machines

MDPI

Article

Investigation of Transfer Learning Method for Motor Fault Detection

Prashant Kumar, Saurabh Singh * and Doug Young Song

Department of AI and Big Data, Woosong University, Daejeon 34606, Republic of Korea; prashant@wsu.ac.kr (P.K.); dysong@wsu.ac.kr (D.Y.S.)
* Correspondence: singh.saurabh@wsu.ac.kr

Abstract: Industry 4.0 is propelling modern industries forward due to its reliability, stability, and performance. Electric motors (EMs) are utilized in multiple industries for their efficiency, precise speed and torque control, and robustness. Detecting faults in motors at an early stage is crucial to ensure maximum productivity. Recently, DL has been implemented as a data-driven approach for detecting faults in motors. However, due to the limited availability of labeled fault data, the performance of the DL model is constrained. This issue is addressed by leveraging transfer learning (TL), which uses knowledge from a larger source domain to improve performance in a smaller target domain. In this paper, a multiple fault detection (FD) model for EMs is proposed by combining the ideas of deep convolutional neural networks (CNNs) and TL. A one-dimensional signal-to-image conversion technique is suggested for converting the vibration signal to images, and an Inception-ResNet-v2-inspired FD model is proposed for detecting bearing faults in the motor. The proposed method achieved a mean accuracy of more than 99%.

Keywords: transfer learning (TL); convolutional neural network (CNN); bearing faults; electric motor

check for updates

Academic Editor: Sajad Saraygord Afshari

Received: 21 March 2025
Revised: 12 April 2025
Accepted: 17 April 2025
Published: 17 April 2025

Citation: Kumar, P.; Singh, S.; Song, D.Y. Investigation of Transfer Learning Method for Motor Fault Detection. *Machines* **2025**, *13*, 329. https://doi.org/10.3390/machines13040329

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Modern industries require early FD to ensure uninterrupted production. Since electric motors (EMs) are a key component of modern industries, detecting faults in EMs is both critical and necessary. EMs are used in various industries, such as mining, textiles, and automobiles. Safety is a significant concern in the mining industry, and early fault detection (FD) is essential to ensure high safety standards. In most industries, timely FD is crucial to maintain high safety standards while maximizing production. Despite their robustness, faults in EMs are unavoidable, and studies by the IEEE have shown that 50% of the total faults in EMs are due to bearing faults [1]. The objective of FD is to detect faults early. With the rise of smart manufacturing, data-driven FD techniques have garnered attention from researchers [2]. An intelligent feature learning capability from the sensor data is the driving force of data-driven FD techniques [3].

For FD in EMs, machine learning (ML) algorithms such as Artificial Neural Network (ANN), Support Vector Machine (SVM), and k-Nearest Neighbors (kNNs) are traditionally used in conjunction with signal processing tools such as fast Fourier transform (FFT), wavelet transforms, and short-time Fourier transform (STFT) [1,4,5]. An ML-based FD approach for induction motors using stator currents and a vibration signal is proposed in [6]. For FD in [7], a ball bearing FD was constructed employing statistical features from vibration data as an input to an ANN and SVM model. In [8], a combined approach of SVM and kNNs is used for monitoring the rotational machines with the help of a vibration

signal. The effectiveness of FD techniques hinges on proficient feature extraction, and the quality of the feature extraction significantly influences the FD techniques' performance. However, determining the necessary features for extraction is a challenging process that necessitates both prior knowledge and skill. It can be difficult to discern which features are critical, especially when considering the vast number of available options [1]. Therefore, it is crucial to appropriately choose the features, as the success of the FD model is highly dependent on this selection process.

Deep learning (DL) is a renowned area in artificial intelligence, and its application is immense in image processing, speech recognition, and natural language processing [9,10]. Its innate capacity to extract features can help it overcome the shortcomings of traditional ML methods [11]. Researchers have utilized DL techniques for FD in addition to other applications. When it comes to FD in EMs, common DL approaches include the sparse autoencoder (SAE), the deep belief network (DBN), and the convolutional neural network (CNN) [12–15]. One advantage of using DL is its ability to automatically extract and select features, which renders it a fitting option for FD in EMs.

The large labeled data requirement by DL is one of its downsides. A large volume of labeled fault data is difficult to accumulate, and thus, this hinders the performance of the DL-based FD method. The current FD technique that uses DL is limited in terms of the number of hidden layers it can have, which can negatively impact its performance since the model is relatively shallow [12]. These shallow DL models do not utilize the full capacity of DL models. The benchmark CNN models feature hundreds of layers that have been trained on millions of tagged images, such as the ImageNet dataset, which has about ten million images. However, in industrial applications, limited labeled fault monitoring data are available. In recent times, researchers are looking toward the application of TL to resolve this deficiency.

Models with more than 51 weighted layers have not been widely investigated in the past, while models with more than 100 weighted layers have not been much studied. The performance of FD models with more than 100 weighted layers has not been analyzed due to the challenges associated with deep network training. Despite the challenges associated with training deep models, such as the limited number of hidden layers, these problems can be addressed by leveraging transfer learning (TL). This technique permits the use of pre-trained models from other applications, effectively eluding the need for extensive training of deep models from scratch. Moreover, the high depth of the network provides numerous advantages. Deep structures learn scattered and small representations that are more remarkable than what shallow models learn [16,17]. As a result, it is preferable to learn a valid representation of information using deep structures rather than shallow models. The systems' depth allows for greater feature learning, and TL promotes faster convergence. Furthermore, the deep structure allows for good domain adaptability [18]. The shallow structure of the model can impede its ability to adapt to different domains and effectively represent data and be susceptible to new data.

In recent times, the availability of computational power has motivated researchers to implement DL in FD. This paper presents an FD model motivated by TL and CNN for bearing faults in EMs. A TL approach using the Inception-ResNet-v2 [19] structure, named as IRV2-CNN model, is used for FD. Inception-ResNet-v2 model has excellent performance on the ImageNet dataset, which has millions of data. Here are the main contributions of the proposed techniques:

- The primary contribution of this paper is the introduction of a TL approach for detecting bearing faults in EMs.
- The proposed technique uses the layers of the pre-trained Inception-ResNet-v2 model for feature extraction with TL application.

