

Bearing Fault Diagnosis Using a Vector-Based Convolutional Fuzzy Neural Network

Cheng-Jian Lin ^{1,2,*} , Chun-Hui Lin ³ and Frank Lin ¹

¹ Department of Computer Science and Information Engineering, National Chin-Yi University of Technology, Taichung 411, Taiwan

² College of Intelligence, National Taichung University of Science and Technology, Taichung 404, Taiwan

³ Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan 701, Taiwan

* Correspondence: cjlin@ncut.edu.tw; Tel.: +886-4-23924505

Abstract: The spindle of a machine tool plays a key role in machining because the wear of a spindle might result in inaccurate production and decreased productivity. To understand the condition of a machine tool, a vector-based convolutional fuzzy neural network (vector-CFNN) was developed in this study to diagnose faults from signals. The developed vector-CFNN mainly comprises a feature extraction part and a classification part. The feature extraction phase encompasses the use of convolutional layers and pooling layers, while the classification phase is facilitated through the deployment of a fuzzy neural network. The fusion layer plays an important role by being placed between the feature extraction and classification parts. It combines the characteristics and then passes the feature information to the classification part to improve the model's performance. The developed vector-CFNN was experimentally evaluated against existing fusion methods; vector-CFNN required fewer parameters and achieved the highest average accuracy (99.84%) in fault diagnosis relative to conventional neural networks, fuzzy neural networks, and convolutional neural networks. Moreover, vector-CFNN achieved superior fault diagnosis using spindle vibration signals and required fewer parameters relative to its counterparts, indicating its feasibility for online spindle vibration monitoring.

Keywords: spindle vibration; vector convolutional neural network; feature fusion; fault diagnosis



Citation: Lin, C.-J.; Lin, C.-H.; Lin, F. Bearing Fault Diagnosis Using a Vector-Based Convolutional Fuzzy Neural Network. *Appl. Sci.* **2023**, *13*, 3337. <https://doi.org/10.3390/app13053337>

Academic Editor: Panagiotis G. Asteris

Received: 8 February 2023

Revised: 25 February 2023

Accepted: 3 March 2023

Published: 6 March 2023



Copyright: © 2023 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Machine tools are vital in modern industry, and the spindle impacts processing performance and the accuracy of machine tools. Unexpected failure of spindle bearings might result in financial loss. However, no universally accepted method exists for determining the condition of a machine tool spindle. Therefore, a robust method is required to detect bearing failures early to prevent costly repairs and machine downtime.

Researchers have proposed numerous methods for diagnosing faults from spindle vibration signals, such as statistical methods, conventional machine learning methods, and deep learning methods. Statistical methods include the fast Fourier transform [1], short-time Fourier transform, Wiener process [2], and Markov model methods [3]. Machine learning methods include those based on artificial neural networks (ANNs) [4], extreme learning machines [5], and neuro-fuzzy networks. For instance, a trained ANN model and an adaptive neuro-fuzzy inference system (ANFIS) were developed for diagnosing faults in bearings. The features used for training were extracted by time- and frequency-domain analysis of the vibration signals. The experimental results revealed that the ANFIS-based framework was superior to the ANN-based framework in diagnosing fault severity [6]. Other scholars combined a statistical method and a machine learning method for classifying the condition of milling tools. For instance, a discrete wavelet transform was employed to extract features from vibration signals and a decision tree was performed to select significant features. Subsequently, the features were identified by two support vector machine

(SVM) kernel functions: C-support vector and ν -support vector classifiers. The results revealed that C-support vector classification yields an accuracy of 94.5%, which is higher than that of ν -support vector classification [7]. In addition, fuzzy neural network (FNN) offers the advantages of both neural networks and fuzzy logic, making it a powerful hybrid tool. FNN integrates expert knowledge into the system and adopts human-like fuzzy reasoning, which makes the system outputs easier to interpret. A singular value decomposition and FNN were adopted to extract and diagnose the fault features in diesel engine crankshaft bearing [8]. While the aforementioned techniques exhibit greater versatility in comparison to models based on physical principles, most of them rely on manual feature extraction. Moreover, conventional machine learning methods cannot accurately interpret large amounts of data. In this respect, deep learning approaches, which can extract features automatically, have recently been employed in prediction and classification tasks.

