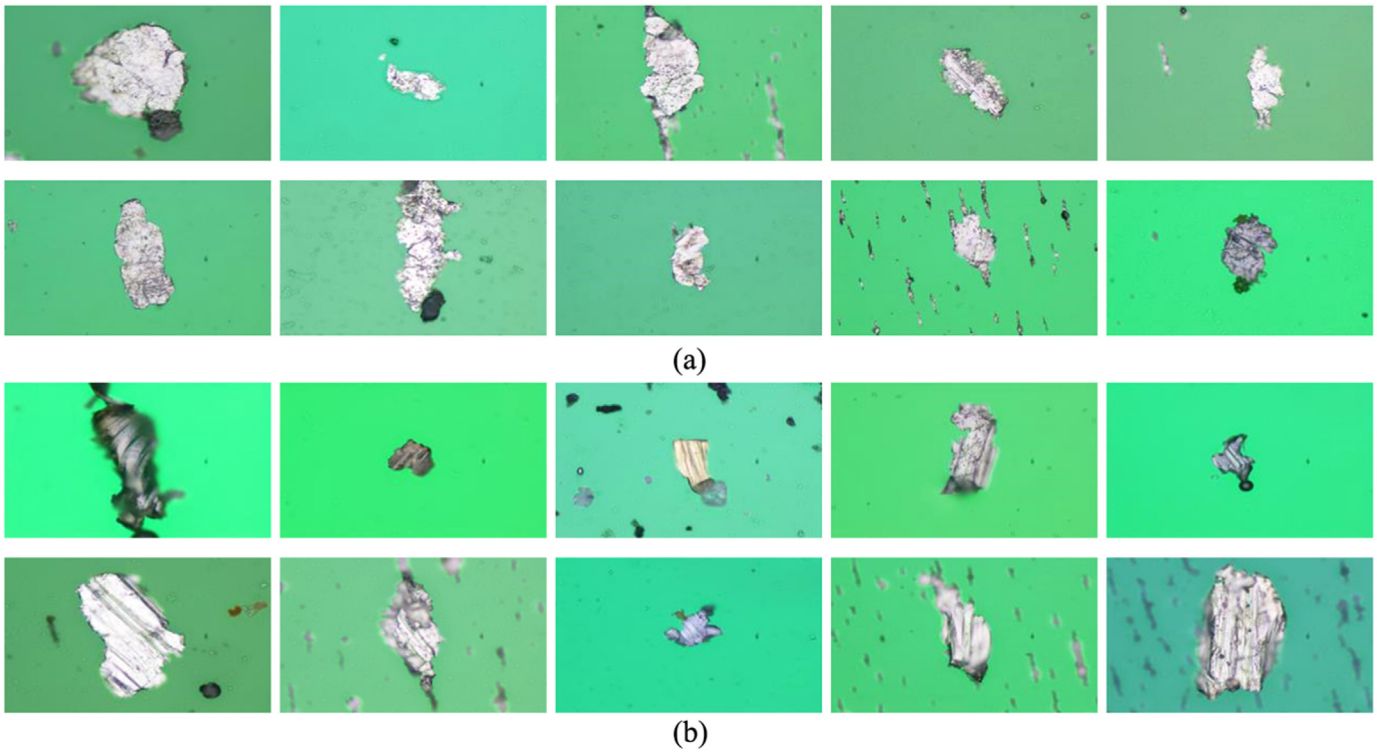

Fig. 8. Test samples for the CNN particle identification model: (a) fatigue wear particles; (b) severe sliding wear particles.