- Data for the analysis of the proposed model were collected from the experimental setup under different loading conditions and fault conditions.
- This model obliterates the requirement of handcrafted features.

This paper is divided into various sections. Section 2 discusses related work as well as the intricacies of feature learning with a pre-trained network. The mechanism for detecting faults is detailed in Section 3. The details of the experimental work and the results are covered in Section 4. Finally, Section 5 concludes the proposed work.

2. Related Work

2.1. FD Based on DL

Early FD in industries can facilitate reliable and optimum production. The FD techniques are of four types, namely, model-based, data-driven, signal-based, and hybrid methods [20]. The data-driven FD methods have seen a surge in usage due to more fault data availability.

A multiple FD approach for EMs based on a deep CNN model using an adaptive gradient optimizer is proposed in [13]. In [21], a bearing FD of the electric locomotive is propounded using an auto-encoder for dimensionality reduction, convolutional deep belief network, and Gaussian visible units for feature learning. Also, an exponential moving average is deployed for performance improvement of the deep network. In [22], a bearing fault diagnosis is propounded using a DBN-based hierarchical network. The authors in [23] proposed an adaptive CNN approach for bearing defect detection. An adaptive deep CNN model is used to study a bearing defect detection technique in [24]. In [25], the authors propounded a bearing FD using vibration data from multiple sensors and a CNN. The research in [12] reveals that most of the developed models for FD in EMs are reasonably shallow in depth, which restricts the fault classification performance. In this paper, a new IRV2-CNN (Inception-ResNet-v2) model is proposed with higher depth and better fault classification.

2.2. Feature Learning Using Pre-Trained Network

In order to make the best use of a DL model, a significant number of labeled data are typically required. However, acquiring a large number of labeled fault data can be a difficult task. In comparison to some of the state-of-the-art CNN models trained on ImageNet, DL models for FD are often shallow. For efficient training of deeper models, the data requirement is humongous. So, the training of deep models for FD is a daunting task and nearly impossible.

The potential solution to this problem can be to use a deep pre-trained model trained on ImageNet as a feature extractor [26]. The pre-trained models can be used where data are scarce. Researchers have used this concept in multiple domains like bio-medical, image processing, and many more. In [27], the authors have used the concept of feature transferring for cancer drug sensitivity prediction. With the use of magnetic resonance images of the brain, a deep TL approach using the CNN-based ResNet34 model is employed to categorize the different states of the brain [28]. The authors of [29] incorporated the notion of TL to construct a DL framework for breast cancer detection and classification in breast cytology pictures. A deep TL approach has found much success in the medical sector, where reliability and high accuracy are of the utmost concern. Similarly, TL can also be used for fault classification and identification in induction motors. In [30], the authors have used a TL-based deep CNN model based on ResNet-50 for FD. For FD on two bearing datasets and pump datasets, an FD approach based on the TL-based FD model based on the LeNet-5 structure is proposed in [31]. The TL approach has immense potential. However, its application in FD in EMs is scarce. By considering the potential of TL, this

paper proposes an IRV2-CNN model inspired by the Inception-ResNet-v2 structure for multiple FD in EMs. Our assumption is that the Inception-ResNet-v2 feature extraction layer will perform well on FD too. Also, the proposed technique is compared with the prevailing state-of-the-art approaches for performance investigation.

3. IRV2-CNN Model for FD

The proposed TL-based FD model is described in detail in this section. For transforming one-dimensional vibration signals into images, a simple approach is proposed. The construction of the IRV2-CNN model based on Inception-ResNet-v2 is described, as well as the method for evaluating its performance.

3.1. One-Dimensional Vibration Signal Conditioning

The visual cortex of the brain serves as inspiration for the CNN model. The CNN is a concept that is based on image processing. The vibration-based signals are one-dimensional time-domain signals, and Inception-ResNet-v2 takes two-dimensional images as input. So, one-dimensional current signal-to-image conversion is required. In [32], the authors used the gray-scale images generated from the empirical wavelet transform for fault classification. Bi-spectrum analysis is utilized in [33] to generate images from a one-dimensional signal, and these images were used as input to the FD model. The pre-trained models including Inception-ResNet-v2, ResNet-50, VGG 16, and VGG 19 are trained using the ImageNet dataset, which consists of a vast collection of colorful images in RGB format. As the pre-trained models demand RGB images as input, it is necessary to convert the one-dimensional vibration signals into RGB images before applying the pre-trained network. The default input image size for the Inception-ResNet-v2 structure is $299 \times 299 \times 3$. The conversion of one-dimensional vibration signals collected at 48 kHz into scalogram images is vital for TL application. Scalogram represents the time-frequency distribution of vibration signals and provides valuable insights into the fault characteristics [34]. Using vibration sensors, the one-dimensional vibration signals were gathered at a high sample frequency of 48 kHz. High-frequency components linked to different bearing defects can be captured with enough resolution because of this high sampling rate [35]. After the vibration signal has been preprocessed, it is transformed into a time–frequency representation using the continuous wavelet transform (CWT). The Morlet wavelet has been used because it balances frequency and temporal localization, making it a good choice for evaluating vibration signals that are not stationary [36]. The one-dimensional vibration signal is subjected to the CWT, which results in a time–frequency matrix. The frequency content of the signal is represented by this matrix as it varies over time. The CWT produces a complex-valued output that contains both real and imaginary parts. The magnitude of the CWT coefficients is calculated to produce the scalogram image. The most important characteristic in making a time–frequency map is the magnitude, which shows how strong the frequency components are at each time instance. The wavelet coefficients' magnitudes are scaled on a logarithmic basis to improve the contrast between various frequency bands. This stage is especially crucial because some bearing defects create low-energy signals at particular frequencies that are easier to see on a logarithmic scale. After that, the magnitude values are rearranged and normalized to create a two-dimensional matrix, or scalogram, with frequency on one axis and time on the other. To make the matrix compatible with the proposed design, it is then resized to the required image dimensions. A rich, detailed depiction of the signal's frequency content over time is provided by transforming one-dimensional vibration signals recorded at 48 kHz into scalogram images, which enables efficient bearing FD.

