



27th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2023)

A Novel approach using WGAN-GP and Conditional WGAN-GP for Generating Artificial Thermal Images of Induction Motor Faults

Shahd Hejazi^{a,*}, Michael Packianather^a, Ying Liu^a

^a Cardiff University, School of Engineering, Cardiff CF24 3AA, United Kingdom

Abstract

This paper proposes a novel approach for generating artificial thermal images for induction motor faults using Wasserstein Generative Adversarial Network with Gradient Penalty (WGAN-GP) and Conditional Wasserstein Generative Adversarial Network with Gradient Penalty (cWGAN-GP) frameworks. Traditional fault classification methods based on vibration signals often require extensive preprocessing and are more susceptible to noise. In contrast, thermal images offer easier classification and require less preprocessing. However, challenges arise due to the limited availability of thermal images representing different fault conditions and data confidentiality. To overcome these challenges, this paper introduces the utilisation of WGAN-GP and cWGAN-GP with health condition labels to create high-quality thermal images artificially. The results demonstrate that the cWGAN-GP approach is superior in generating thermal images that closely resemble real images of induction motors under various health conditions with a Maximum Mean Discrepancy (MMD) score of 1.023 compared to 1.078 using WGAN-GP. Furthermore, cWGAN-GP requires less training time (7.25 hours to train all health conditions classes) compared to WGAN-GP (12 hours to train the Inner fault class only) using NVIDIA V100. In addition to using EMD and MMD metrics for quantitative analysis of the GAN model, the evaluation process incorporated the expertise of a pre-trained CNN model, namely AlexNet, to assess cWGAN-GP's discriminative capabilities of the generated samples and their alignment with the real thermal images, which resulted in an overall accuracy of 98.41%. Therefore, these proposed approaches offer a promising solution to address the lack of public datasets containing induction motor thermal images representing different health states. By leveraging these models, it will be feasible to enhance induction motor condition monitoring systems and improve the process of fault diagnosis.

© 2023 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

Peer-review under responsibility of the scientific committee of the 27th International Conference on Knowledge Based and Intelligent Information and Engineering Systems

* Corresponding author.

E-mail address: shahdhejazi@gmail.com

Keywords: induction motor, WGAN-GP, cWGAN-GP, artificial thermal image, condition monitoring

1. Introduction

Induction Motors (IM) are considered essential parts in various industries, and 40%-50% of IM faults are due to rolling bearings. However, bearings are used to hold elements to maintain proper IM rotation. Bearings consist of inner and outer races and some rolling balls inside a cage that keeps balls at equal distances. In fact, faults happen gradually; the earlier the fault detection, the less impact and risk are likely to occur. As faults grow, they can reduce IM capabilities, put workers' lives at risk, and alter operations quality [1]. Bearing vibration signal analysis is the traditional way of fault classification, where raw vibration signals are rarely used; hence, vibration signals need to be preprocessed using either time-domain analysis or frequency-domain analysis [2]. On the other hand, thermal images resulted in more accurate fault classification of up to 100% accuracy with less preprocessing time than vibration signal fault classification, as shown in [3] and [4]. Thermal images are more stable than vibration signals; hence, thermal images are less sensitive to speed fluctuation scenarios, making them more efficient [5].

However, thermal images have certain drawbacks. For instance, the installation cost of cameras and the potential for camera misalignment can result in an inaccurate recognition process [6]. Furthermore, the limited availability of data and imbalanced distribution of thermal images across specific or all health conditions can significantly impact the performance of condition monitoring systems [7]. To address these limitations, various oversampling techniques have been employed to generate additional samples from the minority classes. One such technique is the Synthetic Minority Oversampling Technique (SMOTE), which uses interpolation based on nearest neighbours. Another approach is the Adaptive Synthetic Sampling Technique (ADASYN) [8]. However, it's important to note that oversampling techniques can be susceptible to overfitting and noise creation, mainly when dealing with high-dimensional and sparse data. These techniques may also generate samples that are more similar to the majority class rather than the desired class [9]. Moreover, while improving classification accuracy is a common approach, it may not be effective when the degree of imbalance is high unless more data is added to the training model [10]. Additionally, the expansion of image data through the inclusion of noise and local blur can be seen as an artificial preprocessing technique. However, it is important to note that these methods may not adequately capture the diversity present in the original samples and potentially hinder fault recognition [11]. In contrast, Generative Adversarial Networks (GANs) offer a new and promising approach to sample generation. GANs provide a framework for learning complex features from high-dimensional, imbalanced, and small dataset distributions, and they have been widely utilised in fault diagnosis applications [8, 9, 10, 11].

