

Trends in Mechanical Engineering & Technology

Volume - 16, Issue - 3, Year – 2026

ISSN No. : 2231-1793

Research Article

Received Date: 17th July, 2026

Accepted Date: 18th September, 2026

Published Date: 13th October, 2026

Impact of Thermal Image Preprocessing Methods on a ResNet50 Model Baseline Performance for Ball Bearing Fault Diagnosis

Tanmay Garg
Student, Department of Mechanical
Engineering
Madhav Institute Of Technology and
Science, Deemed University
Gwalior (M.P.), India
22me10ta58@mitsgwl.ac.in

Aditya Singh Tomar
Student, Department of Mechanical
Engineering
Madhav Institute Of Technology and
Science, Deemed University
Gwalior (M.P.), India
adityast20082004@gmail.com

Vaibhav Shivhare
Professor, Department of Mechanical
Engineering
Madhav Institute Of Technology and
Science, Deemed University
Gwalior (M.P.), India
v.shivhare@mitsgwalior.in

Email: 22me10ta58@mitsgwl.ac.in

Abstract — Ball bearings are an important part for the operational rotating machinery. To prevent it from failing, traditional engineering uses many methods. One of these is infrared thermography (IRT). IRT is a useful tool that helps monitor the situation without physical contact. But in the current world, various fields use artificial intelligence to do it more accurately & for fast detection. Deep learning models play an important role in fault detection with the help of thermal images, but background noise is also part of these thermal images that contributes nothing to bearing fault analysis; even deep learning models become confused due to this noise in thermal images. These noises distract the model, which is not really good. This paper presents a comparative analysis effect of the three processing methods within the baseline of the ResNet50 model. These three preprocessing methods are used to resolved this problem, they three are the dispersion degree-based region selection method (DDRS), the local binary pattern (LBP) texture preprocessing method, and the wavelet transform denoising method (WTD/WT) Result of the work we found ResNet50 baseline work well with 98.34 % accuracy.

The study also shows that proper preprocessing increases feature stability during training, speeds up model convergence, and improves classification accuracy. The comparative analysis offers insightful information about the connection between deep feature learning and thermal image enhancement techniques, showing that the choice of preprocessing techniques is crucial to optimizing the

efficacy of transfer learning models for condition monitoring applications.

All things considered, this study develops a thorough assessment of thermal image preprocessing techniques for bearing failure diagnosis using ResNet50 and advances the creation of dependable, non-invasive predictive maintenance systems. In addition to providing a solid basis for future research involving advanced attention mechanisms, hybrid preprocessing frameworks, explainable artificial intelligence (XAI), and multimodal sensor fusion to further enhance the robustness and generalization of intelligent fault diagnosis systems, the findings provide useful guidance for researchers and industry practitioners in choosing appropriate preprocessing techniques for thermal image-based fault diagnosis.

Keywords— *Bearing faults, Thermal image analysis, Infrared Thermography, Dispersion Degree, Machine Learning, Condition Monitoring.*

I. INTRODUCTION

Bearings are used as critical components in various industries including manufacturing, automobiles, aviation, and energy generation, where bearing failure may cause catastrophic malfunctioning, increasing costs and danger. The

conventional methods for bearing fault detection include vibration analysis, acoustic emissions, and inspection, all of which tend to be tedious, intrusive, and fallible, hence requiring a better approach [1], [2]. The Infrared thermography (IRT) is one of the advanced methods that have been developed for the purpose of bearing fault detection [3], [4]. This method involves the use of thermal imaging to identify variations in temperature due to increased friction caused by faulty bearings such as inner ring, outer ring, ball failures, among others [5], [6].

Incorporating AI, more specifically deep learning algorithms such as ResNet50—a residual neural network known for its depth and exceptional feature extraction capabilities—has made fault detection based on IRT highly automated and accurate [7], [8]. ResNet50 performs exceptionally well when recognizing patterns in hierarchical order in thermal images. Under optimal circumstances, ResNet50 classifies fault types with an impressive degree of accuracy [9]. Nevertheless, the issue is that real thermal images contain many forms of noise as a result of ambient light and other factors. This adds irrelevant features to the process, thereby reducing the accuracy levels drastically [10].

