

Article

Research on Vibration Data-Driven Fault Diagnosis for Iron Core Looseness of Saturable Reactor in UHVDC Thyristor Valve Based on CVAE-GAN and Multimodal Feature Integrated CNN

Xiaolong Zhang ¹, Xiaoguang Wei ^{1,2,*}, Lin Zheng ¹, Chenghao Wang ¹ and Huafeng Wang ¹

¹ State Grid Smart Grid Research Institute Co., Ltd., State Key Laboratory of Advanced Power Transmission Technology, Beijing Key Laboratory of High Power Electronic, Beijing 102209, China

² Beijing Institute of Smart Energy, Beijing 102209, China

* Correspondence: weixiaoguang_sgri@163.com

Abstract: The imbalance of data samples and fluctuating operating conditions are the two main challenges faced by vibration data-driven fault diagnosis for the iron core looseness of saturable reactors in UHVDC thyristor valves. This paper proposes a vibration data-driven saturable reactor iron core looseness fault diagnosis strategy named CVG-MFICNN based on CVAE-GAN and MFICNN to overcome the two challenges. This strategy uses a novel 1-D CVAE-GAN model to produce generated samples and expand the training set based on imbalanced training samples. An MFICNN model structure is designed to allow the simultaneous processing of multimodal features such as the SST time-frequency spectrum, time-domain vibration sequence, frequency-domain power spectrum sequence, and time-domain statistics. Using these multimodal features and the MFICNN model, the hidden fault information in vibration data can be effectively mined. An experiment is conducted to collect vibration data of saturable reactors with different faults. Models based on the proposed strategy and other methods are trained and tested using the collected data. The comparison results show that the performance of the proposed CVG-MFICNN approach is significantly superior to that of single-feature CNNs, traditional machine learning methods, and classical image classification CNNs in the application of UHVDC thyristor valve saturable reactor iron core looseness fault diagnosis.

Keywords: UHVDC thyristor valve; saturable reactor; iron core looseness; fault diagnosis; data imbalance; variable operating conditions; CVAE-GAN



Citation: Zhang, X.; Wei, X.; Zheng, L.; Wang, C.; Wang, H. Research on Vibration Data-Driven Fault Diagnosis for Iron Core Looseness of Saturable Reactor in UHVDC Thyristor Valve Based on CVAE-GAN and Multimodal Feature Integrated CNN. *Energies* **2022**, *15*, 9508. <https://doi.org/10.3390/en15249508>

Academic Editor: Ahmed Abu-Siada

Received: 25 October 2022

Accepted: 2 December 2022

Published: 15 December 2022

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



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The ultrahigh voltage direct current (UHVDC) thyristor valve is the core equipment of a UHVDC transmission project [1]. A saturable reactor is a critical component of the UHVDC thyristor valve, which can inhibit the rapid growth of surge currents during the opening and closing progress of the thyristors, bear most of the peak voltage under lightning overvoltage, and protect the thyristors under various working conditions [2]. Under long-term engineering mechanical stress, the tension band of a saturable reactor iron core will undergo permanent plastic deformation, decreasing the tension force. Then, the air gap of the iron core may change. This may lead to the protection capability of the saturable reactor not meeting the design requirements. Abnormal vibration of the iron core can also occur, which may damage the insulation and cause excessive temperature due to the eddy current effect. Through routine inspection and maintenance, it is impossible to identify the iron core state. Therefore, it is necessary to study the monitoring and diagnosis method of saturable reactor iron core looseness faults to determine the state of the iron core and support reasonable operation and maintenance in engineering.

Vibration signals have been widely used in the condition monitoring and diagnosis of power equipment as a physical quantity that directly reflects the structure state of devices [3–6]. Similar to a transformer or an alternative current (AC) reactor, when a saturable reactor works, its iron core vibrates under magnetostriction. The vibration characteristics are directly affected by the iron core's structural state. Therefore, it is theoretically feasible to diagnose the iron core looseness of a saturable reactor by analyzing and processing the vibration signals.

In the early stage, the time domain statistical characteristics of the vibration signal were used to diagnose the equipment status, and the operation status of the equipment was diagnosed by calculating the root mean square (RMS), skewness, kurtosis, and other time domain statistical features [7–9]. However, the time domain characteristics do not change significantly when some faults occur, and the diagnostic methods based on time domain statistical features cannot detect and distinguish the faults in a timely manner, which creates a certain hysteresis and ambiguity in engineering. With the development of frequency domain analysis methods based on the Fourier transform, such as the frequency spectrum, power spectrum, and cepstrum, it is possible to analyze and calculate the characteristics of signals in the frequency domain [10–13]. Some faults do not obviously influence the time domain features of vibration signals but cause significant changes in frequency domain characteristics. Frequency domain analysis can diagnose these faults earlier and more accurately. However, the vibration signals of equipment faults are usually nonstationary signals, which vary both in the time domain and frequency domain [14]. Some research used time-frequency analysis methods such as a short-time Fourier transform and a wavelet transform to calculate the time-frequency spectrum characteristics of the signal to diagnose equipment faults [15–17]. The time-frequency spectrum expresses the signal both in the time domain and the frequency domain, and state information within it is more abundant than that in the simple time domain feature and frequency domain feature [18].

Methods combining statistical features and threshold judgment have been widely used from the early days to the present. With the development of artificial intelligence technology, machine learning algorithms such as k-nearest neighbor (KNN), decision tree (DT), support vector machine (SVM), and artificial neural network (ANN) are combined with the statistical characteristics of signals in the time domain or frequency domain and applied in vibration-based fault diagnosis [19–21]. Compared with threshold judgment methods, machine learning methods do not need to conduct in-depth research on the vibration mechanism of equipment faults, nor do they need to manually design the judgment threshold, which greatly reduces the difficulty for engineering applications to enact fault diagnosis methods [22,23]. However, due to the influence of working conditions, different fault samples based on the statistical characteristics under variable working conditions may be confused and difficult to separate. Traditional machine learning methods cannot solve the problem of fault diagnosis under unknown working conditions. In recent years, research on fault diagnosis based on deep learning has made great progress [24–27]. The deep learning method can automatically realize feature extraction and classify samples and can distinguish between the small differences of samples. With sufficient samples, it has higher accuracy than traditional machine learning methods and has the potential to solve the problem of fault diagnosis under unknown working conditions [28–30].

The imbalance of data samples and the fluctuation of operating conditions are the two main challenges for vibration data-driven fault diagnosis of saturable reactors in UHVDC thyristor valves. The challenges should be solved from the perspectives of expanding data samples and exploring new fault features or diagnosis models.

It is impossible to obtain sufficient vibration data samples of saturable reactors with different iron core looseness faults from engineering fields. Research on dataset extension should be carried out to improve the accuracy and generalization capability of the fault diagnosis method and to solve the problem of imbalanced samples. In recent years, models such as the variational autoencoder (VAE) [31] and generative adversarial networks (GANs) [32] and derived methods based on these models, such as the conditional VAE

(CVAE) and Wasserstein GAN, have been widely used in sample expansion in different fields [33–35]. The CVAE-GAN model combines the advantages of VAE and GAN. It can improve the low generated sample quality of the VAE model and perfect the GAN model's mode collapse and sample distortion. The CVAE-GAN model has been applied in image generation and mechanical vibration signal generation [36,37], and its performance is superior to that of pure VAE, GAN, and their derived methods.

The vibration of a saturable reactor is affected by the structural state and electrical excitation at the same time. The DC transmission project may operate under different loads, and the current passing through a saturable reactor is variable rather than fixed. Because there is no bridge current and no reactor voltage acquisition device in its engineering, electrical excitation data are not available in real time, which increases the difficulty of saturable reactor fault diagnosis. Without electrical excitation data, it is almost impossible to diagnose the iron core looseness of a saturable reactor using traditional machine learning methods combined with time-domain or frequency-domain statistical features. Therefore, it is necessary to study more effective features and fault diagnosis models to solve the challenges of fluctuating and unknown working conditions. Sychrosqueezed wavelet transform (SST) is an improved wavelet transform method that can sharpen the time-frequency domain representation results of signals by relocating the wavelet transform results on the time-frequency domain plane [38]. Compared with the general wavelet transform time-frequency spectrum, the energy distribution of the SST time-frequency spectrum is more concentrated, and the key components are more prominent, which is conducive to improving the fault diagnostic performance. Through convolution, pooling, and other technologies, the deep convolution neural network (DCNN) can avoid the exponential increase in the number of neurons caused by the rise in input feature dimensions and hidden layers [39]. It can build a deeper network structure and use fewer hidden layer parameters. It has a more vital ability to learn high-dimensional input features. The DCNN can accept features of sequences or images as input, and some research uses the original vibration signal sequence or time spectrum as the input feature of the DCNN to diagnose faults and achieve good diagnostic performance [28,40,41]. Furthermore, the DCNN model can be designed and modified into a special structure that allows multiple different features to be input at the same time, thus achieving higher diagnostic performance than the method based on a single feature [36].

This paper presents a fault diagnosis method named CVG-MFICNN for the iron core looseness of a UHVDC thyristor valve saturable reactor. A CVAE-GAN model was trained based on an unbalanced dataset to produce generated samples and expand the original dataset to balance. The vibration signals of the expanded dataset samples are processed from different perspectives. The time-frequency spectrum and frequency-domain sequence of the vibration signals are calculated using the SST and Welch method, respectively. The time-domain series and time-domain statistics are also extracted as vibration signal features. A new MFICNN structure is designed that integrates different multimodal features to diagnose the iron core looseness fault of a UHVDC thyristor valve saturable reactor under variable operating conditions.

The remaining part of this paper is organized as follows: The vibration mechanism of the saturable reactor in a UHVDC thyristor valve is introduced in Section 2. The workflow and key model structure of the proposed CVG-MFICNN method are described in detail in Section 3. In Section 4, the saturable reactor vibration experiment is described, and the performance of the proposed method is tested using manually set unbalanced datasets collected from the vibration experiment and compared with other methods. The work of this paper is summarized in Section 5.

2. Vibration Mechanism of the Saturable Reactor

The wiring diagram of a 6-pulse converter consists of six single thyristor valves, and the location of the saturable reactor in a single valve is shown in Figure 1a. Figure 1b shows the structure of a saturable reactor, mainly composed of a coil, iron cores, clamping boards,

and screw bolts. The coil is covered by cast epoxy resin, and the iron cores are mounted outside the epoxy resin. Clamping boards and screw bolts connect the saturable reactor as a whole structure. Figure 1c shows a cross-sectional view of the iron core. A single iron core comprises two U-shaped laminated silicon steel sheets and a tensioning belt. There is an air gap filled with insulating material between the two U-shaped laminated silicon steel sheets.

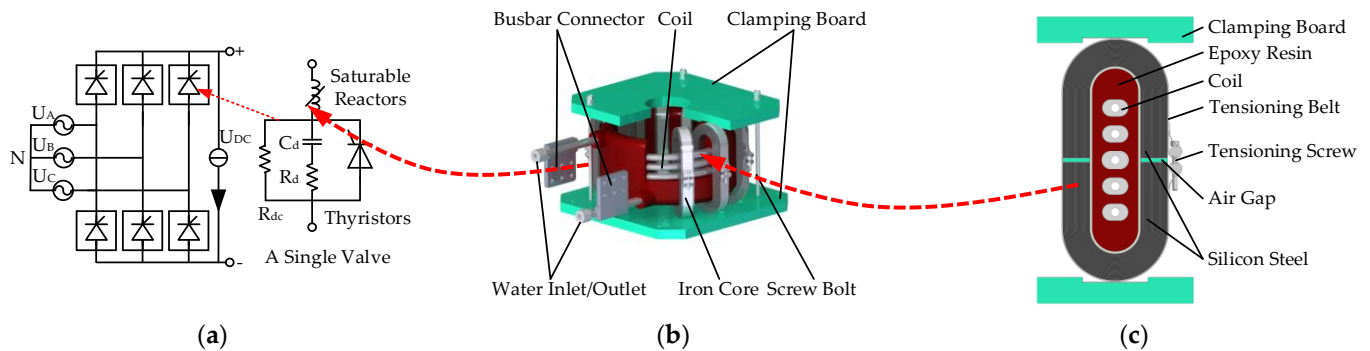


Figure 1. Structure of the UHVDC thyristor valve and saturable reactor: (a) UHVDC thyristor valve appearance; (b) saturable reactor structure; (c) iron core cross-sectional view.