The selection of an appropriate GAN for generating artificial images of thermal induction motor health conditions is a critical aspect of this research. Previous studies in this field have been limited, with only a few papers released recently [12]. Commonly used GAN models in fault diagnoses, such as Deep Convolutional GAN (DCGAN), Auxiliary Classifier GAN (ACGAN), Wasserstein GAN (WGAN), and variational auto-encoding GAN. However, it has been observed that the quality of data generated by the original GAN and improved DCGAN is still relatively low [10, 11]. However, the Wasserstein Generative Adversarial Network with Gradient Penalty (WGAN-GP) has demonstrated improved stability in training, prevention of mode collapse, and generation of high-quality images [13, 14, 15]. WGAN-GP has also proven its effectiveness in fault sample generation [16] and supplementing low-dimensional fault data [17]. The use of Wasserstein distance in WGAN provides a more meaningful measure of the difference between probability distributions and leads to better convergence by avoiding vanishing gradients [18]. Additionally, the training process in WGAN-GP does not require a careful balancing between the Generator and the Discriminator [19]. WGAN-GP has also been employed in the imbalance fault classification of bearings, overcoming convergence issues observed in the original GAN structures. WGAN-GP demonstrated faster convergence within 400 iterations and improved model performance compared to the original WGAN, thanks to the gradient penalty [10].

WGAN-GP was utilised to generate additional vibration signal spectra for imbalanced bearing fault classification problems, demonstrating improved convergence and faster training speed with the gradient penalty [5]. In 2022, GANs and convolutional neural networks were explored for imbalanced vibration signal datasets in induction motors, proving their efficiency. However, there was still room for improvement in utilising labelled data for induction motor fault classification, as models trained on generated data differed in accuracy compared to real data [20]. In January 2023, a paper focused on WGAN-GP for creating vibration signals in the rotor-bearing system, showing high-quality

signal generation and increased diagnostic accuracy [21]. Another paper published in the same month generated thermal images for various health conditions in rotating machinery, achieving good results but suggesting the incorporation of label information in GANs training [5].

The scarcity of induction motor datasets collected under diverse health conditions poses challenges due to data availability, confidentiality, and time constraints. While GANs have been used to generate additional tabular vibration data for condition monitoring, utilising GANs for thermal image synthesis in induction motor condition monitoring is a promising but relatively new research area. Hence, GANs are commonly employed to generate supplementary tabular vibration data, while thermal image condition monitoring offers more accurate results with minimal preprocessing steps due to their lower sensitivity to noise.

Previous studies have primarily focused on creating artificial image models for each fault type separately, leaving a research gap. This paper aims to use conditional GANs to simultaneously create artificial thermal images for different health conditions in induction motors by incorporating health conditions as a new network input. Additionally, the traditional WGAN-GP and conditional WGAN-GP approaches will be compared to evaluate their effectiveness in improving the accuracy of generated thermal images, thereby enhancing the performance of the induction motor's condition monitoring system.

The paper makes several significant contributions to the field of induction motor condition monitoring, including:

1. Generating thermal images synthetically that correspond to various health conditions using WGAN-GP.
2. Examining the effectiveness of training individual WGAN-GP models for each health state.
3. Enhancing the quality of generated images and reducing the required training time by incorporating health state labels through cWGAN-GP.
4. Comparing and contrasting the WGAN-GP and cWGAN-GP approaches using a combined assessment approach.

This paper covers the theoretical background in section 2, the methodology, including model creation and collected dataset in section 3, the results and discussion in section 4, and the conclusion is found in section 5.

2. Theoretical Background

2.1. The introduction of Generative Adversarial Networks (GANs)

GANs are machine learning models that learn the distribution of each class without explicitly separating them into distinct classes like traditional techniques such as decision trees or support vector machines. Instead, GANs focus on generating new data points (x) similar to the training data without considering the relationship between x and y , i.e., $p(x|y)$. Training GANs involve updating the parameters of the Generator (G) and Discriminator (D) using optimisation methods like stochastic gradient descent (SGD), Adam, or RMSProp. The goal is to reach a Nash equilibrium, where D is no longer able to distinguish between real images x and generated fake images ($x' = G(z)$) [22]. GANs have two probability distributions called P_g , the distribution from G's implicit distribution, and P_r , the probability distribution of real images. The Discriminator outputs a number between 0 and 1, representing the probability that the input image is real, with a score close to 1 indicating a real image. The Generator and Discriminator are continuously updated to improve the model's ability to generate data closer to real images and discriminate between real and fake data using the objective Equation(1) [10]:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim P_r(x)} [\log(D(x))] + \mathbb{E}_{z \sim P_z(z)} [\log(1 - D(G(z)))] \quad (1)$$

Equation(1) consists of two terms; G takes a noise vector z sampled from a prior distribution P_z and generates a sample $G(z)$ in the target data distribution. D takes a sample x from either the real data distribution P_r or the generated data distribution P_g (i.e., $D(G(z))$), and outputs a probability score indicating whether the input is a real or fake sample. The first term in the equation is the expected value of the logarithm of D's output on real samples x , while the second term is the expected value of the logarithm of D's output on fake samples $G(z)$ [10].

