

# New Transfer Learning Approach Based on a CNN for Fault Diagnosis <sup>†</sup>

Alasmer Ibrahim <sup>1,\*</sup> , Fatih Anayi <sup>1</sup> and Michael Packianather <sup>2</sup><sup>1</sup> Wolfson Centre for Magnetics, Cardiff University, Cardiff CF24 3AA, UK<sup>2</sup> High-Value Manufacturing Group, Cardiff University, Cardiff CF24 3AA, UK

\* Correspondence: Ibrahim21@cardiff.ac.uk; Tel.: +44-7774221654

<sup>†</sup> Presented at the 1st International Electronic Conference on Machines and Applications, 15–30 September 2022; Available online: <https://iecma2021.sciforum.net>.

**Abstract:** Induction motors operate in difficult environments in the industry. Monitoring the performance of motors in such circumstances is significant, which can provide a reliable operation system. This paper intends to develop a new model for fault diagnosis based on the knowledge of transfer learning using the ImageNet dataset. The development of this framework provides a novel technique for the diagnosis of single and multiple induction motor faults. A transfer learning model based on a VGG-19 convolutional neural network (CNN) was implemented, which provided a quick and fast training process with higher accuracy. Thermal images with different induction motor conditions were captured with the help of an FLIR camera and applied as inputs to investigate the proposed model. The implementation of this task involved the use of a VGG-19 CNN-based pre-trained network, which provides autonomous features learning based on minimum human intervention. Next, a dense-connected classifier was applied to predict the true class. The experimental results confirmed the robustness and reliability of the developed technique, which was successfully able to classify the induction motor faults, achieving a classification accuracy of 99.8%. The use of a VGG-19 network allowed the attributes to be automatically extracted and associated with the decision-making part. Furthermore, this model was further compared with other applications based on related topics; it successfully proved its superiority and robustness.

**Keywords:** VGG-19; induction motor thermal images; fault diagnosis; pretrained model; deep learning networks



**Citation:** Ibrahim, A.; Anayi, F.; Packianather, M. New Transfer Learning Approach Based on a CNN for Fault Diagnosis. *Eng. Proc.* **2022**, *24*, 16. <https://doi.org/10.3390/IECMA2022-12905>

Academic Editor: Antonio J. Marques Cardoso

Published: 16 September 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 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 to Achieving the Performance

Induction motors are the backbone of industry applications because of the production dependent on them. Hence, early maintenance is required to avoid motor breakdown. Monitoring regularly the condition of the motor can achieve improvement in the availability and production system [1]. If an initial defect is not identified, damage could be caused in other motor elements and the system could collapse, leading to massive losses in production. Standard maintenance is required for machines to achieve a high level of production. This maintenance includes condition-monitoring approaches and artificial intelligence techniques based on fault diagnosis.

Induction motor fault diagnosis is a topic that may be researched based on three different categories of attractive research: fault diagnosis based on knowledge, fault diagnosis based on models, and fault diagnosis based on the signal. A hybrid model can be created by combining these different approaches [2]. Some approaches based on modelling and identification are used in the diagnosis phase during the determination process in the industry-based individual component. So, the examination of consistency across anticipated systems can establish the cause of the presented failure. Briefly, some domain techniques have been applied by considering the system pattern in fault diagnosis

approaches based on the signal rather than dealing with system-based models [3]. The knowledge-based fault diagnosis approach, in contrast to the model-based and signal-based fault detection methods, does not require an accurate model or signal pattern in order to perform the diagnosis. In contrast, for establishing a link between raw data and the outcome, a historical process is required in the knowledge-based fault diagnosis approach. The methodology of excitation fault diagnosis is achieved by the installation of multiple sensors precisely for data acquisition. Next, these data are processed to verify the type of the fault. Generally, the methodology of fault diagnosis is performed by following three steps: data collection, extraction and selection the optimum features, and fault classification. In the data collection step, data are captured when the machine is running using a specific sensor, such as a current transformer, to collect the current data and using cameras to capture thermal image data. Traditionally, feature extraction is implemented through an accurate domain, such as the time domain, frequency domain, and time-frequency domain. Next, these features are further processed using feature selection algorithms. In the fault classification step, the obtained features are used to train traditional machine learning classifiers to predict the correct class. With the continuing growth of intelligent fault diagnosis, several fault diagnosis systems, such as expert systems, have emerged [4]. An artificial neural network was developed in [5]; in [6], an efficient application using a KNN was proposed; in [7], another fault diagnosis model applying an SVM classifier was presented; a robust model in [8], based on the use of a random forest classifier, was anticipated to diagnose multiple faults; and a recent model combining the use of an invasive weed optimization algorithm was suggested in [9,10]. These traditional machine learning algorithms have a restricted ability to analyze all data that have been acquired by the sensors [11]. In addition, these approaches use feature extraction and feature selection to generate insufficient classifiers, which are based on handcrafted features and human feature selection that is inadequate. In addition, it has been reported in many studies that the use of handcrafted features with different categorization tasks is a specific task-based approach, which means that features that are used to accurately predict model outcomes under specific conditions are unsuitable for use in other scenarios [12]. Moreover, it is difficult to come up with a collection of attributes that are capable of making accurate predictions under all scenarios. As a result of its formidable capabilities, deep learning (DL) is an effective way to tackle these challenges [13]. DL application is performed without the assistance of human engineers; it investigates the first attributes for the classification stage by exploring them directly from the data that have been collected by the sensors [14]. In addition, in the training process, the architecture of the deep network can automatically select optimum attributes that make an accurate prediction in the classification part. In recent years, DL has become increasingly prominent in the field of computer science due to its increased processing power [15]. Various DL algorithms have been proposed in many science areas, such as computer vision [16], natural language processing [17], and games [18]. In addition, DL has shown itself to be a promising candidate in the area of defect detection [19], for example, the use of a convolutional neural network (CNN) in [20]. In [21], another kind of DL network, the so-called recurrent neural network (RNN), is proposed; in [22], a deep Boltzmann machine (DBM) technique is presented; and a deep belief network (DBN) is considered for creating an efficient model in [22]. Despite the reality that DL models have proven effective applications in machine fault diagnosis interests, there are still issues with this approach. For example, most deep models that have been used in most of the publications cited earlier have a small number of hidden layers. Additionally, when the number and size of hidden layers increase, the number of network parameters is affected, resulting in involving a large amount of data for an efficient training process. However, deep networks with more than 10 hidden layers have not been developed yet, and hyperparameter tuning influences the performance of the model. Hence, the transfer learning (TL) methodology is applied to overcome this problem. This technique can be applied using a deep neural network for extracting high-level features from the original data (raw data) [23]. Moreover, the challenge of fault diagnosis presents opportunities for