Fault diagnosis methods based on deep learning have become a key research topic. In particular, convolutional neural networks (CNNs), which are capable of local feature connection and weight sharing, have been employed for fault diagnosis. For instance, a residual learning algorithm was built in a CNN model for alleviating the information missing during back-propagation. The proposed architecture is able to handle vibration signal with variable length, and minimal prior knowledge of fault diagnosis is needed. The results revealed that the proposed CNN model offers a novel fault diagnosis technique for rotating machinery [9]. Considering the feature extraction ability of deep learning lacks time delay information caused by faults occurrence, a combination of CNN and long short-term memory network (LSTM) model was developed in [10]. Finally, the results were superior to those of CNN, LSTM, ANN, K-nearest neighbor, and support vector machine models. Furthermore, a CNN framework was applied with the proposed slope and threshold adaptive tanh activation function to diagnose bearing faults. The relationship of non-linear features and input signal was established corresponding to the shape of activation function. The experimental results demonstrated that average accuracy in two bearing datasets reached 90.00% and 90.77%, respectively [11]. A one-dimensional (1D) CNN long short-term memory architecture was adopted for predicting the operational life of machining tools. In that architecture, prior to prediction, an ensemble discrete wavelets transform is applied to eliminate the noise of the vibration signals, and statistical feature extraction is performed based on time and frequency domains [12]. A signal-to-image spatial transformation technique was employed to generate grayscale images, thus decreasing the computation time. In addition, they adopted an attention mechanism for improving the accuracy and efficiency of the CNN, and their method achieved a classification accuracy of over 99% [13]. A squeeze-and-excitation-enabled CNN model was proposed for diagnosing faults in bearings. The model assigns learnable weights to each feature extraction channel, allowing the model to focus on the major features. The framework achieved a classification accuracy of over 99% in experiments [14]. A CNN model was integrated with an adaptive batch normalization algorithm to resolve the drawbacks of the high computation time and low versatility of deep CNN [15]. A distribution adaptive deep CNN model was applied for fault diagnosis based on a 1D CNN and achieved an accuracy of 90.29% [16]. A multiscale CNN for feature extraction from signal data was developed for diagnosing faults in bearings [17]. To solve the data imbalance problem, data augmentation techniques, such as permutation and time-warping, have been used to fuse the location features of the input data. The multiscale CNN not only learns better feature expression than conventional CNN by means of multiscale convolution operation but also reduces the number of parameters required and training time taken. Numerous studies have reported that deep learning approaches have outstanding feature extraction ability and high classification accuracy. However, approaches that require many learnable parameters require more powerful hardware to run, which impedes their applicability.

Literature reviews have revealed that spindle vibration signals can be used to diagnose bearing failures; however, those approaches have limitations, such as manual feature extraction and the existence of numerous parameters during model training. Therefore,

a proper and effective deep learning approach that is able to classify the bearing failures from different signal conditions is important to ensure the safety of the automation process. In this study, a vector-based convolutional FNN (vector-CFNN) was developed for fault diagnosis of spindle vibration signals. Moreover, a fusion layer that connects the feature extraction part and the classification part was introduced for fusing the characteristics of feature maps to enhance the model’s classification performance. The contributions of this study are as follows:

- The vector-CFNN model with fewer parameters was proposed to classify bearing failure by using spindle vibration signals.
- The fusion layer was introduced to improve the model’s classification accuracy by fusing the spatial and depth information of feature maps.
- Compared with other methods, vector-CFNN increases accuracy by up to 25% and requires only approximately 1000 learnable parameters.

The remainder of this paper is organized as follows. In Section 2, the structure of CNNs is described. In Section 3, the developed vector-CFNN is detailed. In Section 4, the results of experimental evaluations of the proposed vector-CFNN against existing approaches, such as ANN, FNN, conventional CNN, and CFNN, are presented. Finally, in Section 5, the paper is concluded.

2. Conventional CNN

This section presents LeNet-5 [18], a classical CNN model, as an illustrative example of a prototypical CNN architecture (Figure 1). The LeNet-5 model mainly comprises an input/output layer, two convolutional layers, two pooling layers, two fully connected layers, and a flattening layer that precedes the final fully connected layer. These layers are detailed as follows.

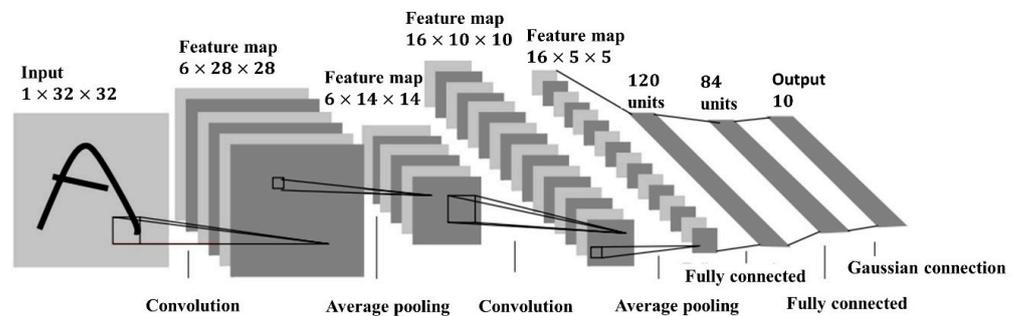


Figure 1. Architecture of LeNet-5.

