

Intelligent Mechanical Fault Diagnosis Using Multisensor Fusion and Convolution Neural Network

Tingli Xie , Xufeng Huang , and Seung-Kyum Choi

Abstract—Diagnosis of mechanical faults in manufacturing systems is critical for ensuring safety and saving costs. With the development of data transmission and sensor technologies, measuring systems can acquire massive amounts of multisensor data. Although deep learning (DL) provides an end-to-end way to address the drawbacks of traditional methods, it is necessary to do deep research on an intelligent fault diagnosis method based on multisensor data. In this article, a novel intelligent diagnosis method based on multisensor fusion (MSF) and convolutional neural network (CNN) is explored. First, a multisignals-to-RGB-image conversion method based on principal component analysis is applied to fuse multisignal data into three-channel red–green–blue (RGB) images. Then, an improved CNN with residual networks is proposed, which can balance the relationship between computational cost and accuracy. Two datasets are used to verify the effectiveness of the proposed method. The results show that the proposed method outperforms other DL-based methods in terms of accuracy.

Index Terms—Convolutional neural network (CNN), feature-level fusion, mechanical fault diagnosis, multisensor fusion (MSF).

NOMENCLATURE

1-D, 2-D	One-dimensional, two-dimensional.
AE	Auto encoder.
BF	Ball fault.
BN	Batch normalization.
CF	Chipped fault.
CNN	Convolutional neural network.
CWT	Continuous wavelet transform.
CWRU	Case Western Reserve University.
CSCoh	Cyclic spectral coherence.
DL	Deep learning.

Manuscript received May 28, 2021; revised July 9, 2021; accepted July 28, 2021. Date of publication August 4, 2021; date of current version February 2, 2022. Paper no. TII-21-2231. (Corresponding author: Seung-Kyum Choi.)

Tingli Xie and Seung-Kyum Choi are with the George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: txie67@gatech.edu; schoi@me.gatech.edu).

Xufeng Huang is with the Department of Industrial and Manufacturing Systems Engineering, University of Michigan-Dearborn, Dearborn, MI 48128 USA (e-mail: xfhuang@umich.edu).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TII.2021.3102017>.

Digital Object Identifier 10.1109/TII.2021.3102017

DBN
DNN
DTL
FFT
GW
GAN
IOF
IRF
LSTM
LFGRU
LReLU
MF
ML
MSF
NO
ORF
PCA
RF
RGB
ReLU
SSL
SAE
SSAE
STFT
TICNN
TCNN
 w_i, b
 \mathbf{X}_{raw}
 \mathbf{X}_{pca}
 \mathbf{W}^T

x_i
 x_j
 n, m
 \mathbf{T}
 S_t, k

N
 $L_g^i(h), \text{PM}^i(a, b)$
 $\text{RGBM}(a, b, C_{rgb})$
 C_{rgb}
 $\mathbf{x}, H(\mathbf{x})$

Deep belief network.
Deep neural network.
Deep transfer learning.
Fast Fourier transformation.
Gear wearing fault.
Generative adversarial network.
Combination fault.
Inner ring fault.
Long short-term memory.
Local feature-based gated recurrent units.
Leaky rectified linear unit.
Missing fault.
Machine learning.
Multisensor fusion.
Normal condition.
Outer ring fault.
Principal component analysis.
Root fault.
Red-green-blue.
Rectified linear unit.
Semisupervised learning.
Stacked auto encoder.
Stacked sparse auto encoder.
Short-time Fourier transform.
CNN with training interference.
Transfer CNN.
Weight of i th layer, the bias.
Raw signals with multisensor.
Time-domain signals with three-channel.
Matrix with three significant features of \mathbf{X}_{raw} .
Signal from sensor i .
 j sample point from signals.
Number of sensors, the sampling points.
Time complexity of the CNN.
Total size of each sample, the size of RGB image.
Total number of samples.
Signal value, the pixel matrix for each channel.
RGB pixel matrix.
Color channels of RGB.
Input vector of the layer, the output of the layer.

F Residual function.
 $\varphi(\cdot)$ Nonlinear activation function.
 $h^{(i)}$ $h^{(i)}$ is the output of the i th hidden unit.

I. INTRODUCTION

ROLLING-ELEMENT bearings and gears are widely used in intelligent manufacturing. Reliability and availability of transmission components are essential to ensure safe and continuous operation in the modern industry [1]. As a result, it is necessary to develop intelligent and accurate fault diagnosis methods for the early detection of potential failure.

Existing research works are mainly model-based methods and data-driven methods. Model-based methods are developed by analyzing the characteristics of signals in a different domain (e.g., time domain, frequency domain, or time-frequency domain), which are effective when single faults occur in the machinery under stable conditions [2]. Different from model-based methods, data-driven methods are achieved by machine learning (ML) algorithms, such as artificial neural networks, support vector machines, random forests, and extreme learning machine [3], [4], which are suitable for complicated fault diagnosis problems. Nevertheless, these methods rely strongly on the hand-crafted feature extraction processes, which are time-consuming and demanding of costly specialized knowledge.

Alternatively, deep learning (DL) provides a promising end-to-end solution to overcome the above drawbacks. Due to the distributed features learning of DL, DL models can automatically learn hierarchy features and correlations among massive data [5]. Zhu *et al.* [6] adopted principal component analysis (PCA) and deep belief network (DBN) for the intelligent fault diagnosis of bearings. This PCA-DBN method was evaluated by the Case Western Reserve University (CWRU) bearing dataset and achieved a test accuracy of 90%, which was inadequate for industry applications. A hybrid framework based on recurrent neural networks was introduced [7]. This algorithm applied local feature-based gated recurrent units (LFGRU) to perform diagnosis and prognosis. Wen *et al.* [8] developed an effective data preprocessing method, which can convert raw vibration signals into gray images. After data preprocessing, a classification model based on convolutional neural network (CNN) was adopted to discriminate faults from converted images. To improve the generalization and accuracy of CNN, Zhu *et al.* [9] presented a novel automated fault diagnosis method by combining capsule network, inception block, and CNN, which can obtain superior performance in diagnosing bearing faults under different loads. Considering the small and unbalanced data, auto encoder (AE), deep transfer learning (DTL), and generative adversarial network (GAN) are also employed for fault diagnosis. Wen *et al.* [10] designed a DTL-based diagnosis model using AE and maximum mean discrepancy, which achieved an accuracy of 99.82% in the CWRU bearing dataset. Li *et al.* [11] introduced a semisupervised learning (SSL) model using AE and unsupervised learning, which can be trained with labeled data and unlabeled data simultaneously. This method achieved an accuracy of 99.99% in the CWRU bearing dataset. Saufi *et al.* [12] adopted stacked sparse AE (SSAE) to deal with limited sample data,

