



Article

Frequency Occurrence Plot-Based Convolutional Neural Network for Motor Fault Diagnosis

Eduardo Jr Piedad ¹, Yu-Tung Chen ², Hong-Chan Chang ² and Cheng-Chien Kuo ^{2,*}

¹ Department of Electrical Engineering, University of San Jose-Recoletos, Cebu City 6000, Philippines; eduardojr.piedad@usjr.edu.ph

² Department of Electrical Engineering, National Taiwan University of Science and Technology, Taipei 106, Taiwan; 100m05003@stud.sju.edu.tw (Y.-T.C.); hcchang@mail.ntust.edu.tw (H.-C.C.)

* Correspondence: cckuo@mail.ntust.edu.tw; Tel.: +886-0920-881-490

Received: 21 August 2020; Accepted: 13 October 2020; Published: 18 October 2020



Abstract: A novel motor fault diagnosis using only motor current signature is developed using a frequency occurrence plot-based convolutional neural network (FOP-CNN). In this study, a healthy motor and four identical motors with synthetically applied fault conditions—bearing axis deviation, stator coil inter-turn short circuiting, a broken rotor strip, and outer bearing ring damage—are tested. A set of 150 three-second sampling stator current signals from each motor fault condition are taken under five artificial coupling loads (0, 25%, 50%, 75% and 100%). The sampling signals are collected and processed into frequency occurrence plots (FOPs) which later serve as CNN inputs. This is done first by transforming the time series signals into its frequency spectra then convert these into two-dimensional FOPs. Fivefold stratified sampling cross-validation is performed. When motor load variations are considered as input labels, FOP-CNN predicts motor fault conditions with a 92.37% classification accuracy. It precisely classifies and recalls bearing axis deviation fault and healthy conditions with 99.92% and 96.13% f-scores, respectively. When motor loading variations are not used as input data labels, FOP-CNN still satisfactorily predicts motor condition with an 80.25% overall accuracy. FOP-CNN serves as a new feature extraction technique for time series input signals such as vibration sensors, thermocouples, and acoustics.

Keywords: fault diagnosis; frequency occurrence plot; convolutional neural network; motor loading; current signal; FOP-CNN; bearing fault; short circuit fault

1. Introduction

Prognostics and health management (PHM) has modernized the industry in terms of equipment reliability, attracting both academia and industry practice [1]. In the PHM strategy, diagnostics and prognostics are two important mechanisms applied in machine condition-based maintenance. A diagnostic mechanism detects, isolates, and identifies the present machine condition. Driven primarily by machines such as motors and generators, modern industries have been improved by advance diagnostics such as better preventive maintenance, improved safety and increased reliability [2]. Deep learning, an emerging branch of artificial intelligence (AI), has been playing an important role in this PHM modernization.

For prognostics and the health management of machines, bearing fault diagnosis is one of the well-known applications of deep learning (DL). The recent survey in [3] and the review in [4] provide comprehensive assessments of different state-of-the-art DL-based machine health monitoring systems applied to bearing fault diagnostics. These systems vary by their different settings; thus, there is always the need to provide alternatives to help AI practitioners choose the best-suited algorithm.

Feature selection is one of the primary concerns for effective deep learning (DL) applications. Vibrations or acoustic signals tend to be widely used features in bearing fault diagnosis [4–6]. Various

deep learning algorithms have been developed for feature extraction. A widely known convolutional neural network has recently evolved into different variants such as the convolutional neural network and convolutional discriminative feature learning [7,8], the discriminative deep belief network [9,10], the convolutional bi-directional long short-term memory (LSTM) network [11], and the ensemble deep convolutional neural network [12]. Deep neural networks with unsupervised learning based on auto-encoders [13] and recurrent neural networks [14] have also been developed for machine health monitoring. Other than vibration or acoustic signal signature features, motor current analysis [15–17] and thermographic images [18–21] are also considered and employed, showing practically good performances. Their feature combinations, such as vibration and current signals [22], are also studied. For practical application, practitioners may not prefer vibration features due to the long-term problem of installing physical sensors. Thermographic devices are also relatively expensive. Thus, the stator current signal seems to be an attractive feature due to its easy installation, more reliable data collection, and the fact that it is comparatively cheaper.

Analyzing timer series data such as vibration and stator current signals uses various feature preprocessing methods. Digital signal processing (DSP) techniques such as wavelet transformation [23–25], frequency spectral analysis [26–28], empirical mode decomposition (EMD) [29–31], Hilbert–Huang transform [32–34], and combined Hilbert and wavelet transform [35,36] have been used for different motor faults, especially for non-DSP practitioners. Frequency transformation, though, tends to be the simplest tool to analyze time series data diagnosis schemes. However, these techniques are selected and configured manually, meaning that they may prove difficult to use.

Recurrence plots are widely used for the analysis and visualization of complex and dynamic systems [37]. Recently, a study [38] reviewed, from the last two decades, the application of recurrence plots in various areas. It was found that this simple recurrence plot technique can represent the dynamic characteristics behind music pieces [39]. A similar plotting scheme is used in this study, using frequency spectra instead of time series signals.

A new method for detecting motor faults is proposed based on frequency occurrence plots and deep learning. The remaining sections of this study are as follows: Materials and Methods, Deep learning Implementation, Results and Discussion, and Conclusion.

2. Motor Dataset

A three-phase 220-V, 2-HP, 4-pole, 1720-rpm squirrel-cage induction motor is used as the test motor for data collection. Table 1 shows the actual motor specifications. A set of 150 three-second healthy motor current signals is collected from this healthy motor, with a sampling frequency of 10,000 Hz. The motor is sampled under five coupled loading variations (0, 25%, 50%, 75% and 100% loads). Another four similar motors are prepared under the same data collection to generate data for four motor fault conditions. With similar motor specifications, these motors are synthetically manipulated to simulate artificial faults, namely bearing axis deviation, stator winding inter-turn short circuiting, a broken rotor strip, and outer ring bearing damage.

Table 1. Test motor specification.

Label	Parameter/Value
Type	AEEF-90-4 Induction motor
Output	2 HP
Pole	4
Insulation	E
Volt	220/380 V
Amp	7 A
r.p.m	1450/1720
Duty type	S1
Cycle (Hz)	50/60
Connection	delta low voltage/ wye high voltage
Manufacturer	TW-141221 Tai Wei Electric Factory., Ltd.

2.1. Synthetic Motor Fault Conditions

There are four typical motor fault conditions—bearing axis misalignment, inter-turn short circuiting, a broken rotor strip and an outer ring bearing fault—synthetically applied in four respective test motors.

2.1.1. Bearing Axis Misalignment

The bearing axis deviation fault happens when a motor is eccentrically coupled to its load. Improper installation, changes or damage to motor bases cause the motor shaft to misalign with the coupling load. Similar to [40], an artificially created eccentricity misalignment experiment with an elevation of 0.5 mm, as illustrated in Figure 1a, is also used to simulate this fault.

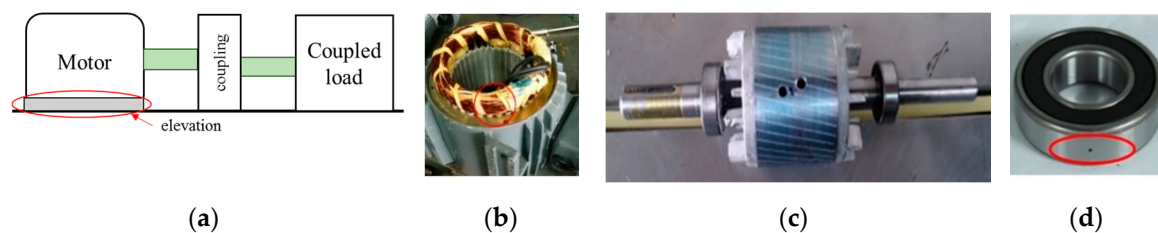


Figure 1. Four synthetically prepared motor fault conditions: (a) bearing axis misalignment; (b) stator inter-turn short circuiting; (c) broken rotor strip; and (d) outer ring bearing damage.