2.1. Convolutional Layer

A convolutional layer [19] utilizes a sliding-window mechanism to perform the dot product operation with the convolution kernel. The convolution results are then calculated by a sigmoid function to obtain feature maps. The convolutional operation [20] can be expressed as follows:

$$O_{r,c}^l = s \left(\sum_{k=1}^n \sum_{i=1}^{k_h} \sum_{j=1}^{k_w} I_{(r-k_h+i,c-k_w+j)}^k \times W_{i,j}^{kl} + b^l \right) \tag{1}$$

$$s(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

where $O_{r,c}^l$ represents the l th feature map; c and r are the column and row of the feature map, respectively; n denotes the number of input channels; k_w and k_h are the width and height, respectively, of a kernel; $W_{i,j}^k$ is the weight of the i th row and j th column convolution kernel in the k th channel; $I_{i,j}^k$ is the input of the i th row and j th column in the k th channel;

b is the bias; and $s(x)$ is the sigmoid activation function. In a deep network, the gradient of the sigmoid function becomes smaller as $|x|$ increases. Therefore, in vector-CFNN, the rectified linear unit (ReLU) function is used in place of the sigmoid function. The ReLU function can be expressed as follows:

$$R(x) = \max(0, x) \tag{3}$$

2.2. Pooling Layer

A pooling layer can conduct in only one channel at a time; however, it condenses the activation levels locally in each channel [21]. Pooling not only reduces the number of dimensions and volume of computation but also decreases the possibility of overfitting because fewer trainable parameters are required, resulting in greater tolerance and reduced distortion. Max pooling and average pooling can be expressed as follows:

$$P_{r,c} = \begin{cases} \text{Max pooling} \left(\text{MAX}(I_{i,j}) \middle| \begin{matrix} r \leq i < r + P_h \\ c \leq j < c + P_w \end{matrix} \right) \\ \text{Average pooling} \left(\sum_{i=r}^{r+P_h} \sum_{j=c}^{c+P_w} I_{i,j} / (P_w \times P_h) \right) \end{cases} \tag{4}$$

where $P_{r,c}$ is the result of the pooling calculation corresponding to the feature map; $I_{i,j}$ is the input feature map; and P_h and P_w are the height and width, respectively, of the pooling size.

2.3. Flattening and Fully Connected Layer

The flattening process involves a conversion of a 2D feature map into a 1D array by concatenating the constituent feature vectors. This long vector is then linked to the fully connected layer, where every neuron is connected to every other neuron, resulting in a high number of parameters in the CNN model.

3. Vector-CFNN

In this section, a prototypical CFNN architecture is described prior to the developed vector-CFNN. Furthermore, various fusion methods used in the fusion layer are discussed.

3.1. CFNN

The structure of a CFNN [22], as depicted in Figure 2, can be divided into two parts: feature extraction and classification. The feature extraction part is similar to a conventional CNN and comprises convolutional layers and pooling layers. The classification part comprises an FNN instead of a fully connected layer. In summary, fuzzy logic provides a symbolic representation of syntactic rules, whereas neurons offer the capability to perform logical inferences, and the numerical membership function serves as an interface between the two. The fuzzy rule used in CFNN defines a nonlinear mapping involving linguistic variables that encapsulate localized process knowledge. Each neuron can be interpreted as a specific fuzzy membership function, while each link represents the weight of the fuzzy rule, which means CFNN is able to use fewer parameters in comparison with conventional neural network [23].

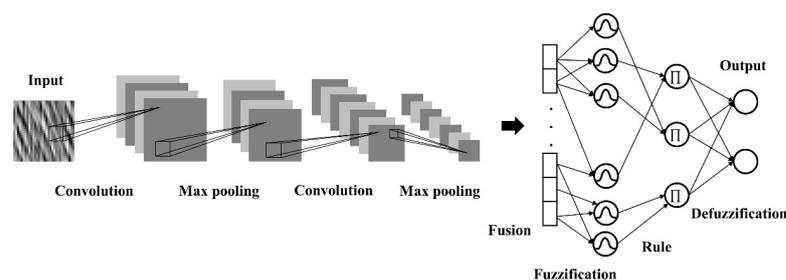


Figure 2. Architecture of CFNN.

3.2. Fuzzification Layer

Fuzzy operation is performed in this layer by using IF-THEN rules. The IF-THEN rule is determined as follows:

$$R_j = IF u_1 \text{ is } S_{1j} \dots \wedge u_n \text{ is } S_{nj}, THEN y_j = w_j \tag{5}$$

where R_j represents a fuzzy rule; S_{ij} represents fuzzy sets; and w_j is a zero-order Takagi–Sugeno–Kang weight. A fuzzy set is governed by a Gaussian membership function that can be expressed as follows:

$$S_{ij} = exp \left\{ \frac{-(u_i - m_{ij})^2}{2\sigma_{ij}^2} \right\} \tag{6}$$

where $exp(\cdot)$ is the exponential function and m_{ij} and σ_{ij} are the mean and standard deviation of a fuzzy set S_{ij} , respectively.