deep transfer learning to deliver potentially useful solutions. Many engineering field and scientific challenges, such as text classification and spam filtering, have shown the transfer learning method's excellence and robustness [24]. Another factor to be considered is that deep transfer learning's layer-by-layer learning structure can allow it to build large data representations, which makes it possible for the performance of the fault diagnosis to be greatly improved, with a reduction in the extraction and training error [25]. According to the knowledge that has been reported, VGG-19 model-based networks that are paired with thermal imaging data have not been the subject of any published research for the purpose of fault diagnosis in induction motors. Hence, this research work proposes a new approach for induction motor fault detection, which is built based on the combination of induction motor thermal images with a pre-trained model as a feature extractor based on the Visual Geometry Group (VGG). The contribution of this research is to present an efficient fault diagnosis application to identify different induction motor conditions. The presented model uses thermal images that are further pre-processed by applying a data augmentation technique, a deep transfer learning model based on a pre-trained (VGG-19) network, and the adjusted densely connected layer for training and classifying the model. This model uses many deep hidden layers to learn hierarchical representations for achieving an accurate model. The performance of the proposed application applying a pretrained model is validated using thermal images of the induction motor.

The remainder of this paper is presented as follows: Related work is presented in Section 2, the proposed model is presented in Section 3, and materials and methods are presented in Section 4. Section 5 reports the results in detail. A discussion is provided in Section 6. Lastly, the conclusion is given in Section 7.

## 2. Related Work

Thermal images of induction motors have been used in several studies and research projects to successfully identify flaws in the motors [26]. However, a few research applications combining thermal images and deep transfer learning (DTL) approaches have been attained for the induction motor fault diagnosis task. Model-based transfer learning is a method for transferring previously learnt model parameters to new datasets in order to improve training efficiency. This technique takes into consideration the correlation between two datasets for further increasing the overall training accuracy by using this technique. A new fault detection model was proposed by Yang [27], which uses a transfer learning network and the trained parameters for a new training model for decreasing both training time and training data. A high-accuracy model using transfer learning was suggested in [28] using sensor data that converted to images. The obtained results achieved a classification accuracy near 100%. In [29], a deep learning model was proposed to solve cross-domain data learning using the vibration data of 48 bearing. The experimental results achieved a classification accuracy of 93%. In [30], another deep learning model was suggested using a 1D signal for machine fault diagnosis and classification. VGG-19 achieved excellent results in the experiments of the gearbox. In addition, a novel model-based CNN for multiple faults of induction was proposed in [23] using the current signal. The results demonstrated that this model outperforms other methods based on the state of the art. In this work, a motor fault diagnosis framework based on a deep transfer learning model is proposed using thermal images.

## 3. Proposed Model

Deep CNNs are used to build the proposed framework for detecting the operational conditions of the induction motor with great precision and accuracy applying thermal images as input. As mentioned earlier, a transfer learning network based on a pre-trained model can help in performance improvement. This paper proposes a pipeline that diagnoses the failure of the induction motor automatically from the thermal images. The procedure of this application is presented in Figure 1, which includes capturing the thermal images, preparing the data, building the pretrained model by applying VGG-19, and classifying the

model. A deep CNN generated by the Oxford Visual Geometry Group (VGG) in [23] was implemented in this work as a pretrained model. This network has 19 layers, and it was trained on ImageNet-based weight. More convolutional layers were added for fine-tuning the thermal images data. First, the pretrained model was tuned after removing some of its layers and replacing them with the output layer that had the same size as the number of motor faults (conditions). The output layer just added was weighted randomly. The earlier layers of this network were frozen in the training process, and the weights were set to ImageNet for reducing the error between the true and predicted labels. The testing dataset was used in the classification stage to validate the robustness of the proposed model based on the induction motor conditions.