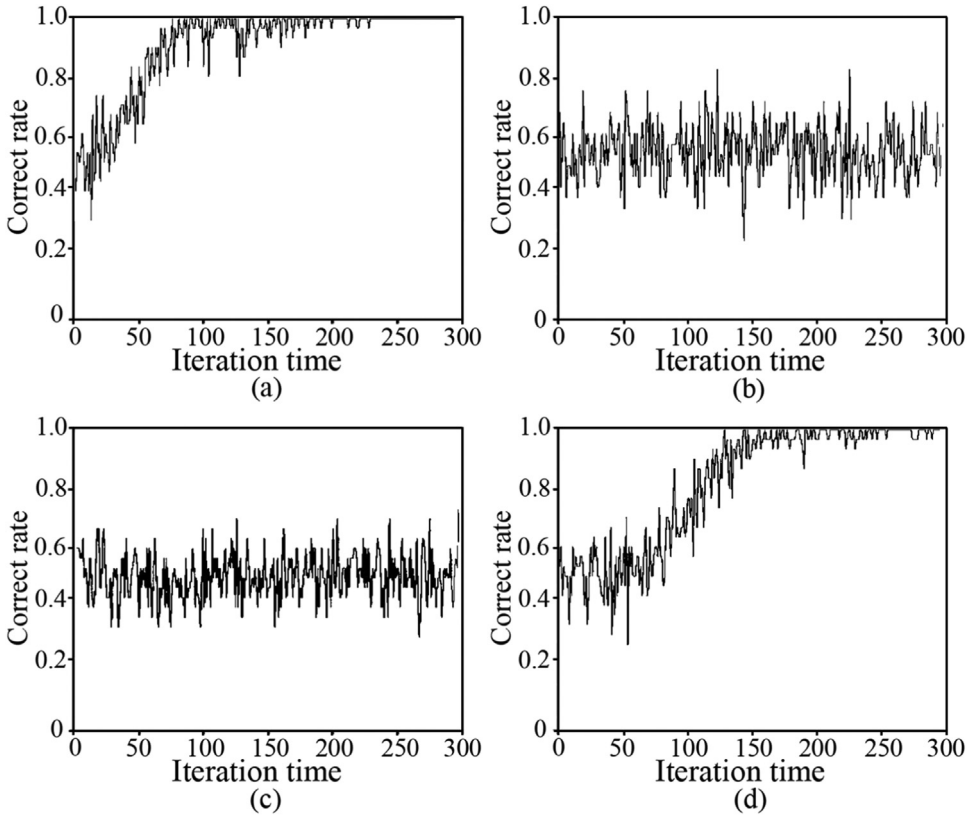

Fig. 9. The correction rates of different networks: (a) the improved LeNet-5 network; (b) the original LeNet-5 network with sigmoid activation function; (c) the original LeNet-5 network with all neurons and layers; (d) the original LeNet-5 network with all connected layer.

Table 7
Wear particle classification based on the CNN.

Wear particle type	Correct recognition	Incorrect recognition	Recognition rate
Fatigue particles	28	7	80%
Severe sliding particles	30	5	85.7%

Compared Fig. 9(b) with Fig. 9(a), the original LeNet-5 network with the sigmoid activation function failed to converge in the training process. The reason for the gradient divergence is that a large amount of images were simultaneously trained and various parameters co-existed in the network. However, the improved network adopts ReLU to effectively avoid this problem.

Compared Fig. 9(c) with Fig. 9(a), the original LeNet-5 network, which preserves all neurons and layers, failed to converge. The reason is that the images of fatigue particles and severe sliding particles have more complex textures than digit images. Therefore, more convolution kernels and neurons on the fully connection layer are required to extract more complex features.

Compared Fig. 9(d) with Fig. 9(a), the convergence rate of the original LeNet-5 network structure, which retains F6 layer, is slow in the training process. The algorithm converges at 120-th iteration, and the training accuracy is stabilized when the training is iterated to about 250-th time. Therefore, the improved network structure is more suitable for practical applications, which can effectively improve the rate of training convergence.

Tested wear particle images were classified using the improved CNN model which has been fully trained to converge by trained samples. The results are shown in Table 7. The identification rates of fatigue particles and sliding particles are 85.7% and 80%, respectively, which are distinctly improved from the values (they are 45.5% and 36.4%, respectively) with BP neural network.