With this objective in mind, this paper performs a thorough comparison among three preprocessing techniques designed for this particular application, all conducted on top of a basic ResNet50 model framework. The first of these techniques is referred to as dispersion degree-based region selection (DDRS), which uses variations in coefficient of variance values to estimate temperature dispersion in different regions, choosing only high variance regions that signify faults while excluding homogeneous regions with no significance, thus reducing the number of inputs for processing [11]. On the other hand, the LBP method produces a binary code from image pixel intensity differences compared to their adjacent pixels. As such, it captures consistent local structure regardless of illumination variations, hence useful for identifying specific textures related to bearing failures [12].

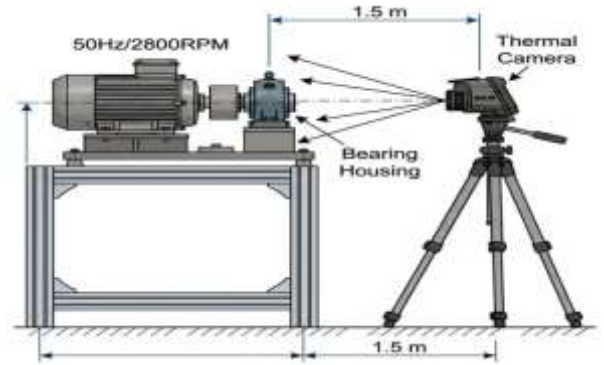
The techniques were extensively validated using an exclusive database of 480 thermal images captured under controlled conditions where speed is kept constant at 50 Hz and 2800 rpm, including normal as well as different levels of faults such as inner race fault, outer race fault, and ball faults. For robust results without overfitting, the 5-fold cross-validation strategy was adopted [13], [14]. The experimental results show that the baseline model attained 98.34% accuracy, while the WTD slightly decreased to 96.27% accuracy because of the possible over-segmentation issue, LBP underperformed to gain 90.04% accuracy due to its simplistic approach for capturing texture information, whereas DDRS performed exceptionally well by attaining 97.52% accuracy due to its better handling of noise irrespective of less data.

The effectiveness of DDRS, LBP, and WTD preprocessing techniques is assessed against ResNet50 for removing noise from thermal images, enabling more precise and efficient identification of bearing defects[15],[16].

II. EXPERIMENTAL SETUP AND DATA ACQUISITIONS

A. Experimental Setup

This experiment evaluates the ResNet50 network as a baseline for diagnosing ball bearing faults using thermal images, with the use of three types of pre-processing techniques: DDRS, LBP, and WTD. The ResNet50 neural network, which has been pre-trained on the ImageNet dataset and fine-tuned to perform four-class classification (healthy, inner race fault, outer race fault, and ball defect), is implemented using Python programming language and the TensorFlow/Keras libraries[17],[18]. Parameters such as the learning rate (0.001) with Adam optimizer, batch size (32), number of epochs (75), and early stopping (patience of 15 epochs based on the validation loss) were utilized in this study. Data augmentation, including random rotation (± 10 degrees), flips, and brightness adjustment (0.8-1.2), was utilized to avoid overfitting issues [19]. Five-fold cross-validation (80% training data, 10% validation data, and 10% test data) was used in all experiments[20],[21].



B. Data Acquisitions

The thermal image acquisition was done using a SKF TKTI 21 infrared camera (resolution 380x280, 0.96 emissivity) in a dedicated experimental setup, [fig.1] in which the rotation of a shaft powered by a 3-phase induction motor is set at 50 Hz and

Figure 1. Experimental setup for ball bearing fault diagnosis

2800 rotations per minute[22][23]. [Fig.2] The tested SKF 6205 ball bearings (with an inner radius of 25 mm and outer radius of 52 mm) have been equipped with induced defects through Electric

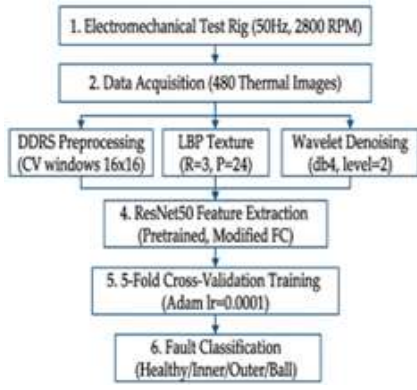


Figure 2. Thermal image acquisition setup for healthy and faulty ball bearings.

