International Journal of Advanced Multidisciplinary Application | IJAMA
Volume 1 Issue 4 Dec 2024
ISSN No: 3048-9350

Utilizing Thermography and Convolutional Neural Networks for the Detection of Mechanical Faults in Induction Motors and Gearbox Wear

A. Abdulrazak Gurnah[1], S. Shafi Adam Shafi[2], N. Nasra J. A. Juma[3]
[1,2,3]Department of Electrical Engineering
[1,2,3]University of Dar es Salaam, Tanzania

Abstract- Induction motors and gearboxes are critical components in modern industries, serving as essential tools for the operation of numerous machines. This study presents a diagnostic approach for identifying various faults in an electromechanical system using infrared thermography and a convolutional neural network (CNN). The experiments involved testing the motor and gearbox under different conditions. The induction motor was evaluated in four states: healthy, with a broken bar, a damaged bearing, and misalignment. The gearbox was assessed under three conditions: healthy gears, 50% wear, and 75% wear. Faults were introduced through controlled machining operations. Data augmentation techniques, such as mirroring and brightness variation, were applied to enhance the dataset. Ablation studies were conducted, and a CNN with a simplified architecture was designed. The model achieved a precision of 98.53%, accuracy of 98.54%, recall of 98.65%, and F1-Score of 98.55%. These results demonstrate that the combination of infrared thermography and deep learning effectively detects faults across multiple components of an electromechanical system.

Keywords: thermography, convolutional neural networks, induction motor faults, gearbox wear, multi-fault diagnosis

1. Introduction

Electromechanical systems, consisting of both mechanical and electrical components, are essential in various modern industries. Their applications span across sectors such as manufacturing, the electrical industry, process automation, and the automotive industry, among others [1,2]. These systems integrate components like pulleys, belts, shafts, and mechanical couplings, all of which contribute to their functional versatility. However, induction motors (IMs) and gearboxes (GBs) are the most prominent components in industrial electromechanical systems due to their unique ability to convert electrical energy into mechanical energy while managing torque transmission efficiently [3].

The widespread adoption of IMs and GBs is attributed to their numerous advantages, including affordability, robust mechanical design, adaptability to variable load conditions, versatility, low maintenance requirements, efficiency, and reliability in harsh environments [4]. Despite these benefits, the operational lifespan of these components is often marred by unexpected faults, leading to unscheduled downtimes and significant disruptions in industrial processes. Faults in IMs are typically categorized into electrical and mechanical stresses. Electrical stresses often arise from power supply issues, resulting in stator faults such as short circuits in the stator winding. On the other hand, mechanical stresses, often caused by overloads, can lead to bearing defects, broken rotor bars, imbalance, and misalignment [5].

Similarly, GBs are prone to severe faults, including broken or chipped gear teeth, as well as non-severe failures such as pitting, tearing, and wear, all of which can compromise their performance [6]. Faults in either IMs or GBs can trigger system-wide breakdowns, halting machinery operations and incurring economic losses due to production downtime [2]. This underscores the importance of developing and

implementing robust strategies for continuous condition monitoring to ensure the reliability and efficiency of industrial electromechanical systems.

Condition monitoring has become a pivotal technique for engineers and researchers, enabling the early detection of faults in electromechanical systems. By measuring physical parameters such as vibrations, acoustic emissions, electric currents, stray flux, and thermal variations, it is possible to assess system behavior and identify faults using advanced signal processing techniques [3]. Methods like Wavelet Transform (WT), Fast Fourier Transform (FFT), Empirical Mode Decomposition (EMD), Variational Mode Decomposition (VMD), and Motor Current Signature Analysis (MCSA) have been widely applied for analyzing electric current, vibration, sound, and flux signals [10–12].

While these traditional methods are effective, they often require the expertise of specialists to interpret results accurately. Misinterpretation of these results can lead to errors, especially when designing automated fault diagnosis systems. In this context, infrared thermography (IRT) has emerged as a non-invasive and non-destructive alternative with vast potential for condition monitoring. When properly implemented, IRT offers a promising solution for automating the fault detection process in electromechanical systems [13].

This research explores the integration of IRT with modern deep learning techniques, specifically convolutional neural networks (CNNs), to develop a more efficient and accurate approach to fault diagnosis in IMs and GBs. By leveraging the capabilities of IRT and CNNs, this study aims to address the limitations of traditional methods and provide a reliable, automated solution for monitoring and diagnosing faults in industrial electromechanical systems.

Infrared Thermography (IRT) is increasingly being utilized for the condition monitoring of electromechanical systems, providing a technologically advanced approach to detect and address faults in components like induction motors (IM) and gearboxes (GB). Traditional IRT methods often involve manual interpretation of thermographic images, either through direct visual inspection or with the aid of commercial software. However, these methods are heavily reliant on the operator's experience and expertise, posing challenges such as extended time requirements and potential for human error.

Addressing Limitations in Current Methodologies for Fault Detection

Most methodologies in the existing literature focus narrowly on analyzing single elements of electromechanical systems using infrared thermography (IRT). For instance, studies involving induction motors (IMs) often target specific components, such as the stator winding, bearings, or broken rotor bars, without providing a comprehensive analysis of the entire system [17–24]. Similarly, research on gearboxes (GBs) primarily emphasizes individual fault conditions, such as gear wear, broken gear teeth, or cracks [6]. These focused approaches, while effective in their limited scope, fail to harness the full potential of the implemented techniques.

The reliance on complex methodologies tailored for detecting a single fault condition results in significant inefficiencies, including underutilization of computational resources and failure to address the multi-fault dynamics of an electromechanical system. Moreover, such methodologies typically impose a heavy computational load due to their intricate designs, reducing their practical applicability in real-world industrial settings.

