UNSUPERVISED FEATURE-PRESERVING CYCLEGAN FOR FAULT DIAGNOSIS OF ROLLING BEARINGS USING UNBALANCED INFRARED THERMAL IMAGING

LUJIALE GUO

FACULTY OF ENGINEERING
UNIVERSITI MALAYA
KUALA LUMPUR

2025

UNSUPERVISED FEATURE-PRESERVING CYCLEGAN FOR FAULT DIAGNOSIS OF ROLLING BEARINGS USING UNBALANCED INFRARED THERMAL IMAGING SAMPLE

LUJIALE GUO

DISSERTATION SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ENGINEERING SCIENCE

FACULTY OF ENGINEERING
UNIVERSITI MALAYA
KUALA LUMPUR

2025

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Lujiale Guo

Matric No: S2137234

Name of Degree: Master Of Engineering Science

Title of Project Paper/Research Report/Dissertation/Thesis (“this Work”):

Unsupervised Feature-Preserving CycleGAN for Fault Diagnosis of Rolling Bearings Using Unbalanced Infrared Thermal Imaging Sample

Field of Study: Signal and System (NEC 520: Engineering and Engineering Trades)

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the Universiti Malaya (“UM”), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate’s Signature Date:19.02.2025

Subscribed and solemnly declared before,

Witness’s Signature Date:19.02.2025

Name:

Designation:

UNSUPERVISED FEATURE-PRESERVING CYCLEGAN FOR FAULT DIAGNOSIS OF ROLLING BEARINGS USING UNBALANCED INFRARED THERMAL IMAGING SAMPLE

ABSTRACT

The fault diagnosis of rolling bearing is of great significance in industrial safety. The method of infrared thermal image combined with neural network could diagnose the fault of rolling bearing in a non-contact manner, however its data in different scenes are often unbalanced and difficult to obtain. In this paper, an unsupervised learning framework named Feature-Preserving Cycle-Consistent Generative Adversarial Networks (FP-CycleGAN) is designed for defect detection in unbalanced rolling bearing infrared thermography sample. Since the classical Cycle-Consistent Generative Adversarial Networks (CycleGAN) not designed to accurately transfer the target features of the image. To avoid this problem, a new discriminator is designed to identify whether the generated image A and B (refer to different conditional bearing image) belongs to two different classes, and a new class loss are proposed. To better extract fault features and perform features migration, the new generator is reconstructed based on the U-Network structure, the transpose convolution method of the up-sampling network is replaced by Bicubic Interpolation to effectively avoid the checkerboard effect of the generated images. The defect detection of the expanded dataset was performed using Residual Network and compared with the pre-expansion data to demonstrate the usability of the generated data and the superiority of the proposed FP-CycleGAN method for rolling bearing defect detection in small samples of infrared thermal images. Finally, the accuracy of the proposed model is 91.52%, which is better than the baseline model (76.81%).

Keywords: Fault diagnosis, rolling bearing, infrared thermal imaging, unbalanced data, generative adversarial networks

CYCLEGAN PEMELIHARAAN CIRI-CIRI TANPA DIAWAS UNTUK DIAGNOSIS KESALAHAN BEARING GELARAN MENGGUNAKAN SAMPEL PENGIMEJIAN TERMA INFRA MERAH TIDAK SEIMBANG

ABSTRAK

Diagnosis kesalahan galas bergolek adalah sangat penting dalam keselamatan industri. Kaedah imej terma inframerah digabungkan dengan rangkaian saraf boleh mendiagnosis kesalahan galas bergolek secara bukan sentuhan, namun datanya dalam adegan yang berbeza selalunya tidak seimbang dan sukar diperoleh. Dalam kertas kerja ini, rangka kerja pembelajaran tanpa pengawasan yang dinamakan Rangkaian Adversarial Generatif Konsisten Kitaran-Pemeliharaan Ciri direka untuk pengesanan kecacatan dalam sampel termografi inframerah galas berguling tidak seimbang. Oleh kerana Rangkaian Adversarial Generatif Konsisten Kitaran (CycleGAN) klasik tidak direka bentuk untuk memindahkan ciri sasaran imej dengan tepat.. Untuk mengelakkan masalah ini, diskriminator baharu direka bentuk untuk mengenal pasti sama ada imej A dan B (merujuk kepada imej galas bersyarat yang berbeza) yang dijana tergolong dalam dua kelas berbeza, dan kehilangan kelas baharu dicadangkan. Untuk mengekstrak ciri kerosakan dengan lebih baik dan melakukan migrasi ciri, penjana baharu dibina semula berdasarkan struktur U-Network, kaedah convtraspose bagi rangkaian pensampelan atas digantikan dengan Interpolasi Bicubic untuk mengelakkan kesan papan dam bagi imej yang dijana dengan berkesan. Pengesanan kecacatan set data yang dikembangkan telah dilakukan menggunakan Rangkaian Sisa dan dibandingkan dengan data pra-pengembangan untuk menunjukkan kebolehgunaan data yang dijana dan keunggulan kaedah FP-CycleGAN yang dicadangkan untuk pengesanan kecacatan galas bergolek dalam sampel kecil imej terma inframerah. Akhir sekali, ketepatan model yang dicadangkan ialah 91.52%, iaitu lebih baik daripada model garis dasar (76.81%).

Keywords: Diagnosis kerosakan, galas bergolek, pengimejan terma inframerah, data tidak seimbang, rangkaian musuh generatif

ACKNOWLEDGEMENTS

Here, I'd like to thank my supervisor, Professor Ir. Dr. Chuah Joon Huang, for his teaching and help in my postgraduate study. He gave me the opportunity to study for this master's degree and gave me valuable experience in code writing and thesis writing. Thanks to my undergraduate tutor, Gu Xiaohui for providing me with experimental data and research help. I would also like to thank my other supervisor, Dr. Wong Jee Keen Raymond, for his help during the whole study period.

In addition, I would also like to thank my roommate, Bingrui Huang, for his support and companionship in my daily study and life. Thanks for the spiritual encouragement given to me by my best friend, Pu Tao, at the final stage of writing and submitting the paper.

I also want to thank the Office staff from the faculty of Engineering of the UM who helped me and the teachers who taught me.

Lujiale Guo

TABLE OF CONTENTS

Abstract iii

Abstrak iv

Acknowledgements vi

Table of Contents vii

List of Figures ix

List of Tables x

List of Symbols and Abbreviations xi

CHAPTER 1: INTRODUCTION 1

1.1 Background 1

1.2 Problem statement 4

1.3 Questions 5

1.4 Objectives 6

1.5 Scopes of the Research 6

1.6 Organization of the Thesis 6

CHAPTER 2: LITERATURE REVIEW 8

2.1 Introduction 8

2.2 Vibration signals analysis 8

2.3 Current analysis 12

2.4 Sound pressure analysis 14

2.5 Rotary encoder analysis 16

2.6 Infrared thermal imaging 19

2.7 Sample imbalance algorithm 22

2.8 Generative algorithm 24

2.9 Convolutional neural network 27

CHAPTER 3: METHODOLOGY 29

3.1 Theoretical background 29

3.1.1 Generative Adversarial Networks 29

3.1.2 Cycle Consistent Generative Adversarial Network 30

3.1.3 Interpolation Algorithm 33

3.1.4 Depthwise Separable Convolution 35

3.2 The proposed approach 37

3.2.1 Architecture of the generator 39

3.2.2 Architecture of the discriminator 43

3.2.3 Loss function 44

3.3 Summary 47

CHAPTER 4: RESULTS 49

4.1 Introduction 49

4.2 Collection and description of laboratory bearing data sets 49

4.3 Sample expansion performance evaluation 53

4.4 Unbalance fault diagnosis results and analysis 57

4.5 Summary 74

CHAPTER 5: CONCLUSION 76

References 78

List of Publications and Papers Presented 102

LIST OF FIGURES

Figure 2.1: Vibration detection test bench..9

Figure 2.2: DLMD overflow..11

Figure 2.3: Current detection test bench..12

Figure 2.4: Sound pressure detection test bench..15

Figure 2.5: Simplified structure of an optical encoder..17

Figure 2.6: Example of an infrared thermal image of the bearing housing. 20

Figure 2.7: Basic Lenet-5 bearing defect detection neural network 21

Figure 2.8: Illustration on transfer learning-based fault diagnosis 23

Figure 2.9: An residual learning unit 28

Figure 3.1: Basic structure of GAN 29

Figure 3.2: Basic structure of CycleGAN..31

Figure 3.3: The framework of the FP-Cyclegan 38

Figure 3.4: The architecture of the ConvNeXt Block..40

Figure 3.5: The architecture of the Generator..41

Figure 3.6: The architecture of the discriminator D_AB..43

Figure 4.1: The rotating machinery test bench 50

Figure 4.2: Thermal image acquisition system 51

Figure 4.3: Single angle and state bearing rain900 data set..52

Figure 4.4: Multi angle and state bearing rain900 data set 53

Figure 4.5: Images generated by different up-sampling methods..56

Figure 4.6: Images generated by different data amplification methods..60

Figure 4.7: The confusion matrix for the classification network..69

LIST OF TABLES

Table 4.1: Health conditions and labels of rolling bearings... 49

Table 4.2: Hyperparameter setting... 53

Table 4.3: Results of different up-sampling methods 55

Table 4.4: Results of different convolution methods 57

Table 4.5: Results of different data amplification method.. 61

LIST OF SYMBOLS AND ABBREVIATIONS

FP-CycleGAN	:	Feature-Preserving Cycle-Consistent Generative Adversarial Networks
CycleGAN	:	Cycle-Consistent Generative Adversarial Networks
MK-MMD	:	Mean Difference Multi-Kernel variant
PBR	:	Personalized binary relevance
HML-KNN	:	Hierarchical multi-label K-nearest neighbor
PHM	:	Prognosis and health management
VAEs	:	Variational Autoencoders
GANs	:	Generative Adversarial Networks
ViTs	:	Vision Transformers
FID	:	Fréchet Inception Distance

CHAPTER 1: INTRODUCTION

1.1 Background

At present, machinery has always played a very important role in various fields, and the rolling bearing system, as its core component, has always been concerned about its safety (Rai & Upadhyay, 2016; Wu et al., 2022). The faults of bearings usually show corrosion, crack or peeling on the inner ring, outer ring, or roller. For the fault detection of bearings, the methods could be summarized as vibration analysis, current characteristic analysis, sound pressure analysis, rotary encoder analysis and infrared thermal image detection, etc (Hakim et al., 2023a). In the past, most of the methods used by people focused on the vibration analysis of rotor bearings (C. Li et al., 2019). However, the vibration analysis has some problems that affect the structure of the equipment, and it is difficult to install sensors (S. Zhang et al., 2020). In addition, because of the long signal transmission path, changeable working conditions and strong noise in practical applications, the processing of vibration signals will be very complicated (B. Sun et al., 2025).

The traditional bearing defect detect method including vibration, current, sound pressure and rotary encoder analysis For example, the vibration analysis is a common bearing fault detection method, which is based on monitoring and analyzing the vibration signals generated by bearings during operation (Dolenc et al., 2016). When the bearing starts to fail, such as cracks or peeling, its vibration characteristics will change obviously. By monitoring these changes, potential failures could be identified and prevented in time. The advantage of this method is that it could be monitored in real time and has good sensitivity to early faults (Bianchini et al., 2011; Prudhom et al., 2017). However, this method requires complex signal processing technology to filter environmental noise and non-fault related vibration of the equipment itself, and also has high requirements on the installation position and mode of the sensor, which may affect the normal operation and maintenance of the equipment (Pacheco-Chérrez et al., 2022). However, none of these methods can avoid the noise caused by vibration or the inconvenience of sensor installation.

Infrared thermal image detection method uses infrared camera to capture the thermal image of bearing during operation (A. Kumar et al., 2024). By analyzing the temperature distribution in thermal images, we could find hot spots caused by excessive friction, poor lubrication or other mechanical failures (AlShorman et al., 2024). Abnormal heating of bearings is usually an obvious signal of fault development, and infrared detection could be used to identify and take measures in time. The advantage of this technology is that it could be detected remotely and without contact, which is very suitable for those applications that are difficult to access or in high temperature environment (Glowacz, 2024). However, the equipment cost of infrared thermal imaging technology is high, and the professional knowledge of operators is high, so proper training is needed to effectively carry out fault analysis.

Deep learning has gained significant attention in recent years for its ability to automate and enhance bearing fault detection systems (F. Dong et al., 2025; Snyder et al., 2025; S. Xie et al., 2025). Traditional methods such as vibration analysis, sound pressure analysis, and infrared thermal imaging often struggle with high noise levels, data complexity, and real-time processing challenges (Glowacz et al., 2025; J. Ma et al., 2024; R & Mutra, 2025). Deep learning techniques, particularly convolutional neural networks (CNNs) and other advanced architectures, have demonstrated superior performance in addressing these issues by automatically extracting relevant features from raw data (D. He et al., 2025; M. Liu et al., 2024).

One of the primary advantages of deep learning in bearing fault detection is its ability to learn hierarchical feature representations directly from the raw sensor data, bypassing the need for manual feature engineering (D. He et al., 2025; P. Wang et al., 2025; T. Yang et al., 2025; Ye et al., 2025). This is particularly beneficial when dealing with large datasets from multiple sensors (Rivas et al., 2024). For instance, deep residual networks (ResNets) have been used in fault diagnosis systems to automatically extract features from vibration signals, significouldtly improving diagnostic accuracy, especially in noisy environments (Han et al., 2024).

Recent studies have explored deep learning's potential to handle multi-sensor fusion for bearing fault detection, where data from various sensor types, such as vibration, acoustic, and current signals, are combined to enhance detection performance (X. Li, Wang, et al., 2024; Ye et al., 2025; W. Zhang et al., 2024). In these systems, deep learning models like long short-term memory networks (LSTMs) and CNNs are used to analyze time-series data, which helps in detecting early-stage bearing faults that may not be identifiable by traditional methods (X. Li, Wang, et al., 2024; Shuming et al., 2025; Z. Xu et al., 2024).

The robustness of deep learning models in fault detection is also demonstrated in their ability to work with limited labeled data through transfer learning techniques (X. Li et al., 2025; Shi et al., 2025; Wei et al., 2025). Transfer learning has proven to be highly effective in adapting pre-trained models to new environments or machines with minimal additional training data (Deveci et al., 2024; Guo et al., 2025; Xiang et al., 2025). This approach has been successfully employed in various studies, such as the work by Lian et al. and S. Tang et al., where transfer learning models were trained on vibration signals and acoustic data, allowing for more generalized fault detection capabilities across different operating conditions (Lian et al., 2025; S. Tang et al., 2024).

Additionally, the integration of deep learning with image-based defect detection has shown great promise (Jiang et al., 2025; P. Kumar & Hati, 2022; J. Li et al., 2022; Y. Liu et al., 2025). For example, methods like the improved YOLOv5 model have been applied to bearing surface defect detection, where images of defects are analyzed using a deep learning framework that excels in object detection tasks (Hu et al., 2024). These models are particularly advantageous in workshops or environments where visual inspection is critical and could significouldtly reduce human error.

Moreover, the development of hybrid models that combine deep learning with other advanced technologies, such as the minimum unscented Kalman filter (MUKF), further enhances fault diagnosis accuracy by efficiently extracting features from multi-sensor signals (H. Tang et al., 2024). These hybrid approaches enable more reliable fault detection, even in challenging environments with strong noise interference (X. Xie et al., 2024; D. Zhao et al., 2024).

In conclusion, deep learning technologies, particularly CNNs, LSTMs, and transfer learning, are revolutionizing bearing fault detection. Their ability to process large, complex datasets with high accuracy, even in noisy or unpredictable conditions, makes them a powerful tool in ensuring the reliability and safety of machinery in various industrial applications.

1.2 Problem statement

With the development of deep learning, it has been applied to the field of industrial target detection and is playing an increasingly important role. Thermal imaging technology is gradually combined with neural network, which could automatically identify faults that are helpful for early warning in a non-contact way, regardless of the speed of rotating machinery, to ensure the reduction of system downtime caused by bearing faults (Hakim et al., 2023b).

However, at present, the thermal image diagnosis method based on CNN could only deal with the same working conditions, but the working environment is changeable in practical application. A key challenge lies in developing a rolling bearing detection network that maintains high generalization performance across diverse operating conditions while being trained on a limited number of samples.

At present, although there has been transfer learning or adversarial generative learning to solve the problem of low generalization of deep neural network, these methods are often not suitable for the latest new technology of non-contact bearing defect detection, or they are often used in other fields, which need further application, especially the application of adversarial generative network to image data (T. Zhang et al., 2022).

Generative Adversarial Network is a technique of generating nonexistent images with the help of original images. It could extract and transfer the features of the original images, to obtain false images with different features in different backgrounds and achieve the effect of expanding data sets. This method is simple and efficient and could save a lot of manpower and resources. However, how to correctly use it to extract features from infrared thermal images and how to prove the authenticity of new images needs to be further explored through experiments.

1.3 Questions

Question 1: According to the infrared thermal image of rotor, what structure of neural network could achieve higher detection efficiency and accuracy?

Questions 2: How does GAN network amplify the infrared thermal image samples of bearings?

Questions 3: How to prove the authenticity and usability of GAN network?

1.4 Objectives

Objectives 1: To select the backbone network for better detection ability, compare and choose the more advantageous activation function, optimizer, and network structure.

Objectives 2: To set up an experimental platform to collect the infrared thermal images of bearings in different environments and select the appropriate GAN network for data amplification according to the diversity of samples.

Objectives 3: To set up an ablation experiments evaluating the performance of various GAN structures and validating the usability of generated images for bearing defect detection.

1.5 Scopes of the Research:

The scope of this study is to train a Generative Adversarial Networks on the infrared thermal imaging of bearings and improve the performance of the damage detection network through the generated images. The scope of this study includes but is not limited to:

1. Collect infrared thermal images of seven different damaged bearings and ensure different angles and distances.
2. The proposed algorithm is written and trained by using pytorch library.
3. Realize other algorithms to generate images.
4. Compared with other methods, the proposed algorithm is tested to ensure that the algorithm has better performance in image generation and damage detection network optimization.

1.6 Organization of the Thesis:

This thesis is organized into five main chapters. i) introduction; ii) literature review; iii) proposed methodology; iv) results and discussion and lastly; v) conclusion and future recommendation. The contents of each section in the thesis are summarized as follows:

Chapter 1: This chapter expounds the background of the subject and the importance of bearing damage detection. It also describes the problem statement that provides the basis for the goal of this study. According to the purpose of the study, this dissertation also explains the scope of the study.

Chapter 2: This chapter profoundly introduces the algorithm of bearing detection. First, it gives an overview of different methods for bearing detection. It also explains the advantages and disadvantages of various bearing detection methods.

Chapter 3: In this chapter, the details of how to design the proposed algorithm are discussed in detail, the principles and advantages of Depthwise Separable Convolution and linear interpolation are explained, and a new loss function is added to the method to restrict the transfer of features in the generated image.

Chapter 4: The comparison of this model with other technologies is presented in this chapter, evaluates the influence of the proposed linear interpolation, separable convolution and new loss function on the results, and analyzes the overall performance of the proposed model by using one subjective and eight different objective evaluation metrics.

Chapter 5: This chapter summarizes the work of this paper, and discusses and optimize the future work.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

Currently, machinery plays a crucial role in various fields, and the rolling bearing system, as its core component, has always been a matter of great concern in terms of safety (J. Ma et al., 2024; Soomro et al., 2024). In complex or high intensity working environments, bearings often suffer from damage such as corrosion, cracks, or spalling in the inner ring, outer ring, or rolling elements, which could potentially lead to industrial accidents (Raouf et al., 2024).

Therefore, conducting fault detection on rolling bearings to proactively identify the damage and enhance system safety is important (R. Liu et al., 2018; Shao et al., 2018; Xiong et al., 2023).

2.2 Vibration signals analysis

Vibration signals are frequently utilized as a method for diagnosing faults in bearings because bearings with different types of damage respond correspondingly to external stimuli while in operation, thus generating corresponding vibration fault signals within the vibration signal (Jin et al., 2025). For example, Dalvand et al. found that the stator current of a typical asynchronous motor involves the fundamental wave of power supply and its harmonics before and after bearing defects. A current noise reduction method based on time shift is proposed. The current residual could be obtained by adding the digital current signal to its own sampling delay representation as an anti-noise component. The amount of sampling delay depends only on the power frequency and sampling rate. This amount is set to eliminate the fundamental wave of the power supply and its odd harmonics. After obtaining the current residual, the characteristic fault frequency of the faulty bearing could be revealed by analyzing the frequency spectrum of the envelope of the current residual (Dalvand et al., 2017; S. Yang et al., 2020). The vibration detection test bench is usually shown in Figure 2.1.

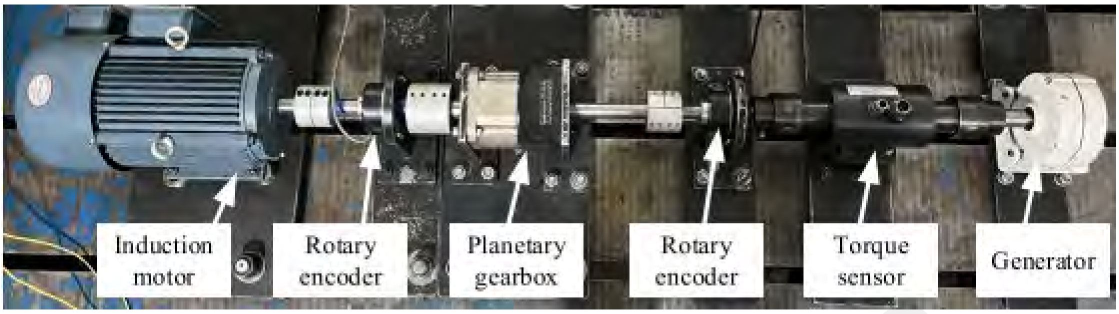


Figure 2.1: Vibration detection test bench

Furthermore, the low signal-to-noise ratio of fault signals in the early stage of motor bearing fault often leads to the low accuracy of fault state identification. Aiming at this problem, the bearing early fault diagnosis method based on differential local mean decomposition (DLMD) and current-vibration signal fusion can obtain better results. In this method, the current signal and vibration signal are decomposed by DLMD respectively, and the decomposed product function (PF) is weighed to reconstruct the signal according to kurtosis value, and the reconstructed signal is fused and normalized to obtain the current. Finally, the Hilbert envelope spectrum analysis of the fused signal is carried out, and the clear fault characteristic frequency is obtained. (B. Zhang et al., 2024)

DLMD is an improvement of traditional local mean decomposition. DLMD differentiates the signal without changing the frequency of the original signal, keeps the frequency with higher energy, and gradually filters out smaller stray interference signals, thus extracting fault characteristic signals. The specific steps of DLMD are as follows:

Perform $k-th$ order differentiation of the original signal to obtain $x^{(k)}(\mathrm{t})$.

Identify all local extremum points n_i of the differentiated signal $x^{(k)}(\mathrm{t})$, and calculate the average of all adjacent local extremum points:

$$m_i = \frac{n_i + n_{i+1}}{2} \tag{1}$$

where n_i is the $i-th$ local extremum point, n_{i+1} is the $(i+1)-th$ local extremum point, and m_i is the $i-th$ mean value. Because the number of n_i characterizes the degree of waveform distortion and the content of impulse components in the $x^{(k)}(\mathrm{t})$, the number of m_i can also reflect the degree of waveform distortion and the content of impulse components in the signal. Connect all adjacent mean points m_i with straight lines and apply a moving average method to smooth the resultant local mean function $m_{11}(t)$ This function reflects the overall trend and periodic variations of the signal. Clearly, $m_{11}(t)$ is influenced by the mean points m_i. However, compared to the number of m_i, the values of m_i have a more significant impact on $m_{11}(t)$. If the differences between each mean point are large, then the fluctuation of the local mean function $m_{11}(t)$ will be large; conversely, if the differences are small, the fluctuation will be small. Therefore, the distribution of m_i has an impact on the subsequent signal decomposition process in the form of the local mean function $m_{11}(t)$.