4. Discussions

A two-level strategy, based on a BP neural network and a 6-layer CNN model, is adopted to identify five types of wear particles; including rubbing, cutting, spherical, fatigue and severe sliding particles. The above experiments clearly confirm that the improved integrated model is able to identify typical wear particles with a high accuracy. In comparison to traditional intelligent algorithms [8,9], the recognition rate of severe sliding and fatigue wear particles is obviously improved. When compared with 3D imaging techniques reported in [11,12], these two methods can identify severe sliding and fatigue particles. However, the equipment such as SEM used in 3D imaging is more expensive and operationally complex than the optical microscope used in the improved integrated model. Thus, the developed system is regarded as an effective approach for wear particle identification in enterprises.

However, it should be mentioned that the proposed method relies on manual collection of particle samples by professional ferrography operators. Non-ideal image acquisition will miss surface textures due to oil stain or defocus blurring on wear particle surfaces. Furthermore, more particle samples need to be captured and the CNN trained to improve the recognition accuracy of the recognition model. With the application of the developed method, more and more particles will be collected and identified to expand the sample base, thus the recognition accuracy will be further improved.

5. Conclusions

In order to improve the accuracy of wear particle automatic recognition, a classification method was developed by integrating BP neural network and CNN algorithm. Main conclusions are drawn as follows:

1) A BP neural network based classification algorithm has been developed by combining with seven features selected from wear particle shape, size and texture. Rubbing, cutting, and spherical particles are identified with an average success rate of about 90%. Fatigue particles and severe sliding particles are separated for further identification.
2) In view of the limitation of the traditional machine learning method in identifying the fatigue and the severe sliding wear particles, a CNN-based wear particle recognition network is established by optimizing the original LeNet-5 network structure to automatically identify the two types of wear particles. The average recognition accuracy of fatigue particles and severe sliding particles is about 83%, which is far higher than 41% identified by the BP neural network.

Based on the hierarchical pattern recognition, the classifier can effectively identify five types of wear particles generated from the friction pair during the wear process.

Acknowledgements

This work is funded by the National Natural Science Foundation of China (No. 51675403) and the International Collaborative Plan of Shaanxi Province, China (No. 2017kw-034). Special thanks to the financial support from the NSFC-Zhejiang Joint Fund for the Integration of Industrialization and Informatization, China (Grant No. U1709215). The author gratefully acknowledges the support of K. C. Wang Education Foundation is a financial organization in China.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.wear.2018.12.087.

References

[1] M. Niemczewska-Wójcik, Wear mechanisms and surface topography of artificial hip joint components at the subsequent stages of tribological tests, Measurement 107 (2017) 89–98.
[2] S. Wang, T.H. Wu, H.K. Wu, et al., Modeling wear state evolution using real time wear debris features, Tribol. Trans. 60 (6) (2017) 1022–1032.
[3] M. Kumar, P.S. Mukherjee, N.M. Misra, Advancement and current status of wear debris analysis for machine condition monitoring: a review, Ind. Lubr. Tribol. 65 (1) (2013) 3–11.
[4] S. Raadnui, Wear particle analysis — utilization of quantitative computer image analysis: a review, Tribol. Int. 38 (10) (2005) 871–878.
[5] Q.A. Memon, M.S. Laghari, Self organizing analysis platform for wear particle, Int. J. Comput. Electr. Autom. Control Inf. Eng. 1 (6) (2007) 1773–1776.
[6] P. Podsiadlo, G.W. Stachowiak, Scale-invariant analysis of wear particle surface morphology I: theoretical background, computer implementation and technique testing, Wear 242 (1–2) (2000) 160–179.
[7] J.Q. Wang, X.L. Wang, A wear particle identification method by combining principal component analysis and grey relational analysis, Wear 304 (1–2) (2013) 96–102.
[8] W. Yuan, K.S. Chin, M. Hua, et al., Shape classification of wear particles by image boundary analysis using machine learning algorithms, Mech. Syst. Signal Process. s 72–73 (2016) 346–358.
[9] J.Q. Wang, J. Bi, L. Wang, et al., A non-reference evaluation method for edge detection of wear particles in ferrograph images, Mech. Syst. Signal Process. 100 (2018) 863–876.
[10] N.K. Myshkin, H. Kong, A.Y. Grigoriev, et al., The use of color in wear debris analysis, Wear 251 (1–12) (2001) 1218–1226.
[11] G.P. Stachowiak, G.W. Stachowiak, P. Podsiadlo, Automated classification of wear particles based on their surface texture and shape features, Tribol. Int. 41 (1) (2008) 34–43.
[12] Z. Peng, T.B. Kirk, Computer image analysis of wear particles in three-dimensions for machine condition monitoring, Wear 223 (1–2) (1998) 157–166.
[13] Y. Lécun, L. Bottou, Y. Bengio, et al., Gradient-based learning applied to document recognition, Proc. IEEE 86 (11) (1998) 2278–2324.
[14] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, in: Proceedings of the International Conference on Neural Information Processing Systems, Curran Associates Inc., 2012, pp.