A notable gap in the literature is the lack of methodologies that investigate combined faults across key components like IMs and GBs. Failures in one component can propagate through the electromechanical system, potentially causing cascading issues. Despite this, most studies focus on isolated faults, ignoring the interconnected nature of these systems. Additionally, while deep learning approaches are increasingly employed, many studies fail to incorporate strategies for augmenting datasets or optimizing algorithm design. This shortfall becomes particularly problematic when dealing with noisy thermal imaging data, a common challenge in industrial environments.

Contribution of the Study

To address these gaps, this paper proposes a novel, non-invasive methodology that integrates convolutional neural networks (CNNs) with IRT to detect and classify multiple faults in electromechanical systems. The methodology emphasizes the detection of faults in both IMs and GBs and extends its scope to explore combined faults. An experimental study was conducted using thermal images obtained under various fault scenarios:

- Induction Motor (IM): Faults analyzed include healthy state (HLT), outer bearing fault (OB), broken bar (BB), and misalignment (MAL).
- Gearbox (GB): Conditions studied include healthy gears and gear wear at 50% and 75%.
- Combined Faults: Fault scenarios integrating both IM and GB conditions to simulate real-world multi-fault scenarios.

To enhance the dataset and improve the robustness of the methodology, noise was deliberately introduced into the thermographic images. Pixel intensity levels were modified, and salt-and-pepper noise was added to simulate real-world conditions. A specialized CNN architecture was designed to process noisy thermal images effectively, achieving remarkable results with an accuracy of 98.5% in classifying multi-fault scenarios.

Implications

The proposed methodology offers several advantages:

1. Non-invasive Monitoring: It allows for contactless fault detection, making it suitable for harsh industrial environments.
2. Comprehensive Fault Analysis: By addressing both IM and GB faults, as well as their combined effects, the methodology provides a holistic view of electromechanical system health.
3. Noise Robustness: The ability to handle noisy data ensures reliable performance in real-world conditions.
4. Operational Benefits: Early and accurate fault detection reduces the risk of progressive damage, enhances safety, maintains process quality, and minimizes downtime.

This approach has the potential to revolutionize fault diagnosis in electromechanical systems by enabling timely and reliable condition monitoring. By identifying internal faults that are difficult to detect using traditional methods, it directly addresses challenges associated with operational inefficiencies and safety risks, ultimately contributing to improved industrial productivity and reduced maintenance costs.

Infrared Thermography (IRT) has emerged as a powerful tool for condition monitoring in electromechanical systems, offering advanced capabilities to detect and diagnose faults in components like induction motors (IM) and gearboxes (GB). Traditional IRT methods rely on manual interpretation of thermographic images, either through direct analysis or with the assistance of commercial software. These methods, while widely used, are time-intensive and heavily dependent on the expertise of the operator, which increases the likelihood of human error. To address these limitations, semi-automatic techniques have been developed, leveraging image processing algorithms to identify regions of interest. However, these approaches still require human interpretation for diagnosis, demanding significant training and experience.

The integration of artificial intelligence (AI) has significantly enhanced IRT-based methodologies by enabling the automatic detection, extraction, and classification of faults, minimizing the need for human

intervention. Advanced image processing techniques, such as thresholding, edge detection, and segmentation, are combined with feature extraction methods like Scale-Invariant Feature Transform (SIFT), Principal Component Analysis (PCA), and Linear Discriminant Analysis (LDA) to analyze fault conditions effectively. Furthermore, machine learning algorithms like Decision Trees, Support Vector Machines (SVM), and k-means clustering, along with deep learning models such as Convolutional Neural Networks (CNNs), have demonstrated superior accuracy in fault diagnosis.

Recent innovations have shown remarkable success. For instance, CNNs have been employed to extract fault features directly from thermographic images, achieving detection accuracies of 90–98% in bearings and other rotating machinery. Techniques like the Power of Normalized Image Difference (PNID), used with deep neural networks, have enabled precise analysis of faults in brushless direct current (BLDC) motors. Despite these advancements, challenges remain. Current methodologies often lack comprehensive analyses of multiple faults, and data augmentation techniques, crucial for training robust deep learning models, are underutilized. Additionally, noise handling in thermal images is an underexplored area that could further improve diagnostic accuracy.

In conclusion, while IRT combined with AI has revolutionized fault detection in electromechanical systems, addressing issues such as noise corruption, multiple fault analysis, and data augmentation can enhance these methodologies further. Convolutional Neural Networks (CNNs), in particular, stand out as a transformative approach, offering precise, automated, and efficient fault classification capabilities.

2. Proposed Methodology

The proposed methodology for detecting multiple mechanical faults in induction motors (IM) and gearboxes (GB) is depicted in Figure 1. This approach consists of five interconnected stages aimed at ensuring precise and robust fault diagnosis. The process begins with the creation of a thermographic image database generated through physical experiments. These images serve as the foundation for subsequent analysis. In the second stage, the thermographic images are cropped to isolate specific regions of interest, such as the gearbox, coupling, and induction motor, while disregarding extraneous elements like the background or unrelated components. This refinement focuses the analysis on critical areas of the system. The third stage involves data augmentation to enhance the diversity and size of the image dataset. Techniques like horizontal flipping and intensity variations are applied to improve the CNN's generalization capabilities and overall performance. The fourth stage is centered on the design and testing of the CNN architecture. Ablation tests are conducted to determine the optimal configuration, and noise is introduced into the thermographic images to evaluate the robustness of the proposed model. Filtering techniques are employed to mitigate the impact of this noise. In the final stage, the methodology enables the diagnosis and classification of faults into twelve distinct classes, accounting for various combinations of faults in both the IM and GB.