The primary vibration source of a saturable reactor is the magnetostrictive vibration of the iron core [42]. Under the operating condition of the thyristor valve, the periodic trapezoidal wave-like current passing through the saturated reactor can be equivalent to the sum of multiple sinusoidal currents.

$$I = \sum_{k=0,1,2,\dots}^n I_k \sin(k\omega t + \varphi_k) \quad (1)$$

I_k and φ_k are the k -th harmonic's current amplitude and phase angle, respectively. ω is the base angle frequency of the 50 Hz current. According to the principle of electromagnetic induction, the magnetic field strength H in the iron core of a saturable reactor is:

$$H = \frac{NI}{L} = \frac{N}{L} \cdot \sum_{k=0,1,2,\dots}^n I_k \sin(k\omega t + \varphi_k) \quad (2)$$

N and L are the turns of a saturable reactor coil and the equivalent length of the iron core magnetic circuit, respectively. Without considering the magnetic saturation effect, the small relative deformation λ of the silicon steel sheet caused by magnetostriction satisfies Equation (3) [43]:

$$\lambda = \frac{\Delta l}{l} = \frac{2\lambda_s}{H_c^2} \int_0^H |H| dH = \frac{\lambda_s}{H_c^2} H^2 \quad (3)$$

Δl and l are the deformation and the original length of the silicon steel sheet, respectively. λ_s is the saturated magnetostriction rate of the silicon steel sheet. H_c is the coercive force, and H is the magnetic field strength in the iron core. By substituting Equation (2) into Equation (3) and consecutively calculating the differential of the deformation to the time twice, the vibration acceleration of the iron core can be obtained as Equation (4):

$$\begin{aligned}
 a &= \frac{d^2 \Delta l}{dt^2} = \frac{\lambda_s l}{H_c^2} \cdot \frac{d^2 (H^2)}{dt^2} = \frac{\lambda_s l}{H_c^2} \cdot \frac{N^2}{L^2} \cdot \frac{d^2 \left(\left(\sum_{k=0,1,2,\dots}^n I_k \sin(k\omega t + \varphi_k) \right)^2 \right)}{dt^2} \\
 &= \frac{\lambda_s l}{H_c^2} \cdot \frac{N^2}{L^2} \cdot \omega^2 \cdot \left(\begin{aligned} &\sum_{k=0,1,2,\dots}^n 2k^2 I_k^2 \cos(2k\omega t + 2\varphi_k) + \\ &\sum_{k_1 \neq k_2=0,1,2,\dots}^n (k_1 k_2 + k_2^2) I_{k_1} I_{k_2} \cos((k_1 + k_2)\omega t + (\varphi_{k_1} + \varphi_{k_2})) + \\ &\sum_{k_1 \neq k_2=0,1,2,\dots}^n (k_1 k_2 - k_2^2) I_{k_1} I_{k_2} \cos((k_1 - k_2)\omega t + (\varphi_{k_1} - \varphi_{k_2})) \end{aligned} \right) \quad (4)
 \end{aligned}$$

where $\frac{\lambda_s l}{H_c^2} \cdot \frac{N^2}{L^2} \cdot \omega^2$ is a constant, its value is related to the magnetostriction rate λ_s , the original length l of the silicon steel sheet, the coercive force H_c , the coil turns N , and the equivalent length L of the iron core. k stands for the harmonic order. k_1 and k_2 are two different orders of harmonics. It can be seen from Equation (4) that there are both odd and even times of 50 Hz components within the vibration signals of a saturable reactor. Magnetization saturation and magnetostriction of silicon steel sheets are both nonlinear physical processes. This will aggravate the complexity of the iron core vibration signal. The looseness fault of an iron core directly affects its air gap size, structural mode, and, ultimately, vibration characteristics.

3. Proposed CVG-MFICNN Method

Figure 2 is the framework of the CVG-MFICNN method proposed in this paper. First, a 1-D (one-dimensional) CVAE-GAN model is trained based on the vibration data samples of the training set to produce generated samples. The generated and original training samples are combined to form an extended training set. Second, the time-frequency spectrum, time-domain vibration sequence, frequency spectrum sequence, and time-domain statistics of vibration signals are extracted. The extended training feature set, validation feature set and testing feature set are formed based on the above features. Then, the extended training feature set is used to train the MFICNN model, and the validation feature set is used to select the model with the best validation accuracy during the total training process. After training and validation, a model can be obtained without overfitting or underfitting. Finally, the performance of the selected model is tested using the testing feature set to evaluate the effectiveness of the proposed method.

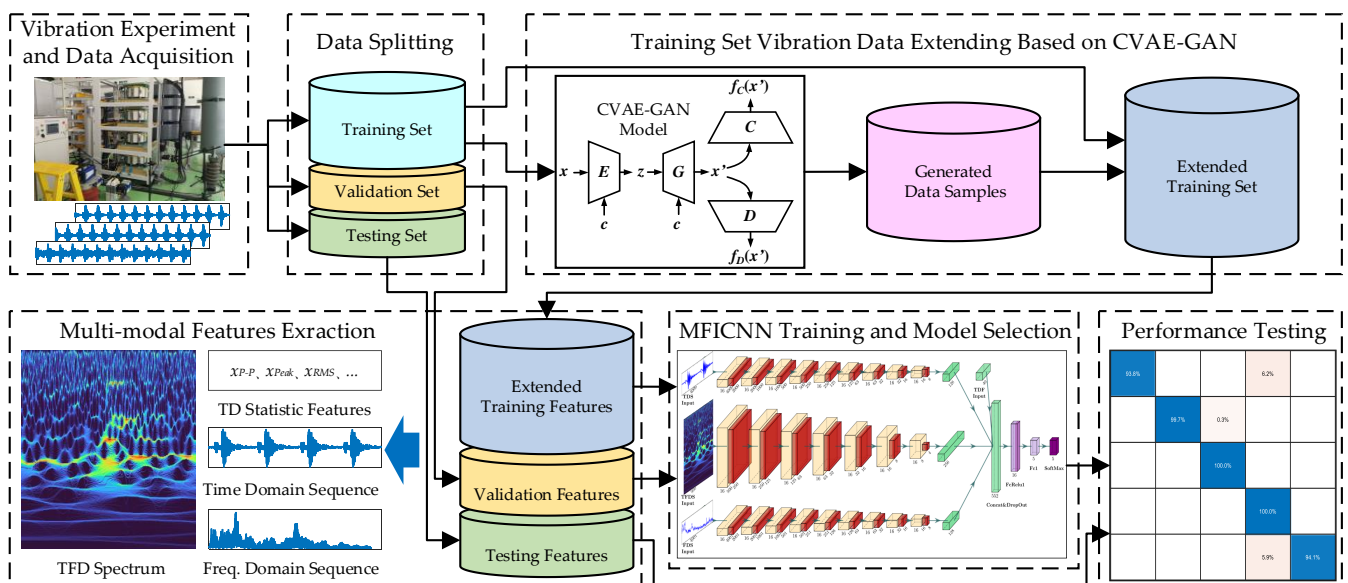


Figure 2. Framework of the proposed CVG-MFICNN diagnosis method.

The processes of training, validation, and testing simulate the actual procedure in engineering applications. The testing corresponds to a practical diagnosis process. Before

testing, only training and validation data are available for model training and selection. There is no data leakage during the whole process.

3.1. 1-D CVAE-GAN Model for Dataset Extending

The method proposed in this paper uses a 1-D CVAE-GAN model to generate vibration time series data to expand the training set. The CVAE-GAN model is composed of CVAE and GAN [37]. As shown in Figure 3, CVAE-GAN includes the following four parts: (1) encoder E; (2) generator G; (3) discriminator D; and (4) classifier C. Encoder E maps real sample x to implicit representation z by learning conditional distribution $P(z|x, c)$ of the real sample, where c is the category of the sample. Generator G learns the distribution of real samples through the gradient provided by discriminator D and samples z to produce generated sample x' through the learned conditional distribution $P(x|z, c)$. Discriminator D learns to distinguish between real and generated samples, and classifier C learns to classify samples.

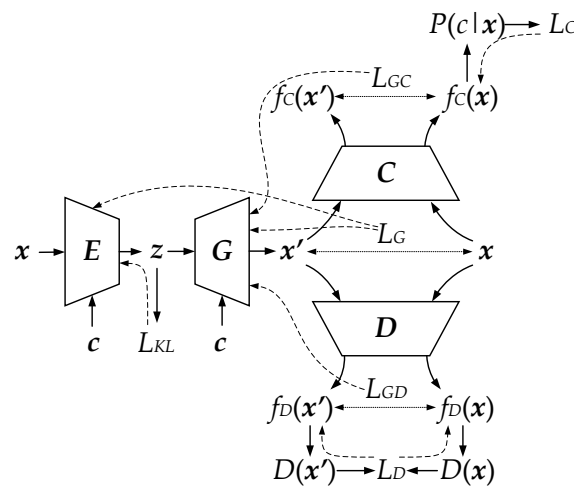


Figure 3. Illustration of a CVAE-GAN model.

The loss functions [37] used for CVAE-GAN training are:

$$\left\{ \begin{array}{l} L_D = -\mathbb{E}_{x \sim P_r} [\log D(x)] - \mathbb{E}_{z \sim P_z} [\log(1 - D(G(z)))] \\ L_C = -\mathbb{E}_{x \sim P_r} [\log P(c|x)] \\ L_G = \frac{1}{2} (\|x - x'\|_2^2 + \|f_D(x) - f_D(x')\|_2^2 + \|f_C(x) - f_C(x')\|_2^2) \\ L_{KL} = \frac{1}{2} (\mu^T \mu + \text{sum}(\exp(\epsilon) - \epsilon - 1)) \\ L_{GD} = \frac{1}{2} \|\mathbb{E}_{x \sim P_r} f_D(x) - \mathbb{E}_{z \sim P_z} f_D(G(z))\|_2^2 \\ L_{GC} = \frac{1}{2} \sum_c \|\mathbb{E}_{x \sim P_r} f_C(x) - \mathbb{E}_{z \sim P_z} f_C(G(z))\|_2^2 \end{array} \right. \quad (5)$$

L_D , L_C , L_G , and L_{KL} are the loss functions of discriminator D, classifier C, generator G, and encoder E, respectively. L_{GD} and L_{GC} are introduced to solve the problem of unstable early training caused by gradient explosion. $D(x)$ is the output of discriminator D. $P(c|x)$ is the classification probability output of classifier C. $f_D(x)$ and $f_C(x)$ are the input vectors of the last fully connected (FC) layer of discriminator D and classifier C, respectively. μ and ϵ are the mean and standard deviation logarithm vectors of hidden layer representation z . L_D and L_C are used to update discriminator D and classifier C, respectively. L_G , L_{GD} , and L_{GC} are used to update generator G. L_{KL} and L_G are used to update encoder E.

A 1-D CVAE-GAN model is designed, and its structure is shown in Figure 4. The 1-D CVAE-GAN model is constructed based on CNN. The encoder, the discriminator, and

the classifier use kernels with a size of 17 and a stride of 5 to construct convolution layers. Through two consecutive convolutions, the size of the feature map input to the final FC layer is reduced to 1/25 of the original vibration data. The decoder applies two consecutive transposed convolution layers constructed using kernels with the same parameters to realize the conversion from code to data. The parameters of each component of the model are shown in Table 1.