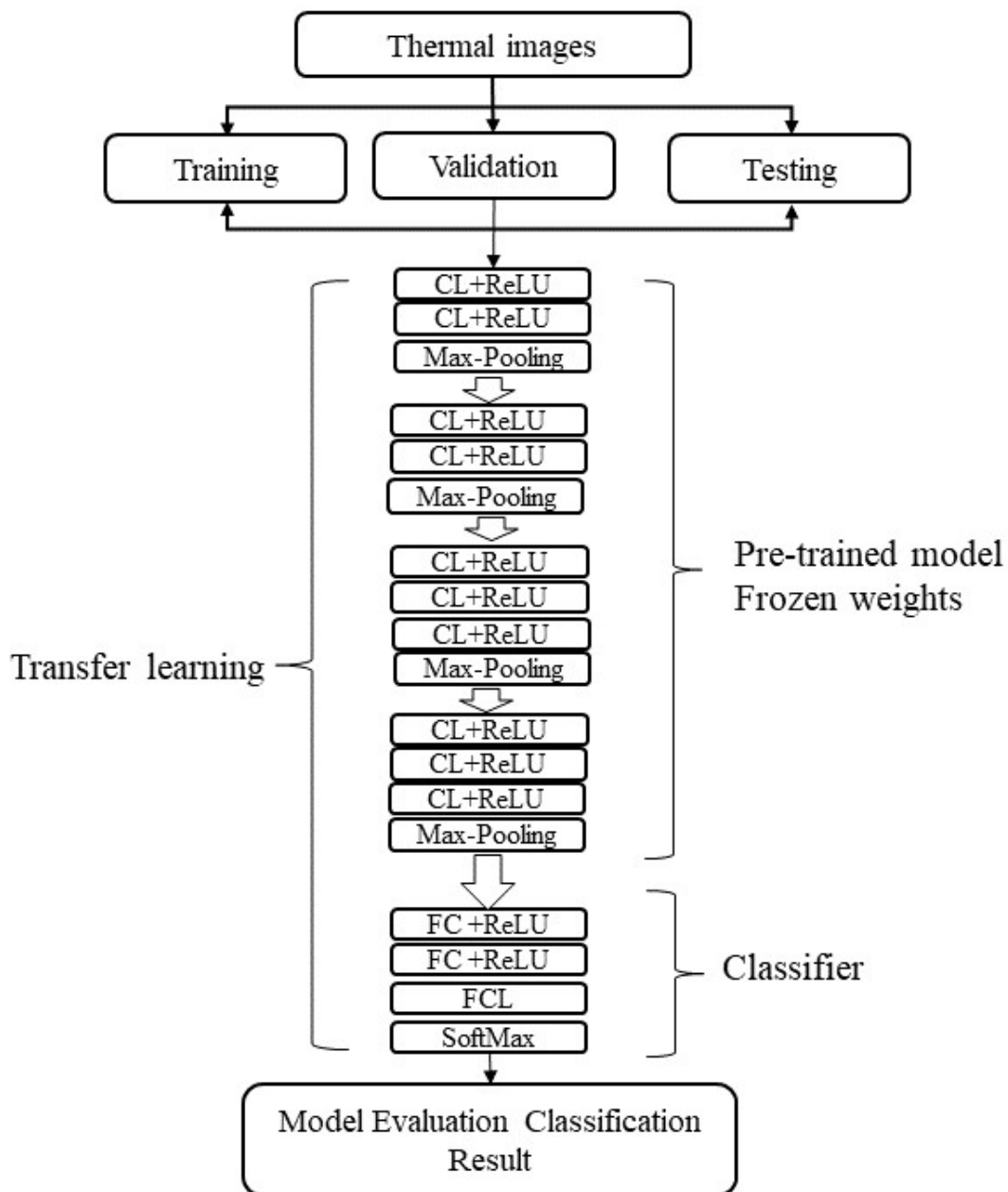


Figure 1. The proposed framework for fault detection applying VGG-19.

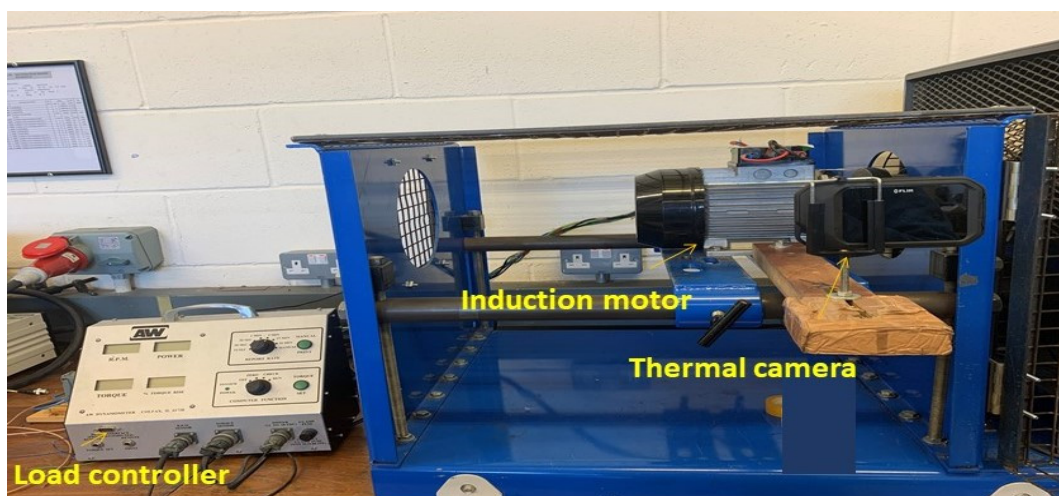
## 4. Materials and Methods

### 4.1. Data Collection

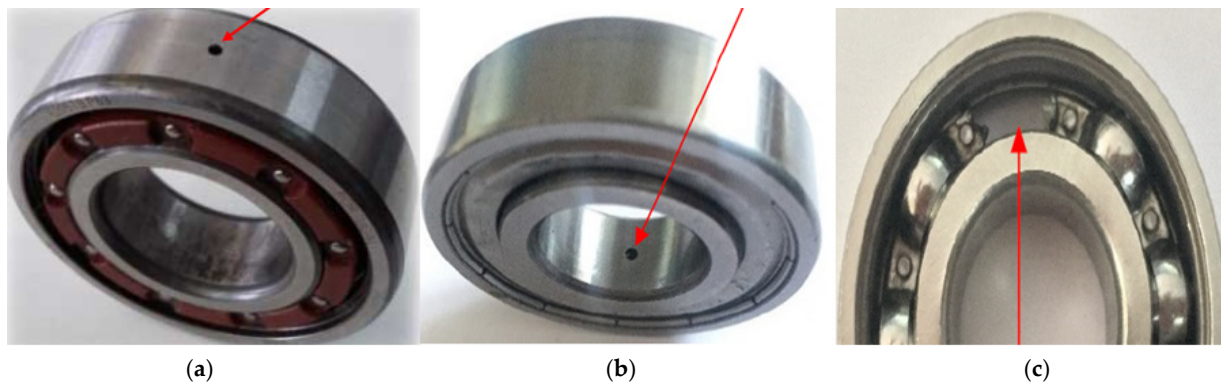
During the course of this investigation, thermal images of the induction motor were carefully acquired with the consideration of both healthy and faulty states, as detailed in Table 1. The motor was tested in the laboratory at two different speeds, namely 1480 and 1380 revolutions per minute. The examinations were performed in the lab of Cardiff University, which is located in the United Kingdom. Figure 2 depicts the testing apparatus. On the bearing inner and outer races, an artificial pit measuring 0.25 cm was created to simulate the inner bearing fault (IBF) and the outer bearing fault (OBF), as seen in Figures 3a and 3b, respectively. To create a ball bearing fault (BBF), as displayed in Figure 3c, a single ball was removed from its cage in the bearing. A broken rotor bar fault (1BRBF), as seen in Figure 4, was accomplished by drilling a cavity into one bar of the motor rotor. This cavity had a certain depth (cm) and a certain in diameter (cm). For the fifth and eighth instances of the broken rotor bar failures, the same technique was used. This means five bars were drilled to generate five broken rotor bar faults (5BRBF), as shown in Figure 4b, and eight bars were bored, as shown in Figure 4c, to produce eight broken rotor bar faults (8BRBF). Multi-induction motor faults were also studied in this research. These faults were specifically produced so that they could be shown as a new state of the motor. As an example, labels 8 and 9 included the inner bearing fault with a broken rotor bar into one label (IBF+1BRBF) and the outer bearing fault with five broken rotor bars in one label (OBF+5BRBF), respectively. In both of these labels, the rotor had a broken rotor bar fault.