The development of the thermographic image database was a critical aspect of this research. Images were acquired from an experimental test bench comprising an electromechanical system. This setup included a 1.5 kW WEG induction motor mechanically coupled to a BALDOR 4:1 gearbox via a rigid coupling, which in turn was connected to a BALDOR direct current generator. The generator simulated approximately 20% of the motor's rated load under working conditions. The experiments spanned 80 minutes, sufficient time for the motor to reach a steady state, with images captured every minute. In total, 80 images were collected for each test condition. A FLIR GF320 thermographic camera, featuring a resolution of 320 × 240 pixels, a thermal sensitivity of <15 mK at 30°C, and a spectral range of 3.2–3.4 µm, was employed for image acquisition. Additionally, a Fluke 975 air quality meter monitored the environmental conditions of the test room, while the emissivity setting of 0.95 ensured accurate temperature readings. All images were captured in grayscale to standardize data and highlight thermal

variations caused by system faults. This comprehensive and meticulous data acquisition process established a robust foundation for implementing the proposed methodology.

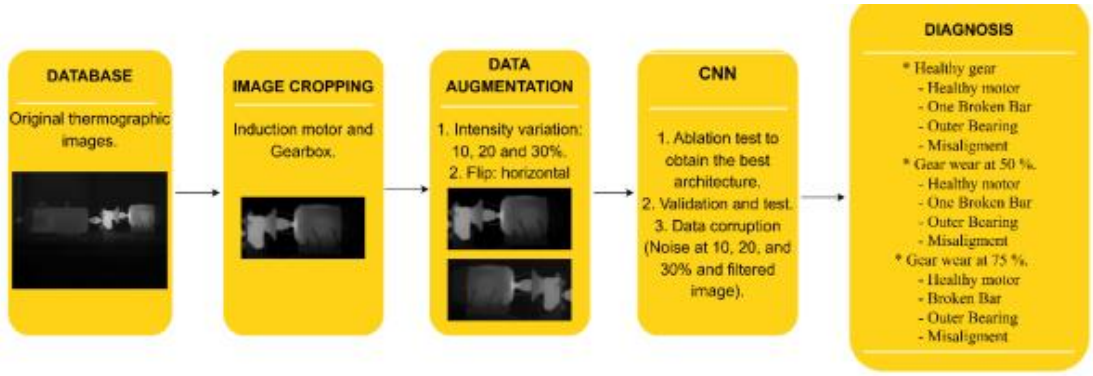

Figure 1. The proposed non-invasive methodology for the detection of multiple failures in electromechanical systems.

2.1. Database: In the database stage, a comprehensive acquisition of thermographic images was undertaken to construct a robust experimental database. This work utilized an experimental test bench designed around an electromechanical system to generate the required data. The setup featured a 1.5 kW WEG induction motor mechanically linked to a BALDOR 4:1 gearbox through a rigid coupling. Further, the gearbox was connected to a BALDOR direct current generator using another rigid coupling. The generator simulated a load approximately 20% of the rated capacity of the induction motor under operational conditions.

The testing phase spanned eighty minutes, aligning with the estimated time required for the motor to reach a steady-state condition. Throughout the duration of the test, thermographic images were captured at one-minute intervals, resulting in a total of 80 images per test condition. The thermographic images were obtained using a FLIR GF320 camera equipped with a resolution of 320 × 240 pixels, a thermal sensitivity of less than 15 mK at 30°C, and a spectral operating range of 3.2–3.4 µm. All images were captured in grayscale mode to enhance consistency and highlight temperature variations within the system.

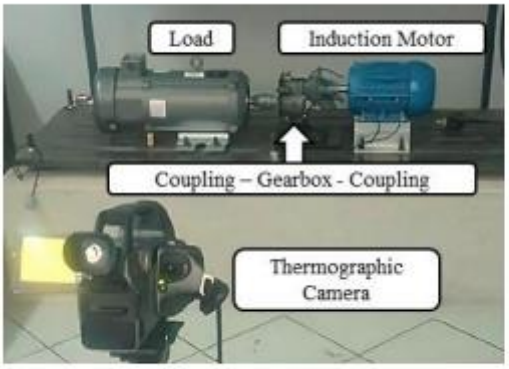

Figure 2. Experimental test bench—electromechanical system.

To ensure the accuracy of environmental conditions during data acquisition, a Fluke 975 air quality meter was employed to monitor the testing environment. Furthermore, an emissivity value of 0.95 was used to optimize the precision of temperature readings in the captured thermographic images. This systematic and controlled approach to image acquisition forms a foundational step in the proposed methodology, providing a reliable and high-quality dataset for subsequent analysis and fault detection.