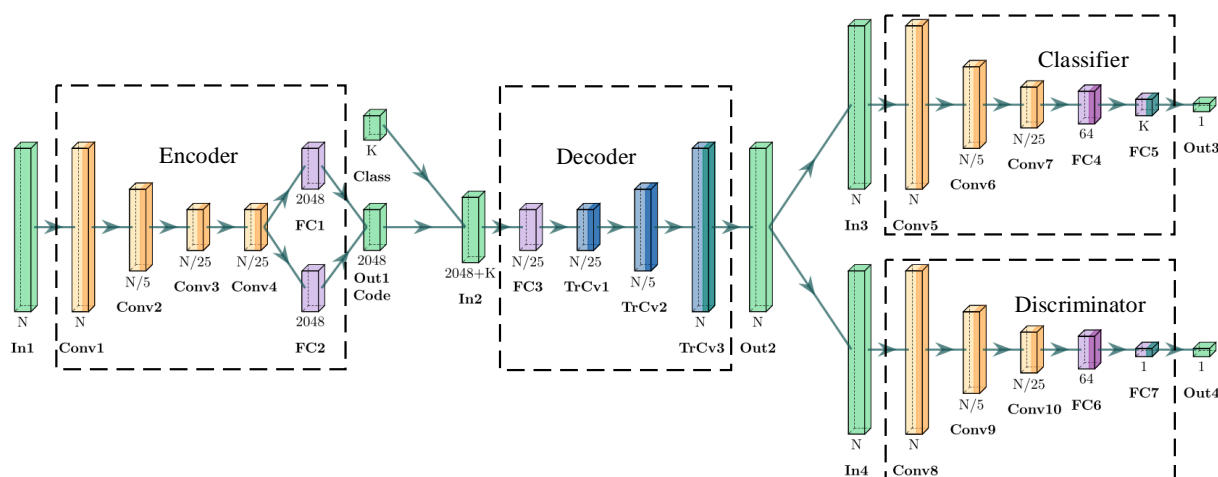


Figure 4. The designed 1-D CVAE-GAN model structure.

Table 1. Parameters of the designed 1-D CVAE-GAN model.

Component	Layer	Filter Size, Stride	Filter/Neuron Number	Output Size
Encoder	Input1	-	-	$N \times 1 \times 1$
	Conv1	5, 1	16	$N \times 1 \times 16$
	Conv2	17, 5	16	$N/5 \times 1 \times 16$
	Conv3	17, 5	32	$N/25 \times 1 \times 32$
	Conv4	3, 1	32	$N/25 \times 1 \times 32$
	Fc1	-	2048	2048
	Fc2	-	2048	2048
Decoder	Input2	-	-	$2048 + K$
	Fc 3	-	$N/25 \times 32$	$N/25 \times 1 \times 32$
	TrCv1	17, 5	16	$N/5 \times 1 \times 16$
	TrCv2	17, 5	16	$N \times 1 \times 16$
	TrCv3	5, 1	1	$N \times 1 \times 1$
Classifier	Input3	-	-	$N \times 1 \times 1$
	Conv5	5, 1	16	$N \times 1 \times 1$
	Conv6	17, 5	16	$N/5 \times 1 \times 16$
	Conv7	17, 5	32	$N/25 \times 1 \times 32$
	Fc4	-	64	64
	Fc5	-	K	K
Discriminator	Input4	-	-	$N \times 1 \times 1$
	Conv8	5, 1	16	$N \times 1 \times 1$
	Conv9	17, 5	16	$N/5 \times 1 \times 16$
	Conv10	17, 5	32	$N/25 \times 1 \times 32$
	Fc6	-	64	64
	Fc7	-	1	1

The encoder consists of four convolution layers and two fully connected layers. A batch normalization layer and a ReLU activation layer are set after every convolution layer. The Fc1 and Fc2 FC layers output the mean and standard deviation vector of the code, respectively. The decoder consists of one FC layer and three transposed convolution layers.

The FC layer is used to input codes and category labels. A ReLU activation layer is set after each of the first two convolution layers, and a sigmoid activation layer is set after the last transposed convolution layer. The discriminator and classifier are composed of three convolution layers and two FC layers. For the discriminator and classifier, all convolution layers and the first FC layer are set with a ReLU activation layer, and the second FC layer is set with a sigmoid layer as the activation function.

Model parameters are randomly initialized before training. For a single training process, the classifier parameters are first updated based on the original samples. Then, the batched original samples are put into the encoder to obtain the mean and standard deviation of the code corresponding to every original sample in the batch. Then, the random codes are generated and input into the decoder together with the category labels to obtain batched generated samples. The generated samples are put into the classifier and discriminator to obtain the classification and discrimination results, respectively. The parameters of the decoder, encoder, and discriminator are updated according to the output of each component. The training of the whole CVAE-GAN model can be performed through multiple iterations of the above processes.

For the sample generation stage, we input the batched original samples and category labels into the encoder and generate random codes according to the mean and standard deviation of the code corresponding to every original sample. The random codes are put into the decoder together with the sample category labels to obtain the batched generated samples to extend the training dataset.

3.2. Features as Input of the MFICNN Model

The time-frequency spectrum, time-domain sequence, frequency-domain sequence, and time-domain statistical code are used as features as the input of the MFICNN model in the proposed method, as shown in Table 2.

Table 2. Features used in the proposed method.

Feature	Method	Size	Remarks
Time-frequency spectrum	SST [38]	$500 \times 500 \times 3$	With a duration of one complete cycle, rescaled and pseudo color processed.
Time-domain sequence	Time-domain signal data segmentation	4000×1	With a duration of two complete cycles.
Frequency-domain sequence	Welch method [44]	4001×1	With a frequency range of 0~50 kHz.
Time-domain statistical code	One-hot method	10×4	The four statistics are the peak value, the peak-peak value, the root mean square value, and the peak factor.

3.3. MFICNN Model

Conventional convolutional neural network models cannot process multiple types of input features, such as images and sequences, simultaneously. In this paper, an MFICNN model structure is designed which can fuse and process multiple types of input features. The model structure is shown in Figure 5. In the figure, TFDS, TDS, FDS, and TDF represent the time-frequency spectrum, time-domain sequence, frequency-domain sequence, and time-domain statistical value code, respectively.

The MFICNN model is constructed using several two-dimensional (2-D) and 1-D convolution pooling layer groups. A convolution pooling layer group is composed of a convolution layer, a batch normalization layer, a maximum pooling layer, and a ReLU activation layer, in order. After the convolution layer, the size of the feature graph remains unchanged. After subsampling by the maximum pool layer, the size of the feature graph is reduced to 1/2 of the original size. The parameters of the above two kinds of convolution and pooling layer groups are shown in Table 3. The data dimension is reduced, and the

key fault characteristics are extracted by passing the original features through multiple consecutively stacked convolution pooling layer groups.

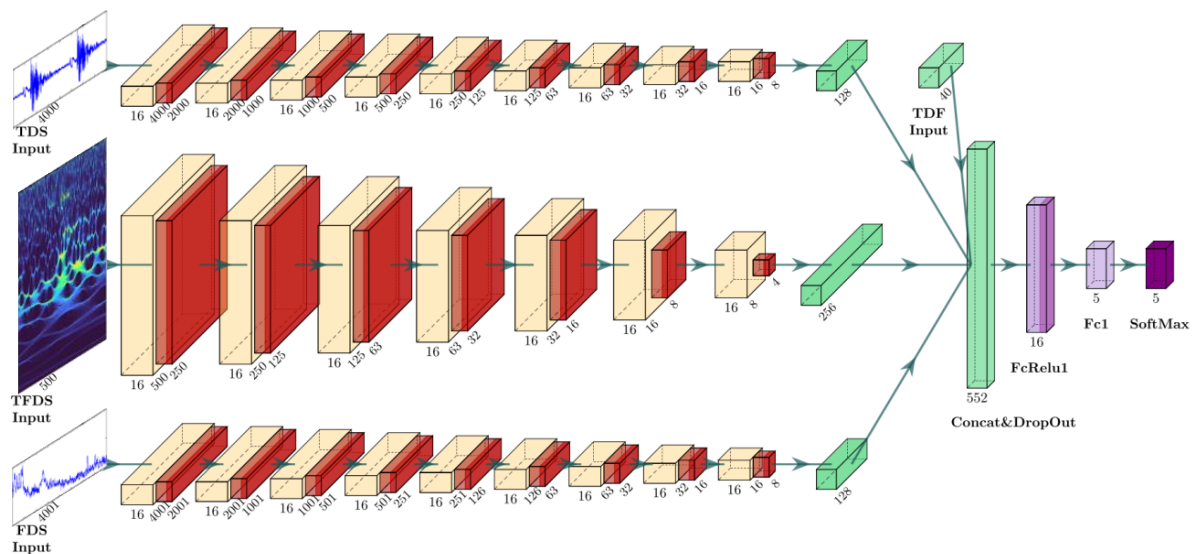


Figure 5. The MFICNN model structure.

Table 3. Parameters of the convolution and pooling layer groups.

Components	Layer	Filter Size, Stride
2-D convolution pooling layer group	Convolution	$5 \times 5, (1, 1)$
	BatchNormalization	-
	MaxPooling	$2 \times 2, (2, 2)$
	ReLU	-
1-D convolution pooling layer group	Convolution	$5, 1$
	BatchNormalization	-
	MaxPooling	$2, 2$
	ReLU	-

The time-frequency spectrum and sequence features are converted into feature graphs with sizes of $4 \times 4 \times 16$ and 8×16 , respectively. The final output feature graphs are flattened into three 1-D sequences and combined with the flattened time-domain statistics codes to form an integrated sequence. The integrated sequence is input to the first FC layer with 16 neurons and a dropout rate of 0.4. The number of neurons in the second FC layer is five, which is consistent with the number of fault types. A softmax layer is set in the end to compute the probability that the samples belong to different categories and output the classification result.

4. Case Study and Discussion

4.1. Saturable Reactor Vibration Experiment

A vibration experiment was conducted to obtain the vibration data of the faulty saturable reactor under different electrical excitations.

Figure 6 shows the vibration experiment wiring diagram. An AC power with a voltage of 220 V is used and connected to a DC power via a transformer. The DC power mainly consists of four sets of IGBTs and is used to charge capacitor C_s . The capacitor C_s is connected to the DC power through switch K_1 , grounded through grounding resistance R_{Cs} and switch K_s . K_s ensures that capacitor C_s is reliably grounded when the experiment is not running. A thyristor valve Thy is set to control the energy supplement for the H-bridge module. Isolation transformers and series reactors are arranged in the H-bridge

module cabinet to supply energy for IGBT drivers. The H-bridge module rectifies the DC voltage for the saturable reactor. The voltage peak value is adjusted by connecting the equivalent stray capacitor to simulate the operating conditions of the saturable reactor. The excitation can simulate the two pulse voltages generated at the moment of opening and closing the UHVDC thyristor valve. The current passing through the saturable reactor is approximately trapezoidal. The voltage and the current waveform are consistent with the actual engineering waveform. Therefore, the operating vibration characteristics of the saturated reactor can be well simulated by conducting this vibration experiment.

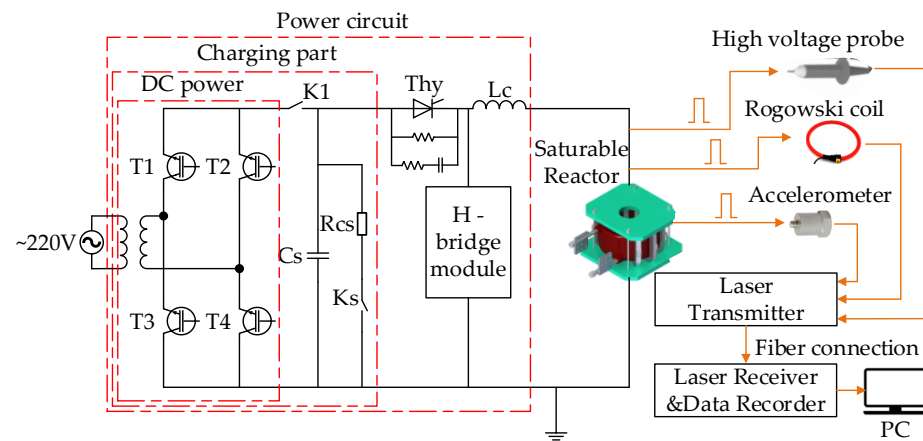


Figure 6. Wiring diagram of the saturable reactor vibration experiment.

A high-voltage probe and a Rogowski coil are used to measure the voltage and current of the saturable reactor. The voltage, current, and vibration data are synchronously collected using a high-voltage isolated data acquisition system. The isolated data acquisition system uses laser fiber communication to ensure reliable insulation between the saturable reactor and the data acquisition device and prevent high voltage from causing safety accidents or damaging the data acquisition device. The high-voltage probe, Rogowski coil, and piezoelectric accelerometer are connected to the laser transmitter of the acquisition system and isolated from each other. The piezoelectric accelerometer and the laser transmitter use batteries for the power supply. The laser transmitter is connected to the receiving device of the data acquisition system through laser fibers, and the laser receiver and data recorder are connected to a PC for data storage.