**Table 1.** Motor conditions.

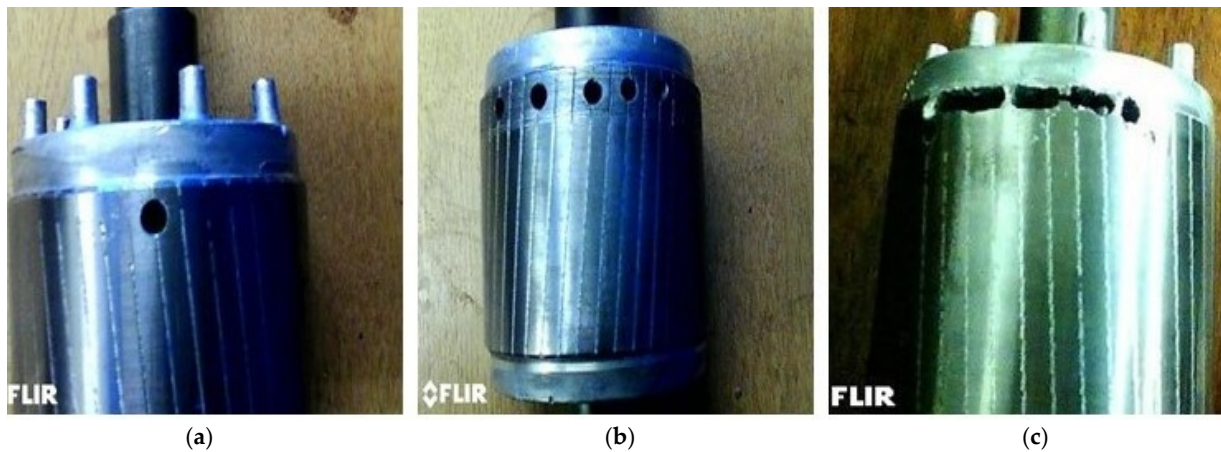
Fault Mode	Motor Load (rpm)	Image No.	Class Label
Normal motor	1480/1380	250	1
IBF	1480/1380	250	2
OBF	1480/1380	250	3
BBF	1480/1380	250	4
1BRBF	1480/1380	250	5
5BRBF	1480/1380	250	6
8BRBF	1480/1380	250	7
IBF + 1BRBF	1480/1380	250	8
OBF + 5BRBF	1480/1380	250	9
BBF + 8BRBF	1480/1380	250	10



**Figure 2.** Experimental test rig.

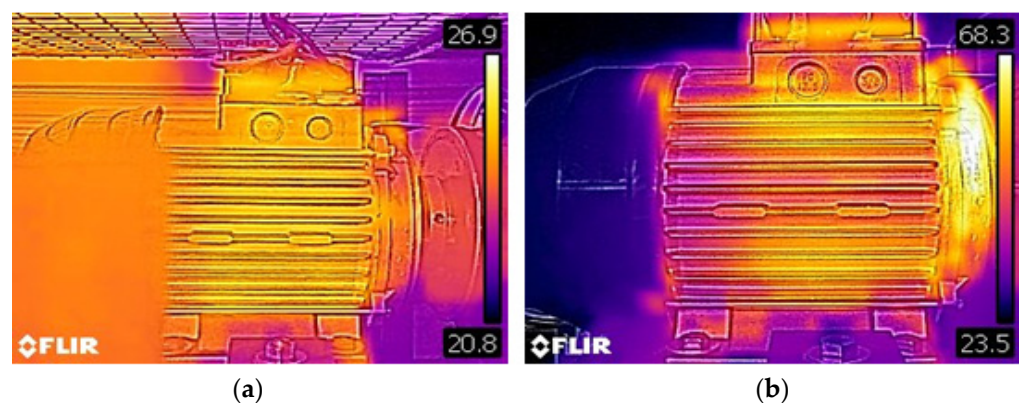


**Figure 3.** Artificial bearing faults: (a) outer bearing fault, (b) inner bearing fault, and (c) ball bearing fault.



**Figure 4.** Artificial rotor faults: (a) one broken rotor bar fault, (b) five broken rotor bar faults, and (c) eight broken rotor bar faults.

An FLIR thermal camera was used to capture thermal images, and it was positioned appropriately on the test rig a certain distance away from the induction motor center, as shown on the test rig. Several camera locations were tried before the final one was selected, considering the quality of each one factored. Thermal images were captured at two speeds after the engine had run for 15 min, and then, the images were stored in JPEG format with a pixel size of  $320 \times 240$ , as displayed in Figure 5.



**Figure 5.** Thermal images: (a) healthy motor run at a speed of 1480 rpm and (b) the motor has an inner bearing fault and run at a speed of 1380 rpm.

#### 4.2. Deep Convolutional Neural Network Architecture Based on VGG-19

When a convolutional neural network is built from scratch, there will be some pros and cons, considering the large amount of data. However, the use of pre-trained models can achieve promising results due to the limitation of the dataset.