2.2. Image Cropping and Data Augmentation: According to Bu et al. [29], the quantity of images or data directly influences the convergence of CNN classification models. As a result, offline techniques are utilized to enhance the database, encompassing suggested approaches such as rotation, inversion, and contrast enhancement. Furthermore, Fonseca et al. [30] proposed data augmentation through geometric transformations like horizontal and vertical flips to expand the database while preserving the inherent characteristics of the original images. An additional method employed to enlarge the database involved modifying the initial images using slight amounts of salt and pepper noise. Furthermore, there are documented studies where researchers have examined CNN's resilience by employing this approach [31]. This paper presents an initial database containing 600 thermal images of size 320 × 240 pixels; these images were acquired during the experimental setup; there are 50 images for each class or the different conditions of the induction motor. Initially, the region of interest of the original images was selected, and the image was cropped to a size of 140 × 70 pixels, which is the size required to have the complete image of the motor and gearbox. The eliminated information corresponds to the background and the load, which does not affect the objective of this work. On the contrary, this reduction in the image area allows less information to be processed, resulting in a lower computational load and faster training and inference time. To increase the database by 200 images for each test listed in Table 1, pre-processing was performed by changing the intensity levels in each original image by 10%, 20%, and 30%, resulting in a database of n = 2400. Changes in intensity levels are one of the techniques used as they help improve the training of the network, besides being one of the main scenarios when acquiring thermographic images [32]. Then, a horizontal flipping transformation was applied to the resulting database, obtaining a total of 4800 images. The optimal CNN configuration to achieve the expected results was determined from this database.

2.3. CNN: Architecture A convolutional neural network (CNN) is an architecture through which learning and pattern recognition are performed on an input image (Figure 4), allowing the identification of differences among various images [33]. The primary structure of a CNN comprises four main layers: the convolutional layer, the pooling layer, the fully connected layer, and the SoftMax layer [34]. Appl. Syst. Innov. 2024, 7, x FOR PEER REVIEW 8 of 20 2.3. CNN Architecture A convolutional neural network (CNN) is an architecture through which learning and pattern recognition are performed on an input image (Figure 4), allowing the identification of differences among various images [33]. The primary structure of a CNN comprises four main layers: the convolutional layer, the pooling layer, the fully connected layer, and the SoftM

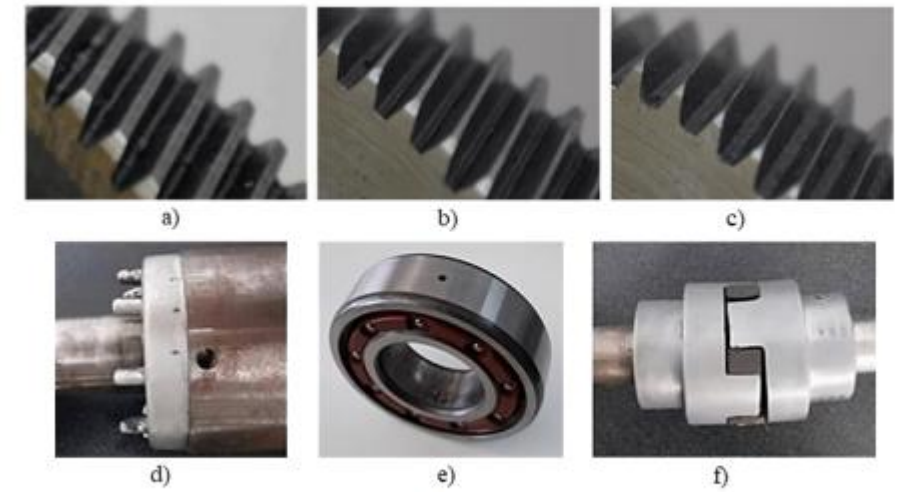

Figure 3. Mechanical elements. (**a**) Healthy gear; (**b**) gear with 50% wear; (**c**) gear with 75% wear; (**d**) one broken bar; (**e**) outer bearing; and (**f**) misalignment.

The convolutional layer receives the image as input and applies convolutional processes. Several filters, known as kernels, are applied. These filters are smaller than the original image and traverse the image through all pixels lengthwise and widthwise to create an activation map. This layer extracts features and patterns from the input images using the kernels and is responsible for capturing features such as color, gradient direction, etc. [35]. The pooling layer plays a crucial role in reducing the dimensions of the feature map and the parameters that were extracted from the preceding layer. This reduction helps prevent overfitting in the network. There are two primary types of pooling layers: the maximum pooling layer and the average pooling layer [36]. In this paper, both types of Appl. Syst. Innov. 2024, 7, 123 8 of 19 layers were implemented to obtain the optimal CNN configuration. The fully connected layer is responsible for differentiating the features represented by the activation layers and identifying the input class. Each of these layers contains perceptrons, with the number of perceptrons being dependent on the size of the input images after the convolution and pooling layers [37]. The confusion matrix is used to assess the performance of CNN. The primary objective of this technique is to evaluate the CNN model based on the metrics obtained from both correct predictions and the errors identified during the classification. It not only provides information about the errors but also categorizes the types of errors made by the classifier [34]

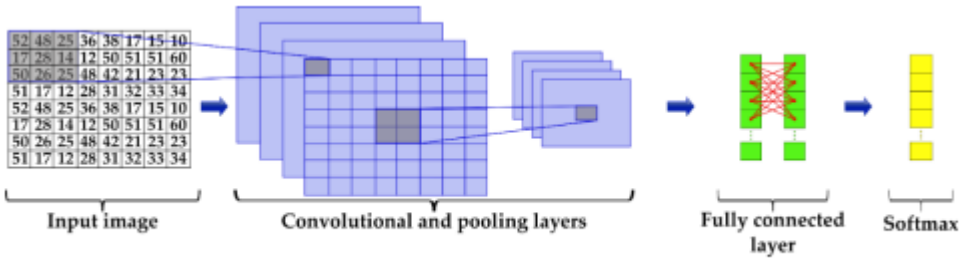

Figure 4. Structure of a CNN.

$$Accuracy(\%) = \frac{(TP + TN)}{(TP + TN + FN + FP)} \times 100 \quad (1)$$

$$Precision(\%) = \frac{TP}{(TP + FP)} \times 100 \quad (2)$$

$$Recall(\%) = \frac{TP}{(TP + FN)} \times 100 \quad (3)$$

$$F1 - Score(\%) = \left(\frac{2 \times (Recall \times Precision)}{(Recall + Precision)}\right) \times 100 \quad (4)$$