2.1.2. Stator Inter-Turn Short Circuiting

The aging insulation of the stator coil due to the long operation period of motors is often believed to be the primary reason for motor overheating. In severe cases, it causes short-circuits between turns of the same phase or even in different phases. To simulate this motor stator turn-to-turn short circuit fault, two adjacent turns of the stator winding of a test motor are intentionally short-circuited by breaking their insulation and allowing them to make contact, as shown in Figure 1b.

2.1.3. Broken Rotor Strip Fault

Excessive current due to long-term overloading is often seen as the reason of a broken rotor strip fault. This fault is synthetically simulated by drilling directly into one side of the rotor bar similar to [41]. Taking into consideration its serious impact to the motor, the first drill underwent an experimental detection. After checking that the motor is still running well, a second drill is performed as shown in Figure 1c.

2.1.4. Outer Ring Bearing Fault

Lastly, outer ring damage is a common bearing fault. This bearing fault increases machine vibration whenever a bearing ball passes over a damaged area [42]. In this simulation, a hole is drilled in the outer ring of a test motor. Electrical conducting heat is applied to the hole, making sure that its residue is removed and no physical deformations are present after drilling, as shown in Figure 1d.

2.2. Data Collection

The experimental simulation of previous test motors produces a total of 3750 time series data from 150 three-second current signals of five motor conditions under five loading variations. A set of three full-cycle sample waves of current signals plotted in Figure 2 shows five motor conditions under three different motor loadings. It can be observed that the healthy motor upper-bounded all other fault-conditioned motors in terms of current magnitude in all three motor loadings. This may be because healthy motors have less energy dissipation than faulty motors.

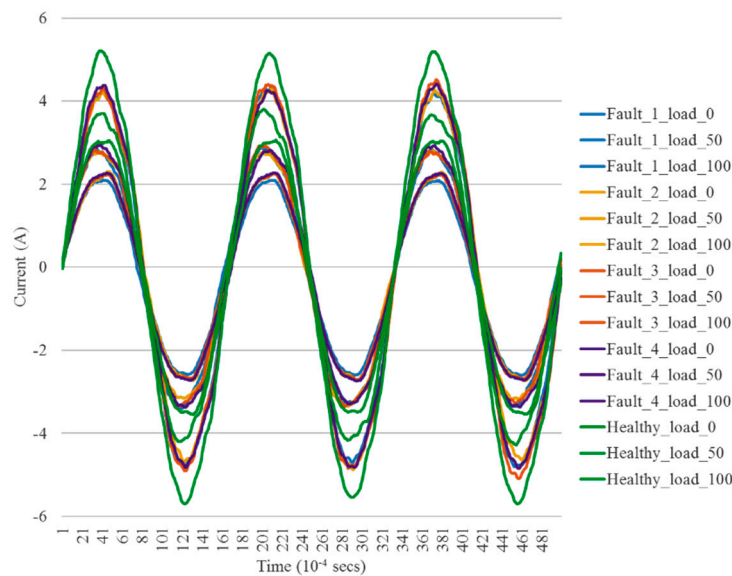


Figure 2. Three-cycle time series samples of stator current signal from five motor conditions with three loading variations—0%, 50% and 100% coupled loading.

3. Data Preprocessing

With the generated raw motor data, two data processing techniques are used before learning. First, a signal data transformation, from the time domain to the frequency spectrum, is performed using a frequency transformation. Second, novel frequency occurrence plot (FOP) image generation is employed to convert the frequency spectra into FOPs. These plots will then serve as inputs for the convolutional neural network (CNN) model.

3.1. Fast Frequency Transform

Fourier analysis is a widely known tool when converting a time series signal into frequency spectrum representation, and vice versa. It has a form called discrete-time Fourier transform (DTFT) that analyzes discrete-time samples whose intervals have units of time. Given a sampling signal $X(j)$, $j = 0, 1, \dots, N - 1$ with N sampling points, a sequence in (1), as a function of frequency n , gives the complex Fourier amplitudes. The expression in (2) is a principle N th root of unity in a complex Fourier series. For this motor digital signal application, this DTFT is employed using discrete Fourier transform (DFT).

$$A(n) = \frac{1}{N} \sum_{j=0}^{N-1} X(j) W_N^{-jn} \quad (1)$$

$$W_N = e^{2\pi i/N} \quad (2)$$

Fast Fourier transform (FFT) is simply an algorithm that performs DFT [43]. With the previously generated time series motor current signals, FFT is used to transform these into frequency spectra. An example of a frequency spectrum of a healthy motor signal is shown in Figure 3. FFT is performed using SciPy library [44] to convert the time series data into the frequency spectrum. Three data preprocessing techniques are performed to avoid potential noise and to ease the learning process of the proposed fault classification system.

First, data clipping helps truncate the converted data within the 0–500 Hz range. It was assumed here that all motor faults would not operate beyond 500 Hz. Increasing the frequency range decreases the FOP image resolution, which may give a poor CNN classification performance. Second, a 90% percentile clipping is performed to change the magnitude of less significant frequencies into zero, thus avoiding possible noise. Finally, the magnitude of the operating frequency (60 Hz) and its side

band frequencies (around 59–61 Hz) are far greater than the other modal frequencies; thus, a log function normalization is performed. Figure 4 plots a sample of a pre-processed healthy motor condition dataset of frequency spectra.

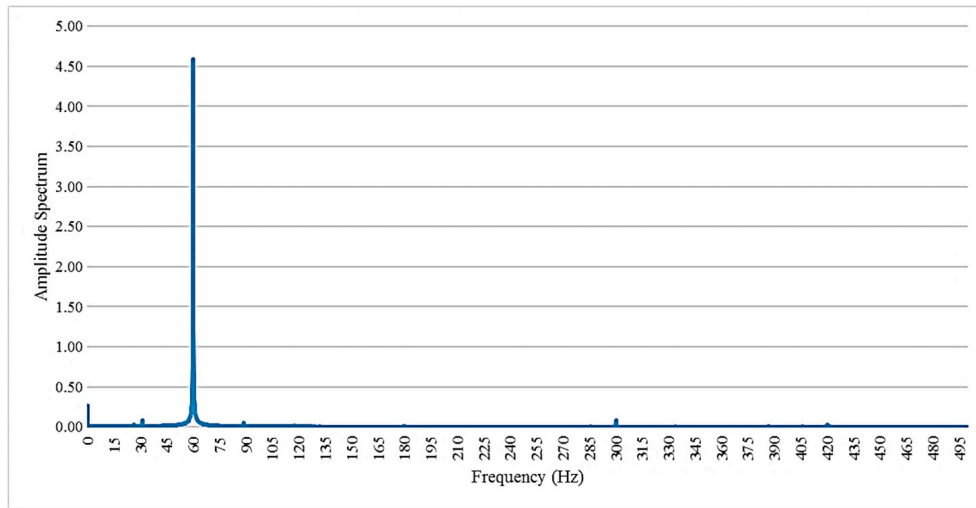


Figure 3. Frequency transformation of a healthy motor condition.