2.2. Wasserstein GAN (WGAN)

WGAN differs from other GANs by not using a sigmoid at the end of the model and using the Wasserstein distance metric (EMD) as its loss function instead of the Jensen-Shannon Divergence (JSD) used in traditional GAN

models. In WGAN, D is called a "critic" because it evaluates the quality of generated samples by assigning them a score or "criticism" rather than classifying inputs as real or fake [13]. Equation(2) consists of two parts. In the first part, the critic applies the function f to a real image x from the real probability distribution. In the second part, x is taken from G's output, which is generated from a latent noise vector, and then the critic is applied to the generated image. The critic is constrained with $\max_{\|f\|_{L \leq 1}}$, ensuring the function is Lipschitz continuous. This constraint is important for the critic to differentiate between real and generated samples. The critic estimates the Wasserstein distance between the real and generated data distributions, guiding G to generate more realistic samples. The critic aims to maximise the expression, while G aims to minimise this distance [18].

$$\max_{\|f\|_{L \leq 1}} \mathbb{E}_{x \sim P_r} [f(x)] - \mathbb{E}_{x \sim P_g} [f(x)] \quad (2)$$

2.3. Wasserstein GAN with Gradient Penalty (WGAN-GP)

WGAN-GP addresses the limitations of weight clipping in regular WGAN. Instead of weight clipping, Gradient Penalty (GP) is used to enforce the Lipschitz constraint on the critic. WGAN-GP, introduced in 2017 by Arjovsky et al. [13, 18], improves stability, resolves mode collapse, and optimises hyperparameters in training. It measures the difference between generated and real images using the Wasserstein distance metric. Additionally, the algorithm includes a gradient penalty term in the critic for smoothness. The number of generator and critic iterations, as well as the strength of the gradient penalty, can be adjusted using a lambda term [23].

2.4. Conditional WGAN-GP (cWGAN-GP)

cWGAN-GP is an extension of traditional GANs called Conditional Generative Adversarial Networks (CGAN). It introduces an additional input, denoted as y , to the network, which can represent additional information such as class names, data from another model, vectors, or images. This conditional factor adds a new dimension to the min-max game between G and D [24].

The objective function of cWGAN-GP, as shown in Equation(3), involves D outputting a high value when given real data point x conditioned on a label y drawn from the true distribution P_r [22, 24]. The second part of the objective function calculates the expected value of the logarithm of D's output when given a fake data point generated by G using a noise vector z drawn from a prior distribution $P(z)$ conditioned on the same label y [24]. The goal is to optimise G and D to minimise this objective function, generating high-quality conditional samples.

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim P_r(x)} [\log(D(x|y))] + \mathbb{E}_{z \sim P_z(z)} [\log(1 - G(z|y))] \quad (3)$$

2.5. GAN Performance Evaluation Measures: Generated Images Similarity Assessment Approaches

Evaluating GAN performance is challenging due to the lack of standardised metrics and the subjectivity of human visual evaluation, although visual similarity assessment is used in collaboration with other quantitative metrics [5, 7]. This section explores the utilisation of non-visual quantitative metrics for assessing image similarity, providing insights into the quality, diversity, and similarity of generated images using GAN compared to real images. Additionally, the section discusses the utilisation of pre-trained CNN models for non-visual image similarity assessment.

2.5.1 Generated Images Similarity Assessment: GAN Similarity Metrics

- **Fréchet Inception Distance (FID)**

FID was introduced by Heusel et al. in 2017. It measures the distance between the real distribution and the distribution generated by the trained model. It is computed using Equation(4), where μ_r and μ_g are the mean value for real images and generated images, respectively, and C_r and C_g are the covariance of the image features [7]. A lower FID indicates a better model with images closer to real ones. Hence, FID fits a Gaussian distribution to the

hidden activation of InceptionNet for each image set and computes the Fréchet Distance (also known as Wasserstein-2 distance) between the Gaussians [25].

$$FID(P_r, P_g) = \|\mu_r - \mu_g\| + \text{Tr}(C_r + C_g - 2(C_r C_g))^{1/2} \quad (4)$$

- **Maximum Mean Discrepancy (MMD)**

Measures the dissimilarities between generated and real images by capturing independent samples from each distribution. It quantifies the distance between the actual distribution and the generated distribution, with a lower score indicating better model performance [14]. Equation(5) presents the MMD score using the Gaussian kernel. P_r and P_g represents the real and generated image distribution, respectively, while x and y are samples drawn from these distributions. The first term captures the similarity of samples within the real distribution P_r , the second term measures the similarity between samples from the real and generated distributions P_r and P_g , and the third term assesses the similarity of samples within the generated distribution P_g [26].

$$MMD(P_r, P_g) = \mathbb{E}_{x, x' \sim P_r}[k(x, x')] - 2\mathbb{E}_{x, x' \sim P_r, y \sim P_g}[k(x, y)] + \mathbb{E}_{y, y' \sim P_g}[k(y, y')] \quad (5)$$

- **Earth Mover's Distance (EMD)**

Also known as the Wasserstein distance, EMD measures the distance between two probability distributions [26]. It represents the minimum amount of work or effort needed to transform one distribution into another [15]. Equation(6) defines EMD, where γ is a transport plan specifying the amount of mass to be transported from each point in P_r to each point in P_g . $\Pi(P_r, P_g)$ represents the set of all joint distributions, and $\gamma(x, y)$ indicates the amount of work needed to transform the distributions P_r into P_g from point x to point y . The Wasserstein distance is calculated as the infimum (greatest lower bound) of the expected distance $\|x - y\|$ between randomly sampled pairs of points (x, y) from γ [19].

$$W(P_r, P_g) = \inf_{\gamma \in \Pi(P_r, P_g)} \mathbb{E}_{(x, y) \sim \gamma}[\|x - y\|] \quad (6)$$