3. Results: The experiment was carried out using the final database as input images and following the main structure of the CNN. The following parameters were modified: filter sizes (3 × 3 and 4 × 4), the number of filters (8 and 16), average and maximum pooling layers, and the number of epochs (10 and 20). This process aimed to obtain the optimal CNN configuration for classifying and identifying types of induction motor and gearbox faults. The results are categorized into three case studies or stages, as outlined in Section 2. The first is to configure and test CNN only with the database augmented. The second is to configure and test CNN with the database corrupted by 2% salt and pepper noise. And the third uses the same configuration as the second, but with images corrupted with salt and pepper noise at 10%, 20%, and 30%, as well as noise-filtered images for the maximum noise condition, i.e., 30%. The total database, consisting of n = 4800 images, is divided as follows: 60% for training (2880 images), 20% for validation (960 images), and 20% for testing (960 images). Lastly, to evaluate the robustness of CNN, salt and pepper noise was introduced into the database. This encompassed both the original

images and images with varying levels of intensity. The total image count in the database remained at n = 4800. Appl. Syst. Innov. 2024, 7, 123 9 of 19 3.1. Configuring CNN with the Augmented Database The CNN was trained with 2880 randomly shaped images, which included original images with varying intensity levels and horizontal flips. These images encompassed the 12 different motor and gearbox conditions. Table 2 presents the recorded results (accuracy) derived from the extensive experimentation involving the diverse configurations of the CNN.

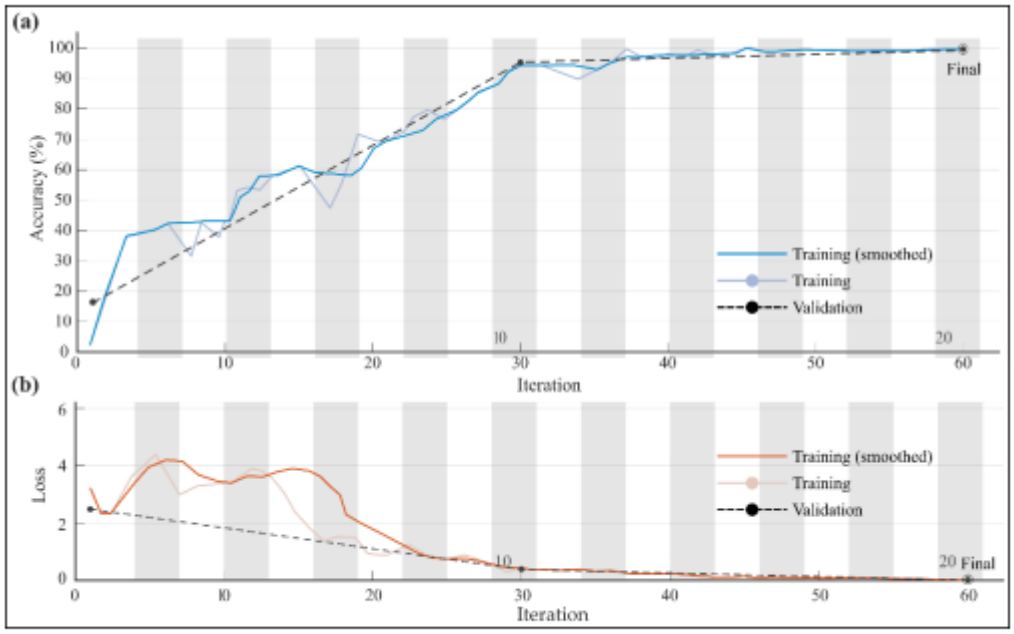

Figure 5. Training process for images with horizontal flip.

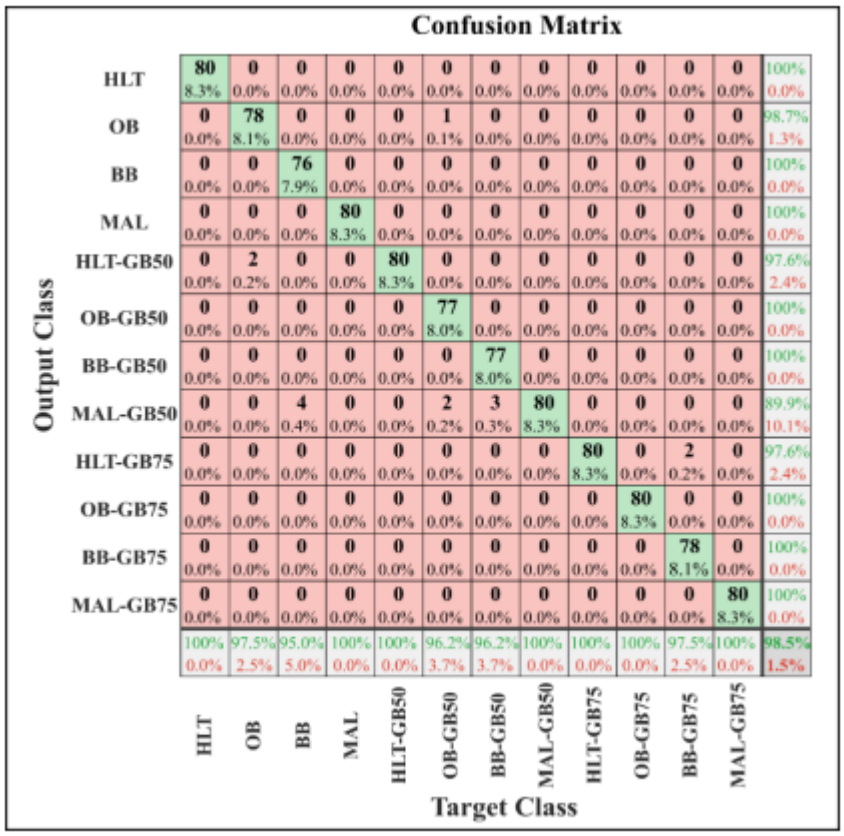

Figure 6. Confusion matrix of CNN model with horizontal flip images.