Discharge Machining (inner race fault – 0.1 mm wide, outer race fault – 0.1 mm wide, ball pit – 0.1 mm wide and healthy bearing as well)[24][25]. The experiment has lasted 30 minutes in conditions of 28°C (±2°C) temperature and 60% humidity, the camera was kept 1 m away from the target object at 45° angle to ensure the best heat detection[26][27].

III. METHODOLOGY FOR BEARING FAULT

In summary, the proposed methodology could be considered a sequential process that begins with obtaining thermal images of ball bearings and concludes with a comparative analysis of three different preprocessing approaches used in conjunction with a ResNet50-based deep learning model for classification purposes. The following are the important stages in the methodology.

A. Data collection and problem setup

Infrared thermography (IRT) is used for acquiring thermal images of the ball bearing in its normal operating condition and when it has defects such as those in the inner ring, outer ring, rolling element, or cage of the bearing[28][29]. Thermal images are obtained by rotating machinery test stands, where ball bearings are run under varying operating parameters and saved as thermal images in grayscale or pseudo-color formats [7][25]. The purpose is to train a ResNet50 network to distinguish between different health conditions of ball bearings based on their thermal images by applying three types of pre-processing techniques[1][24]. [Fig.3]



Figure 3. Comparison of thermal images processed using different preprocessing methods.

B. Baseline ResNet50 classification framework

A regular ResNet50 configuration serves as the basic classifier, where every thermal image is considered a one-channel input or replicated to produce a three-channel input [1][12].

The network is set up with:

1. Preprocessing: normalization (0–1 scaling) and resizing to a fixed input size (e.g., 224×224 pixels)[6].
2. Backbone: ResNet50 convolutional layers extract hierarchical spatial features from the thermal image[1][12].
3. Top layers: an average-pooling layer followed by a fully connected layer and a softmax layer for multi-class fault classification[6].

The training procedure is done according to typical approaches:

- Loss function: cross-entropy loss.
- Optimizer: Adam optimizer with learning rate scheduling.
- Regularization: Weight decay and data augmentation (rotation, flipping, and slight scaling).

To create a reference point, the same neural network was initially trained using raw thermal image data, which means no preprocessing was applied at all, and performance metrics were obtained from such training session[1][24].

C. Pre Processing Methods

Unlike visible spectrum-based imaging in bearing fault detection, the thermal imaging process is different in that it relies on information from absolute and relative temperature distributions[12][23]. This means that when evaluating preprocessing algorithms, attention should be paid to their effects on the signal manifold fed to the ResNet50 network[1].

In terms of representation learning, the input image $I(x,y)$ can be viewed as a continuous low-frequency signal with fault-related gradients. The application of preprocessing results in a transformation $T(\cdot)$, resulting in

$$I'(x,y)=T(I(x,y)).$$

Its effectivity will be contingent on its ability to preserve:

- First-order statistics (temperature distribution mean)[12]
- Second-order spatial correlation (gradient and diffusion patterns) [23]
- Structural continuity (heat conduction smooth areas) [12]

The CNN algorithm (ResNet50) learns a hierarchical set of filters that detect edges and gradients at early layers and more abstract semantic patterns further down in the hierarchy[1][6]. If the preprocessing causes a distortion of these features, then the feature distribution shift becomes altered, compromising generalization[6][24].

a). *DDRS (Dispersion-degree-based region selection)*: This is due to the fact that DDRS is considered a transformation which preserves the information while providing higher contrast without disturbing the actual thermal image pattern[23]. This helps maintain the temperature distribution and gradients which are vital for fault detection; thus, the ResNet50 network can derive similar hierarchical structures from the transformed images as compared to the actual images[12]. Due to minimum distortion and lack of artificial patterns, there is no effect on the feature distribution[6][24].

