
Research on Rolling Bearing Fault Diagnosis Based on DRS Frequency Spectrum Image and Deep Learning

Zhuoxian Li, Hao Wang, Jiatai Chen, Zhixin Zhou and Wei Chen

College of Energy and Mechanical Engineering, Shanghai University of Electric Power, Shanghai 201306, China.
E-mail: whouco@163.com

(Received 6 January 2023; accepted 17 April 2023)

Deep learning is gradually being widely used in fault diagnosis now, because deep learning networks are more advantageous in processing data, especially image data. However, research using frequency spectra image of fault signals as inputs to deep learning networks are extremely rare in the field of fault diagnosis. Therefore, a brand-new intelligent fault diagnosis method is proposed in this paper which combines discrete random separate (DRS) frequency spectrum images with deep learning networks (DRSFSL-DL). To investigate the fault diagnosis effects of the method mentioned above, several deep learning networks are utilized for comparisons, such as GoogLeNet, residual network, and Inception_ResNet_v2. The vibration fault frequency spectrum images processed by the DRS method are input to train several deep learning networks. Under the same circumstance of deep learning networks, the fault diagnosis using the DRS frequency spectrum image (DRSFSL), is also compared to the fault diagnosis using traditional frequency spectrum, including the power spectrum density (PSD) and cepstrum. The fault diagnosis results show that the proposed method has a better classification accuracy than the PSD image and cepstrum image, with the same deep learning networks. The fault diagnosis accuracy can reach up to 100.00% for some deep learning networks with better generalization performance than the PSD image and cepstrum image. Lastly, the proposed method is further verified using the brand-new bearing fault dataset, and excellent accuracy and generalization ability are achieved.

1. INTRODUCTION

Fault diagnosis of operating machinery and equipment is an essential part of safe industrial manufacturing, and the ability to accurately diagnose faults directly affects the safety of modern machinery and even life safety. However, the time domain signals collected from industrial equipment with faults are usually uneven, containing a large amount of noise and various frequency signals, making it challenging to analyze and diagnose the faults directly.

Frequency spectrum analysis is one of the most commonly used traditional fault diagnosis methods. The typical frequency spectrum analysis methods currently include PSD and cepstrum,¹ which are very critical to traditional fault diagnosis. The power spectrum represents signal power variation with frequency. To realize fault diagnosis during rotor vibration, Cheng-Wei Fei et al.² proposed a PPSE-SVM fault diagnosis model based on Process Power Spectrum Entropy(PPSE), using PPSE to extract the PPSE values of the fault data as fault feature vectors, and verified experimentally that the method had a good generalization capability. Cepstrum analysis is a secondary analysis technique resulting from Fourier inversion of the logarithmic values of the power spectrum. Xining Zhang et al.³ proposed an improved local cepstrum based on the cepstrum and local cepstrum, the noise in the vibration signal and the influence of non-harmonic components on the main signal are reduced, and accurate detection of early faults was effectively made in gearboxes and rolling bearings.

Using vibration images as network input for a fault diagnosis is also a common method in engineering as follows: 1) the vibration image includes a time-domain image; 2) frequency-domain images; and 3) a time-frequency-domain image. Among these three images, the time domain and time-frequency-domain images of vibration signals are more common,⁴ especially for the time-frequency domain images. Traditional machine learning methods have significant drawbacks in fault diagnoses, such as the need for a large number of professionals to participate in the feature extraction work and the inability of the classifier to extract features of higher dimensionality.⁵⁻⁷ In this context, deep learning methods have emerged with the ability to overcome the problems encountered in the machine learning process.⁷ Deep learning networks have powerful feature-learning capabilities and can automatically learn multidimensional features from the original data allowing more useful information to be extracted. Through model training, deep learning can automatically select more typical feature types based on the training data, which helps to make accurate predictions in the subsequent pattern recognition steps.⁸ Although many deep learning-based fault diagnosis methods already exist and have achieved good results, most of these diagnosis methods use time-frequency-domain images as network training inputs.⁹⁻¹¹

Zhang R et al.¹² selected some fault data of the Case Western Reserve University (CWRU) bearing dataset and used the time-domain image of these data as the training input to the neural network. Zhang A et al.¹³ converted the collected

time-domain data into two-dimensional images by regularizing them as input to a deep recurrent network. Cao P et al.¹⁴ converted the one-dimensional time series into two-dimensional grayscale images and then input to the deep convolutional neural network (DCNN) for training. In addition, there are also many methods using time-frequency domain images as neural network inputs.¹⁵ Shao S et al.⁵ used the continuous wavelet transform of the fault signals as the input of the convolutional neural network (CNN) and achieved good results. Du Y et al.¹⁶ used the time-frequency domain image obtained from the short-time Fourier transform (STFT) as the training input of the deep residual network and verified the effectiveness of the method. Yu H et al.¹⁷ proposed a fault diagnosis model based on the deep belief network (DBN) used to predict the residual useful life of a hydraulic pump.

Existing research using the frequency domain image or the frequency spectrum image is based on traditional frequency domain transform methods, such as the Fourier Transform (FT), Hilbert Huang Transform (HHT), etc. A deep transfer learning online diagnosis method was proposed by Mao W. et al.,¹⁸ aiming to achieve real-time detection of bearing incipient faults. This diagnosis model is constructed based on the HHT as the input to the vgg16 network, and finally, the effectiveness of the method was verified by the PHM dataset. To achieve the remaining useful life prediction of rolling bearings, Mao W. et al.¹⁹ proposed a new RUL prediction method based on deep learning and transfer learning, using HHT as input data, and got good results on the PHM dataset. However, deep learning fault diagnosis methods that use other frequency spectrum images as the input of deep learning networks are very rare. The discrete random separate (DRS) method²⁰ is a frequency domain transform method. Its core idea is to decompose the fault signal into many small segments and obtain the transfer function to separate the deterministic and random components of the fault signal.

To improve the fault diagnosis efficiency of frequency domain image combined with deep learning, based on CWRU bearing datasets, the DRS frequency spectrum image will be used as the input of several common deep learning networks. As comparisons, the PSD image and cepstrum image of the fault signals will be applied as the inputs of deep learning networks as well.

The paper is arranged as follows: the theoretical basis, the frame of the DRSFSI-DL method, the fault diagnosis dataset, the fault diagnosis based on the DRSFSI-DL, the generalization of the DRSFSI-DL fault diagnosis method, the validation of the DRSFSI-DL method and conclusions at the end.