3.2. Transfer Learning

Transfer learning is a process in which a trained model for one activity is fine-tuned and adjusted for another activity. It is a well-known technique in DL. It helps in reducing the computational cost and resources required to train a complete DL model. A model pre-trained on a large base dataset is used in TL to transfer knowledge that will be used on the target dataset and task. The training of a model with random initialization is a cumbersome task, and it takes a lot of computational power and time to achieve good results. The initialization of the model with the pre-trained model allows faster convergence and less computational time. The pre-trained models are trained on a large dataset using hyperparameters that have been carefully chosen. The pre-trained model's few layers can be used without modification or can be slightly modified to adapt to the new task. TL uses knowledge from the source task (TS) in the source domain (DS) to improve the target task (TT) in the target domain (DT).

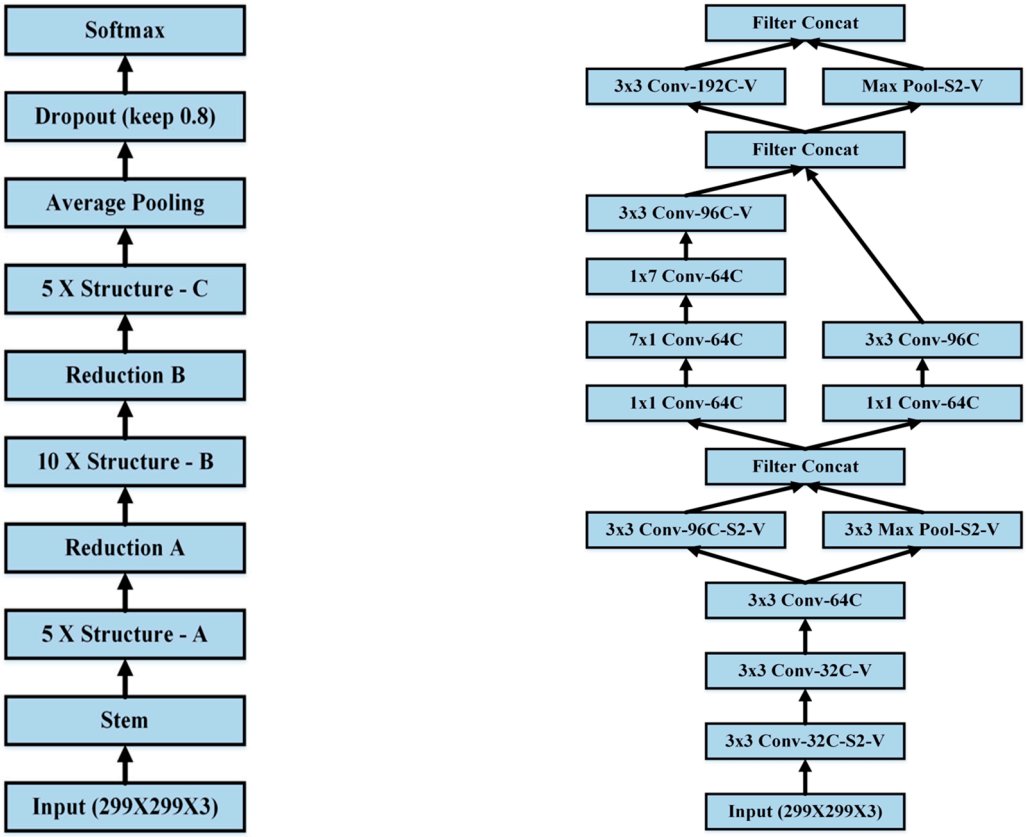
The capabilities of TL can be harnessed in the FD system for higher accuracy and reliability. It facilitates the application of the pre-trained model for faster convergence. Several deep TL models have been studied for FD problems under various ambient settings in recent years [37,38]. The authors of [38] proposed an application of TL in combination with a deep CNN for multiple FD in EMs. In [37], the authors propounded an FD technique based on the DL technique and TL and verified it on multiple datasets. Studies have shown that TL-based DL models are a viable solution for FD in EMs.

3.3. IRV2-CNN Model (Inception-ResNet-v2) Structure

The Inception-ResNet-v2 [19] model is a cutting-edge CNN-based model trained on the ImageNet database. It comprises 164 layers and can segregate images into 1000 distinct classes. In comparison to the millions of images in the ImageNet dataset, the number of tagged fault data are smaller for fault classification. FD can be improved by combining TL with the Inception-ResNet-v2 model. The Inception-ResNet-v2 model, which was trained on the ImageNet dataset, was used in the present work to achieve faster performance and high accuracy in multiple FD in EMs utilizing the knowledge transfer technique. The Inception-ResNet-v2 model, owing to its deep architecture, can extract essential features, and our assumption is that the deep structure of Inception-ResNet-v2 will work well in FD in EMs.