2.5.2 Generated Images Similarity Assessment: CNN Deep Learning Classification Models

Classification problems involve mapping inputs to outputs, typically achieved through supervised learning. After training a classification model, its quality of learning is evaluated by testing it on unseen data and predicting the respective classes. Alternatively, pre-trained CNN models can be used to assess the similarity between generated and real images. This involves training the pre-trained model on an artificial image dataset and evaluating its performance on a real dataset [27]. Transfer learning offers an efficient approach by leveraging knowledge from a source domain to a target domain, allowing customisation of the CNN architecture based on the selected dataset [1]. Various CNN architectures have been employed, including ResNet152V2, MobileNetV2, and AlexNet [27]. Evaluation of deep transfer learning models commonly utilises accuracy Equation(7), precision Equation(8), and recall Equation(9) metrics. These Metrics rely on the correct classification of true positives (TP) and true negatives (TN), representing accurate identification of positive and negative instances. Additionally, they consider false positives (FP) and false negatives (FN), which refer to the incorrect classification of negative and positive instances, respectively .

$$\text{Accuracy} = (TP + TN) / (TP + FN + TN + FP) \quad (7)$$

$$\text{Precision} = TP / (TP + FP) \quad (8)$$

$$\text{Recall} = TP / (TP + FN) \quad (9)$$

3. Methodology

3.1. Dataset Used

The data used in this study were captured in the Wolfson Magnetics Laboratory at Cardiff University School of Engineering, using a Forward Looking InfraRed (FLIR) thermal camera positioned 30cm from the centre of the

housing. The camera was connected to a computer to capture images of six artificially induced faults and one healthy or normal condition, as shown in Fig. 1. A total of 120 images was captured under three load types, resulting in 360 images per condition [1, 28].

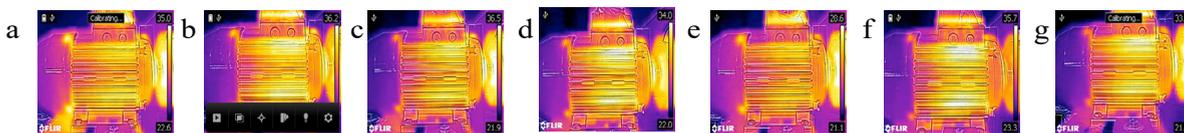


Fig. 1. Thermal images for all the faults and healthy condition: (a) 8bars; (b)Inner; (c) Outer; (d) Ball; (e) 4bars; (f) Healthy; and (g) 1bar.

3.2. Models creation

This paper has trained four WGAN-GP models for Inner, Outer, 8bars and healthy. Then all health conditions were trained together using the cWGAN-GP model to generate high-quality artificial thermal images of the induction motor under various health conditions. Two resolutions were generated: 32x32 as a baseline for quick and easy comparisons between different GAN architectures and 128x128 for higher resolution outputs. In the following sections, the proposed WGAN-GP and cWGAN-GP frameworks will be presented.

3.2.1. Proposed Wasserstein GAN with Gradient Penalty (WGAN-GP) Framework

Fig. 2 illustrates the overall WGAN-GP framework for generating thermal images of induction motors under various health conditions. The framework involves training D and G to produce realistic images. D distinguishes between real and fake images, while G generates images to fool D. The loss function is based on the Wasserstein distance between the distributions of real and fake images with a gradient penalty to control D's power. Training alternates between D and G until convergence. The Generator is a neural network that takes a 100-dimensional latent vector as input and uses transpose convolutional layers to generate images of size $C \times 128 \times 128$, where C is the number of channels. The model uses a main module consisting of several convolutional layers followed by a Tanh activation function to generate the image. The output of the main module is then passed through the Tanh function to normalise the pixel values between -1 and 1. D is a neural network with three layers of filters (256, 512, 1024), taking images of size 128x128, with C channels and outputting a single value indicating real or fake. The input image is passed through a sequential module and then flattened to be fed into a fully connected layer.

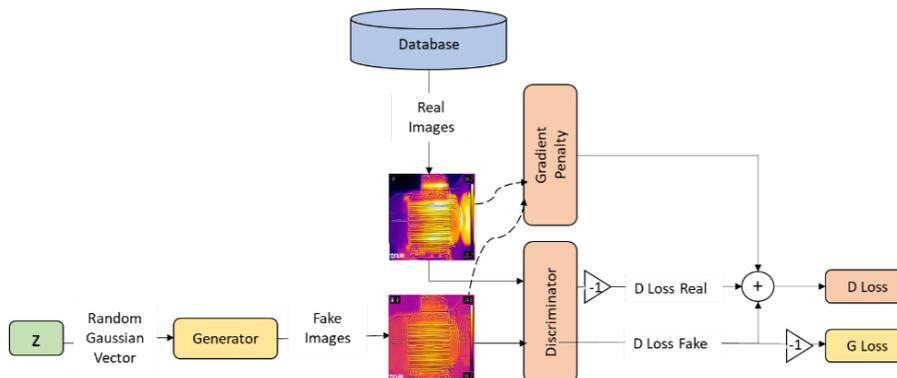
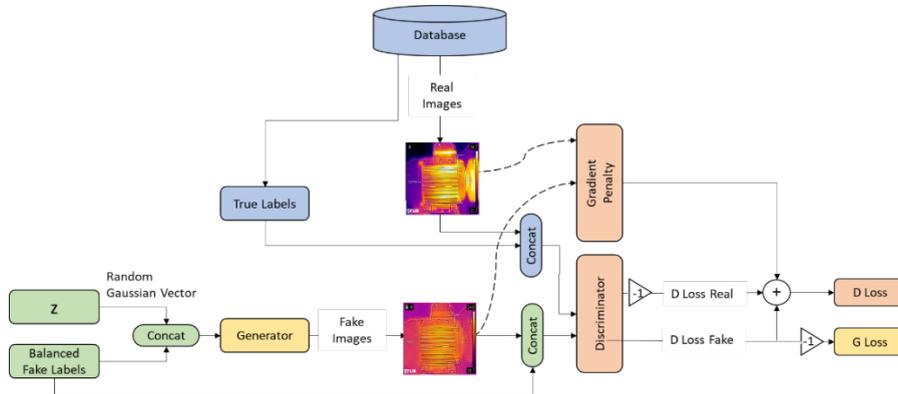