1097–1105.
[15] C. Szegedy, W. Liu, Y. Jia, et al., Going deeper with convolutions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 2015, pp. 1–9.
[16] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556, 2014.
[17] K. He, X. Zhang, S. Ren, et al., Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 2016, pp. 770–778.
[18] Z. Peng, T.B. Kirk, Automatic wear-particle classification using neural networks, Tribol. Lett. 5 (4) (1998) 249–257.
[19] J. Wang, X. Wang, A wear particle identification method by combining principal component analysis and grey relational analysis, Wear 304 (1–2) (2013) 96–102.
[20] T. Wu, Y. Peng, S. Wang, et al., Morphological feature extraction based on multi-view images for wear debris analysis in on-line fluid monitoring, Tribol. Trans. 60 (3) (2016) 408–418.
[21] R. Hecht-Nielsen, Theory of the backpropagation neural network, Neural Netw. 1 (1) (1988) (445-445).
[22] Y.R. Ding, Y.J. Cai, P.D. Sun, et al., The use of combined neural networks and genetic algorithms for prediction of river water quality, J. Appl. Res. Technol. 12 (3) (2014) 493–499.
[23] N. Otsu, Threshold selection method from gray-level histograms, IEEE Trans. Syst. Man Cybern. 9 (1) (1979) 62–66.
[24] B. Leng, K. Yu, Y. Liu, et al., Data augmentation for unbalanced face recognition training sets, Neurocomputing 235 (2017) 10–14.
[25] E.A. Smirnov, D.M. Timoshenko, S.N. Andrianov, Comparison of regularization methods for imagenet classification with deep convolutional neural networks, Aasri Procedia 6 (1) (2014) 89–94.
[26] B.I. Cirstea, L. Likformansulem, Improving a deep convolutional neural network architecture for character recognition, Electron. Imaging 17 (2016) 1–7.
[27] L.I. Song, Z. Wei, B. Zhang, et al., Target recognition using the transfer learning-based deep convolutional neural networks for SAR images, J. Univ. Chin. Acad. Sci. 35 (1) (2018) 75–83.
[28] K. He, X. Zhang, S. Ren, et al., Delving deep into rectifiers: surpassing human-level performance on ImageNet classification, in: Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 2015, pp. 1026–1034.
[29] P.T.D. Boer, D.P. Kroese, S. Mannor, et al., A tutorial on the cross-entropy method, Ann. Oper. Res. 134 (1) (2005) 19–67.
[30] D.R. Wilson, T.R. Martinez, The general inefficiency of batch training for gradient descent learning, Neural Netw. Off. J. Int. Neural Netw. Soc. 16 (10) (2003) 1429–1451.
[31] B. Dong, D.Q. Ren, X. Zhang, Stochastic parallel gradient descent based adaptive optics used for a high contrast imaging coronagraph, Res. Astron. Astrophys. 11 (8) (2011) 997–1002.
[32] John Duchi, Elad Hazan, Yoram Singer, Adaptive subgradient methods adaptive subgradient methods for online learning and stochastic optimization, J. Mach. Learn. Res. 12 (7) (2011) 257–269.