2. THE THEORETICAL BASIS

2.1. DRS Method

The DRS method can be used to separate the random component of the signal from the deterministic components. The core idea of the DRS method is to find the transfer function between the original signal and the delayed signal after a certain time delay. The principle is as follows: an original signal is separated into many segments, and the time interval of each

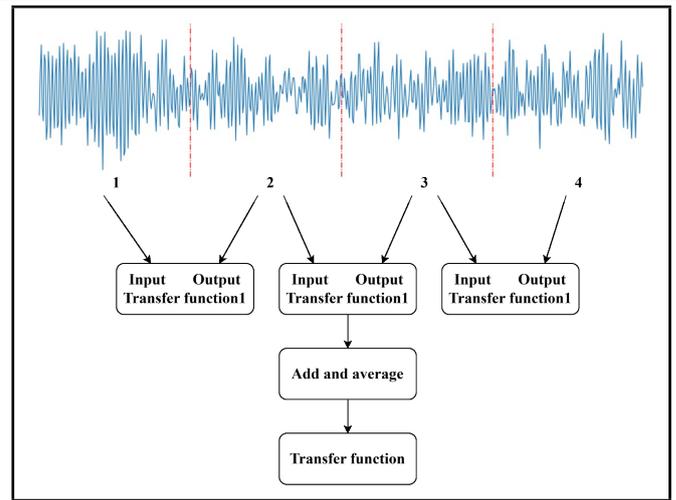


Figure 1. The schematic of the DRS transfer function calculation.

piece is the same. If there is a deterministic component in part and its delayed signal, then these two segments show a high correlation. On the contrary, they exhibit a low correlation, and the transfer function reflects the level of correlation. The averaging transfer function solution process is shown in Fig. (1), where the input signal is non-delayed, and the output signal is delayed,²¹ for example, segment 2 is the delayed signal and segment 1 is non-delayed.

As shown in Fig. 1, the average value of the sum of the transfer functions between each segment signal and its delayed signal is finally to be obtained. The equation for solving the transfer function is shown in Eq. (1).

$$H_1(f) = \frac{G_{XY}(f)}{G_{XX}(f)}; \tag{1}$$

where, $G_{xx}(f)$ denotes the product of the Fourier Transform between the input signal and its conjugate signal, $G_{xy}(f)$ denotes the Fourier Transformed product between the input to output signals, and $H_1(f)$ is the transfer function.

And $G_{xy}(f)$ is derived from Eq. (2).

$$G_{XY}(f) = \frac{1}{N} \sum_{i=1}^N F[y_i(t)]F^*[x_i(t)]; \tag{2}$$

Where, N is the number of segments of the signal segmentation, and F denotes the Fourier Transform, $*$ denotes the conjugate complex for complex numbers, $x_i(t)$ is the non-delayed signal, $y_i(t)$ is the delayed signal.

$G_{xx}(f)$ is derived from Eq. (3).

$$G_{XX}(f) = \frac{1}{N} \sum_{i=1}^N F[x_i(t)]F^*[x_i(t)]; \tag{3}$$

Finally, the transfer function is again converted to the time domain using the Fourier inverse transform, as shown by Eq. (4).

$$h(t) = F^{-1}[H_1(f)]; \tag{4}$$

Where, F^{-1} represents the Fourier inverse transform.

The original signal can be filtered using the transfer function to remove deterministic or random components from the raw

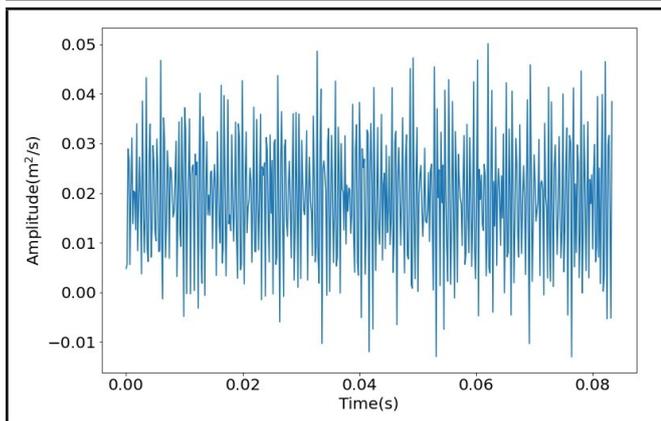


Figure 2. Time domain image of 0.014 inches of the outer race fault.

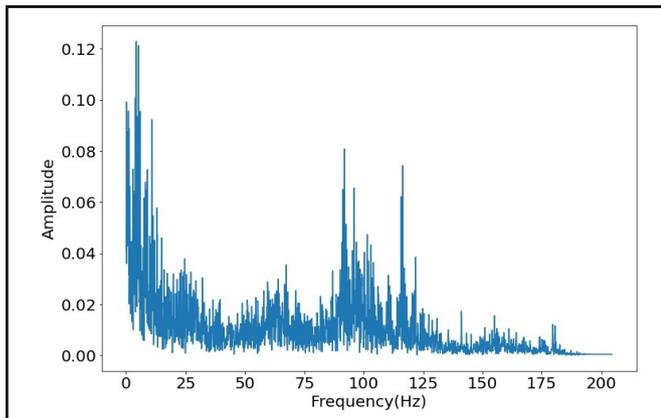


Figure 3. Corresponding DRS frequency spectrum image.

signal. The filtering process is usually performed using time-domain convolution or frequency-domain multiplication. The filtering method of time domain convolution is used here, ie., $(x(t) * h(t), x(t))$ is the whole input signal), after the convolution, the time domain function is converted to the frequency domain by Fourier transform again, and then output as DRS frequency spectrum image with a fixed length. As shown in Fig. 2 and Fig. 3, the time domain diagram of the bearing fault signal and its corresponding DRS frequency spectrum image represents the 0.014 inches of the outer race fault.

2.2. Deep Learning Networks

Five kinds of deep learning network structures are considered to examine the effect of DRSFSI-DL for fault diagnosis. These deep learning models are GoogLeNet, ResNet50, ResNet101v2, ResNet152v2, and Inception_ResNet.v2. According to the internal structures of these five deep learning networks, the following results from the Inception structure, residual module, and Inception_Res module are introduced respectively.