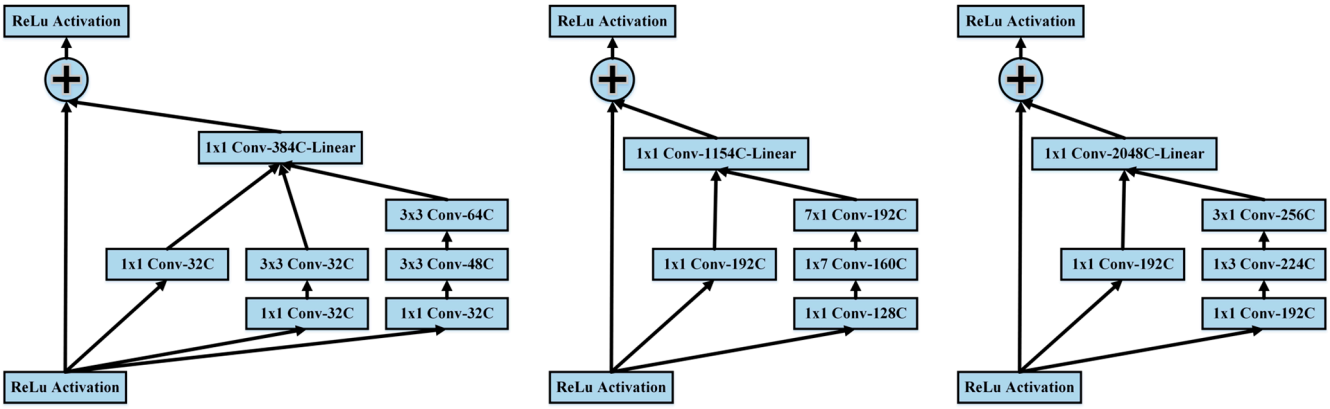
3.3.1. Inception-ResNet-v2 Structure

Inception-ResNet-v2 structure has 164 layers and is a profound network. The comprehensive assembly of Inception-ResNet-v2 is given in Figure 1a. The structure of different modules of Inception-ResNet-v2 like the stem is given in Figure 1b; different inception modules (Structure A, Structure B, and Structure C) are given in Figure 2; and reduction modules (Reduction A and Reduction B) are given in Figure 3. V represents Valid padding in the blocks, and S2 denotes a stride of 2. Batch normalization is used on the top of the convolutional layer, and its use was avoided at the top of the residual summations [19]. The Stem block is the structure's most important layer. It is a layer that comes before the Inception module. The convolution kernel size in the Stem block is 3×3, and the stride is 2 (S2); as a result, the feature map will become smaller, and the parameter values will rightfully decrease as shown in Figure 1b. The Inception module's advantage is the combination of the ResNet architecture and the Inception layer. So, it is known as the Inception-ResNet block. As shown in Figure 2, the Inception-ResNet has three blocks, which are referred to as structures A, B, and C. Reduction blocks isolate the space between the Inception-ResNet blocks because of parameter reduction.



(**a**) (**b**)

Figure 1. (**a**) Overall scheme for the Inception-Resnet-v2 network; (**b**) composition of the Stem [19].



(**a**) (**b**) (**c**)

Figure 2. Scheme for interior grid modules of the Inception-ResNet-v2 network; (**a**) 35 × 35 grid modules; (**b**) 17 × 17 grid modules; (**c**) 8 × 8 grid modules [19].

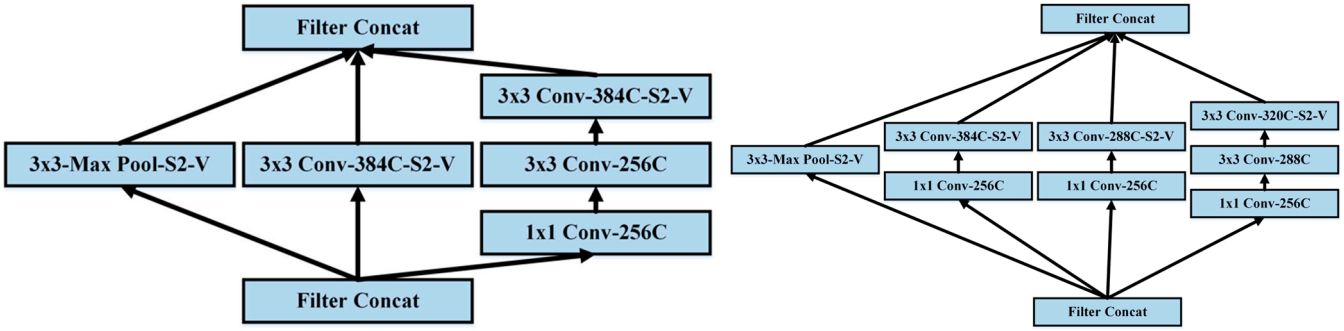


Figure 3. Scheme for interior grid modules of the Inception-ResNet-v2 network; (**a**) Reduction A; (**b**) Reduction B [19].

3.3.2. TL Using Inception-ResNet-v2 Structure

The presented methodology's structure is based on the Inception-ResNet-v2 structure (trained on the ImageNet dataset) as shown in Figure 1. Figure 4 presents the schematic diagram of the proposed methodology, and Figure 5 shows the framework of the proposed Inception-ResNet-v2 based FD model. It is worth mentioning that the pre-trained Inception-ResNet-v2 model has 164 layers of depth. The Inception-ResNet-v2-based FD model's deeper network and stronger feature extraction layers would aid in boosting performance and achieving high FD accuracy. The proposed model uses the Inception-ResNet blocks (enclosed in the lock symbol in Figure 5) of the pre-trained Inception-ResNet-v2 model to extract features from images generated from a one-dimensional current signal. The depth of the network facilitates optimum feature extraction. These features would then be used to train fault classification algorithms. The classifier (wrapped by the unlocking symbol in Figure 5) was fed the retrieved bottleneck features, which included one fully connected layer, a dropout layer, and an output layer with a SoftMax activation function and four nodes. The fully connected layers' weights were chosen at random. The remaining layers were fine-tuned, and the weights of the Inception-ResNet-v2 blocks were the same as for the pre-trained Inception-ResNet-v2 model (contained by the lock symbol in Figure 5). The classifier was trained utilizing a categorical cross-entropy loss function and an adaptive gradient optimizer with an initial learning rate of 0.01. Figure 5 shows the block diagram for the proposed work.

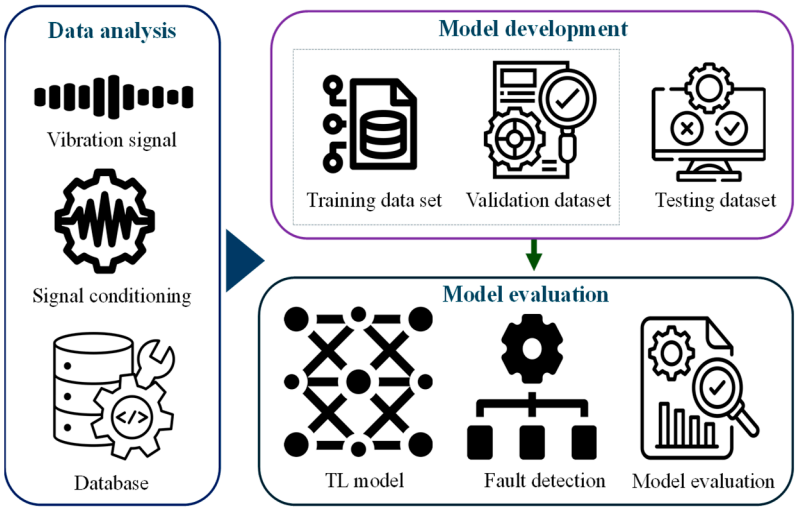


Figure 4. Schematic diagram of the proposed methodology.