Transfer learning models can help us use existing machine learning algorithms. Many techniques can be used to perform transfer learning, for example, by reprocessing the model for feature extraction, which means that only the fully connected classifier is trained.

According to recent reports, the VGG-19 CNN architecture achieves great accuracy when it is processed on the weight of ImageNet. To train the VGG-19 model, it uses the ImageNet dataset of 1.2 million general object images from 1000 different object categories [31]. Moreover, this network contains 19 layers consisting of fully connected, max-pooling, and convolutional layers. The trained convolution base is used with a densely connected classifier. The standard version of VGG-19 is displayed in Figure 6. Using convolutional layers can apply a convolution operation across an image (feature map) and perform the operation at each location, transferring the result to the next layer in the process [32]. Convolutional filters are trainable feature extractors with a  $3 \times 3$  size, and each convolutional layer is followed by a rectified linear unit (ReLU) activation function and a max-pooling procedure. A ReLU is now the most widely used nonlinear activation function, and it can be defined as given in the following equation:

$$f(x) = \max(0, x), \tag{1}$$

where  $x$  is the neuron input.

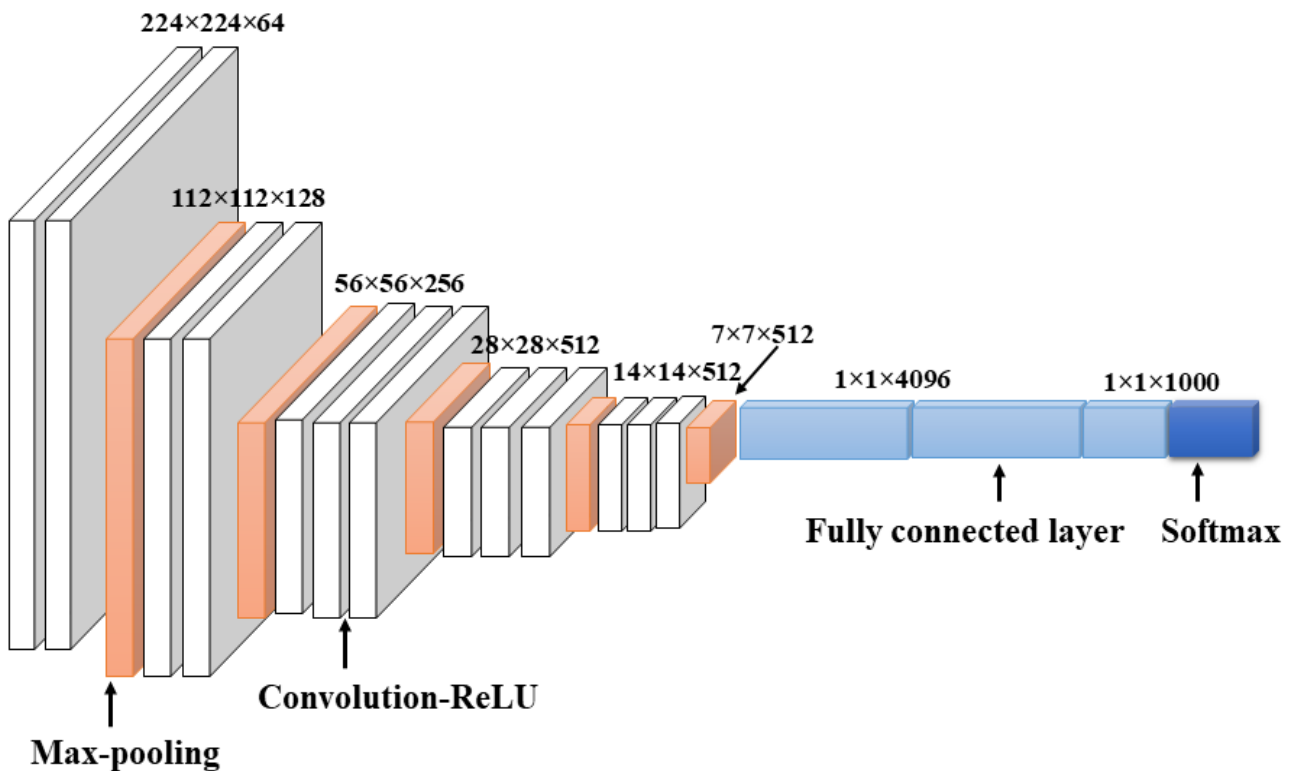


Figure 6. The pre-trained VGG-19 structure.

After down-sampling, the max-pooling layer is applied to the model with a filter size of  $2 \times 2$ . Each neuron in the densely connected layer receives input from all the neurons in the previous layer. The activation function of this densely connected layer must be specified depending on the class type [33].

#### 4.3. Data Pre-Processing and Augmentation

A dataset of 10 classes, each with 500 images, was created based on various motor conditions. The original image was cropped to fit within the bounding box, and the resulting images were resized to  $224 \times 224$ , which is the same input size as required by the classification network. Next, the images were pre-processed on the other layer using a class imbalance technique and a data augmentation technique. By controlling the image magnification, horizontal flip, rotation, translation, and orientation, the overall outcome was influenced with further model improvement.

#### 4.4. Model Evaluation

This work is presented to create a unique application that uses thermal images and a pre-trained CNN-based model as a feature extractor that was developed using Python software installed on a 2GHz GPU PC. The ImageNet dataset's weight was used to train this model. Next, the energy of this data was transferred to the classification part for model prediction. The VGG-19 algorithm was trained using ReLU activation and dropout. Categorical cross-entropy (CE) and SoftMax function (s) were applied as it was a multi-class classification task. The error rate between the original and predicted values was simply achieved [34].

The categorical cross-entropy (CE) and SoftMax function (s) are calculated using the following formulas:

$$CE = - \sum_i^C t_i \log(f(s)_i), \quad (2)$$

where  $t_i$  is the ground truth and  $f(s)_i$  is the standard SoftMax.

$$f(s)_i = \frac{e^{s_i}}{\sum_i^C e^{s_j}}, \quad (3)$$

where  $s_i$  presents the given the class and  $s_j$  is the scores derived from the net for each class.