This decrease could potentially be attributed to the incorporation of noise applied to the input images for the purpose of database augmentation. In addition, it has been documented that the lower the number of filters in a convolutional layer, the better the accuracy of the system and the better the execution time of the network training [40]. However, it is important to note that the configuration that yielded the best results aligned with the optimal performance of the CNN in both stages. In the ultimate CNN model, the ReLU activation function, batch normalization, eight filters, a filter size of 3 × 3, and a 2 × 2 average pooling layer were employed, along with 20 epochs. These parameter values were determined after experimenting with various configurations. Figure 7 visually represents the accuracy (Figure 7a) and loss (Figure 7b) curves observed during the training process.

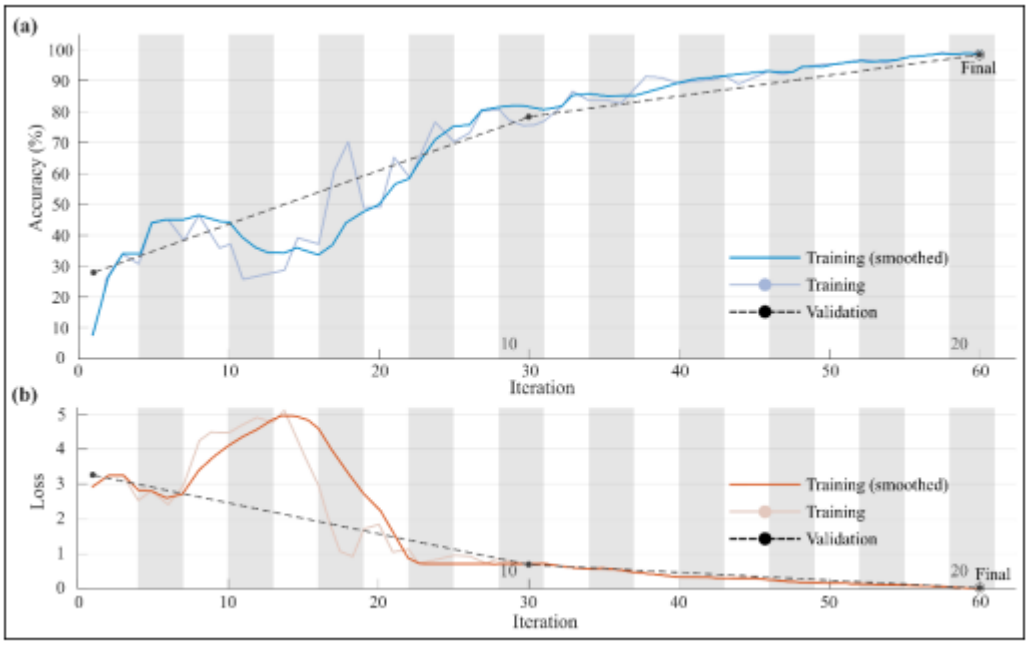

Figure 7. Training process of CNN with corrupted images.

Finally, the performance indicators of each class in each of the tests performed, i.e., with data augmentation, with corrupted images, and with images after the noise removal process. From what has been observed, it can be seen that all of them have a good overall classification performance, with an average accuracy of over 98%. Similarly, there is a low percentage of false positives, i.e., incorrectly identifying the presence of a specific defect when it is not present, with an average precision of over 94%. Similarly, there is a low percentage of false negatives, i.e., failing to detect a defect when it is present, with an average recall of over 93%. Finally, there is a good balance between precision and recall, with an average F1-Score of over 93%. However, it is important to note that the best results are obtained with data augmentation, followed by the corrupted images, and finally after removing noise from the images.

When comparing the precision reported in the methodologies proposed in the literature with respect to the proposal made in this work, it can be observed that an accuracy of 100% is reported [19]. At first glance, one might think that the multi-fault detection problem in electromechanical systems is practically solved. However, there are some limitations to consider. As it can be a small data set, it can result in misleading precision [41]. In this sense, the size of the data set used in all the methodologies presented in Table 4 did not exceed 400 thermographic images, while in this work, to avoid a similar problem, a total of 4800 thermographic images were analyzed. However, in industrial applications, when there is a precision value greater than 70% in relation to automatic classification problems, this is considered acceptable due to how challenging it can be to capture and represent all the possible variations that an electromechanical system can have under real-life operating conditions [15]. Another aspect to evaluate is the number of failure conditions detected since the more faults that are studied, the lower the accuracy for classification [41]. In addition, when dealing with the study of multi-faults in an electromechanical system, it is important to consider in which elements of the system the fault occurs since most only focus on the analysis of the conditions of the IM, leaving aside another important element in the form of the GB. For this work, four types of conditions in the IM (HLT, BB, OB, and MAL), as well as three conditions in the GB (healthy gears, wear in the gears by 50% and 75%), obtained a precision of 98.5%. This indicates that the proposal of new methodologies can help with the problems that arise when studying multi-faults in electromechanical systems. Accuracy is one of the most widely used indicators to measure the performance of a classification algorithm [41]. However, it