TABLE I
SUMMARY OF ML- AND DL-BASED MODELS FOR FAULT DIAGNOSIS

Input	Model	Publication
FFT signal	DTL-AE	Wen et al. 2019 [10]
	SSL-AE	Li et al. 2020 [11]
Vibration signal	PCA-DBN	Zhu et al. 2019 [6]
	GAN	Zhou et al. 2020 [13]
	1-D CNN	Zhang et al. 2018 [19]
Vibration and acoustic signal	1-D CNN	Wang et al. 2020 [20]
Gray-scale vibration image	2D CNN (LeNet-5)	Wen et al. 2018 [8]
Gray-scale STFT image	CNN+ Capsule Network	Zhu et al. 2019 [9]
Gray-scale CWT image	TCNN (VGG-16)	Shao et al. 2019 [14]
Gray-scale kurtogram image	SSAE	Saufi et al. 2020 [12]
Vibration signal+ Reshaped 2-D matrix	PCA+GBDT	Liu et al. 2021 [21]
	LFGRU	Zhao et al. 2018 [7]
	2-D CNN	Xia et al. 2018 [16]
	1-D atrous CNN	Li et al. 2020 [22]
FFT signal+ Reshaped 2-D matrix	2-D CNN	Azamfar et al. 2020 [18]
Statistical features+ Reshaped 2-D matrix	SAE	Liu et al. 2018 [23]
Hilbert transform+ Reshaped 2-D matrix	CNN+LSTM	Wang et al. 2019 [24]
Gray-scale current image	2-D CNN	Hoang. 2020 [25]
Gray-scale CSCoh image	2-D CNN	Chen et al. 2019 [26]
RGB vibration image	TCNN (ResNet-50)	Wen et al. 2020 [15]

which achieved a classification performance of 100% and 99% on two gearbox datasets. Similarly, a novel GAN-based fault diagnosis method was proposed [13], which applied a global optimization GAN to generate features of unbalanced faults. Shao *et al.* [14] used a pretrained deep convolutional network called VGG-16 to classify different faults, which achieved state-of-the-art results on three datasets. Similarly, Wen *et al.* [15] proposed a new Transfer CNN (TCNN) using ResNet-50 as a feature extractor, named TCNN (ResNet-50), for fault diagnosis. Table I provides a summary of the ML and DL models for fault diagnosis.

Nonetheless, most fault diagnosis methods only use information of a single-channel signal or arbitrarily convert the 1-D vibration data into a square matrix as the input of the DL model [16]. Recent research has shown that a diagnosis system using multiple sensor resources with sensor fusion techniques can provide improved and robust diagnostic results [17], [18]. Therefore, this study aims to develop an intelligent fault diagnosis method to provide an end-to-end and accurate diagnosis of key faults in different rotating machines. This method is based on multisensor fusion (MSF) and CNN, which is suitable for diagnosis problems with multisource signals and has advantages in the improvement of classification accuracy. The main contributions of this article are summarized in the following.

- 1) A feature-level fusion method using PCA and signal-to-image conversion method is adopted to fuse and convert

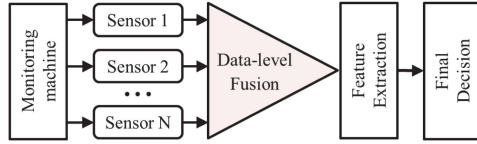


Fig. 1. Flowchart of data-level fusion.

- multisensor signals into red–green–blue (RGB) fused images.
- 2) An improved CNN using residual learning is proposed, which could automatically extract the representative features and fully utilize the multisensor data.
 - 3) Different experimental cases, including the Konstruktions- und Antriebstechnik (KAT) bearing dataset and Gearbox dataset, are used to validate the effectiveness of the proposed method.

The rest of this article is organized as follows. Section II introduces the background of multisensor data fusion and CNN. Section III demonstrates the details of the proposed approach. Section IV presents the experiments and testing results to demonstrate the applicability of the proposed method. Finally, Section V concludes this article.

II. BACKGROUND

A. Fault Diagnosis Systems With Multisensor Fusion

In the modern industry, monitoring big data of machines holds the following four characteristics [27]:

- 1) massive data;
- 2) low-value density;
- 3) heterogeneous or multisensor data;
- 4) high-speed data streams.

The massive information available in the modern industry motivates the application of MSF and intelligent algorithms for addressing specific industrial needs and requirements [28]. The algorithms that only use information from single-channel signals are no longer suitable for intelligent diagnosis systems. Therefore, fault diagnosis systems with MSF have become a research hotspot recently. Specifically, MSF can be divided into three levels: data-level fusion, feature-level fusion, and decision-level fusion. The data-level fusion model can achieve high accuracy performance with less loss of data, as shown in Fig. 1. Xia *et al.* [16] proposed a novel fault diagnosis method using CNN and data-level fusion, which fused the information from vibration signals to 2-D matrices. Chen *et al.* [26] developed a deep CNN based data-level fusion for fault diagnosis, which combined the signals from vertical and horizontal vibration sensors. Li *et al.* [22] demonstrated a novel adaptive data fusion method based on 1-D CNN and adaptive-size convolution kernel, which can fully utilize the information from the three orthogonal directions of vibration sensors. Considering current signals, Azamfar *et al.* [18] introduced a 2-D CNN architecture based on current signature analysis for fault diagnosis, which directly used raw frequency data from multiple current sensors. Unfortunately, the data-level fusion model lacks error correction ability, which results in inferior performance if the sensors are working unsymmetrically or

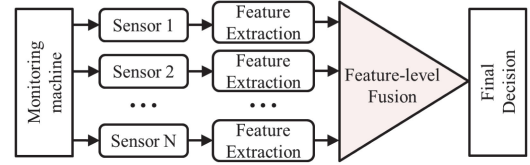


Fig. 2. Flowchart of feature-level fusion.

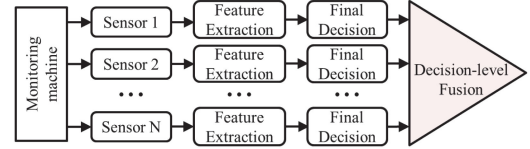


Fig. 3. Flowchart of decision-level fusion.

with different types. In real time or online applications, different sensors are not completely symmetrical due to disturbances or faulty installation. Also, it is invalid to directly fuse the raw data from different types of sensors (e.g., thermal sensor and acoustic sensor) because of the different units and dimensions (1-D with 2-D).

The feature-level fusion model carries out the dimensionality reduction of massive data, which is suitable for real-time fault diagnosis, as shown in Fig. 2. Compared with data-level fusion, information from feature-level fusion is more refined and comprehensive to demonstrate characteristics of data. Recently, feature-level fusion has been frequently combined with the DL model to obtain better diagnosis results. Liu *et al.* [23] adopted the SAE and feature-level fusion for gearbox fault diagnosis, which can capture more abstract features from multisensor acquisition channels. Wang *et al.* [24] designed an end-to-end fault diagnosis framework using CNN and long short-term memory (LSTM), which effectively achieved multiscale features fusion of motor vibration and stator current signals. Similarly, Wang *et al.* [20] investigated the multimodel sensor fusion using 1-D CNN for fault diagnosis, which extracted and fused the features from an accelerometer and a microphone. Liu *et al.* [21] introduced a novel train diagnosis model using PCA and ensemble learning, which achieved higher accuracy compared with the traditional algorithms.

During decision-level fusion, final decisions are first completed by using extracted features from each sensor signal, as shown in Fig. 3. The decision-level fusion model can fuse multimodal data and obtain exact fusion results, which has the advantage of strong robustness [17]. To date, the decision-level fusion method is rarely combined with the DL model for fault diagnosis. Hoang and Kang [25] suggested a current signals-based bearing fault diagnosis method using CNN and decision-level fusion, which can eliminate duplicate information from current phases and achieve higher performance.