3.3. Rule Layer

The firing strength of a fuzzy rule is determined via a product operation of each membership function, and the following equation formally expresses the algebraic product operation:

$$R_j = \prod_{i=1}^n M_{ij} \tag{7}$$

3.4. Defuzzification Layer

This layer carries out a defuzzification operation, in which the crisp values are computed by aggregating the results of each rule. Subsequently, to determine the output probability for classification, the softmax function is applied. The defuzzification and softmax equations are, respectively, presented as follows:

$$d_i = \sum_{j=1}^r R_j w_{ij} \tag{8}$$

$$y_i = Softmax(d_i) = f(d_i) \tag{9}$$

$$f(d_i) = \frac{e^{d_i}}{\sum_{j=1}^n e^{d_j}} \tag{10}$$

where d is the crisp output computed by the defuzzification operation; R_j denotes the firing strength of j th fuzzy rule; w represents the learnable weight; y_i is i th class probability; and $f(\cdot)$ is the function for normalizing the value to be within 0 and 1.

3.5. Vector-Based Convolution

Although the CFNN can reduce many learnable parameters by replacing fully connected layers, matrix kernels can also produce a great deal of parameter redundancy by the convolution operation [24]. Vector-CFNN employs a vector-based convolution instead of a conventional convolution. For instance, a convolutional layer with a kernel size of $(k \times k, c)$ can be convoluted by a kernel size of $(k \times 1, m)$ followed by a kernel size of $(1 \times k, c)$. Figure 3 presents a schematic of a vector-based convolution. It can effectively reduce parameters and floating-point operations per second without suffering the loss of recognition accuracy rate.

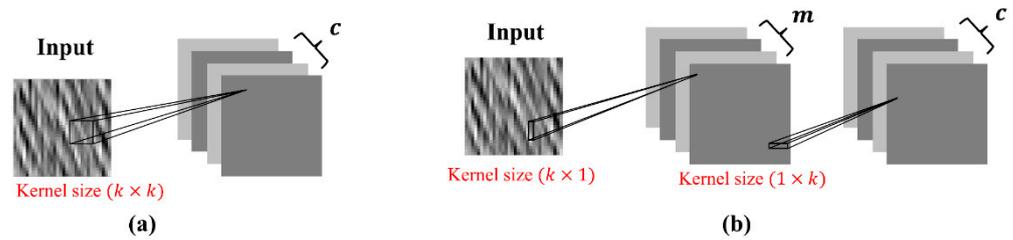


Figure 3. Schematic of (a) conventional convolution and (b) vector-based convolution.

3.6. Fusion Layer

The fusion layer is mainly used to characterize features from feature maps to increase classification accuracy. In this study, five fusion methods (global average pooling (GAP) [25], global max pooling (GMP), channel average pooling (CAP), channel max pooling (CMP) [26], and network mapping) were adopted. GAP inherently reflects the spatial information of feature maps, which indicates the robustness of the interpretation of input images. By contrast, GMP can enhance the local spatial information of feature maps, which is also important in classification tasks. Furthermore, the generalization ability of a network can be significantly improved by learning the discriminative features among channels. Finally, network mapping assigns each feature element a weight and fuses all elements into a new feature. Network mapping can be expressed using the following equation:

$$f_j = \sum_{i=1}^n w_{zi} \times x_i \quad (11)$$

where x is a feature element; f is the network mapping result; w is the feature's weight; and n is the number of features.

4. Experimental Results

4.1. Data Acquisition

The bearing data provided by Case Western Reserve University were used in experiments [27]. According to the specifications of the bearing data, the faults were intentionally introduced to the motor bearings via an electrical discharge machining procedure. The faults were seeded at the inner raceway, outer raceway, or rolling element (i.e., ball). The actual test conditions of the motor and the bearing fault status were documented for each experiment. The apparatus is depicted in Figure 4, and the 12-kHz vibration signals of normal bearing, fault in the inner race, fault in the ball, and fault in the outer race are illustrated in Figure 5a–d, respectively. The various signal conditions were applied as input data in different deep learning models (ANN, FNN, CNN, CFNN, and vector-CFNN) and the output was the fault classification results. Finally, the accuracy and the parameters used for each model were compared with the vector-CFNN model.

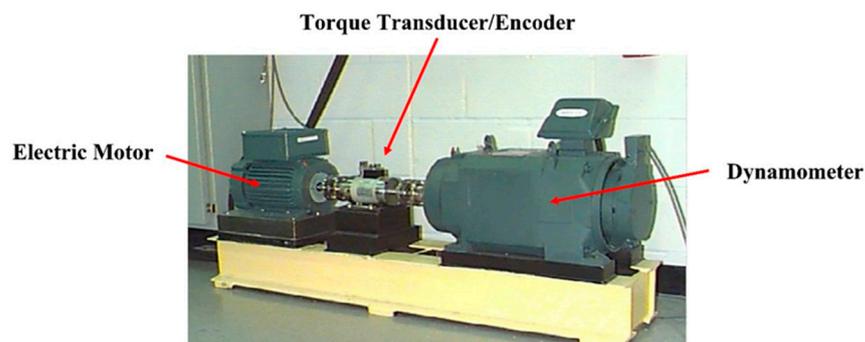


Figure 4. Apparatus used in Case Western Reserve University.