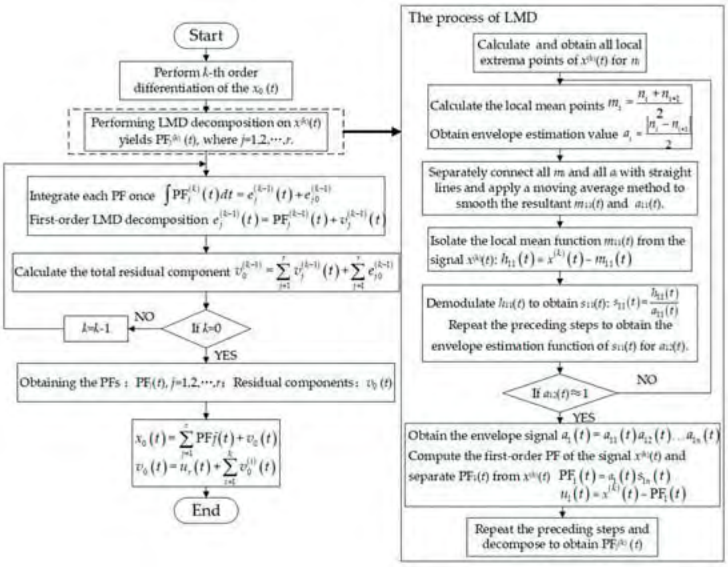


Figure 2.2: DLMD Overflow

The specific DLMD process is shown in figure 2.2

In addition, many methods based on vibration analysis have been developed for other kinds of bearings. For example, for planetary bearings, when they are running, they need to bear the large torque load borne by planetary gear sets, which is prone to failure, but the kinematic characteristics and vibration characteristics of planetary bearings are complex, which often lead to poor fault diagnosis results. To solve this problem, the order spectrum analysis method based on dual rotary encoders can be used. In this method, the double rotary encoders are installed on the input shaft and the output shaft of the planetary gearbox respectively. Firstly, the zero-crossing detection method is used to identify the encoder pulse occurrence time, estimate the torsional vibration and resample in the angular domain. Then, the torsional angular displacement of each encoder pulse is estimated by using the angle difference of the two encoders. Accordingly, the torsional angular velocity and acceleration are estimated by using the first and second time derivatives of torsional angular displacement respectively. Next, the angular displacement, velocity and acceleration are resampled in the angular domain according to the encoder pulse generation time which is evenly distributed at a constant angle. Finally, the resampled torsional angular displacement, velocity and acceleration are Fourier transformed respectively, and the corresponding order spectra are obtained. The relative rotation between the two rotary encoders only reflects the torsional vibration of the planetary gearbox. This characteristic makes the proposed method free from the additional modulation effect caused by the rotation of the planet carrier and the interference of input and output devices. In addition, the order spectrum method further makes this method suitable for non-stationary signal analysis under the condition of time-varying speed, which can successfully detect local faults on the inner ring, outer ring and roller of planetary bearings under the condition of constant and time-varying speed. (Feng et al., 2025)

However, vibration signal analysis faces several limitations, such as difficulties in sensor installation, the potential impact on equipment structure during signal acquisition, the variability of working conditions, and the presence of strong noise in practical applications.

2.3 Bearing defect detect by current analysis

Most machines in industry are driven by induction motors. Under extreme working conditions, both electrical and mechanical systems may fail prematurely. Bearing fault is the most common fault in motor (40%), followed by stator fault (37%) and rotor fault (10%). The result of these faults is increased operating cost and prolonged downtime. Different from bearings in mechanical devices, bearing defect detection in motors often includes vibration analysis, axial magnetic flux analysis, lubricating oil debris analysis, partial discharge and motor current characteristics analysis. Current detection test bench is usually shown as Figure 2.3.

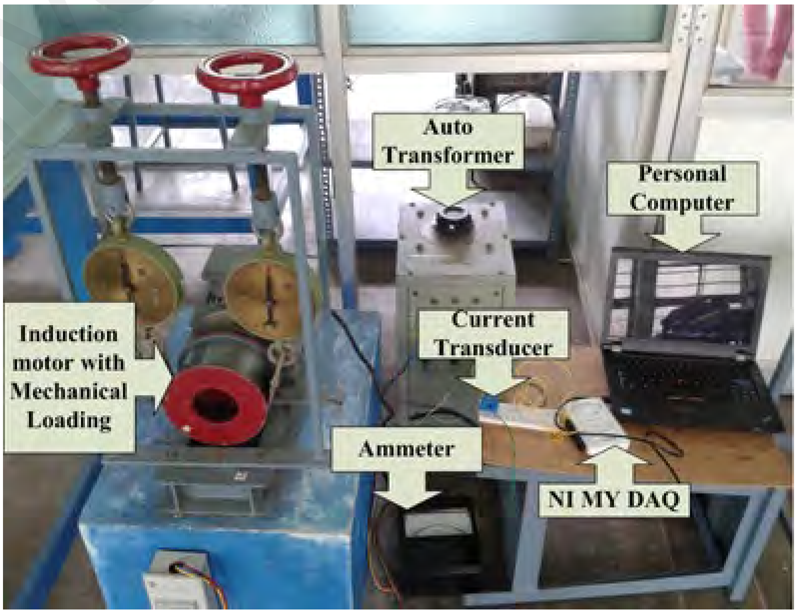


Figure 2.3: Current detection test bench

The current characteristic analysis detects the bearing fault by analyzing the current signal of the motor (Singh & Kumar, 2017). This method takes advantage of the abnormal fluctuation of motor current when the bearing fails. Because bearing damage will lead to changes in motor load and magnetic field distribution, and then affect current characteristics, monitoring these changes could effectively identify bearing problems (Blodt et al., 2008). The advantage of the current characteristic analysis method is non-invasive and does not need direct contact with bearings, thus reducing the physical interference to equipment. The analysis of motor current characteristics is usually carried out by Fast Fourier Transform (FFT) (Marcelo et al., 2012).

However, the traditional FFT has some shortcomings, such as poor resolution, spectrum leakage and inability to provide time-frequency relationships. Therefore, it is difficult to find out when the fault occurred, and if the amplitude of the fault is low compared with the noise generated by the machine, it is difficult to find out the fault components by using the traditional FFT. In order to overcome these shortcomings of FFT, many advanced signal processing technologies have been realized. It is a convenient and accurate method to use stator current spectrum subtraction and various wavelet decomposition, and the static wavelet transform based on stator current spectrum subtraction has a good indication for any type of fault, but different types of current spectrum may be suitable for different types of bearing faults, which limits the generalization of this method (Deekshit Kompella et al., 2018).

In addition, for the fault detection and diagnosis of asynchronous motors, some studies have proved that the influence of bearing fault on motor current can be ignored. On the contrary, the more likely reason why faults can be detected in current is that they will produce fluctuating resistance torque, which will work immediately, while radial displacement takes time to integrate into perceptible displacement (Stack et al., 2004).

This is true even when the speed changes step by step. The method of detecting bearing defects based on instantaneous power factors which changes with torque oscillation can deal with this situation well. This method is a new method based on instantaneous power factor. By modeling bearing defects, it is shown that the possibility of radial movement is very small, and the more likely reason why faults appear in current is that they will produce fluctuating resistance torque, which will be reflected in stator current. The results prove that the detection effect of instantaneous power factor as a signal is better than instantaneous power (Ibrahim et al., 2008).

However, this method requires high accuracy of current measuring equipment and stability of motor running state and may require special hardware and software support.

2.4 Bearing defect detect by sound pressure analysis

At present, a large number of research results on detecting and diagnosing bearing defects by vibration and acoustic methods have been published, most of which have been published in the past twenty years (Tandon & Choudhury, 1999). For example, the fault diagnosis method of multi-signal fusion is one of the current research trends, which can improve the reliability of diagnosis results. A new fusion demodulation method is constructed by using two signal demodulation methods to extract the characteristic frequency of single channel signal. Then, the characteristic spectrum of multi-channel signals is fused to extract the final characteristic frequency, which successfully reduces the content of noise components in the characteristic spectrum and highlights the fault characteristic frequency (Zheng et al., 2025). Sound pressure detection test bench usually be designed as Figure 2.4.

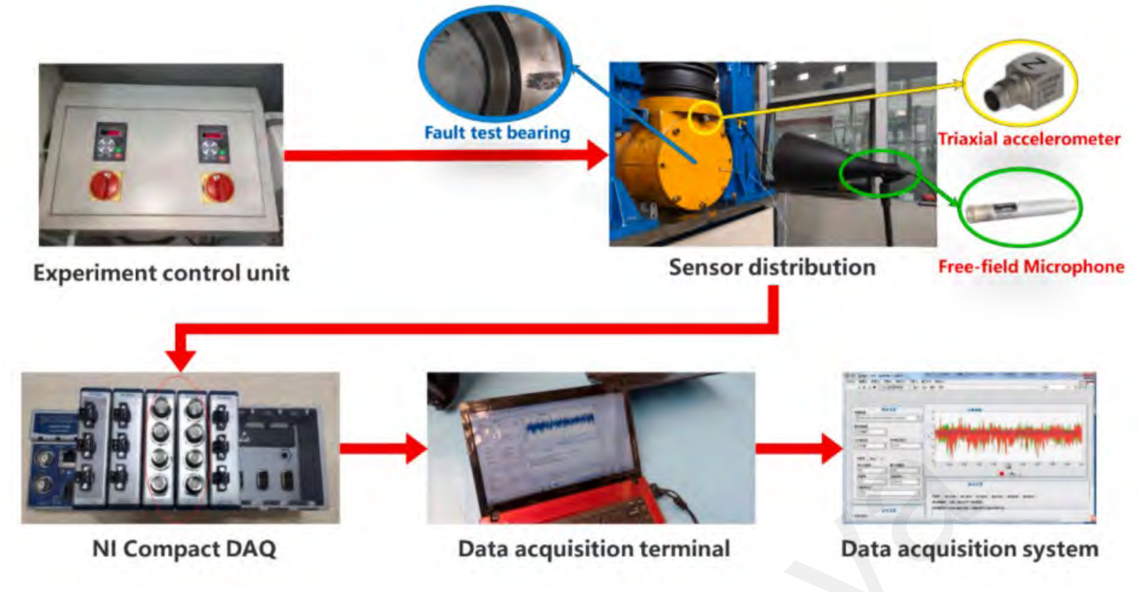


Figure 2.4: Sound pressure detection test bench

Furthermore, the vibration analysis (VA) and acoustic emission (AE) methods have also achieved good results in predicting the severity of rolling bearing defects, because the signals obtained from bearings by sensors based on vibration methods will be distorted by other faults and mechanical noise of equipment. AE is a nondestructive testing technology for structural health monitoring. Because of its inherent high signal-to-noise ratio, its application in bearing defect diagnosis is gaining development momentum as an alternative diagnostic tool. Rao and Ratnam focus on the same type of seed defects on the outer ring of cylindrical roller bearings with radial load and run the defective bearings at different speeds and loads. A method of bearing defect detection based on AE and vibration probe data is designed and achieved good results (Rao & Ratnam, 2015).

The sound pressure analysis uses acoustic sensors to capture the sound waves generated when the bearing is running and detects the abnormal state of the bearing by analyzing the frequency and amplitude of these sound waves (J. Yang et al., 2024). When the bearing is damaged, such as cracks or peeling, it will produce sound waves with specific frequencies, which are obviously different from those in normal operation (Z. Zhang & Wu, 2024). Sound pressure analysis method could find problems in the initial stage of bearing damage, which is very suitable for those application scenarios that require high equipment safety (Guan et al., 2024).

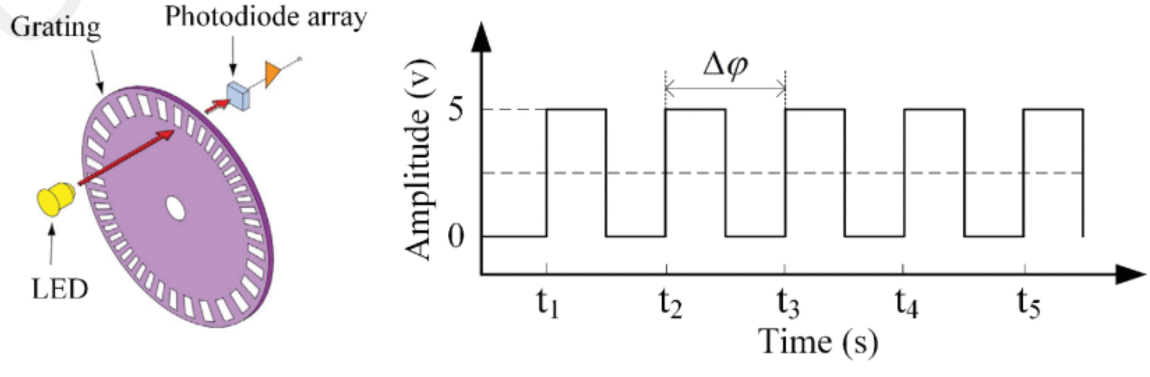
However, the method based on sound pressure is sensitive to environmental noise and needs to be carried out in a relatively quiet environment to avoid wrong diagnosis.

2.5 Bearing defect detect by rotary encoder analysis

The rotary encoder analysis relies on the encoder installed on the shaft to monitor the rotating speed and position of the bearing (Pang et al., 2024). By accurately measuring the rotation parameters of bearings, we could identify the slight changes caused by bearing faults. These changes include slight misalignment or imbalance of bearings, which may be early signs of bearing failure. The advantage of rotary encoder analysis method lies in its high accuracy, which is especially suitable for high-precision mechanical systems that need fine control of speed and position (Huang et al., 2024).

When the rolling bearing has a local fault, the change of contact stiffness between the rolling body and the raceway at the fault position will lead to the regular fluctuation of instantaneous angular velocity (IAS) (Y. Wang et al., 2020), that is, the optical encoder contains abundant fault information, which can be used for condition monitoring of the failed rolling bearing. Compared with vibration, IAS has the advantages of low noise, direct correlation with mechanical dynamics, no need for periodic calibration and short transmission path (Miao et al., 2020). In addition, compared with visual measurement and acoustic emission, the optical encoder has the advantage of no external installation, and can extract IAS signals from the control system. Therefore, condition monitoring based on IAS has become hot research in the field of fault diagnosis.

In recent years, because gears directly transmit torque, IAS signals are mainly used in gear-related fault diagnosis (M. Zhao et al., 2018). Although IAS signal contains abundant rolling element bearing (REB) fault information, REB does not directly transmit torque, that is, IAS fluctuation caused by REB fault requires a certain radial load. Therefore, IAS fluctuation caused by REB fault is weak (Renaudin et al., 2010). In addition, the estimated IAS signal also includes some interference components, such as measurement error (Coats & Randall, 2014), encoder installation error and speed trend component caused by speed change; Because the regular IAS fluctuation caused by REB fault is weak under low load conditions, these interference components cannot be ignored for feature extraction of REB fault. Aiming at the installation error of encoder, cross-correction (Deng et al., 2013) method and self-correction (Zeng et al., 2020) scheme are proposed, but cross-correction scheme needs high-precision reference devices, such as ring laser or polyhedral autocollimator, and self-correction method must provide high-precision installation of reference devices. It should be noted that in some cases, such as industrial robots, the above installation requirements and harsh operating conditions may not be met. On the other hand, although the angle-based IAS is a stationary signal under the condition of variable speed, the amplitude of the velocity trend component will change with the speed, and the conventional Qualcomm filtering may not be able to reduce the interference component, because the fault spectrum related to REB may be at a low frequency under the condition of variable speed. The simplified structure of an optical encoder is shown as figure 2.5 (a) and voltage output is shown as figure 2.5 (b).



(a) (b)

Figure 2.5: (a) Simplified structure of an optical encoder. (b) Voltage output.

Savitzky-Golay (SG) filter is an effective tool for reducing noise while maintaining the shape and height of waveform peaks (Schafer, 2011). However, the effectiveness of the SG filter depends on the parameters (filter order P and fitting length 2M+1), that is, the optimal parameters determination of the SG filter for different conditions remains an open problem, which is often dependent on experience and prior knowledge.

The SG filter is a widely used polynomial fitting tool, which can eliminate the interference components by setting suitable parameters (fitting length $2M+1$ and polynomial order P), it is defined as:

$$y[n] = \sum_{m=n-M}^{m=n+M} h[n-m]x[m] \tag{2}$$

where $y[n]$ and $x[m]$ indicate the fit data and the raw data, the convolution coefficient $h[n-m] = (B^T B)^{-1} B^T$, $B^T = [M_a^0, M_b^1, \cdots, M_b^p]^T$, $M_b^p = (-M)^p, \cdots, (-1)^p, 0, 1^p, \cdots$, M^p, $M_a^0 = -M, \cdots, -1, 1, 1, \cdots, M, P = 1, 2, \cdots$.It should be noted that h depends on P and $2M+1$, that is, the same weighting coefficients will be obtained at each group of $2M+1$ samples, and so we can think the SG filtering as a shift-invariant discrete convolution process (X. Chen et al., 2024). The SG filter requires that $2M+1$ is odd and $P \leq 2M$.

In order to effectively extract the characteristics of fault REB from IAS signals under variable speed conditions, an adaptive SG filter scheme can be adopted. Firstly, the parameter decomposition structure (PDS) is introduced, which can obtain high-precision optimal parameters with low computational cost. Secondly, the improved diagnostic characteristic index can be used to evaluate the effectiveness of SG filters with different parameters under the condition of harmonic interference, which can be used to supplement the traditional diagnostic characteristic index.

However, this method requires very high mechanical installation of equipment, and the installation and maintenance costs of encoders are relatively high.

2.6 Infrared thermal imaging

Infrared thermal imaging is a non-contact and non-invasive temperature measurement technology, which has the advantage of not changing the surface temperature and displaying the real-time temperature distribution. This technology has been applied in many industrial and/or research fields, including meteorology, environment, medicine, architecture, engineering and so on, in which temperature is the key parameter. The principle of measurement is based on the fact that any physical object will radiate energy at infrared wavelength (that is, in the infrared range of electromagnetic spectrum). The thermal imager can measure and visualize the infrared radiation caused. Therefore, the surface temperature distribution is recorded in the form of thermal maps. Based on this characteristic, thermal images are currently used in the field of machine condition monitoring and diagnosis (Younus & Yang, 2012).

Infrared thermal imaging is a commonly used thermal measurement and diagnosis method, which is based on the measurement of infrared rays emitted by solids at a given temperature (Khor et al., 2024; Seong & Kim, 2024). The obtained spectrum is used to estimate the surface temperature of the elements under consideration (AlShorman et al., 2024; Glowacz, 2024). Compared with the contact sensor that provides single point measurement, thermal imaging technology could estimate and visualize the temperature distribution on the surface of mechanical system (Touret et al., 2018).

Related literature has proved that infrared thermal imaging could detect several faults and situations in rotating machinery to some extent, such as rotor imbalance, misalignment, loose coupling, insufficient lumbriculid and damage of rolling bearings (Janssens et al., 2019). For some cases, it is not difficult to see the advantages of infrared thermal imaging, such as the detection of lubricant shortage. One of the main uses of lubricants is to control friction. If there is too little lubricant in the bearing housing, excessive friction will occur, resulting in heat that can be observed by infrared thermal imager. In addition, if there is too much lubricant in the bearing, agitation will occur and additional heat will be observed. However, in some cases, methods based on infrared thermal imaging can only provide suboptimal results, such as detecting the damage from rolling bearings. For example, the outer raceway damage can only be detected 45% of the time using the system based on infrared thermal imaging (Liao et al., 2023). Figure 2.6 shows an example of an infrared thermal image of the bearing housing.

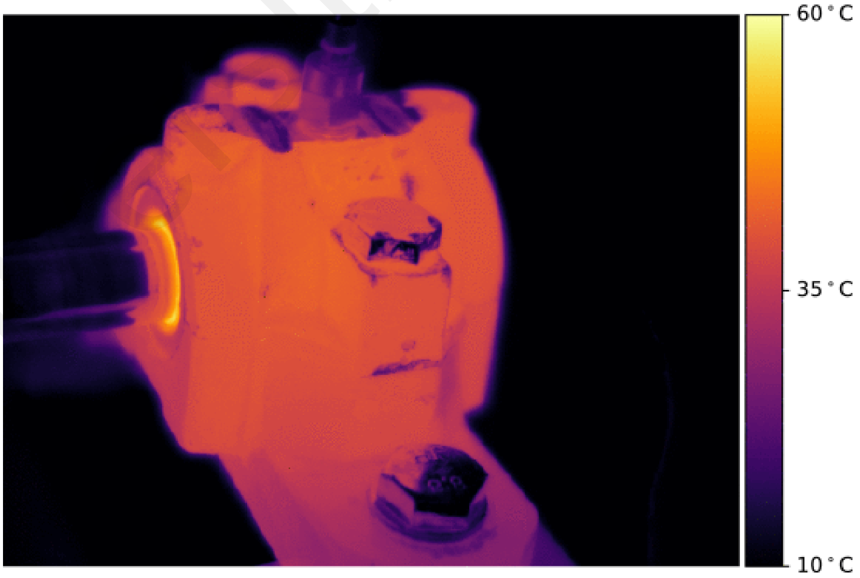


Figure 2.6: Example of an infrared thermal image of the bearing housing

Recently, the combination of infrared thermal imaging and artificial intelligence has gained increasing attention in the non-destructive fault diagnosis of rotating machinery. Compared to vibration signals, infrared thermal imaging offers advantages such as easy installation, non-contact usage, and high precision (Shao et al., 2021). Choudhary et al. proposed a method for detecting bearing defects using infrared thermal imaging by combining artificial neural networks with convolutional neural networks. Their approach addresses the challenge of early bearing failure detection under different bearing conditions (Choudhary et al., 2021). This method firstly extracts continuous frames from IRT data, and then features the bearing fault based on ANN. Multiple parameters are used to interpret the time domain signal. A group of 15 such statistical time-domain parameters or features, namely mean value, root mean square value, standard deviation, shape factor, kurtosis, skewness, peak amplitude, crest factor, pulse factor, signal-to-noise ratio, variance, vitality, entropy, marginal factor and signal-to-noise ratio-to-distortion ratio, are extracted from thermal images of six data sets with different bearing conditions based on expert knowledge. The extracted features are then normalized in the range of 0 to 1, and then the classification accuracy is improved by neighborhood component analysis (NCA). The CNN model used is based on the LeNet-5 model architecture. This method is famous for its excellent performance in fault diagnosis of rotating machinery because of its pixel-level feature extraction characteristics. Basic Lenet-5 bearing defect detection neural network is shown as Figure 2.7.