The experimental field, the appearance of the saturable reactor used in the experiment, and the installation of the accelerometer are shown in Figure 7. To ensure dynamic insulation, the reactor iron core and fixing screw bolts are equipotent to the midpoint of the coil, and all discrete metal parts are wired together. The reactor iron cores are installed on the epoxy resin structure outside the coil and fixed by the tensioning belt. The iron core looseness faults at various levels are simulated by adjusting the torque of the tensioning belt screw bolt. Considering the accuracy of vibration measurement and insulation requirements, we stuck the piezoelectric accelerometer to the surface of the clamping board above the iron core to measure the acceleration along the iron core length direction.

The vibration experiment is conducted under different peak currents to collect vibration data of saturable reactors with iron core looseness faults of various levels. Specific experimental parameters are shown in Table 4. In the table, a, b, c, d, and e represent five levels of iron core looseness faults.

Figure 8 shows the saturable reactor's voltage, current, and vibration data with the iron core in the normal state (NS) and a peak current of 1280 A. Figure 8a represents the time domain waveform of voltage, current, and vibration acceleration of five consecutive cycles, and Figure 8b represents the first valve opening and closing process in Figure 8a. Saturable reactor vibration is caused by current change, so the macro period of vibration is also 50 Hz, which is the same as that of the current. In the process of single valve opening and closing, the vibration excitation of a saturated reactor caused by current change and

magnetostriction is similar to two consecutive shocks with a certain time interval. Due to the higher current change rate di/dt during the closing process, the corresponding shock excitation is stronger, and the resulting vibration response is more intense than the opening process.

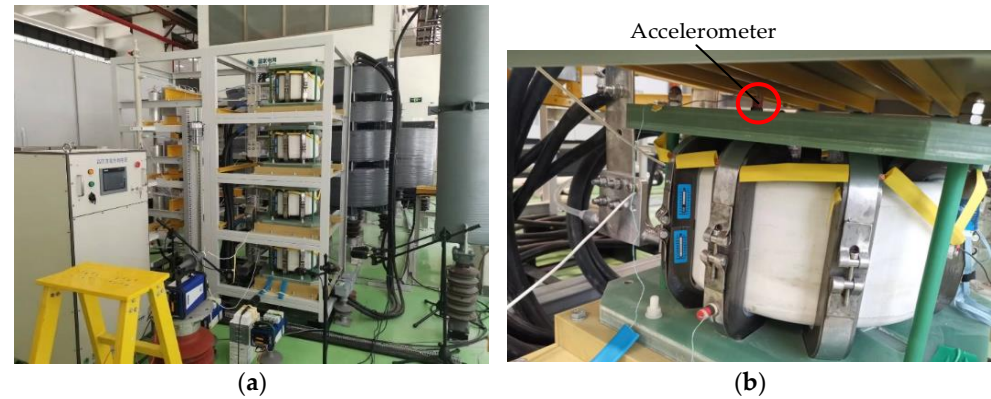


Figure 7. Saturable reactor vibration experiment: (a) experimental field; (b) saturable reactor appearance and installation of the accelerometer.

Table 4. Parameters of the vibration experiment.

Parameter	Fault	Value
Tension belt screw torque	a. Screw off (SO)	Tension belt screw removed
	b. Totally loose (TL)	0 N·m
	c. Extremely loose (EL)	6 N·m
	d. Slightly loose (SL)	8 N·m
	e. Normal state (NS)	10 N·m
Current peak value	All faults	270 A, 335 A, 465 A, 600 A, 730 A, 865 A, 1000 A, 1130 A, 1280 A
Cooling water temperature	All faults	25
Accelerometer response frequency	All faults	3 Hz~30 KHz
Sampling rate	All faults	1 MHz

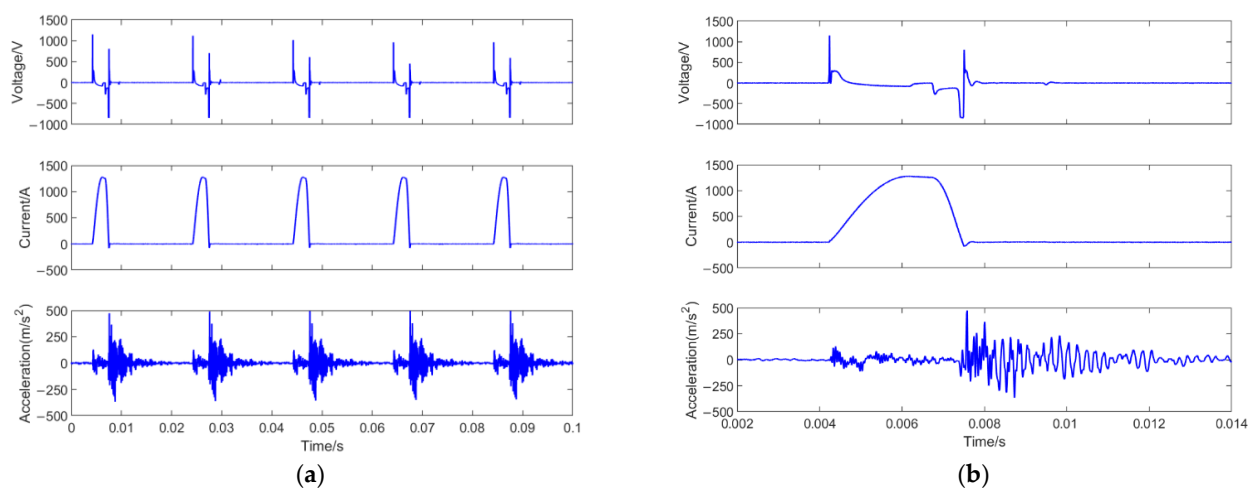


Figure 8. Experimental waveforms of the saturable reactor in the normal state with $I_p = 1280$ A: (a) Time domain waveforms of five consecutive cycles; (b) Time domain waveforms of the single valve opening and closing process.

Figure 9 shows two periodic vibration time domain waveforms of saturated reactors with different loose cores under different operating conditions. The fault state and load affect the vibration signal of the saturable reactor at the same time. It is difficult to distinguish the fault state of the saturable reactor iron core by purely analyzing the characteristics of the vibration signal without load information.

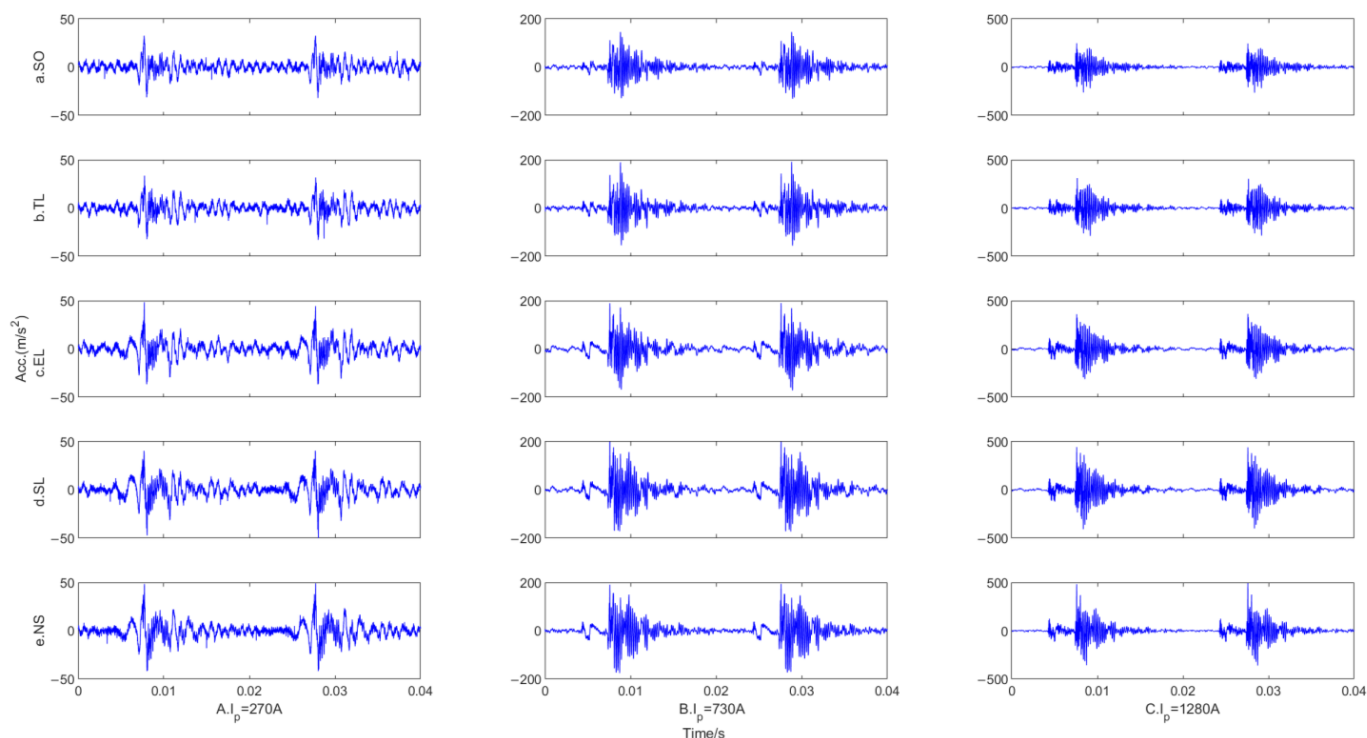


Figure 9. Comparison of the time domain vibration waveform of saturable reactors with different loose iron cores under various operating conditions. From top to bottom, a~e represent the five iron core looseness faults in Table 4. From left to right, A, B, and C represent that the current peak values are 270 A, 730 A, and 1280 A, respectively.

4.2. Dataset Splitting and Extending

The collected vibration data are resampled at 100 kHz and cut into samples with a length of 0.25 s to form the original dataset. The dataset is divided into a training set, a validation set, and a testing set with a ratio of 50%:25%:25%. To create datasets with unbalanced sample distribution, one can randomly remove a certain number of faulty samples of four types except for the normal state type so that the proportion of faulty and normal samples is 2, 4, 8, 15, and 30. The above training sets are defined as training sets I, II, III, IV, and V. The number of different load samples in each dataset is consistent. The specific division of datasets is shown in Table 5.

The CVAE-GAN model was trained separately using each unbalanced training set. The model parameters are set as $N = 25,000$ and $K = 5$ since the length of a single vibration data sample is 25,000. Every unbalanced training set is expanded to balance so that the proportion between normal and faulty samples is 1:1.

Figure 10 shows the power spectra of the original and generated samples in training set V or produced based on training set V. The blue lines represent the real samples, and the red lines represent the generated samples. In general, the power spectrum curve of the generated sample coincides with the real sample, which means that the CVAE-GAN model has learned the key features of the real sample data. Due to the limited number of samples used to train the CVAE-GAN model, there are local differences in certain positions. This can increase the difference among training samples, alleviate the overfitting problem, and improve the model's generalization capability to a certain extent.

Table 5. Datasets organization.