In this article, we used the feature-level fusion combined with PCA, which not only does not require specialized knowledge but also achieves superior performance.

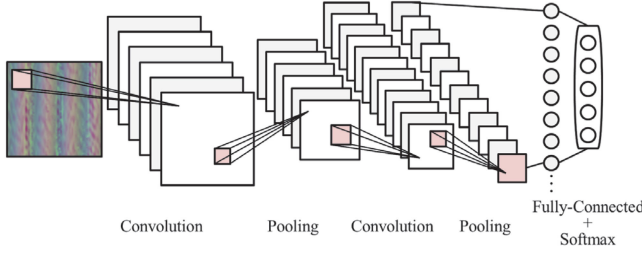


Fig. 4. Basic structure of CNN for image classification.

B. Convolutional Neural Network

CNN [29] is a significant class of deep neural network (DNN), which can perform end-to-end image classification by automatic extraction of local features. Much well-known work has been reported in the construction of CNN models, such as ResNet [30], EfficientNet [31], 1-D CNN [31], [32], and so on. Though a great number of variants of the CNN model have been developed, the basic structure of CNN for image classification is shown in Fig. 4, which contains a convolution layer, pooling layer, fully connected layer, and Softmax layer.

The Convolution layer utilizes the convolution operation to filter redundant resources in the original images and strengthen the important task-related information, which can perform automatic high-level features extraction. In addition, an activation function [e.g., the rectified linear unit (ReLU) [33]] is usually deployed after the convolution operation, which is beneficial for the acceleration of convergence and alleviation of the vanishing gradient problem. The pooling layer is regularly adopted with the convolution layer alternately, which is adapted for the completion of translation-invariant characteristics and dimensionality reduction of parameters. After feature mapping by the convolution layer and pooling layer, the fully connected layer is finally designed to convert the features into 1-D vectors and to accomplish the classification for different tasks. In the Softmax layer, the representative vectors from the fully connected layer are reshaped and mapped into a probability distribution for classification. Finally, the whole CNN is trained and adjusted by the back-propagation algorithm based on the gradient optimization algorithm. After the training process, the parameters of CNN are optimized. In this way, an optimal CNN is obtained. It can be applied to final image classification.

III. PROPOSED METHOD

A. Multisignals-to-RGB-Image Conversion Method

In order to convert raw signals into images for faults classification, time-frequency imaging methods, including continuous wavelet transform (CWT), short-time Fourier transform (STFT), etc., are applied in some researches [34]. However, these methods are not only time-consuming but also demanding of practical knowledge and experience. Additionally, these studies do not pay much attention to multisource signal problems in fault diagnosis. In this section, an effective multisignals-to-RGB-image conversion method is proposed for data preprocessing,

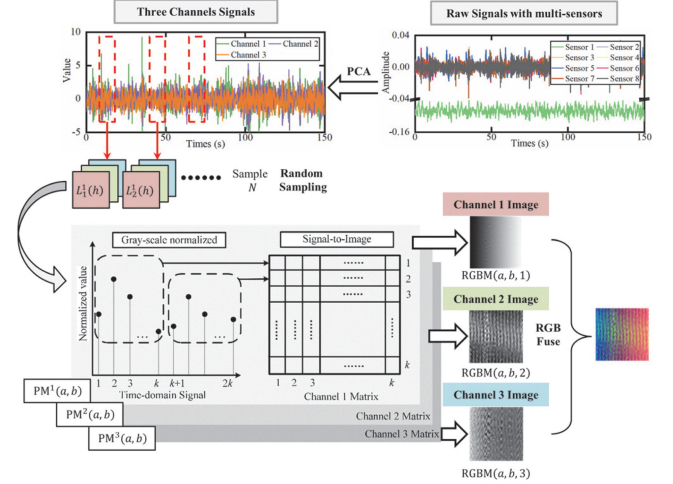


Fig. 5. Flowchart of the proposed multisignals-to-RGB-image conversion method.

which can convert multisensor signals into images and obtain feature-level fused information.

Fig. 5 depicts the framework of the proposed multisignals-to-RGB-image conversion approach. The rest of Section III-A illustrates the detailed steps as follows.

Step 1: Multiple measurement data (i.e., multisensor signals) are collected from the mechanical system. Specifically, the raw signals are resampled to the highest sampling frequency among them using a polyphase antialiasing filter, if they have different sampling frequencies. In this way, each signal has the same length (i.e., sampling points). Step 1 takes $O(1)$ time.

The multisensor raw signals are expressed by

$$\mathbf{X}_{\text{raw}} = \{x_{ij} : x_{ij} \in R^{n \times m}\}, i = 1, 2, \dots, n, j = 1, 2, \dots, m \quad (1)$$

where \mathbf{X}_{raw} is a matrix with n rows and m columns, $x_i \in R^{1 \times m}, i = 1, 2, \dots, n$ is the raw signal from sensor i , n is the number of sensors, m is the number of sampling points, and $x_j \in R^{n \times 1}, j = 1, 2, \dots, m$ represents one sample.

Step 2: To reduce the impact of invalid data and prepare for subsequent transformations, the multisensor raw signals are preprocessed for dimension reduction using PCA. Notably, no trainable parameters need to be trained during this step. Indeed, to get the three-channel RGB image values, only the first three principal components are retained and transformed to three-channel time-domain signals.

The signal value of \mathbf{X}_{raw} matrix is approximated using PCA as follows:

$$\mathbf{X}_{\text{pca}} = \mathbf{W}^T \mathbf{X}_{\text{raw}} = \{x_{ij} : x_{ij} \in R^{3 \times m}\}, i = 1, 2, 3, j = 1, 2, \dots, m \quad (2)$$

where $\mathbf{X}_{\text{pca}} \in R^{3 \times m}$ is the three-channel time-domain signals and $\mathbf{W}^T \in R^{3 \times n}$ is the matrix with three significant features of \mathbf{X}_{raw} .

After this step, the modified three-channel signals are generated, which have a lower dimension and contain only primary

information compared to raw signals. The time complexity of Step 2 is $O(\min\{n^3 + m \cdot n^2, m^3 + n \cdot m^2\})$.

Step 3: Randomly split the three-channel time-domain signals and generate the signal samples for different fault types. Notably, the total size of each sample is $S_t = k \times k \times 3$, where $k \times k \in [4, m]$ is the size of the RGB image. The splitting point of each sample is first randomly sampled from uniform distributions in $[1, m - k \times k]$ without overlap. After that, the signal samples with length $k \times k$ are generated according to the corresponding splitting points.

Step 4: Convert the samples of the segmented signal in each channel to a pixel matrix and normalize their values.

Given the total number of samples N and the signal value of $L_g^i(h) \in \mathbf{X}_{pca}, g = 1, 2, \dots, N, h = 1, 2, \dots, k^2$, the pixel matrix for each channel $PM^i(a, b), a = 1, \dots, k, b = 1, \dots, k, i = 1, 2, 3$ can be formulated by

$$PM^i(a, b) = \frac{L^i((a-1) \times k + b) - \min(L^i)}{\max(L^i) - \min(L^i)} \times 255, i = 1, 2, 3. \quad (3)$$

In this way, each value of the pixel matrix is normalized between 0 and 255. The total time complexity of Step 3–Step 4 is $O(N)$.