Fig. 2. The Proposed Wasserstein GAN with Gradient Penalty (WGAN-GP) Framework

3.2.2. Proposed Conditional Wasserstein GAN with Gradient Penalty (WGAN-GP)

To enhance the generation of induction motor thermal images under different health conditions, we introduce cWGAN-GP. This approach incorporates label information into the model inputs, enabling G and D networks to generate class-specific images that benefit from patterns of other classes, leading to faster convergence. Fig. 3 illustrates the cWGAN-GP framework, which includes a condition vector as input for both G and D networks, representing the image class. The loss function incorporates the Wasserstein distance and a gradient penalty term for

smoothness in D. The Generator in the cWGAN-GP framework is similar to WGAN-GP, with the addition of a one-hot encoded condition vector representing all health state classes. It takes a concatenated input of the latent vector and



condition vector, using transpose convolutional layers to generate $C \times 128 \times 128$ images. The model includes a main module with convolutional layers and a Tanh activation function. The Discriminator in cWGAN-GP is similar to WGAN-GP, with the addition of a condition vector concatenated with the input image, resulting in an input shape of (10, 128, 128) after combining the label information. The output of D is a single value indicating real or fake.

Fig. 3. The proposed Conditional Wasserstein GAN with Gradient Penalty (cWGAN-GP) Framework

4. Results and discussion

4.1 Generated Images Similarity Assessment: Visual Quality Assessment

The WGAN-GP and cWGAN-GP approaches yielded promising results in generating realistic thermal images. The cWGAN-GP approach, which incorporates class information, demonstrated further improvements in image generation, allowing for better control over the generated images. Fig. 4 showcases generated images using the cWGAN-GP approach, exhibiting a resolution of 128×128 and belonging to seven different health condition classes. Meanwhile, Fig. 5 presents generated images using the WGAN-GP approach, with a resolution of 128×128 and belonging to the healthy condition class. Visually, both sets of generated images demonstrate a high degree of variability and closely resemble real motor thermal images. However, additional quantitative assessment is needed.

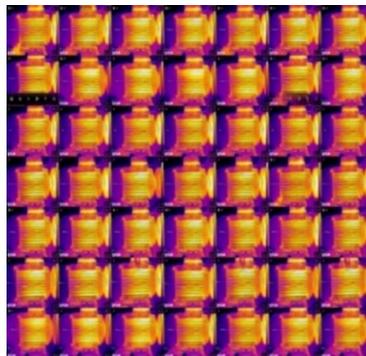


Fig. 4. Generated Images with Resolution 128×128 using cWGAN-GP. Each row represents a different health condition class from row one to row seven, representing 8 bars, Inner, Outer, Ball, 4 bars, Healthy, and 1 bar, respectively.

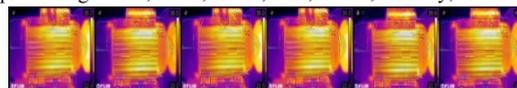


Fig. 5. Generated images (class: Healthy) with resolution 128×128 using WGAN-GP

4.2 Generated Images Similarity Assessment: GAN Similarity Metrics (MMD, EMD)

Different evaluation scores serve different purposes in assessing image generation, considering both similarities and diversities between real and generated images which was discussed in section 2.5. FID measures image similarities by utilising the Inception network to extract and compare features. However, FID scores can be misleading if the

Inception network is biased or mismatched, particularly when working with datasets lacking visual diversity or exhibiting high similarity. In contrast, EMD and MMD are robust metrics that focus on comparing distributions. They allow for evaluating the similarity between generated samples and the real data distribution, even when visual appearances may appear similar to the human eye. EMD measures the distance between probability distributions, while MMD quantifies the distance between sets of data, capturing their statistical properties [26, 29, 30, 31].

Thermal images present unique challenges for human visual perception, necessitating a comprehensive evaluation. This section compares the WGAN-GP and cWGAN-GP approaches for generating thermal images of induction motors under various health conditions. Two evaluation metrics, EMD and MMD, are used to assess the similarity between generated and real images. The experiments were conducted using NVIDIA T4 and NVIDIA V100 GPUs with different training times and epochs. The NVIDIA V100 GPU demonstrated superior performance, processing nearly three times faster than the NVIDIA T4 GPU. All experiments were performed on Google Colab Pro, utilising the allocated GPUs.