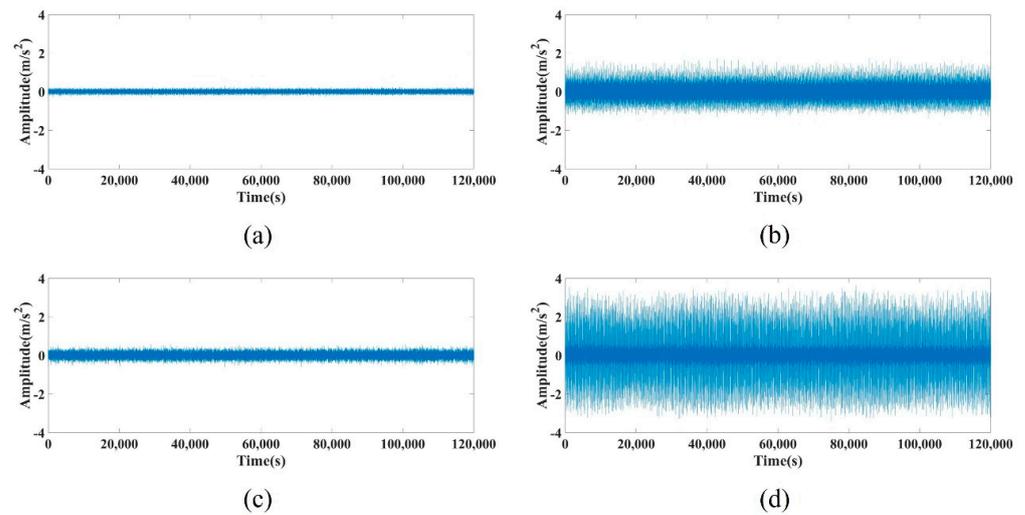


Figure 5. Vibration signals of (a) normal bearing and fault in (b) inner race, (c) ball, and (d) outer race under 1797-rpm motor speed.

4.2. Data Pre-Processing

To determine the condition of a bearing, a total of sixty-four 12-kHz vibration signals were pre-processed to establish a bearing fault diagnosis model. Each signal was split into several nonoverlapping fragments containing 1024 points. Subsequently, each fragment was formed into a 32×32 image as the CNN input data. The data pre-processing procedure is illustrated in Figure 6. The total number of fragments in each category is listed in Table 1, and the corresponding fragments and processed images are presented in Figure 7.

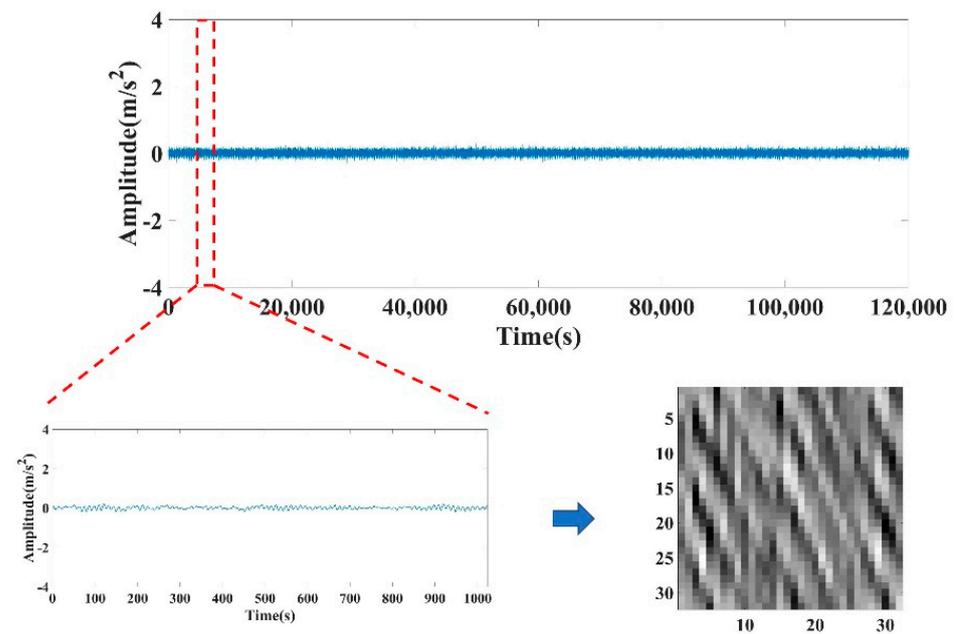


Figure 6. Data pre-processing procedure.