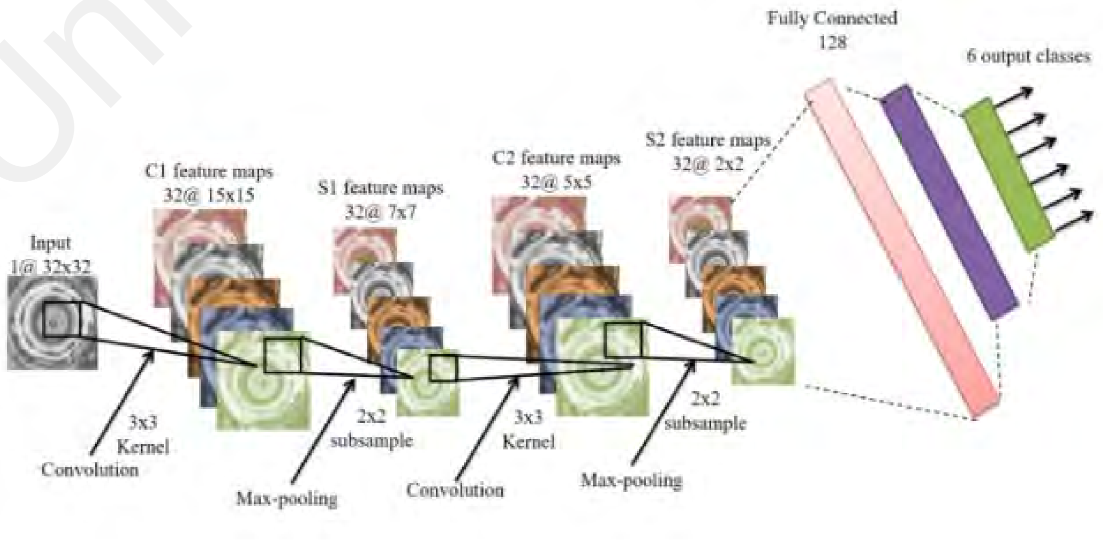


Figure 2.7: Basic Lenet-5 bearing defect detection neural network

He et al. proposed a framework that combines a convolutional autoencoder with an enhanced convolutional neural network to achieve high diagnostic accuracy of small-labeled infrared thermal images for rotor bearings (Zhiyi et al., 2020). Firstly, the infrared thermal image is measured to characterize various health states of the rotor bearing system. Secondly, an enhanced convolutional neural network is constructed by using exponential linear units and random pools. Then, the model parameters of CAE pre-trained with unlabeled thermal images are transmitted to initialize the enhanced convolutional neural network. Finally, the enhanced convolutional neural network is trained by using small label thermal image to further adjust the model parameters. The collected thermal images are used to test the diagnostic performance of the proposed method. Most of the past data-driven defect detection methods are based on the premise that the training data samples are large, average, and easily available.

However, in practical engineering applications, many defect samples under different conditions are very difficult to obtain, and deep learning models trained from samples in the laboratory usually do not work very well in practical engineering applications.

2.7 Sample imbalance algorithm

In industrial defect detection, it is often difficult to obtain the same amount of fault samples as healthy samples and the same characteristics as in real engineering, which leads to unsatisfactory results when the trained system is applied to complex conditions (G. Liu & Wu, 2024; Y. Liu & Lai, 2023; H. Wang & Zhang, 2024). Many solutions have been proposed for this problem (Mao et al., 2019; Ren et al., 2023).

Transfer learning was introduced in 1995 to solve the lifelong learning problem by reusing previous knowledge, which relaxed the restriction that training and test data must follow the same distribution. Transfer learning is an effective method to solve new problems by using knowledge learned from different but related tasks. Its idea is inspired by human learning mechanism, and it can reuse knowledge or skills summarized from tasks in other fields. For example, if a person learns to ride a bike, the skills of riding a bike will help him ride a motorcycle. The core idea of transfer learning is to find the common features of two or more related but different learning tasks and apply shared knowledge to solve new problems. In the actual bearing fault diagnosis scenario, the amount of normal data collected from bearings is usually far more than the amount of fault data, because laws and regulations usually prohibit long-term operation of actual machines with faulty bearings. This problem of insufficient fault data will significantly reduce the diagnostic performance of deep neural network]. Therefore, it is very important to introduce migration learning to solve the challenge of insufficient data. In fault diagnosis based on migration learning, the diagnosis model is pre-trained by using the source domain data and tasks, and then the learned knowledge (such as fault characteristics or model parameters) is reused in the target domain fault diagnosis tasks (X. Chen et al., 2023). Classical transfer learning-based fault diagnosis method is shown as Figure 2.8

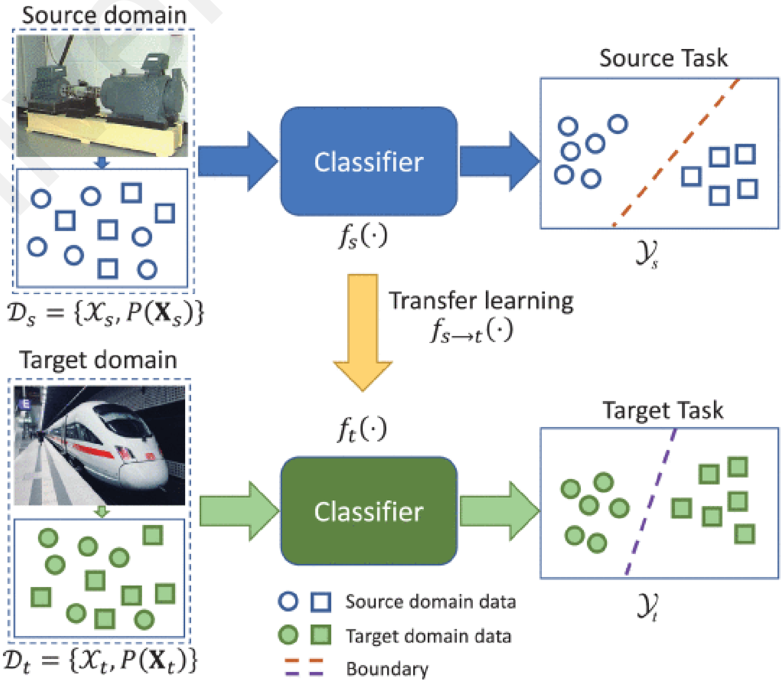


Figure 2.8: Illustration on transfer learning-based fault diagnosis.

Xiao et al. proposed a cross-domain bearing diagnosis framework based on transferable features and watershed embedding discriminative distribution adaptation to compensate for the imbalance between two real rolling bearing datasets (X. Yu et al., 2023). Zhu et al. designed a migration learning method based on the Maximum Mean Difference Multi-Kernel variant (MK-MMD), enabling pseudo label learning to resist data imbalance and improve prediction accuracy (W. Zhu et al., 2023). Ma et al. proposed two multi-label learning algorithms, personalized binary relevance (PBR) and hierarchical multi-label K-nearest neighbor (HML-KNN), for prognosis and health management (PHM) of rolling bearings, which exhibit fast modeling and high global information analysis capability in bearing fault analysis, respectively, and could determine the type of compound faults using only a single fault sample (X. Ma et al., 2021). However, most of the solutions for imbalanced data based on transfer learning or label learning are unable to explore the deeper differences between different features and have limited effects. In contrast, generative networks based on deep learning have the advantage of extending sample diversity and could amplify data in an unsupervised or semi-supervised mode (Y. Dong et al., 2024; Pan et al., 2024; Xia et al., 2022).

2.8 Generative algorithm

Classical generative models include Variational Autoencoders (VAEs) (Kingma & Welling, 2022) and Generative Adversarial Networks (GANs) (Goodfellow et al., 2014), which typically generate data of higher complexity and closer to the true distribution than VAEs (J. Liu, Zhang, et al., 2022). GAN consists of two networks, a generator, and a discriminator, which could cleverly compute the difference between the distribution of the generated data and the distribution of the original data, so that the generated data could learn the features of the original data. Since the creation of GAN, various variants of GAN have emerged. Mirza and Osindero proposed CGAN, which generates type-specific data by adding a conditional vector to the input noise vector (Mirza & Osindero, 2014). The goal of GAN training is to find the Nash equilibrium between the generator and the discriminator, but traditional GANs are often unstable and prone to pattern collapse due to the large oscillations between the generator and the discriminator during adversarial training. Moreover, the generator often generates a single output or non-ideal output to fool the discriminator, resulting in training failure. Radford et al. proposed DCGAN, which introduces convolutional neural networks to build generators and discriminators that could extract deeper features and generate higher quality images (Radford et al., 2016). To solve the problem that when using JS Divergence, the value of JS Divergence for two distributions without overlapping is always Log2, and it is impossible to calculate the relationship between the two distributions in this state, Arjovsky et al. introduced the Wasserstein GAN. The Wasserstein distance was introduced on top of the traditional GAN to replace the original KL divergence and JS divergence. The Wasserstein divergence is used to calculate the distance between the original distribution and the generated distribution by using the idea of minimizing the earth mover's distance, which makes the GAN more stable and could generate higher quality data (Arjovsky et al., 2017). Later, in response to the problem that supervised learning algorithms require a large amount of data from two distributions corresponding to each other, but the data in the situation are often unbalanced, J.-Y. Zhu et al. proposed a model called CycleGAN, which transforms the features in the domain of two different distributions by constructing two sets of generators and discriminators, and introduced a loss function called "cycle consistency loss" which separates the key features from the rest of the data and fixes one to enable the other to be transformed. This enables the style migration of images. Probably the best-known example is their unsupervised training and image transformation on two image datasets of horses and zebras (J.-Y. Zhu et al., 2020).

Recently, GAN has been gradually applied to solve the problem of engineering fault imbalance data. For example, Zhou et al. combined autoencoder (AE) and GAN to generate fault features for several different fault samples to improve the accuracy of fault diagnosis (F. Zhou et al., 2020). Chen et al. proposed a Joint Atrium and Scar Segmentations generative adversarial network (JAS-GAN) to segment unbalanced atrial targets from late gadolinium-enhanced cardiac magnetic resonance (LGE CMR) images in an end-to-end mode, producing better segmentation performance (J. Chen et al., 2022). For health monitoring of civil structures, Luleci et al. proposed a deep convolutional GAN incorporating Wasserstein loss (CycleWDCGAN-GP), which allows the undamaged and damaged acceleration data to be transformed between the two domains and could be used for the possible response of healthy structures to potentially damaged conditions, solving the problem of scarcity and imbalance of SHM data (Luleci et al., 2023).

Compared with the traditional fault diagnosis method for rolling bearings, GAN abandons the influence of human subjective factors on the results, and avoids the problem of poor applicability caused by the fact that fault data are only suitable for specific fault types. The generated large amount of data can make the fault diagnosis results based on data-driven algorithm closer to the actual working conditions. At the same time, different from the traditional neural network and deep learning methods, the antagonistic relationship between the discriminator and the generator eliminates the need for variational lower bound or approximate reasoning in the process of data generation, and avoids the calculation of partition function caused by repeated application of Markov chain learning mechanisms. GAN, as a new generation model, solves the problems of low efficiency and inaccurate data generated by traditional generation model, and at the same time, it has the idea of discriminating model, and greatly improves the accuracy and speed of generating data by using the confrontation between generator and discriminator. At the same time, the limitations of the traditional generation model on the dimension of generated samples and loss function have also been solved in this new generation model, which also gives GAN a very high degree of freedom in model design and greatly enhances the possibility of its practical application. In just six years after the advent of GAN, a large number of innovations and application methods have emerged, which also makes its application in the industrial field possible. Generally, the steps of generating pseudo-real data of rolling bearings by using GAN method are as follows: constructing and training to generate confrontation network model. The data generated by the countermeasure are compared with the real data respectively, and the similarity between the generated data and the real data is verified (Huo et al., 2022).

However, datasets generated using GAN often suffer from the problem of lack of realism and usability, and it is challenging to generate data with a natural distribution of fault features. There is still a lack of research on imbalanced data from infrared thermal images of bearings, and Cyclegan could transform features and generate images between thermal images of bearings with different types of faults, while maintaining the natural distribution of the images. However, further research and experimental validation is needed for how to generate high quality images using generative loss, adversarial loss, cycle consistency loss and identity loss to preserve and migrate target and non-target features in both domains while avoiding pattern collapse and gradient disappearance.

2.9 Convolutional neural network

The convolution layer uses convolution kernels to map and extract the features of the input signal, and the activation function enhances the learning ability of the network by adding nonlinear mapping. Down sampling is carried out in the pool layer to reduce the dimension of the feature map and the amount of calculated data; The full connection layer fuses and classifies the extracted features to reduce the loss of feature information; The output layer, also known as SoftMax classifier, transforms the output features into probability distribution for classification and discrimination. With the development of CNN, many classic models have emerged, such as LeNet, ResNet, AlexNet, DenseNet, Vgg, GoogleNet and so on. On the whole, there are three advantages of CNN. The first advantage is the local connection. Each neuron is no longer connected with all neurons in the previous layer, but only connected with a few neurons, which reduces the number of parameters. The second advantage is weight sharing, a group of connections can share the same weight, which also reduces the network parameters; The last advantage is down sampling, which can be used to reduce the number of samples in each layer, further reduce the number of parameters and improve the robustness of the model (X. Li, Ma, et al., 2024).

The bearing fault detection method based on Resnet classifier and model-based data enhancement can achieve good results. This method constructs a four-degree-of-freedom dynamic model to describe the bearing system. The dynamic model is identified by comparing the simulation and experimental results. Then, a large amount of data under different conditions can be generated, and a training data set can be constructed on this basis, and the Resnet classifier can be trained to classify the bearing state. In addition, in order to narrow the gap between the simulated data and the actual data, the envelope signal is used instead of the original signal in the training process (Qian et al., 2022). Classical ResNet block is shown as Figure 2.9.

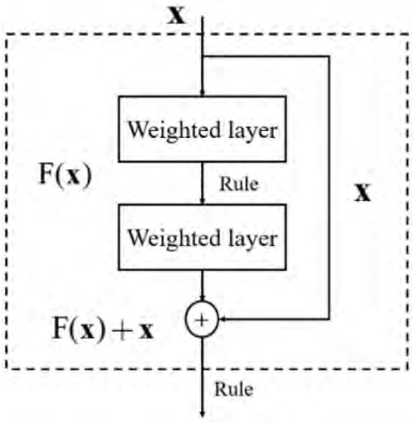


Figure 2.9: A residual learning unit.

CHAPTER 3: METHODOLOGY

3.1 Theoretical background

3.1.1 Generative Adversarial Networks

The basic structure of GAN is shown in Figure 3.1, where x represented the real sample and obeyed the prior distribution of P_r. z represented the random noise input into the network and obeyed the prior distribution of P_z. Generally, P_z often used Gaussian distribution and Uniform distribution. G represented the Generator, and the input z could be mapped into a fake sample $G(z)$ and obeyed the distribution P_g. Then $G(z)$ and x were input discriminator (D) to judge whether they conformed to the prior distribution P_r, achieving direct confrontation between G and D and reaching Nash equilibrium by constantly adjusting parameters.

The objective function of GAN was:

$$\min_G \max_D L(G,D) = E_{x\sim P_r}[logD(x)] + E_{z\sim P_z}\left[log\left(1 - D\big(G(z)\big)\right)\right] \qquad (3)$$

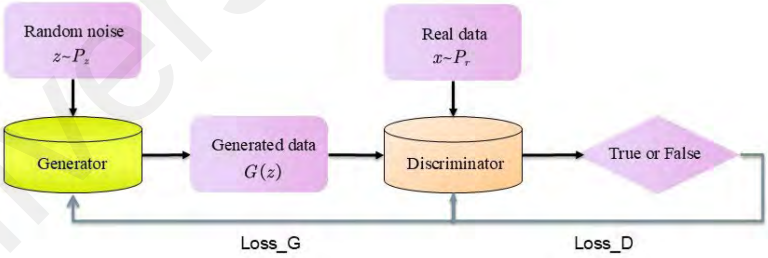


Figure 3.1: Basic structure of GAN

Here, the loss of G is $\log\left(1 - D\big(G(z)\big)\right)$, and the loss of D is $-\left(\log\big(D(x)\big)\right) + \log\left(1 - D\big(G(z)\big)\right)$. G hoped that the generated data would deceive D, meaning that $D\big(G(z)\big)$ was close to 1. D hoped to detect the data generated by G, meaning that $D\big(G(z)\big)$ was close to 0, so that adversarial training could be realized and P_g could be made as close as possible to P_r. In other words, the generated samples conformed to the real sample distribution as much as possible (Mirza & Osindero, 2014).

In the design of GAN, the input to the generator was random noise z, which was passed through a network to generate fake data. This design meant that GAN did not consider the specific features of the input data when generating samples, and therefore could not perform feature transformation for a specific domain. The goal of GAN was to make the generated fake data as similar as possible to the real data distribution P_r, rather than transforming or mapping the features of a specific input data. Since the generator G generated data based on noise and did not rely on the features of a specific input data sample, GAN could not guarantee that certain characteristics of the input data would be preserved during the generation process. As a result, GAN could not directly perform feature transformation of bearing image because its generation process was inherently random and did not learn or map features in a directed manner based on specific input images. This meant that while GAN could generate samples that resemble real data, the features of the generated samples would not align with the features of the input data, leading to the generated data not being usable as real data.

3.1.2 Cycle Consistent Generative Adversarial Network

In the field of image processing, the main purpose of CycleGAN was to exchange features between two images of different domains. It contained two generators, G_AB and G_BA, which could convert image A to B and image B to A respectively. It also contained two discriminators, D_A and D_B, which determined the authenticity of the generated and original images. The basic structure was shown in Figure 3.2.

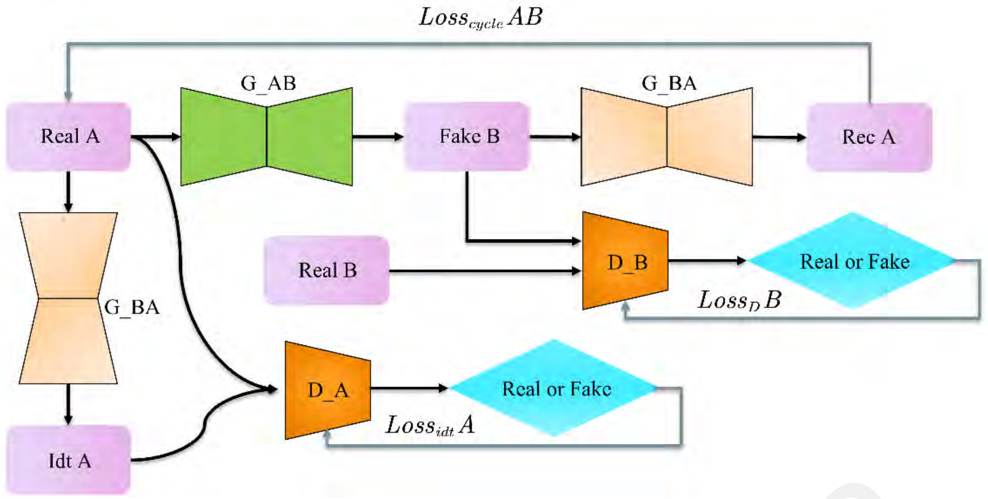


Figure 3.2: Basic structure of CycleGAN

The core idea was that Real A was desired to have not only the target features of the sample B distribution, but also the non-target features of the A distribution, except for the target features, after Real A had been transformed into Fake B by G_{AB}. To achieve this, Cycle Consistency Loss and Identity Loss were introduced.

Cycle Consistency Loss was achieved by feeding Fake B back into generator G_{BA}, producing an image Rec A that was as similar as possible to the original Real A, which could be described as:

$$G_{BA}\left(G_{AB}(A)\right) = A \tag{4}$$

However, it was difficult to keep the non-target features unchanged only by Cycle Consistency Loss, because G_{AB} might have learned the non-target features of non-input Real A, and the universal non-target features from other real pictures in the source domain could have deceiving D_A as well. To solve this problem, Identity Loss is introduced.

Identity Loss was realized by directly inputting Real A into generator G_{BA} to generate picture Idt A, and constraining Idt A to have only the characteristics of the input picture

Real A, without mixing the characteristics of other pictures from the source domain, which could be expressed as:

$$G_{BA}(A) = A \tag{5}$$

When training the discriminator, the parameters of G_{AB} and G_{BA} were fixed and only the parameters of D_A and D_B were adjustable. The target was then changed from Minimize Loss to Maximize Loss.

CycleGAN could be applied to bearing data generation by training a model that mapped laboratory data to real-world application data. Bearings were critical components in many mechanical systems, and gathering large amounts of real-world bearing data could be expensive, time-consuming, and sometimes impractical due to the difficulty in replicating certain operating conditions or failure scenarios. However, laboratory data, which was easier to collect and could cover a wide range of controlled conditions, was used as a source domain. CycleGAN could then have learned a mapping from this laboratory data (which might not perfectly represent real-world variations) to the real application data, effectively generating synthetic but realistic bearing data. This synthetic data could have been used to overcome the scarcity of real-world data, helping to avoid the challenges of collecting large datasets under realistic conditions. By training the CycleGAN model to transform laboratory data into realistic application data, the generator learned to produce realistic bearing data that mimicked the real-world distribution, even though the model had been trained on the laboratory data. This process enabled the creation of a large and diverse dataset of real-world-like bearing data, which could then be used to train models for fault detection, predictive maintenance, or other tasks. Additionally, by using this augmented dataset, a model trained on the generated data is likely to generalize better to real-world applications, since it could learn from a broader range of conditions and scenarios that were difficult to capture through traditional data collection methods. Therefore, CycleGAN helped bridge the gap between the lab and the field by generating realistic, diverse data that improved the performance and generalization of machine learning models in real-world bearing applications.

However, due to the complexity of the data and CycleGAN's original design not being for data augmentation, constructing a model that could transform target features without altering other background features requires further refinement. This involves ensuring that the model focuses on the desired features—like fault conditions—while preserving background elements such as operating conditions. Achieving this required additional constraints or modified loss functions to prevent changes to non-target features, along with fine-tuning Cycle Consistency Loss and Identity Loss to maintain the right balance. The main contribution of this research is to try to build such a model.

3.1.3 Interpolation Algorithm

In the past, it was found that the up-sampling process of GAN for image generation often used transpose convolution for image reconstruction, but this frequently led to a tessellation effect (uneven overlap) in the generated images. The cause of this phenomenon was often considered to be frequency artifacts since the kernel size of the convolution was not divisible by the stride (C. Dong et al., 2016). This problem could be effectively avoided by using interpolation methods, such as nearest-neighbor interpolation, bilinear interpolation and bicubic interpolation, of which bicubic interpolation was more computationally intensive but produced the best results.

Linear interpolation played a crucial role in various domains of image processing, medical imaging, geospatial data modeling, and computational interpolation techniques. In the domain of image demosaicing, traditional interpolation methods, such as bilinear interpolation, were widely used but suffered from fixed weight limitations, leading to suboptimal color reconstruction. The DTDeMo model introduces convolution interpolation blocks (CIBs) with trainable interpolation parameters to enhance the interpolation process while maintaining the original pixel arrangement, thus achieving superior demosaicing results (Hou et al., 2024). Similarly, in airway segmentation, a data-centric deep learning approach leverages big interpolated data to improve segmentation accuracy. The Interpolation-Split method utilized interpolation to refine the input data, ensuring high-quality feature extraction and boosting segmentation performance without excessive computational overhead (Cheung et al., 2024).

In the field of MRI reconstruction, interpolation was vital for refining k-space data and enhancing image quality. The Faster Fourier Convolution-based Single-to-Group Network (FAS-Net) incorporates k-space interpolation to improve spatial domain reconstruction, effectively mitigating full-image artifacts and optimizing 3D MRI reconstruction (X. Liu et al., 2024). In geospatial data processing, machine learning-based spatial interpolation techniques provide a robust alternative to traditional kriging methods. A machine-learning-driven block discretization technique enables efficient spatial interpolation, reducing computational complexity while maintaining high accuracy in mineral resource estimation and geospatial modeling (Nwaila et al., 2024). Furthermore, in digital volume correlation (DVC), cubic B-spline interpolation was commonly employed to achieve sub-voxel accuracy. The introduction of a direct cubic B-spline interpolation method provides a fuzzy, weightless alternative to traditional interpolation approaches, minimizing RAM usage while ensuring robust and precise displacement computations (D. Li et al., 2024).