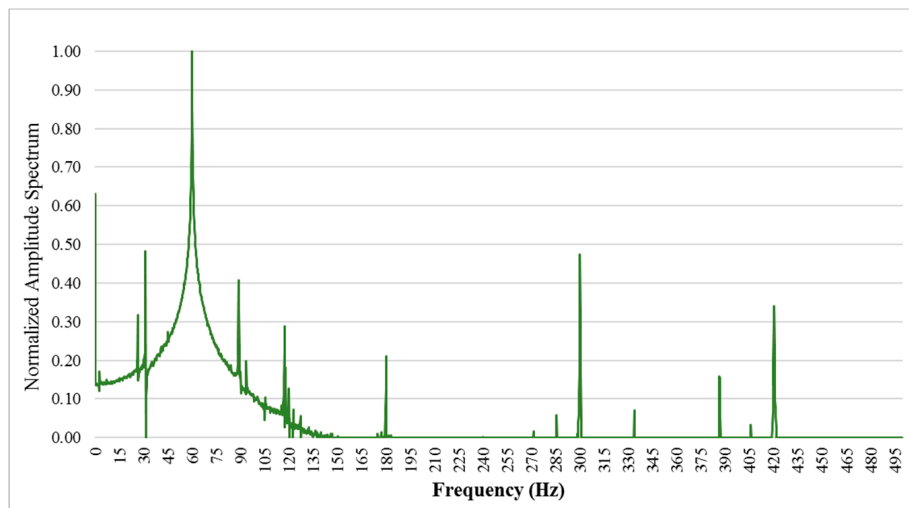


Figure 4. Frequency transformation of healthy motor condition with data preprocessing.

The presence of other frequencies with noticeable amplitudes is often believed to have been caused by motor faults. Identifying these frequencies for each motor fault type is most likely difficult due to the complex frequency spectra, as perceived in Figure 5. Accordingly, differences among five motor conditions can be observed, but these seem difficult to distinguish simply by using human visual recognition.

3.2. Frequency Occurrence Plots

Let metric space M be defined and let $A(i) \in M$ denote the i th point of the previously defined frequency spectrum A . A frequency occurrence plot is defined in (3) where the spectrum $A(i) = A(j)$.

$$FOP(i, j) = (A(i) - A(j)^T) / \epsilon \tag{3}$$

Here, ϵ is the mapping resolution, which scales the difference of two identical signals in the i th row and j th column. The matrix $FOP(i, j)$ are transformed into a color map. This initially normalizes

and scales the data, then maps it into an RGB color map using the Matplotlib library [45]. Thus, ε has no effect in the color mapping, but is useful for visualization purposes. After a series of trials, we use $\varepsilon = 0.001$ which produces distinctive occurrence plots.

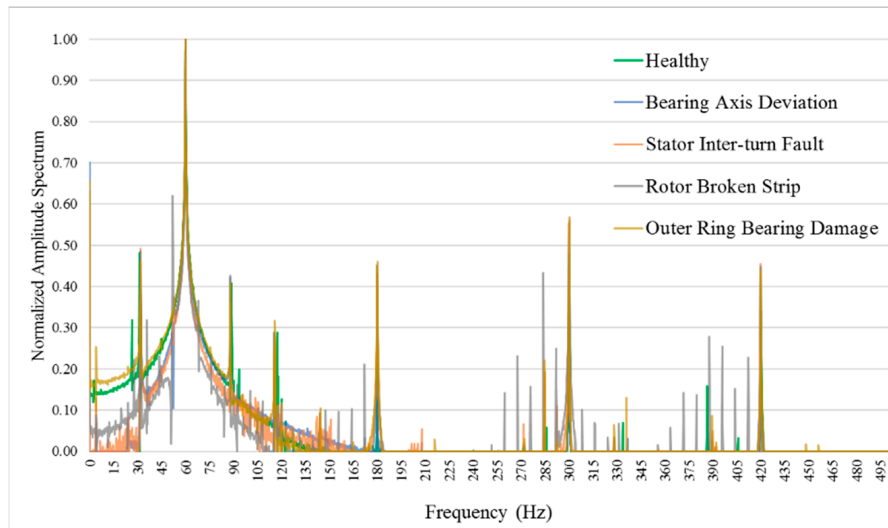


Figure 5. Frequency occurrence plots of five motor fault conditions under five loading conditions.

Figure 6 displays a sample illustration of how a frequency occurrence plot (FOP) is produced from a sample in the range of the 500 Hz frequency spectrum. Each plot has a 217×217 resolution. Brightly colored lines, vertically and horizontally, represent higher magnitude in the frequency spectrum. Accordingly, the motor operating frequency at 60 Hz has the brightest color. The other bright line represents other frequencies with significant magnitudes, which may have been caused by different motor faults.

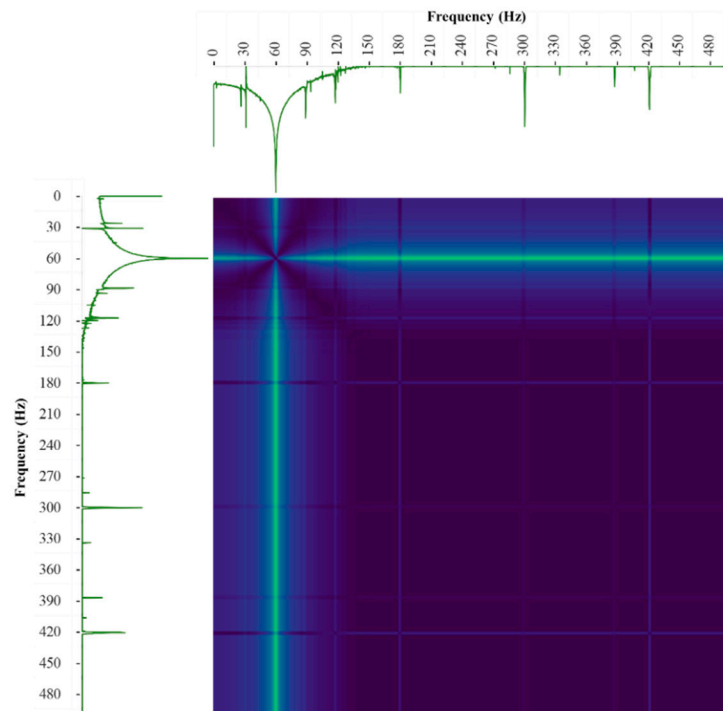


Figure 6. An illustration of how to produce frequency occurrence plots from a frequency spectrum with 500 Hz range.

4. Deep Learning Implementation

A convolutional neural network (CNN), a powerful deep learning tool for image recognition, is used to learn and classify faults from the generated FOPs in this study. Initially, FOPs are converted into three images, each representing its red, green, and blue (RGB) color features with the same image size. These serve as initial inputs for a CNN model.

4.1. Convolutional Neural Network

The architecture of the employed sequential-based CNN is shown in Figure 7. The CNN model initially inputs three extracted color images from the original FOP. It has various stages such as convolution, max pooling, dense, flatten, and dropout that lead to its final fault classification. Convolution layers use a filter matrix to obtain convolved feature maps by performing convolution operations over an array of input image pixels. On the other hand, the max pooling layer applies a moving two-dimensional window to the incoming matrix and outputs its maximum value to down-sample it, reduce its dimension and generalize its internal feature. We use a 2×2 window for two max pool layers, thus reducing their output by half. The dense layer is simply a linear operation where each input is connected to every output with weights. The first dense layer has a huge number of output units, so dropout was performed. Flattening, which is simply a method of linearizing a two-dimensional array, was also used. Dropout is a popular and well-known regularization technique that reduces the risk of overfitting. It is applied and tuned in different values per layer. Finally, another dense layer is added, which serves as its output layer. Here, five output units represent five motor conditions.