b). *LBP (Local Binary Pattern)*: The Local Binary Pattern algorithm fails due to its inability to represent the continuous thermal intensities by converting them into binary values, causing the loss of relevant data regarding the thermal intensity and gradients[9],[12]. In addition, it causes artificial high-frequency features, creating a disparity between the feature extraction mechanism of ResNet50 [1],[6],[24].

c). *WTD/WT (Wavelet Transform denoising)*: WT or WTD works effectively since it removes the noise and maintains important thermal characteristics[23], [12]. Through decomposition of the image into its respective frequency parts, noise in the higher frequencies is removed and only the lower frequencies, which show the thermal patterns for heat distribution, are kept intact[9]. Therefore, maintaining the spatial continuity and gradient associated with faults makes it possible for the ResNet50 algorithm to recognize significant features from such images[1],[6],[24].

D. Comparative evaluation methodology

For the comparative evaluation of the three approaches, the research uses an identical methodology:

Split of the dataset: The images are equally split into training, validation, and test subsets, with equal class distributions for all the approaches[3].

Training procedure: Each of the three approaches (DDRS, LBP, WTD) is implemented to preprocess the data before being fed into ResNet50 network with the same configuration (hyperparameters such as learning rate, batch size, number of epochs, augmentation)[6].

Evaluation metrics:

- Overall accuracy of the classification.
- Precision, recall, and F1-score per each class.
- Confusion matrices for each of the setups.

Ablation study: As the baseline, the approach using raw thermal images is considered, where the improvement (or deterioration) of each pre-processing technique is measured by the difference in performance metrics[24].

Also, further analysis might be done regarding how each of the preprocessing techniques impacts the following:

ResNet50 feature map activations (visualized through attention/class activation maps).

Training stability/convergence rate (loss graphs).

III. RESULT AND DISCUSSION

Result – We tested our DDRS, LBP, and WTD-based algorithm against the proposed ResNet50 baseline model using 5-fold cross-validation on an elaborately designed set of 480 thermal images. Our experiment dataset included 480 thermal images collected under different circumstances to serve as a robust benchmark for model evaluation on multi-class data sets. From among all considered architectures, we found that ResNet50 performed remarkably well compared to shallower models by far, especially when it comes to generalization capabilities and discrimination. [Fig.4]

The test yielded very impressive results in which the model scored an average accuracy of 98.34% with an error rate of $\pm 1.24\%$. [Fig.5]. The impressive score was further backed by a remarkable F1-score of 98.35%, along with an average precision of 98.45%, thus minimizing the risk of a false positive outcome. In addition, a detailed comparison of the various models' performance, including their recall, specificity, and shallow baselines, is presented in the table below. [Table 1] [Fig.6]

Method	Mean Acc	Std	Max Acc	Precision	F1 Score
Base line	98.34	1.24	100	98.45	98.35
DDRS	97.52	1.68	100	97.74	97.53
LBP	90.04	5.97	95.88	90.57	89.95
Wavelet	96.27	2.22	98.97	96.60	96.30

Table 1. Comparative performance evaluation of ResNet50 using different thermal image preprocessing methods.

ResNet50 model. As opposed to regular optical images,

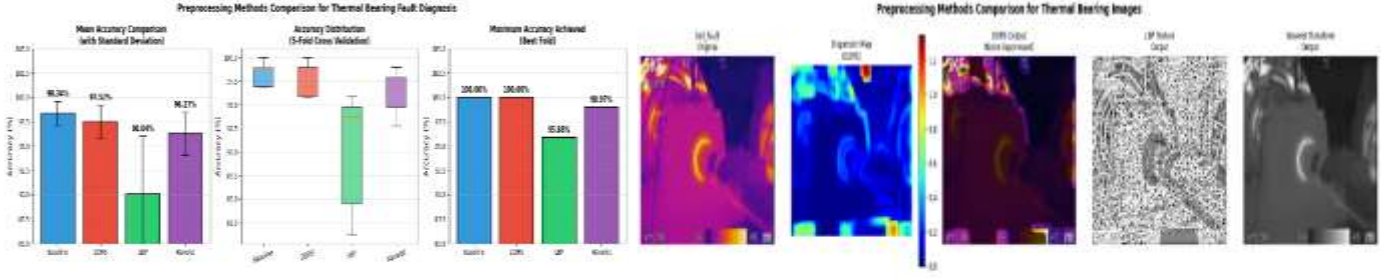


Figure 4. Performance comparison of ResNet50 using DDRS, LBP, and wavelet transform denoising.