2.2.1. GoogLeNet network model

GoogLeNet was proposed by the Google team in 2014 and won first place in the Classification Task in the ImageNet competition. The Inception structure was first introduced into GoogLeNet.²² The Inception structure can fuse feature information at different scales to improve the network diagnosis ac-

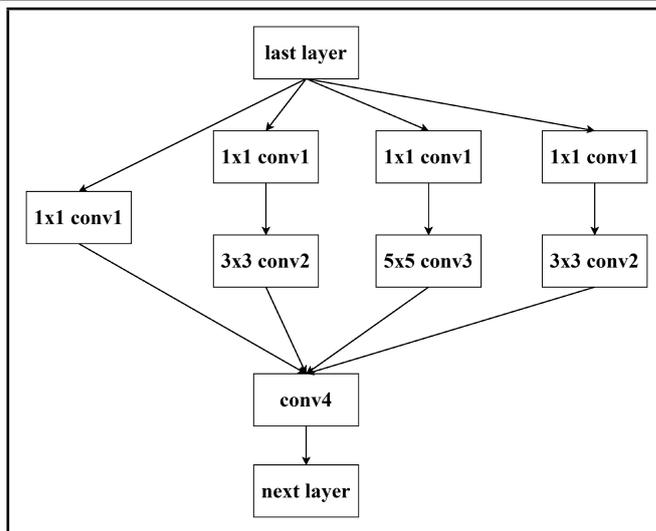


Figure 4. Inception structure.

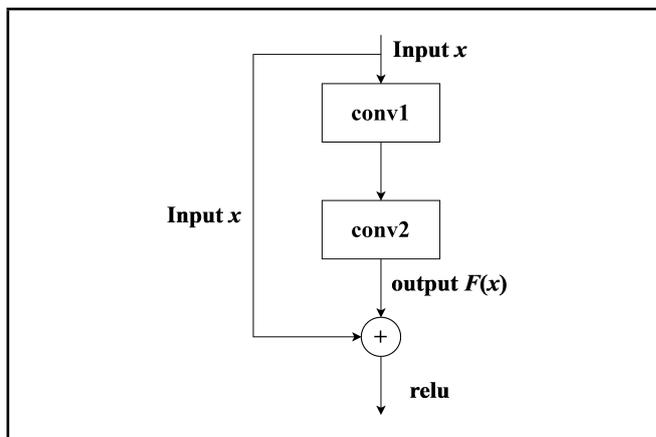


Figure 5. Residual module.

curacy. The underlying structure of the Inception structure is shown in Fig. 4.

As shown in Fig. 4, when the data from the last layer are input to the Inception module, it will be delivered to each of the four convolutional layers with different convolutional kernel sizes. In this way, the width and the adaptability of the network to the scale of the data will be increased. Additionally, the feature information of different scales will be fused too, and the multi-scale information of fault signals can be learned because of the difference of the perceptual fields of all the branches.

2.2.2. Residual Network (ResNet) model

ResNet proposed by Microsoft Labs in 2015, won first place in the ImageNet competition for classification tasks and first place for target detection.²³ The residual structure is the most critical component of the residual series networks. Among the ResNet model family, ResNet50, ResNet101v2, and ResNet152v2²⁴⁻²⁶ will be used as deep learning networks in this paper. The residual design can solve the problem of overfitting caused by too many layers of the network, which significantly increases the depth of networks and improves the classification ability of the network model.²⁷ Figure 5 below shows the schematic diagram of the residual structure.²⁸

When x is the input of the residual module, $F(x)$ is ob-

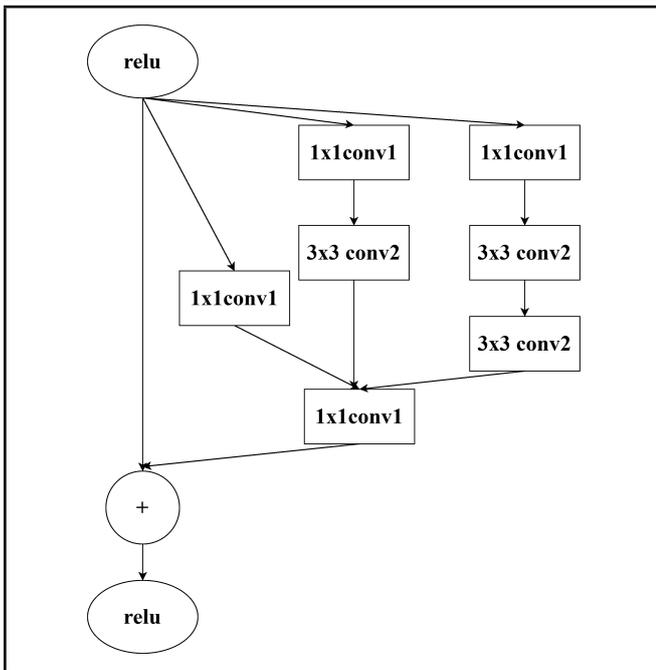


Figure 6. Inception_Res module.

tained after two convolutional layers. The final output result is the sum of input x and output result $F(x)$. If the $F(x) + x$ obtained after adding these two layers is worse than the original input x , it makes $F(x)$ zero. That is to say, after two layers of convolution pooling, the output $F(x) + x$ remains the same as x and does not become worse. And if $F(x) + x$ is better than the original input x , it remains the $F(x) + x$ as the final output result.

2.2.3. Inception_ResNet_v2 model

The main component modules of the Inception_ResNet_v2 network used in this paper are several different types of Inception_Res modules, a new network component structure obtained by combining the Inception structure and the residual module mentioned in the previous section, the principle of which is shown in Fig. 6.²⁹ The performance of the network model can be improved again to a certain extent based on the ResNet model.

The hardware choices for this article are as follows: GPU is RTX3090*1 with 24 GB of video memory. CPU is a 15-core AMD EPYC 7543 32-Core Processor with 80 GB of memory.

3. THE FRAME OF THE DRSFSI-DL METHOD

As shown in Fig. 7, the proposed fault diagnosis method combining DRS frequency spectrum images with deep learning (DRSFSI-DL) consists of the following three main steps. The first step is to select data to be trained from the datasets. The second step is to perform DRS transformation on these data to obtain two-dimensional frequency spectrum images. The third step is to use these spectrum images to train standard deep-learning network models and classify them according to predefined labels.

