International Journal of Basic and Applied Sciences, 15 (1) (2026) 68-78

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/08hkh236
Research paper

Convolutional Neural Network with Mix-Up Data Augmentation for Ball Bearing Fault Diagnosis

Dillip Kumar Baral [1], Krishna Chandra Patra [2] *, Ambuja Behera [3], Nameet Kumar Sethy [4], Dhiren Kumar Behera [5], Rabinarayan Sethi [5]

[1] *Research Scholar, Department of Architecture, Veer Surendra Sai University of Technology, Burla, Odisha, India*
[2] *PhD research scholar, BPUT University, Rourkela, Odisha, India*
[3] *GITA Autonomous College, BPUT, Odisha*
[4] *Research Scholar, Indra Gandhi Institute of Technology, Sarang, BPUT*
[5] *Mechanical Engineering, IGIT, Sarang, Odisha, India*
**Corresponding author E-mail: kcpmechcvrce@gmail.com*

Received: July 2, 2025, Accepted: August 13, 2025, Published: January ,12 2026

Abstract

Ball bearings are vital for rotating machinery, requiring reliable fault diagnosis. Traditional approaches rely on manual feature extraction, requiring specialised expertise. Deep learning reduces human input but struggles with capturing global input context, integrating statistical features, and computational costs. This work introduces a CNN-based fault diagnosis method with Mix-up augmentation. Vibration signals are transformed into 2D time-frequency images via Continuous Wavelet Transform (CWT) to retain temporal-spectral information. Mix-up enhances dataset diversity, improving model robustness. CNNs then classify fault type and severity using these augmented inputs. Evaluat-ed on experimental and CWRU datasets, the approach surpasses state-of-the-art methods in accuracy and stability. Combining CWT's de-tailed analysis, Mix-up's data enrichment, and CNNs' automated feature extraction resolves prior limitations, delivering an efficient solution for industrial fault detection. The framework ensures reliable, resource-effective automation, advancing predictive maintenance in rotating machinery.

Keywords: *Convolutional Neural Networks; Continuous Wavelet Transform; Fault Diagnostics; Data Augmentation; Scalogram.*

1. Introduction

Some industrial equipment must operate in severe environments on a continual basis, and key components such as bearings frequently fail. Ball bearings are one of the most important components of many industrial machines, and their condition has a significant impact on how well they perform [1]. Furthermore, the quickly expanding sectors and rising demand for hybrid and electric vehicles predict that this rate of utilisation will continue to rise. During the lifetime of an electrical machine, electrical or mechanical failures will arise [2]. As a result, research into ball bearing defect diagnosis technology is critical for the sake of manufacturing process safety and avoiding financial losses. Many common intelligent approaches have been effectively applied in fault diagnostic research because of the development of machine learning. Its primary applications are the Extraction of signal features and fault classification [3].

Bearing vibration signals normally contain enough fault details; however, they are mostly nonlinear and nonstationary. As a result, signal feature extraction is critical [4]. Time-frequency based signal analysis is a strong signal processing method that can study both the time and frequency domains simultaneously, as in Empirical Mode Decomposition (EMD) [5], Wavelet Transform (WT) [6], as well as the Short-Time Fourier Transform (STFT) [7]. EMD can break down the signal into intrinsic modal function components with varying sizes in an adaptive manner. In this approach, however, there is a modal confusion issue. Despite the fact that STFT can do a study of the frequency and time of a signal, its time resolution is fixed, so it is unable to accurately reflect a change in the vibration signal that occurs suddenly [8]. The wavelet transform (WT) is a type of time-frequency assessment in which the time frame shrinks as the signal frequency rises and vice versa. WT is frequently used because it expands STFT and successfully compensates for its flaws. In the realm of fault diagnosis, Yan et al. [9] summarised the use of Continuous Wavelet Transform (CWT), Second-Generation Wavelet Transform (SGWT), and Wavelet Packet Transform (WPT) To accurately diagnose bearing faults, researchers investigated real-time bearing vibration signal fault detection approaches, including traditional methods, machine learning, and deep learning methods. Traditional methods necessitate feature extraction, dimension reduction, and categorisation of the bearing vibration signal, all of which result in sophisticated mathematical models. Machine learning (ML) methods such as k-nearest neighbours (KNNs) [10], adaptive neuro-fuzzy inference system (ANFIS) [11], fuzzy cognitive networks (FCNs) [12], the multi-agent system (MAS) approach using intelligent classifiers [13], and the support vector

machine (SVM) [14] have been introduced to automate the process. Multi-layer neural networks are used in the deep learning model to extrapolate and comprehend the signal input's specifics. Deep learning can solve complex, multi-dimensional issues in enormous volumes of data in a way that shallow learning can't [15]. Many deep learning models, such as the Long Short-Term Memory (LSTM) [16], Gated Recurrent Unit Network (GRUN) [17], Deep Belief Network (DBN) [18], Convolutional Neural Network (CNN) and Auto-encoder using wavelet [19] have been applied to fault diagnosis research due to their great productivity, adaptability, and universality. In the investigation of ball bearing fault detection, it has received increased attention. A deep CNN structural model is created [20] that can classify ball bearing defects automatically. [21] proposes a method for diagnosing rolling bearing faults using CNN and Cyclic Spectral Coherence (CSCoh), which enhances Performance in detecting faults. In [22], Hierarchical Symbolic Analysis (HSA) and a convolutional neural network (CNN) are coupled for bearing failure diagnostics to simplify the network design. However, the aforementioned method requires a huge number of samples for training, which is tough to come by in real-world engineering. Furthermore, the deep network takes a long time to develop. Data augmentation is the method of choice for training on comparable but distinct examples to the training data. In data analysis, data augmentation refers to ways for increasing the amount of data by adding slightly modified copies of current data, such as horizontal reflections, modest rotations, and moderate scalings, or by creating new synthetic data from existing data. While data augmentation regularly improves generalisation [23], it is dataset-dependent, necessitating the application of expert knowledge. Furthermore, data augmentation implies that all of the instances in the area belong to the same class and does not model the proximity relationship between examples of various classes [24]. Machine problems are frequently caused by bearing failures. These faults are difficult to detect and anticipate using traditional, non-machine learning based methods due to their complexity. Machine learning promises to provide better answers for problems that typically need a lot of manual fine-tuning or cannot be solved at all using traditional methods. In this research, a CNN with Mix-up Data Augmentation was used to diagnose bearing defects, taking into account the benefits and drawbacks of all of the previous work. First and foremost, CWT is utilised to extract preliminary features. To reduce the classification loss with validation data, the validation dataset, other than the training and test datasets, is prepared using Mix-up data augmentation. The CNN with Mix-up Data Augmentation model of a network is built using transfer learning from ResNet-18. ResNet-18 is an image classification model pre-trained on the ImageNet dataset [25]. More than a million photos from the ImageNet database were used to train ResNet-18. The network can classify photos into 1000 different object categories, including keyboards, mice, pencils, and a variety of animals. As a result, the network has learned a variety of rich feature representations for a variety of images. Finally, the network model is trained using the previously retrieved characteristics to classify ball bearing faults. This approach fully utilises CNN's excellent feature extraction capabilities, addressing the difficulty of extracting in-depth signal features as well as the challenge of acquiring large-scale data in real-world situations. The following is a breakdown of the paper's structure: The theory of CWT, Mix-up Data Augmentation and CNN is covered in Section 2. Section 3 presents the suggested CNN with the Mix-up Data Augmentation model. The full experimental process, as well as the data analysis, is introduced in Section 4. Finally, Section 7 has the conclusion.