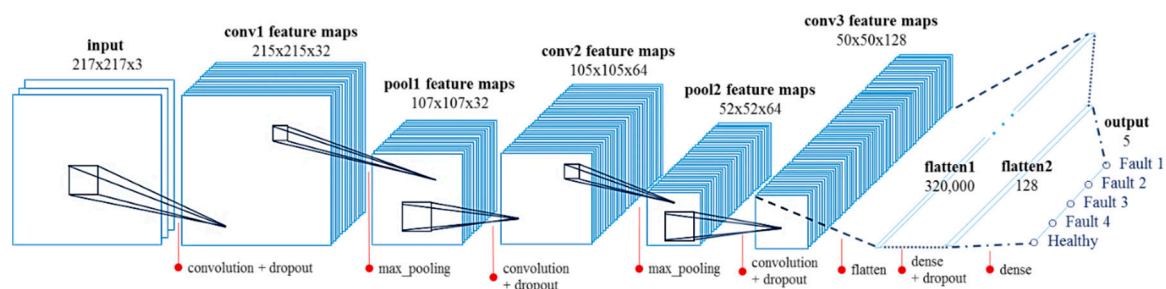


Figure 7. A proposed convolutional neural network (CNN) architecture with actual parameters.

This model is trained under batch gradient descent and the Adam optimizer. For epoch optimization, the Adam optimizer is a widely used optimization method for deep learning applications and is favorably chosen over other stochastic optimization methods [46].

4.2. Supervised Learning

The supervised learning of CNN is summarized in Figure 8. There are five steps performed in this implementation—model selection, model training and testing, model performance comparison, test scenarios, and performance validation.

4.2.1. Model Selection

First, the CNN model architecture and its hyper-parameters are chosen during the model selection, based on the brute-force method. This manual selection is still a limitation of this study because finding its optimal value tends to be computationally expensive and complex.

4.2.2. Model Training and Testing

Simultaneously, the FOPs generated from the previous section are split into training and testing datasets. Using the selected model parameters, model training is first performed by learning the patterns

and features from the training FOPs and evaluating their training performance. Then, another set of FOPs, also called testing FOPs, test the trained CNN model, and evaluate its testing performance. In the third stage, a comparison between training and testing performances is performed to observe the presence of overfitting.

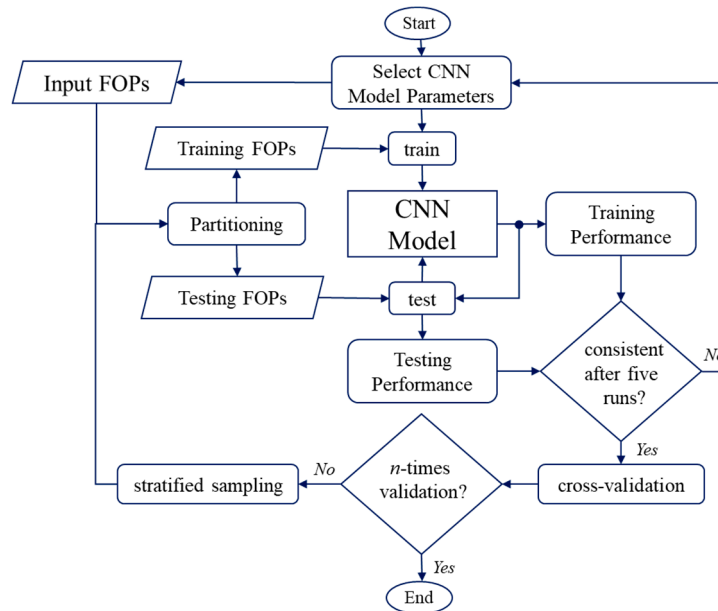


Figure 8. A supervised deep learning pseudocode for CNN model with validation.

4.2.3. Model Performance Evaluation

The supervised training and testing of CNN are evaluated using training and testing datasets, respectively. There are two typical modes of evaluation. First, the loss function of the CNN model is determined, usually in the form of a loss of function graph. This measures the consistency between the predicted value and actual label of the input FOPs during the training phase based on the theoretical functions used by CNN. The robustness of the model increases as the loss value decreases. To determine whether the model has overfitting issues, we also consecutively determine the loss function of the testing dataset and compare it with the loss function of the training dataset in every epoch. Categorical cross entropy (CCE) is commonly used and is shown to have a robust performance, even with synthetically generated noisy labels [47].

$$Accuracy = \frac{\text{Number of correctly classified FOPs}}{\text{Total number of FOPs}} \% \quad (4)$$

$$F \text{ score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (5)$$

Second, the classification performance accuracy in (4) is used to determine the model prediction accuracy. To further evaluate its performance in terms of positive and false negative classification, the F-score in (5) is also used. In addition, a confusion matrix provides a visualization of each classification performance.

4.2.4. Test Scenarios

A motor usually operates in different loading conditions throughout its whole operation. In actual practice, monitoring motor load values may have practical applications for a company. This study thus simulates two test scenarios. First, a simulation is performed when motor loading condition is available. Five separate train–test CNN models are simulated, corresponding to the five motor loading conditions.

Each model is simulated using 600 and 150 frequency occurrence plots (FOPs) for training and testing simulations, respectively. All five CNN models are combined into a single, generalized model.

Another case is performed when the motor loading condition is assumed to be unavailable. Only one CNN model, using the entire dataset at once, is learned, without data labeling, based on motor loading conditions. A total of 3750 frequency occurrence plots (FOPs) are used. It only learns 3000 and 750 FOPs for training and testing simulations, respectively. Both cases have similar partitions and equal numbers of FOPs.

4.2.5. Performance Validation

Testing datasets are used to evaluate the same model chosen from the previous step. To avoid model bias, each testing dataset is completely different from its training dataset. After testing, the testing performance is compared with the previous training performance. This is repeated via fivefold cross-validation with stratified sampling data partition, as shown in Figure 9.

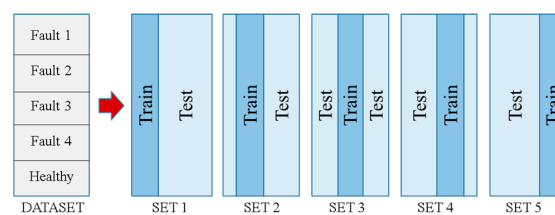


Figure 9. Five sets of cross-validation stratified sampling data partitions.

The frequency occurrence plot based convolutional neural network (FOP-CNN) algorithm is developed as shown in Algorithm 1. The FOP-CNN core is built primarily in Keras, an open-source neural network Python library [48]. The scikit-learn platform in [49] is used to implement the performance evaluation. In this second stage, the motor current signal data undergo signal transformation.

Algorithm 1. Frequency occurrence plot-based (FOP)-CNN pseudocode implemented in Python programming language.

Step No.	Function
1	<i># initialization</i> import libraries load motor current time series signal
2	<i># data preprocessing</i> perform Fast Fourier Transform data statistical treatment frequency spectrum truncation
3	<i># frequency occurrence plot (FOP)</i> color-mapping of the frequency spectrum to generate FOPs
4	<i># data partitioning</i> splitting of FOPs dataset into n = 5 sets of training & testing datasets with 4:1 ratio
5	<i># model selection and training</i> for i = 1: n training of CNN using train FOP datasets determine train loss function
6	<i># model training and validation</i> testing of CNN using test datasets determine test loss function
7	<i># model overall performance evaluation</i> measure average classification accuracy measure average f-score, precision, and recall

Figure 10 shows a matrix of FOPs where all five motor fault conditions under five different loading conditions are compared. A motor operating frequency of 60 Hz with its sidebands is noticeable in all graphs, since all test motors have identical specifications. The healthy motor seems to have the clearest plots, while the rotor coil turn-to-turn and outer ring bearing damage faults tends to share similar messy plots. In addition, it can be observed that there are variations in the occurrence plots for each motor fault condition under different motor loadings. For example, the healthy motor tends to have a clearer plot when it is under no load or a full load. Similar observations can be made for bearing axis misalignment (Fault 1). Motors with other fault conditions tend to be messy at any loading condition. A smoother plot is expected for the healthy motor condition.