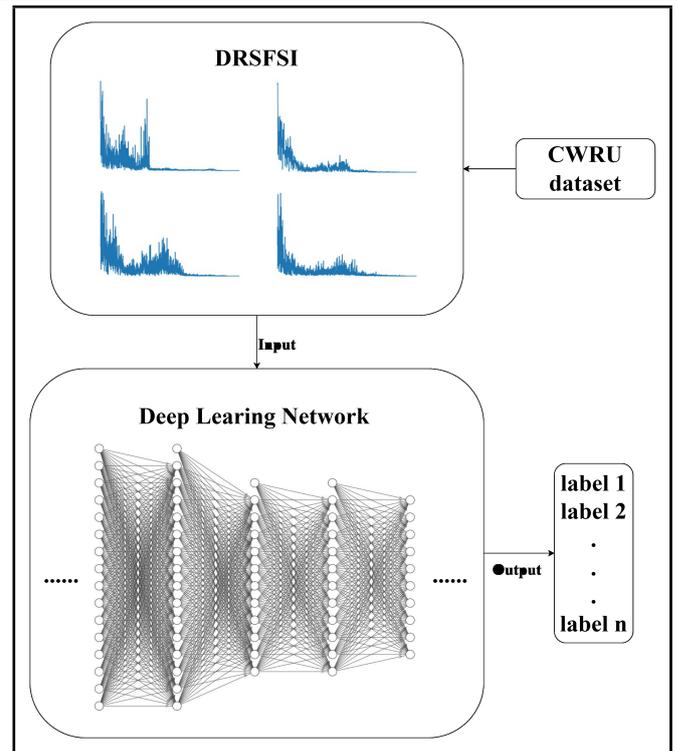


Figure 7. The steps of DRSFSI-DL.

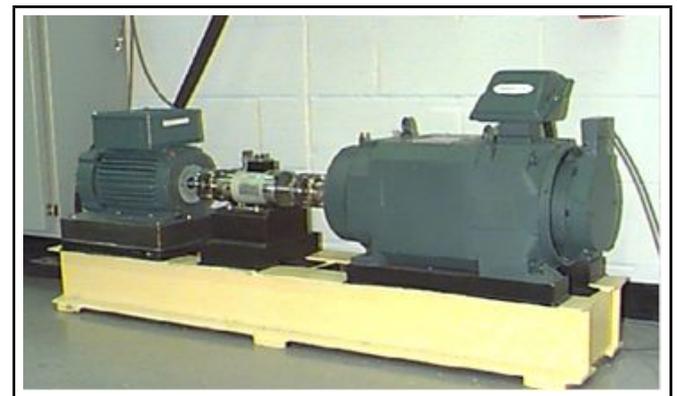


Figure 8. The CWRU bearing test rig.

4. THE FAULT DIAGNOSIS DATASET

4.1. Dataset Description

The rolling bearing fault dataset used in this paper is the well-known CWRU bearing dataset, which can be mainly sampled at 48 kHz and 12 kHz on the test stand shown in Fig. 8, including nine failure states and one normal state, which are normal, 0.007-inch inner race fault, 0.007-inch outer race fault, 0.007-inch ball fault, 0.014-inch inner race fault, 0.014-inch outer race fault, 0.014-inch ball fault, 0.021-inch inner race fault, 0.021-inch outer race fault, and 0.021-inch ball fault. Each kind of fault contains sampling data at four different loads, 0 hp, 1 hp, 2 hp, and 3 hp, whose corresponding motor speeds are 1792 rpm, 1772 rpm, 1750 rpm, and 1730 rpm, respectively. In addition, three faults are collected at 6 o'clock, 12 o'clock, and 3 o'clock according to the different outer ring sensor mounting positions for the same fault data of the outer ring of the rolling bearing.

Table 1. Classification of the CWRU dataset.

Datasets	A	B
Load (hp)	0	1
Rotation speed(rpm)	1792	1772
Numbers of Fault	10	10
Categories of Fault	X	X

4.2. The Fault Dataset Used in This Paper

The types of faults are classified according to their fault states. Then the bearings are classified into two operating datasets, A and B, based on the different loads subjected to the test rig during operation, and the bearing speeds are 1792 rpm and 1772 rpm. The loads are 0 hp and 1 hp, as shown in Table 1. And dataset A is used to demonstrate the effectiveness of the proposed method, and dataset B is used to verify the generalization of the proposed method.

Where, X indicates 10 different kinds of fault, ie. IF (0.007, 0.014, 0.021), BF (0.007, 0.014, 0.021), OF(0.007, 0.014, 0.021) and Normal. IF indicates inner race fault, BF indicates ball fault, OF indicates outer race fault, and the number in brackets indicates the size of the corresponding fault in inches.

Then, the one-dimensional data of each classification is transformed into two-dimensional frequency spectrum images by the corresponding frequency domain transformation. In the process of frequency domain transformation (including DRS, PSD and cepstrum), the time domain data is first segmented in equal length. Then the frequency domain transformation is performed on the 1D data of each size. Multiple spectrum images are obtained by this method and divided into a train set and test set according to the ratio of 3:1. The accuracy rates discussed later are the averaging accuracy rates obtained on the test set after the network is trained three times.

5. THE FAULT DIAGNOSIS BASED ON THE DRSFSI-DL

In this part, the best length of the fault signal segment will be determined according to the fault diagnosis efficiency of the DRSFSI-DL method, using the same deep learning networks. Then the best length of the signal segment can be applied to verify the effects of fault diagnosis of the DRSFSI-DL method, compared with the PSD images and cepstrum images.

5.1. Comparison of the Classification Accuracy of Different Deep Learning with Different Data Lengths

Five network structures, namely, GoogLeNet, ResNet50, ResNet101v2, ResNet152v2, and Inception_ResNet_v2, are used here as the network models for deep learning. In addition, four different segment lengths are selected for the raw data based on the use of the CWRU bearing datasets, i.e., data segment lengths of 8192, 4096, 2048, and 1024, were processed by the DRSFSI-DL method. The optimal data segment length is determined by comparing the training classification accuracy on the five deep networks with different data segment length cases. Note that the above datasets are all from dataset A.

Table 2. Performance of images with different data segment lengths for various DL networks.

Data Length	GoogLeNet	ResNet50	ResNet101v2	ResNet152v2	Inception_ResNet_v2
1024	95.44%	98.10%	98.45%	98.70%	98.98%
2048	97.56%	94.53%	97.65%	97.65%	98.09%
4096	98.70%	99.11%	99.51%	100.00%	100.00%
8192	96.40%	98.45%	98.19%	98.71%	99.48%

Table 3. Comparison of fault diagnosis of frequency spectrum image with DL networks using dataset A.