Dataset	Number of Samples Belonging to Different Categories					Unbalance Proportion: Normal:Faulty
	a. Screw off (SO)	b. Totally Loose (TL)	c. Extremely Loose (EL)	d. Slightly Loose (SL)	e. Normal State (NS)	
Training I	300	300	300	300	600	2:1
Training II	150	150	150	150	600	4:1
Training III	75	75	75	75	600	8:1
Training IV	40	40	40	40	600	15:1
Training V	20	20	20	20	600	30:1
Validation	300	300	300	300	300	1:1
Testing	300	300	300	300	300	1:1

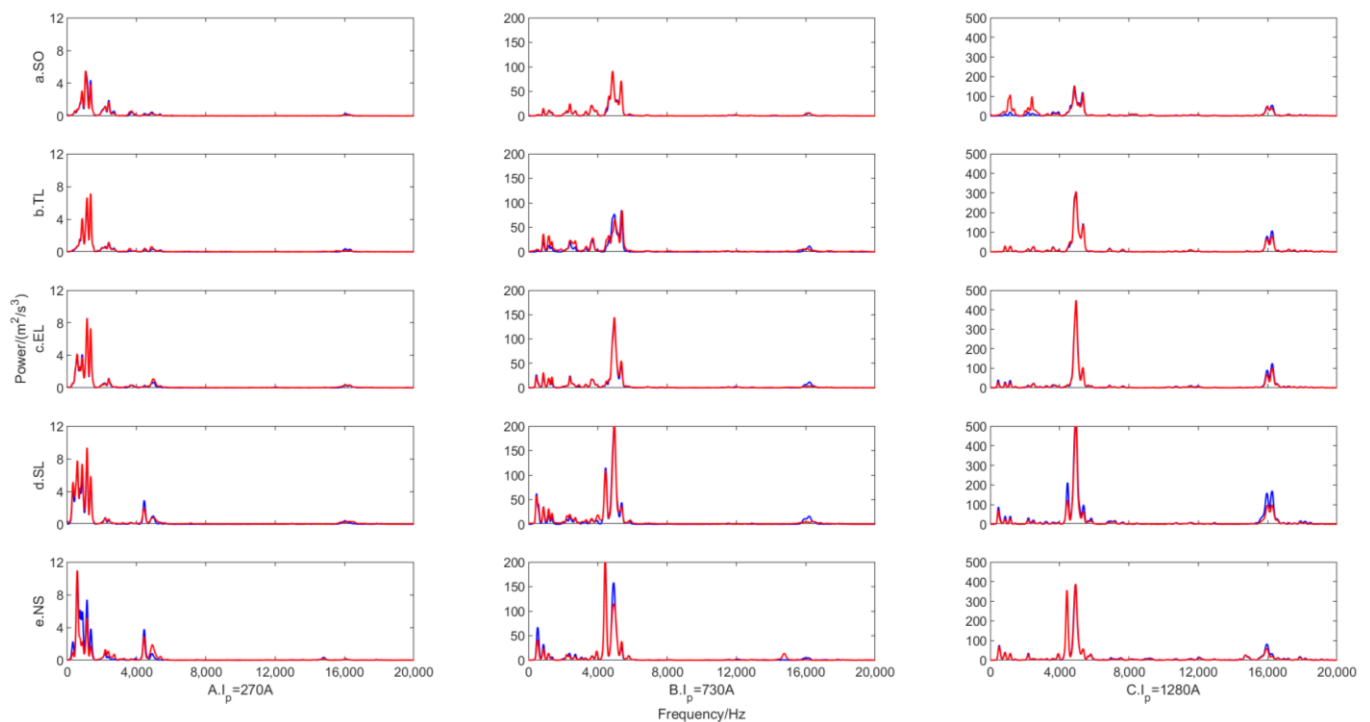


Figure 10. Power spectra of the original and generated samples. The blue line represents the real sample, and the red line represents the generated sample. From top to bottom, a–e represent the five iron core looseness faults in Table 4. From left to right, A, B, and C represent that the current peak values are 270 A, 730 A, and 1280 A, respectively.

4.3. Feature Extraction

4.3.1. Time-frequency Domain Feature Extraction

For a given signal $x(t)$, its continuous wavelet transform $W_s(a, b)$ is defined as:

$$W_s(a, b) = \int x(t) a^{-\frac{1}{2}} \overline{\psi\left(\frac{t-b}{a}\right)} dt \quad (6)$$

ψ is the selected mother wavelet. a and b are the scale parameter and translation parameter, respectively. For any (a, b) meeting $W_s(a, b) \neq 0$, the instantaneous frequency $\omega_s(a, b)$ of signal $x(t)$ is:

$$\omega_s(a, b) = -i(W_s(a, b))^{-1} \frac{\partial}{\partial b} W_s(a, b) \quad (7)$$

$W_s(a, b)$ is converted from the time-scale plane to the time-frequency plane according to the mapping relationship of $(b, a) \rightarrow (b, \omega_s(a, b))$, which is called sychrosqueezing. a and ω are divided into several small boxes so that $a_k - a_{k-1} = (\Delta a)_k$, $\omega_l - \omega_{l-1} = \Delta \omega$. The

sychrosqueezed transform $T_s(\omega_l, b)$ is only determined by $W_s(a, b)$ in consecutive boxes with a center of ω_l and a width of $\Delta\omega$ [38]:

$$T_s(\omega_l, b) = (\Delta\omega)^{-1} \sum_{a_k: |\omega(a, b) - \omega_l| \leq \Delta\omega/2} W_s(a_k, b) a_k^{-3/2} (\Delta a)_k \quad (8)$$

where a_k is the k -th discrete scale and ω_l is the l -th discrete angular frequency. Through the above transformation, the wavelet factor spectrum can be squeezed along the scale axis, and the energy distribution can be more concentrated. Therefore, the SST time-frequency spectrum has higher energy aggregation than the general wavelet factor spectrum.

The time-frequency spectra of the vibration samples are calculated using the SST method. The vibration signals are linearly normalized to the range of $(-0.5, 0.5)$ according to the peak value before the time-frequency spectrum calculation. Considering the periodicity of the vibration signal in the time domain, we only retain the time-frequency spectrum within one single period for a single data sample. The time-frequency spectrum's center of the time dimension is defined based on its peak position in the time domain to ensure phase consistency. Because the original SST time-frequency spectrum can only highlight the main frequency components, the original SST time-frequency spectrum is processed with a logarithm of $T_{st} = \log(T_s + \alpha)$ to better express the nonmajor components that are submerged in the background area, and the parameter α is set as $\alpha = 0.0005$. Every single time-frequency spectrum matrix is linearly normalized according to its maximum and minimum values and then pseudocolor processed using the Turbo color map to improve its distinguishability.

Figure 11 shows the SST time-frequency spectrum features of saturable reactors with different fault states operating under different loads. It can be seen from the figure that both the operating condition and fault state affect the color and shape of the internal vortices in the time-frequency spectrum. With increasing load, the concentration degree of the signal in the frequency domain increases, and the difference between the hot spot area and the background area of the time-frequency spectrum increases as well. However, both the overall brightness of the image and the proportion of red, yellow, and green areas decrease. By comparing the time-frequency spectra of different fault samples under the same operating conditions, we find that the color distribution position and the shape, number, and position of the internal vortices all vary. The time-frequency spectrum in the areas circled by the red solid ovals change significantly. In general, the time-frequency spectra of different faulty samples under different operating conditions are quite different, which can be roughly distinguished by human eyes and used as input of CNN for fault classification.

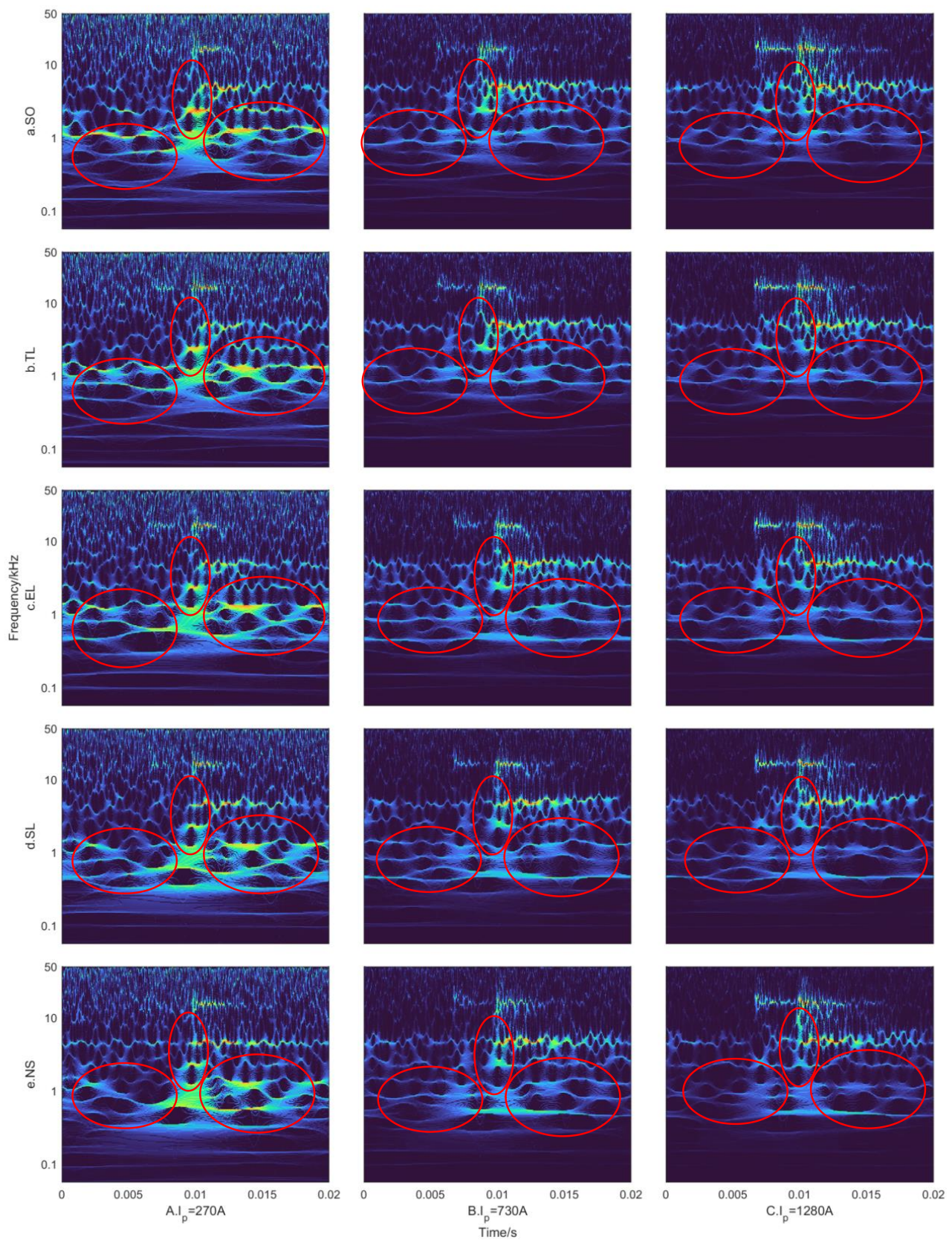


Figure 11. SST time-frequency spectra extracted from samples of different faults and loads. From top to bottom, a~e represent the five iron core looseness faults in Table 4. From left to right, A, B, and C represent that the current peak values are 270 A, 730 A, and 1280 A, respectively.

4.3.2. Time Domain Sequence Feature Extraction

The time-domain vibration waveforms of the samples are shown in Figure 8. The time-domain vibration sequence segment phase is consistent with the SST time-frequency spectrum. A two-periodic time-domain segment with a duration of 0.04 s and a length of 4000 is intercepted for each sample. The vibration time domain sequence is normalized according to Equation (9) so that the original vibration signal with a mean value of 0 is converted into a sequence with a mean value of 0.5 and maximum and minimum values not exceeding the range of (0.05, 0.95).

$$x_{norm}(t) = \frac{0.45 \cdot x(t)}{\max(\max(x(t)), \text{abs}(\min(x(t))))} + 0.5 \quad (9)$$

4.3.3. Frequency Domain Sequence Feature Extraction

A power spectrum sequence with a length of 4001 is calculated using the Welch method for each complete sample with a length of 0.25 s. Figure 10 shows the power spectrum curves of real and generated samples in different fault states under different loads. To improve the power spectrum data distribution and the resolving ability of nonmajor frequency components, we convert the spectrum sequence into a decibel value according to $x_{dB}(f) = 10 \log_{10}(x_{original}(f))$. Before being input into the neural network, the power spectrum sequence should be linearly normalized to the range of (0.05, 0.95) according to the maximum and minimum values according to Equation (10).

$$x_{norm}(f) = \frac{0.9 \cdot (x_{dB}(f) - \min(x_{dB}(f)))}{\max(x_{dB}(f)) - \min(x_{dB}(f))} + 0.05 \quad (10)$$

4.3.4. Time Domain Statistic Code Extraction

Four time-domain statistics, the peak value, peak-peak value, root mean square value, and peak factor, are added as value features to retain the absolute value information of vibration signals, as the above three features are all normalized before being input into the neural network. These features are processed into structured code using the one-hot method. In this case, each feature is converted into a code with a length of 10. We divided the distribution range of features in the dataset into ten sections at equal intervals and set the code value at the same position as the section where the feature is located as 1, and the other 9 code values are set as 0, as shown in Figure 12.