The bicubic interpolation used the values of the 16 points around the point to be sampled to interpolate three times, considering the rate of change of the values of each neighbouring point in addition to the directly adjacent points. The formula was:

$$B(X,Y) = \sum_{i=0}^{3}\sum_{j=0}^{3} a_{ij} \times W(i) \times W(j) \tag{6}$$

Where $B(X,Y)$ was the target pixel, i and j were the rows and columns of the points to be sampled, and $W(x)$ is the Bicubic function:

$$W(x) = \begin{cases} (a+2)|x|^3 - (a+3)|x|^2 + 1 & for|x| \leqslant 1 \\ a|x|^3 - 5a|x|^2 + 8a|x| - 4a & for 1 < |x| < 2 \\ 0 & otherwise \end{cases} \tag{7}$$

Where a was the weight, often taken as -0.5

3.1.4 Depthwise Separable Convolution

CycleGAN contained multiple generators and discriminators, which often led to an increase of computing resources and time, especially in more complex networks. How to improve the training speed of the model without reducing the model effect was a problem that needs to be considered.

Depthwise Separable Convolution was a form of convolution with factorization and was probably best known for its application in Mobile Nets, proposed by Howard et al. in 2017. Compared to traditional 3D convolution, Depthwise Separable Convolution reduced computational effort by splitting the standard convolution into two steps: depth convolution and point convolution (Howard et al., 2017).

Depthwise Separable Convolution (DWSC) had been widely applied in various deep learning models to enhance computational efficiency while maintaining high performance across different domains. In medical image segmentation, DSML-UNet leverages depthwise separable convolution with multiscale large kernels to increase the receptive field while minimizing computational complexity, leading to improved segmentation accuracy in spine, skin, and lung datasets (B. Wang et al., 2024). Similarly, the DCDS-Net integrates DWSC with residual connections and densely connected blocks to enhance feature learning in the diagnosis of gastrointestinal diseases, demonstrating superior classification accuracy (Asif et al., 2024). The MDSU-Net model further incorporated multi-attention mechanisms alongside DWSC to improve feature fusion and reduce model complexity for medical image segmentation (Y. Zhou et al., 2024). In medical imaging reconstruction, DP-GAN+B utilizes DWSC in a generative adversarial network to efficiently generate high-quality CT volumes from X-ray images, significantly reducing the number of model parameters while preserving image fidelity (Xing et al., 2024). Beyond medical applications, depthwise separable convolution was also adopted in agriculture, as evidenced by MDSCIRNet, a novel deep learning model for potato leaf disease detection that integrates DWSC with Transformer networks, achieving state-of-the-art classification accuracy (Catal Reis & Turk, 2024). These studies collectively highlight the versatility and efficiency of DWSC in deep learning models across different domains.

The computation number C and the parametric number P of the standard 3D Convolution are:

$$C = D_k \times D_k \times M \times N \times D_F \times D_F \tag{8}$$

$$P = D_k \times D_k \times M \times N \tag{9}$$

where D_F×D_F×M was input feature maps, D_k×D_k×M is kernel size, D_F×D_F×N was output feature maps.

The computation C_{DS} and the parametric number P_{DS} of the Depthwise Separable Convolution were:

$$C_{DS} = C_{depth} + C_{point}$$

$$= D_k \times D_k \times M \times D_F \times D_F + M \times N \times D_F \times D_F \qquad (10)$$

$$P_{DS} = P_{depth} + P_{point} = D_k \times D_k \times M + M \times N \qquad (11)$$

where C_{depth} and P_{depth} is the computation and parametric numbers respectively.

Computational volume analysis of standard convolution with depthwise separable convolution:

$$\frac{C_{DS}}{C} = \frac{D_k \times D_k \times M \times D_F \times D_F + M \times N \times D_F \times D_F}{D_k \times D_k \times M \times N \times D_F \times D_F} = \frac{1}{N} + \frac{1}{D_K^2} \qquad (12)$$

This shows that with a 3x3 convolutional kernel, the parameters could be reduced by a factor of 8 to 9 with almost no loss of accuracy.

3.2 The proposed approach

The framework of the proposed unsupervised rotor bearing unbalance fault diagnosis method was shown in Figure 3.3.

The main steps were as follows:

Step 1: For the network master process, the real images Real A and Real B enter the generators G_AB and G_BA respectively to generate the target dummy images Fake B and Fake A. The generated images then entered the generators G_BA and G_AB to generate the reconstructed images Rec A and Rec B. Real A and RecA were used to calculate the loss cycle A, similar to Real B and Rec B. This loss helped the Generator maintain the non-target feature of the source domain data. Real A was directly input into Generator_BA to get idt A and calculate the loss idt A, similarly to Real B. This loss helped the Generator use the input data to generate images, rather than directly outputting without considering the input. As for the design of loss gan, it followed the same design as normal GAN. The generator was trained by driving the generated image closer to the real image, and the discriminator was trained by driving the generated image to be distinguishable from the real image.

Step 2: For the input of the new discriminator D_AB, each generated fake image and the corresponding real image are extracted into the image buffer with Query probability (total probability 1). Real A and Fake A composed the image buffer A, from which a random image was extracted and named New A. New B was generated by the same process. By classifying the data in buffer A and buffer B into corresponding categories, D_AB helped the generator generate images with target feature.

Step 3: For the loss function, Real A was compared with Fake A using the discriminator D_A and the loss value was calculated using BCE. The process for Image B was the same as for A. In addition, New A and New B entered the discriminator D_AB for binary classification, and the loss value was calculated using the cross-entropy function.

Step 4: All the losses were summed for gradient feedback and network parameter update, and finally the trained network was used to amplify the data and calculate the final accuracy and other metrics.

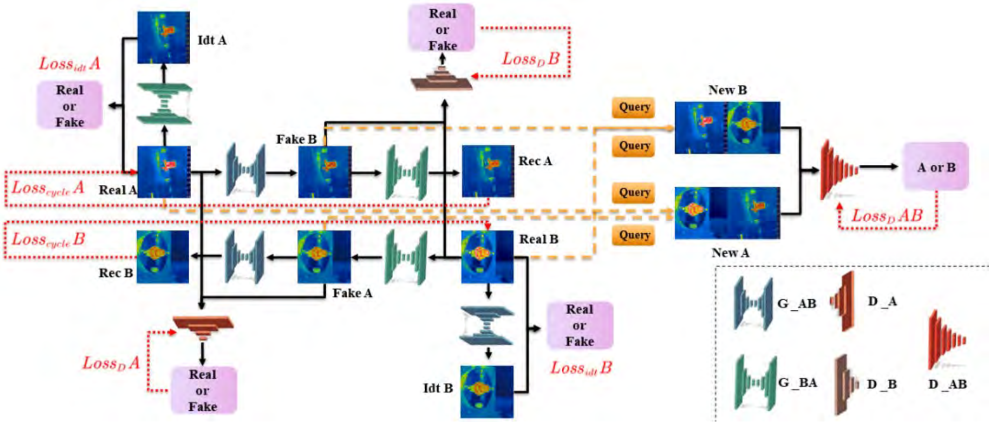


Figure 3.3: The framework of the FP-Cyclegan

3.2.1 Architecture of the generator.

With the development of Transformer frameworks based on self-attentive mechanisms, excellent vision backbone frameworks such as Vision Transformers (ViTs), Swin Transformers, etc. had emerged in the field of recognition vision and seem to be gradually replacing the old ConvNets (Dosovitskiy et al., 2021; Z. Liu et al., 2021).

However, Z. Liu et al. found that, with the same key components of the network, ConvNets still had advantages that could not be replaced by the self-attentive mechanism and re-examined the design space to propose a network framework called ConvNeXt to achieve results comparable to those of Transformer (Z. Liu et al., 2022). Inspired by ConvNeXt, a U-shaped network was designed based on the ConvNeXt module for the generators of Cyclegan networks.

ConvNeXt had been widely adopted in various domains due to its advanced convolutional architecture and performance improvements. In the field of computer vision, Yu et al. proposed InceptionNeXt, an extension of ConvNeXt that introduces an inception-style decomposition to large-kernel depthwise convolutions, achieving improved speed and accuracy for image classification tasks (W. Yu et al., 2024). In another study, Benchallal et al. developed a ConvNeXt-based semi-supervised learning approach for weed classification, effectively leveraging consistency regularization to improve classification performance with limited labeled data (Benchallal et al., 2024). In the aerospace sector, Yang et al. integrated ConvNeXt within an encoder-decoder framework to predict aerodynamic wall quantities for hypersonic vehicles, demonstrating superior accuracy compared to traditional computational fluid dynamics (CFD) models (Y. Yang et al., 2024). Moreover, Ramos et al. explored the application of ConvNeXt for enhanced image captioning, showing that it outperformed existing CNN-based and transformer-based models in generating accurate textual descriptions of images (Ramos et al., 2024). In medical imaging, Gulsoy et al. introduced FocalNeXt, a ConvNeXt-augmented FocalNet architecture designed for lung couldcer classification from CT scoulds, achieving state-of-the-art accuracy and robustness in couldcer detection tasks (Gulsoy & Baykal Kablan, 2025). Additionally, in image generation, Verma et al**.** incorporated ConvNeXt-V2 into a denoising diffusion probabilistic model, significouldtly improving image diversity and quality by leveraging ConvNeXt's robust feature extraction capabilities (Verma et al., 2024). These studies highlighted ConvNeXt's adaptability and efficiency across multiple disciplines, ranging from computer vision and healthcare to aerospace and generative modeling.

The architecture of the ConvNeXt Block was shown in Figure 3.4.

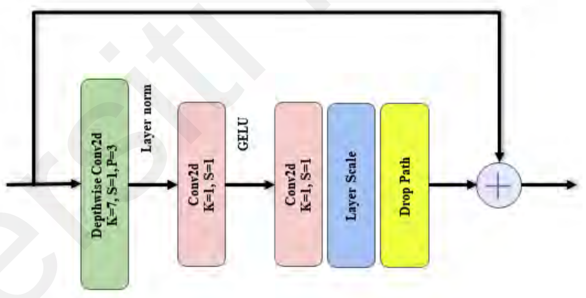


Figure 3.4: The architecture of the ConvNeXt Block

The input was first passed through a depthwise separable convolutional layer with a kernel size of 7, the step size of 2, and the padding of 3, and then through two 1x1 convolutional layers. Before the output, a scaling factor called layer scale was applied to adjust the expressiveness of the layer output and a drop path is introduced to randomly discard network paths to prevent overfitting and improve generalization.

The architecture of the generator is shown in Figure 3.5.

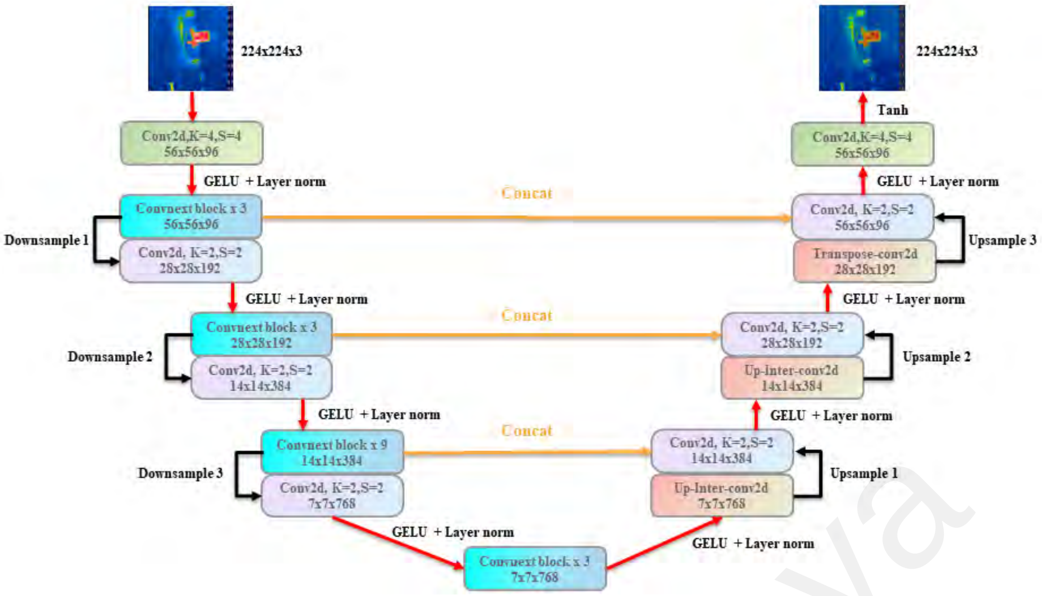


Figure 3.5: The architecture of the Generator

First, the pre-processing process scaleed and center-cropped the bearing thermal image to the size of 224x224x3 as input.

For the down-sampling process, the image was first projected into a 56x56x96 feature map by a convolutional layer with a kernel size and stride of 4. The feature map was then downsampled three times, with each downsample module consisting of multiple ConvNeXt Blocks and a Conv2d layer. The final feature map of size 7x7x768 entered a bottleneck layer consisting of three ConvNeXt Blocks. The same number of blocks was maintained as in ConvNeXt-T [3,3,9,3].

For the up-sampling process, the output of the bottleneck layer was passed through three up sample modules, each consisting of a linear Interpolation or Transpose Covn2d layer and a Conv2d layer. Network Layer Fusion was also introduced, where the corresponding feature map sizes from the down-sampling and up-sampling processes were concatenated in the channel dimension to improve the expressiveness of the model and the perception of different levels of features in the image. The final feature map was reconstructed back to an input size of 224x224x3 by a Transpose Covn2d layer with the kernel size and step size of 4 and output by Tanh as the activation function.

In recent studies, the concat mechanism had been widely used to fuse features from different layers of neural networks, improving the performance in various image and scene classification tasks. For instance, the study by Liu et al. integrates deep feature fusion using concatenation techniques to combine multi-scale features, achieving enhanced results in plant leaf recognition (Hu et al., 2018). Similarly, MCFNet employs a multi-layer concatenation fusion network to improve medical image fusion, demonstrating that concatenation-based feature aggregation enhances classification accuracy (Liang et al., 2019). In another work, the authors leverage concatenation to fuse features from different convolutional layers for scene classification in remote sensing applications, enabling a more detailed representation of input data (C. Ma et al., 2019). Additionally, the two-stage deep feature fusion model proposed by Liu et al. utilized concatenation to combine information from both intermediate and fully connected layers to improve the model's performance in remote sensing classification (Y. Liu et al., 2018). Finally, a multiscale convolutional neural network also applies feature concatenation to effectively merge global and local features for scene classification, further enhancing the accuracy of the model (J. Liu, Sun, et al., 2022).

Throughout the generator process, a smoother GELU and a more feature-focused LayerNorm were used as the activation function and normalization method after each layer of the network to improve the model's generalizability and training speed.

In deep learning, both GELU and LayerNorm had become critical components in enhancing model performance and efficiency (Hendrycks & Gimpel, 2023; Kiliçarslan & Celik, 2024; Lee, 2023). The GELU activation function had been extensively analyzed for its smoothness, differentiability, and superior performance over traditional activation functions like ReLU. In particular, GELU had been found to improve training stability and computational efficiency in various deep neural network architectures. On the other hand, LayerNorm was crucial for stabilizing training, especially in large pre-trained models such as BERT, by normalizing input distributions across layers (ValizadehAslani & Liang, 2024; J. Xu et al., n.d.). For example, Zhang et al. introduce an optimized version, Root Mean Square Layer Normalization (RMSNorm), which reduced computational overhead while maintaining the benefits of traditional LayerNorm (B. Zhang & Sennrich, n.d.). Together, GELU and LayerNorm (and its variants) had become fundamental in modern deep learning for improving both performance and efficiency.

3.2.2 Architecture of the discriminator.

The structure of the new discriminator D_AB was shown in Figure 3.6.

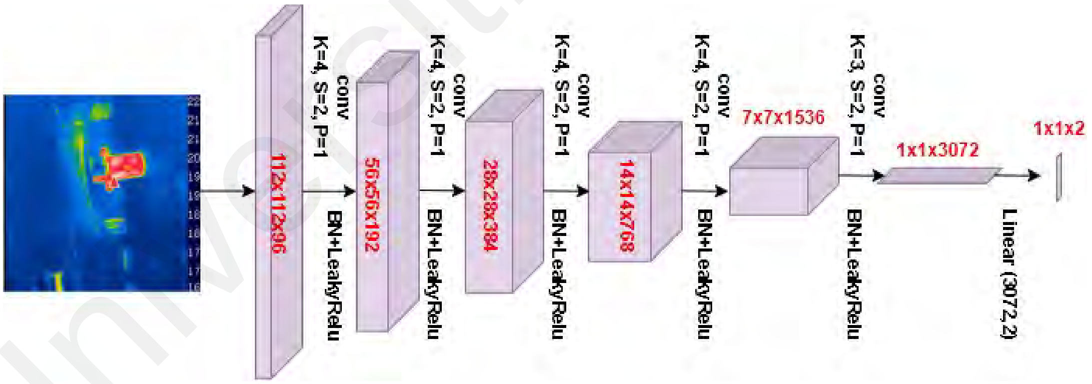


Figure 3.6: The architecture of the discriminator D_AB

The input to D_AB was the same as D_A and D_B, both pre-processed 224x224x3 images of the centre of the rotor bearing thermal image. The difference was that the new D_AB finally compresses the feature map to 1x1x3072 and outputs it as a 4x1x1x2 matrix with a fully-connected layer, where 4 was the batch size and 2 was the target score for binary classification. The reason for this was to use CrossEntropyLoss instead of

MSELoss to avoid the problem of non-decreasing gradients when the output was close to 0 or 1.

3.2.3 Loss function

The loss function of the whole model was:

$$Loss_{FP-Cyclegan} = Loss_{gan} + Loss_{cycle} + \lambda Loss_{idt} + Loss_{class} \tag{13}$$

where $Loss_{gan}$ was the classical GAN network adversarial loss. For the FP-Cyclegan network, there are two generators and two discriminators, so its expression was:

$$\begin{aligned} Loss_{gan} &= \mathcal{L}_{gan}(G_{AB}, D_B, A, B) + \mathcal{L}_{gan}(G_{BA}, D_A, A, B) \\ &= \mathbb{E}_{a \sim P_{data}(a)}[\log D_A(a)] + \mathbb{E}_{b \sim P_{data}(b)}\left[\log\left(1 - D_A\left(G_{BA}(b)\right)\right)\right] \\ &+ \mathbb{E}_{b \sim P_{data}(b)}[\log D_B(b)] + \mathbb{E}_{a \sim P_{data}(a)}\left[\log\left(1 - D_B\left(G_{AB}(a)\right)\right)\right] \end{aligned} \tag{14}$$

The goal of the generator was to bring the defective features of the generated image closer to the distribution of defective features in the target domain, and the goal of the discriminator was to be able to identify whether the image contains defective features in the source domain or in the target domain.

The expression of the optimization objective in a mathematical formula was:

$$G_{AB}^*, G_{BA}^*, D_A^*, D_B^* = arg \min_{G_{AB}^*, G_{BA}^*} \max_{D_A, D_B} \mathcal{L}(G_{AB}, G_{BA}, D_A, D_B) \tag{15}$$

In addition, MSE Loss was used to calculate the loss of both, with the generated image label set to 0 and the original image label set to 1. The mathematical formula was:

$$\min_{D_B} \frac{1}{N} \sum_{i=1}^{N} MSE(D_B(b), 1) \tag{16}$$

$$\min_{D_A} \frac{1}{N} \sum_{i=1}^{N} MSE(D_A(a), 1) \tag{17}$$

$$\min_{G_{AB}} \frac{1}{N} \sum_{i=1}^{N} MSE\big(D_B\big(G_{AB}(a)\big), 0\big) \tag{18}$$

$$\min_{G_{BA}} \frac{1}{N} \sum_{i=1}^{N} MSE\big(D_A\big(G_{BA}(b)\big), 0\big) \tag{19}$$

The expression for $Loss_{cycle}$ was:

$$\begin{aligned} Loss_{cycle} = {} & \mathbb{E}_{a \sim P_{data}(a)}\big[\big\|G_{BA}\big(G_{AB}(a)\big) - a\big\|\big] \\ & + \mathbb{E}_{b \sim P_{data}(b)}\big[\big\|G_{AB}\big(G_{BA}(b)\big) - b\big\|\big] \end{aligned} \tag{20}$$

$Loss_{cycle}$ was implemented as L1 Loss, which, when minimizing $Loss_{cycle}$, was equivalent to making the input image as consistent as possible with the image reconstructed by the two generators. This enables the generated image to acquire the fault features in the target domain while keeping the other non-target features in the source domain as constant as possible.

The expression for $Loss_{idt}$ was:

$$Loss_{idt} = \mathbb{E}_{a \sim P_{data}(a)}[\|G_{BA}(a) - a\|] + \mathbb{E}_{b \sim P_{data}(b)}[\|G_{AB}(b) - b\|] \tag{21}$$

$Loss_{idt}$ was implemented as L1 Loss. When minimizing $Loss_{idt}$, the generated images tried to keep the non-target features of that original input image while obtaining the defective features of the target domain, instead of the non-target features of other images in the source domain.

The expression for $Loss_{class}$ was:

$$\begin{aligned} Loss_{class} &= \mathcal{L}_{class}(G_{AB}, D_{AB}, A, B) + \mathcal{L}_{class}(G_{BA}, D_{AB}, A, B) \\ &= \mathbb{E}_{newa \sim P_{data}(newa)}[\log D_{AB}(newa)] \\ &+ \mathbb{E}_{newb \sim P_{data}(newb)}\left[\log\left(1 - D_{AB}\left(G_{BA}(newb)\right)\right)\right] \\ &+ \mathbb{E}_{newb \sim P_{data}(newb)}[\log D_{AB}(newb)] \\ &+ \mathbb{E}_{newa \sim P_{data}(newa)}\left[\log\left(1 - D_{AB}\left(G_{AB}(newa)\right)\right)\right] \end{aligned} \tag{22}$$

where $P_{data}(newa)$ was the joint distribution of the original image A distribution and the generated image A distribution, and $P_{data}(newb)$ was the same.

$$P_{data}(newa) = \mu P_{data}(a) + (1 - \mu) P_{data}[G_{BA}(b)] \quad b \sim P_{data}(b) \tag{23}$$

$$P_{data}(newb) = \mu P_{data}(b) + (1 - \mu) P_{data}[G_{AB}(a)] \quad a \sim P_{data}(a) \tag{24}$$

where μ was the ratio of the original images to the generated images in the new joint distribution, used to control the degree of transfer of the target features.

In addition, $Loss_{class}$ was implemented as Cross Entropy Loss instead of MSE Loss because the objective of the new discriminator D_AB could be seen as a logistic regression problem to classify images from $newa$ and $newa$ as binary categories. MSE Loss was more suitable for linear regression problems.