The Adam optimizer, which is an extension of the stochastic gradient descent, was chosen to implement the presented work. This optimizer can update the weights of the neurons by using backpropagation techniques, where the derivative of the error is calculated with respect to each weight. The key to using this optimizer is to achieve an optimum weight with the maximum accuracy and minimum loss. Some evaluation matrices have been used to assess the proposed application, as given in the following equations:

$$\text{Specificity} = \frac{TN}{FP + TN}, \quad (4)$$

$$\text{Overall accuracy} = \frac{TP + TN}{TP + FP + TN + FN}, \quad (5)$$

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (6)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN}, \quad (7)$$

$$F1\_score = 2 \times \frac{\text{Precision} \times \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}}, \quad (8)$$

where TP is the true-positive prediction, FP is the false-positive prediction, TN presents the true-negative prediction, and FN presents the false-negative prediction.

## 5. Results

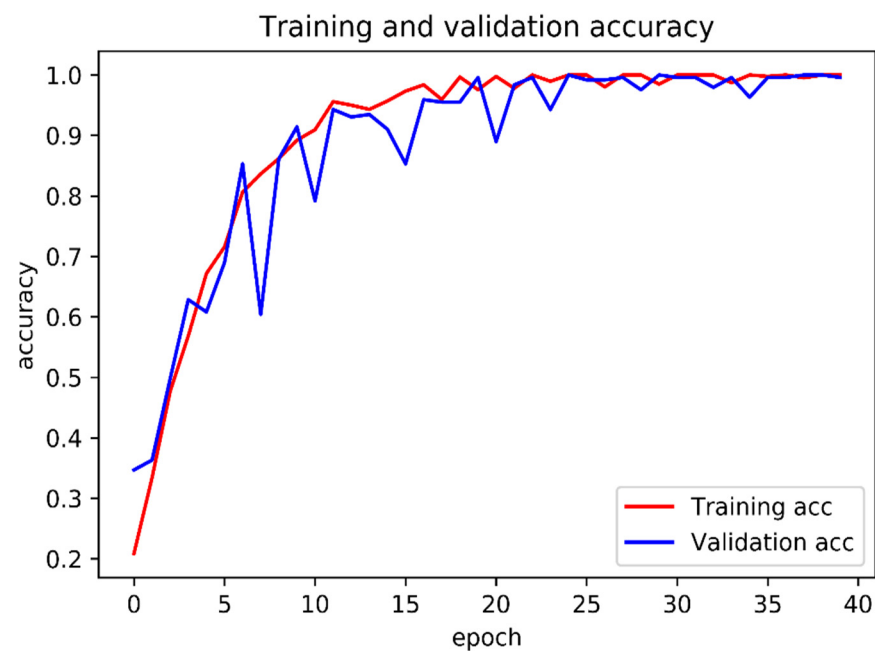
The proposed model of fault diagnosis was investigated using induction motor thermal images. Thermal images with different motor conditions were input into the VGG-12 pre-trained model to process the energy of the extracted features for predicting the correct class. The classification result based on the use of this model is provided in Table 2, where



the model was trained with a categorical cross-entropy function and the Adam optimizer on a SoftMax classifier. From the achieved results, it can be concluded that the proposed pre-trained network VGG-19 with trained ImageNet accomplished satisfactory application of diagnosing induction motor faults. The experimental results have attained an average accuracy of 99.8%, with a training loss equal to 0.0144. The other evaluation metrics presented the same outcome manner. The validation accuracy and loss based on the same epoch number are displayed in Figures 7 and 8, respectively.

**Table 2.** Model classification results.

Train the Model with VGG-19			
	Score	Batch Size	Epochs
Specificity (%)	99.9	64	40
Accuracy (%)	<b>99.8</b>	<b>64</b>	<b>40</b>
Precision (%)	98	64	40
Sensitivity (%)	97.0	64	40
F1-score (%)	94.2	64	40



**Figure 7.** VGG-19-based model accuracy curve.

Compared to current methods, the suggested method is more accurate than the published deep learning method (CNN). A new pre-trained model was proposed in [35], which uses different 48 bearing faults. This model had an average accuracy of 93%. In [36], a fault diagnosis model based on transfer learning-based knowledge was proposed by Kumar and achieved an accuracy of 99.40%. This model was further trained by the K-nearest neighbor classifier, support vector machine, and random forest and achieved an accuracy of 78.60%, 90%, and 89.40%, respectively. A novel and accurate deep learning framework for fault diagnosis was presented by Shao in [37]. This model was investigated on three different mechanical datasets, including the gearbox and bearing datasets. The model achieved an accuracy from 94.8% to 99.64%. In [38], a pre-trained VGG-19 model was proposed by Wen for fault diagnosis. This model converted famous time domain signals from CWRU to images and processed them using VGG-19 and the SoftMax classifier. A prediction accuracy of 99.175% was achieved by the presented model. Another fault diagnosis model was suggested by Grover for rolling element bearings in [39]. This model used four pretrained models, namely Alexnet, VGG-19, Google Net, and ResNet-50.

VGG-19 achieved a validation accuracy of 99.7% by applying the Adam optimizer and 94.1% by applying the SGD optimizer with few epoch numbers. In [40], a novel transfer learning model for fault detection was proposed. This model was tested on current signals and achieved an accuracy of 99.4%. In this work, the proposed model achieved a higher classification performance compared to the aforementioned models; it achieved an accuracy of 99.8%.