DL Models	DRSFSI	PSD images	Cepstrum images
GoogLeNet	98.70%	87.27%	73.29%
ResNet50	99.11%	95.33%	92.68%
ResNet101v2	99.51%	96.36%	92.68%
ResNet152v2	100.00%	97.61%	93.15%
Inception_ResNet_v2	100.00%	98.18%	94.84%

When the DRS frequency spectrum images of the different data lengths of dataset A is used to train the various deep learning neural networks, the accuracy of classification of the five deep learning networks using the four kinds of data segment length can be seen in Table 2. Among the five deep learning networks, ie. GoogLeNet, ResNet50, ResNet101v2, ResNet152v2, and Inception_ResNet_v2, the deep learning networks at the end always yielded the best classification accuracy. The best classification accuracy can be obtained for the same network model when the DRSFSI with a segment data length of 4096 is used as training input, and the highest classification accuracy of 100.00% can be obtained for ResNet152v2 and Inception_ResNet_v2 deep learning networks for classification. Therefore, the segment data length of 4096 will be determined for the following part. Note that the results obtained by using PSD and cepstrum are consistent with Table 2, which is not mentioned here.

5.2. Fault Diagnosis Using DRSFSI and Two Other Frequency Spectrum Images with DL Networks

To reflect the superiority of DRSFSI-DL for fault classification, the proposed DRSFSI method is also compared with PSD images and cepstrum images. Through the use of dataset A, the classification effects of three frequency spectrum images with five different DL networks can be obtained, as shown in Table 3.

According to Table 3, when the conventional frequency spectrum images are used for deep learning networks, the fault diagnosis effect of the PSD images is much better than that of the cepstrum images. However, the proposed DRSFSI is far superior to both conventional frequency spectrum images. And using DRSFSI, the highest accuracy rate will be achieved for all the deep learning networks. The latter four models reach more than 99%, while GoogLeNet can also get 98.70%. The DRSFSI combined with ResNet152v2 and Inception_ResNet_v2 networks can achieve the highest fault diagnosis accuracy, 100.00%. Fig. 9 and Fig. 10 show the variation process of the classification accuracy and training loss rate of the Inception_ResNet_v2 network over 50 training epochs. It can be seen from the figures that in the process of network training, the classification accuracy and loss rate of the model

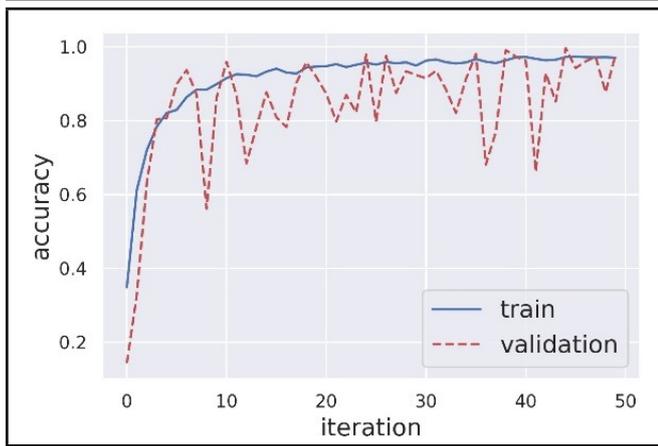


Figure 9. Training accuracy.

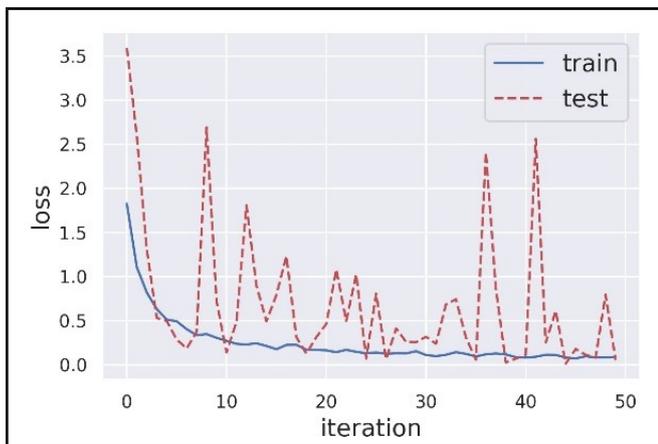


Figure 10. Training loss.

has reached a relatively high level around the 10th epoch. Although there are fluctuations in the later training epochs, the overall results are excellent.

Figure 11 denotes the confusion matrix of PSD images for dataset A, visualizing the classification results of the PSD with the Inception_ResNet_v2 network, and the results of classification accuracy are not good enough. Fig. (12) denotes the confusion matrix of DRSFSI-DL for dataset A, visualizing the classification results of the DRSFSI with the Inception_ResNet_v2 network. It can be seen that the highest classification accuracy, 100.00%, can be achieved with the use of the Inception_ResNet_v2 network when the DRSFSI is used as the classification dataset image, proving the effectiveness of the proposed DRSFSI-DL method.

6. THE GENERALIZATION OF THE DRSFSI-DL FAULT DIAGNOSIS METHOD

As mentioned above, the three kinds of frequency spectrum images have been applied to train the five deep learning networks considering dataset A. To verify the generalizability of the proposed DRSFSI-DL for fault diagnosis, these trained deep networks are also used directly to diagnose dataset B, which is different from dataset A. The classification results on the five trained networks are shown in Table 4.

From Table 4, it can be seen that when using the five types of deep networks trained using dataset A to diagnose the faults

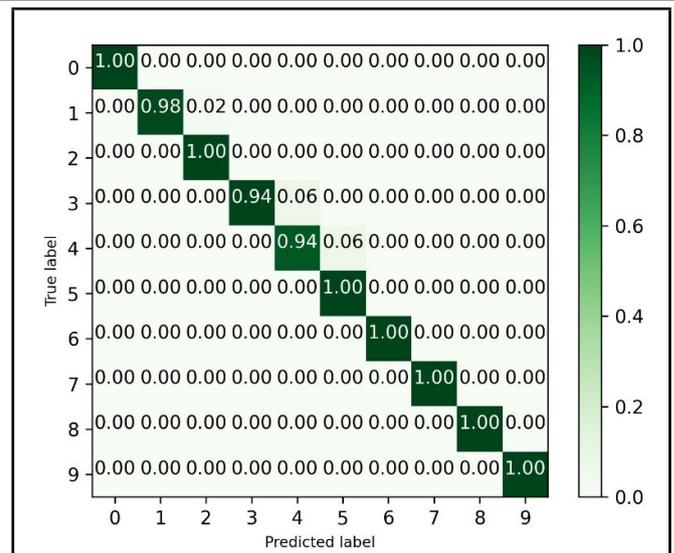


Figure 11. The confusion matrix of PSD images for dataset A.