In summary, if the features contained in the original input images A and B were represented by F_A and F_B. Use $F_{G_{BA}(B)}$ and $F_{G_{AB}(A)}$ to represent the features contained in the generated images Fake A and Fake B. with PF_A and PF_B representing the defective features of images A and B. with NF_A and NF_B representing the non-defective features of images A and B. Then the input image features were:

$$F_A = (PF_A, NF_A) \tag{25}$$

$$F_B = (PF_B, NF_B) \tag{26}$$

The goal of FP-Cyclegan's output image features was:

$$F_{G_{BA}(B)} = (PF_B, NF_A) \tag{27}$$

$$F_{G_{AB}(A)} = (PF_A, NF_B) \tag{28}$$

3.3 Summary

In summary, this paper proposes a fault diagnosis framework called Feature Preserving CycleGAN. The core of the framework is the introduction of a new discriminator D_AB based on CycleGAN to constrain the non-target features in the source domain, which in turn migrates the target features in the source domain and reconstructs the generator structure. It solves the problem of data imbalance in the infrared thermal images of faulty rotor bearings and enhances the stability of the network and the realism of the generated images during the adversarial generation process. The main contributions of this paper are as follows:

A new unsupervised fault diagnosis method called FP-CycleGAN is proposed. A new discriminator D_AB is introduced to efficiently migrate the target features while maintaining the non-target features during the transformation of the bearing thermal image from the source domain A to the target domain B. The total Loss function is reconstructed, and a new Loss-class is added.

The generator is reconstructed to provide better feature extraction capability and data processing speed. It is possible to reduce the higher data volume caused by D_AB and the difficulty of bearing thermal image reconstruction in the new framework, which improves the operation speed of the network and the quality of the generated images.

Using the rotor bearing infrared thermal image fault simulation experiment bench to extract data and conduct extensive experiments, to verify the effectiveness of the proposed data amplification and fault diagnosis method.

CHAPTER 4: RESULTS

4.1 Introduction

In this chapter, the experimental results of image amplification and damage detection by the proposed method are given. At first, the design and operation details of collecting experimental data of bearing thermal image are introduced. Then, the ablation analysis of different modules of the proposed model is given and the numerical results of evaluation are given. The data sets with different data amplification methods are ablated and the index results are analyzed.

4.2 Collection and description of laboratory bearing data sets.

The faulty rotor bearing specimens used in the text are seven kinds of bearings with different failure points and degrees of damage, with a dimensional size of 650 mm outer ring radius and GCr15 material, and health condition names with corresponding labels as shown in Table 4.1.

Table 4.1: Health conditions and labels of rolling bearings

Health Condition	Amount (Multi dimension)	Amount (single dimension)	Label
F-N-0.6	300	400	Label 1
F-N-0.8	896+300	400	Label 2
F-N-1.0	300	400	Label 3
F-Q-0.8	300	400	Label 4
F-W-0.6	300	400	Label 5
F-W-0.8	300	400	Label 6
F-W-1.0	300	400	Label 7

The label N indicates damage to the inner ring, Q indicates damage to the sphere, W indicates damage to the outer ring and the number indicates the degree of damage.

The infrared thermal image acquisition process of this experiment was carried out at the State Key Laboratory of Mechanical Behavior and System Safety of Traffic Engineering Structures, Shijiazhuang Tiedao University. The experimental data came from the rotating machinery test bench shown in Figure 4.1, mainly consisting of a driven motor, rotor shaft, bearing under test and drive belt. The infrared thermal imaging camera used is A315 made by FLIR, USA, which could collect the temperature field information of the faulty bearing at various points in the rotation, with a thermal sensitivity less than or equal to 0.05°C and a maximum acquisition temperature of 1200°C. The resolution was set to 370×240 pixels when shooting the camera is initially installed at 40cm from the fixed rotor shaft, and the position will change as the experiment progresses.

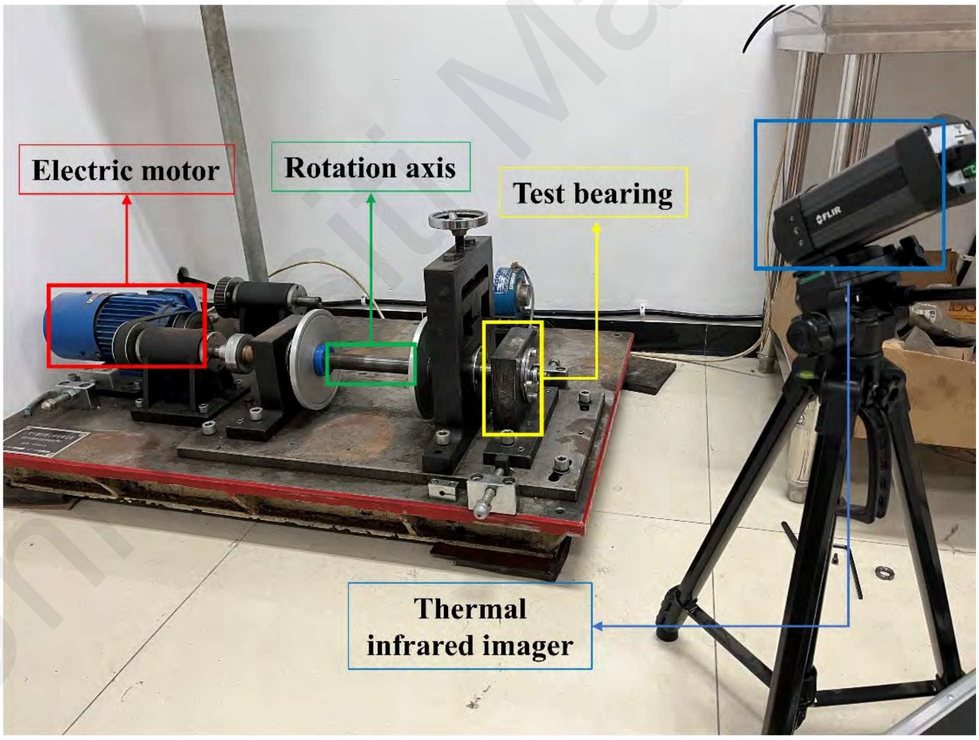
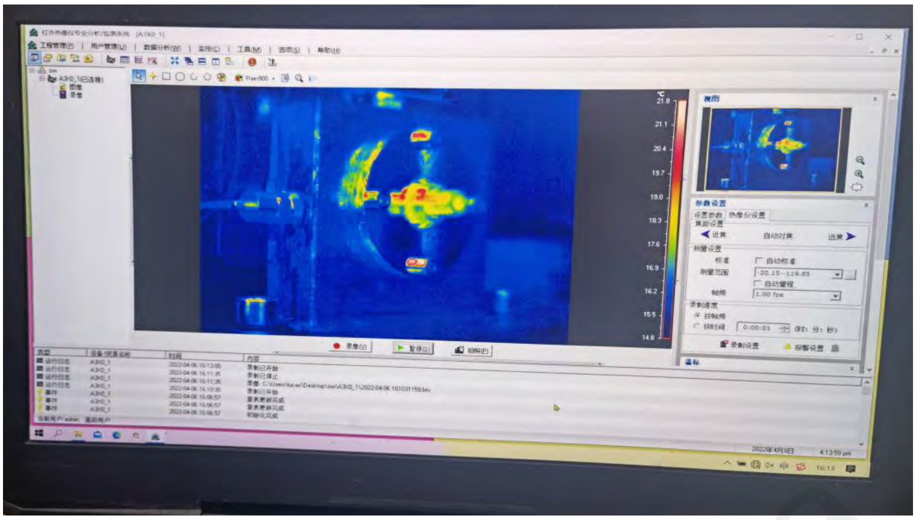


Figure 4.1: The rotating machinery test bench

The shooting software uses the professional analysis/detection system (BM_IR) of infrared thermal imager matched with A315. The interface is shown in Figure 4.2, the measuring range is -20.15-119.85, and the shooting frame rate is set to 1fps.



校准 ☐ 自动校准

测量范围 -20.15--119.85

☐ 自动量程

帧频 1.00 fps

● 按帧频

○ 按时间 0:00:01 (时: 分: 秒)

Figure 4.2: Thermal image acquisition system

In the first stage of the experiment, thermal imaging images were taken at five speeds (1,000, 2,000, 3,000, 4,000 and 5,000 rev/min) for each healthy condition of the bearing at a constant distance, and 80 images were taken for each speed condition, i.e., 400 experimental images for each bearing, for a total of 2,800 images. After the shooting, BTV files were obtained, and the files were transcoded to obtain a dataset of 2,800 images in the rain900 style, as shown in Figure 4.3. They will be used as the data for a single working condition in the laboratory to participate in the training of defect detection and data generation. In the proposed FP-CycleGAN model, the data of each kind of defects in this part will be used as the data in Real A to participate in the training.

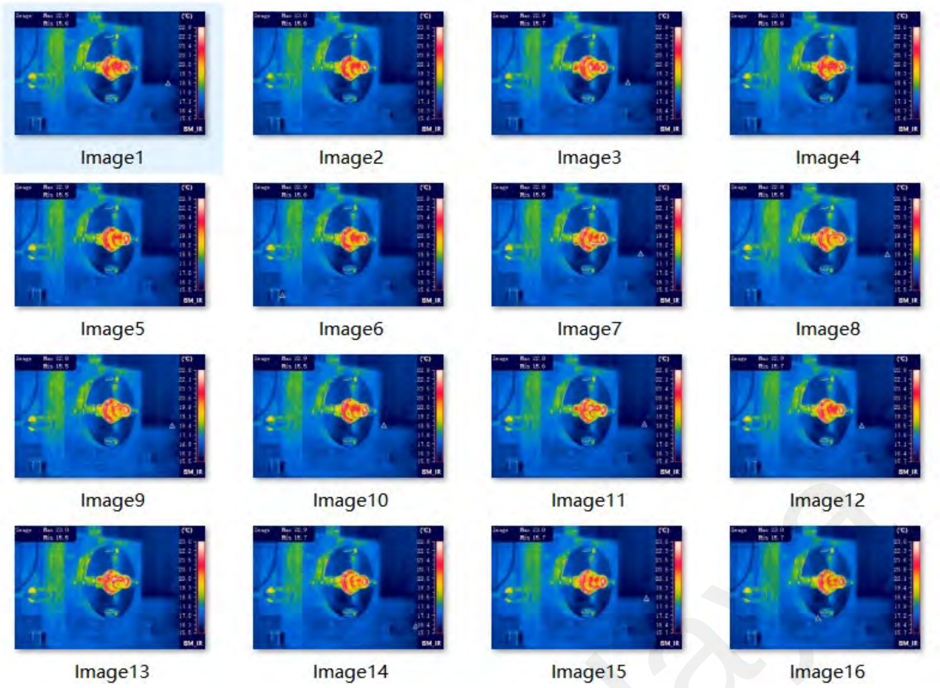


Image1 Image2 Image3 Image4

Image5 Image6 Image7 Image8

Image9 Image10 Image11 Image12

Image13 Image14 Image15 Image16

Figure 4.3: Single angle and state bearing rain900 data set

In the second phase of the experiment, 896 images of a bearing with health condition F-N-0.8 at a random speed were collected at a random distance and angle between the thermal imaging camera and the faulty bearing for data amplification. They will be used as data to simulate the complex working conditions of reality and participate in the training of data generation. In the proposed FP-CycleGAN model, this part of data will be used as Real B to guide the data in Real A to generate Fake A data under multiple working conditions.

400 images of each of the seven healthy conditions of the bearing at random speeds were acquired at random distances and angles between the thermal imaging camera and the faulty bearing. This was used to construct a test set to simulate the data diversity in a real application. It was also converted to the rain900 style as shown in Figure 4.4.

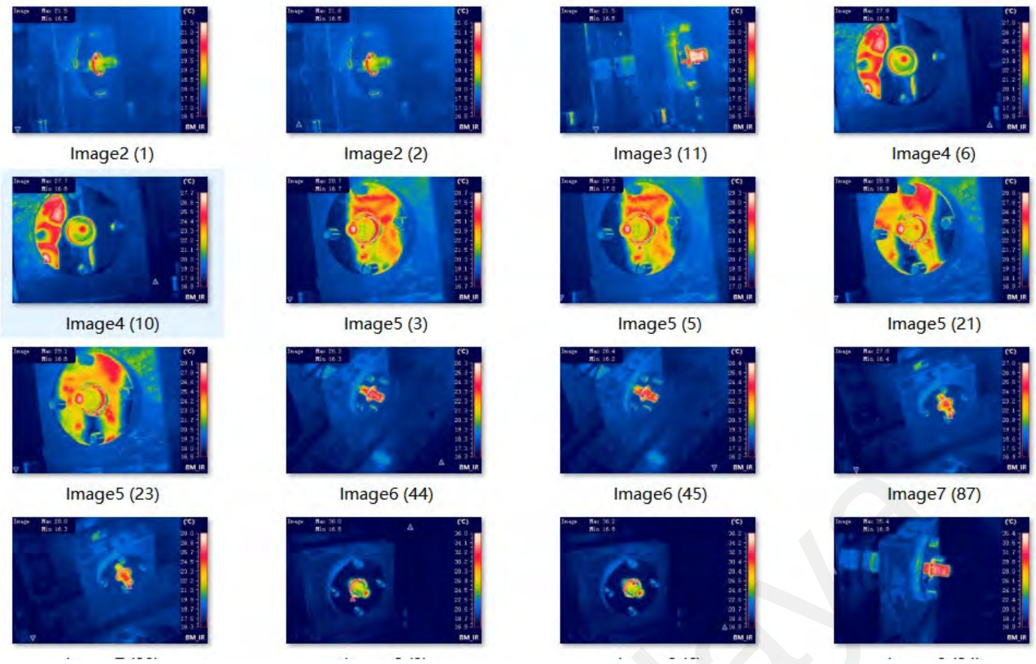


Figure 4.4: Multi angle and state bearing rain900 data set

4.3 Sample expansion performance evaluation

The experimental algorithms in this paper were designed using Python 3.7.3 and Pytorch Stable (2.0.1). The experimental platform consists of an Intel Core i5-12400 CPU, NVIDIA RTX 3090 GPU and 32G RAM.

The hyperparameters of the FP-CycleGAN framework are set as shown in Table 4.2.

Table 4.2: Hyperparameter setting.

Hyperparameter	Value
Adam Learning Rate	0.0001
Adam momentum term	0.5
Batch size	4
Number of epochs	300
Number of epochs with decay	300
Load image size	260
Crop image size	224
Q	0.9
λ	10

where Q is the proportion of real images in the image buffer used for discriminator D_AB training, and λ is the weight of $Loss_{cycle}$. These hyperparameters were determined experimentally.

The experimental scheme for the sample expansion was to use 896 F-N-0.8 bearing thermal images with multi-state information as domain A, and 400 of one of the other six single-state bearing thermal images as domain B, converting them in turn. The aim is to transform the defective features of the thermal image of domain B with those of domain A to obtain the 896 generated multi-state images of domain B, thus enabling the expansion of the imbalance data.

To fully demonstrate the superiority of the proposed approach, it is necessary to investigate in turn the impact of the different modules of the FP-CycleGAN system on the system. A variety of evaluation metrics are used to quantify the different performances of the proposed model. The Fréchet Inception Distance (FID) value is a metric used to quantify the dissimilarity between images generated by a Generative Adversarial Network (GAN) and real images (Bynagari, 2019). It measures the distance between the distributions of feature representations extracted from a pre-trained neural network, typically an Inception network. A lower FID score indicates that the generated images are closer to the real images in terms of their high-level features, implying better performance of the GAN in generating realistic images.

However, this metric could only be used to evaluate whether the generated images match the target domain distribution but could not measure whether the migration of the target feature distribution could be performed while preserving to the maximum extent the non-target feature distribution in the source domain. So, in the experiment, it is necessary to strive for both lowering the FID value and improving the accuracy of the final classification network.

The time metric Time is used to evaluate the amount of data and the speed at which the model could be run. Models with different up-sampling methods and convolution methods are combined with the proposed best models for comparison. All methods are tested in ten runs to overcome particularity and contingency. All tables with an asterisk (*) represent the optimal approach in the proposed FP-CycleGAN framework. Table 4.3 shows the operation time and FID values for the FP-CycleGAN with different up-sampling methods.

The comparison shows that the interpolation algorithm could improve the quality of the images and avoid the tessellation effect as much as possible. Bicubic interpolation gives the best results for the generated images, although it takes a little longer, but the quality is improved more. The interpolation algorithm works better in the first half of the up-sampling (step1 and step2), whereas it reduces the quality of the resulting image in the latter half.

Table 4.3: Results of different up-sampling methods.

Up-simple method	Time (s/epoch)	FID score
Transposed Convolution	97.75	65.7604
Nearest neighbour interpolation (Step 1,2)	91.57	62.7155
Bilinear interpolation (Step 1,2)	98.14	62.5304
Bicubic interpolation* (Step 1,2)	103.95	59.7446
Bicubic interpolation (Step 1,2,3)	105.45	60.5247
Bicubic interpolation (Step 1,2,3,4)	120.50	85.0308
Bicubic interpolation (Step 3,4)	102.60	94.3045

Figure 4.5 shows the images generated by the different up-sampling methods.

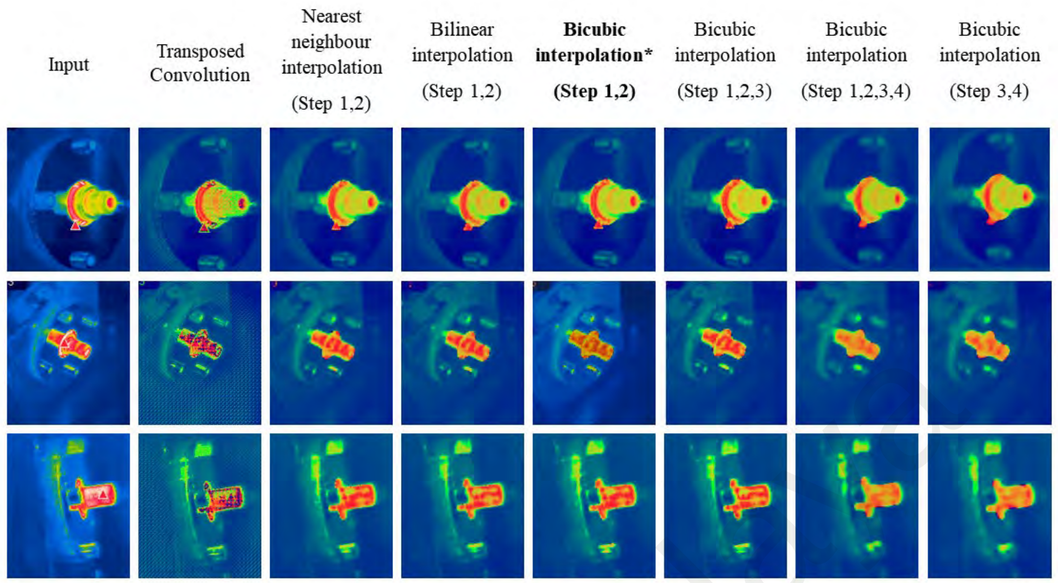


Figure 4.5: Images generated by different up-sampling methods.

Bicubic interpolation outperforms other up-sampling methods in the early stages (Step 1 and Step 2) of the FP-CycleGAN process may be due to its ability to generate smoother and higher-quality images. This method considers more surrounding pixels compared to simpler techniques like bilinear or nearest-neighbor interpolation, resulting in a more refined and detailed image that avoids the tessellation effect (the blocky or pixelated artifacts that could appear during image up-sampling). The improved image quality is particularly important in thermal imaging, where subtle differences in temperature need to be preserved. However, Bicubic interpolation’s advantage diminishes in the later stages (Step 3 and beyond), where it may lead to blurring or distortion, negatively impacting the image quality. Thus, while it requires more time, Bicubic interpolation's superior quality in the initial steps makes it the best choice for improving image clarity and detail in FP-CycleGAN's image generation process.

Table 4.4 shows the operation time and FID values for different convolution methods.

The results show that the computation time of the generator structure with depth wise separable convolution is significouldtly less than that with normal convolution, and the generated images are slightly better than the normal convolution structure.

Depthwise Convolution demonstrates a clear advantage over the original convolution method. This improvement is primarily due to its computational efficiency and ability to preserve important image features. Depthwise Convolution reduces the number of parameters by applying a separate filter to each individual input channel rather than using a single filter across all channels, as in the traditional convolution method. This leads to faster processing times and a reduction in computational complexity, which is crucial when working with large datasets or performing multiple iterations in the FP-CycleGAN. Despite the reduced complexity, Depthwise Convolution maintains or even enhances the quality of the generated images, as evidenced by the improved Fréchet Inception Distance (FID) values. This makes Depthwise Convolution not only more efficient but also more effective in preserving the quality of thermal images used for fault diagnosis of rolling bearings, thereby providing a better balance between computational cost and diagnostic accuracy.

Table 4.4: Results of different convolution methods.

Up-simple method	Time (s/epoch)	FID score
Deep wise*	103.95	59.7446
Original	114.6	60.8022

4.4 Unbalance fault diagnosis results and analysis

After determining the optimal FP-CycleGAN data augmentation framework, further comparative experiments for the diagnosis of thermal images of unbalanced faulty bearings are required to demonstrate that the generated images successfully perform the target defect feature transfer and improve the generalizability of the data.

The specific experimental scheme is:

Construction of a data-free augmented training set: 400 single-angle and 15 multi-angle images for each bearing defect category except for the label F-N-0.8. The category labelled F-N-0.8 contains 300 plus 400 multi-angle images. In this case, the trained neural network often has low generalization in the test set, because the training lacks data with sufficient generalization. This simulates the problem of often unbalanced data in practical engineering applications, where a small amount of data or a single piece of data often makes it difficult to obtain good detection results.

Construction of the training set after data amplification: each bearing defect category, except for the label F-N-0.8, contains 400 single-angle and 15 multi-angle real images, as well as 896 generated multi-angle dummy images. The category labelled F-N-0.8 contains 300 plus 400 multi-angle real images.

Construction of the test set: Each category in the test set contains 400 images of random rotation speeds and angles to simulate the diversity of data in real applications.

Besides the proposed method FP-CycleGAN, the classical CycleGAN with ResNet block and UNet as backbones and the FP-CycleGAN without D_AB were also used for comparison experiments, thus proving the superiority of the proposed method.

After amplification of the data by different methods, ResNet18 was used uniformly for classification training and the results were compared on the test set, ten times for each training and the average was taken.

Figure 4.6 presents a detailed comparison of the results obtained from various data amplification methods applied to thermal images of unbalanced faulty bearings, aiming to demonstrate the efficacy of the proposed FP-CycleGAN model. The comparison includes the classical CycleGAN with two different backbones—ResNet and UNet—and the FP-CycleGAN method without the discriminator component D_AB. The purpose of this comparison is to assess the performance of these methods in transferring the defect features effectively and improving the model’s generalization capabilities.

The image is structured as follows:

First row (Input): Displays the raw thermal images of the bearings, which serve as the initial data before amplification.

Second row (CycleGAN-Unet): Shows the results from the CycleGAN method using a UNet backbone for data amplification. The output images here reflect how this model handles feature transfer but may exhibit less accuracy in maintaining image realism compared to the FP-CycleGAN method.

Third row (CycleGAN-Resnet): Displays the results from the CycleGAN method using a ResNet backbone. This row highlights the effects of using ResNet as a backbone in amplification the images, with a focus on how it influences the feature transfer and realism of the generated images.

Fourth row (FP-CycleGAN without D_AB): Presents the output of the FP-CycleGAN method without the D_AB component. This is an attempt to assess the impact of omitting the D_AB part on the quality of the generated images, showing that the exclusion may reduce the model's ability to maintain high-quality image generation.

Fifth row (FP-CycleGAN): Shows the results of the proposed FP-CycleGAN approach. As the primary method under evaluation, FP-CycleGAN outperforms the other models, addressing issues such as pattern collapse and ensuring the generated images retain more realism and better defect feature transfer.

Each image in the columns represents the results for a different set of labels and amplification parameters, such as F-N-0.6, F-N-1.0, F-Q-0.8, and F-W with varying values (e.g., 0.6, 0.8, 1.0). These variations simulate real-world conditions where the fault characteristics of bearings may vary due to different rotational speeds and angles.