Step 5: Fuse the three matrices into one RGB pixel matrix to generate the RGB images. The feature-level fused information is obtained in this step, which takes $O(N)$ time.

The RGB pixel matrix $RGBM(a, b, C_{rgb}), C_{rgb} = 1, 2, 3$ is given by

$$RGBM(a, b, C_{rgb}) = (PM^i(a, b), C_{rgb}), \quad \text{where } C_{rgb} = i = 1, 2, 3. \quad (4)$$

Step 6: These RGB images represented by RGB pixel matrices $RGBM(a, b, C_{rgb})$ are divided into training, validation, and testing portions, which takes $O(1)$ time. The RGB images from the training dataset are first fed into the proposed CNN model to adjust the hyperparameters of the model. To prevent overfitting during the training process, it is necessary to periodically evaluate the model after each training epoch. Indeed, the validation dataset is used to evaluate the generalization error during the training phase, allowing the model hyperparameters to be updated consequently. The training phase will be stopped when the loss or validation metrics (i.e., accuracy and F1-score) decrease slightly or even start to increase. In this context, the optimal model can be selected for the final assessment. The testing dataset is then used to evaluate the performance of the trained model.

B. Improved CNN With Residual Learning

Given the input shape of $64 \times 64 \times 3$ and the categories of N_{class} , the structure of our improved CNN model is shown in Fig. 6.

The Conv block and residual block are first designed to extract high-level features of the RGB image automatically, which are then merged with multiple convolution and batch normalization layers to enhance the valuable task-related information

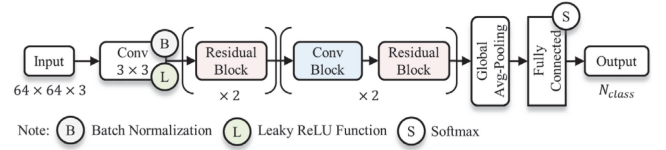


Fig. 6. Structure of the proposed CNN.

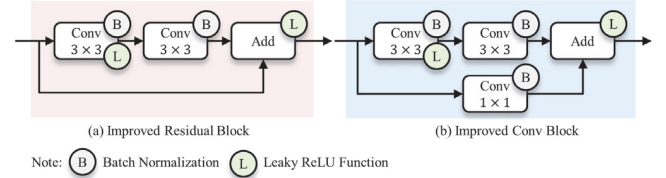


Fig. 7. Details of improved residual block and Conv block.

and avoid overfitting. Finally, the global avg-pooling layer is employed for global dimension reduction of feature parameters and achievement of translation-invariant characteristics. The fully connected layer is applied to convert the features from the global avg-pooling layer into 1-D vectors and to realize the classification for different tasks with the Softmax layer. The computational complexity of the proposed CNN is calculated by [35], which can be given by

$$\mathbf{T} = O(3^2 \cdot (3 + 16^2 \cdot 4 + 16 \cdot 32 + 32^2 \cdot 3 + 32 \cdot 64 + 64^2 \cdot 3)) = O(37467). \quad (5)$$

Notably, the computational complexity mentioned above represents the theoretical time complexity instead of actual running time (i.e., training time and testing time), which depends on the number of parameters of the CNN model (i.e., the weights of the convolutional layers).

Residual learning allows the information (i.e., features) to transmit from one layer directly to the following layers. A basic residual learning block is shown in [30], which can be defined as

$$H(x) = F(x, \{\mathbf{w}_i\}) + x \quad (6)$$

where x is the input vector of the layer, $H(x)$ denotes the output of the layer, and F is the residual function.

Given the nonlinear activation function $\varphi(\cdot)$, the output of the first layer $\varphi(\cdot)$ can be defined as $F(x, \{\mathbf{w}_i\}) = \mathbf{w}_2 \varphi(\mathbf{w}_1^T x) + \mathbf{b}$, where \mathbf{w}_2 and \mathbf{b} are the weight and bias of the second layer, respectively. Then, the operation $F(x, \{\mathbf{w}_i\}) + \mathbf{b} + x$ is performed by an identity shortcut connection and elementwise addition, which add neither redundant parameters nor computational cost. In this context, the degradation problem is addressed.

The details of residual block and Conv block in the proposed CNN are shown in Fig. 7. The original residual network adopted the ReLU to perform a nonlinearity function (i.e., activation function). The ReLU can effectively address the vanishing gradient problem suffering from sigmoidal nonlinearities. The ReLU function can be given by the following equation:

$$h^{(i)} = \max(w_i x, 0) = \begin{cases} w_i x & w_i x > 0 \\ 0 & \text{else} \end{cases} \quad (7)$$

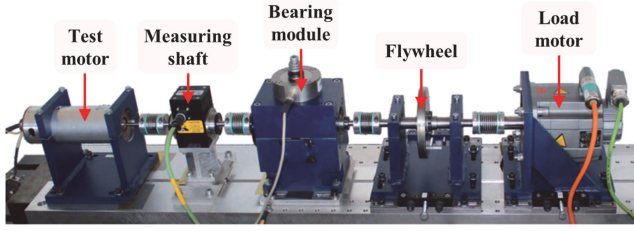


Fig. 8. Test rig of the KAT bearing dataset [37].

where $h^{(i)}$ is the output of the i th hidden unit, w_i is the weight of the i th hidden unit, and x is the input.

However, it may not be active if the gradient is 0, which will result in slow training or overfitting. To alleviate potential problems caused by ReLU, we used leaky rectified linear unit (LReLU) [36] as the activation function. The LReLU function can be defined by the following equation:

$$h^{(i)} = \max(w_i x, 0) = \begin{cases} w_i x & w_i x > 0 \\ \alpha w_i x & \text{else} \end{cases} \quad (8)$$

where $\alpha = 0.01$. The LReLU function enables a small amount of information to flow when the input is less than zero, which is potentially more robust during optimization.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, two case studies are used to demonstrate the capability of the proposed method, including the KAT bearing dataset and Gearbox dataset. All experiments were implemented in MATLAB R2020b and run on Windows 10 with an Nvidia RTX 2080 GPU.

A. Description of Experimental Datasets

1) *KAT Bearing Dataset*: The KAT bearing dataset is provided by the KAT-DataCenter of the Chair of Design and Drive Technology at Paderborn University [37]. The mechanical setup of the test rig in the KAT bearing dataset is shown in Fig. 8, which contains five basic components: a test drive motor, a measurement shaft, the bearing module, a flywheel, and a load motor.