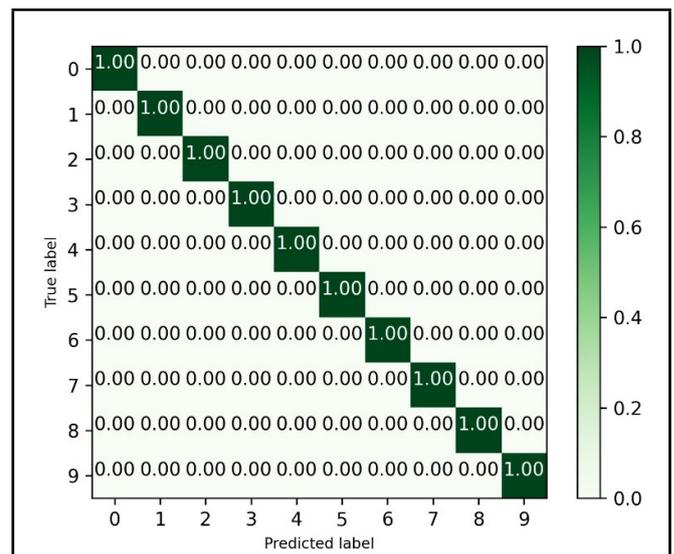


Figure 12. The confusion matrix of DRSFSI for dataset A.

Table 4. Comparison of three frequency spectrum images under dataset B.

DL Models	DRSFSI	PSD images	Cepstrum images
GoogLeNet	75.16%	58.22%	56.16%
ResNet50	80.63%	60.36%	61.59%
ResNet101v2	83.10%	62.53%	65.36%
ResNet152v2	85.36%	62.87%	67.33%
Inception_ResNet_v2	87.56%	65.70%	67.12%

in working dataset B directly, the fault diagnosis accuracy of all three frequency spectrum images decreases to some extent. However, the diagnosis effect of the PSD images and cepstrum images is worse. In contrast, the generalization of DRSFSI-DL is more robust. The accuracy decreases relatively less, showing that the DRSFSI-DL method has a reasonably strong generalization ability. As can be seen between Table 3 and Table 4, the average classification accuracy of DRSFSI, PSD images and cepstrum images in deep learning using dataset A is 99.46%, 94.95% and 87.64% respectively. Under dataset B, the average classification accuracy is 82.36%, 61.94% and 63.51%, respectively, and the decline rate is 17.10%, 33.10% and 24.13% in turn. DRSFSI has stronger robustness and generalization abil-

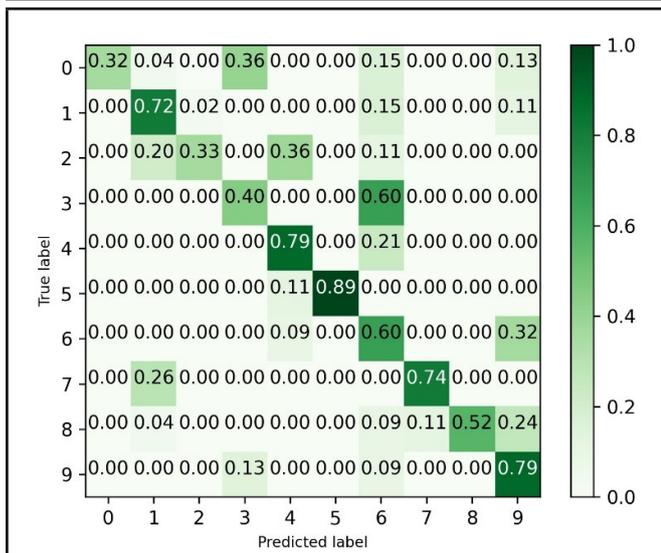


Figure 13. The confusion matrix of the PSD image for dataset B.

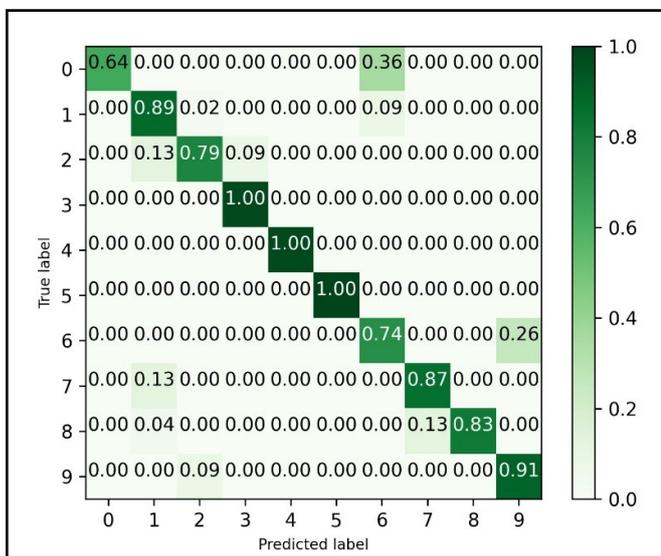


Figure 14. The confusion matrix of DRSFSFI for dataset B.

ity.

Figure 13 shows the results of the DRSFSFI-trained network under dataset A for classifying fault data of dataset B. It offers its better generalization performance and proves the effectiveness of the proposed DRSFSFI-DL method. Figure 14 shows the network trained by PSD images in dataset A for classifying fault data in dataset B. This illustrates its poor generalization performance, with a decrease in classification accuracy of nearly 30%.

7. THE VALIDATION OF THE DRSFSFI-DL FAULT DIAGNOSIS METHOD

In this section, the validity of the proposed DRSFSFI-DL method is verified by using the bearing fault dataset of the University of Ottawa, which collects the vibration signals of bearings with different health statuses over time. These signals can be generally classified into two cases: the first case is the bearing health status, which involves health, inner race fault and outer race fault. The second case is the operating conditions,

Table 5. Classification of the University of Ottawa dataset.

Categories of Fault	Health	Inner Fault	Outer Fault
Numbers of Fault	3		

Table 6. Performance of spectrum images with different DL networks.