The key takeaway from this comparison is that FP-CycleGAN significouldtly improves upon previous methods, effectively mitigating the problem of pattern collapse while preserving the realism of the thermal images. This is crucial in practical applications where accurate fault diagnosis requires high-quality images that could generalize well across different fault conditions and variations. The experimental results support the superiority of the FP-CycleGAN framework in handling the challenges posed by unbalanced data and ensuring reliable performance in bearing fault detection.

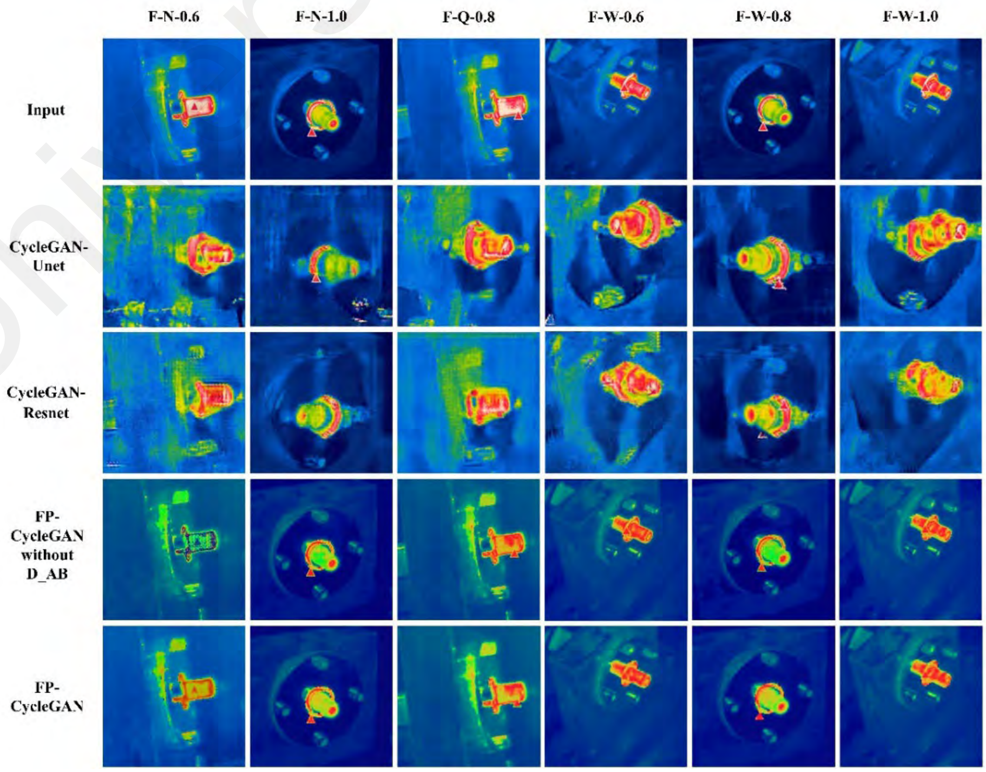


Figure 4.6: Images generated by different data amplification methods.

Compared with the previous methods, the proposed method has been improved in precision, F1 score and recall. We also use the mean precision, top-1 accuracy, mean f1 score to show the different result.

Table 4.5 shows the results of the classification network ResNet18 after different methods of data enhancement. The comparison of the results shows that the accuracy, recall and F1 scores of the classification network trained on the proposed FP-CycleGAN amplified dataset have improved significouldtly on the test set.

By comparison, the diagnostic accuracy of the classification network trained by all methods is 100% for the bearing images labeled F-N-0.8, but for the other six categories, the proposed method has a great improvement in the accuracy.

Table 4.5: Results of different data amplification methods

Data Amplification Method	Label	Pre (%)	Rec (%)	F1 (%)	Mean Pre (%)	Top-1 Acc (%)	Mean F1(%)
No Amplification	F-N-0.6	78.67	71.00	74.64	76.81	72.29	72.41
	F-N-0.8	55.79	100.00	71.62			
	F-N-1.0	77.88	62.50	69.35			
	F-Q-0.8	91.82	61.75	73.84			
	F-W-0.6	70.00	80.50	74.88			
	F-W-0.8	65.48	69.25	67.31			
	F-W-1.0	97.99	61.00	75.19			

Data Amplification Method	Label	Pre (%)	Rec (%)	F1 (%)	Mean Pre (%)	Top-1 Acc (%)	Mean F1(%)
CycleGAN-Unet	F-N-0.6	100.00	62.50	76.92	77.49	73.32	73.36
	F-N-0.8	79.52	100.00	88.59			
	F-N-1.0	61.05	64.25	62.61			
	F-Q-0.8	62.17	71.50	66.51			
	F-W-0.6	59.31	90.00	71.50			
	F-W-0.8	98.43	62.50	76.45			
	F-W-1.0	81.97	62.50	70.92			
CycleGAN-ResNet	F-N-0.6	100.00	62.50	76.92	83.75	80.00	80.38
	F-N-0.8	68.26	100.00	81.14			
	F-N-1.0	96.51	76.00	85.03			
	F-Q-0.8	59.40	79.00	67.81			
	F-W-0.6	86.49	80.00	83.12			
	F-W-0.8	75.57	75.00	75.28			
	F-W-1.0	100.00	87.50	93.33			
FP-CycleGAN without D_AB	F-N-0.6	99.67	75.00	85.59	89.40	87.50	87.72
	F-N-0.8	72.07	100.00	83.77			
	F-N-1.0	76.59	87.50	81.68			
	F-Q-0.8	97.77	87.50	92.35			
	F-W-0.6	89.51	87.50	88.50			
	F-W-0.8	100.00	87.50	93.33			
	F-W-1.0	90.21	87.50	88.83			

Data Amplification Method	Label	Pre (%)	Rec (%)	F1 (%)	Mean Pre (%)	Top-1 Acc (%)	Mean F1(%)
FP-CycleGAN	F-N-0.6	100.00	75.25	85.88	91.52	87.93	88.61
	F-N-0.8	61.07	100.00	75.83			
	F-N-1.0	83.73	87.50	85.57			
	F-Q-0.8	99.43	87.50	93.09			
	F-W-0.6	98.37	90.25	94.13			
	F-W-0.8	100.00	87.50	93.33			
	F-W-1.0	98.04	87.50	92.47			

The table presents the results of different data augmentation methods applied to the classification network ResNet18. Each method is evaluated using various metrics, including Precision (Pre), Recall (Rec), F1 score (F1), Mean Precision (Mean Pre), Top-1 Accuracy (Top-1 Acc), and Mean F1 score (Mean F1). Here's the analysis of each method and its results:

1. The "No Amplification" method refers to the scenario where no data augmentation is applied to the training dataset. In this case, the classification network ResNet18 is trained directly on the original data without any synthetic data generation or transformations. As a result, the model's performance is relatively limited, with lower Precision, Recall, and F1 scores compared to methods involving data augmentation. Specifically, for the F-N-0.8 label, although the Recall is perfect (100%), the Precision is only 55.79%, indicating that the model is overly optimistic in predicting positive instances but fails to accurately identify many true positives. This lack of augmentation leads to poorer generalization, as

the model struggles to distinguish between different categories, yielding suboptimal results in terms of both accuracy and F1 score across most labels.

2. When applying the CycleGAN-Unet method for data augmentation, we could observe its ability to amplify thermal image data. This approach uses UNet as the generator architecture, addressing the issue of transferring different defect features in the data and generating images with multi-angle characteristics, thus expanding the diversity of the dataset. Specifically, the CycleGAN-Unet method effectively preserves defect features during image generation, but slightly lags behind FP-CycleGAN in terms of image realism (X. Sun et al., 2024).

From the results, we see that CycleGAN-Unet achieves 100% accuracy for the F-N-0.6 class, but the recall rate is only 62.5%. This indicates that the model is overly optimistic in predicting positive instances but fails to accurately identify many of the true positives. In the F-N-0.8 class, although the recall is 100%, precision is lower at 79.52%, suggesting that while the generated images transfer most of the defect features, there are still noticeable deviations in the synthetic images, potentially impacting overall classification performance.

Another notable characteristic of CycleGAN-Unet is that it outperforms the baseline model (i.e., the "no augmentation" approach) across several metrics. Specifically, in the F-Q-0.8 category, it achieves a precision of 62.17% and recall of 71.5%. However, despite the more balanced generation of images, some finer details might not be fully recovered, leading to a lack of realism in the generated images, which could affect classification accuracy.

In conclusion, while CycleGAN-Unet is effective at generating images with multi-angle features and diversifying the dataset, its lower image realism might limit its practical performance. With further optimization and integration with other techniques, this method holds potential for improving both the quality of the generated images and the overall model performance.

3. When applying the CycleGAN-ResNet method for data augmentation, we observe its ability to generate multi-angle thermal image data, thereby addressing the challenge of transferring defect features in the dataset. This approach uses ResNet as the generator architecture, which is known for its deep learning capability and feature extraction prowess. The CycleGAN-ResNet method works effectively in amplifying the diversity of the dataset, producing synthetic images with varying characteristics that reflect different fault conditions. Specifically, the use of ResNet helps preserve important defect features during image generation, contributing to better fault identification compared to simpler methods (K. He et al., 2016).

From the results, CycleGAN-ResNet achieves 100% accuracy for the F-N-0.6 class, demonstrating its ability to generate realistic images for this category. However, the recall rate is only 62.5%, suggesting that while the model is effective at identifying positive instances, it may miss some of the true positives, thus not fully capturing the diversity of fault features. In the F-N-0.8 class, although recall remains at 100%, the precision is slightly reduced to 68.26%. This indicates that while the generated images accurately transfer many defect features, some of the synthetic images exhibit deviations from real-world fault patterns, which could affect classification accuracy.

In the F-Q-0.8 class, CycleGAN-ResNet achieves a precision of 59.4% and recall of 79%, showing a good balance in transferring defect features. However, similar to other methods, some subtle details may not be fully recovered in the generated images, which could reduce the realism of the synthetic data and, in turn, impact the final classification performance.

Notably, CycleGAN-ResNet outperforms the baseline model (i.e., the "no augmentation" approach) across several fault categories, demonstrating the effectiveness of using ResNet for feature extraction and improving dataset diversity. However, while the method shows promise, further refinements are needed to improve the precision and recall for certain classes, especially in cases where subtle defect details are critical for accurate fault detection.

In conclusion, CycleGAN-ResNet is a robust data augmentation technique that generates diverse and defect-rich images, significouldtly improving model performance compared to the baseline. While it enhances feature transfer and image diversity, optimizing its ability to recover finer image details and improving overall image realism could further boost the method's practical applicability in bearing fault detection.

4. When FP-CycleGAN without D_AB method is used for data expansion, we observe significant improvement in image generation, especially in defect feature transfer and image diversity. The FP-CycleGAN method aims to generate more realistic images using the optimized CycleGAN, without relying on the additional discriminator D_AB for feature-specific guidance. In this method, the generator learns to transmit defect features based on classical discriminators A and B, thus enhancing the quality and diversity of the generated data set.

From the results, FP-CycleGAN without D_AB achieves a precision of 99.67% and a recall of 75.00% for the F-N-0.6 class, resulting in an F1 score of 85.59%. These metrics demonstrate that the model is highly effective at generating images that match real fault characteristics, though it still struggles somewhat in identifying all true positives. The precision and recall for the F-N-0.8 class are 72.07% and 100%, respectively, leading to an F1 score of 83.77%. Although recall is perfect, the reduced precision indicates that there are still some synthetic images deviating from the real fault features, which could affect classification accuracy.

In the F-Q-0.8 class, FP-CycleGAN without D_AB performs exceptionally well, with a precision of 97.77% and recall of 87.50%, resulting in an F1 score of 92.35%. These results highlight the method's ability to generate diverse, yet accurate, images across various fault categories. However, as with other methods, there may still be subtle image quality issues, particularly in terms of fine-grained defect details, which could affect performance in highly demanding classification tasks.

In the F-W-0.6 and F-W-1.0 categories, FP-CycleGAN without D_AB achieves strong precision rates of 89.51% and 90.21%, respectively, while maintaining consistent recall values of 87.50%. These results demonstrate the method's overall robustness in generating synthetic images across a range of fault conditions. However, the method's lack of the additional discriminator (D_AB) for feature-specific guidance might still limit its ability to fully capture complex fault details, leading to slightly lower precision in certain scenarios.

Overall, FP-CycleGAN without D_AB demonstrates impressive performance in fault detection tasks, with high precision and recall rates across various classes. Its ability to generate diverse and defect-rich images is clear, but the lack of D_AB's guidance may slightly impact the realism and precision of some generated images, particularly in complex fault categories. Despite this, the method shows great potential in improving the overall model performance for defect detection, with further optimization possible through the integration of more sophisticated feature extraction techniques.

5. When applying the FP-CycleGAN method for data augmentation, we observe its significant impact on improving the diversity and quality of thermal images for bearing

fault diagnosis. FP-CycleGAN leverages a novel discriminator, D_AB, to preserve non-target features in the source domain while efficiently migrating the target features to the target domain. This method addresses the issue of data imbalance in infrared thermal images of faulty rotor bearings and enhances the realism of the generated images, ensuring that critical fault features are not lost during the image transformation process.

From the results, FP-CycleGAN demonstrates superior performance, achieving 100% precision for the F-N-0.6 class, with a recall rate of 75.25%. This indicates that the model is highly accurate in detecting defects, but it could potentially miss a few true positives, as evidenced by the slightly lower recall. In the F-N-0.8 class, FP-CycleGAN achieves 100% recall, although the precision is reduced to 61.07%, which suggests that while the model successfully captures all positive instances, some of the generated images might include irrelevant features, lowering the overall precision.

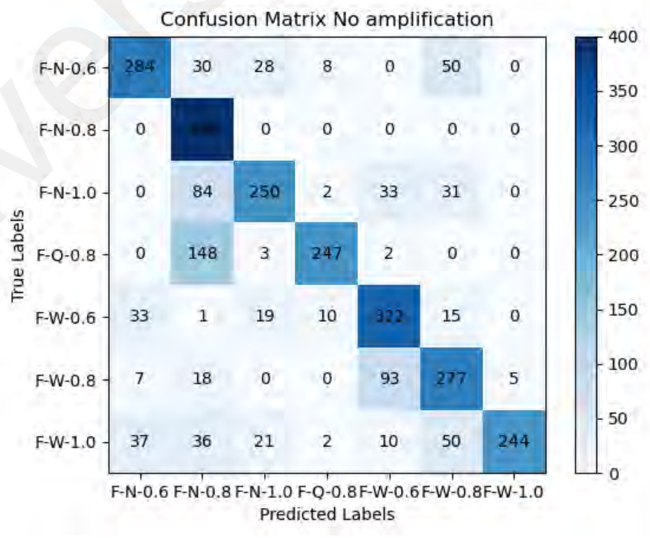
In the F-N-1.0 class, FP-CycleGAN achieves a precision of 83.73% and recall of 87.50%, showing a good balance between detecting positive instances and minimizing false positives. This class demonstrates the model's ability to preserve defect characteristics while generating high-quality images that aid in fault detection. Moreover, FP-CycleGAN performs exceptionally well in the F-Q-0.8 class, achieving 99.43% precision and 87.50% recall, reflecting its ability to generate highly realistic fault images while maintaining classification accuracy.

One of the key advantages of FP-CycleGAN, particularly when compared to the baseline model (no augmentation) and other methods like CycleGAN-Unet and CycleGAN-ResNet, is its ability to balance precision and recall across different fault categories. It shows improved results across several fault types, including F-W-0.6 (98.37% precision, 90.25% recall), F-W-0.8 (100% precision, 87.50% recall), and F-W-1.0 (98.04%

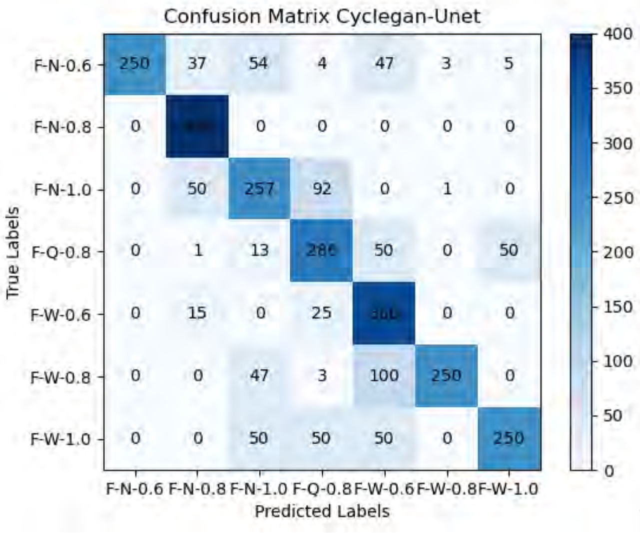
precision, 87.50% recall), indicating that the model could effectively preserve defect features while reducing data imbalance and generating more realistic thermal images.

In conclusion, FP-CycleGAN, with its novel discriminator D_AB, demonstrates remarkable performance in fault diagnosis tasks by generating high-quality, realistic thermal images and enhancing the dataset's diversity. The method shows significant improvements over baseline models and other augmentation techniques, achieving higher precision and recall in multiple fault categories. While the precision may occasionally suffer in some classes, FP-CycleGAN's overall effectiveness makes it a promising approach for bearing fault diagnosis and thermal image data augmentation. Further optimization could further improve its precision, especially in terms of generating images with even finer details for fault classification.

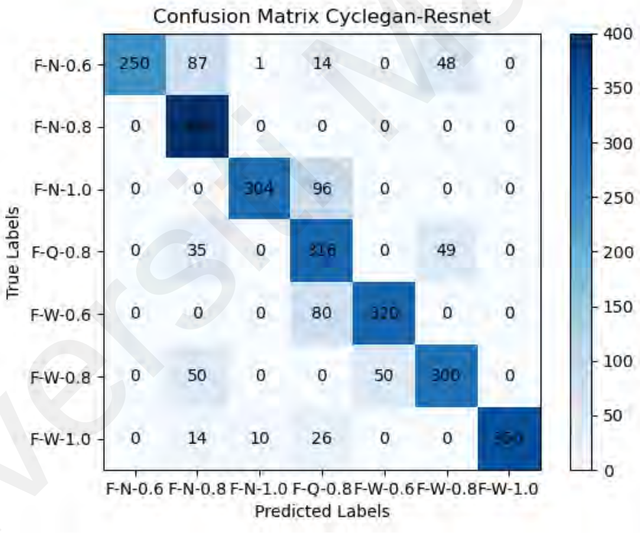
Figure 4.7 shows the confusion matrix for the classification network trained on the original and different data amplification datasets.



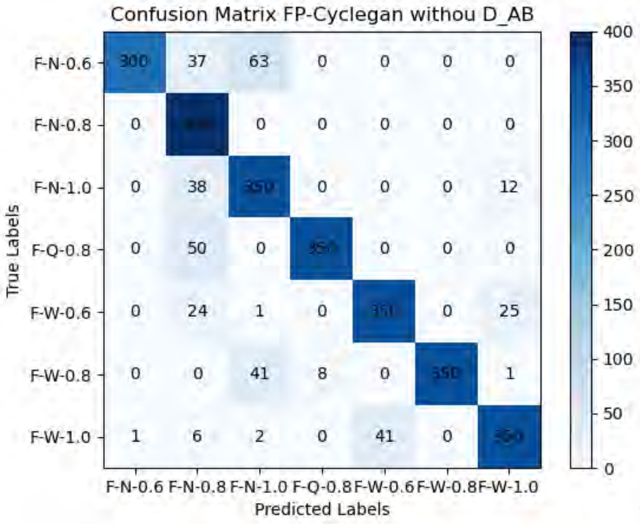
(a)


Confusion Matrix Cyclegan-Unet
True Labels
Predicted Labels
F-N-0.6 250 37 54 4 47 3 5
F-N-0.8 0 0 0 0 0 0
F-N-1.0 0 50 257 92 0 1 0
F-Q-0.8 0 1 13 286 50 0 50
F-W-0.6 0 15 0 25 360 0 0
F-W-0.8 0 0 47 3 100 250 0
F-W-1.0 0 0 50 50 50 0 250
F-N-0.6 F-N-0.8 F-N-1.0 F-Q-0.8 F-W-0.6 F-W-0.8 F-W-1.0
400 350 300 250 200 150 100 50 0

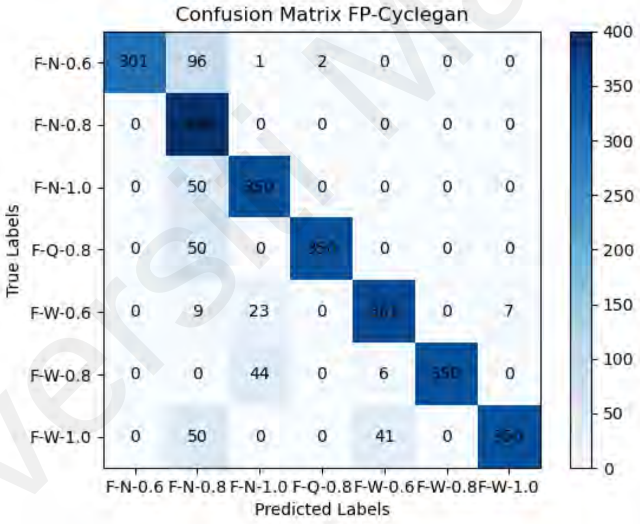
(b)


Confusion Matrix Cyclegan-Resnet
True Labels
Predicted Labels
F-N-0.6 250 87 1 14 0 48 0
F-N-0.8 0 0 0 0 0 0
F-N-1.0 0 0 304 96 0 0 0
F-Q-0.8 0 35 0 316 0 49 0
F-W-0.6 0 0 0 80 320 0 0
F-W-0.8 0 50 0 0 50 300 0
F-W-1.0 0 14 10 26 0 0 350
F-N-0.6 F-N-0.8 F-N-1.0 F-Q-0.8 F-W-0.6 F-W-0.8 F-W-1.0
400 350 300 250 200 150 100 50 0

(c)



(d)



(e)

Figure 4.7: The confusion matrix for the classification network. (a) No amplification. (b) Cyclegan-Unet. (c) Cyclegan-Resnet. (d) FP-Cyclegan- Without D_AB. (e) FP-Cyclegan

The five confusion matrices shown represent the performance of different data augmentation methods used to improve the classification of bearing thermal images. Each matrix corresponds to a different model trained using the following methods:

a. No Amplification:

The confusion matrix for the "No Amplification" method shows that the model struggles with classifying most categories correctly. There are several misclassifications, particularly for the F-N-0.6 and F-W-0.6 classes. The F-N-0.8 label achieves perfect accuracy with 400 correct predictions, but other labels exhibit considerable misclassification. For example, F-Q-0.8 has 148 instances misclassified as F-N-0.8, F-N-1.0 has 84 instances misclassified as F-N-0.8. The overall performance is poor in terms of precision and recall across most categories, suggesting the lack of augmented data hampers the model's ability to generalize.

b. CycleGAN-Unet:

With the CycleGAN-Unet method, the model shows a notable improvement. The misclassifications are reduced, especially for F-Q-0.8, where fewer instances are misclassified into other labels. F-N-0.8 still has perfect recall (400), and F-N-0.6 and F-N-1.0 also show fewer misclassifications compared to the "No Amplification" method. However, F-W-0.8 still experiences some misclassifications, although not as severe. This indicates that the CycleGAN-Unet method provides better feature transfer and class balance but still leaves room for improvement in the finer details of fault detection.

c. CycleGAN-ResNet:

The CycleGAN-ResNet method demonstrates further improvements. It maintains the perfect classification for F-N-0.8 and improves precision and recall in categories like F-N-1.0 and F-Q-0.8. The F-W-1.0 categories see fewer errors compared to previous models. However, its performance in F-W-0.6 is worse, and 320 correct samples are reduced. We guess this may be caused by the defects between F-W-0.6 and F-Q-0.8 are similar. The overall balance between precision and recall is better, which is indicative of the robust performance of the ResNet backbone in handling the thermal image features and generating more realistic data.

d. FP-CycleGAN without D_AB:

The FP-CycleGAN without D_AB model shows strong performance. The misclassifications in F-N-0.6 and F-N-0.8 are much lower than in earlier models. Although the F-N-1.0 class still faces some challenges (38 +12 misclassifications), the precision and recall for F-Q-0.8 and F-W-1.0 are improved. However, F-W-0.6 still faces some misclassifications. The absence of the D_AB discriminator affects the model's ability to maintain all non-target features, resulting in slightly reduced accuracy in some fault classes. Nevertheless, the model still performs well and offers a good balance between speed and image realism.

e. FP-CycleGAN:

The full FP-CycleGAN model demonstrates the best overall performance. Precision and recall are high for all classes, especially for F-W-0.6 (361 out of 400 correctly classified). The F-N-0.6 and F-N-1.0 categories show excellent results, with relatively few misclassifications. F-W-1.0 and F-Q-0.8 also perform well, with minimal errors. The results indicate that FP-CycleGAN, with the full implementation of the D_AB discriminator, significouldtly improves image quality and classification performance.