The test drive motor is a 425-W Hanning synchronous motor, which is controlled by a standard industrial inverter (KEB Combivert 07F5E 1D-2B0A) with a switching frequency of 16 kHz. In this experiment, the motor current and vibration signals were measured synchronously with the temperature of oil in the bearing module, the rotary speed, the load torque on the torque-measuring shaft, and the radial force on the bearing [25]. Notably, the motor currents were measured in two phases at a sampling rate of 64 kHz by the current transformer (LEM type CKSR 15-NP, Bessel-Filter fourth order with a filter frequency of 25 kHz). The acceleration of the bearing module (vibration signals) was measured by a piezoelectric accelerometer (Model No. 336C04, PCB Piezotronics Inc.) with a sampling rate of 64 kHz. Also, the radial force on the bearings and the load torque at the measurement shaft were measured with a sampling rate of 4 kHz by a measuring box K11 (Lorents-Messtechnik)

TABLE II
WORKING CONDITIONS IN THE KAT BEARING DATASET

Name of setting	Rotating speed (rpm)	Load torque (N·m)	Radial force (N)
B1	1500	0.1	1000
B2	1500	0.7	400
B3	1500	0.7	1000
B4	900	0.7	1000

TABLE III
HEALTH CONDITIONS IN THE KAT BEARING DATASET

Types	Description	Number of samples
NO	Normal condition	2000
IRF	Real damage occurs in the inner ring	2000
ORF	Real damage occurs in the outer ring	2000

TABLE IV
EXAMPLES OF CONVERTED RGB IMAGES IN THE KAT BEARING DATASET UNDER THE B1 CONDITION

Types	NO	IRF	ORF
Channel 1			
Channel 2			
Channel 3			
RGB images			

and a Torque Transducer Model 305, respectively. Therefore, the KAT bearing dataset contains multisensor signals from five sensors (two phases of currents, one vibration signal, one force, and one torque). For consistency of signal length, the force and torque signals were resampled to 64 kHz using a polyphase antialiasing filter. The experiments were performed under four different working conditions, as shown in Table II.

The KAT bearing dataset consists of measurements from 32 different bearing experiments, and 15 datasets can be categorized into three health classifications as shown in Table III, which include normal condition (NO), inner ring fault (IRF), and outer ring fault (ORF). Each fault type consists of 2000 samples, and the entire dataset contains 6000 images. Table IV shows examples of converted RGB images in the KAT bearing dataset under the B1 condition.

2) *Gearbox Dataset*: The Gearbox dataset was collected from the Drivetrain Dynamics Simulator [7], [14]. The mechanical setup of the test rig in the Gearbox dataset is shown in Fig. 9, which contains four basic units: a motor, a planetary gearbox, a parallel gearbox, and a brake. In the experiments, the vibration signals in the three orthogonal directions (i.e., X, Y, and Z) around the planetary and parallel gearboxes were measured, and the vibration and torque signals of the motor were also measured. The seven vibration signals were all measured by vibration sensors (Model 608A11, frequency range of 0.5–10 kHz), and the torque signal was measured by a torque sensor (Model

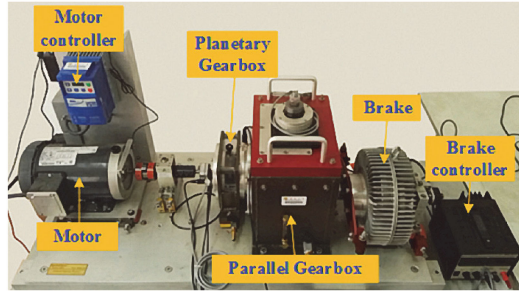


Fig. 9. Test rig of the Gearbox dataset [7], [14].

TABLE V
WORKING CONDITIONS IN THE GEARBOX DATASET

Name of setting	Rotating frequency	Load configuration
G1	20 Hz	0 V
G2	30 Hz	2 V

TABLE VI
HEALTH CONDITIONS IN THE GEARBOX DATASET

Categories	Types	Description	Number of samples
Health state	NO	Normal condition	500
Gear faults	CF	A crack occurs in the gear feet	500
	MF	One of the gear feet is missing	500
	RF	A crack occurs in the root of gear feet	500
	GW	Wear occurs in the surface of the gear	500
Bearing faults	BF	A crack occurs in the ball	500
	IRF	A crack occurs in the inner ring	500
	ORF	A crack occurs in the outer ring	500
	IOF	A crack occurs in both inner and outer ring	500

FT293, measuring range: ± 5 V). These signals were collected by a 20-channel data acquisition instrument with a sampling rate of 1 kHz and a 512-s sampling window. Therefore, the Gearbox dataset contains multisensor signals from eight sensors. The experiments were conducted with two different working conditions, as shown in Table V.

The Gearbox dataset contains two sub-datasets. One is the gear sub-dataset that includes one health state, which is NO, and four gear faults, which are chipped fault (CF), missing fault (MF), root fault (RF), and gear wearing fault (GW). The other is the bearing sub-dataset that includes one health state, which is NO, and four bearing faults, which are ball fault (BF), IRF, ORF, and combination fault (IOF), as shown in Table VI. Each sub-dataset contains five different health conditions.

Each fault type consists of 500 samples, and the gear and bearing sub-datasets each contain 2500 images. To test the proposed method when dealing with mixed faults, gear faults and bearing faults are combined to form a mixture dataset that includes four kinds of gear failure, four kinds of bearing failure, and one health state. Each health condition contains 500 samples, and the whole dataset contains 4500 samples.

TABLE VII
EXAMPLES OF CONVERTED RGB IMAGES IN THE GEARBOX DATASET UNDER THE G1 CONDITION

Categories	Health	Gear faults				Bearing faults			
Types	NO	CF	MF	RF	GW	BF	IRF	ORF	IOF
Channel 1									
Channel 2									
Channel 3									
RGB images									

TABLE VIII
DETAILED DESCRIPTIONS OF COMPARED METHODS

Method	Input data	Model
Raw-TICNN [19]	Raw vibration signal	1-D CNN
Matrix-CNN [16]	Reshaped 2-D matrix	2-D CNN
Gray-LeNet-5 [8]	Gray-scale vibration image	LeNet-5
CWT-VGG16 [14]	Gray-scale CWT image	TCNN (VGG16)
RGB-ResNet [15]	RGB vibration image	TCNN (ResNet-50)

TABLE IX
EVALUATION METRICS

Metrics	Calculation Equations
Accuracy	$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$
Recall	$Recall = \frac{TP}{TP + FN}$
Precision	$Precision = \frac{TP}{TP + FP}$
F1-Score	$F1-Score = \frac{2 * Precision * Recall}{Precision + Recall}$

where TP, TN, FP, and FN represent the number of true positives, true negatives, false positives, and false negatives, respectively.

Table VII shows examples of converted RGB images in the Gearbox dataset under the G1 condition.

B. Compared Methods and Evaluation Metrics

Table VIII shows detailed descriptions of the compared methods. Different formulations of input data are selected, including raw vibration signal, reshaped 2-D matrix, gray-scale vibration image, gray-scale CWT image, and RGB vibration image, which are different from the proposed method. Particularly, the size of the image is set to be 64, while the size of the data points of raw signals is set to be 4096. The computational complexities of the compared methods are calculated by [35]. The time complexities of Raw-TICNN, Matrix-CNN, Gray-LeNet-5, CWT-VGG16, and RGB-ResNet are $O(90772)$, $O(2117120)$, $O(97408)$, $O(138357544)$, and $O(25636712)$, respectively. The CWT-VGG16 and RGB-ResNet, which are based on TCNN, have higher time complexity compared with other methods.