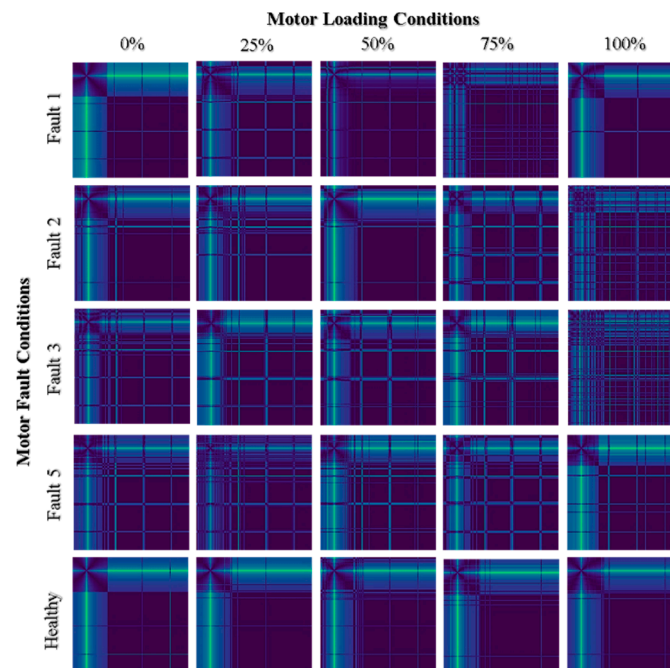


Figure 10. Sample frequency occurrence plots of five motor fault conditions under five loading conditions.

4.2.6. Computer Simulation Specifications

This supervised train–test simulation is performed in an Intel(R) Core (TM) i5-4590 CPU processor at 3.30 GHz with 16 GB of RAM installed memory. Since this study is concerned with supervised learning, computational speed is not the primary objective and can be sped up with newer specifications.

5. Results and Discussion

The CCE loss functions of five models according to five motor couple loadings are shown in Figure 11a–e. All five models tend to converge to a CCE loss value less than 0.25 at each epoch. The presence of early convergence occurred in some runs in the early epochs. Applying dropouts help the model to get away from early convergence, which is often believed to be caused by local optimum convergence. On the other hand, when the motor loading condition is not used as an input label, the model in Figure 12 tends to converge at a slightly higher loss value of 0.50 after five cross-validation runs. All of the models are still converging, but training them further seems to produce no significant changes in their performances and will only lead to a greater risk of overfitting.

Figure 13a shows the average loss function difference between the combined five models when the motor loading condition is used as an input label, and the model when the motor loading condition is not used. It is evident that FOP-CNN tends to predict better than when the load condition is not used. Figure 13b also presents their classification performances. This average classification accuracy graph of both cases shows similar converging performances. Both cases have identical FOP-CNN

parameters as shown in Table 2. Furthermore, this further verifies the observed graphical differences across motor loading conditions (see Figure 10). Simulating separate models, as performed in the first case model, may have avoided the difficulty caused by these differences. However, both cases still reach practical accuracies of 92% and 80%, respectively.

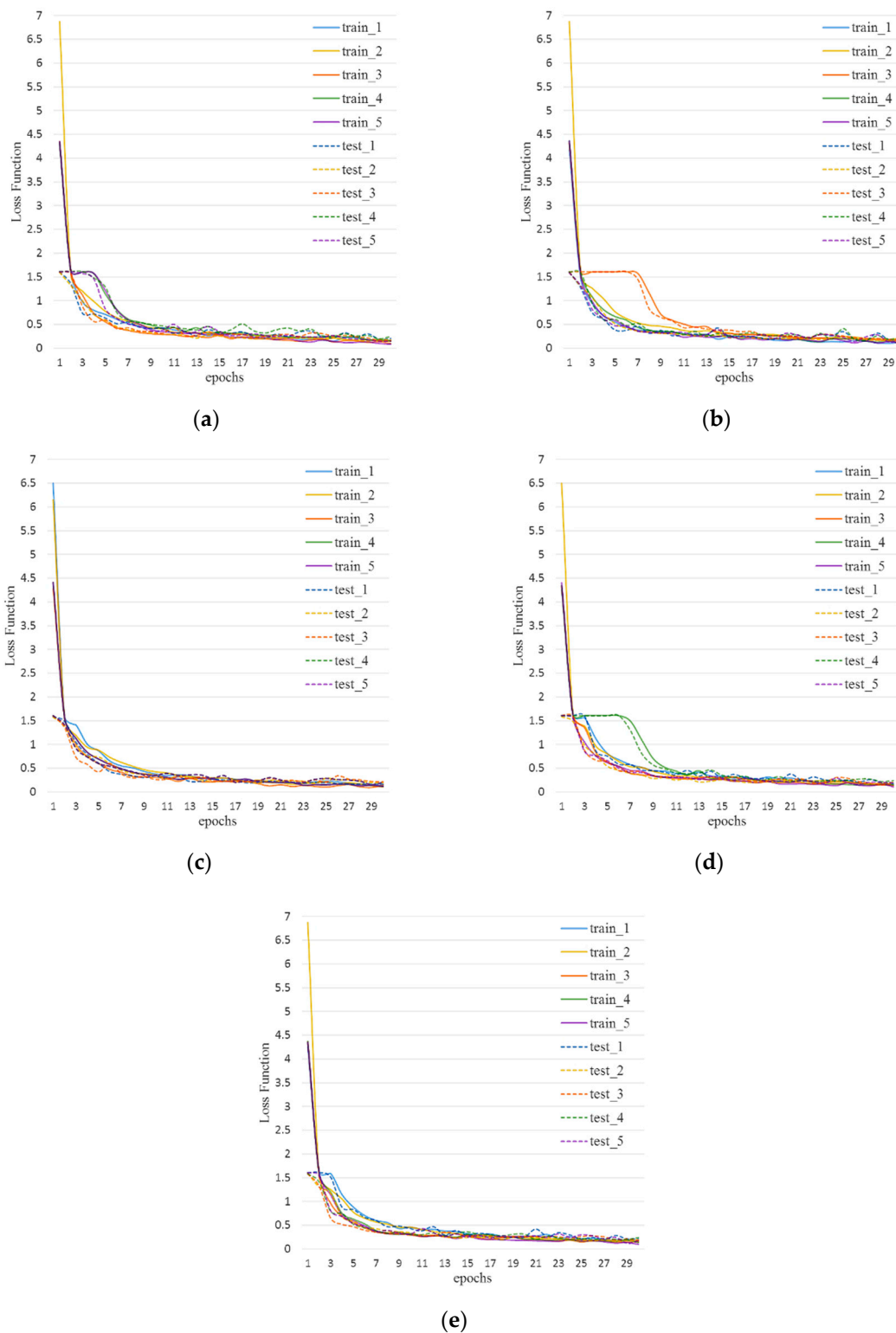


Figure 11. Five train–test categorical cross entropy (CCE) loss function graphs (a–e) of five motor loading conditions, respectively, with five train–test runs of cross-validation.