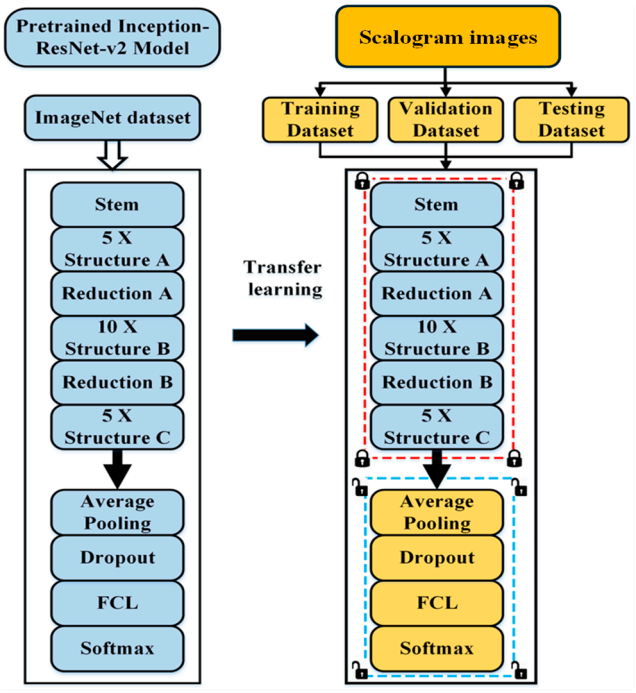


Figure 5. Architecture of the proposed IRV2-CNN model-based FD.

4. Experimental Investigation

4.1. Experimental Setup

The experimental setup comprises a 2 hp EM, torque transducer/encoder, dynamometer, and control electronics [39]. The different bearing faults were emulated by single-point flaws using electro-discharging machining. The vibration data were collected using accelerometers at a 48 kHz sampling frequency. The vibration data were acquired under different health conditions like inner race fault (FIR), outer race fault (FOR), and ball bearing fault (FBB). Based on the dataset, different health states of the bearing like FIR, FOR, FBB, and normal (N) were considered for the analysis. This dataset was utilized to develop a database and develop a TL-based FD methodology. The proposed approach utilizes the database developed using the vibration data obtained under various operating conditions as well as states like FIR, FOR, FBB, and N. The data for the FD model were acquired from the experimental setup shown in Figure 6. The Inception-ResNet-v2 structure requires an input image, and one-dimensional vibration signals were converted to an image of this size using the technique described in the previous section. The proposed FD model should perform a four-class segregation task as there are four bearing states. The overall layout is shown in Figure 5.

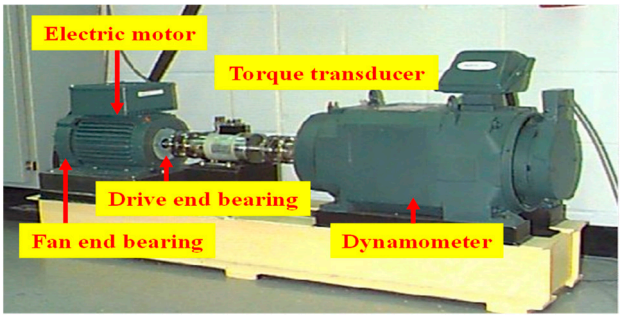


Figure 6. Experimental setup for vibration data collection under different bearing faults [39].

4.2. Results

The proposed method aims to identify four different health states based on the vibration data of an EM, and it uses a deep architecture with residual connections to achieve accurate FD. The performance of the proposed Inception-ResNet-v2 model is evaluated using a confusion matrix (Figure 7), which shows the predicted labels on the horizontal axis and the true labels of the samples on the vertical axis. To reduce overfitting, the proposed Inception-ResNet-v2 model uses regularization techniques such batch normalization and dropout layers. To make sure the results are consistent and not unique to a particular train–test configuration, we also conducted many training and testing cycles utilizing shuffled data splits. The confusion matrix demonstrates that the chances of misclassification of certain states, such as FIR, FOR, FBB, and N, were minimal. The mean accuracy of the proposed model was 99.80%, indicating its effectiveness in accurately identifying the different states of the EM. This accuracy measurement helps in analyzing the overall system performance on overall available classes. Training and testing of the developed framework were implemented multiple times to achieve the best result. Also, for the in-depth analysis, the performance indices like precision (p), recall (r), and F-1 score (F1) were calculated. The values of p, r, and F1, as shown in Table 1, were over 0.99, which renders the high performance of the model. The chance of misclassification between two states was below 0.01, which is vital for developing an effective FD system. The deep architecture of the model also helps to boost the domain adaptability of the model too.

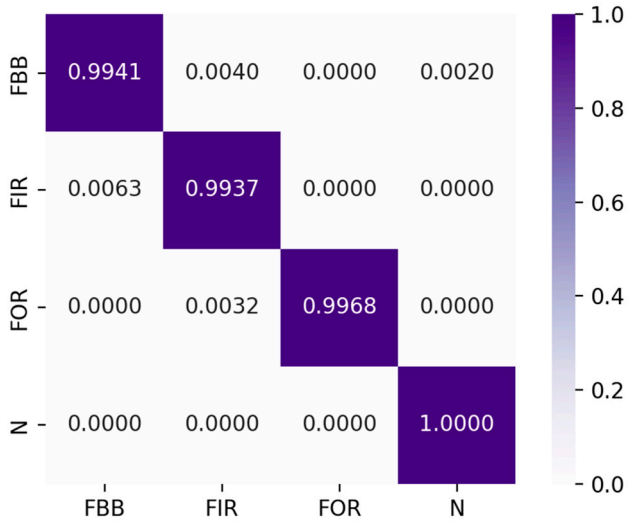


Figure 7. Confusion matrix for the proposed model.

Table 1. Performance indices of the proposed model.