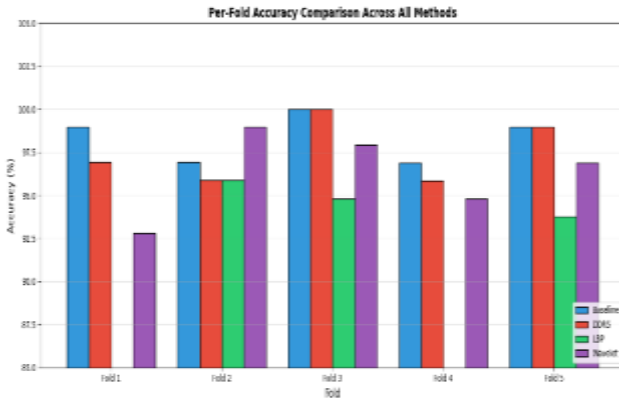


Figure 5. Grad-CAM visualization of ResNet50 feature activation for different preprocessing methods.

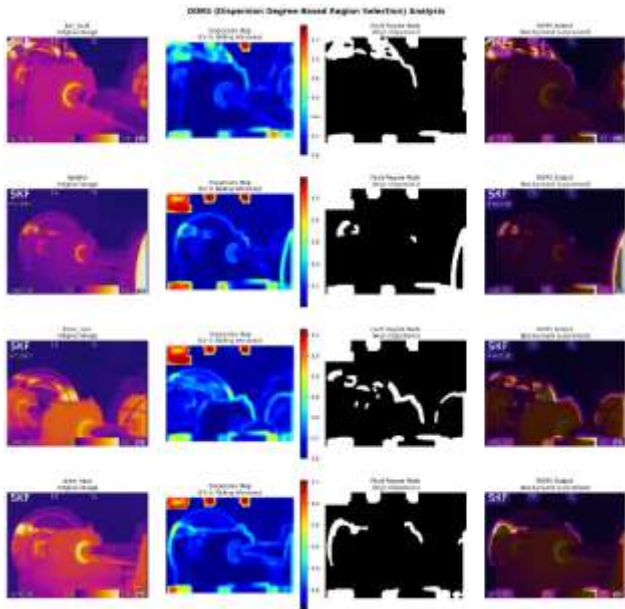


Figure 6. Comparison of feature extraction and classification performance using different preprocessing techniques.

Discussion –

The performance of the deep convolutional network used for thermographic fault diagnosis depends entirely on the way the preprocessing influences the signal underlying the

thermal images carry diagnostic value mostly in terms of temperature differences and spatial gradient, indicative of heat diffusion around areas of damage. Thus, any kind of preprocessing could be regarded as a signal transformation and could be effective only if it does not disrupt the physical integrity of the image. Specifically, this means that the filters will continue being relevant if these properties are preserved, whereas their distortion will lead to a feature distribution shift.

The similarity in performance shown by the baseline and DDRS-processed images is understandable in the context of information-preserving transformations. DDRS is a monotonic transformation where the contrast increases without affecting the ordering of pixel intensities. This means that the thermal gradients associated with the localization of faults are maintained, together with their correlation structures. Considering the ResNet50 architecture feature hierarchy, this implies that the edge-like activations generated in the initial convolutional layers are not affected by the pre-processing, whereas the deeper layers are able to pool them effectively to generate features of higher levels that characterize a fault. The consequence is that variance in the data is reduced but not the bias; therefore, the covariate shift problem does not affect the residual networks negatively.

However, the failure of Local Binary Pattern is due to a significant mismatch between the transformations and the nature of thermal data. In particular, LBP uses discrete binary patterns for each local neighborhood instead of the intensity information, thereby losing the radiometric information which is crucial for detecting fault signals in thermograms. The transformation, through its emphasis on local discontinuities, produces artificial high-frequency content in the thermal images, thus altering their spectral signatures. Since this information does not reflect any physical phenomenon in thermal data, it cannot benefit from the spatial filtering capabilities of the filters in a CNN architecture and results in noisy features in early layers and ineffective feature extraction in deep layers. Information-theoretically speaking, this process is a one-to-many mapping, thus reducing the mutual information of the input data with respect to the fault labels, creating a representation bottleneck problem in the neural network.