To evaluate the performance of the proposed approach, accuracy and F1-score are used, as shown in Table IX. Notably, the

TABLE X
DIAGNOSIS PERFORMANCE OF DIFFERENT METHODS UNDER DIFFERENT WORKING CONDITIONS ON THE KAT DATASET

Methods	Testing accuracy (%) (Mean)					Average F1-score (%)
	B1	B2	B3	B4	Average	
Proposed Method	100	100	99.99	99.99	99.99	99.99
Raw-TICNN	96.53	93.56	97.34	96.67	96.03	96.03
Matrix-CNN	96.80	95.87	94.94	90.33	94.49	95.04
Gray-LeNet-5	93.94	88.79	89.08	91.24	90.76	91.45
CWT-VGG16	98.16	98.70	98.66	96.48	97.99	97.99
RGB-ResNet	99.66	99.60	99.70	99.17	99.53	99.53

The bold values represent not only the performance of the proposed method, but also the best results.

testing accuracy is used for the evaluation of different methods, which is calculated using the testing dataset in Section III Step 6. Similarly, the F1-score is also determined by the testing dataset, and it is the harmonic mean of the recall and precision. The higher the values of the testing accuracy and F1-score are, the more accurate the model is.

C. Diagnostic Performances on KAT Bearing Dataset

During the training procedure, Adam was used to optimize with a batch size of 64. The initial learning rate was defined as $5e-4$ and decreased by a factor of 0.5 every 8 epochs. In our experiments, the total number of training epochs was defined as 30. Unless otherwise stated, all experiments were performed 10 times to avoid contingency during the test, and the average values were regarded as the final results of classification for analysis. Also, the dataset is randomly split into subsets of training, validation, and testing. The ratios of each subset are defined as 70%, 15%, and 15%, which are commonly used.

Table X presents the testing results of different methods under different working conditions on the KAT dataset. From the diagnosis performance shown in Table X, it can be seen that the proposed method has the best performance, achieving testing accuracies of 100%, 100%, 99.99%, and 99.99% in four different working conditions, respectively. The proposed method also achieves superior performance with an average F1-score of 99.99%, while the performance of other methods only has F1-scores of 96.03%, 95.04%, 91.45%, 97.99%, and 99.53%, respectively. The corresponding confusion matrix of the best result of the proposed method on the B4 working condition is shown in Fig. 10(a). The y-axis is the actual label, and the x-axis is the predicted label. As shown in Fig. 10(a), all types of health conditions have a test accuracy of 100%, which means the proposed method can effectively recognize each fault on the KAT bearing dataset with a 70% training ratio. Therefore, the effectiveness of the proposed method on the KAT bearing dataset with a full training ratio is initially validated. The bold values in Table X and Table XI represent not only the proposed method, but also the best results.

Due to the dependency on the amount of training data for the DL-based method, it is difficult for the DL-based method

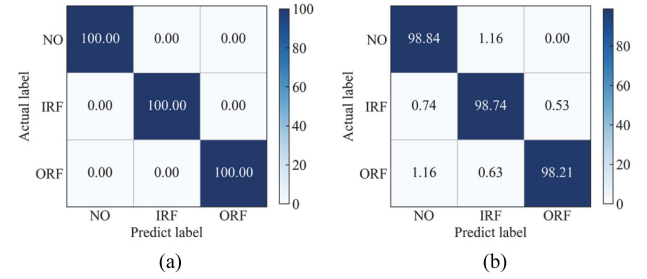


Fig. 10. Confusion matrix of the proposed method on the KAT dataset under B4 condition with (a) 70% training ratio and (b) 5% training ratio.

TABLE XI
TESTING ACCURACY WITH 5% TRAINING RATIO ON THE KAT DATASET

Methods	Testing accuracy (%) (Mean)					Average F1-score (%)
	B1	B2	B3	B4	Average	
Proposed Method	97.09	96.96	94.79	97.76	96.65	96.68
Raw-TICNN	86.69	80.42	89.23	87.18	86.63	86.65
Matrix-CNN	82.16	65.96	72.87	65.09	71.52	71.71
Gray-LeNet-5	71.47	68.09	71.01	61.54	68.02	68.04
CWT-VGG16	85.63	84.04	87.64	82.25	84.89	84.92
RGB-ResNet	92.18	92.62	93.33	86.26	91.10	91.21

The bold values represent not only the performance of the proposed method, but also the best results.

to obtain satisfying results with limited data. Therefore, we used a 5% training ratio to further validate the effectiveness and robustness of the proposed method. During the training procedure, Adam was used to optimize with a batch size of 64. Due to the limited data, the initial learning rate was defined as $1e-3$ and decreased by a factor of 0.5 every 25 epochs. In this experiment with reduced training samples, the total number of training epochs was defined as 100.

Table XI presents the testing results of different methods with 5% training ratio on the KAT dataset. From the classification results shown in Table XI, the proposed method can accurately classify bearing operating states with a limited training ratio, with an average testing accuracy of 96.65% and an F1-score of 96.68%, outperforming all other methods. Specifically, the average testing accuracies of Gray-LeNet-5, CWT-VGG16, and RGB-ResNet based on the signal-to-image method are 68.02%, 84.86%, and 91.10%, respectively, inferior to that of the proposed method. Although the Raw-TICNN has perfect bearing diagnostic performance on the CWRU dataset [19], which is based on raw signal and 1-D CNN, its accuracy is 10.02% lower than that of the proposed method with 5% training ratio on the KAT dataset. In addition, the performance of matrix-CNN based on data-level fusion is still inferior to the proposed method, with an average testing accuracy of only 71.52%. Thus, even with the limitation on the training dataset, the proposed method outperforms other CNN-based models.

The corresponding confusion matrix of the best result of the proposed method on the B4 working condition is given in Fig. 10(b). It demonstrates that the NO has the highest misdiagnosis rate; 1.16% of ORF and 0.74% of IRF are misclassified as

TABLE XII

DIAGNOSIS PERFORMANCE OF DIFFERENT METHODS UNDER VARIOUS WORKING CONDITIONS ON THE GEARBOX DATASET

Methods	Testing accuracy (%) (Mean)						Average F1-score (%)	
	Bearing		Gear		Mixture			
	G1	G2	G1	G2	G1	G2		
Proposed Method	99.97	99.97	100	100	99.97	99.97	99.98	99.98
Raw-TICNN	99.84	97.60	99.05	99.97	98.29	98.28	98.84	98.86
Matrix-CNN	99.55	99.94	95.76	95.23	95.23	96.44	97.03	97.12
Gray-LeNet-5	99.12	99.28	89.33	91.52	95.91	95.88	95.17	95.25
CWT-VGG16	99.89	99.68	99.44	99.52	99.59	99.13	99.54	99.54
RGB-ResNet	99.84	99.55	92.27	96.32	99.57	99.70	97.87	97.96

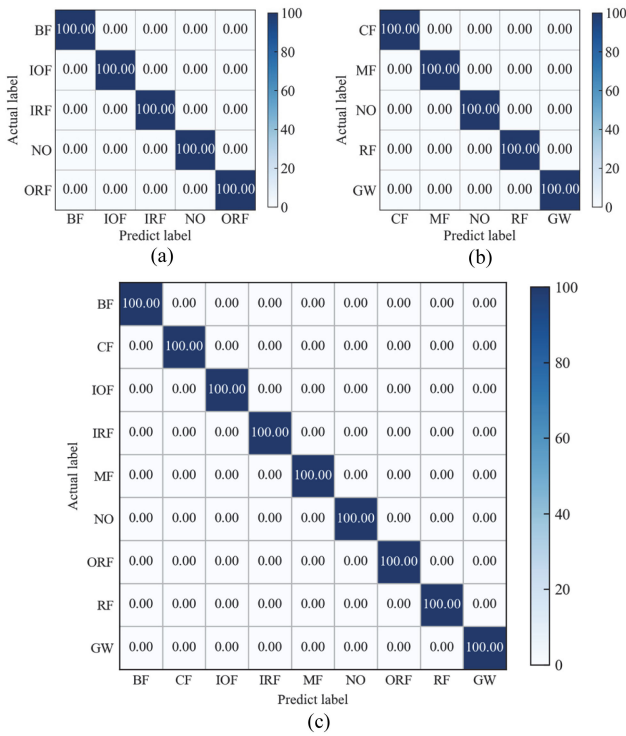


Fig. 11. Confusion matrices of the proposed method on the Gearbox dataset under G1 condition.