2. Background

An intelligent diagnosis approach for ball bearing defects is proposed in this research. To begin, the original vibration signals are converted into scalogram pictures using CWT. Then use a data augmentation technique known as a Mix-up. CNN gets a training image as a result of the Mix-up. Finally, the in-depth features of the scalogram photos are extracted using CNN. The following are the essential theories of CWT, CNN, and Mix-up.

2.1. Time-frequency analysis with CWT

The continuous wavelet transform (CWT) is a method that allows the translation and scale parameters of wavelets to vary continuously, resulting in an overcomplete representation of a signal [26]. Preliminary feature extraction is performed using CWT, which converts the original one-dimensional time-domain signals into two-dimensional scalogram pictures.

2.1.1. Scalograms

The CWT is shown graphically by scalograms. Instead of sinusoidal functions, CWTs are a wavelet-based linear time-frequency representation. Because it incorporates a scale variable in addition to the time variable, CWT is effective for nonstationary and transient signals. The following integral represents the continuous wavelet transform of a function at a scale ($a > 0$) and translational value $b \in L^2(r)$.

$$WT_f(b,a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t)\psi^* \left(\frac{t-b}{a}\right) dt \quad (1)$$

Where the signal dilates or compresses based on the scale factor a. b Is the time factor, $f(t)$ is the original signal, ψ Denotes the analyzing mother wavelet function, and the complex conjugate is ψ^*.

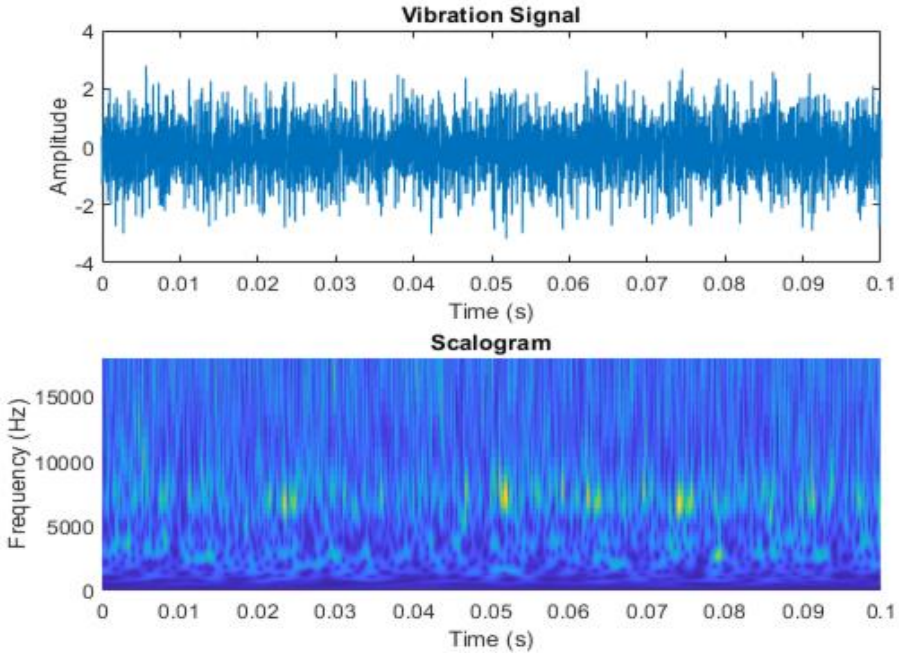


Fig. 1: Displays A Scalogram for the Normal State, When the Bearing Is in Good Working Order.

Wavelets have been widely used in the identification of machinery faults. Figure 1 shows a scalogram for the normal state of a vibrational response of a rolling element with a Morlet wavelet (Gabor wavelet) basis. Much research has been conducted to determine the efficacy of specific wavelets and their capacity to match a signal. In both simple and complicated functions, the Haar, Daubechies, Biorthogonal, Coiflets, Symlets, Morlet, Mexican Hat, and Meyer wavelets could be used. There is currently no clear approach for determining the appropriate wavelet to utilise, and this remains an open subject among researchers [27]. The Morlet wavelet (Gabor wavelet) is chosen for this paper because it is similar to the impulse component of symptomatic failures in many mechanical systems [28], and is defined as

$$\psi_\sigma(t) = C\pi^{-0.25}\exp(-0.5t^2)\left(e^{i\sigma t} - K\right) \tag{2}$$

The admissibility criteria define $K = e^{-0.5\sigma^2}$, and the normalisation constant C Is:

$$C = \left(1 + e^{-\sigma^2} - 2e^{-0.75\sigma^2}\right)^{-1/2}$$

Morlet wavelets have a number of advantages when it comes to time-frequency analysis. First, the Morlet wavelet has a Gaussian structure in the frequency response. The lack of ridges reduces ripple effects that could be mistaken for vibrations. This is a risk that plateau-shaped filters may pose. Second, the Morlet wavelet convolution results preserve the original signal's temporal resolution. Third, compared to other approaches, wavelet convolution is more computationally efficient and requires less code because it involves the fewest number of operations. The FFT (fast Fourier transform) is used in the majority of cases [29].

2.2. Mix-up data augmentation

Many machine learning algorithms, such as image classification, use data augmentation to virtually increase the training sample size and avoid overfitting [30]. Mix-up is a data augmentation routine that is both easy and data agnostic. Mix-up, in a nutshell, creates virtual training instances.

$$\begin{aligned} \tilde{x} &= ax_i + bx_j \\ \tilde{y} &= ay_i + by_j \end{aligned} \tag{3}$$

Where, $a + b = 1$

(x_i, y_i) and (x_j, y_j) are two randomly selected samples from our training data belongs to the same classes. As a result, Mix-up broadens the training distribution by factoring in the assumption that linear interpolations of feature vectors should result in linear interpolations of the related targets. while other advanced data augmentation techniques exist (e.g., those based on Generative Adversarial Networks), Mix-up was chosen for its simplicity, low computational overhead, and proven effectiveness in broadening training distribution. GANs are known for their training instability, difficulty in convergence, and issues like mode collapse (where the generator produces only a limited variety of samples). Mix-up bypasses these challenges entirely as it does not involve a generative adversarial process. Mix-up takes only a few lines of code to implement and has a low computational overhead. Mix-up allows a new state-of-the-art performance in the CWRU bearing classification datasets (section 5), despite its simplicity. Furthermore, when learning from corrupt labels, Mix-up improves the durability of neural networks.

2.3. Convolutional neural network

Convolutional neural networks were first introduced by Yann LeCun and Yoshua Bengio in 1995 [31]. A neural network is a collection of neurons that are connected; at the same time, each connection in a neural network has a weight associated with it, which, when multiplied

by the input value, indicates the relevance of this relationship in the neuron. The activation function of each neuron determines the cell's output. The activation function is utilised to add non-linearity to the network's modelling capabilities. CNN has made significant progress in the study of image recognition as a standard means for obtaining data characteristics using deep learning models. Figure 2 illustrates the overall structure of the CNN. The convolution layer, pooling layer, and fully-connected layer make up the majority of its internal hidden layer structure.