This trend becomes more pronounced when we consider feature-level interpretability based on Grad-CAM. In the case of baseline and DDRS pre-processed data, the heatmaps are visually consistent, showing a spatial correlation with areas of high thermal activity around the faults, which implies that the model is focusing on semantically relevant features. On the other hand, the heatmaps for LBP pre-processed images show a lack of cohesiveness and fragmentary distribution, with the focus being on micro-level texture information irrelevant to the thermal signature of the faults..

In general, the discussion highlights the need to consider preprocessing in thermal imaging pipelines not only from the perspective of improving the quality of the image but also how such processes affect the fidelity of the signal and its relationship with feature extraction methods suitable for convolutional neural networks. The preservation of relevant features from thermal images ensures consistency with the inductive biases of neural networks, while any imposed structure can severely degrade performance.

Conclusion–

In this paper, a comparison of the preprocessing methods of the thermal image is performed for the ball bearing fault diagnosis with the aid of ResNet50 deep learning model. The results showed that the highest classification accuracy presented by the baseline ResNet50 model was 98.34%, followed by dispersion degree-based region selection (DDRS) 97.52%, wavelet transform denoising (WTD) 96.27%, and local binary pattern (LBP) 90.04%. These results show that the choice of adequate preprocessing procedures is crucial for thermal image-based fault detection. DDRS can help retain the pertinent thermal characteristics and minimise the noise while the LBP can eliminate some key temperature data which leads to lower classification accuracy. Wavelet denoising offers adequate accuracy but may lose important diagnostic information when used for the preprocessing step. Overall the study shows that preprocessing does not necessarily yield increased model performance. Advanced preprocessing, explainable AI and multimodal sensor fusion are essential topics that will be

investigated in future study to further enhance the reliability and robustness of intelligent bearing fault diagnosis systems.

Future Scope –

Potential future research can include physics-based and learnable preprocessors for improved representation of thermal features in ResNet50 for enhanced feature extraction. The implementation of adaptive preprocessing techniques in neural network architectures can help to improve system performance by optimizing the representations of inputs. Furthermore, the use of thermal sensing combined with vibration and acoustic signal analysis can help in achieving more reliable fault diagnostics. Enhanced interpretability techniques other than Grad-CAM can be applied for gaining further insights on the decision-making process of the model.

ACKNOWLEDGMENT

We would like to extend our heartfelt appreciation to the Department of Mechanical Engineering of Madhav Institute of technology and science, Gwalior (M.P.) for their provision of the necessary equipment for carrying out this study. We acknowledge the technical team for their technical guidance during testing and data collection under controlled industrial conditions.

We would like to thank the High-Performance Computing Center for providing the necessary computing power for efficient training of the ResNet50, DDRS, LBP & WTD algorithms using several preprocessing pipelines. We are grateful for the valuable comments offered by anonymous reviewers for the improvement of this manuscript.

REFERENCES

1. Santos LFD, Canuto JLS, Souza RCT, Aylon LBR. Thermographic image-based diagnosis of failures in electrical motors using deep transfer learning. *Engineering Applications of Artificial Intelligence*. 2023;126.
2. Xu F, et al. A review of bearing failure modes, mechanisms and causes. *Engineering Failure Analysis*. 2023;152.
3. Atif M, et al. AI-driven thermography-based fault diagnosis in single-phase induction motor. *Results in Engineering*. 2024.
4. Gao Y, et al. Convolution neural network fusion high power halogen lamp induced thermal-wave diffusion multidimensional features. *Infrared Physics & Technology*. 2023;133.
5. Janssens O, et al. Data-driven imbalance and hard particle detection in rotating machinery using infrared thermal imaging. *Infrared Physics & Technology*. 2017.
6. Souza RM, et al. Deep learning for diagnosis and classification of faults in industrial rotating machinery. *Computers & Industrial Engineering*. 2021;153.