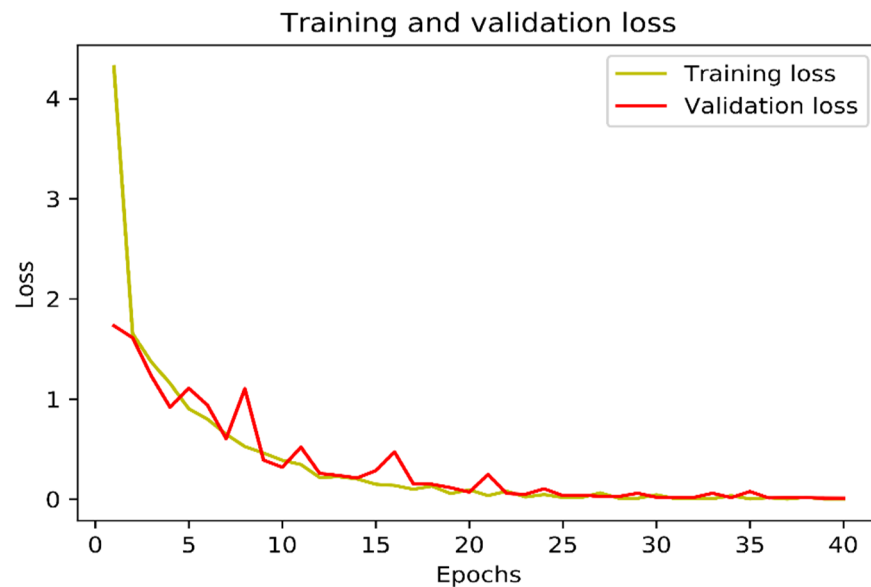


Figure 8. VGG-19-based model loss curve.

## 6. Discussion

Although many automated machine learning techniques have been introduced for diagnosing induction motor faults, there is still a lack of solutions to predict motor conditions. Compared to other methods, it is successfully used in the content of this work, indicating it is superior since it requires less human involvement, while producing accurate predictions. Moreover, training a model via transfer learning yields better results. Transfer learning has also compensated the limitation of traditional techniques and deep CNNs.

## 7. Conclusions

In this work, a new application for fault diagnosis based on a deep learning network was proposed. This application combined thermal images, a deep transfer learning network, and a densely connected classifier. This application applied VGG-19 to construct the first model features from the images directly to achieve a fast and robust classification model. The highest classification accuracy of 99.80% was achieved by combining the suggested pre-trained network with the densely connected classifier.

Concisely, the classification method's accuracy is reasonable, suggesting this model could be used to identify induction motor defects using thermal imaging data. The VGG-19 network's resilience will be evaluated in future studies using more comprehensive data, not just using a transfer learning model, but also fine-tuning the network's properties. Other deeper networks based on detection time will be built considering the prediction challenge.

**Author Contributions:** A.I., F.A. and M.P. contributed equally to the conception of the idea; collecting data, A.I., F.A. and M.P.; analysing experimental results, A.I. and F.A.; investigating, A.I.; writing—original draft preparation, A.I.; writing—review and editing, A.I. and M.P.; supervision, A.I. All authors have read and agreed to the published version of the manuscript.

**Funding:** This paper is part of the Ph.D. research of the corresponding author, A. Ibrahim, who is sponsored by the Ministry of Higher Education in Libya.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** All authors would like to thank Cardiff University/School of Engineering for achieving this successful work. The corresponding author A. Ibrahim acknowledges the Ministry of Higher Education in Libya for the Scholarship.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Ayas, S.; Ayas, M.S. A novel bearing fault diagnosis method using deep residual learning network. *Multimedia Tools Appl.* **2022**, *81*, 22407–22423. [\[CrossRef\]](#)
2. Guo, L.; Li, N.; Jia, F.; Lei, Y.; Lin, J. A recurrent neural network based health indicator for remaining useful life prediction of bearings. *Neurocomputing* **2017**, *240*, 98–109. [\[CrossRef\]](#)
3. Pons-Llinares, J.; Antonino-Daviu, J.A.; Riera-Guasp, M.; Bin Lee, S.; Kang, T.-J.; Yang, C. Advanced Induction Motor Rotor Fault Diagnosis Via Continuous and Discrete Time–Frequency Tools. *IEEE Trans. Ind. Electron.* **2015**, *62*, 1791–1802. [\[CrossRef\]](#)
4. Park, Y.M.; Kim, G.-W.; Sohn, J.-M. A logic based expert system (LBES) for fault diagnosis of power system. *IEEE Trans. Power Syst.* **1997**, *12*, 363–369. [\[CrossRef\]](#)
5. Li, B.; Chow, M.-Y.; Tipsuwan, Y.; Hung, J. Neural-network-based motor rolling bearing fault diagnosis. *IEEE Trans. Ind. Electron.* **2000**, *47*, 1060–1069. [\[CrossRef\]](#)
6. Baraldi, P.; Cannarile, F.; Di Maio, F.; Zio, E. Hierarchical k-nearest neighbours classification and binary differential evolution for fault diagnostics of automotive bearings operating under variable conditions. *Eng. Appl. Artif. Intell.* **2016**, *56*, 1–13. [\[CrossRef\]](#)
7. Kang, M.; Kim, J.; Kim, J.-M. Reliable fault diagnosis for incipient low-speed bearings using fault feature analysis based on a binary bat algorithm. *Inf. Sci.* **2015**, *294*, 423–438. [\[CrossRef\]](#)
8. Wang, Z.; Zhang, Q.; Xiong, J.; Xiao, M.; Sun, G.; He, J. Fault Diagnosis of a Rolling Bearing Using Wavelet Packet Denoising and Random Forests. *IEEE Sens. J.* **2017**, *17*, 5581–5588. [\[CrossRef\]](#)
9. Mehrabian, A.R.; Lucas, C. A novel numerical optimization algorithm inspired from weed colonization. *Ecol. Inform.* **2006**, *1*, 355–366. [\[CrossRef\]](#)
10. Ibrahim, A.; Anayi, F.; Packianather, M.; Alomari, O.A. New Hybrid Invasive Weed Optimization and Machine Learning Approach for Fault Detection. *Energies* **2022**, *15*, 1488. [\[CrossRef\]](#)
11. Qian, C.; Zhu, J.; Shen, Y.; Jiang, Q.; Zhang, Q. Deep Transfer Learning in Mechanical Intelligent Fault Diagnosis: Application and Challenge. *Neural Process. Lett.* **2022**, *54*, 2509–2531. [\[CrossRef\]](#)
12. Han, T.; Jiang, D.; Wang, N. The Fault Feature Extraction of Rolling Bearing Based on EMD and Difference Spectrum of Singular Value. *Shock Vib.* **2016**, *2016*, 5957179. [\[CrossRef\]](#)
13. Tamilselvan, P.; Wang, P. Failure diagnosis using deep belief learning based health state classification. *Reliab. Eng. Syst. Saf.* **2013**, *115*, 124–135. [\[CrossRef\]](#)
14. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [\[CrossRef\]](#)
15. Chaturvedi, I.; Ong, Y.-S.; Arumugam, R.V. Deep transfer learning for classification of time-delayed Gaussian networks. *Signal Process.* **2015**, *110*, 250–262. [\[CrossRef\]](#)
16. Gonzalez, T.F. ImageNet Classification with Deep Convolutional Neural Networks. In *Handbook of Approximation Algorithms and Metaheuristics*; Chapman and Hall/CRC: New York, NY, USA, 2007; pp. 1–1432. [\[CrossRef\]](#)
17. Socher, R.; Bengio, Y.; Manning, C.D. Deep Learning for NLP (without Magic). In Proceedings of the NAACL-HLT 2013—The 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Atlanta, GA, USA, 9–15 June 2013; pp. 1–3.
18. Wu, L.; Baldi, P. Learning to play Go using recursive neural networks. *Neural Netw.* **2008**, *21*, 1392–1400. [\[CrossRef\]](#)
19. Gan, M.; Wang, C.; Zhu, 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**, *72–73*, 92–104. [\[CrossRef\]](#)
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. Cho, H.C.; Knowles, J.; Fadali, M.S.; Lee, K.S. Fault Detection and Isolation of Induction Motors Using Recurrent Neural Networks and Dynamic Bayesian Modeling. *IEEE Trans. Control Syst. Technol.* **2010**, *18*, 430–437. [\[CrossRef\]](#)
22. Shao, H.; Jiang, H.; Zhang, X.; Niu, M. Rolling bearing fault diagnosis using an optimization deep belief network. *Meas. Sci. Technol.* **2015**, *26*, 115002. [\[CrossRef\]](#)
23. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. In Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, 7–9 May 2015; pp. 1–14.
24. Panigrahi, S.; Nanda, A.; Swarnkar, T. A Survey on Transfer Learning. *Smart Innov. Syst. Technol.* **2021**, *194*, 781–789. [\[CrossRef\]](#)
25. Zuo, L.; Jing, M.; Li, J.; Zhu, L.; Lu, K.; Yang, Y. Challenging tough samples in unsupervised domain adaptation. *Pattern Recognit.* **2021**, *110*, 107540. [\[CrossRef\]](#)