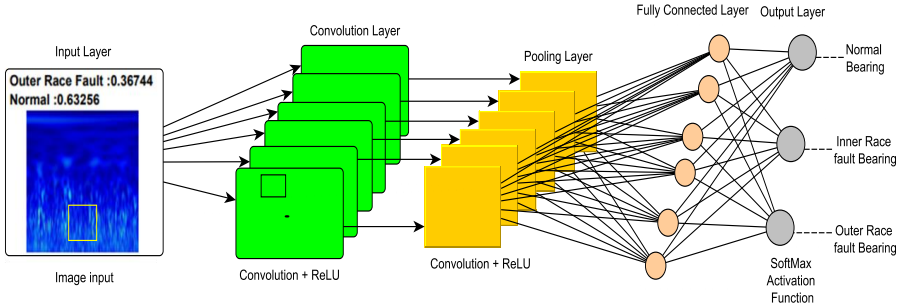


Fig. 2: Illustrates CNN's Overall Structure.

2.3.1. Convolution layer

A CNN's main building block is the Conv layer. A set of learnable filters makes up the parameters. Every filter is modest in terms of size (width and height), but it covers the entire depth of the input volume. During the forward pass, each filter is slid (convolved) over the width and height of the input volume, and dot products between the filter's entries and the input are computed at any place. It creates a 2-dimensional activation map that shows the filter's reactions at all spatial positions. The network will, intuitively, learn filters that activate when they see a specific visual feature. The following is a definition of the 2D convolution formula:

$$f * g(x,y) = \int_{\theta=-\infty}^{\infty} \int_{\phi=-\infty}^{\infty} f(\theta,\phi) g(x-\theta, y-\phi) d\theta d\phi \quad (4)$$

Where $f(\theta,\phi)$ Is the activation function. Each CONV layer has a set of filters that form a separate 2-dimensional activation matrix. It creates the output volume by stacking these activation maps along the depth dimension.

2.3.2. Pooling layer

The addition of a pooling layer minimises the representation's spatial dimension, reduces the number of parameters and computations in the network, and so controls overfitting. Using the MAX operation, the Pooling Layer operates independently on each depth slice of the input and resizes it spatially. The following is the mathematical formula for pooling [32]:

$$x_k^{(m)} = f\left(\alpha_k^m \text{down}\left(x_k^{m-1}\right) + b_k^m\right) \quad (5)$$

Where the down sampling function is down(·), α_k^m Is the network's weight and b_k^m Stands for additive bias.
The outputs can be passed into the CNN's last phase, which is a fully connected layer (FC), when numerous stacks of previous layers have been finished. Simply said, the Fully Connected Layer is made up of feed-forward neural networks. The max-pooling layer's findings are used as input for mapping to the FC layer's data labels. As a result, vibration image analysis and prediction are a sequence of representations of the signal in its unprocessed state. The vibration signals are converted to scalogram images using CWT. Then, CNN is used to learn and classify the scalogram image features and predict a label-based classification.
The Proposed CNN With Mix-Up Data Augmentation
For rolling bearing defect diagnosis, a network architecture based on CNN with Mix-up Data Augmentation is built in this research. The design of a CNN with a Mix-up Data Augmentation model and bearing fault diagnosis using the suggested model are the primary topics of this chapter.

2.4. Design of CNN with mix-up data augmentation for bearing fault diagnosis

The CNN with Mix-up Data Augmentation model is built in this section, and its structure is shown in Figure 3. To achieve the goal of picture classification, the model employs a transfer learning ResNet-18 model to obtain 2-D image features, and then the classifier is trained using the retrieved characteristics. The method of transfer learning is applied in this system, and the ResNet-18 is used to extract features. Without raising the training error %, networks with a large number of layers (even thousands) can be easily trained. Using identity mapping, ResNets can aid with the vanishing gradient problem. Resnet-18 with Mix-up Validation further improves classification performance. Table 1 shows its shape and properties [25]. The issue of manual CNN building, as well as Inadequate data samples produced in actual engineering, can be efficiently solved by applying relevant knowledge gathered in the pre-trained network. To some extent, the ResNet-18 network uses residual blocks to avoid the problem of vanishing gradients.

Table 1: Resnet-18's Structure and Properties

Layer Name	Activations	Parameters
Image Input	$224 \times 224 \times 3$	$--$
Convolution1	$112 \times 112 \times 64$	$\begin{pmatrix}\text{Weights}\\ \text{Stride}\\ \text{Padding}\end{pmatrix} = \begin{pmatrix}7 \times 7 \times 3 \times 64\\ [22]\\ [3333]\end{pmatrix}$
Convolution2_x	$56 \times 56 \times 64$	$\begin{pmatrix}\text{Weights}\\ \text{Stride}\\ \text{Padding}\end{pmatrix} = \begin{pmatrix}3 \times 3 \times 64 \times 64\\ [11]\\ [1111]\end{pmatrix} \times 4$

Convolution3_x	$28 \times 28 \times 128$	$\begin{pmatrix}\text{Weights}\\ \text{Stride}\\ \text{Padding}\end{pmatrix} =$	$\begin{pmatrix}3 \times 3 \times 128 \times 128\\ [11]\\ [1111]\end{pmatrix} \times 4$
Convolution4_x	$14 \times 14 \times 256$	$\begin{pmatrix}\text{Weights}\\ \text{Stride}\\ \text{Padding}\end{pmatrix} =$	$\begin{pmatrix}3 \times 3 \times 256 \times 256\\ [11]\\ [1111]\end{pmatrix} \times 4$
Convolution5_x	$7 \times 7 \times 512$	$\begin{pmatrix}\text{Weights}\\ \text{Stride}\\ \text{Padding}\end{pmatrix} =$	$\begin{pmatrix}3 \times 3 \times 512 \times 512\\ [11]\\ [1111]\end{pmatrix} \times 4$
2-D Global Average Pooling	$1 \times 1 \times 512$	$- -$	
Fully Connected	$1 \times 1 \times 1000$	Weights =	1000×512
SoftMax	$1 \times 1 \times 1000$	$- -$	

In practical engineering, there is a lack of data on bearing defects, and it's challenging to extract features in traditional intelligent diagnosis approaches. To improve estimation performance, this research provides a Mix-up Data Augmentation approach with CNN for detecting ball bearing faults. To monitor the classification loss with validation data, the validation dataset, other than the training and test datasets, are prepared using Mix-up data augmentation. This strategy can be beneficial to CNN's superiority in acquiring data features and the generalisation capabilities of Mix-up Data Augmentation. Mix-up training is simple to implement and has a low computational burden. Figure 3 explains how to implement Mix-up training in MATLAB. The actions taken to run the suggested model are as follows:

- On two independent experimental platforms, vibration signals from defective bearings are gathered.
- The vibration signals are first divided into segments. The gathered vibration signals are then converted into scalogram pictures of $224 \times 224 \times 3$ Size using CWT. Finally, the scalogram images are divided into training and test samples, with each sample receiving its label.
- To begin, use the Mix-up data augmentation from the previous phase to generate a new generic validation dataset.
- Load the ResNet-18 network that has been pre-trained. Then, using the fully connected layer, feed the samples generated in step 3 into the model and receive training and test images' high-level feature representations. The data in the output layer is classified using the SoftMax activation function.