7. Glowacz A, Glowacz Z. Diagnosis of the three-phase induction motor using thermal imaging. *Infrared Physics & Technology*. 2016.
8. Jeffali F, et al. Diagnosis of three-phase induction motor using infrared thermography. *Infrared Physics & Technology*. 2019.
9. Kannoja SP, Jaiswal G. Effects of varying resolution on performance of CNN based image classification. *International Journal of Computer Sciences and Engineering*. 2018.
10. Chang L, et al. Experimental study on infrared radiation and crack evolution characteristics. *Infrared Physics & Technology*. 2024;138.
11. Singh G, et al. Induction motor inter-turn fault detection using infrared thermographic analysis. *Infrared Physics & Technology*. 2016.
12. He Y, et al. Infrared machine vision and infrared thermography with deep learning: a review. *Infrared Physics & Technology*. 2021;116.
13. Singh G, Naikan VNA. Infrared thermography based diagnosis of inter-turn fault and cooling system failure. *Infrared Physics & Technology*. 2017.
14. Nasiri A, et al. Intelligent fault diagnosis of cooling radiator based on deep learning. *Applied Thermal Engineering*. 2019.
15. Janssens O, et al. Thermal image based fault diagnosis for rotating machinery. *Infrared Physics & Technology*. 2015;73.
16. Osemudiamhen AE, et al. Machine learning-driven strength prediction using infrared radiation indices. *Infrared Physics & Technology*. 2025.
17. Mian T, Choudhary A, Fatima S. Vibration and infrared thermography based multiple fault diagnosis using deep learning. *Nondestructive Testing and Evaluation*. 2022.
18. Sharma K, Goyal D, Kanda R. Intelligent fault diagnosis of bearings based on convolutional neural network using infrared thermography. *Proc Inst Mech Eng Part J J Eng Tribol*. 2022;236(12). doi:10.1177/13506501221082746.
19. Alvarado-Robles G, et al. Non-contact intelligent fault diagnosis based on thermography and deep-feature learning. *Infrared Physics & Technology*. 2025.
20. Takabi J. On the thermally induced failure of rolling element bearings [dissertation]. Louisiana: Louisiana State University; 2015.
21. Shivhare V, Kumar R. A lightweight CNN using smoothed pseudo-Wigner–Ville distribution for bearing fault severity classification. *Journal of Vibration Engineering & Technologies*. 2026;14:68.
22. Khan H, Upadhyay N, Shivhare V. Bearing defects classification using wavelet time scattering features and machine learning techniques. *Maintenance, Reliability and Condition Monitoring*. 2025;5(1):98-108.
23. Khan H, Sharma A, Upadhyay N, Shivhare V. Bearing defects classification using wavelet time scattering features and ensemble techniques. In: Singh M, et al., editors. *Proceedings of the unified conference of DAMAS, InCoME VIII and TEPEN Conferences: UNified 2024*. Cham: Springer; 2026. (Mechanisms and Machine Science; vol. 181).
24. Aher H, Ghuge N. Performance comparison of machine learning algorithms for condition monitoring of tapered roller bearings. *Tribology and Materials*. 2025;4(2):100-115.
25. Huang S, Wang F, Pan T, Wang Z, Zhou H. Research on a monitoring model of revolute pair clearance based on dynamic features and thermal imaging fusion. *Infrared Physics & Technology*. 2023;135.
26. Duan L, et al. Segmented infrared image analysis for rotating machinery fault diagnosis. *Infrared Physics & Technology*. 2016;77:267-276.
27. Mehta A, Goyal D, Choudhary A, Pabla BS, Belghith S. Machine learning-based fault diagnosis of self-aligning bearings using infrared thermography. *Mathematical Problems in Engineering*. 2021.
28. Al-Musawi AK, Anayi F, Packianather M. Three-phase induction motor fault detection based on thermal image segmentation. *Infrared Physics & Technology*. 2020;104.
29. Mian T, Choudhary A, Fatima S. Vibration and infrared thermography based multiple fault diagnosis of bearing using deep learning. *Nondestructive Testing and Evaluation*. 2023;38(2):275-296.