Table 1. Total number of fragments in each category.

Category	Numbers of Fragments
Normal bearing	1657
Fault in inner race	1893
Fault in ball	1894
Fault in outer race	3324
Total	8768

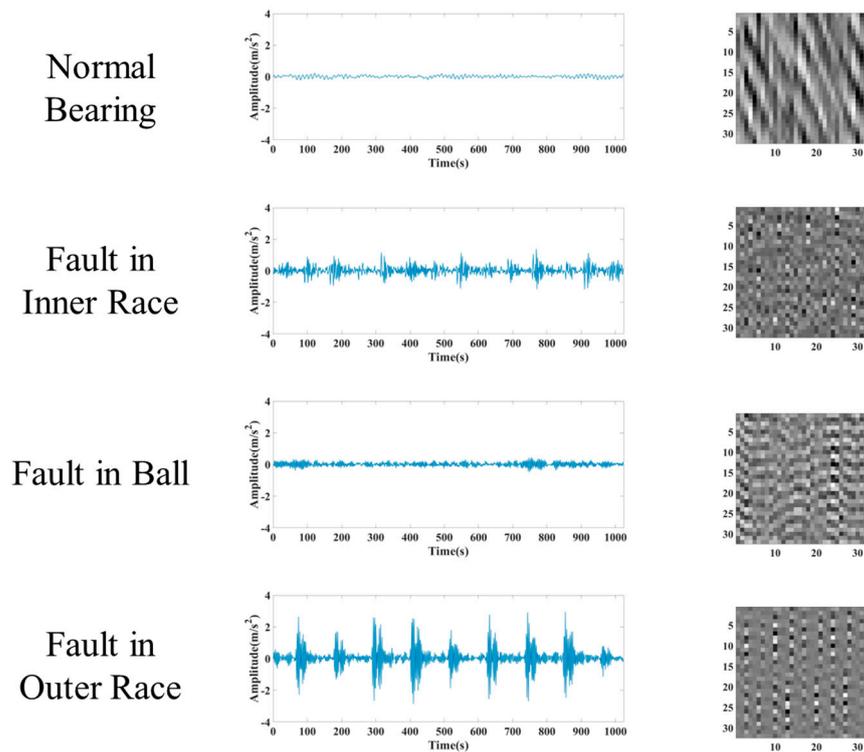


Figure 7. A processed fragment and the corresponding images in each category.

4.3. Evaluation of Fault Diagnosis

The vibration signals were randomly divided into training (70%) and testing (30%) datasets to assess the stability and generalization capabilities of the vector-CFNN. Additionally, during model training, the training data (70% of datasets) were split into 8:2 for model training and validating. Following that, k-fold cross-validation was employed to validate the efficiency of the model across various subsets. The architecture of the developed vector-CFNN is displayed in Table 2, and the layer of the fusion method was adjusted corresponding to GAP, GMP, CAP, CMP, and network mapping. Finally, the performance of the vector-CFNN was compared with that of ANN, FNN, conventional CNN, and CFNN. The total parameters used in the models are provided and improvement in the accuracy of different fusion methods is also presented in Table 3.

Table 2. The architecture of vector-CFNN.

	Layer	Kernel Size
Feature extraction	Convolution1	(9, 1), channel = 6
	Convolution2	(1, 9), channel = 6
	Max_pooling1	(2, 2)
	Convolution3	(9, 1), channel = 6
	Convolution4	(1, 9), channel = 6
	Max_pooling2	(2, 2)
Classification	Fusion Method Fuzzy Layer Output Layer	

Table 3. Accuracy of fault diagnosis and parameters used in each model.

Model	Fusion Method	Lowest Accuracy	Highest Accuracy	Average Accuracy	Parameter
ANN	-	71.77%	77.32%	75.78%	533,008
FNN	-	82.38%	93.98%	88.76%	132,368
CNN	-	91.62%	93.77%	92.43%	22,854
CFNN	-	98.68%	98.8%	98.80%	3890
	GAP	95.92%	99.86%	97.68%	3578
	GMP	95.61%	99.75%	98.06%	3578
	CAP	91.82%	99.32%	96.17%	3674
	CMP	95.57%	99.58%	98.11%	3674
	Network Mapping	99.25%	99.67%	99.49%	6106
Vector-CFNN	-	97.76%	99.61%	98.89%	1910
	GAP	99.81%	99.91%	99.84%	1214
	GMP	98.73%	99.39%	99.11%	1214
	CAP	99.52%	99.79%	99.68%	1502
	CMP	99.63%	99.78%	99.71%	1502
	Network Mapping	99.61	99.82	99.69%	5278