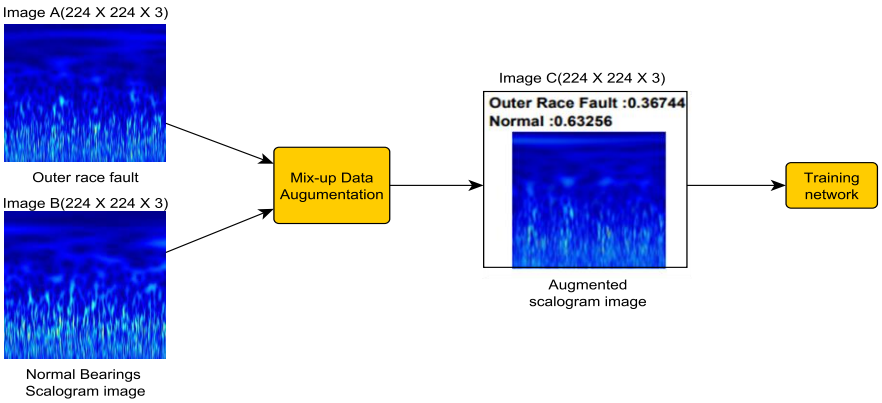


Fig. 3: Proposed Architecture of CNN With Mix-Up Data Augmentation Model for Bearing Fault Diagnosis, Utilizing Transfer Learning from Resnet-18.

The hyperparameters dropout and learning rate must be specified and optimised for the CNN. A Dropout layer is another common feature of CNNs. Dropout helps thin the network by preventing overfitting and reducing training errors. All units that make it through the dropout make up the remaining connections. Dropout is set to 0.01 for this architecture. The stochastic gradient descent with momentum (SGDM) algorithm was used to optimise the other hyperparameter, learning rate. For this work, GPU computing was employed to speed up the training process and reduce the amount of time it took. The analysis was carried out using the following processing system: RYZEN 5 3600 4.1GHz CPU with 16 GB RAM and NVIDIA GTX 1660 graphics card.

3. Experimental Setup

We created an experimental setup and gathered three classes of bearing data, as shown in Figure 4 to evaluate the flexibility and applicability of the created model in ball bearing defect diagnostics. Validate the suggested model once more using the Bearing Data Centre dataset from Case Western Reserve University (CWRU) [33].

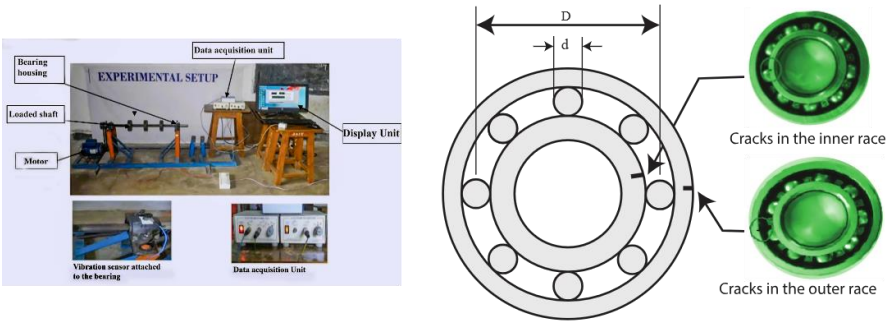


Fig. 4: A Test Rig with A Variety of Bearing Failures.

Our experimental design yielded three datasets: a normal set, an inner race fault set, and an outer race fault set. All of the data was collected at a sample rate of 48828 Hz for 12 seconds. The data was collected using a single-channel radial accelerometer and a USB-6002 NI DAQ Device. HRB-6202 is a deep groove ball bearing with an element number of 8 and dimensions of 15mm inner diameter, 35mm outer diameter, and 11mm breadth is employed in the test setup.

3.1. Dataset from the experiments

Table 2 shows the total scalogram images generated from the data set: Inner race fault 406, Normal Bearing 351, and Outer race fault 757. 70 per cent of the images from each label are picked at random as training samples, 20 per cent as test samples, and the remaining 10 per cent as validation samples. Figure 5 shows the scalogram images of the three health conditions described in Table 2.

Table 2: Description of the Experimental Dataset

Types of faults in bearings	Total number of samples	No of samples in the validation dataset	Class Type
Normally or a healthy status	351	42	Normal
Inner race fault	406	42	Inner Race Fault
Outer race fault	757	80	Outer Race Fault

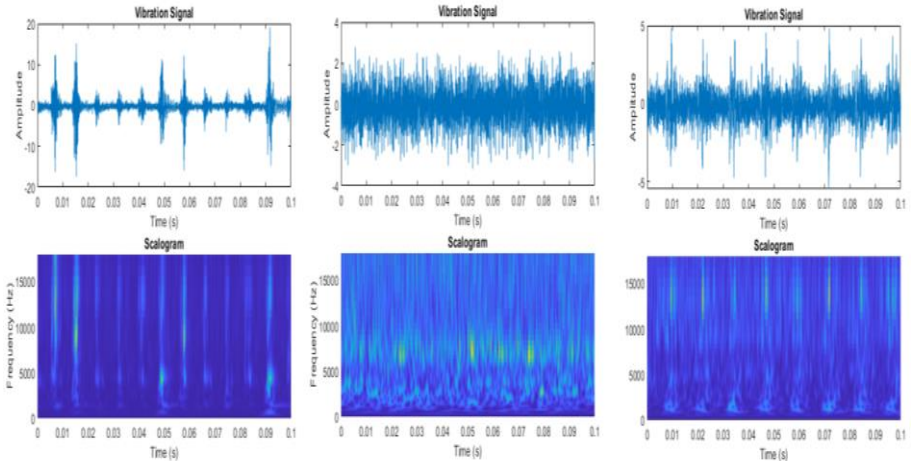


Fig. 5: Scalogram Images of Three Health Conditions on the Experiment Dataset: (A) Inner Race Fault; (B) Normal; (C) Outer Race Fault.

3.2. The experiment's results

To begin, use the proposed Mix-up data augmentation to create all synthetic images. Figure 6: Display the synthesised images to show how much of each image class is used. The three bearing failure classes' constructed training samples are used to train the CNN model with Mix-up data augmentation. When the training is completed, the test and validation samples corroborate the trained model. The experiment is performed ten times to demonstrate the stability of the suggested model. Figure 7 depicts the model's Mix-up training performance and ROC. Figure 7(a) shows that Mix-up training is successfully learnt because the validation losses, like the training loss, decrease with each epoch. The Receiver Operator Characteristic (ROC) curve is a binary classification issue evaluation metric. It's a probability curve that displays the TPR (True positive rate) against the FPR (False positive rate) at different threshold levels, thereby separating the ‘signal' from the 'noise.' The Area Under the Curve (AUC) is a summary of the ROC curve that measures a classifier's ability to distinguish between classes. The average diagnosis performance of all trials is shown in Table 3. On each trial, achieve 100% accuracy in bearing dataset classification using new validation data collected through Mix-up data augmentation. Figure 8 shows the Confusion matrix of a fault diagnosis.