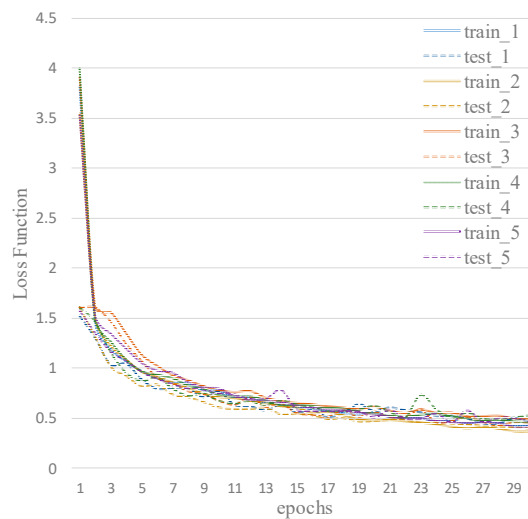


Figure 12. CCE loss graph with five train-test runs of cross-validation for the second case when motor load condition is ignored as input label.

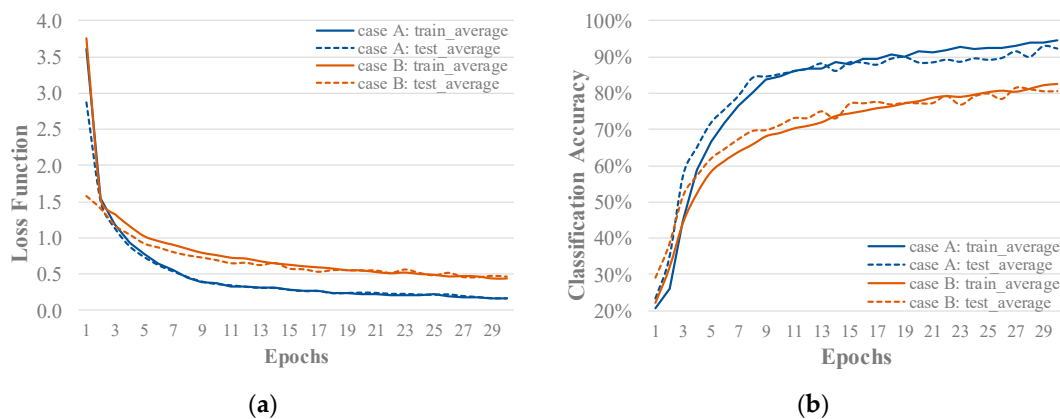


Figure 13. (a) An average loss function graph of two FOP-CNN case models when motor loading condition is used (Case A), and when it is not (Case B); (b) an average classification graph of two FOP-CNN case models when motor loading condition is used (Case A), and when it is not (Case B).

Table 2. FOP-CNN Hyper-parameters.

Batch Size	Epochs	Dropout 1	Dropout 2	Dropout 3	Dropout 4
Test data size	30	25	30	30	40

The classification reports of both cases are also taken. Each motor fault condition has 150 balance data. The first case model can classify bearing misalignment (Fault 1) with perfect recall (100%) and almost perfect precision (99.88%), as shown in Table 3. This seems intuitive since the energy loss caused by this fault has a direct effect on the motor’s current signature. It also precisely classifies healthy motors with 100% precision, and has little difficulty in recalling other faults with 92.22% recall. Two motor fault conditions—stator inter-turn fault (Fault 2) and broken rotor strip (Fault 3)—both have particularly good performances, with F-scores greater than 87.70% and 94.74%, respectively. However, outer bearing ring damage (Fault 4) performs the worst, with an F-score of 83.38%. When predicting Fault 4, there is a strong confusion in terms of predicting it as Fault 2, as shown in the confusion matrix of Figure 14a. Compared to other motor fault conditions, this fault may prove difficult to predict using FOP-CNN based only on the motor’s current signature. It seems that this fault is greatly confused with the motor stator inter-turn fault condition, with a 15% accuracy on average. The prediction of a healthy motor condition is also confused with Fault 4, with a 5.20% average accuracy. This means that

the synthetic physical damage inflicted to the outer ring bearing may have insignificant energy loss and may be due to friction. This small amount of energy loss may have no direct effect on and lead no changes in the motor’s current signature.

Table 3. Classification report when loading conditions are used.

Machine Fault	Precision (%)	Recall (%)	F1-Score (%)	Test Support
Fault 1	99.88	100.00	99.92	150
Fault 2	82.95	93.29	87.70	150
Fault 3	97.36	92.44	94.74	150
Fault 4	83.91	83.42	83.38	150
Healthy	100.00	92.22	96.13	150
Average	92.37	92.37	92.37	750

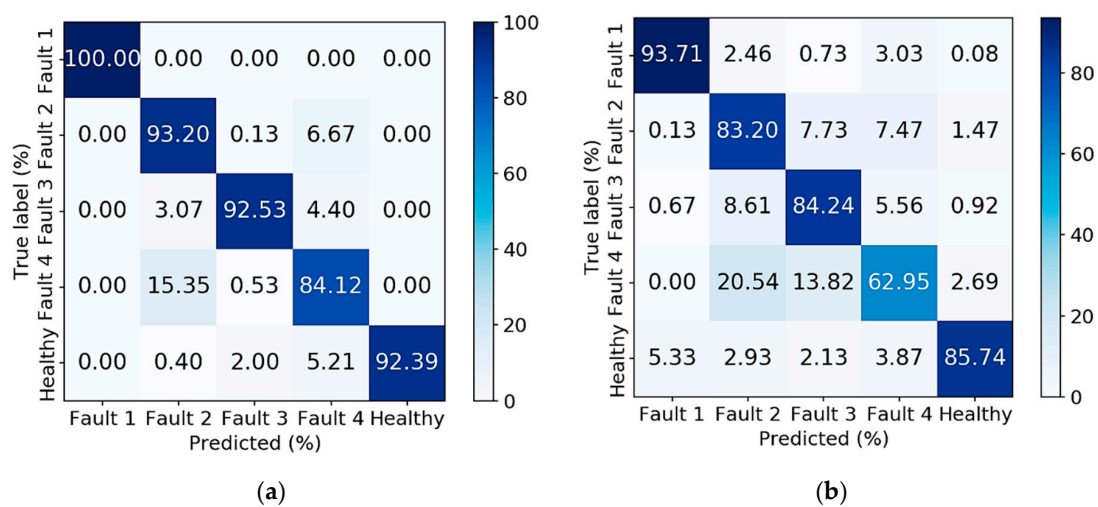


Figure 14. (a) Confusion matrix performance of FOP-CNN when motor load condition is used as an input label; and (b) confusion matrix performance when motor load condition is not used.

In the second case, the FOP-CNN model still performs like the previous case model, but with relatively lower classification accuracy. Predictions of bearing misalignment faults and the healthy state have better F-scores than the other fault conditions, as shown in Table 4. The classification of the bearing axis misaligned fault (Fault 1) also has the highest precision, with 93.40%. Fault 4 has the least recall, with only 62.60%, which seems to affect the overall performance. The model can classify all motor fault conditions with at least 83.20% accuracy, except when classifying outer bearing ring damage (Fault 4), where it only has a 62.95% accuracy, as shown in its confusion matrix in Figure 14b. When predicting Fault 4, this is often confused with the stator inter-turn fault (Fault 2) and broken rotor strip (Fault 3) conditions. This performance is relatively similar to the previous case model where prediction Fault 4 is worst-performing class.

Table 4. Classification report when loading conditions are not used.

Machine Fault	Precision (%)	Recall (%)	F1-Score (%)	Test Support
Fault 1	93.40	86.60	90.00	150
Fault 2	71.20	83.00	76.60	150
Fault 3	78.00	84.60	81.00	150
Fault 4	75.80	62.60	68.40	150
Healthy	87.00	85.80	86.40	150
Average	81.00	80.25	80.25	750

The performances of other motor fault detection algorithms are shown in Table 5. Comparatively, the proposed FOP-CNN performs competitively with the other known algorithms. It is important to take note that these algorithms have different case settings; thus, the comparison based on classification accuracy seems ungrounded and is difficult to justify. Moreover, note that the dataset used for FOP-CNN is identical to empirical wavelet transform convolutional neural network (EWT-CNN) [50], but with more samples collected.