NO. The ORF has the lowest testing accuracy of 98.21%. For the fine-grained classification, the proposed method is significantly better than the compared methods under the limited training ratio.

D. Diagnostic Performances on Gearbox Dataset

During the training procedure, Adam was used to optimize with a batch size of 64. The initial learning rate was defined as $1e-3$ and decreased by a factor of 0.5 every 8 epochs. In our experiments, the total number of training epochs was defined as 30.

Table XII presents the testing results of different methods under various working conditions on the Gearbox dataset. According to Table XII, the proposed method has the best performance, with an average testing accuracy of 99.98% and an F1-score of 99.98%. Also, the proposed method is able to accurately classify not only health conditions of signal components (i.e., bearing

and gear) but also that of mixture datasets. Compared with the signal-to-image based methods (i.e., Gray-LeNet-5, CWT-VGG16, and RGB-ResNet), the proposed method achieves an average accuracy improvement of 4.81%, 0.44%, and 2.11%, respectively. The proposed method is 2.95% more accurate compared to matrix-CNN, which means the proposed method using CNN and MSF can effectively handle information from multiple sensors.

The confusion matrices of the best result of the proposed method on the G1 working condition are shown in Fig. 11. From these matrices, all categories of health conditions have testing accuracies of 100%. Thus, these results demonstrate the proposed method's robustness and effectiveness for the diagnosis of the Gearbox dataset.

V. CONCLUSION

This article presented a novel fault diagnosis method based on multisensor data fusion and improved CNN. Two different fault datasets with multisensor data and different working conditions, including the KAT bearing dataset and Gearbox dataset, were used to validate the effectiveness of the proposed method. The conclusions can be illustrated as follows.

First, a novel signal conversion approach was developed to convert the time-domain fault signals to the RGB format, which is based on PCA and can fuse 2-D features of raw multisensor data. The PCA used in the proposed multisignals-to-RGB-image conversion method can generate low-dimension signals with primary information. After that, the low-dimension signals were converted into RGB images, which were fused with the feature-level information. Compared to other signal pre-processing methods, the proposed multisignals-to-RGB-image conversion method has the advantage of easy deployment and improvement of faults classification. Second, combining with residual learning and LReLU, an improved CNN was developed, which can balance the relationship between computational cost and accuracy. With better feature extraction layers and activation function, the proposed CNN would improve the final classification accuracy on fault diagnosis. Third, two different fault datasets, the KAT bearing dataset and Gearbox dataset, were used to validate the effectiveness of the proposed method. For the KAT bearing dataset, the average testing accuracy and the F1-score of the proposed method were 99.99% and 99.99%, respectively, which are higher than the results of other models. Furthermore, when using a limited training dataset, the proposed method achieved the highest average testing accuracy of 96.65%, which shows that the proposed method has stronger robustness than other methods. For the Gearbox dataset, the proposed method achieved an average testing accuracy of 99.98%, while the average F1-score was 99.98%. The results showed that the proposed method can make a tradeoff between high accuracy and relatively low computational cost. Thus, the trained model can be directly used for fault diagnosis of different machines measured under different working conditions and multisensor datasets.

As part of future work, the proposed method will be further tested for mechanical faults with multiple sensors. In addition,

the uncertainty sources in the fault diagnosis domain will be considered and quantified based on Bayesian learning. For instance, the proposed method will give incorrect results when detecting unknown fault types. Combined with Bayesian learning, the proposed method will output the confidence interval and the probability of prediction simultaneously, which is beneficial for real applications. In addition, the multimodal fusion schemes [21] and multiscale CNN [38] will be investigated, aiming to make the proposed method handle information from sensors with different sampling frequencies and modals (e.g., 2-D image and 1-D signal).