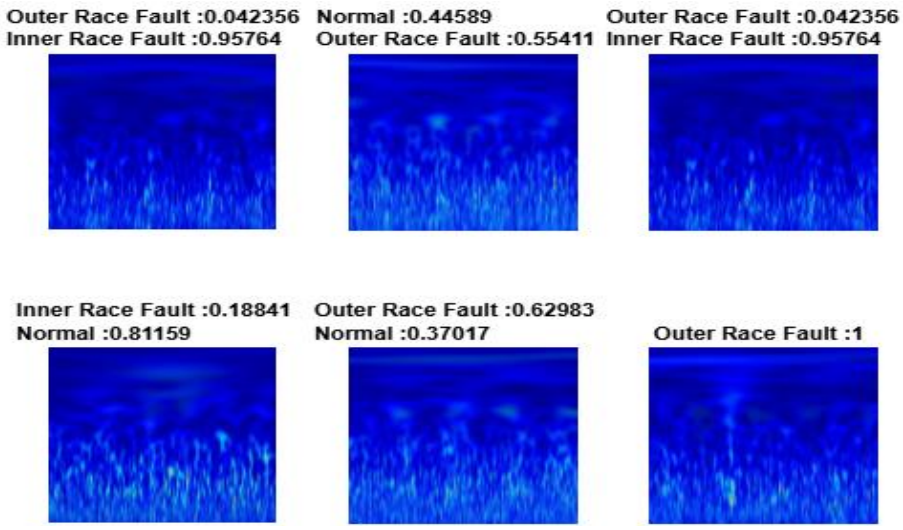
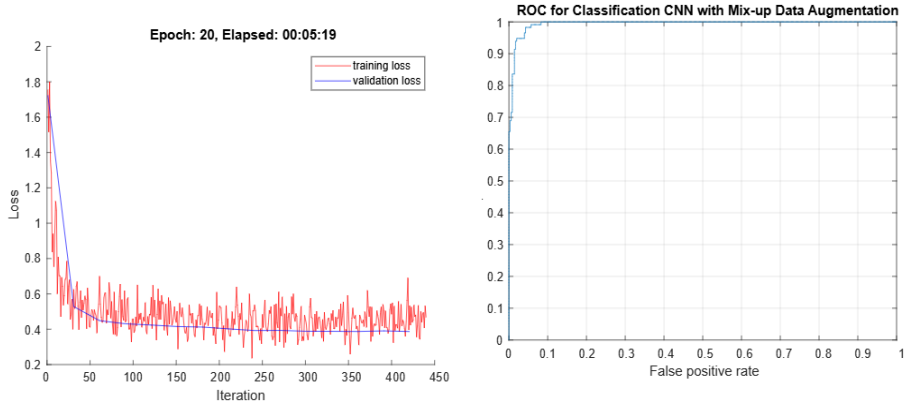


Fig. 6: The Synthesised Images Show How Much of Each Image Class Is Used.



(a) Training progress (b) ROC curve

Fig. 7: Shows the Model's Training Performance as Well as the ROC Curve.

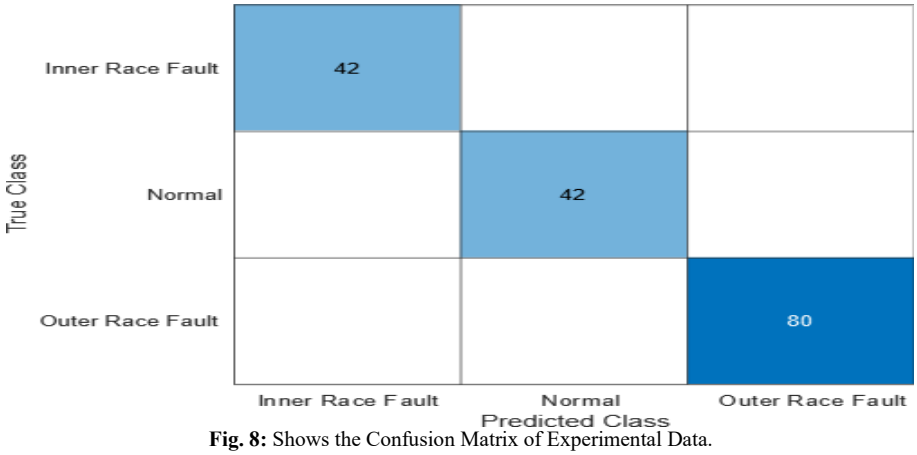


Fig. 8: Shows the Confusion Matrix of Experimental Data.

Table 3: Average Diagnosis Performance

Accuracy	1
Overall-precision	1
Overall-recall	1
fl_score	1
AUC	0.9944

4. CWRU Dataset

The second set of data from tests used in this study came from the Bearing Data Centre at CWRU. The tests used a 2 hp electric motor to obtain accelerometer data on the bearings, as illustrated in Figure 9(a). The motor shaft is supported by the bearings. Inner raceway artificial faults (IR), rolling ball faults (BF), and outer raceway artificial faults (OR) were seeded using electro-discharge machining (EDM). The outer raceway faults varied in diameter and location. A motor with a horsepower range of 0 to 3 is included in the data. At noon, the accelerometers were magnetically linked to the casing. The damage diameters of the inner race and ball are set to 0.1778, 0.3556, 0.5334, and 0.7112 mm, respectively, using the EDM's single-point harm approach. At 6 o'clock, however, the outer ring's damage diameters are the same as before. Bearing vibration data in normal conditions and 12 fault states is gathered.

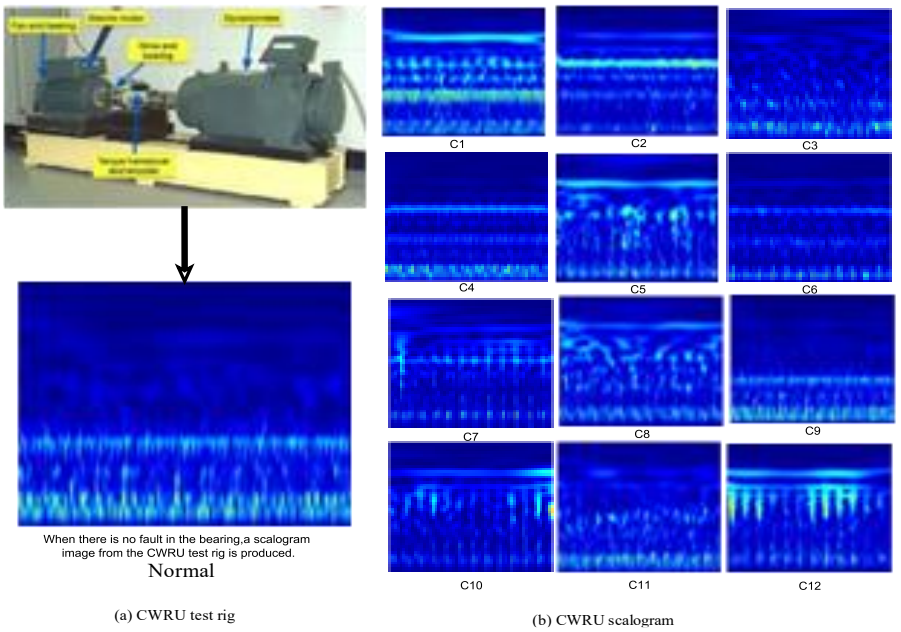
Table 4: CWRU Dataset's Description

Bearing fault types	Diameters of faults(mm)	Total number of samples	Number of samples used in the model	Class Type
Normal status	0	289	48	Normal
Fault in the inner race	0.1778	264	66	C1
	0.3556	264	66	C2
	0.5334	264	66	C3
	0.7112	264	66	C4
Fault in the ball	0.1778	264	66	C5
	0.3556	264	66	C6
	0.5334	264	66	C7
	0.7112	264	66	C8
Fault in the outer race	0.1778	504	66	C9
	0.3556	480	66	C10
	0.5334	480	66	C11
	0.7112	432	66	C12

Signals obtained at a frequency of 12kHz were utilised in this study to look into 13 potential health concerns that can damage the fan end bearing. This method can better approximate different fault sites and bearing damage severity. As demonstrated in Figure 9(b), CWT was used to construct scalogram images from this data on different classes. Table 4 shows the scalogram images of the 13 health issues of bearings in sequence. The Morlet wavelet is a single-frequency sine function with a Gaussian envelope that is used to generate scalogram images using CWT.