State	FIR	FOR	FBB	N
p	0.99	1.0	0.99	1.0
r	0.99	1.0	0.99	1.0
F1	0.99	1.0	0.99	1.0

The proposed model's performance was compared to that of a shallow CNN model with five convolution layers and a fully connected layer. The shallow CNN model achieved a mean accuracy of 90.60%, which is lower than that of the proposed model. The performance indices like precision (p), recall (r), and F-1 score (F1) are given in Table 2. The performance indices were between 0.86 and 0.95, which indicates that the chances of misclassification were on the higher side. The proposed model was also compared with the

other established techniques. The receiver operating characteristic (ROC) curve is key metric for evaluating the model's performance. The diagonal dashed line represents a random classifier (area under the curve (AUC) = 0.5), and it displays the true positive rate (TPR) against the false positive rate (FPR) at different threshold levels. A model is considered to be highly accurate in differentiating between classes if its AUC is near 1.0. It is evident from Figure 8 that the model achieved outstanding classification accuracy across all the classes, with AUC values ranging from 0.98 to 1.00. The bearing states, such as FIR and N, had a precise AUC of 1.00, which indicates that the model accurately separated these categories and produced neither false positives nor false negatives. With an AUC of 0.99, the bearing state FOR came in second, showing almost flawless classification. The somewhat lower AUC of 0.98 for the bearing state FBB indicates a modest improvement margin, because of small misclassifications. The model's strong sensitivity and low misclassification rate are further supported by the close clustering of the ROC curves in the upper-left corner. Table 3 presents a comparison of the accuracy of the proposed DL model with other established models that employ different approaches, including deep belief network (DBN) [22], hierarchical CNN (HCNN) [40], empirical mode decomposition-based CNN (CNN-EMD) [41], deep CNN (DCNN) [23], adaptive deep CNN (ADCNN) [24], and multi-CNN (MCNN) [25].

Table 2. Performance indices of shallow CNN model.

State	FIR	FOR	FBB	N
p	0.93	0.94	0.90	0.83
r	0.93	0.94	0.84	0.90
F1	0.93	0.94	0.87	0.86

Table 3. Comparison of the proposed method with DL and CNN-based FD method.

Methods	Mean Accuracy (%)
IRV2-CNN	99.80
DBN	99.03
HCNN	92.60
CNN-EMD	99.7
ADCNN	98.1
MCNN	99.41
DCNN	99.70

Also, the proposed model was compared with the TL-based FD models like the VGG19-based FD model (VGG19-FD) [38], the ResNet-50-based FD model (ResNet50-FD) [30], and the LeNet-5-based FD model (LeNet5-FD) [31]. The accuracy of the proposed model and the number of weighted layers was compared with these models and is tabulated in Table 4. In [38], the VGG19-based FD model was used, and it achieved an accuracy of 99.40%. In [30], a pre-trained ResNet-50 was used for TL and achieved an average accuracy of 99.38% (average of accuracies in three cases). In [31], a pre-trained LeNet-5 was used for TL and achieved an average accuracy of 99.66% (average of accuracies in three cases). The proposed model achieved a mean accuracy of 99.80%. It can be seen from Tables 3 and 4 that the proposed structure outperformed the other established FD models. According to Table 3, the proposed DL model had an accuracy of 99.80%, which is higher than that for the other established models. The DCNN model has an accuracy of 99.70%, which is less than that of the proposed model. The accuracy of other models such as the MCNN, ADCNN, HCNN, and DBN was also lower than that of the proposed model. The deep residual architecture of the IRV2-CNN model, which efficiently extracts intricate patterns

from the vibration data and improves feature learning, is responsible for the great accuracy attained by that model. By allowing for deeper representation learning and mitigating the vanishing gradient problem, residual connections enhance classification accuracy across all bearing states. The model's robustness is further supported by the good precision, recall, and F1 scores (>0.99) as well as the low misclassification seen in the confusion matrix. Strong sensitivity and specificity are also indicated by the AUC values, which ranged from 0.98 to 1.00, particularly in important fault states like FOR and FIR. The proposed model's superior performance can be attributed to its deep architecture, which facilitates better feature learning, leading to improved performance and domain adaptability. Additionally, the proposed model provides an end-to-end learning solution for multiple FD.

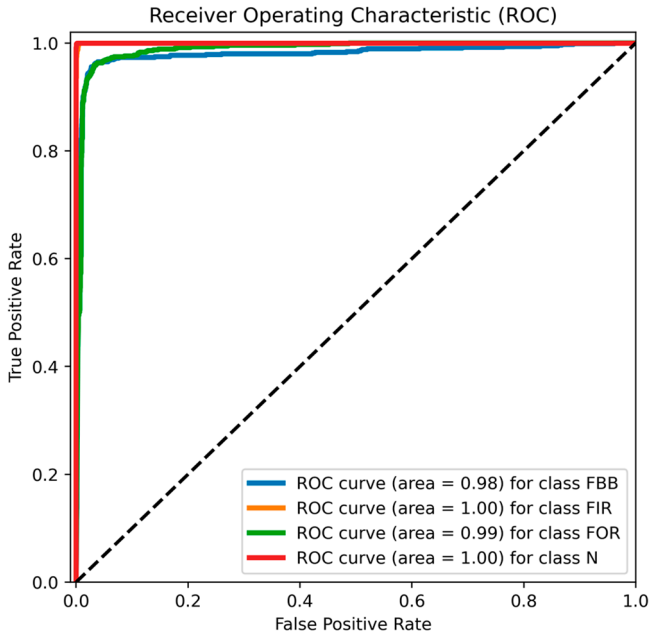


Figure 8. ROC curve for the proposed model.

Table 4. Comparison of the proposed method with TL-based FD methods.

Methods	Mean Accuracy (%)
IRV2-CNN	99.80
VGG19-FD	99.40
ResNet50-FD	99.38
LeNet5-FD	99.66

5. Conclusions

This paper proposes a new and reliable method for detecting bearing faults in an EM. The end-to-end learning solution is provided by the deep structure based on the Inception-ResNet-v2 structure. Despite the deep structure, the application of TL reduces the training time and makes FD faster and smoother. The contributions of this paper are listed below:

- Domain adaptability owing to deep architecture;
- Less training time due to TL;
- Fast decision making;
- End-to-end learning solution;
- Computationally viable;
- Efficient performance with high accuracy.