The average accuracy of ANN and FNN were 75.78% and 88.76%, respectively; however, the average accuracy of CNN reached 92.43%, which revealed a 3.67% improvement. The images formed by the time domain signals (Figure 7) have been clearly distinguished by their image pattern, as the CNN series algorithms showed superior performance than those of ANN and FNN because of their capacities for strong correlation and characteristic extraction in image classification tasks. By contrast, CFNN, a combination of CNN and FNN, increased the classification ability, and achieved around 96~99% average accuracy when using different fusion methods due to the fuzzy logic's ability to mimic human reasoning. Although network mapping fusion method showed 99.49% average accuracy in CFNN, it needed more parameters than other fusion methods. In order to decrease the parameter usage in a model, the developed vector-CFNN was evaluated against CFNN; the accuracy of vector-CFNN improved by 2.16%, 1.05%, 3.51%, 1.60%, and 0.20% while using GAP, GMP, CAP, CMP, and network mapping fusion methods, respectively. Moreover, vector-CFNN reduced the number of parameters by 66.07% in comparison with CFNN.

Additionally, various sizes of damage diameter classification (including 0.007, 0.014, and 0.021 inches) provided by the bearing dataset of Western Reserve University were examined by vector-CFNN. The lowest, highest, and average accuracy were 95.13%, 98.02%, and 96.69%, respectively, and the confusion matrix is given in Figure 8. Overall, while diagnosing the bearing faults, without the addition of a fusion method, the average accuracy of vector-CFNN was 98.89%, and with fusion methods, the average accuracy of vector-CFNN was over 99%, which was higher than that of other models. Moreover, in classifying bearing faults based on different fault diameters, the vector-CFNN can also reach over 95% average accuracy, which showed the capability of fault diagnosis. Specifically, the

advantages of vector-CFNN are that few parameters are needed while retaining a high accuracy rate, and the fusion methods can greatly improve the network classification performance in diagnosing faults.



Figure 8. A confusion matrix of vector-CFNN for various sizes of damage diameter classification.

5. Conclusions

This study developed vector-CFNN for diagnosing faults in spindles from vibration signals. The use of vector-based convolution and FNN reduced the number of parameters required during model learning and improved classification accuracy. Furthermore, five fusion methods were used to improve the feature characterization ability of the model. In experiments, vector-CFNN with GAP required only approximately 1000 parameters and achieved the highest average accuracy (99.84%) when evaluated against ANN, FNN, CNN, and CFNN. In future research, the model can be embedded into portable machines owing to the fewer parameters required in vector-CFNN; moreover, various kernel sizes and bearing motor speeds will be considered. Furthermore, different bearing databases can be used to verify the robustness of the proposed model.

Author Contributions: Conceptualization, C.-J.L.; methodology, C.-J.L. and C.-H.L.; software, C.-J.L., C.-H.L. and F.L.; data curation, C.-H.L. and F.L.; writing—original draft preparation, C.-J.L. and C.-H.L.; funding acquisition, C.-J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Science and Technology of the Republic of China, grant number MOST 111-2218-E-005-009.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The authors confirm that the data supporting the findings of this study are available within the paper.

Acknowledgments: The authors would like to thank the funding support of the Ministry of Science and Technology of the Republic of China.

Conflicts of Interest: The authors declare that they have no conflict of interest.

References

- Garzón, C.V.; Moncayo, G.B.; Alcantara, D.H.; Morales-Menendez, R. Fault Detection in Spindles Using Wavelets—State of the Art. *IFAC Pap.* **2018**, *1*, 450–455. [[CrossRef](#)]
- Zhai, Q.; Ye, Z.S. RUL Prediction of Deteriorating Products Using an Adaptive Wiener Process Model. *IEEE Trans. Ind. Inform.* **2017**, *13*, 2911–2921. [[CrossRef](#)]