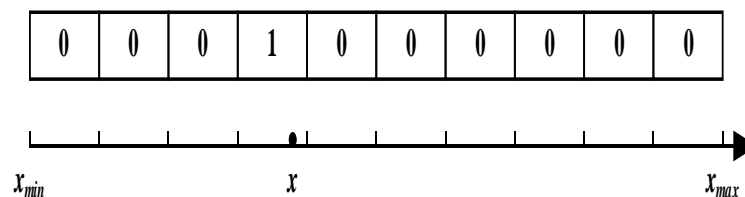


Figure 12. One-hot processing of a statistical feature.

4.4. Model Training and Testing

The F1 score [45] is used to evaluate the performance of the fault diagnosis methods. Equation (11) is the definition of the F1 score, where P is the precision rate, R is the recall rate, TP is the number of samples correctly diagnosed for a specific type of fault, FP is the number of samples wrongly diagnosed as a specific type of fault, and FN is the number of wrongly diagnosed samples for a specific type of fault. The precision ratio P indicates the proportion of a specific type of sample in the diagnosis results that truly belongs to this type of fault. The recall ratio R represents the correctly diagnosed proportion of all samples belonging to a specific type of fault in the set. The F1 score combines the P and R values and more comprehensively considers the method's performance. Diagnosing the iron core looseness fault of the UHVDC thyristor saturable reactor is a multiclass classification

problem. For models trained based on a certain method, this paper calculates the F1 score of all fault classifications and uses the minimum and average values of all fault classifications F1 scores to characterize the method performance.

$$\begin{cases} P = \frac{TP}{TP + FP} \\ R = \frac{TP}{TP + FN} \\ F1 = 2 \cdot \frac{P \cdot R}{P + R} \end{cases} \quad (11)$$

The MFICNN model is trained using the imbalanced training set I~V and the corresponding CVAE-GAN extended training sets. The Adam optimizer is used for model training. The learning rate is set to 0.0005, the maximum number of iterations is set to 1000, the number of small batch samples is set to 64, and the single training model selection strategy is set as selecting the model with the best validation loss. Figure 13 shows the model training process using the CVG-MFICNN method based on training set I. The solid blue line represents the training accuracy, the dotted blue line represents the validation accuracy, the solid red line represents the training loss, and the dotted red line represents the validation loss. In the early stage, the model's accuracy continues to rise with the increase in training iterations. At 248 iterations, the training accuracy reached 100% and remained almost unchanged. The validation accuracy exceeded 95% but still fluctuated with the increase in iterations. The changing trend of the loss and the accuracy rate is the opposite. After 694 iterations, the accuracy of model validation reached the maximum value of 99.47%, and the validation loss reached the minimum value. After that, the accuracy and loss of model validation continued to fluctuate, but it could not obtain better validation performance than the model of iteration 694, which indicates that the model was overfitted after 694 iterations. Therefore, the model corresponding to iteration 694 is selected as the final model obtained from this training process and then used for performance testing.

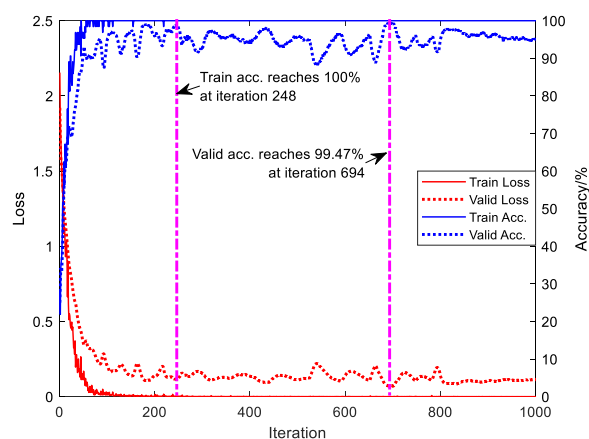


Figure 13. Model training process using the CVG-MFICNN method based on training set I.

To illustrate the advantage of MFICNN over the single-feature neural network, we trained and tested single-feature neural network models of the time-frequency spectrum, time-domain sequence, and frequency-domain sequence for all training sets. 1-D CNN models and 2-D CNN models are separately trained based on the sequence features and the time-frequency spectrum feature. The above neural network model parameters are consistent with the parameters in the MFICNN model proposed in this paper, and the training parameters and model selection strategy are the same as those of the MFICNN model training process.

The testing results of models trained using different methods and based on different training sets are shown in Table 6. In the table, I~V represent five training sets with different

unbalanced proportions. TDS-1DCNN, FDS-1DCNN, and TFD-CNN represent three single-feature CNN methods. The three features are the time-domain sequence feature, frequency-domain sequence feature, and time-frequency spectrum feature. MFICNN represents the method using the MFICNN model designed in this paper. CVG represents the combination of the above methods with CVAE-GAN data extension. To exclude the influence of the model training randomness on the testing results, 30 models were trained for each method. The average values of the minimum F1 scores and average F1 scores of the 30 models for each method were used for performance comparison. A higher F1 score indicates better method performance. The results in Table 6 are also illustrated in Figures 14 and 15.

Table 6. Testing results of models trained by different methods and based on different training sets.

Method	Min. F1 Score					Ave. F1 Score				
	I	II	III	IV	V	I	II	III	IV	V
TDS-1DCNN	0.586	0.514	0.495	0.477	0.421	0.741	0.685	0.631	0.626	0.583
CVG-TDS-1DCNN	0.604	0.586	0.563	0.554	0.444	0.767	0.723	0.700	0.688	0.619
FDS-1DCNN	0.838	0.781	0.762	0.702	0.598	0.924	0.889	0.878	0.828	0.768
CVG-FDS-1DCNN	0.856	0.840	0.819	0.797	0.680	0.924	0.925	0.903	0.902	0.823
TFDS-CNN	0.865	0.839	0.792	0.766	0.653	0.927	0.920	0.894	0.863	0.801
CVG-TFDS-CNN	0.911	0.905	0.890	0.905	0.849	0.955	0.954	0.947	0.951	0.916
MFICNN	0.902	0.871	0.828	0.791	0.701	0.949	0.932	0.907	0.872	0.812
CVG-MFICNN	0.965	0.947	0.921	0.918	0.863	0.983	0.973	0.959	0.958	0.927

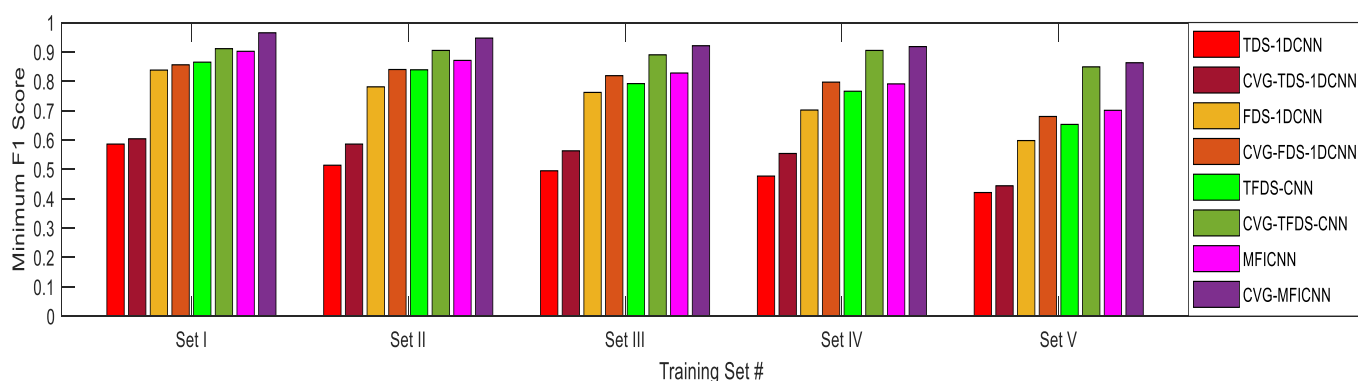


Figure 14. Comparison of minimum testing F1 scores of models trained using different methods and based on different training sets.

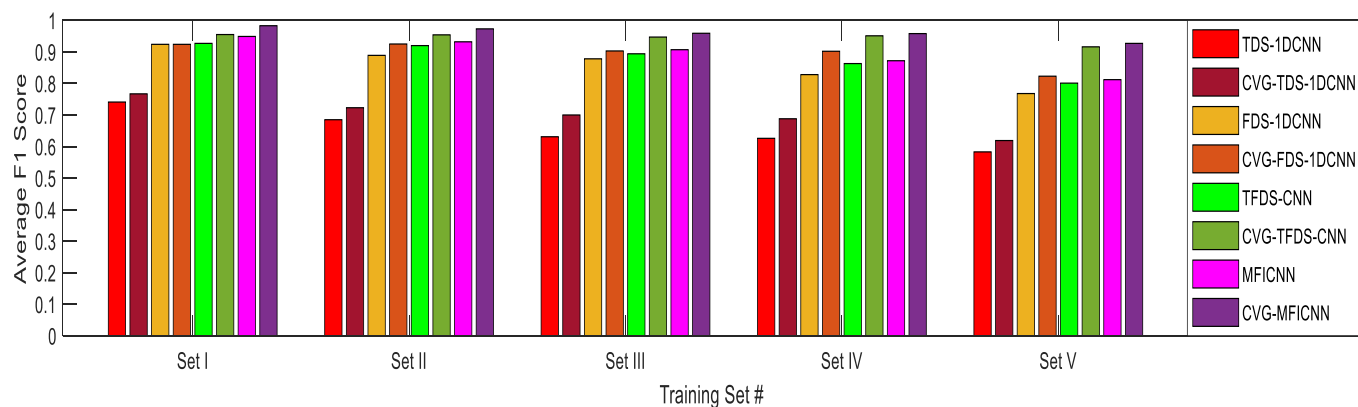


Figure 15. Comparison of average testing F1 scores of models trained using different methods and based on different training sets.

Table 6 shows that CVAE-GAN data extension can improve the performance of models trained based on all original unbalanced training sets for all methods. Especially for the

III, IV, and V training sets with an imbalance ratio of greater than 8, the performance improvement is more significant.

Compared to the other two single-feature methods, the TDS-1DCNN method has the worst performance. The performance of the FDS-1DCNN method is better than that of the TDS-1DCNN method. Combined with CVAE-GAN data extension, the average F1 score of the CVG-FDS-1DCNN method on datasets I~IV exceeds 0.900, which is significantly better than that of the TDS-1DCNN and CVG-TDS-1DCNN methods. The TFDS-CNN method has the best performance. Combined with CVAE-GAN data extension, the average F1 score of the CVG-TFDS-CNN method on all datasets exceeds 0.900, especially on datasets I~IV, and the average F1 score is close to or exceeds 0.950.

The performance of the MFICNN method is better than that of any method based on a single feature and CNN. The F1 scores of models trained based on the MFICNN and CVG-MFICNN methods are higher than those trained based on the TFDS-CNN and CVG-TFDS-CNN methods, respectively. The average F1 score of models trained using the CVG-MFICNN method on training set I reaches 0.983. For the extremely imbalanced training set V with an imbalance ratio of 30, the average F1 score of models trained based on the CVG-MFICNN method still reaches 0.927.

To represent the diagnosis result of the CVG-MFICNN method more clearly, we take training sets I and V as examples to compare the testing confusion charts of the TFDS-CNN, CVG-TFDS-CNN, MFICNN, and CVG-MFICNN methods, as shown in Figure 16. The vertical axis of a confusion chart represents the real fault categories, and the horizontal axis represents the diagnosed fault categories. The five fault categories are defined in Table 4. Paired comparisons of A vs. B, C vs. D, E vs. F, and G vs. H show that combining CVAE-GAN data extension can improve the performance of the models trained using unbalanced training sets. We find that the performance of models trained using the most imbalanced training set V is improved significantly from paired comparisons of C vs. D and G vs. H. In addition, by comparing A vs. E, B vs. F, C vs. G, and D vs. H, it can be found that the number of samples wrongly classified by the MFICNN method for a single fault category is less than that of the TFDS-CNN method, so the diagnostic performance of the MFICNN method is better.