In addition, we found that for the results of FP-CycleGAN, many other categories were wrongly classified into the category of F-N-0.8, which is acceptable and understandable, because all multi-angle data are generated by the multi-angle data of F-N-0.8 combined with each single-angle data. After adding D_AB, more emphasis is placed on the information from the field of F-N-0.8.

FP-CycleGAN performs best across all categories, with the highest precision and recall values, indicating its effectiveness in enhancing image realism and fault detection. CycleGAN-ResNet offers solid performance with fewer misclassifications compared to CycleGAN-Unet, especially for more complex categories. CycleGAN-ResNet and FP-CycleGAN without D_AB show improvements over the baseline (No Amplification) but still lag behind FP-CycleGAN in terms of precision and recall balance.

The results highlight the importance of data augmentation in improving the generalization of the model and ensuring accurate fault detection across various fault categories.

4.5 Summary

In this chapter, we have detailed the experimental results of applying the FP-CycleGAN model for image amplification and bearing damage detection. The proposed method demonstrated significant improvements over traditional approaches by effectively handling data imbalance and enhancing image quality. By employing advanced techniques such as Bicubic interpolation and depthwise separable convolutions, the model achieved superior thermal image generation, retaining critical defect features while improving image clarity. The experiments were conducted using real bearing thermal images with varying fault conditions, where the FP-CycleGAN model utilized both single and multi-angle data to expand the training set, ensuring better generalization for fault detection. Furthermore, the comparison with other models, including CycleGAN with ResNet and UNet backbones, highlighted the advantages of the FP-CycleGAN framework in terms of both computational efficiency and image realism. Overall, the proposed method significouldtly outperformed previous algorithms, offering a robust solution for bearing fault diagnosis in real-world applications. The results support the potential of FP-CycleGAN for improving the reliability and accuracy of fault detection in industrial systems

CHAPTER 5: CONCLUSION

This research proposes an unsupervised learning framework based on CycleGAN for detecting defects in small sample rolling bearing thermal images. Infrared thermal images could be used to detect defects in a non-contact way, thus avoiding the shortcomings of traditional methods such as vibration signal, current characteristics and sound pressure analysis in bearing detection.

The proposed model is mainly based on Cyclegan, and the generator and loss function are greatly optimized. The framework firstly reconstructs the generator for the highly complex rolling bearing defect thermal images to improve the extraction capability of the network for different dimensional features and the reconstruction capability of the images. To avoid the loss of features in the target domain of the generated images, a new D_AB is designed to identify whether the generated image A and the generated image B belong to two different classes, and a new class loss is proposed to ensure that the generated images should keep the fault features of the target domain. Thus, the accurate transfer of fault features is achieved while the non-fault features of the image are highly maintained. In other words, it solves the problem of lack of realism and non-fault features in the generated images, which are often found in adversarial generation networks. Experiments show that the proposed FP-CycleGAN algorithm outperforms the conventional CycleGAN-based feature transfer algorithms in terms of stability and accuracy. It is noteworthy that it provides a new research idea in maintaining the pattern non-collapse and the realism of the generated images while performing feature transfer.

Future research work will focus on the processing time of the model and data collection. Although the performance of the proposed method has been greatly improved by adding a new loss function and reconstructing the generator, compared with other methods, this process needs more time and computing resources. Therefore, how to reasonably design the architecture of generator and discriminator needs more experiments to explore. In addition, how to rationally allocate the weights of different loss functions, gradient return and parameter optimization also needs further experiments. For data collection and verification, although thousands of multi-angle thermal images of bearings have been collected, they are still not enough to be generalized to any possible application scenarios in practice. Therefore, more data collection experiments should be designed or cooperated with other research groups.

REFERENCES

AlShorman, O., Irfan, M., Abdelrahman, R. B., Masadeh, M., Alshorman, A., Sheikh, M. A., Saad, N., & Rahman, S. (2024). Advancements in condition monitoring and fault diagnosis of rotating machinery: A comprehensive review of image-based intelligent techniques for induction motors. *Engineering Applications of Artificial Intelligence*, *130*, 107724. https://doi.org/10.1016/j.engappai.2023.107724

Arjovsky, M., Chintala, S., & Bottou, L. (2017). *Wasserstein GAN* (No. arXiv:1701.07875). arXiv. http://arxiv.org/abs/1701.07875

Asif, S., Zhao, M., Tang, F., & Zhu, Y. (2024). DCDS-Net: Deep transfer network based on depth-wise separable convolution with residual connection for diagnosing gastrointestinal diseases. *Biomedical Signal Processing and Control*, *90*, 105866. https://doi.org/10.1016/j.bspc.2023.105866

Benchallal, F., Hafiane, A., Ragot, N., & Canals, R. (2024). ConvNeXt based semi-supervised approach with consistency regularization for weeds classification. *Expert Systems with Applications*, *239*, 122222. https://doi.org/10.1016/j.eswa.2023.122222

Bianchini, C., Immovilli, F., Cocconcelli, M., Rubini, R., & Bellini, A. (2011). Fault Detection of Linear Bearings in Brushless AC Linear Motors by Vibration Analysis. *IEEE Transactions on Industrial Electronics*, *58*(5), 1684–1694. https://doi.org/10.1109/TIE.2010.2098354

Blodt, M., Granjon, P., Raison, B., & Rostaing, G. (2008). Models for Bearing Damage Detection in Induction Motors Using Stator Current Monitoring. *IEEE Transactions on Industrial Electronics*, *55*(4), 1813–1822. https://doi.org/10.1109/TIE.2008.917108

Bynagari, N. B. (2019). GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium. *Asian Journal of Applied Science and Engineering*, *8*(1), 25–34. https://doi.org/10.18034/ajase.v8i1.9

Catal Reis, H., & Turk, V. (2024). Potato leaf disease detection with a novel deep learning model based on depthwise separable convolution and transformer networks. *Engineering Applications of Artificial Intelligence*, *133*, 108307. https://doi.org/10.1016/j.engappai.2024.108307

Chen, J., Yang, G., Khan, H., Zhang, H., Zhang, Y., Zhao, S., Mohiaddin, R., Wong, T., Firmin, D., & Keegan, J. (2022). JAS-GAN: Generative Adversarial Network Based Joint Atrium and Scar Segmentations on Unbalanced Atrial Targets. *IEEE Journal of Biomedical and Health Informatics*, *26*(1), 103–114. https://doi.org/10.1109/JBHI.2021.3077469

Chen, X., Guo, Y., Li, C., & Kang, W. (2024). Optical Encoder-Based Feature Extract of Faulty Rolling Element Bearing Under Variable-Speed Conditions. *IEEE Transactions on Instrumentation and Measurement*, *73*, 1–10. https://doi.org/10.1109/TIM.2024.3417604

Chen, X., Yang, R., Xue, Y., Huang, M., Ferrero, R., & Wang, Z. (2023). Deep Transfer Learning for Bearing Fault Diagnosis: A Systematic Review Since 2016. *IEEE Transactions on Instrumentation and Measurement*, *72*, 1–21. https://doi.org/10.1109/TIM.2023.3244237

Cheung, W. K., Pakzad, A., Mogulkoc, N., Needleman, S. H., Rangelov, B., Gudmundsson, E., Zhao, A., Abbas, M., McLaverty, D., Asimakopoulos, D., Chapman, R., Savas, R., Janes, S. M., Hu, Y., Alexander, D. C., Hurst, J. R., & Jacob, J. (2024). Interpolation-split: A data-centric deep learning approach with big interpolated data to boost airway segmentation performance. *Journal of Big Data*, *11*(1), 104. https://doi.org/10.1186/s40537-024-00974-x

Choudhary, A., Mian, T., & Fatima, S. (2021). Convolutional neural network based bearing fault diagnosis of rotating machine using thermal images. *Measurement*, *176*, 109196. https://doi.org/10.1016/j.measurement.2021.109196

Coats, M. D., & Randall, R. B. (2014). Single and multi-stage phase demodulation based order-tracking. *Mechanical Systems and Signal Processing*, *44*(1–2), 86–117. https://doi.org/10.1016/j.ymssp.2013.09.016

Dalvand, F., Dalvand, S., Sharafi, F., & Pecht, M. (2017). Current Noise Cancellation for Bearing Fault Diagnosis Using Time Shifting. *IEEE Transactions on Industrial Electronics*, *64*(10), 8138–8147. https://doi.org/10.1109/TIE.2017.2694397

Deekshit Kompella, K. C., Venu Gopala Rao, M., & Srinivasa Rao, R. (2018). Bearing fault detection in a 3 phase induction motor using stator current frequency spectral subtraction with various wavelet decomposition techniques. *Ain Shams Engineering Journal*, *9*(4), 2427–2439. https://doi.org/10.1016/j.asej.2017.06.002

Deng, F., Chen, J., Wang, Y., & Gong, K. (2013). Measurement and calibration method for an optical encoder based on adaptive differential evolution-Fourier neural networks. *Measurement Science and Technology*, *24*(5), 055007. https://doi.org/10.1088/0957-0233/24/5/055007

Deveci, B. U., Celtikoglu, M., Albayrak, O., Unal, P., & Kirci, P. (2024). Transfer Learning Enabled Bearing Fault Detection Methods Based on Image Representations of Single-Dimensional Signals. *Information Systems Frontiers*, *26*(4), 1345–1397. https://doi.org/10.1007/s10796-023-10371-z

Dolenc, B., Boškoski, P., & Juričić, Đ. (2016). Distributed bearing fault diagnosis based on vibration analysis. *Mechanical Systems and Signal Processing*, *66–67*, 521–532. https://doi.org/10.1016/j.ymssp.2015.06.007

Dong, C., Loy, C. C., He, K., & Tang, X. (2016). Image Super-Resolution Using Deep Convolutional Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *38*(2), 295–307. https://doi.org/10.1109/TPAMI.2015.2439281

Dong, F., Zhu, M., Wang, Y., Chen, Z., Dai, Y., Xi, Z., Du, T., & Xu, M. (2025). AI-enabled rolling triboelectric nanogenerator for bearing wear diagnosis aiming at digital twin application. *Nano Energy*, *134*, 110550. https://doi.org/10.1016/j.nanoen.2024.110550

Dong, Y., Jiang, H., Yao, R., Mu, M., & Yang, Q. (2024). Rolling bearing intelligent fault diagnosis towards variable speed and imbalanced samples using multiscale dynamic supervised contrast learning. *Reliability Engineering & System Safety*, *243*, 109805. https://doi.org/10.1016/j.ress.2023.109805

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2021). *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale* (No. arXiv:2010.11929). arXiv. http://arxiv.org/abs/2010.11929

Feng, Z., Gao, T., Yu, X., Zhang, Y., Chen, X., Yang, Y., & Du, M. (2025). Planet bearing fault diagnosis via double encoder signal analysis. *Mechanical Systems and Signal Processing*, *224*, 111978. https://doi.org/10.1016/j.ymssp.2024.111978

Glowacz, A. (2024). Ventilation diagnosis of minigrinders using thermal images. *Expert Systems with Applications*, *237*, 121435. https://doi.org/10.1016/j.eswa.2023.121435

Glowacz, A., Sulowicz, M., Zielonka, J., Li, Z., Glowacz, W., & Kumar, A. (2025). Acoustic fault diagnosis of three-phase induction motors using smartphone and deep learning. *Expert Systems with Applications*, *262*, 125633. https://doi.org/10.1016/j.eswa.2024.125633

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). *Generative Adversarial Networks* (No. arXiv:1406.2661). arXiv. http://arxiv.org/abs/1406.2661

Guan, B., Bao, X., Qiu, H., & Yang, D. (2024). Enhancing bearing fault diagnosis using motor current signals: A novel approach combining time shifting and CausalConvNets. *Measurement*, *226*, 114049. https://doi.org/10.1016/j.measurement.2023.114049

Gulsoy, T., & Baykal Kablan, E. (2025). FocalNeXt: A ConvNeXt augmented FocalNet architecture for lung cancer classification from CT-scan images. *Expert Systems with Applications*, *261*, 125553. https://doi.org/10.1016/j.eswa.2024.125553

Guo, W., Li, F., Zhang, P., & Luo, L. (2025). A stage-related online incremental transfer learning-based remaining useful life prediction method of bearings. *Applied Soft Computing*, *169*, 112491. https://doi.org/10.1016/j.asoc.2024.112491

Hakim, M., Omran, A. A. B., Ahmed, A. N., Al-Waily, M., & Abdellatif, A. (2023a). A systematic review of rolling bearing fault diagnoses based on deep learning and transfer learning: Taxonomy, overview, application, open challenges, weaknesses and recommendations. *Ain Shams Engineering Journal*, *14*(4), 101945. https://doi.org/10.1016/j.asej.2022.101945

Hakim, M., Omran, A. A. B., Ahmed, A. N., Al-Waily, M., & Abdellatif, A. (2023b). A systematic review of rolling bearing fault diagnoses based on deep learning and transfer learning: Taxonomy, overview, application, open challenges, weaknesses and recommendations. *Ain Shams Engineering Journal*, *14*(4), 101945. https://doi.org/10.1016/j.asej.2022.101945

Han, S., Sun, S., Zhao, Z., Luan, Z., & Niu, P. (2024). Deep Residual Multiscale Convolutional Neural Network With Attention Mechanism for Bearing Fault Diagnosis Under Strong Noise Environment. *IEEE Sensors Journal*, *24*(6), 9073–9081. https://doi.org/10.1109/JSEN.2023.3345400

He, D., Zhang, Z., Jin, Z., Zhang, F., Yi, C., & Liao, S. (2025). RTSMFFDE-HKRR: A fault diagnosis method for train bearing in noise environment. *Measurement*, *239*, 115417. https://doi.org/10.1016/j.measurement.2024.115417

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. https://doi.org/10.1109/CVPR.2016.90

Hendrycks, D., & Gimpel, K. (2023). *Gaussian Error Linear Units (GELUs)* (No. arXiv:1606.08415). arXiv. https://doi.org/10.48550/arXiv.1606.08415

Hou, J., Gendy, G., Chen, G., Wang, L., & He, G. (2024). DTDeMo: A Deep Learning-Based Two-Stage Image Demosaicing Model With Interpolation and Enhancement. *IEEE Transactions on Computational Imaging*, *10*, 1026–1039. https://doi.org/10.1109/TCI.2024.3426360

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications* (No. arXiv:1704.04861). arXiv. http://arxiv.org/abs/1704.04861

Hu, J., Chen, Z., Yang, M., Zhang, R., & Cui, Y. (2018). A Multiscale Fusion Convolutional Neural Network for Plant Leaf Recognition. *IEEE Signal Processing Letters*, *25*(6), 853–857. https://doi.org/10.1109/LSP.2018.2809688

Hu, J., Yang, H., He, J., Bai, D., & Chen, H. (2024). EHA-YOLOv5: An Efficient and Highly Accurate Improved YOLOv5 Model for Workshop Bearing Rail Defect Detection Application. *IEEE Access*, *12*, 81911–81924. https://doi.org/10.1109/ACCESS.2024.3412425

Huang, X., Zhang, J., & Cheng, M. (2024). Fault Detection of Servo Motor Bearing Based on Speed Signal Under Variable-Speed Conditions. *IEEE Transactions on Instrumentation and Measurement*, *73*, 1–12. https://doi.org/10.1109/TIM.2024.3381274

Huo, L., Qi, H., Fei, S., Guan, C., & Li, J. (2022). A Generative Adversarial Network Based a Rolling Bearing Data Generation Method Towards Fault Diagnosis. *Computational Intelligence and Neuroscience*, *2022*, 1–21. https://doi.org/10.1155/2022/7592258

Ibrahim, A., El Badaoui, M., Guillet, F., & Bonnardot, F. (2008). A New Bearing Fault Detection Method in Induction Machines Based on Instantaneous Power Factor. *IEEE Transactions on Industrial Electronics*, *55*(12), 4252–4259. https://doi.org/10.1109/TIE.2008.2003211

Janssens, O., Loccufier, M., & Van Hoecke, S. (2019). Thermal Imaging and Vibration-Based Multisensor Fault Detection for Rotating Machinery. *IEEE Transactions on Industrial Informatics*, *15*(1), 434–444. https://doi.org/10.1109/TII.2018.2873175

Jiang, Z., Zhang, G., & Gao, Y. (2025). Two-dimensional quad-stable Gaussian potential stochastic resonance model for enhanced bearing fault diagnosis. *Applied Mathematical Modelling*, *137*, 115657. https://doi.org/10.1016/j.apm.2024.115657

Jin, X., Zhou, W., Ma, J., Su, H., Liu, S., & Gao, B. (2025). Analysis on the vibration signals of a novel double-disc crack rotor-bearing system with single defect in inner race. *Journal of Sound and Vibration*, *595*, 118729. https://doi.org/10.1016/j.jsv.2024.118729

Khor, W., Chen, Y. K., Roberts, M., & Ciampa, F. (2024). Non-contact, portable, and stand-off infrared thermal imager for security scanning applications. *AIP Advances*, *14*(4), 045314. https://doi.org/10.1063/5.0188862

Kiliçarslan, S., & Celik, M. (2024). Parametric RSigELU: A new trainable activation function for deep learning. *Neural Computing and Applications*, *36*(13), 7595–7607. https://doi.org/10.1007/s00521-024-09538-9

Kingma, D. P., & Welling, M. (2022). *Auto-Encoding Variational Bayes* (No. arXiv:1312.6114). arXiv. http://arxiv.org/abs/1312.6114

Kumar, A., Kumar, R., Xiang, J., Qiao, Z., Zhou, Y., & Shao, H. (2024). Digital twin-assisted AI framework based on domain adaptation for bearing defect diagnosis in the centrifugal pump. *Measurement*, *235*, 115013. https://doi.org/10.1016/j.measurement.2024.115013

Kumar, P., & Hati, A. S. (2022). Dilated convolutional neural network based model for bearing faults and broken rotor bar detection in squirrel cage induction motors. *Expert Systems with Applications*, *191*, 116290. https://doi.org/10.1016/j.eswa.2021.116290

Lee, M. (2023). *GELU Activation Function in Deep Learning: A Comprehensive Mathematical Analysis and Performance* (No. arXiv:2305.12073). arXiv. https://doi.org/10.48550/arXiv.2305.12073

Li, C., De Oliveira, J. V., Cerrada, M., Cabrera, D., Sanchez, R. V., & Zurita, G. (2019). A Systematic Review of Fuzzy Formalisms for Bearing Fault Diagnosis. *IEEE Transactions on Fuzzy Systems*, *27*(7), 1362–1382. https://doi.org/10.1109/TFUZZ.2018.2878200

Li, D., Cheng, B., & Xiang, S. (2024). Direct cubic B-spline interpolation: A fuzzy interpolating method for weightless, robust and accurate DVC computation. *Optics and Lasers in Engineering*, *172*, 107886. https://doi.org/10.1016/j.optlaseng.2023.107886

Li, J., Liu, Y., & Li, Q. (2022). Intelligent fault diagnosis of rolling bearings under imbalanced data conditions using attention-based deep learning method. *Measurement*, *189*, 110500. https://doi.org/10.1016/j.measurement.2021.110500

Li, X., Ma, Z., Yuan, Z., Mu, T., Du, G., Liang, Y., & Liu, J. (2024). A review on convolutional neural network in rolling bearing fault diagnosis. *Measurement Science and Technology*, *35*(7), 072002. https://doi.org/10.1088/1361-6501/ad356e

Li, X., Teng, W., Wang, L., Hu, J., Su, Y., Peng, D., & Liu, Y. (2025). Trend-constrained pairing based incremental transfer learning for remaining useful life prediction of bearings in wind turbines. *Expert Systems with Applications*, *263*, 125731. https://doi.org/10.1016/j.eswa.2024.125731

Li, X., Wang, Y., Yao, J., Li, M., & Gao, Z. (2024). Multi-sensor fusion fault diagnosis method of wind turbine bearing based on adaptive convergent viewable neural networks. *Reliability Engineering & System Safety*, *245*, 109980. https://doi.org/10.1016/j.ress.2024.109980

Lian, Y., Wang, J., Li, Z., Liu, W., Huang, L., & Jiang, X. (2025). Residual attention guided vision transformer with acoustic-vibration signal feature fusion for cross-domain fault diagnosis. *Advanced Engineering Informatics*, *64*, 103003. https://doi.org/10.1016/j.aei.2024.103003

Liang, X., Hu, P., Zhang, L., Sun, J., & Yin, G. (2019). MCFNet: Multi-Layer Concatenation Fusion Network for Medical Images Fusion. *IEEE Sensors Journal*, *19*(16), 7107–7119. https://doi.org/10.1109/JSEN.2019.2913281

Liao, Y., Jiang, X., Zhang, Z., Zheng, H., Li, T., & Chen, Y. (2023). The Influence of Wind Speed on the Thermal Imaging Clarity Based Inspection for Transmission Line Conductors. *IEEE Transactions on Power Delivery*, *38*(3), 2101–2109. https://doi.org/10.1109/TPWRD.2022.3232738

Liu, G., & Wu, L. (2024). Incremental bearing fault diagnosis method under imbalanced sample conditions. *Computers & Industrial Engineering*, *192*, 110203. https://doi.org/10.1016/j.cie.2024.110203

Liu, J., Sun, W., Zhao, X., Zhao, J., & Jiang, Z. (2022). Deep feature fusion classification network (DFFCNet): Towards accurate diagnosis of COVID-19 using chest X-rays images. *Biomedical Signal Processing and Control*, *76*, 103677. https://doi.org/10.1016/j.bspc.2022.103677

Liu, J., Zhang, C., & Jiang, X. (2022). Imbalanced fault diagnosis of rolling bearing using improved MsR-GAN and feature enhancement-driven CapsNet. *Mechanical Systems and Signal Processing*, *168*, 108664. https://doi.org/10.1016/j.ymssp.2021.108664

Liu, M., Wang, H., Du, L., Ji, F., & Zhang, M. (2024). Bearing-DETR: A Lightweight Deep Learning Model for Bearing Defect Detection Based on RT-DETR. *Sensors*, *24*(13), 4262. https://doi.org/10.3390/s24134262

Liu, R., Yang, B., Zio, E., & Chen, X. (2018). Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mechanical Systems and Signal Processing*, *108*, 33–47. https://doi.org/10.1016/j.ymssp.2018.02.016