3. Yuwono, M.; Qin, Y.; Zhou, J.; Guo, Y.; Celler, B.G.; Su, S.W. 2 Automatic Bearing Fault Diagnosis Using Particle Swarm Clustering and Hidden Markov Model. *Eng. Appl. Artif. Intell.* **2016**, *47*, 88–100. [CrossRef]
4. Lall, P.; Deshpande, S.; Nguyen, L. ANN Based RUL Assessment for Copper-Aluminum Wirebonds Subjected to Harsh Environments. In Proceedings of the 2016 IEEE International Conference on Prognostics and Health Management (ICPHM), Ottawa, ON, Canada, 20–22 June 2016; pp. 1–10.
5. Li, X. Remaining Useful Life Prediction of Bearings Using Fuzzy Multimodal Extreme Learning Regression. In Proceedings of the 2017 International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC), Shanghai, China, 16–18 August 2017; pp. 499–503.
6. Ertunç, H.M.; Ocak, H.; Aliustaoglu, C. ANN- and ANFTIS-Based Multi-Staged Decision Algorithm for the Detection and Diagnosis of Bearing Faults. *Neural Comput. Appl.* **2013**, *22*, 435–446. [CrossRef]
7. Madhusudana, C.K.; Gangadhar, N.; Kumar, H.K.; Narendranath, S. Use of Discrete Wavelet Features and Support Vector Machine for Fault Diagnosis of Face Milling Tool. *SDHM Struct. Durab. Health Monit.* **2018**, *12*, 111–127.
8. Jingbo, G.; Yifan, H. Research on Fault Diagnosis Based on Singular Value Decomposition and Fuzzy Neural Network. *Shock. Vib.* **2018**, *2018*, 8218657.
9. Zhang, W.; Li, X.; Ding, Q. Deep residual learning-based fault diagnosis method for rotating machinery. *ISA Trans.* **2019**, *95*, 295–305. [CrossRef]
10. Huang, T.; Zhang, Q.; Tang, X.; Zhao, S.; Lu, X. A novel fault diagnosis method based on CNN and LSTM and its application in fault diagnosis for complex systems. *Artif. Intell. Rev.* **2022**, *55*, 1289–1315. [CrossRef]
11. Zhang, T.; Liu, S.; Wei, Y.; Zhang, H. A novel feature adaptive extraction method based on deep learning for bearing fault diagnosis. *Measurement* **2021**, *185*, 110030. [CrossRef]
12. Niu, J.; Liu, C.; Zhang, L.; Liao, Y. Remaining Useful Life Prediction of Machining Tools by 1D-CNN LSTM Network. In Proceedings of the 2019 IEEE Symposium Series on Computational Intelligence (SSCI), Xiamen, China, 6–9 December 2019; pp. 1056–1063.
13. Wang, H.; Xu, J.; Yan, R.; Sun, C.; Chen, X. Intelligent Bearing Fault Diagnosis Using Multi-Head Attention-Based CNN. *Procedia Manuf.* **2020**, *49*, 112–118. [CrossRef]
14. Wang, H.; Xu, J.W.; Yan, R.Q.; Gao, R.X. A New Intelligent Bearing Fault Diagnosis Method Using SDP Representation and SE-CNN. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 2377–2389. [CrossRef]
15. Fu, C.; Lv, Q.; Lin, H.C. Development of Deep Convolutional Neural Network with Adaptive Batch Normalization Algorithm for Bearing Fault Diagnosis. *Shock. Vib.* **2020**, *2020*, 8837958. [CrossRef]
16. Yuying, Z.; Yi, L.; Jing, Z. Research on Numerical Control Machine Fault Diagnosis Based on Distribution Adaptive One-Dimensional Convolutional Neural Network. In Proceedings of the Journal of Physics: Conference Series, Hangzhou, China, 14–16 May 2021.
17. Han, S.; Oh, S.; Jeong, J. Bearing Fault Diagnosis Based on Multiscale Convolutional Neural Network Using Data Augmentation. *J. Sens.* **2021**, *2021*, 6699637. [CrossRef]
18. Lecun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-Based Learning Applied to Document Recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [CrossRef]
19. Lin, C.-J.; Jhang, J.-Y. Bearing Fault Diagnosis Using a Grad-CAM-Based Convolutional Neuro-Fuzzy Network. *Mathematics* **2021**, *9*, 1502. [CrossRef]
20. Lin, C.-J.; Lin, C.-H.; Wang, S.-H. Integrated Image Sensor and Light Convolutional Neural Network for Image Classification. *Math. Probl. Eng.* **2021**, *2021*, 5573031. [CrossRef]
21. Galety, M.G.; Mukthar, F.H.A.; Maaroo, R.J.; Rofoo, F. Deep Neural Network Concepts for Classification Using Convolutional Neural Network: A Systematic Review and Evaluation. *Technium* **2021**, *3*, 58–70. [CrossRef]
22. Lin, C.-J.; Jhang, J.-Y.; Young, K.-Y. Parameter Selection and Optimization of an Intelligent Ultrasonic-Assisted Grinding System for SiC Ceramics. *IEEE Access* **2020**, *8*, 195721–195732. [CrossRef]
23. Ellery, A. Artificial intelligence through symbolic connectionism—A biomimetic rapprochement. In *Biomimetic Technologies: Actuators, Robotics & Integrated Systems*; Ngo, D., Ed.; Woodhead Publishing: Cambridge, UK, 2015.
24. Ou, J.; Li, Y. Vector-Kernel Convolutional Neural Networks. *Neurocomputing* **2019**, *330*, 253–258. [CrossRef]
25. Lin, M.; Chen, Q.; Yan, S. Network in Network. *arXiv* **2013**, arXiv:1312.4400. [CrossRef]
26. Ma, Z.; Chang, D.; Xie, J.; Ding, Y.; Wen, S.; Li, X.; Si, Z.; Guo, J. Fine-Grained Vehicle Classification with Channel Max Pooling Modified CNNs. *IEEE Trans. Veh. Technol.* **2019**, *68*, 3224–3233. [CrossRef]
27. Bearing Data Center: Seeded Fault Test Data. Available online: <https://engineering.case.edu/bearingdatacenter/welcome> (accessed on 2 March 2023).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.