Table 1 provides a detailed performance comparison of the WGAN-GP and cWGAN-GP approaches. Initially, WGAN-GP trained models for each health condition, starting with a 32x32 resolution for the Inner fault class as a baseline. Subsequently, the scope expanded to include four health condition classes. Training duration and epochs varied for each condition, ranging from 18.5 hours for Inner and Outer faults to 36 hours for 8-bar faults. The training was terminated based on plateaued evaluation metrics and visually acceptable generated images. While training individual models resulted in high-quality 128x128 images, it required substantial time that varies per health condition; for instance, the Outer fault condition model took more than one day.

Table 1. Comparison of GPU types, training time, epochs, FID, MMD, EMD, resolution, class name, and method used for generating synthetic images. The table includes results for four different classes using WGAN-GP and all seven classes using Conditional WGAN-GP.

Method	Class Name	Resolution	EMD	MMD	Epochs	Training Time (Hours)	GPU Type
WGAN-GP	Inner	32x32	0.32	0.24	5000	4.5	NVIDIA T4
	Inner	128x128	4.64	1.10	10000	12	NVIDIA V100
	Healthy		4.70	1.07	5000	18.5	NVIDIA T4
	Outer		4.72	1.10	5000	18.5	NVIDIA T4
	8 bars		4.59	1.04	10000	36	NVIDIA T4
Average (8bars, Inner, Outer, Healthy)(128x128)			4.663	1.078			
Conditional-WGAN-GP	Inner	32x32	0.29	0.21	10000	11	NVIDIA T4
	8 bars		0.18	0.59			
	Outer		0.13	0.81			
	Ball		0.18	0.25			
	4 bars		0.15	0.66			
	Healthy		0.21	0.30			
	1bar		0.21	0.31			
	All Classes		0.12	0.09			
	Inner	128x128	4.83	1.07	10000	7.25	NVIDIA V100
	8 bars		4.78	1.02			
	Outer		4.74	1.01			
	Ball		4.80	1.26			
	4 bars		4.75	1.08			
Healthy		4.88	0.99				
1 bar		5.08	1.23				
All Classes		4.70	1.13				
Average (8bars, Inner, Outer, Healthy)(128x128)			4.816	1.023			

In contrast, the cWGAN-GP approach trained all fault types together, reducing overall training time and increasing efficiency. WGAN-GP evaluation indices represent the average of four conditions, while cWGAN-GP evaluation indices represent the average of all conditions. The EMD metric quantifies dissimilarity in terms of spatial alignment and intensity variations. WGAN-GP achieved a lower average EMD score of 4.663 for four conditions compared to cWGAN-GP's score of 4.816 for all conditions, indicating a slightly higher degree of similarity between the generated and real images in terms of spatial alignment and intensity characteristics.

The MMD metric compares the mean feature representations of real and generated image distributions. cWGAN-GP obtained a lower MMD score of 1.023, suggesting a better capture of real image characteristics, while WGAN-GP had a slightly higher score of 1.078. Thus, the cWGAN-GP approach outperforms the WGAN-GP approach in capturing the distribution and characteristics of real images. Additionally, the cWGAN-GP approach's advantage lies in training all fault types together, reducing the overall training time and increasing methodology efficiency.

Considering the better similarity scores (MMD) achieved by the cWGAN-GP approach and its reduced training time, it can be concluded that the cWGAN-GP approach is superior to the WGAN-GP approach in generating thermal images that closely resemble real images of induction motors under various health conditions while also being more efficient in terms of training time.

4.3 Generated Images Similarity Assessment: Pre-Trained AlexNet Classification

To further enhance the evaluation process, a pre-trained CNN model called AlexNet was used, using the Stochastic Gradient Descent with Momentum (SGDM) optimiser and 0.0001 Learning Rate (LR) with seven classes (health conditions) and 56.8M total learnable parameters. These parameters include weights and biases associated with the layers in the network. By leveraging the knowledge and features learned by AlexNet from large-scale image classification tasks, we can evaluate the generated samples based on their classification accuracy or other relevant metrics. This approach enables us to assess the discriminative capabilities of the generated samples and their alignment with the real data distribution. In which an artificial dataset was generated using cWGAN-GP has 288 images per health condition, then was divided randomly into 80% training and 20% validation, then tested on unseen original lab-collected images and resulted in 98.41% overall classification accuracy, 98.41% precision and 98.49% recall. However, Table 2 shows the accuracy per health condition type as follows:

Table 2. Accuracy per health condition using AlexNet

8 bars	Inner	Outer	Ball	4 bars	Healthy	1 bar
100%	95.83%	100%	100%	93.06 %	100%	100%