The proposed method achieved high accuracy and performed faster despite the deep structure due to TL. It offers a practical method for detecting faults using raw current signals. The proposed FD technique outperformed different state-of-the-art FD techniques based on TL using pre-trained VGG19, ResNet-50, and LeNet-5 models. Also, the proposed method repressed the requirement for manual feature extraction and selection. The developed technique can efficiently detect bearing faults and can also be extended to other faults in the future.

Notwithstanding these advantages, the proposed strategy has some drawbacks too. The evaluation was conducted using the CWRU dataset, which, while standard and widely used, represents controlled laboratory conditions. The model's performance in actual industrial settings with higher noise levels and operational variability has to be examined. The model was trained and verified on data collected in a controlled laboratory setting. Furthermore, even if the model performs well, its very complex structure might make it difficult to deploy in contexts with limited resources, like edge devices. Furthermore, more research is needed to determine whether the current method can be extended to complex or previously unknown issue categories, as it is intended to identify a predetermined set of fault types. This study's emphasis on diagnostic performance meant that metrics of computational complexity like FLOPs, FPS, and parameter counts were left out, but they will be covered in further research. Future research will concentrate on expanding the assessment to a variety of industrial datasets, investigating the model's resilience to noise, carrying out in-depth ablation investigations, and refining the architecture for implementation in resource-constrained settings like edge computing platforms.

Author Contributions: Conceptualization, P.K., S.S. and D.Y.S.; methodology, P.K.; software, P.K.; validation, P.K., S.S. and D.Y.S.; formal analysis, P.K. and S.S.; investigation, P.K.; resources, P.K. and S.S.; writing—original draft preparation, P.K. and S.S.; writing—review and editing, P.K., S.S. and D.Y.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data will be made available on request.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Kumar, P.; Hati, A.S. Review on Machine Learning Algorithm Based Fault Detection in Induction Motors. *Arch. Comput. Methods Eng.* **2021**, *28*, 1929–1940. [CrossRef]
2. Antonino-Daviu, J.; Aviyente, S.; Strangas, E.G.; Riera-Guasp, M. Scale Invariant Feature Extraction Algorithm for the Automatic Diagnosis of Rotor Asymmetries in Induction Motors. *IEEE Trans. Ind. Inform.* **2013**, *9*, 100–108. [CrossRef]
3. Dai, X.; Gao, Z. From Model, Signal to Knowledge: A Data-Driven Perspective of Fault Detection and Diagnosis. *IEEE Trans. Ind. Inform.* **2013**, *9*, 2226–2238. [CrossRef]
4. Kudelina, K.; Raja, H.A.; Naseer, M.U.; Autsou, S.; Asad, B.; Vaimann, T.; Kallaste, A. Study of Bearing Currents in Induction Machine: Diagnostic Possibilities, Fault Detection, and Prediction. *Electr. Eng.* **2024**, *106*, 7089–7102. [CrossRef]
5. Kudelina, K.; Raja, H.A. Neuro-Fuzzy Framework for Fault Prediction in Electrical Machines via Vibration Analysis. *Energies* **2024**, *17*, 2818. [CrossRef]
6. Ali, M.Z.; Shabbir, M.N.S.K.; Liang, X.; Zhang, Y.; Hu, T. Machine Learning-Based Fault Diagnosis for Single- and Multi-Faults in Induction Motors Using Measured Stator Currents and Vibration Signals. *IEEE Trans. Ind. Appl.* **2019**, *55*, 2378–2391. [CrossRef]
7. Kankar, P.K.; Sharma, S.C.; Harsha, S.P. Fault Diagnosis of Ball Bearings Using Machine Learning Methods. *Expert Syst. Appl.* **2011**, *38*, 1876–1886. [CrossRef]
8. Andre, A.B.; Beltrame, E.; Wainer, J. A Combination of Support Vector Machine and k-nearest Neighbors for Machine Fault Detection. *Appl. Artif. Intell.* **2013**, *27*, 36–49. [CrossRef]
9. Razzaq, K.; Shah, M. Machine Learning and Deep Learning Paradigms: From Techniques to Practical Applications and Research Frontiers. *Computers* **2025**, *14*, 93. [CrossRef]