Table 5. Comparison with other algorithms based on [50].

Methods	Testing Accuracy (%)
ANN [51]	81.8
DBN [51]	96.4
SVM [31]	89.8
Sparse filter [52]	92.2
ADCNN [53]	96.2
EWT-CNN [50]	97.4
FOP-CNN (<i>proposed</i>)	92.4

The difference between training and testing accuracies is the most common indicator to analyze the presence of overfitting—an especially important property to determine whether the train–test learning algorithm is robust and reliable. The higher the difference, the greater the learning generalization, thus making the method more unreliable. In Table 6, the proposed FOP-CNN is shown to be more robust than the best-performing algorithm, with a 13-fold lower learning difference. This means that the model tends to be more generalized, meaning that it can predict motor faults more reliably and accurately. It is commonly known in the literature that the higher the dataset, the more reliable and accurate the learning model tends to be, which is the case for the proposed algorithm.

Table 6. Comparison with the best performing algorithm.

Methods	Testing Accuracy (%)	Training Accuracy (%)	Learning Difference (%)	Dataset
EWT-CNN [50]	97.4	91	6.4	900
FOP-CNN (<i>proposed</i>)	92.4	92	0.4	3750

6. Conclusions

A novel motor fault diagnosis is successfully performed using only motor stator current signals and a frequency occurrence plot-based convolutional neural network (FOP-CNN). Five motor fault conditions—bearing axis deviation, stator coil turn-to-turn short circuit fault, broken rotor strip, outer bearing ring damage, and healthy motors—are considered and simulated under five motor loading conditions: 0%, 25%, 50%, 75% and 100% coupled loads. The diagnosis is also evaluated under two case scenarios—when the motor loading condition is considered as a label and when it is not. It was found that FOP-CNN tends to have a more robust performance when the motor load condition is available and is considered as an input label of the model. However, FOP-CNN still performed satisfactorily when the loading condition was not considered as an input label. Both cases provide users with an option of whether to install motor-coupled load monitoring or not.

FOP-CNN easily predicts the bearing axis deviation fault and healthy motor conditions. It can also satisfactorily predict stator coil turn-to-turn short circuit faults, broken rotor strips, and outer bearing ring damage faults. On the other hand, when the motor loading condition is not available, FOP-CNN can still predict all motor fault conditions satisfactorily, except the outer bearing ring damage fault. Future research on motor fault diagnosis based on other signals generated by vibration sensors and thermocouples can use FOP-CNN. This deep learning model also paves the way for new feature extraction techniques for time series applications.

Author Contributions: E.J.P. contributed to the conceptualization, data curation, formal analysis, investigation, methodology, project administration, validation, visualization, writing original draft, review, and editing. Y.-T.C. contributed to the conceptualization, data curation, formal analysis, investigation, methodology, software, and validation. H.-C.C. contributed to the funding acquisition, investigation, methodology, project administration, resources, and validation. C.-C.K. is the corresponding author and contributed to the conceptualization, data curation, formal analysis, funding acquisition, methodology, project administration, resources, review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Lee, J.; Wu, F.; Zhao, W.; Ghaffari, M.; Liao, L.; Siegel, D. Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications. *Mech. Syst. Signal Process.* **2014**, *42*, 314–334. [[CrossRef](#)]
- An, D.; Kim, N.H.; Choi, J. Practical options for selecting data-driven or physics-based prognostics algorithms with reviews. *Reliab. Eng. Syst. Saf.* **2015**, *133*, 223–236. [[CrossRef](#)]
- Zhao, R.; Yan, R.; Chen, Z.; Mao, K.; Wang, P.; Gao, R.X. Deep learning and its applications to machine health monitoring: A survey. *arXiv* **2015**, arXiv:1612.07640.
- Zhang, S.; Zhang, S.; Wang, B.; Habetler, T.G. Machine learning and deep learning algorithms for bearing fault diagnostics—A comprehensive review. *arXiv* **2019**, arXiv:1901.08247.
- Glowacz, A.; Glowacz, W.; Glowacz, Z.; Kozik, J. Early fault diagnosis of bearing and stator faults of the single-phase induction motor using acoustic signals. *Measurement* **2018**, *113*, 1–9. [[CrossRef](#)]
- Glowacz, A. Acoustic based fault diagnosis of three-phase induction motor. *Appl. Acoust.* **2018**, *137*, 82–89. [[CrossRef](#)]
- Sun, W.; Zhao, R.; Yan, R.; Member, S.; Shao, S.; Chen, X. Convolutional discriminative feature learning for induction motor fault diagnosis. *IEEE Trans. Ind. Inform.* **2017**, *13*, 1350–1359. [[CrossRef](#)]
- Jing, L.; Zhao, M.; Li, P.; Xu, X. A convolutional neural network-based feature learning and fault diagnosis method for the condition monitoring of gearbox. *Measurement* **2017**, *111*, 1–10. [[CrossRef](#)]
- Ma, M.; Sun, C.; Chen, X. Discriminative deep belief networks with ant colony optimization for health status assessment of machine. *IEEE Trans. Instrum. Meas.* **2017**, *66*, 3115–3125. [[CrossRef](#)]
- Shao, H.; Jiang, H.; Zhang, H.; Liang, T. Electric locomotive bearing fault diagnosis using a novel convolutional deep belief network. *IEEE Trans. Ind. Electron.* **2018**, *65*, 2727–2736. [[CrossRef](#)]
- Zhao, R.; Yan, R.; Wang, J.; Mao, K. Learning to monitor machine health with convolutional bi-directional LSTM networks. *Sensors* **2017**, *17*, 273. [[CrossRef](#)]
- Li, S.; Liu, G.; Tang, X.; Lu, J.; Hu, J. An ensemble deep convolutional neural network model with improved D-S evidence fusion for bearing fault diagnosis. *Sensors* **2017**, *17*, 1729. [[CrossRef](#)] [[PubMed](#)]
- Ren, L.; Sun, Y.; Cui, J.; Zhang, L. Bearing remaining useful life prediction based on deep autoencoder and deep neural networks. *J. Manuf. Syst.* **2018**, *48*, 71–77. [[CrossRef](#)]
- Zhao, R.; Wang, D.; Yan, R.; Mao, K.; Shen, F.; Wang, J. Machine health monitoring using local feature-based gated recurrent unit networks. *IEEE Trans. Ind. Electron.* **2018**, *65*, 1539–1548. [[CrossRef](#)]
- Ebrahimi, B.M.; Faiz, J. Feature extraction for short-circuit fault detection in permanent-magnet synchronous motors using stator-current monitoring. *IEEE Trans. Power Electron.* **2010**, *25*, 2673–2682. [[CrossRef](#)]
- Houssin, E.; Bouchikhi, E.; Choqueuse, V.; Benbouzid, M. Induction machine faults detection using stator current parametric spectral estimation. *Mech. Syst. Signal Process.* **2015**, *52–53*, 447–464.
- Singh, S.; Kumar, N. Detection of bearing faults in mechanical systems using stator current monitoring. *IEEE Trans. Ind. Inform.* **2017**, *13*, 1341–1349. [[CrossRef](#)]
- Schulz, R.; Verstockt, S.; Vermeiren, J.; Loccuferp, M.; Stockman, K.; van Hoecke, S. Thermal imaging for monitoring rolling element bearings. In Proceedings of the 12th International Conference on Quantitative InfraRed Thermography (QIRT 2014), Bordeaux, France, 7–11 July 2014; pp. 1–13.
- Janssens, O.; van de Walle, R.; Loccufer, M.; van Hoecke, S. Deep learning for infrared thermal image based machine health monitoring. *IEEE/ASME Trans. Mechatron.* **2018**, *23*, 151–159. [[CrossRef](#)]
- Singh, G.; Kumar, T.C.A.; Naikan, V.N.A. Induction motor inter turn fault detection using infrared thermographic analysis. *Infrared Phys. Technol.* **2016**, *77*, 277–282. [[CrossRef](#)]