REFERENCES

- [1] P. Gangsar and R. Tiwari, "Signal based condition monitoring techniques for fault detection and diagnosis of induction motors: A state-of-the-art review," *Mech. Syst. Signal Process.*, vol. 144, no. 4, Oct. 2020, Art. no. 106908.
- [2] Z. Chen, K. Gryllias, and W. Li, "Intelligent fault diagnosis for rotary machinery using transferable convolutional neural network," *IEEE Trans. Ind. Informat.*, vol. 16, no. 1, pp. 339–349, Jan. 2020.
- [3] Q. Hu, X.-S. Si, Q.-H. Zhang, and A.-S. Qin, "A rotating machinery fault diagnosis method based on multi-scale dimensionless indicators and random forests," *Mech. Syst. Signal Process.*, vol. 139, May 2020, Art. no. 106609.
- [4] Z. Chen, K. Gryllias, and W. Li, "Mechanical fault diagnosis using convolutional neural networks and extreme learning machine," *Mech. Syst. Signal Process.*, vol. 133, Nov. 2019, Art. no. 106272.
- [5] B. Rezaeianjouybari and Y. Shang, "Deep learning for prognostics and health management: State of the art, challenges, and opportunities," *Measurement*, vol. 163, no. 4, Oct. 2020, Art. no. 107929.
- [6] J. Zhu, T. Hu, B. Jiang, and X. Yang, "Intelligent bearing fault diagnosis using PCA-DBN framework," *Neural Comput. Appl.*, vol. 32, no. 14, pp. 10773–10781, Jul. 2020.
- [7] R. Zhao, D. Wang, R. Yan, K. Mao, F. Shen, and J. Wang, "Machine health monitoring using local feature-based gated recurrent unit networks," *IEEE Trans. Ind. Electron.*, vol. 65, no. 2, pp. 1539–1548, Feb. 2018.
- [8] L. Wen, X. Li, L. Gao, and Y. Zhang, "A new convolutional neural network-based data-driven fault diagnosis method," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5990–5998, Jul. 2018.
- [9] Z. Zhu, G. Peng, Y. Chen, and H. Gao, "A convolutional neural network based on a capsule network with strong generalization for bearing fault diagnosis," *Neurocomputing*, vol. 323, pp. 62–75, Jan. 2019.
- [10] L. Wen, L. Gao, and X. Li, "A new deep transfer learning based on sparse auto-encoder for fault diagnosis," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 49, no. 1, pp. 136–144, Jan. 2019.
- [11] X. Li, X. Li, and H. Ma, "Deep representation clustering-based fault diagnosis method with unsupervised data applied to rotating machinery," *Mech. Syst. Signal Process.*, vol. 143, Sep. 2020, Art. no. 106825.
- [12] S. R. Saufi, Z. A. B. Ahmad, M. S. Leong, and M. H. Lim, "Gearbox fault diagnosis using a deep learning model with limited data sample," *IEEE Trans. Ind. Informat.*, vol. 16, no. 10, pp. 6263–6271, Oct. 2020.
- [13] F. Zhou, S. Yang, Y. Fujita, D. Chen, and C. Wen, "Deep learning fault diagnosis method based on global optimization GAN for unbalanced data," *Knowl.-Based Syst.*, vol. 187, Jan. 2020, Art. no. 104837.
- [14] S. Shao, S. McAleer, R. Yan, and P. Baldi, "Highly accurate machine fault diagnosis using deep transfer learning," *IEEE Trans. Ind. Informat.*, vol. 15, no. 4, pp. 2446–2455, Apr. 2019.
- [15] L. Wen, X. Li, and L. Gao, "A transfer convolutional neural network for fault diagnosis based on ResNet-50," *Neural Comput. Appl.*, vol. 32, no. 10, pp. 6111–6124, May 2020.
- [16] M. Xia, T. Li, L. Xu, L. Liu, and C. W. de Silva, "Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks," *IEEE/ASME Trans. Mechatron.*, vol. 23, no. 1, pp. 101–110, Feb. 2018.
- [17] T. Meng, X. Jing, Z. Yan, and W. Pedrycz, "A survey on machine learning for data fusion," *Inf. Fusion*, vol. 57, no. 1, pp. 115–129, May 2020.
- [18] M. Azamfar, J. Singh, I. Bravo-Imaz, and J. Lee, "Multisensor data fusion for gearbox fault diagnosis using 2-D convolutional neural network and motor current signature analysis," *Mech. Syst. Signal Process.*, vol. 144, Oct. 2020, Art. no. 106861.
- [19] W. Zhang, C. Li, G. Peng, Y. Chen, and Z. Zhang, "A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load," *Mech. Syst. Signal Process.*, vol. 100, pp. 439–453, Feb. 2018.
- [20] X. Wang, D. Mao, and X. Li, "Bearing fault diagnosis based on vibro-acoustic data fusion and 1D-CNN network," *Measurement*, vol. 173, no. 6, Mar. 2021, Art. no. 108518.
- [21] Z. Liu, M. Zhang, F. Liu, and B. Zhang, "Multidimensional feature fusion and ensemble learning-based fault diagnosis for the braking system of heavy-haul train," *IEEE Trans. Ind. Informat.*, vol. 17, no. 1, pp. 41–51, Jan. 2021.
- [22] S. Li, H. Wang, L. Song, P. Wang, L. Cui, and T. Lin, "An adaptive data fusion strategy for fault diagnosis based on the convolutional neural network," *Measurement*, vol. 165, Dec. 2020, Art. no. 108122.
- [23] J. Liu, Y. Hu, Y. Wang, B. Wu, J. Fan, and Z. Hu, "An integrated multi-sensor fusion-based deep feature learning approach for rotating machinery diagnosis," *Meas. Sci. Technol.*, vol. 29, no. 5, Mar. 2018, Art. no. 055103.
- [24] J. Wang, P. Fu, L. Zhang, R. X. Gao, and R. Zhao, "Multilevel information fusion for induction motor fault diagnosis," *IEEE/ASME Trans. Mechatron.*, vol. 24, no. 5, pp. 2139–2150, Oct. 2019.
- [25] D. T. Hoang and H. J. Kang, "A motor current signal-based bearing fault diagnosis using deep learning and information fusion," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 6, pp. 3325–3333, Jun. 2020.
- [26] H. Chen, N. Hu, Z. Cheng, L. Zhang, and Y. Zhang, "A deep convolutional neural network based fusion method of two-direction vibration signal data for health state identification of planetary gearboxes," *Measurement*, vol. 146, pp. 268–278, Nov. 2019.
- [27] Y. Lei, B. Yang, X. Jiang, F. Jia, N. Li, and A. K. Nandi, "Applications of machine learning to machine fault diagnosis: A review and roadmap," *Mech. Syst. Signal Process.*, vol. 138, Apr. 2020, Art. no. 106587.
- [28] A. Diez-Oliván, J. D. Ser, D. Galar, and B. Sierra, "Data fusion and machine learning for industrial prognosis: Trends and perspectives towards industry 4.0," *Inf. Fusion*, vol. 50, pp. 92–111, Oct. 2019.
- [29] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.
- [30] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Proc. Eur. Conf. Comput. Vis.*, Amsterdam, The Netherlands, 2016, pp. 630–645.
- [31] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. 36th ICML*, Long Beach, CA, USA, 2019, pp. 6105–6114.
- [32] B. Zhao, X. Zhang, H. Li, and Z. Yang, "Intelligent fault diagnosis of rolling bearings based on normalized CNN considering data imbalance and variable working conditions," *Knowl.-Based Syst.*, vol. 199, 2020, Art. no. 105971.
- [33] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proc. 14th Int. Conf. Artif. Intell. Statist.*, Fort Lauderdale, FL, USA, 2011, pp. 315–323.
- [34] J. Wang, Z. Mo, H. Zhang, and Q. Miao, "A deep learning method for bearing fault diagnosis based on time-frequency image," *IEEE Access*, vol. 7, pp. 42373–42383, 2019.
- [35] K. He and J. Sun, "Convolutional neural networks at constrained time cost," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Boston, MA, USA, 2015, pp. 5353–5360.
- [36] A. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," in *Proc. 30th Int. Conf. Mach. Learn.*, Atlanta, GA, USA, 2013, pp. 1–3.
- [37] C. Lessmeier, J. K. Kimotho, D. Zimmer, and W. Sextro, "Condition monitoring of bearing damage in electromechanical drive systems by using motor current signals of electric motors: A benchmark data set for data-driven classification," in *Proc. Eur. Conf. Prognostics Health Manage. Soc.*, 2016, pp. 5–8.
- [38] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 9, pp. 1904–1916, Sep. 2015.



diagnosis.

Tingli Xie received the B.S. degree in mechanical design manufacturing and automation and the M.S. degree in mechanical engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2016 and 2019, respectively. She is currently working toward the Ph.D. degree in mechanical engineering with the Georgia Institute of Technology, Atlanta, GA, USA.

Her current research interests include deep learning, uncertainty quantification, and fault



His current research interests include deep learning, design under uncertainty, additive manufacturing, and fault diagnosis.

Xufeng Huang received the B.S. degree in mechanical design manufacturing and automation from the Huazhong University of Science and Technology, Wuhan, China, in 2016, and the M.S. degree in mechatronic engineering from Guangxi University, Nanning, China, in 2019. He is currently working toward the Ph.D. degree in industrial and systems engineering with the Department of Industrial and Manufacturing Systems Engineering, University of Michigan-Dearborn, Dearborn, MI, USA.



Seung-Kyum Choi received the Ph.D. degree in uncertainty quantification techniques for the analytical certification of complex engineered systems from the Mechanical and Materials Engineering Department, Wright State University, Dayton, OH, USA.

He is currently the Director of the Center for Additive Manufacturing Systems (CAMS), Georgia Tech, Atlanta, GA, USA, where he has responsibilities for developing research and educational programs in additive manufacturing.

directly began at Georgia Tech in Fall 2006 as an Assistant Professor. His reliability-based synthesis methods are critical in the development of new engineering components that are comprised of cellular material structures when those structures have inherent uncertainty associated with their geometry, loads, and material properties. His current research interests include developing robust simulation and decision-support tools that can help manage complex engineered systems that involve the design of additively manufactured parts and multifunctional products.