DL models	DRSFSFI
ResNet50	98.06%
Inception_ResNet_v2	98.63%

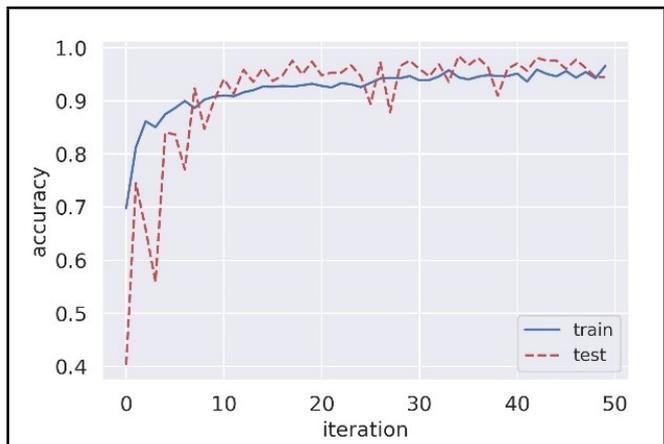


Figure 15. Training accuracy.

which involves increasing rotation speed, decreasing rotation speed, decreasing rotation speed after increasing rotation speed and increasing rotation speed after decreasing rotation speed. Therefore, a total of 12 different bearing working conditions are set. The data collection in each case contains two methods, the vibration signal collected by using the accelerometer and the speed data collected by using the encoder. The sampling frequency for all data is 200,000 Hz and the sampling time is 10 seconds.

For the different bearing fault datasets of the University of Ottawa, this section will mainly classify them according to their health status, i.e., health, inner race fault and outer race fault. The acceleration signal measured by the accelerometer is used as the time domain signal for the fault diagnosis, and the classification ratio of the training set to the test set is 3:1 during the training process. The specific classification is shown in Table 5.

After classifying the dataset, the fault data under the corresponding operating conditions were converted into DRS frequency spectrum images and classified by deep network models. The ResNet50 network and Inception_ResNet_v2 network are only used to verify for simplicity, with the specific classification accuracy shown in Table 6.

As shown in the above table, the proposed DRSFSFI-DL fault diagnosis method was effectively validated on the bearing fault dataset of the University of Ottawa with a maximum of 98.63%. Fig. 15 and Fig. 16 show the accuracy and loss rate during the training process, respectively. Overall, the classification accuracy achieved on this dataset is slightly lower than that achieved on the CWRU dataset, but it can still effectively show that the method has good fault diagnosis capability and generalization performance, because the new data has not been used for training the deep networks.

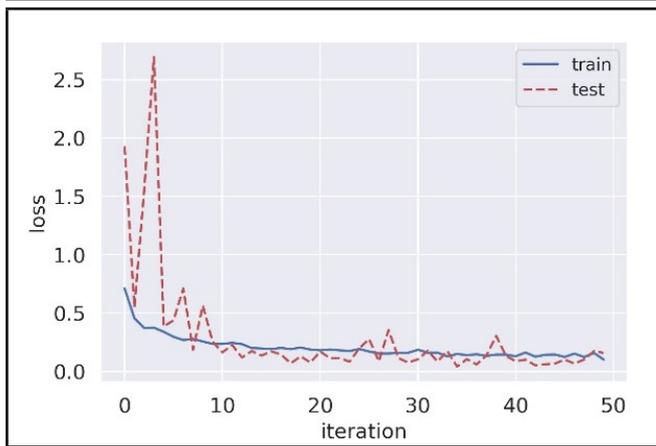


Figure 16. Training loss.

8. CONCLUSIONS

In this paper, a new fault diagnosis method called DRSFSI-DL has been proposed, the DRS frequency spectrum images were used as the input of five existing deep learning networks and compared with the PSD images and cepstrum images. The main conclusions are as follows.

1. The DRS frequency spectrum images can be obtained from fault signal using DRS method. To find the best data segment length of fault data for DRSFSI used to train deep learning networks, four different data segment lengths of 8192, 4096, 2048, and 1024 respectively are investigated through the use of five deep network models. It was found that the best data segment length is 4096, which can give the best diagnosis accuracy.
2. The original fault data were converted into the DRSFSI, the PSD image and cepstrum image respectively as the input of deep learning networks, and five kinds of deep learning network models are trained for the fault diagnosis. The fault diagnosis accuracy of the DRSFSI-DL method proposed was much higher than those of PSD images and cepstrum images.
3. In order to check the generalization ability of the DRSFSI-DL method, the deep learning network models trained by dataset A was directly used to diagnose the faults dataset B, which has something to do with dataset A. The results showed that the generalization ability of the DRSFSI-DL method was much better than those of PSD images and cepstrum images.
4. To further verify the reliability of the proposed method, a brand-new fault dataset which has nothing to do with the CWRU dataset, has been diagnosed using the DRSFSI-DL method. It also showed the better fault diagnosis accuracy with better generalization ability for the DRSFSI-DL method. In the future, the authors will strive to improve the DRSFSI-DL network considering transfer learning to achieve higher accuracy in the case of insufficient fault data under cross-working conditions.

ACKNOWLEDGEMENTS

The work was supported by the non-carbon energy conversion and utilization institute under the Shanghai Class IV Peak Disciplinary Development Program.

REFERENCES

- ¹ Smith, W. A. and Randall, R. B. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study, *Mechanical Systems and Signal Processing*, **64-65** 100–131, (2015). <https://doi.org/10.1016/j.ymssp.2015.04.021>
- ² Fei, C.-W., Bai, G.-C., Tang, W.-Z. and Ma, S. Quantitative Diagnosis of Rotor Vibration Fault Using Process Power Spectrum Entropy and Support Vector Machine Method, *Shock and Vibration*, **2014**(2014). <https://doi.org/10.1155/2014/957531>
- ³ Zhang, X., Zhou, R. and Zhang, W. Improved local cepstrum and its applications for gearbox and rolling bearing fault detection, *Measurement Science and Technology*, **30**(7), (2019). <https://doi.org/10.1088/1361-6501/ab1235>
- ⁴ Huang, L., Huang, H. and Liu, Y. A Fault Diagnosis Approach for Rolling Bearing Based on Wavelet Packet Decomposition and GMM-HMM, *International Journal of Acoustics and Vibration*, **24**(2), 199–209, (2019). <https://doi.org/10.20855/ijav.2019.24.21120>
- ⁵ Shao, S., McAleer, S., Yan, R. and Baldi, P. Highly Accurate Machine Fault Diagnosis Using Deep Transfer Learning, *Ieee Transactions on Industrial Informatics*, **15**(4), 2446–2455, (2019). <https://doi.org/10.1109/tii.2018.2864759>
- ⁶ Zhu, Z., Peng, G., Chen, Y. and Gao, H. A convolutional neural network based on a capsule network with strong generalization for bearing fault diagnosis, *Neurocomputing*, **323** 62–75, (2019). <https://doi.org/10.1016/j.neucom.2018.09.050>
- ⁷ Wen, L., Li, X. and Gao, L. A transfer convolutional neural network for fault diagnosis based on ResNet-50, *Neural Computing & Applications*, **32**(10), 6111–6124, (2020). <https://doi.org/10.1007/s00521-019-04097-w>
- ⁸ LeCun, Y., Bengio, Y. and Hinton, G. Deep learning, *Nature*, **521**(7553), 436–444, (2015). <https://doi.org/10.1038/nature14539>
- ⁹ Zhao, B., Zhang, X., Zhan, Z. and Pang, S. Deep multi-scale convolutional transfer learning network: A novel method for intelligent fault diagnosis of rolling bearings under variable working conditions and domains, *Neurocomputing*, **407** 24–38, (2020). <https://doi.org/10.1016/j.neucom.2020.04.073>