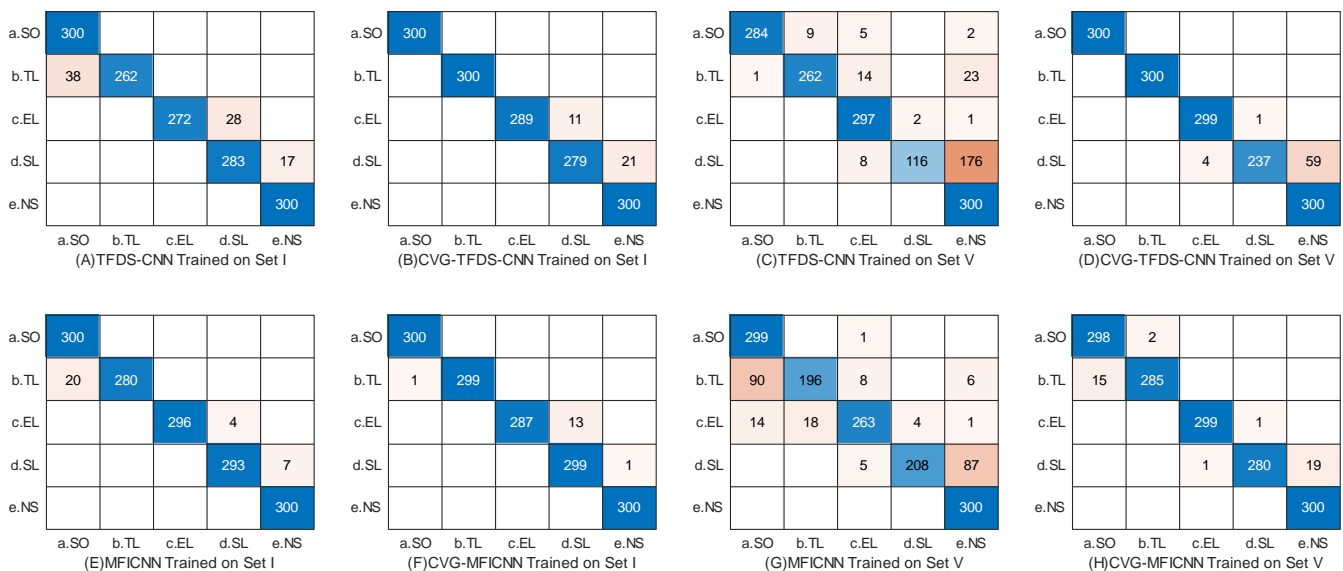


Figure 16. Testing confusion charts of models trained using the TFDS-CNN, CVG-TFDS-CNN, MFICNN, and CVG-MFICNN methods based on training sets I and V.

In summary, the CVG-MFICNN method proposed in this paper can effectively improve the diagnostic performance of the models trained based on unbalanced training sets by extending the datasets using the CVAE-GAN method. The performance of the

severely unbalanced training sets III, IV, and V is improved rather significantly. Using multimodal features can achieve a higher minimum and average F1 score than other methods based on single features such as the time-frequency spectrum, time-domain sequence, and frequency-domain sequence.

4.5. Comparison with Other Methods

To further illustrate the performance advantages of the CVG-MFICNN method, traditional machine learning models such as KNN, DT, SVM, fully connected neural network (FCNN), and classical convolutional neural network models such as LeNet-5 [39], AlexNet [46], VGGNet-16 [47], ResNet-50 [48] were trained, tested, and compared in this chapter.

The time domain and frequency domain statistics [49] are taken as the input features for the traditional machine learning models. Time domain statistics include the peak-to-peak value, the average amplitude, the root mean square value, the kurtosis factor, the skewness factor, and the peak factor. Frequency domain statistics include the average frequency domain amplitude, the gravity frequency, the standard deviation of frequency, the skewness factor of frequency amplitude, and the kurtosis factor of frequency domain amplitude. The SST time-frequency spectrum is taken as the feature for the classical convolutional neural network models. To adapt to the classical network structure, the SST time-frequency spectrum image is scaled before being input into the neural networks. The neuron number in the last fully connected layer of all CNN models is changed to five, which is the number of iron core looseness fault categories of the saturable reactor in the UHVDC thyristor valve.

The performance comparison result is shown in Table 7. The TFSF represents the time domain and frequency domain statistics, and the TFDS stands for the SST time-frequency spectrum. Similar to the comparison in Table 6, 30 models were trained for each method, and the average value of the minimum F1 scores and average F1 scores of the 30 models were used for method performance comparison. The computational costs and parameter numbers of different method models are also illustrated in the table. The computational cost is obtained by training models of different methods via the same platform with a CPU of i5-12600KF and a GPU of Nvidia RTX 3090. The computational cost is defined as the time of the total model training process with 1000 iterations for other methods, and the CVG-MFICNN method needs an additional 60 h to extend the training set. The results in Table 7 are also illustrated in Figures 17 and 18.

Table 7. Performance comparison of traditional ML and classical CNN methods.

Method	Computational Cost	Para. Number	Min. F1 Score					Ave. F1 Score				
			I	II	III	IV	V	I	II	III	IV	V
TFSF-KNN	-	-	0.472	0.323	0.328	0.338	0.325	0.648	0.508	0.477	0.456	0.520
TFSF-DT	-	-	0.333	0.316	0.369	0.370	0.010	0.571	0.481	0.528	0.472	0.347
TFSF-SVM	-	-	0.634	0.541	0.336	0.489	0.499	0.702	0.623	0.548	0.599	0.579
TFSF-FCNN	-	-	0.392	0.357	0.366	0.351	0.276	0.550	0.529	0.523	0.524	0.452
TFDS-LeNet-5	16 min 43 s	61.8 k	0.675	0.837	0.763	0.786	0.792	0.837	0.893	0.859	0.858	0.832
TFDS-AlexNet	17 min 40s	56.8 M	0.672	0.807	0.757	0.786	0.743	0.853	0.915	0.902	0.891	0.863
TFDS-VGGNet-16	25 min 8 s	134.2 M	0.530	0.620	0.578	0.682	0.495	0.702	0.788	0.754	0.795	0.734
TFDS-ResNet-50	43 min 0s	23.5 M	0.728	0.710	0.591	0.627	0.467	0.857	0.843	0.777	0.785	0.707
CVG-MFICNN	60 h 20 min 28 s	70.3 k	0.965	0.947	0.921	0.918	0.863	0.983	0.973	0.959	0.958	0.927

The minimum and average F1 scores of all traditional machine learning methods are low. The performance of the SVM is the best, but its average F1 score is only 0.5~0.7, and the minimum F1 score is 0.3~0.6. The performance of KNN, decision tree, and FCNN is inferior to that of SVM, the minimum F1 score is in the range of 0.3~0.4, and the average F1 score is approximately 0.5.

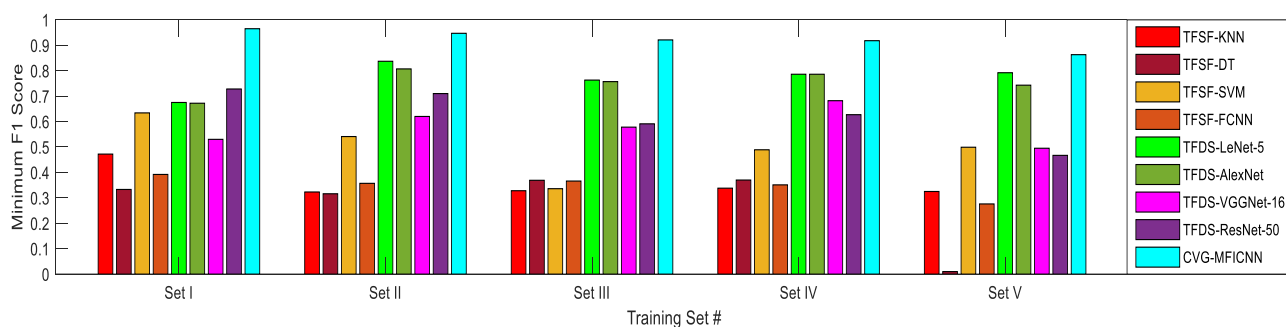


Figure 17. Comparison of minimum testing F1 scores of models trained using the proposed CVG-MFICNN method and existing methods based on different training sets.

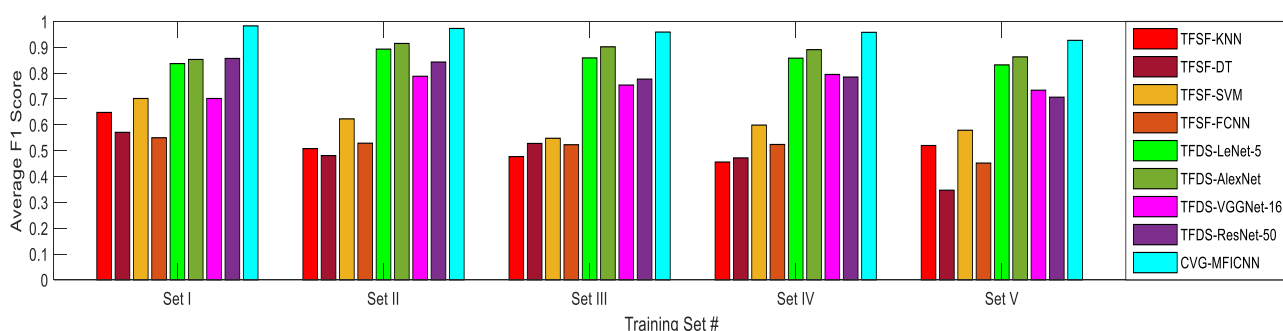


Figure 18. Comparison of average testing F1 scores of models trained using the proposed CVG-MFICNN method and existing methods based on different training sets.

Among the classical CNN methods, AlexNet has the best performance. The average F1 scores of AlexNet models trained on all imbalanced training sets exceed 0.850, and the average F1 scores of models trained on training sets II and III exceed 0.900. The performance of LeNet-5, which has the simplest structure, was the second best. The average F1 scores of the LeNet-5 models trained on training sets I, II, and III exceed 0.870, and the average F1 scores of the LeNet-5 models trained on training sets IV and V exceed 0.810. The performance of ResNet-50 and VGGNet-16 is poor. Only the average F1 scores of ResNet-50 models trained on training sets I and II exceed 0.800. The average F1 scores of ResNet-50 models trained on other training sets and VGGNet-16 models trained on all training sets do not exceed 0.800.

Through comparison, it can be found that the performance of the CVG-MFICNN method proposed in this paper is far superior to the traditional machine learning methods and the classical CNN methods. The model training computational cost of the CVG-MFICNN method is similar to that of the LeNet-5, AlexNet, and VGGNet-16 but over 50% less than that of the ResNet-50. The CVG-MFICNN method costs additional computational resources of approximately 60 h for training set expansion during the model preparation stage, but this is acceptable because the model preparation is conducted only once before the model deployment. On the other hand, the models trained using the CVG-MFICNN method have significantly fewer parameters than most classical CNN models, which makes the models easier to deploy in the actual production environment.

5. Conclusions

This paper proposes a CVG-MFICNN fault diagnosis method for the iron core looseness of saturable reactors in UHVDC thyristor valves. The method can achieve high-accuracy saturable reactor iron core looseness fault diagnosis without operating condition information under variable line loads based on training sets with various imbalance proportions. Its performance is significantly superior to traditional machine learning and classical CNN methods. The work of this paper is summarized as follows:

1. A saturable reactor vibration experiment was conducted to obtain the vibration data. Saturable reactors with five types of iron core looseness faults operated under various loads using the comprehensive experimental platform, and the vibration acceleration, voltage, and current data were synchronously acquired. The vibration signals are similar to the response of periodic continuous two-time shock, and the frequency distribution range is approximately 0~20 kHz.
2. To solve the problem of imbalanced data in engineering, a 1-D CVAE-GAN model is designed and used to extend the original training sets with various imbalance proportions. The testing results show that the performance of the models trained on imbalanced training sets can be improved significantly by extending the training sets using the designed 1-D CVAE-GAN model, especially for severely unbalanced training sets.
3. The time-frequency spectrum features of vibration signals are extracted based on the SST method, and the identifiability of the time-frequency spectrum is improved by logarithmic calculation and pseudocolor processing. The testing results show that the method's performance using the SST time-frequency spectrum feature is superior to that of other methods using the time-domain sequence feature or frequency-domain sequence feature.
4. An MFICNN model is designed, which allows simultaneous input of the time-frequency spectrum feature, the time-domain sequence feature, the frequency-domain sequence feature, and the time-domain statistical code feature. Its diagnostic performance is superior to the method using any single type of feature. Combined with the CVAE-GAN data extension model, a CVG-MFICNN method for iron core looseness fault diagnosis of a saturable reactor in a UHVDC thyristor valve is proposed. Its performance is significantly better than traditional machine learning and classical CNN methods. The problems of imbalanced data and variable unknown operating conditions in saturable reactor iron core looseness fault diagnosis can be solved using the proposed method.