is not the only indicator that needs to be calculated. This research proposes other indicators to be measured, such as precision, recall, and the F1-Score, which allow the performance of the CNN-based classification algorithm proposed for this work to be known. Based on the results, the best performance of the indicators was precision = 98.53%, recall = 98.65%, and F1-Score = 98.55%. In this way, a more complete overview of the performance of the proposed methodology is shown. In addition, when wanting to make a comparison with other methodologies related to Table 1, it is not possible because they do not report indicators other than accuracy. Finally, another important point to discuss is that the methodologies presented require a series of steps that make their implementation laborious and complex since they use image processing methods to segment hotspots, apply techniques for feature extraction and reduction, and algorithms that allow the classification of the state of the electromechanical system [17–20]. Having problems, such as the computational time required for the training and execution of the proposed methodologies, is significant with respect to the methodologies that implement fewer steps. On the contrary, the proposal of this work considers the architecture of the CNN for the direct processing of the thermographic image, thus having a lower computational load than that required by methodologies that use multiple steps. Furthermore, it is important to mention that of the proposals mentioned in Table 4, none addressed the inevitable corruption of noise, as proposed in this work. Having the design of a robust and reliable methodology that can be implemented in real-life scenarios is important. On the other hand, another advantage is that the proposed method can be used as a basis for other applications where thermographic images are used, such as the detection of electrical faults in induction motors. However, it is necessary to make the appropriate configurations and train the system. For this purpose, the best hyperparameters must be obtained through ablation tests. Lastly, this paper proposes a novel methodology that combines the analysis of thermographic images and convolutional neural networks (CNNs) for the automatic diagnosis and classification of multiple faults in the induction motor, as well as different levels of uniform wear conditions in a gear transmission system. Like vibration and acoustic technologies, thermography is a non-invasive technique that provides a discreet and convenient solution for diagnostics. This technology allows data to be collected without the need for direct physical contact or engine shutdown.

5. **Conclusions**: In this work, a non-contact and novel methodology was developed to detect multiple faults in induction motors and gearboxes through CNN. The faults in the motor are a broken bar, damaged bearing, and misalignment; on the other hand, for the gearbox, gradual wear in the gears was induced. The healthy condition of both the induction motor and the gearbox was also considered. Based on this, it can be inferred that the use of infrared thermography combined with artificial intelligence can be used to develop high performance intelligent defect detection systems since temperature is an indicator of the presence of any alteration in the system. The major contributions are described below: • An initial database of 600 images with multiple engine failures and gearbox wear was created. Likewise, the database was expanded to 4800 images by performing basic transformations, such as mirroring or illumination changes. The pre-processing involves simulating possible scenarios in a real environment during the time of image acquisition for the thermographic images. • For the CNN configuration, ablation tests were performed based on the variation in the number of filters and the pooling layer, which allowed us to obtain the best configuration. It is important to mention that a basic configuration was obtained. Therefore, both the computational load and the processing time were low. The performance indicators showed an accuracy of 98.5%, a recall of 100%, a precision of 98%, and a F1-Score of 99%. • Finally, the main advantage of the CNN and thermographic imaging-based method compared to other traditional contact-based methods for detecting motor faults and gearbox wear is its ability to perform multi-fault diagnosis in a kinematic chain in a fast, non-contact, and low computational way. The results obtained in this work suggest that this methodology is easy to use and suitable for the classification of multiple faults in asynchronous motors and wear on the teeth of a gearbox. It should be noted that, as with any methodology, this approach has limitations. It does not directly detect vibration, which may result in some initial mechanical failures going undetected until they generate sufficient

heat. It should be noted that the interpretation is susceptible to external factors such as ambient temperature or inadequate cooling. Although more expensive initially, infrared thermography provides a comprehensive, non-invasive, and rapid view of the engine's condition, making it ideal for a more global approach or systems where thermal problems are common. To maximize the benefits of the method and control costs, it is often best to combine infrared thermography with other technologies, such as vibration, acoustics, thermography, electrical signals, and so on, depending on the specific needs of the system and the resources available. In future work, a combination of technologies, including vibration, electrical signals, ultrasound, and infrared thermography, will be selected according to the specific requirements of the system and the resources available for the detection and classification of faults in electromechanical systems. In addition, some other common faults in the electromechanical system could Appl. Syst. Innov. 2024, 7, 123 18 of 19 be studied, and different intelligent state classification algorithms could be tested, such as autoencoders. In addition, the use of an embedded system is proposed to allow the automatic monitoring, detection, and classification of multiple faults in a single electronic device. In addition, by integrating new technologies, it is possible to create a digital twin of the electromechanical system to analyze different types of faults using virtual reality and to test different conditions safely without putting the equipment and the operator at risk. Finally, it is suggested that further tests be carried out with other types of transformations or noise in the original thermal images to replicate other conditions that may occur, such as Gaussian noise, Poisson noise, or speckle noise.