5. Conclusion

This study explored and compared two frameworks, namely WGAN-GP and cWGAN-GP, for generating artificial thermal images of induction motors with different health conditions. The evaluation process of comparing the similarity between the real images and the artificially created images included visual quality assessment, evaluation using GAN similarity metrics (MMD and EMD), and classification using a pre-trained AlexNet model. Both approaches produced high-quality thermal images that closely resembled real motor images when visually and qualitatively evaluated. Quantitatively, the generated images were evaluated using two similarity metrics, EMD and MMD. While WGAN-GP achieved a slightly better EMD score of 4.663 for four conditions compared to cWGAN-GP's score of 4.816 for all conditions, cWGAN-GP obtained a lower MMD score of 1.023, indicating a closer resemblance to real images in terms of statistical properties. This suggests that the generated images from cWGAN-GP exhibit similar texture, shape, and overall distribution as observed in the real images. To further validate the generated images, a pre-trained AlexNet model was utilised for classification on the cWGAN-GP dataset, which achieved an overall classification accuracy of 98.41% as well as higher accuracy rates for some health conditions. In conclusion, the cWGAN-GP approach proved to be superior in generating thermal images that closely resemble real images of induction motors with various health conditions. Its ability to incorporate class information facilitated faster convergence, pattern recognition, and diversity in image generation. The proposed approach achieved a higher similarity MMD score, reduced training time, and demonstrated high classification accuracy on real datasets, whereby highlighting its effectiveness and efficiency. These findings have contributed to the field of thermal image generation and demonstrate the potential for applications in motor condition monitoring and fault diagnosis. Future research will prioritise practical applications and advancements in fault diagnosis in real-world scenarios. Specifically, it will explore adapting the models to handle noise in thermal images, like camera rotation, and utilising data augmentation techniques to generate scenario-specific thermal images. These efforts, which aim to streamline condition monitoring systems to enhance fault detection and diagnosis, will offer reliable alternatives or support to vibration-based methods in induction motor condition monitoring.

Acknowledgements

This work was supported by the Saudi Arabian Cultural Bureau in London, the Saudi Arabian Ministry of Education, and Cardiff University.