26. Chen, W.; Qiu, Y.; Feng, Y.; Li, Y.; Kusiak, A. Diagnosis of wind turbine faults with transfer learning algorithms. *Renew. Energy* **2021**, *163*, 2053–2067. [[CrossRef](#)]
27. Weiss, K.; Khoshgoftaar, T.M.; Wang, D.D. A survey of transfer learning. *J. Big Data* **2016**, *3*, 9. [[CrossRef](#)]
28. Yang, Q. Highly Accurate Machine Fault Diagnosis Using Deep Transfer Learning. *IEEE Trans. Ind. Inform.* **2019**, *15*, 2446–2455.
29. Wang, P.; Ananya; Yan, R.; Gao, R.X. Virtualization and deep recognition for system fault classification. *J. Manuf. Syst.* **2017**, *44*, 310–316. [[CrossRef](#)]
30. Zhou, J.; Yang, X.; Zhang, L.; Shao, S.; Bian, G. Multisignal VGG19 Network with Transposed Convolution for Rotating Machinery Fault Diagnosis Based on Deep Transfer Learning. *Shock Vib.* **2020**, *2020*, 8863388. [[CrossRef](#)]
31. Zhang, X.; Zou, J.; He, K.; Sun, J. Accelerating Very Deep Convolutional Networks for Classification and Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *38*, 1943–1955. [[CrossRef](#)]
32. Handa, A.; Garg, P.; Khare, V. Masked Neural Style Transfer using Convolutional Neural Networks. In Proceedings of the 2018 International Conference on Recent Innovations in Electrical, Electronics & Communication Engineering (ICRIEECE), Bhubaneswar, India, 27–28 July 2018; pp. 2099–2104. [[CrossRef](#)]
33. Mocsari, E.; Stone, S.S. Densely Connected Convolutional Networks. *Am. J. Vet. Res.* **1978**, *39*, 1442–1446.
34. Jamin, A.; Humeau-Heurtier, A. (Multiscale) Cross-Entropy Methods: A Review. *Entropy* **2020**, *22*, 45. [[CrossRef](#)]
35. Wang, P.; Gao, R.X. Transfer learning for enhanced machine fault diagnosis in manufacturing. *CIRP Ann.* **2020**, *69*, 413–416. [[CrossRef](#)]
36. Kumar, P.; Hati, A.S. Transfer learning-based deep CNN model for multiple faults detection in SCIM. *Neural Comput. Appl.* **2021**, *33*, 15851–15862. [[CrossRef](#)]
37. Shen, F.; Chen, C.; Yan, R.; Gao, R.X. Bearing fault diagnosis based on SVD feature extraction and transfer learning classification. In Proceedings of the 2015 Prognostics and System Health Management Conference (PHM), Beijing, China, 21–23 October 2015. [[CrossRef](#)]
38. Wen, L.; Li, X.; Gao, L. A New Transfer Learning Based on VGG-19 Network for Fault Diagnosis. In Proceedings of the 2019 IEEE 23rd International Conference on Computer Supported Cooperative Work in Design (CSCWD), Porto, Portugal, 6–8 May 2019; pp. 205–209. [[CrossRef](#)]
39. Grover, C.; Turk, N. A novel fault diagnostic system for rolling element bearings using deep transfer learning on bispectrum contour maps. *Eng. Sci. Technol. Int. J.* **2022**, *31*, 101049. [[CrossRef](#)]
40. 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, 9–12 June 2020. [[CrossRef](#)]