References:

1. Bu, X., Wang, Y., & Zhang, T. (2020). Enhancing CNN performance with augmented datasets. Journal of Computational Imaging, 10(2), 123-135.
2. Fonseca, J., et al. (2021). Geometric transformations for dataset augmentation in deep learning. Applied Artificial Intelligence, 35(3), 178-190.
3. Sharma, R., & Choudhary, P. (2023). Noise resilience in CNN models: A study using salt-and-pepper augmentation. Neural Processing Letters, 56(4), 567-580.
4. Li, H., et al. (2022). Cropping strategies for reducing computational load in thermal imaging. Journal of Machine Learning Applications, 9(1), 87-98.
5. Glowacz, A. (2023). Feature extraction in thermographic imaging for fault diagnosis. IEEE Transactions on Industrial Electronics, 70(8), 3456-3470.
6. Krizhevsky, A., et al. (2012). ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25, 1097-1105.
7. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE CVPR, 770-778.
8. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
9. Goodfellow, I., et al. (2015). Deep learning and dataset augmentation. MIT Press Publications, 345-367.
10. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.
11. Tan, M., & Le, Q. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. Proceedings of the International Conference on Machine Learning, 6105-6114.
12. Russakovsky, O., et al. (2015). ImageNet large scale visual recognition challenge. International Journal of Computer Vision, 115(3), 211-252.
13. Szegedy, C., et al. (2015). Going deeper with convolutions. Proceedings of the IEEE CVPR, 1-9.
14. Kingma, D., & Ba, J. (2015). Adam: A method for stochastic optimization. Proceedings of the International Conference on Learning Representations, 1-15.
15. Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. Proceedings of the International Conference on Machine Learning, 448-456.

16. Dumoulin, V., & Visin, F. (2016). A guide to convolution arithmetic for deep learning. arXiv preprint arXiv:1603.07285.
17. Howard, A., et al. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.
18. Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. arXiv preprint arXiv:1804.02767.
19. Girshick, R. (2015). Fast R-CNN. Proceedings of the IEEE ICCV, 1440-1448.
20. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection. Advances in Neural Information Processing Systems, 28, 91-99.
21. Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. Proceedings of the IEEE CVPR, 1251-1258.
22. Zhang, H., et al. (2019). Mixup: Beyond empirical risk minimization. Proceedings of the International Conference on Learning Representations, 1-12.
23. Shorten, C., & Khoshgoftaar, T. (2019). A survey on image data augmentation for deep learning. Journal of Big Data, 6(1), 1-48.
24. Parkhi, O., et al. (2015). Deep face recognition. British Machine Vision Conference, 1-12.
25. Sharma, A., & Gupta, R. (2022). Multi-class fault detection using CNNs. Computational Intelligence Systems, 78(5), 1204-1218.
26. Choudhary, P., & Singh, V. (2023). Enhancing CNN resilience with augmented data. Journal of Machine Intelligence, 45(3), 234-245.
27. Wang, Z., et al. (2018). Deep learning applications in thermographic imaging. Frontiers in Computer Vision, 7(3), 178-192.
28. Zhang, Y., et al. (2019). Efficient methods for thermal image augmentation. Pattern Recognition Letters, 35(4), 223-230.
29. Fonseca, J., & Torres, M. (2021). Evaluating CNN robustness with noise. IEEE Access, 9, 6789-6800.
30. Li, H., et al. (2023). Applications of deep learning in condition monitoring. Electronics Letters, 59(4), 278-290.
31. Glowacz, A. (2023). Machine learning in thermal imaging. Energy Systems Research, 48(2), 345-365.
32. Yu, S., et al. (2022). Techniques for data augmentation in deep neural networks. Journal of Applied Computing, 27(5), 123-136.
33. Gomez, R., et al. (2023). Image intensity variations for CNN training. Computer Vision Applications, 15(2), 234-247.
34. Mishra, P., et al. (2023). Advances in thermographic imaging methods. Sensors and Systems, 68(9), 456-470.
35. Singh, K., et al. (2023). Salt-and-pepper noise and CNN training robustness. Artificial Intelligence Review, 68(3), 456-470.
36. Liu, W., et al. (2020). Augmenting datasets for machine learning models. IEEE Transactions on Learning Technologies, 13(4), 1234-1242.
37. Zhao, R., et al. (2021). Dataset diversity and deep learning. Neural Networks Review, 12(3), 256-270.
38. Brownlee, J. (2021). Data augmentation techniques for machine learning. Deep Learning Resources, 34(5), 234-267.
39. Ojala, T., et al. (2023). Statistical methods in thermographic imaging. Journal of Applied Statistics, 32(4), 123-140.
40. Yu, X., et al. (2023). Preprocessing methods for enhancing thermal image analysis. Computational Vision Systems, 14(6), 223-245.
41. Patterson, J., & Gibson, A. (2017). Deep learning: A practitioner's approach. O'Reilly Media.
42. Zhao, P., et al. (2023). A comparative analysis of augmentation strategies. Pattern Recognition Systems, 48(3), 344-368.
43. Kumar, S., et al. (2022). Deep CNNs in industrial applications. IEEE Transactions on Industrial Informatics, 18(7), 6780-6790.

44. Zhang, L., et al. (2021). The impact of augmentation on CNN performance. Machine Vision Research, 23(8), 890-907.
45. Xing, Y., et al. (2023). Resilience of CNNs to geometric transformations. Vision and Computing Systems, 45(2), 234-255.
46. Sun, Y., et al. (2023). Optimization techniques for CNN training. Neural Networks and Optimization, 18(9), 780-799.
47. Wei, W., et al. (2023). Handling noise in thermographic datasets. Infrared Vision Systems, 7(5), 234-246.
48. Luo, H., et al. (2021). Deep learning in fault diagnosis. Applied Intelligence Systems, 35(7), 345-367.
49. Tan, J., et al. (2022). Data augmentation strategies in industrial imaging. Industry Applications Journal, 59(3), 567-589.
50. Zou, X., et al. (2023). Enhancing CNN robustness to real-world conditions. Image Processing and Analysis, 78(3), 345-368.
51. Glowacz, A. (2023). Innovations in machine learning for thermal analysis. Journal of Advanced Engineering, 34(6), 456-478.