References

- [1] S. Hejazi, M. Packianather, and Y. Liu, “Novel Preprocessing of Multimodal Condition Monitoring Data for Classifying Induction Motor Faults Using Deep Learning Methods,” in *2022 IEEE 2nd International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC)*, Dec. 2022, pp. 1–6. doi: 10.1109/iSSSC56467.2022.10051321.
- [2] V. Sinitin, O. Ibrayeva, V. Sakovskaya, and V. Eremeeva, “Intelligent bearing fault diagnosis method combining mixed input and hybrid CNN-MLP model,” *Mech. Syst. Signal Process.*, vol. 180, no. June, p. 109454, Nov. 2022, doi: 10.1016/j.ymsp.2022.109454.
- [3] A. Choudhary, T. Mian, and S. Fatima, “Convolutional neural network based bearing fault diagnosis of rotating machine using thermal images,” *Measurement*, vol. 176, no. February, p. 109196, May 2021, doi: 10.1016/j.measurement.2021.109196.
- [4] M. Khanjani and M. Ezoji, “Electrical fault detection in three-phase induction motor using deep network-based features of thermograms,” *Measurement*, vol. 173, no. July 2020, p. 108622, Mar. 2021, doi: 10.1016/j.measurement.2020.108622.
- [5] H. Shao, W. Li, B. Cai, J. Wan, Y. Xiao, and S. Yan, “Dual-Threshold Attention-Guided Gan and Limited Infrared Thermal Images for Rotating Machinery Fault Diagnosis Under Speed Fluctuation,” *IEEE Trans. Ind. Informatics*, pp. 1–10, 2023, doi: 10.1109/TII.2022.3232766.
- [6] P. Gangsar and R. Tiwari, “Signal based condition monitoring techniques for fault detection and diagnosis of induction motors: A state-of-the-art review,” *Mech. Syst. Signal Process.*, vol. 144, p. 106908, 2020, doi: 10.1016/j.ymsp.2020.106908.
- [7] S. Niu, B. Li, X. Wang, and H. Lin, “Defect Image Sample Generation With GAN for Improving Defect Recognition,” *IEEE Trans. Autom. Sci. Eng.*, vol. 17, no. 3, pp. 1–12, 2020, doi: 10.1109/TASE.2020.2967415.
- [8] X. Liu, T. Li, R. Zhang, D. Wu, Y. Liu, and Z. Yang, “A GAN and Feature Selection-Based Oversampling Technique for Intrusion Detection,” *Secur. Commun. Networks*, vol. 2021, 2021, doi: 10.1155/2021/9947059.
- [9] J. Engelmann and S. Lessmann, “Conditional Wasserstein GAN-based oversampling of tabular data for imbalanced learning,” *Expert Syst. Appl.*, vol. 174, no. ML, 2021, doi: 10.1016/j.eswa.2021.114582.
- [10] B. Han, S. Jia, G. Liu, and J. Wang, “Imbalanced Fault Classification of Bearing via Wasserstein Generative Adversarial Networks with Gradient Penalty,” *Shock Vib.*, vol. 2020, 2020, doi: 10.1155/2020/8836477.
- [11] H. Fan, J. Ma, X. Zhang, C. Xue, Y. Yan, and N. Ma, “Intelligent data expansion approach of vibration gray texture images of rolling bearing based on improved WGAN-GP,” *Adv. Mech. Eng.*, vol. 14, no. 3, pp. 1–11, 2022, doi: 10.1177/16878132221086132.
- [12] W. Wu, K. Cao, C. Li, C. Qian, C. Change, and L. Reed, “TransGaGa : Geometry-Aware Unsupervised Image-to-Image Translation,” 2019.
- [13] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, “Improved Training of Wasserstein GANs,” Mar. 2017, [Online]. Available: <http://arxiv.org/abs/1704.00028>
- [14] Z. Pan, W. Yu, X. Yi, A. Khan, F. Yuan, and Y. Zheng, “Recent Progress on Generative Adversarial Networks (GANs): A Survey,” *IEEE Access*, vol. 7, pp. 36322–36333, 2019, doi: 10.1109/ACCESS.2019.2905015.
- [15] X. Gao, F. Deng, and X. Yue, “Data augmentation in fault diagnosis based on the Wasserstein generative adversarial network with gradient penalty,” *Neurocomputing*, vol. 396, no. xxxx, pp. 487–494, Jul. 2020, doi: 10.1016/j.neucom.2018.10.109.
- [16] R. Wang, S. Zhang, Z. Chen, and W. Li, “Enhanced generative adversarial network for extremely imbalanced fault diagnosis of rotating machine,” *Meas. J. Int. Meas. Confed.*, vol. 180, no. April, p. 109467, 2021, doi: 10.1016/j.measurement.2021.109467.
- [17] H. Zhong, S. Yu, H. Trinh, Y. Lv, R. Yuan, and Y. Wang, “Fine-tuning Transfer Learning based on DCGAN Integrated with Self-attention and Spectral Normalization for Bearing Fault Diagnosis,” *Measurement*, vol. 210, no. January, p. 112421, 2023, doi: 10.1016/j.measurement.2022.112421.
- [18] M. Arjovsky, S. Chintala, and L. Bottou, “Wasserstein GAN,” 2017, [Online]. Available: <http://arxiv.org/abs/1701.07875>
- [19] M. Arjovsky, S. Chintala, and L. Bottou, “Wasserstein generative adversarial networks,” *34th Int. Conf. Mach. Learn. ICML 2017*, vol. 1, pp. 298–321, 2017.
- [20] H. C. Chang, Y. C. Wang, Y. Y. Shih, and C. C. Kuo, “Fault Diagnosis of Induction Motors with Imbalanced Data Using Deep Convolutional Generative Adversarial Network,” *Appl. Sci.*, vol. 12, no. 8, 2022, doi: 10.3390/app12084080.
- [21] J. Ma, X. Jiang, B. Han, J. Wang, Z. Zhang, and H. Bao, “applied sciences Dynamic Simulation Model-Driven Fault Diagnosis Method for Bearing under Missing Fault-Type Samples,” 2023.
- [22] Y. Pang, J. Lin, T. Qin, and Z. Chen, “Image-to-Image Translation: Methods and Applications,” *IEEE Trans. Multimed.*, vol. 24, pp. 3859–3881, 2022, doi: 10.1109/TMM.2021.3109419.
- [23] J. Wang and C. Wang, “High Performance WGAN-GP based Multiple- category Network Anomaly Classification System”.
- [24] X. He, Z. Chang, L. Zhang, H. Xu, H. Chen, and Z. Luo, “A Survey of Defect Detection Applications Based on Generative Adversarial Networks,” *IEEE Access*, vol. 10, no. September, pp. 113493–113512, 2022, doi: 10.1109/ACCESS.2022.3217227.
- [25] R. Chen, W. Huang, B. Huang, F. Sun, and B. Fang, “Reusing Discriminators for Encoding: Towards Unsupervised Image-To-Image Translation,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 8165–8174, 2020, doi: 10.1109/CVPR42600.2020.00819.
- [26] A. Borji, “Pros and cons of GAN evaluation measures,” *Comput. Vis. Image Underst.*, vol. 179, pp. 41–65, 2019, doi: 10.1016/j.cviu.2018.10.009.
- [27] H. H. N. Alrashedy, A. F. Almansour, D. M. Ibrahim, and M. A. A. Hammoudeh, “BrainGAN: Brain MRI Image Generation and Classification Framework Using GAN Architectures and CNN Models,” *Sensors*, vol. 22, no. 11, 2022, doi: 10.3390/s22114297.
- [28] A. K. Al-Musawi, F. Anayi, and M. Packianather, “Three-phase induction motor fault detection based on thermal image segmentation,” *Infrared Phys. Technol.*, vol. 104, p. 103140, Jan. 2020, doi: 10.1016/j.infrared.2019.103140.

- [29]H. Alqahtani, M. Kavakli-Thorne, and G. Kumar, “An Analysis of Evaluation Metrics of GANs,” *Int. Conf. Inf. Technol. Appl.*, vol. 7, no. July, 2019.
- [30]M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, “GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium,” *Adv. Neural Inf. Process. Syst.*, vol. 2017-Decem, no. Nips, pp. 6627–6638, Jun. 2017, [Online]. Available: <http://arxiv.org/abs/1706.08500>
- [31]Y. Chen, X. Xu, and J. Jia, “Domain Adaptive Image-to-Image Translation,” in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2020, pp. 5273–5282. doi: 10.1109/CVPR42600.2020.00532.