21. Glowacz, A.; Glowacz, Z. Diagnostics of stator faults of the single-phase induction motor using thermal images, MoASoS and selected classifiers. *Measurement* **2016**, *93*, 86–93. [CrossRef]
22. Ali, M.Z.; Shabbir, M.N.S.K.; Liang, X.; Zhang, Y.; Hu, T. Machine learning based fault diagnosis for single- and multi-faults in induction motors using measured stator currents and vibration signals. *IEEE Trans. Ind. Appl.* **2019**, *55*, 2378–2391. [CrossRef]
23. Bouzida, A.; Touhami, O.; Ibtouen, R.; Belouchrani, A.; Fadel, M.; Rezzoug, A. Fault diagnosis in industrial induction machines through discrete wavelet transform. *IEEE Trans. Ind. Electron.* **2011**, *58*, 4385–4395. [CrossRef]
24. Cusidó, J.; Romeral, L.; Ortega, J.A.; Rosero, J.A.; Espinosa, A.G. Fault detection in induction machines using power spectral density in wavelet decomposition. *IEEE Trans. Ind. Electron.* **2008**, *55*, 633–643. [CrossRef]
25. Lou, X.; Loparo, K.A. Bearing fault diagnosis based on wavelet transform and fuzzy inference. *Mech. Syst. Signal Process.* **2004**, *18*, 1077–1095. [CrossRef]
26. El, M.; Benbouzid, H. A review of induction motors signature analysis as a medium for faults detection. *IEEE Trans. Ind. Electron.* **2000**, *47*, 984–993.
27. Strangas, E.G.; Aviyente, S.; Zaidi, S.S.H. Time-frequency analysis for efficient fault diagnosis and failure prognosis for interior permanent-magnet AC Motors. *IEEE Trans. Ind. Electron.* **2008**, *55*, 4191–4199. [CrossRef]
28. Thomson, W.T.; Fenger, M. Current signature analysis to detect induction motor faults. *IEEE Ind. Appl. Mag.* **2001**, *7*, 26–34. [CrossRef]
29. Georgoulas, G.; Loutas, T.; Stylios, C.D.; Kostopoulos, V. Bearing fault detection based on hybrid ensemble detector and empirical mode decomposition.pdf. *Mech. Syst. Signal Process.* **2013**, *41*, 510–525. [CrossRef]
30. Lei, Y.; Lin, J.; He, Z.; Zuo, M.J. A review on empirical mode decomposition in fault diagnosis of rotating machinery. *Mech. Syst. Signal Process.* **2013**, *35*, 108–126. [CrossRef]
31. Zhang, X.; Liang, Y.; Zhou, J. A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM. *Measurement* **2015**, *69*, 164–179. [CrossRef]
32. Espinosa, A.G.; Rosero, J.A.; Cusidó, J.; Romeral, L.; Ortega, J.A. Fault detection by means of Hilbert—Huang transform of the stator current in a PMSM with demagnetization. *IEEE Trans. Energy Convers.* **2010**, *25*, 312–318. [CrossRef]
33. Yu, X.; Ding, E.; Chen, C.; Liu, X.; Li, L. A novel characteristic frequency bands extraction method for automatic bearing fault diagnosis based on hilbert huang transform. *Sensors* **2015**, *15*, 27869–27893. [CrossRef] [PubMed]
34. Rai, V.K.; Mohanty, A.R.Ā. Bearing fault diagnosis using FFT of intrinsic mode functions in Hilbert—Huang transform. *Mech. Syst. Signal Process.* **2007**, *21*, 2607–2615. [CrossRef]
35. Konar, P.; Chattopadhyay, P. Multi-class fault diagnosis of induction motor using Hilbert and Wavelet transform. *Appl. Soft Comput. J.* **2015**, *30*, 341–352. [CrossRef]
36. Wang, D.; Miao, Q.; Fan, X.; Huang, H. Rolling element bearing fault detection using an improved combination of Hilbert and Wavelet transforms. *J. Mech. Sci. Technol.* **2009**, *23*, 3292–3301. [CrossRef]
37. Marwan, N.; Romano, M.C.; Thiel, M.; Kurths, J. Recurrence plots for the analysis of complex systems. *Phys. Rep.* **2007**, *438*, 237–329. [CrossRef]
38. Marwan, N. Historical review of recurrence plots. *Eur. Phys. J. Spec. Top.* **2008**, *164*, 1–11. [CrossRef]
39. Fukino, M.; Hirata, Y.; Aihara, K. Coarse-graining time series data: Recurrence plot of recurrence plots and its application for music. *Chaos* **2016**, *26*, 023116. [CrossRef]
40. Morinigo-sotelo, D.; Duque-perez, O.; Perez-alonso, M. Practical aspects of mixed-eccentricity detection in PWM voltage-source-inverter-fed induction motors. *IEEE Trans. Ind. Electron.* **2010**, *57*, 252–262. [CrossRef]
41. Didier, G.; Ternisien, E.; Caspary, O.; Razik, H. Fault detection of broken rotor bars in induction motor using a global fault index. *IEEE Trans. Ind. Appl.* **2006**, *42*, 79–88. [CrossRef]
42. Knight, A.M.; Bertani, S.P. Mechanical fault detection in a medium-sized induction motor using stator current monitoring. *IEEE Trans. Energy Convers.* **2005**, *20*, 753–760. [CrossRef]
43. Bracewell, R.N.; Bracewell, R.N. *The Fourier Transform and Its Applications*; McGraw-Hill: New York, NY, USA, 1986; Volume 31999.
44. Jones, E.; Oliphant, T.; Peterson, P. *Open Source Scientific Tools for Python*; 2014; SciPy. Available online: <http://www.scipy.org> (accessed on 20 August 2020).
45. Hunter, J.D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. [CrossRef]
46. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980.

47. Zhang, Z.; Sabuncu, M.R. Generalized Cross Entropy Loss for Training Deep Neural Networks with Noisy Labels. In Proceedings of the Advances in Neural Information Processing Systems 31 (NIPS 2018), Montreal, QC, Canada, 3–8 December 2018.
48. Chollet, F. *Keras*; 2015; Keras: Simple Flexible Powerful. Available online: <https://keras.io> (accessed on 20 August 2020).
49. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
50. Shao, H.; Jiang, H.; Zhang, X.; Niu, M. Rolling bearing fault diagnosis using an optimization deep belief network. *Meas. Sci. Technol.* **2015**, *26*, 115002. [[CrossRef](#)]
51. Lei, Y.; Jia, F.; Lin, J.; Xing, S.; Ding, S.X. An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data. *IEEE Trans. Ind. Electron.* **2016**, *63*, 3137–3147. [[CrossRef](#)]
52. Guo, X.; Chen, L.; Shen, C. Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis. *Meas. J. Int. Meas. Confed.* **2016**, *93*, 490–502. [[CrossRef](#)]
53. Hsueh, Y.M.; Ittangihal, V.R.; Wu, W.B.; Chang, H.C.; Kuo, C.C. Fault diagnosis system for induction motors by CNN using empirical wavelet transform. *Symmetry* **2019**, *11*, 1212. [[CrossRef](#)]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).