4.1. The experiment's results

To begin, produce all synthetic images using the proposed Mix-up data augmentation. Figure 10: Shows how much of each image class is used by displaying the synthesised images. The training samples produced from the thirteen bearing failure classes are used to train the CNN model with Mix-up data augmentation. When the training is finished, the test sample confirms the trained model. The experiment is carried out a total of ten times to show that the suggested model is stable. The model's ROC curve is shown in Figure 11(a). Figure 11(b) shows the images in the test dataset that were incorrectly categorised. Table 5 shows the average diagnosis performance of all trials. On the CWRU bearing dataset, the average accuracy utilising Mix-up data augmentation is 92.17 per cent. SqueezeNet transfer learning without data augmentation was assessed in the CWRU dataset to determine the model's efficiency. Using the pre-trained SqueezeNet network, an average accuracy of 84.57 percent was attained.



(a) CWRU test rig (b) CWRU scalogram

Fig. 9: CWRU Dataset Test Rig and Scalogram Images of Thirteen Health Conditions.

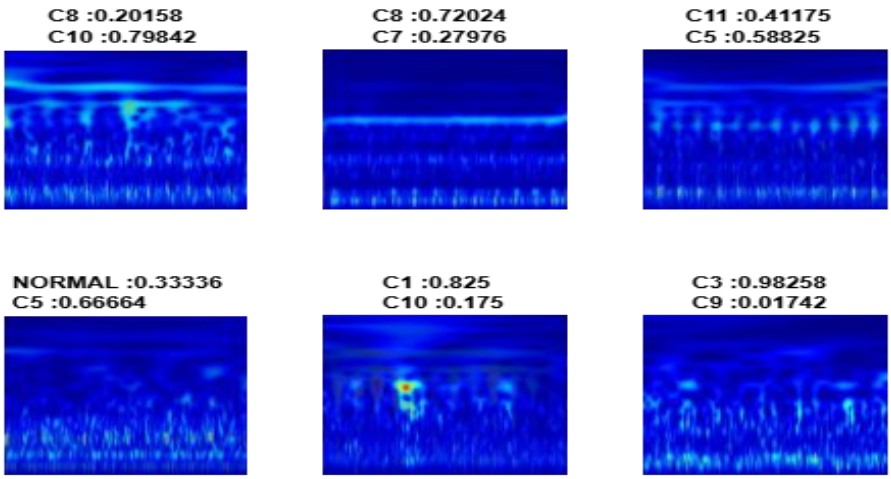
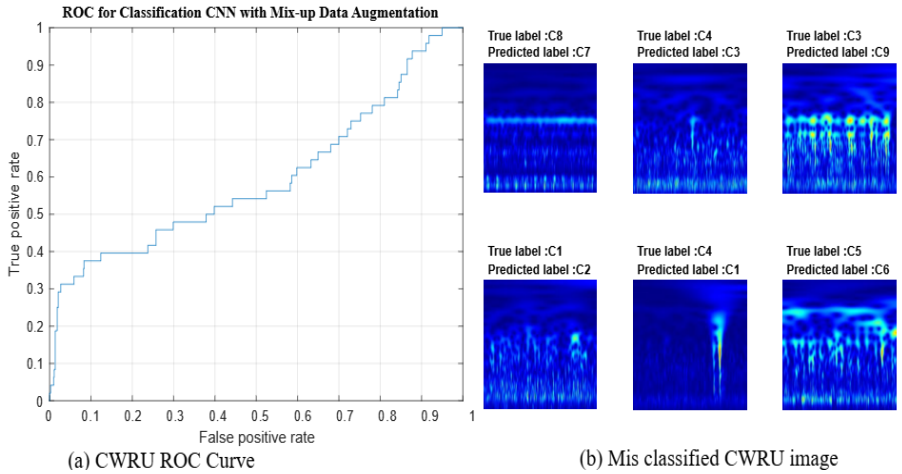


Fig. 10: Display The Synthesised Images to Show How Much of Each Image Class Is Used from the CWRU Dataset

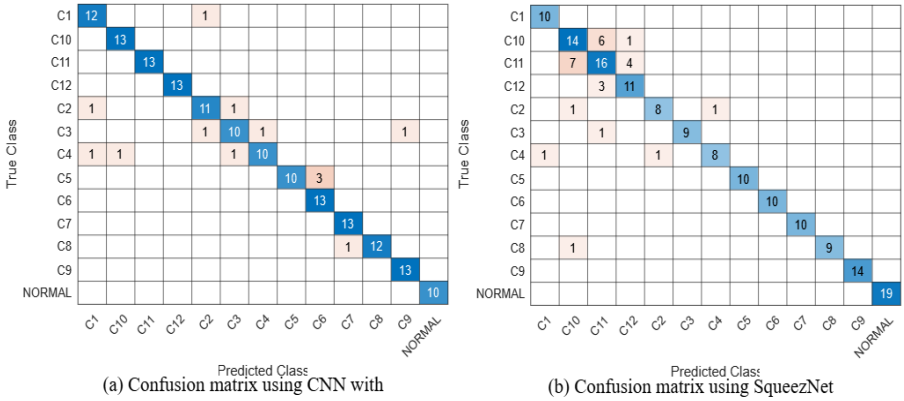
Table 5: Average Diagnosis Performance Using the CWRU Dataset

Accuracy	0.9217
Overall-precision	0.9265
Overall-recall	0.9231
f1_score	0.9248
AUC	0.5937



(a) CWRU ROC Curve (b) Mis classified CWRU image

Fig. 11: The Model's ROC Curve for CWRU Datasets and Misclassified Images in the Test Dataset.



(a) Confusion matrix using CNN with Mix-up data augmentation

(b) Confusion matrix using SqueezNet without Mix-up data augmentation

Fig. 12: CNN with Mix-Up Data Augmentation and CNN Without Data Augmentation Were Used to Create A Confusion Matrix from the CWRU Datasets.

5. Analysis of Other Methods

We compare the suggested model to standard CNN models like ResNet-18 and SqueezeNet to demonstrate its advantages. Scalogram images are the usual CNN's inputs, while Softmax is the standard CNN's classifier. The test is repeated ten times using the experimental test dataset on these two models to guarantee that the experiment is fair and the findings are trustworthy. Figure 13 depicts the diagnostic accuracy of each approach. Table 6 summarises the accuracy of diagnosis and the overall training time of the three approaches. As can be shown, in five trials, the recommended method's diagnosis accuracy outperformed the other two models, and the training period is significantly less than the standard CNN model. The suggested model outperforms existing pre-trained networks such as ResNet-18 and SqueezeNet, according to the findings. It has been demonstrated that the problem of obtaining in-depth characteristics from a range of data can be efficiently solved with this technique, as well as the difficulty of meeting the demands of huge samples for CNN training and that it has outstanding stability. Many papers on deep learning-based fault diagnosis algorithms have been published in recent years. Advanced approaches PNN-SFAM [34], CNN-HMM [35], BPNN [35], DGNN [36], and DAFD [37] are compared to the suggested method to illustrate its uniqueness in the subject of diagnosing bearing defects. Table 7 illustrates the average diagnostic accuracy of each approach. The following is a list of additional hyperparameters to consider. The mini-batch size is 25, and the number of epochs is 20. The learning rate begins at 0.001, decays at 0.01, and the models are trained using the stochastic gradient descent with momentum (SGDM) optimiser with a momentum of 0.9 [25]. When the findings of the suggested approach are compared to those of other methods, it is clear that the proposed strategy improves the accuracy of diagnosis, demonstrating its efficacy.