- ¹⁰ Han, T., Liu, C., Yang, W. and Jiang, D. Deep transfer network with joint distribution adaptation: A new intelligent fault diagnosis framework for industry application, *Isa Transactions*, **97** 269–281, (2020). <https://doi.org/10.1016/j.isatra.2019.08.012>
- ¹¹ Jiao, J., Zhao, M., Lin, J. and Liang, K. Residual joint adaptation adversarial network for intelligent transfer fault diagnosis, *Mechanical Systems and Signal Processing*, **145** (2020). <https://doi.org/10.1016/j.ymssp.2020.106962>
- ¹² Zhang, R., Tao, H., Wu, L. and Guan, Y. Transfer Learning With Neural Networks for Bearing Fault Diagnosis in Changing Working Conditions, *Ieee Access*, **5** 14347–14357, (2017). <https://doi.org/10.1109/access.2017.2720965>
- ¹³ Zhang, A., Wang, H., Li, S., Cui, Y., Liu, Z., Yang, G. and Hu, J. Transfer Learning with Deep Recurrent Neural Networks for Remaining Useful Life Estimation, *Applied Sciences-Basel*, **8** (12), (2018). <https://doi.org/10.3390/app8122416>
- ¹⁴ Cao, P., Zhang, S. and Tang, J. Preprocessing-Free Gear Fault Diagnosis Using Small Datasets With Deep Convolutional Neural Network-Based Transfer Learning, *Ieee Access*, **6** 26241–26253, (2018). <https://doi.org/10.1109/access.2018.2837621>
- ¹⁵ Lu, T., Yu, F., Han, B. and Wang, J. A Generic Intelligent Bearing Fault Diagnosis System Using Convolutional Neural Networks With Transfer Learning, *Ieee Access*, **8** 164807–164814, (2020). <https://doi.org/10.1109/access.2020.3022840>
- ¹⁶ Du, Y., Wang, A., Wang, S., He, B. and Meng, G. Fault Diagnosis under Variable Working Conditions Based on STFT and Transfer Deep Residual Network, *Shock and Vibration*, **2020**(2020). <https://doi.org/10.1155/2020/1274380>
- ¹⁷ Yu, H., Tian, Z., Li, H., Xu, B. and An, G. A Novel Deep Belief Network Model Constructed by Improved Conditional RBMs and its Application in RUL Prediction for Hydraulic Pumps, *International Journal of Acoustics and Vibration*, **25**(3), 373–382, (2020). <https://doi.org/10.20855/ijav.2020.25.31669>
- ¹⁸ Mao, W., Ding, L., Tian, S. and Liang, X. Online detection for bearing incipient fault based on deep transfer learning, *Measurement*, **152** (2020). <https://doi.org/10.1016/j.measurement.2019.107278>
- ¹⁹ Mao, W., He, J. and Zuo, M. J. Predicting Remaining Useful Life of Rolling Bearings Based on Deep Feature Representation and Transfer Learning, *Ieee Transactions on Instrumentation and Measurement*, **69**(4), 1594–1608, (2020). <https://doi.org/10.1109/tim.2019.2917735>
- ²⁰ Borghesani, P., Pennacchi, P., Randall, R. B. and Ricci, R. Order tracking for discrete-random separation in variable speed conditions, *Mechanical Systems and Signal Processing*, **30** 1–22, (2012). <https://doi.org/10.1016/j.ymssp.2012.01.015>
- ²¹ Randall, R. B. and Antoni, J. Rolling element bearing diagnostics—A tutorial, *Mechanical Systems and Signal Processing*, **25**(2), 485–520, (2011). <https://doi.org/10.1016/j.ymssp.2010.07.017>
- ²² Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A. A. and Aaai. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning, *31st AAAI Conference on Artificial Intelligence*, San Francisco, CA, (2017).
- ²³ Shafiq, M. and Gu, Z. Deep Residual Learning for Image Recognition: A Survey, *Applied Sciences-Basel*, **12**(18), (2022). <https://doi.org/10.3390/app12188972>
- ²⁴ Li, B. and He, Y. An Improved ResNet Based on the Adjustable Shortcut Connections, *Ieee Access*, **6** 18967–18974, (2018). <https://doi.org/10.1109/access.2018.2814605>
- ²⁵ Jeon, W.-S., Rhee, S.-Y. and Ieee. Single Image Super Resolution using Residual Learning, *International Conference on Fuzzy Theory and Its Applications (iFUZZY)*, New Taipei City, TAIWAN, (2019).
- ²⁶ Tao, Z., Bing-qiang, H., Huiling, L., Hongbin, S., Pengfei, Y. and Hongsheng, D. F-18-FDG-PET/CT Whole-Body Imaging Lung Tumor Diagnostic Model: An Ensemble E-ResNet-NRC with Divided Sample Space, *Biomed Research International*, **2021** (2021). <https://doi.org/10.1155/2021/8865237>
- ²⁷ Zhang, K., Sun, M., Han, T. X., Yuan, X., Guo, L. and Liu, T. Residual Networks of Residual Networks: Multi-level Residual Networks, *Ieee Transactions on Circuits and Systems for Video Technology*, **28**(6), 1303–1314, (2018). <https://doi.org/10.1109/tcsvt.2017.2654543>
- ²⁸ Lim, B., Son, S., Kim, H., Nah, S., Lee, K. M. and Ieee. Enhanced Deep Residual Networks for Single Image Super-Resolution, *30th IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Honolulu, HI, (2017).
- ²⁹ Thomas, A., Harikrishnan, P. M., Palanisamy, P. and Gopi, V. P. Moving Vehicle Candidate Recognition and Classification Using Inception-ResNet-v2, *44th Annual IEEE-Computer-Society International Conference on Computers, Software, and Applications (COMPSAC)*, Electr Network, (2020).