10. Archana, R.; Jeevaraj, P.S.E. Deep Learning Models for Digital Image Processing: A Review. *Artif. Intell. Rev.* **2024**, *57*, 11. [CrossRef]
11. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*, Illustrated edition; The MIT Press: Cambridge, MA, USA, 2016; ISBN 978-0-262-03561-3.
12. Zhao, R.; Yan, R.; Chen, Z.; Mao, K.; Wang, P.; Gao, R.X. Deep Learning and Its Applications to Machine Health Monitoring. *Mech. Syst. Signal Process.* **2019**, *115*, 213–237. [CrossRef]
13. Kumar, P.; Hati, A.S. Deep Convolutional Neural Network Based on Adaptive Gradient Optimizer for Fault Detection in SCIM. *ISA Trans.* **2021**, *111*, 350–359. [CrossRef]
14. Kumar, P.; Raouf, I.; Song, J.; Kim, H.S. Multi-Size Wide Kernel Convolutional Neural Network for Bearing Fault Diagnosis. *Adv. Eng. Softw.* **2024**, *198*, 103799. [CrossRef]
15. Raouf, I.; Kumar, P.; Soo Kim, H. Deep Learning-Based Fault Diagnosis of Servo Motor Bearing Using the Attention-Guided Feature Aggregation Network. *Expert Syst. Appl.* **2024**, *258*, 125137. [CrossRef]
16. Zhong, G.; Ling, X.; Wang, L.-N. From Shallow Feature Learning to Deep Learning: Benefits from the Width and Depth of Deep Architectures. *WIREs Data Min. Knowl. Discov.* **2019**, *9*, e1255. [CrossRef]
17. Bengio, Y. Learning Deep Architectures for AI. *MAL* **2009**, *2*, 1–127. [CrossRef]
18. Cohen, N.; Sharir, O.; Shashua, A. On the Expressive Power of Deep Learning: A Tensor Analysis. In Proceedings of the 29th Annual Conference on Learning Theory, New York, NY, USA, 6–9 July 2016.
19. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A.A. Inception-v4, Inception-Resnet and the Impact of Residual Connections on Learning. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 12 February 2017.
20. Wen, L.; Li, X.; Gao, L.; Zhang, Y. A New Convolutional Neural Network-Based Data-Driven Fault Diagnosis Method. *IEEE Trans. Ind. Electron.* **2018**, *65*, 5990–5998. [CrossRef]
21. Shao, H.; Jiang, H.; Zhang, H.; Liang, T. Electric Locomotive Bearing Fault Diagnosis Using a Novel Convolutional Deep Belief Network. *IEEE Trans. Ind. Electron.* **2018**, *65*, 2727–2736. [CrossRef]
22. Gan, M.; Wang, C. Construction of Hierarchical Diagnosis Network Based on Deep Learning and Its Application in the Fault Pattern Recognition of Rolling Element Bearings. *Mech. Syst. Signal Process.* **2016**, *72*, 92–104. [CrossRef]
23. Fuan, W.; Hongkai, J.; Haidong, S.; Wenjing, D.; Shuaipeng, W. An Adaptive Deep Convolutional Neural Network for Rolling Bearing Fault Diagnosis. *Meas. Sci. Technol.* **2017**, *28*, 095005. [CrossRef]
24. Guo, X.; Chen, L.; Shen, C. Hierarchical Adaptive Deep Convolution Neural Network and Its Application to Bearing Fault Diagnosis. *Measurement* **2016**, *93*, 490–502. [CrossRef]
25. Xia, M.; Li, T.; Xu, L.; Liu, L.; de Silva, C.W. Fault Diagnosis for Rotating Machinery Using Multiple Sensors and Convolutional Neural Networks. *IEEE/ASME Trans. Mechatron.* **2018**, *23*, 101–110. [CrossRef]
26. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How Transferable Are Features in Deep Neural Networks? *arXiv* **2014**, arXiv:1411.1792.
27. Dhruba, S.R.; Rahman, R.; Matlock, K.; Ghosh, S.; Pal, R. Application of Transfer Learning for Cancer Drug Sensitivity Prediction. *BMC Bioinform.* **2018**, *19*, 497. [CrossRef]
28. Talo, M.; Baloglu, U.B.; Yıldırım, Ö.; Rajendra Acharya, U. Application of Deep Transfer Learning for Automated Brain Abnormality Classification Using MR Images. *Cogn. Syst. Res.* **2019**, *54*, 176–188. [CrossRef]
29. Khan, S.; Islam, N.; Jan, Z.; Ud Din, I.; Rodrigues, J.J.P.C. A Novel Deep Learning Based Framework for the Detection and Classification of Breast Cancer Using Transfer Learning. *Pattern Recognit. Lett.* **2019**, *125*, 1–6. [CrossRef]
30. Wen, L.; Li, X.; Gao, L. A Transfer Convolutional Neural Network for Fault Diagnosis Based on ResNet-50. *Neural Comput. Applic* **2020**, *32*, 6111–6124. [CrossRef]
31. Xu, G.; Liu, M.; Jiang, Z.; Shen, W.; Huang, C. Online Fault Diagnosis Method Based on Transfer Convolutional Neural Networks. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 509–520. [CrossRef]
32. Hsueh, Y.-M.; Ittangihal, V.R.; Wu, W.-B.; Chang, H.-C.; Kuo, C.-C. Fault Diagnosis System for Induction Motors by CNN Using Empirical Wavelet Transform. *Symmetry* **2019**, *11*, 1212. [CrossRef]
33. Lu, C.; Wang, Y.; Ragulskis, M.; Cheng, Y. Fault Diagnosis for Rotating Machinery: A Method Based on Image Processing. *PLoS ONE* **2016**, *11*, e0164111. [CrossRef]
34. Zaman, W.; Ahmad, Z.; Siddique, M.F.; Ullah, N.; Kim, J.-M. Centrifugal Pump Fault Diagnosis Based on a Novel SobelEdge Scalogram and CNN. *Sensors* **2023**, *23*, 5255. [CrossRef] [PubMed]
35. El Bouharrouti, N.; Morinigo-Sotelo, D.; Belahcen, A. Multi-Rate Vibration Signal Analysis for Bearing Fault Detection in Induction Machines Using Supervised Learning Classifiers. *Machines* **2024**, *12*, 17. [CrossRef]
36. Łuczak, D. Machine Fault Diagnosis through Vibration Analysis: Continuous Wavelet Transform with Complex Morlet Wavelet and Time–Frequency RGB Image Recognition via Convolutional Neural Network. *Electronics* **2024**, *13*, 452. [CrossRef]
37. Shao, S.; McAleer, S.; Yan, R.; Baldi, P. Highly Accurate Machine Fault Diagnosis Using Deep Transfer Learning. *IEEE Trans. Ind. Inform.* **2018**, *15*, 2446–2455. [CrossRef]

38. Kumar, P.; Hati, A.S.; Padmanaban, S.; Leonowicz, Z.; Chakrabarti, P. Amalgamation of Transfer Learning and Deep Convolutional Neural Network for Multiple Fault Detection in SCIM. In Proceedings of the 2020 IEEE International Conference on Environment and Electrical Engineering and 2020 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe), Madrid, Spain, 9 June 2020; pp. 1–6.
39. Loparo, K.A.; Case Western Reserve University Bearing Data Center. *Bearings Vibration Data Sets*; Case Western Reserve University: Cleveland, OH, USA, 2012; pp. 22–28.
40. Lu, C.; Wang, Z.; Zhou, B. Intelligent Fault Diagnosis of Rolling Bearing Using Hierarchical Convolutional Network Based Health State Classification. *Adv. Eng. Inform.* **2017**, *32*, 139–151. [CrossRef]
41. Xie, Y.; Zhang, T. Fault Diagnosis for Rotating Machinery Based on Convolutional Neural Network and Empirical Mode Decomposition. *Shock. Vib.* **2017**, *2017*, 3084197. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.