Liu, X., Pang, Y., Sun, X., Liu, Y., Hou, Y., Wang, Z., & Li, X. (2024). Image Reconstruction for Accelerated MR Scan With Faster Fourier Convolutional Neural Networks. *IEEE Transactions on Image Processing*, *33*, 2966–2978. https://doi.org/10.1109/TIP.2024.3388970

Liu, Y., Chen, Y., Li, X., Zhou, X., & Wu, D. (2025). MPNet: A lightweight fault diagnosis network for rotating machinery. *Measurement*, *239*, 115498. https://doi.org/10.1016/j.measurement.2024.115498

Liu, Y., & Lai, K. W. C. (2023). The Performance Index of Convolutional Neural Network-Based Classifiers in Class Imbalance Problem. *Pattern Recognition*, *137*, 109284. https://doi.org/10.1016/j.patcog.2022.109284

Liu, Y., Liu, Y., & Ding, L. (2018). Scene Classification Based on Two-Stage Deep Feature Fusion. *IEEE Geoscience and Remote Sensing Letters*, *15*(2), 183–186. https://doi.org/10.1109/LGRS.2017.2779469

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., & Guo, B. (2021). *Swin Transformer: Hierarchical Vision Transformer using Shifted Windows* (No. arXiv:2103.14030). arXiv. http://arxiv.org/abs/2103.14030

Liu, Z., Mao, H., Wu, C.-Y., Feichtenhofer, C., Darrell, T., & Xie, S. (2022). *A ConvNet for the 2020s* (No. arXiv:2201.03545). arXiv. http://arxiv.org/abs/2201.03545

Luleci, F., Necati Catbas, F., & Avci, O. (2023). CycleGAN for undamaged-to-damaged domain translation for structural health monitoring and damage detection. *Mechanical Systems and Signal Processing*, *197*, 110370. https://doi.org/10.1016/j.ymssp.2023.110370

Ma, C., Mu, X., & Sha, D. (2019). Multi-Layers Feature Fusion of Convolutional Neural Network for Scene Classification of Remote Sensing. *IEEE Access*, *7*, 121685–121694. https://doi.org/10.1109/ACCESS.2019.2936215

Ma, J., Hu, S., Fu, J., & Chen, G. (2024). A hierarchical attention detector for bearing surface defect detection. *Expert Systems with Applications*, *239*, 122365. https://doi.org/10.1016/j.eswa.2023.122365

Ma, X., Hu, Y., Wang, M., Li, F., & Wang, Y. (2021). Degradation State Partition and Compound Fault Diagnosis of Rolling Bearing Based on Personalized Multilabel Learning. *IEEE Transactions on Instrumentation and Measurement*, *70*, 1–11. https://doi.org/10.1109/TIM.2021.3091504

Mao, W., Liu, Y., Ding, L., & Li, Y. (2019). Imbalanced Fault Diagnosis of Rolling Bearing Based on Generative Adversarial Network: A Comparative Study. *IEEE Access*, *7*, 9515–9530. https://doi.org/10.1109/ACCESS.2018.2890693

Marcelo, C., Pablo, J., & Ignacio, J. (2012). Fault Diagnosis of Induction Motors Based on FFT. In S. Salih (Ed.), *Fourier Transform—Signal Processing*. InTech. https://doi.org/10.5772/37419

Miao, Y., Zhao, M., Liang, K., & Lin, J. (2020). Application of an improved MCKDA for fault detection of wind turbine gear based on encoder signal. *Renewable Energy*, *151*, 192–203. https://doi.org/10.1016/j.renene.2019.11.012

Mirza, M., & Osindero, S. (2014). *Conditional Generative Adversarial Nets* (No. arXiv:1411.1784). arXiv. http://arxiv.org/abs/1411.1784

Nwaila, G. T., Zhang, S. E., Bourdeau, J. E., Frimmel, H. E., & Ghorbani, Y. (2024). Spatial Interpolation Using Machine Learning: From Patterns and Regularities to Block Models. *Natural Resources Research*, *33*(1), 129–161. https://doi.org/10.1007/s11053-023-10280-7

Pacheco-Chérrez, J., Fortoul-Díaz, J. A., Cortés-Santacruz, F., María Aloso-Valerdi, L., & Ibarra-Zarate, D. I. (2022). Bearing fault detection with vibration and acoustic signals: Comparison among different machine leaning classification methods. *Engineering Failure Analysis*, *139*, 106515. https://doi.org/10.1016/j.engfailanal.2022.106515

Pan, H., Li, B., Zheng, J., Tong, J., Liu, Q., & Deng, S. (2024). Research on roller bearing fault diagnosis based on robust smooth constrained matrix machine under imbalanced data. *Advanced Engineering Informatics*, *62*, 102667. https://doi.org/10.1016/j.aei.2024.102667

Pang, B., Wang, B., Sun, Z., & Hao, Z. (2024). Torsional and lateral vibration analysis of wind turbine generator bearing outer ring fault considering unbalanced magnetic pull.

Engineering Failure Analysis, *161*, 108251. https://doi.org/10.1016/j.engfailanal.2024.108251

Prudhom, A., Antonino-Daviu, J., Razik, H., & Climente-Alarcon, V. (2017). Time-frequency vibration analysis for the detection of motor damages caused by bearing currents. *Mechanical Systems and Signal Processing*, *84*, 747–762. https://doi.org/10.1016/j.ymssp.2015.12.008

Qian, L., Pan, Q., Lv, Y., & Zhao, X. (2022). Fault Detection of Bearing by Resnet Classifier with Model-Based Data Augmentation. *Machines*, *10*(7), 521. https://doi.org/10.3390/machines10070521

R, M., & Mutra, R. R. (2025). Fault classification in rotor-bearing system using advanced signal processing and machine learning techniques. *Results in Engineering*, *25*, 103892. https://doi.org/10.1016/j.rineng.2024.103892

Radford, A., Metz, L., & Chintala, S. (2016). *Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks* (No. arXiv:1511.06434). arXiv. http://arxiv.org/abs/1511.06434

Rai, A., & Upadhyay, S. H. (2016). A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings. *Tribology International*, *96*, 289–306. https://doi.org/10.1016/j.triboint.2015.12.037

Ramos, L., Casas, E., Romero, C., Rivas-Echeverría, F., & Morocho-Cayamcela, M. E. (2024). A Study of ConvNeXt Architectures for Enhanced Image Captioning. *IEEE Access*, *12*, 13711–13728. https://doi.org/10.1109/ACCESS.2024.3356551

Rao, V. V., & Ratnam, C. (2015). *A Comparative Experimental Study on Identification of Defect Severity in Rolling Element Bearings using Acoustic Emission and Vibration Analysis*. *37*(2).

Raouf, I., Kumar, P., & Soo Kim, H. (2024). Deep learning-based fault diagnosis of servo motor bearing using the attention-guided feature aggregation network. *Expert Systems with Applications*, *258*, 125137. https://doi.org/10.1016/j.eswa.2024.125137

Ren, Z., Lin, T., Feng, K., Zhu, Y., Liu, Z., & Yan, K. (2023). A Systematic Review on Imbalanced Learning Methods in Intelligent Fault Diagnosis. *IEEE Transactions on Instrumentation and Measurement*, *72*, 1–35. https://doi.org/10.1109/TIM.2023.3246470

Renaudin, L., Bonnardot, F., Musy, O., Doray, J. B., & Rémond, D. (2010). Natural roller bearing fault detection by angular measurement of true instantaneous angular speed. *Mechanical Systems and Signal Processing*, *24*(7), 1998–2011. https://doi.org/10.1016/j.ymssp.2010.05.005

Rivas, A., Delipei, G. K., Davis, I., Bhongale, S., Yang, J., & Hou, J. (2024). A component diagnostic and prognostic framework for pump bearings based on deep learning with data augmentation. *Reliability Engineering & System Safety*, *247*, 110121. https://doi.org/10.1016/j.ress.2024.110121

Schafer, R. W. (2011). On the frequency-domain properties of Savitzky-Golay filters. *2011 Digital Signal Processing and Signal Processing Education Meeting (DSP/SPE)*, 54–59. https://doi.org/10.1109/DSP-SPE.2011.5739186

Seong, G., & Kim, D. (2024). An Intelligent Ball Bearing Fault Diagnosis System Using Enhanced Rotational Characteristics on Spectrogram. *Sensors*, *24*(3), 776. https://doi.org/10.3390/s24030776

Shao, H., Jiang, H., Zhang, H., & Liang, T. (2018). Electric Locomotive Bearing Fault Diagnosis Using a Novel Convolutional Deep Belief Network. *IEEE Transactions on Industrial Electronics*, *65*(3), 2727–2736. https://doi.org/10.1109/TIE.2017.2745473

Shao, H., Li, W., Xia, M., Zhang, Y., Shen, C., Williams, D., Kennedy, A., & de Silva, C. W. (2021). Fault Diagnosis of a Rotor-Bearing System Under Variable Rotating Speeds Using Two-Stage Parameter Transfer and Infrared Thermal Images. *IEEE Transactions on Instrumentation and Measurement*, *70*, 1–11. https://doi.org/10.1109/TIM.2021.3111977

Shi, H., Cao, S., Zuo, H., Ma, J., & Lin, C. (2025). Deep subdomain adversarial network with self-supervised learning for aero-engine high speed bearing fault diagnosis with unknown working conditions. *Measurement, 241*, 115668. https://doi.org/10.1016/j.measurement.2024.115668

Shuming, Y., Changlin, X., Yuqiang, C., Biao, W., Xunyi, M., & Zinuo, W. (2025). Data-Driven Fault Diagnosis for Rolling Bearings Based on Machine Learning and Multisensor Information Fusion. *IEEE Sensors Journal*, *25*(2), 3452–3464. https://doi.org/10.1109/JSEN.2024.3499365

Singh, S., & Kumar, N. (2017). Detection of Bearing Faults in Mechanical Systems Using Stator Current Monitoring. *IEEE Transactions on Industrial Informatics*, *13*(3), 1341–1349. https://doi.org/10.1109/TII.2016.2641470

Snyder, Q., Jiang, Q., & Tripp, E. (2025). Integrating self-attention mechanisms in deep learning: A novel dual-head ensemble transformer with its application to bearing fault diagnosis. *Signal Processing, 227*, 109683. https://doi.org/10.1016/j.sigpro.2024.109683

Soomro, A. A., Muhammad, M. B., Mokhtar, A. A., Md Saad, M. H., Lashari, N., Hussain, M., Sarwar, U., & Palli, A. S. (2024). Insights into modern machine learning approaches for bearing fault classification: A systematic literature review. *Results in Engineering*, *23*, 102700. https://doi.org/10.1016/j.rineng.2024.102700

Stack, J. R., Habetler, T. G., & Harley, R. G. (2004). Fault Classification and Fault Signature Production for Rolling Element Bearings in Electric Machines. *IEEE Transactions on Industry Applications*, *40*(3), 735–739. https://doi.org/10.1109/TIA.2004.827454

Sun, B., Sheng, Z., Song, P., Sun, H., Wang, F., Sun, X., & Liu, J. (2025). State-of-the-Art Detection and Diagnosis Methods for Rolling Bearing Defects: A Comprehensive Review. *Applied Sciences*, *15*(2), 1001. https://doi.org/10.3390/app15021001

Sun, X., Ding, H., Li, N., Dong, X., Liao, Y., Liu, Z., & Bai, X. (2024). Intelligent fault diagnosis of rolling bearings under small samples based on lightweight UNet with attention-fused residual block. *Measurement Science and Technology*, *35*(7), 075002. https://doi.org/10.1088/1361-6501/ad3a08

Tandon, N., & Choudhury, A. (1999). A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings. *Tribology International*, *32*(8), 469–480. https://doi.org/10.1016/S0301-679X(99)00077-8

Tang, H., Tang, Y., Su, Y., Feng, W., Wang, B., Chen, P., & Zuo, D. (2024). Feature extraction of multi-sensors for early bearing fault diagnosis using deep learning based on minimum unscented kalman filter. *Engineering Applications of Artificial Intelligence*, *127*, 107138. https://doi.org/10.1016/j.engappai.2023.107138

Tang, S., Ma, J., Yan, Z., Zhu, Y., & Khoo, B. C. (2024). Deep transfer learning strategy in intelligent fault diagnosis of rotating machinery. *Engineering Applications of Artificial Intelligence*, *134*, 108678. https://doi.org/10.1016/j.engappai.2024.108678

Touret, T., Changenet, C., Ville, F., Lalmi, M., & Becquerelle, S. (2018). On the use of temperature for online condition monitoring of geared systems – A review. *Mechanical Systems and Signal Processing*, *101*, 197–210. https://doi.org/10.1016/j.ymssp.2017.07.044

ValizadehAslani, T., & Liang, H. (2024). *LayerNorm: A key component in parameter-efficient fine-tuning* (No. arXiv:2403.20284). arXiv. https://doi.org/10.48550/arXiv.2403.20284

Verma, A., Badal, T., & Bansal, A. (2024). Advancing Image Generation with Denoising Diffusion Probabilistic Model and ConvNeXt-V2: A novel approach for enhanced diversity and quality. *Computer Vision and Image Understanding*, *247*, 104077. https://doi.org/10.1016/j.cviu.2024.104077

Wang, B., Qin, J., Lv, L., Cheng, M., Li, L., He, J., Li, D., Xia, D., Wang, M., Ren, H., & Wang, S. (2024). DSML-UNet: Depthwise separable convolution network with multiscale large kernel for medical image segmentation. *Biomedical Signal Processing and Control*, *97*, 106731. https://doi.org/10.1016/j.bspc.2024.106731

Wang, H., & Zhang, X. (2024). Fault Diagnosis Using Imbalanced Data of Rolling Bearings Based on a Deep Migration Model. *IEEE Access*, *12*, 5517–5533. https://doi.org/10.1109/ACCESS.2024.3350785

Wang, P., Song, Y., Wang, X., & Xiang, Q. (2025). MD-BiMamba: An aero-engine inter-shaft bearing fault diagnosis method based on Mamba with modal decomposition and bidirectional features fusion strategy. *Measurement*, *242*, 115870. https://doi.org/10.1016/j.measurement.2024.115870

Wang, Y., Tang, B., Qin, Y., & Huang, T. (2020). Rolling Bearing Fault Detection of Civil Aircraft Engine Based on Adaptive Estimation of Instantaneous Angular Speed. *IEEE Transactions on Industrial Informatics*, *16*(7), 4938–4948. https://doi.org/10.1109/TII.2019.2949000

Wei, Y., Xiao, Z., Chen, X., Gu, X., & Schröder, K.-U. (2025). A bearing fault data augmentation method based on hybrid-diversity loss diffusion model and parameter transfer. *Reliability Engineering & System Safety*, *253*, 110567. https://doi.org/10.1016/j.ress.2024.110567

Wu, G., Yan, T., Yang, G., Chai, H., & Cao, C. (2022). A Review on Rolling Bearing Fault Signal Detection Methods Based on Different Sensors. *Sensors*, *22*(21), 8330. https://doi.org/10.3390/s22218330

Xia, X., Pan, X., Li, N., He, X., Ma, L., Zhang, X., & Ding, N. (2022). GAN-based anomaly detection: A review. *Neurocomputing*, *493*, 497–535. https://doi.org/10.1016/j.neucom.2021.12.093

Xiang, L., Bing, H., Li, X., & Hu, A. (2025). A frequency channel-attention based vision Transformer method for bearing fault identification across different working conditions. *Expert Systems with Applications*, *262*, 125686. https://doi.org/10.1016/j.eswa.2024.125686

Xie, S., Shen, C., Wang, D., Shi, J., Huang, W., & Zhu, Z. (2025). A new lifelong learning method based on dual distillation for bearing diagnosis with incremental fault types. *Advanced Engineering Informatics*, *65*, 103136. https://doi.org/10.1016/j.aei.2025.103136

Xie, X., Zhang, L., Wang, J., Chen, G., & Yang, Z. (2024). Adaptive minimum noise amplitude deconvolution and its application for early fault diagnosis of rolling bearings. *Applied Acoustics*, *220*, 109962. https://doi.org/10.1016/j.apacoust.2024.109962

Xing, X., Li, X., Wei, C., Zhang, Z., Liu, O., Xie, S., Chen, H., Quan, S., Wang, C., Yang, X., Jiang, X., & Shuai, J. (2024). DP-GAN+B: A lightweight generative adversarial network based on depthwise separable convolutions for generating CT volumes. *Computers in Biology and Medicine*, *174*, 108393. https://doi.org/10.1016/j.compbiomed.2024.108393

Xiong, J., Liu, M., Li, C., Cen, J., Zhang, Q., & Liu, Q. (2023). A Bearing Fault Diagnosis Method Based on Improved Mutual Dimensionless and Deep Learning. *IEEE Sensors Journal*, 1–1. https://doi.org/10.1109/JSEN.2023.3264870

Xu, J., Sun, X., Zhang, Z., Zhao, G., & Lin, J. (n.d.). *Understanding and Improving Layer Normalization*.

Xu, Z., Chen, X., Li, Y., & Xu, J. (2024). Hybrid Multimodal Feature Fusion with Multi-Sensor for Bearing Fault Diagnosis. *Sensors*, *24*(6), 1792. https://doi.org/10.3390/s24061792

Yang, J., Li, Z., Zhang, P., Zhang, K., & Xu, Y. (2024). Motor Current Time-Varying Quadratic Phase Coupling Analysis and Its Application in Traction Motor Fault Detection Under Varying-Speed Condition. *IEEE Sensors Journal*, *24*(8), 12877–12886. https://doi.org/10.1109/JSEN.2024.3371491

Yang, S., Gu, X., Liu, Y., Hao, R., & Li, S. (2020). A general multi-objective optimized wavelet filter and its applications in fault diagnosis of wheelset bearings. *Mechanical Systems and Signal Processing*, *145*, 106914. https://doi.org/10.1016/j.ymssp.2020.106914

Yang, T., Xu, M., Chen, C., Wen, J., Li, J., & Han, Q. (2025). DSTF-Net: A novel framework for intelligent diagnosis of insulated bearings in wind turbines with multi-source data and its interpretability. *Renewable Energy*, *238*, 121965. https://doi.org/10.1016/j.renene.2024.121965

Yang, Y., Yao, S., Xue, Y., Zhao, W., & Wu, C. (2024). A variable fidelity approach for predicting aerodynamic wall quantities of hypersonic vehicles using the ConvNeXt encoder-decoder framework. *Aerospace Science and Technology*, *155*, 109605. https://doi.org/10.1016/j.ast.2024.109605

Ye, M., Yan, X., Hua, X., Jiang, D., Xiang, L., & Chen, N. (2025). MRCFN: A multi-sensor residual convolutional fusion network for intelligent fault diagnosis of bearings in noisy and small sample scenarios. *Expert Systems with Applications*, *259*, 125214. https://doi.org/10.1016/j.eswa.2024.125214

Younus, A. M. D., & Yang, B.-S. (2012). Intelligent fault diagnosis of rotating machinery using infrared thermal image. *Expert Systems with Applications*, *39*(2), 2082–2091. https://doi.org/10.1016/j.eswa.2011.08.004

Yu, W., Zhou, P., Yan, S., & Wang, X. (2024). InceptionNeXt: When Inception Meets ConvNeXt. *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 5672–5683. https://doi.org/10.1109/CVPR52733.2024.00542

Yu, X., Yin, H., Sun, L., Dong, F., Yu, K., Feng, K., Zhang, Y., & Yu, W. (2023). A New Cross-Domain Bearing Fault Diagnosis Framework Based on Transferable Features and Manifold Embedded Discriminative Distribution Adaption Under Class Imbalance. *IEEE Sensors Journal*, *23*(7), 7525–7545. https://doi.org/10.1109/JSEN.2023.3248950

Zeng, Q., Feng, G., Shao, Y., Devitt, J., Gu, F., & Ball, A. (2020). An accurate instantaneous angular speed estimation method based on a dual detector setup. *Mechanical Systems and Signal Processing*, *140*, 106674. https://doi.org/10.1016/j.ymssp.2020.106674

Zhang, B., Li, H., Kong, W., Fu, M., & Ma, J. (2024). Early-Stage Fault Diagnosis of Motor Bearing Based on Kurtosis Weighting and Fusion of Current–Vibration Signals. *Sensors*, *24*(11), 3373. https://doi.org/10.3390/s24113373

Zhang, B., & Sennrich, R. (n.d.). *Root Mean Square Layer Normalization*.

Zhang, S., Zhang, S., Wang, B., & Habetler, T. G. (2020). Deep Learning Algorithms for Bearing Fault Diagnostics—A Comprehensive Review. *IEEE Access*, *8*, 29857–29881. https://doi.org/10.1109/ACCESS.2020.2972859

Zhang, T., Chen, J., Li, F., Zhang, K., Lv, H., He, S., & Xu, E. (2022). Intelligent fault diagnosis of machines with small & imbalanced data: A state-of-the-art review and possible extensions. *ISA Transactions*, *119*, 152–171. https://doi.org/10.1016/j.isatra.2021.02.042

Zhang, W., Xu, Q., Hu, Y., Xu, C., & Luo, L. (2024). Attention-Based Two-Stage Multi-Sensor Feature Fusion Method for Bearing Fault Diagnosis. *IEEE Transactions on Industry Applications*, *60*(6), 8709–8721. https://doi.org/10.1109/TIA.2024.3443232

Zhang, Z., & Wu, L. (2024). Graph neural network-based bearing fault diagnosis using Granger causality test. *Expert Systems with Applications*, *242*, 122827. https://doi.org/10.1016/j.eswa.2023.122827

Zhao, D., Cai, W., & Cui, L. (2024). Adaptive thresholding and coordinate attention-based tree-inspired network for aero-engine bearing health monitoring under strong noise. *Advanced Engineering Informatics*, *61*, 102559. https://doi.org/10.1016/j.aei.2024.102559

Zhao, M., Jia, X., Lin, J., Lei, Y., & Lee, J. (2018). Instantaneous speed jitter detection via encoder signal and its application for the diagnosis of planetary gearbox. *Mechanical Systems and Signal Processing*, *98*, 16–31. https://doi.org/10.1016/j.ymssp.2017.04.033

Zheng, Z., Song, D., Zhang, W., & Jia, C. (2025). A fault diagnosis method for bogie axle box bearing based on sound-vibration multiple signal fusion. *Applied Acoustics*, *228*, 110336. https://doi.org/10.1016/j.apacoust.2024.110336

Zhiyi, H., Haidong, S., Xiang, Z., Yu, Y., & Junsheng, C. (2020). An intelligent fault diagnosis method for rotor-bearing system using small labeled infrared thermal images and enhanced CNN transferred from CAE. *Advanced Engineering Informatics*, *46*, 101150. https://doi.org/10.1016/j.aei.2020.101150

Zhou, F., Yang, S., Fujita, H., Chen, D., & Wen, C. (2020). Deep learning fault diagnosis method based on global optimization GAN for unbalanced data. *Knowledge-Based Systems*, *187*, 104837. https://doi.org/10.1016/j.knosys.2019.07.008

Zhou, Y., Kang, X., Ren, F., Lu, H., Nakagawa, S., & Shan, X. (2024). A multi-attention and depthwise separable convolution network for medical image segmentation. *Neurocomputing*, *564*, 126970. https://doi.org/10.1016/j.neucom.2023.126970

Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2020). *Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks* (No. arXiv:1703.10593). arXiv. http://arxiv.org/abs/1703.10593

Zhu, W., Shi, B., & Feng, Z. (2023). A Transfer Learning Method Using High-Quality Pseudo Labels for Bearing Fault Diagnosis. *IEEE Transactions on Instrumentation and Measurement*, *72*, 1–11. https://doi.org/10.1109/TIM.2022.3223146