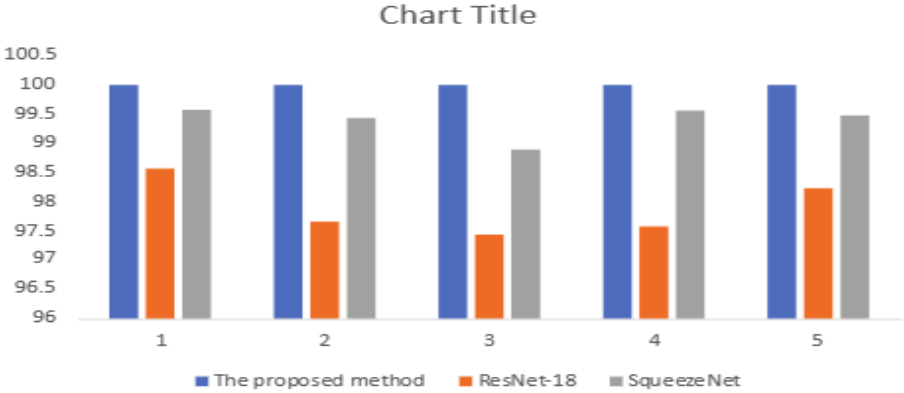


Fig. 13: Displays the Diagnostic Accuracy of Each Method.

Table 6: Comparing the Three Techniques' Diagnostic Accuracy and Overall Training Time

Models	Accuracy (%)	Training Time in sec
CNN with Mix-up	100	319
Data Augmentation	98.56	423
ResNet-18	99.44	383

Table 7: The Proposed Methods and Numerous Other Approaches' Average Accuracy

Category	Methods	Accuracy (%)
Machine learning	PNN-SFAM [34]	97.82
	BPNN [35]	77.86
	KNN, Decision tree, Random forest [10]	97
Deep learning	CNN-HMM [35]	98.125
	DGNN [36]	97.81
	DAFD [37]	94.73
	CNN with MMD [38]	74 – 81
	enhanced deep transfer auto-encoder [39]	90.42
CNN with Data	Proposed Model	100

6. Conclusion

For fault diagnosis of rolling bearings, a new deep neural network model, CNN with Mix-up Data Augmentation, is developed in this study. To begin, we create scalogram images of vibration signals using CWT. The acquired photos are then used to train the suggested model. Finally, the position and harshness of the ball bearing fault are determined. The results show that this method's diagnostic accuracy can reach 92.17 per cent in the case of the CWRU dataset and 100 per cent for the experimental dataset, demonstrating the model's flexibility and applicability. By comparing the proposed model to transfer learning models such as ResNet-18 and SqueezeNet, it has been demonstrated that it can resolve the prior method's issues with deep feature extraction as well as the limited sample issue in actual engineering and that it is extremely stable. The effectiveness of the suggested technique is further proven by comparing it to advanced approaches, PNN-SFAM [34], CNN-HMM [35], BPNN [35], DGNN [36], DAFD [37], CNN with MMD [38] and enhanced deep transfer auto-encoder [39]. However, the proposed method's accuracy needs to be improved for relatively noisy data sources. The Wavelet Transform (WT) is an effective method for removing noise from a wide range of signals. Combining WT with other noise-reduction strategies could result in even more noise reduction. Singular Vector Decomposition (SVD), like WT, is an effective noise reduction technique [40]. Future work will explore hybrid noise-reduction techniques, such as combining Wavelet Transform (WT) with Singular Vector Decomposition (SVD), to further enhance the model's performance on noisy datasets. Another critical area for future investigation involves adapting Mix-up for imbalanced datasets. Mix-up can be adapted for imbalanced datasets through a novel mechanism called Balanced-MixUp. This technique involves simultaneously performing both regular (instance-based) and balanced (class-based) sampling of training data. The two resulting sets of samples are then "mixed-up" to create a more balanced training distribution. This approach allows a neural network to learn effectively without heavily under-fitting the minority classes. Experimental results have shown that Balanced-MixUp outperforms other conventional sampling schemes and loss functions designed for imbalanced data. The exploration of such advanced Mix-up strategies will further enhance the model's robustness and applicability in real-world scenarios with varying data distributions [41]. We can further improve the CWRU performance estimate using K-fold cross-validation. As a result, the proposed model's structure will need to be improved in the future. Furthermore, we used single-fault vibration signals for model training in this study, and no samples of compound faults were provided to match the actual condition, posing a barrier to the suggested model's application in real-world engineering. We've decided to pursue it as a possible future research topic.

Conflicts of Interest

There is no Conflict of interest.

Ethical

This study does not involve human participants or animals.

References

[1] D.-T. Hoang and H.-J. Kang (2019) A survey on deep learning based bearing fault diagnosis. Neurocomputing335, pp. 327-335. https://doi.org/10.1016/j.neucom.2018.06.078.

[2] Faults in Electrical Machines and their Diagnosis, pp. 1-22.

[3] M. Kang, M. Islam, J. Kim, J. Kim, and M. Pecht (2016) A hybrid feature selection scheme for reducing diagnostic performance deterioration caused by outliers in data-driven diagnostics. IEEE Transactions on Industrial Electronics63 (5), pp. 3299-3310. https://doi.org/10.1109/TIE.2016.2527623.

[4] Q. Hu, X.-S. Si, A.-S. Qin, Y. Lv, and Q. Zhang (2020) Machinery fault diagnosis scheme using redefined dimensionless indicators and mrmr feature selection. IEEE Access 8, pp. 40313-40326. https://doi.org/10.1109/ACCESS.2020.2976832.

[5] Randall, R., & Antoni, J. Why EMD and similar decompositions are of little benefit for bearing diagnostics. *Mechanical Systems and Signal Processing.* 2023 https://doi.org/10.1016/j.ymssp.2023.110207. https://doi.org/10.1016/j.ymssp.2023.110207.

[6] Yan, R., Shang, Z., Xu, H., Wen, J., Zhao, Z., Chen, X., & Gao, R. Wavelet transform for rotary machine fault diagnosis:10 years revisited. *Mechanical Systems and Signal Processing.* 2023 https://doi.org/10.1016/j.ymssp.2023.110545.

[7] Alexakos, C., Karnavas, Y., Drakaki, M., & Tziafettas, I. A Combined Short Time Fourier Transform and Image Classification Transformer Model for Rolling Element Bearings Fault Diagnosis in Electric Motors. *Mach. Learn. Knowl. Extr.*. 2021; 3. https://doi.org/10.3390/make3010011.

[8] Yuan, L., Lian, D., Kang, X., Chen, Y., Zhai, K.: Rolling bearing fault diagnosis based on convolutional neural network and support vector machine. IEEE Access 8, 137395-137406 (2020). https://doi.org/10.1109/ACCESS.2020.3012053.

[9] Yan, R., Gao, R.X., Chen, X.: Wavelets for fault diagnosis of rotary machines: A review with applications. Signal processing 96, 1-15 (2014). https://doi.org/10.1016/j.sigpro.2013.04.015.