Author Contributions: Conceptualization, X.Z., X.W. and H.W.; methodology, X.Z., L.Z. and X.W.; software, X.Z.; validation, X.Z., L.Z. and C.W.; formal analysis, X.Z.; investigation, X.Z.; resources, X.W., C.W. and H.W.; data curation, X.Z.; writing—original draft preparation, X.Z.; writing—review and editing, X.W., L.Z. and H.W.; visualization, X.Z.; supervision, H.W.; project administration, X.W.; funding acquisition, X.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Beijing Natural Science Foundation (Grant No. 3212039).

Data Availability Statement: Not applicable.

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

References

1. Tang, G.; He, Z.; Pang, H. R&D and application of voltage sourced converter based high voltage direct current engineering technology in China. *J. Mod. Power Syst. Clean Energy* **2014**, *2*, 1–15.
2. Kunpeng, Z.; Junzheng, C.; Wenmin, O.; Baokui, S.; Chong, G.; Hongzhou, L. *Design of 6250A/±800kV UHVDC Converter Valve*; IET: Hong Kong, 2017; pp. 1–6.
3. Borucki, S. Diagnosis of technical condition of power transformers based on the analysis of vibroacoustic signals measured in transient operating conditions. *IEEE Trans. Power Deliv.* **2012**, *27*, 670–676. [[CrossRef](#)]
4. Saponara, S.; Fanucci, L.; Bernardo, F.; Falciani, A. Predictive diagnosis of high-power transformer faults by networking vibration measuring nodes with integrated signal processing. *IEEE Trans. Instrum. Meas.* **2016**, *65*, 1749–1760. [[CrossRef](#)]
5. García, B.; Burgos, J.C.; Alonso, Á. Winding deformations detection in power transformers by tank vibrations monitoring. *Electr. Power Syst. Res.* **2005**, *74*, 129–138. [[CrossRef](#)]
6. Hou, P.; Ma, H.; Ju, P. Intelligent Diagnosis Method for Mechanical Faults of High-Voltage Shunt Reactors Based on Vibration Measurements. *Machines* **2022**, *10*, 627. [[CrossRef](#)]
7. Caesarendra, W.; Kosasih, B.; Tieu, A.K.; Moodie, C.A. Circular domain features based condition monitoring for low speed slewing bearing. *Mech. Syst. Signal Pr.* **2014**, *45*, 114–138. [[CrossRef](#)]

8. Randall, R.B. *Vibration-Based Condition Monitoring: Industrial, Automotive and Aerospace Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2021.
9. Hong, L.; Dhupia, J.S. A time domain approach to diagnose gearbox fault based on measured vibration signals. *J. Sound Vib.* **2014**, *333*, 2164–2180. [[CrossRef](#)]
10. Hong, K.; Huang, H.; Fu, Y.; Zhou, J. A vibration measurement system for health monitoring of power transformers. *Measurement* **2016**, *93*, 135–147. [[CrossRef](#)]
11. Hu, Y.; Huang, H.; Zheng, J.; Li, L. Power Transformer Fault Diagnosis Based on Controlled Switching Vibration Analysis. In Proceedings of the 2019 14th IEEE Conference on Industrial Electronics and Applications, Xi'an, China, 19–21 June 2019; pp. 678–683.
12. Rahman, M.M.; Uddin, M.N. Online unbalanced rotor fault detection of an IM drive based on both time and frequency domain analyses. *IEEE Trans. Ind. Appl.* **2017**, *53*, 4087–4096. [[CrossRef](#)]
13. Ribeiro Junior, R.F.; Areias, I.A.D.S.; Gomes, G.F. Fault detection and diagnosis using vibration signal analysis in frequency domain for electric motors considering different real fault types. *Sens. Rev.* **2021**, *41*, 311–319. [[CrossRef](#)]
14. Yuen, C.K. On the smoothed periodogram method for spectrum estimation. *Signal Process.* **1979**, *1*, 83–86. [[CrossRef](#)]
15. Zhang, S.; Ji, Y.; Jiang, X. Research on Method of Mechanical State Characteristics Diagnosis Based on STFT and RVM for Transformer Winding. In Proceedings of the 2020 IEEE 3rd International Conference on Electronics Technology (ICET), Chengdu, China, 8–12 May 2020; pp. 271–277.
16. Duan, R.; Wang, F. Fault diagnosis of on-load tap-changer in converter transformer based on time–frequency vibration analysis. *IEEE Trans. Ind. Electron.* **2016**, *63*, 3815–3823. [[CrossRef](#)]
17. Huerta-Rosales, J.R.; Granados-Lieberman, D.; Garcia-Perez, A.; Camarena-Martinez, D.; Amezcua-Sanchez, J.P.; Valtierra-Rodriguez, M. Short-circuited turn fault diagnosis in transformers by using vibration signals, statistical time features, and support vector machines on FPGA. *Sensors* **2021**, *21*, 3598. [[CrossRef](#)]
18. Lilly, J.M.; Olhede, S.C. Generalized Morse Wavelets as a Superfamily of Analytic Wavelets. *IEEE Trans. Signal Process.* **2012**, *60*, 6036–6041. [[CrossRef](#)]
19. Valtierra-Rodriguez, M. Fractal dimension and data mining for detection of short-circuited turns in transformers from vibration signals. *Meas. Sci. Technol.* **2019**, *31*, 025902. [[CrossRef](#)]
20. Sun, W.; Chen, J.; Li, J. Decision tree and PCA-based fault diagnosis of rotating machinery. *Mech. Syst. Signal Process.* **2007**, *21*, 1300–1317. [[CrossRef](#)]
21. Shantisagar, K.; Jegadeeshwaran, R.; Sakthivel, G.; Manghai, T.A. Vibration Based Tool Insert Health Monitoring Using Decision Tree and Fuzzy Logic. *Struct. Durab. Health Monit.* **2019**, *13*, 303. [[CrossRef](#)]
22. Awadallah, M.A.; Morcos, M.M. Application of AI tools in fault diagnosis of electrical machines and drives—an overview. *IEEE Trans. Energy Convers.* **2003**, *18*, 245–251. [[CrossRef](#)]
23. Tiboni, M.; Remino, C.; Bussola, R.; Amici, C. A Review on Vibration-Based Condition Monitoring of Rotating Machinery. *Appl. Sci.* **2022**, *12*, 972. [[CrossRef](#)]
24. Do, T.; Tuyet-Doan, V.; Cho, Y.; Sun, J.; Kim, Y. Convolutional-neural-network-based partial discharge diagnosis for power transformer using UHF sensor. *IEEE Access* **2020**, *8*, 207377–207388. [[CrossRef](#)]
25. Jiang, A.; Yan, N.; Wang, F.; Huang, H.; Zhu, H.; Wei, B. Visible Image Recognition of Power Transformer Equipment Based on Mask r-cnn. In Proceedings of the 2019 IEEE Sustainable Power and Energy Conference (iSPEC), Beijing, China, 21–23 November 2019; pp. 657–661.
26. Akpudo, U.E.; Hur, J. D-dCNN: A novel hybrid deep learning-based tool for vibration-based diagnostics. *Energies* **2021**, *14*, 5286. [[CrossRef](#)]
27. Jiang, X.; Yang, S.; Wang, F.; Xu, S.; Wang, X.; Cheng, X. OrbitNet: A new CNN model for automatic fault diagnostics of turbomachines. *Appl. Soft Comput.* **2021**, *110*, 107702. [[CrossRef](#)]
28. Mitiche, I.; Nesbitt, A.; Conner, S.; Boreham, P.; Morison, G. 1D-CNN based real-time fault detection system for power asset diagnostics. *IET Gener. Transm. Distrib.* **2020**, *14*, 5766–5773. [[CrossRef](#)]
29. Abdeljaber, O.; Avci, O.; Kiranyaz, S.; Gabbouj, M.; Inman, D.J. Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks. *J. Sound Vib.* **2017**, *388*, 154–170. [[CrossRef](#)]
30. Hasegawa, T.; Saeki, M.; Ogawa, T.; Nakano, T. Vibration-based fault detection for flywheel condition monitoring. *Procedia Struct. Integr.* **2019**, *17*, 487–494. [[CrossRef](#)]
31. Kingma, D.P.; Welling, M. Auto-encoding variational bayes. *arXiv* **2013**, arXiv:1312.6114.
32. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. *Commun. ACM.* **2020**, *63*, 139–144. [[CrossRef](#)]
33. Dixit, S.; Verma, N.K.; Ghosh, A.K. Intelligent fault diagnosis of rotary machines: Conditional auxiliary classifier GAN coupled with meta learning using limited data. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 3517811. [[CrossRef](#)]
34. Vahdat, A.; Kautz, J. NVAE: A deep hierarchical variational autoencoder. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 19667–19679.
35. An, J.; Cho, S. Variational autoencoder based anomaly detection using reconstruction probability. *Spec. Lect. IE* **2015**, *2*, 1–18.
36. Zhang, L.; Zhang, H.; Cai, G. The Multi-class Fault Diagnosis of Wind Turbine Bearing Based on Multi-source Signal Fusion and Deep Learning Generative Model. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 3514212.

37. Bao, J.; Chen, D.; Wen, F.; Li, H.; Hua, G. CVAE-GAN: Fine-Grained Image Generation through Asymmetric Training. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2745–2754.
38. Daubechies, I.; Lu, J.; Wu, H. Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool. *Appl. Comput. Harmon. A* **2011**, *30*, 243–261. [[CrossRef](#)]
39. LeCun, Y.; Boser, B.; Denker, J.; Henderson, D.; Howard, R.; Hubbard, W.; Jackel, L. Handwritten digit recognition with a back-propagation network. *Adv. Neural Inf. Process. Syst.* **1989**, *2*, 396–404.
40. Sharma, P.; Chandan, S.; Shaw, R.N.; Ghosh, A. *Vibration-Based Diagnosis of Defect Embedded in Inner Raceway of Ball Bearing Using 1D Convolutional Neural Network*; Elsevier: Amsterdam, The Netherlands, 2021; pp. 25–36.
41. Toh, G.; Park, J. Review of vibration-based structural health monitoring using deep learning. *Appl. Sci.* **2020**, *10*, 1680. [[CrossRef](#)]
42. Zhang, P.; Li, L. Vibration and noise characteristics of high-frequency amorphous transformer under sinusoidal and non-sinusoidal voltage excitation. *Int. J. Elec. Power* **2020**, *123*, 106298. [[CrossRef](#)]
43. Weiner, M.M. Magnetostrictive Offset and Noise in Flux Gate Magnetometers. *IEEE Trans. Magn.* **1969**, *5*, 98–105. [[CrossRef](#)]
44. Welch, P. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Trans. Audio Electroacoust.* **1967**, *15*, 70–73. [[CrossRef](#)]
45. Chinchor, N. MUC-4 evaluation metrics. In Proceedings of the 4th Conference on Message Understanding (MUC4 '92), McLean, VA, USA, 16–18 June 1992; pp. 22–29.
46. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Commun. Acm.* **2017**, *60*, 84–90. [[CrossRef](#)]
47. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
48. 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 (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
49. Helmi, H.; Forouzantabar, A. Rolling bearing fault detection of electric motor using time domain and frequency domain features extraction and ANFIS. *IET Electr. Power Appl.* **2019**, *13*, 662–669. [[CrossRef](#)]