[10] Khraisat, A., Uddin, M., Halder, R., Aryal, S., & Uddin, A. Enhancing K-nearest neighbor algorithm: a comprehensive review and performance analysis of modifications. *J. Big Data.* 2024; 11. https://doi.org/10.1186/s40537-024-00973-y.
[11] Karnavas, Y.L., Chasiotis, I.D., Vrangas, A.: Fault diagnosis of squirrel-cage induction motor broken bars based on a model identification method with subtractive clustering. In: 2017 IEEE 11th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED), pp. 304-310 (2017). IEEE. https://doi.org/10.1109/DEMPED.2017.8062372.
[12] Wu, W. Automotive Motor Fault Diagnosis Model Integrating Machine Learning Algorithm and Fuzzy Control Theory. International Journal of Fuzzy Systems. 2025 https://doi.org/10.1007/s40815-024-01926-6.
[13] Drakaki, M., Karnavas, Y.L., Karlis, A.D., Chasiotis, I.D., Tzionas, P.: Study on fault diagnosis of broken rotor bars in squirrel cage induction motors: A multi-agent system approach using intelligent classifiers. IET Electric Power Applications 14(2), 245-255 (2019). https://doi.org/10.1049/iet-epa.2019.0619.
[14] Hosseini, N., Toshani, H., & Abdi, S. A Projection-Based Support Vector Machine Algorithm for Induction Motors' Bearing Fault Detection. *2023 IEEE 14th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED).* 2023 https://doi.org/10.1109/SDEMPED54949.2023.10271436.
[15] Schmidhuber, J.: Deep learning in neural networks: An overview. Neural networks 61, 85-117 (2015). https://doi.org/10.1016/j.neunet.2014.09.003.
[16] Qiao, M., Yan, S., Tang, X., Xu, C.: Deep convolutional and lstm recurrent neural networks for rolling bearing fault diagnosis under strong noises and variable loads. IEEE Access 8, 66257-66269 (2020). https://doi.org/10.1109/ACCESS.2020.2985617.
[17] Li, X., Jiang, H., Xiong, X., Shao, H.: Rolling bearing health prognosis using a modified health index based hierarchical gated recurrent unit network. Mechanism and Machine Theory 133, 229-249 (2019). https://doi.org/10.1016/j.mechmachtheory.2018.11.005.
[18] Zhao, H., Chen, H., Chen, J., Li, Y., Xu, J., & Deng, W. Intelligent Diagnosis Using Continuous Wavelet Transform and Gauss Convolutional Deep Belief Network. IEEE Transactions on Reliability. 2023; 72. https://doi.org/10.1109/TR.2022.3180273.
[19] Haidong, S., Hongkai, J., Xingqiu, L., Shuaipeng, W.: Intelligent fault diagnosis of rolling bearing using deep wavelet auto-encoder with extreme learning machine. Knowledge-Based Systems 140, 1-14 (2018). https://doi.org/10.1016/j.knosys.2017.10.024.
[20] Hoang, D.-T., Kang, H.-J.: Rolling element bearing fault diagnosis using convolutional neural network and vibration image. Cognitive Systems Research 53, 42-50 (2019). https://doi.org/10.1016/j.cogsys.2018.03.002.
[21] Chen, Z., Mauricio, A., Li, W., Gryllias, K.: A deep learning method for bearing fault diagnosis based on cyclic spectral coherence and convolutional neural networks. Mechanical Systems and Signal Processing 140, 106683 (2020). https://doi.org/10.1016/j.ymssp.2020.106683.
[22] Yang, Y., Zheng, H., Li, Y., Xu, M., Chen, Y.: A fault diagnosis scheme for rotating machinery using hierarchical symbolic analysis and convolutional neural network. ISA transactions 91, 235-252 (2019). https://doi.org/10.1016/j.isatra.2019.01.018.
[23] Simard, P., Lecun, Y., Denker, J.: Transformation invariance in pattern recognition -tangent distance and tangent propagation (2000). https://doi.org/10.1002/1098-1098(2000)11:3<181::AID-IMA1003>3.0.CO;2-E.
[24] Zhang, H., Cisse, M., Dauphin, Y.N., Lopez-Paz, D.: mixup: Beyond empirical risk minimization. arXiv preprint arXiv:1710.09412 (2017).
[25] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770-778 (2016). https://doi.org/10.1109/CVPR.2016.90.
[26] Yan, R., Gao, R.X., Chen, X.: Wavelets for fault diagnosis of rotary machines: A review with applications. Signal processing 96, 1-15 (2014). https://doi.org/10.1016/j.sigpro.2013.04.015.
[27] Verstraete, D., Ferrada, A., Droguett, E.L., Meruane, V., Modarres, M.: Deep learning enabled fault diagnosis using time-frequency image analysis of rolling element bearings. Shock and Vibration 2017 (2017). https://doi.org/10.1155/2017/5067651.
[28] Lin, J., Qu, L.: Feature extraction based on morlet wavelet and its application for mechanical fault diagnosis. Journal of sound and vibration 234(1), 135-148 (2000). https://doi.org/10.1006/jsvi.2000.2864.
[29] Cohen, M.X.: A better way to define and describe morlet wavelets for time-frequency analysis. bioRxiv (2018). https://doi.org/10.1101/397182.
[30] Inoue, H.: Data augmentation by pairing samples for images classification (2018).
[31] LeCun, Y., Bengio, Y., *et al.*: Convolutional networks for images, speech, and time series. The handbook of brain theory and neural networks 3361(10), 1995 (1995).
[32] Bouvrie, J.: Notes on convolutional neural networks (2006).
[33] Center, B.D.: Case western reserve university seeded fault test data. Case Western Reserve University: Cleveland, OH, USA (2015).
[34] Ali, J.B., Saidi, L., Mouelhi, A., Chebel-Morello, B., Fnaiech, F.: Linear feature selection and classification using pnn and sfam neural networks for a nearly online diagnosis of bearing naturally progressing degradations. Engineering Applications of Artificial Intelligence 42, 67-81 (2015). https://doi.org/10.1016/j.engappai.2015.03.013.
[35] Wang, S., Xiang, J., Zhong, Y., Zhou, Y.: Convolutional neural network-based hidden markov models for rolling element bearing fault identification. Knowledge-Based Systems 144, 65-76 (2018). https://doi.org/10.1016/j.knosys.2017.12.027.
[36] Li, X., Zhang, W., Ding, Q.: Cross-domain fault diagnosis of rolling element bearings using deep generative neural networks. IEEE Transactions on Industrial Electronics 66(7), 5525-5534 (2018). https://doi.org/10.1109/TIE.2018.2868023.
[37] Lu, W., Liang, B., Cheng, Y., Meng, D., Yang, J., Zhang, T.: Deep model-based domain adaptation for fault diagnosis. IEEE Transactions on Industrial Electronics 64(3), 2296-2305 (2016). https://doi.org/10.1109/TIE.2016.2627020.
[38] Yang, B., Lei, Y., Jia, F., Xing, S.: An intelligent fault diagnosis approach based on transfer learning from laboratory bearings to locomotive bearings. Mechanical Systems and Signal Processing 122, 692-706 (2019). .https://doi.org/10.1016/j.ymssp.2018.12.051.
[39] Zhiyi, H., Haidong, S., Lin, J., Junsheng, C., Yu, Y.: Transfer fault diagnosis of bearing installed in different machines using enhanced deep auto-encoder. Measurement 152, 107393 (2020). https://doi.org/10.1016/j.measurement.2019.107393.
[40] Patil, R.: Noise reduction using wavelet transform and singular vector decomposition. Procedia Computer Science 54, 849–853 (2015). https://doi.org/10.1016/j.procs.2015.06.099.
[41] Galdran, A., Carneiro, G., González Ballester, M.A. (2021). Balanced-MixUp for Highly Imbalanced Medical Image Classification. In: de Bruijne, M., *et al.* Medical Image Computing and Computer Assisted Intervention – MICCAI 2021. MICCAI 2021. Lecture Notes in Computer Science(), vol 12905. Springer, Cham. https://doi.org/10.1007/978-3-030-87240-3_31.