*Proceedings of the 10th World Congress on Electrical Engineering and Computer Systems and Sciences (EECSS'24)*
*Barcelona, Spain - August 19 - 21, 2024*
*Paper No. EEE 113 (The number assigned by the OpenConf System)*
*DOI: 10.11159/eee24.113*

# Classifying Induction Motor Faults Using Spectrogram Images with Deep Transfer Learning

**Merve Ertargin[1], Ozal Yildirim[2], Ahmet Orhan[3]**
[1]University of Munzur
Department of Electrical and Electronics Engineering, Tunceli, Türkiye
merveboydak@munzur.edu.tr;
[2] University of Firat
Department of Artificial Intelligence and Data Engineering, Elazig, Türkiye
ozalyildirim@firat.edu.tr;
[3]University of Firat
Department of Electrical and Electronics Engineering, Elazig, Türkiye
aorhan@firat.edu.tr

***Abstract*** **-** For industrial systems to be reliable and efficient, motor faults detection is important. Conventional methods of fault detection are typically expensive and time-consuming. The ability to detection motor faults has advanced significantly in recent years thanks to the application of deep learning techniques. Deep learning algorithms have the ability to automatically extract complicated features from big data sets, which can speed up and improve the accuracy of motor faults detection. Since collecting motor acoustic data is cost-effective and easy, it is advantageous to use it in fault detection. In this study, motor acoustic signals were converted into spectrograms and faults in induction motors were detected using a transfer learning approach with pre-trained models. Totally 8-class fault detection was performed with an accuracy rate of 91.52% using VGG16 and 92.11% using VGG19.

***Keywords*****:** electric motor fault detection, acoustic data, VGG16, VGG19.

## 1. Introduction

Computer-aided software in motor faults detection most commonly utilizes motor vibration signals as input data [1-3]. Diagnosing faults with motor vibrations is efficient and does not require significant time. However, it necessitates the placement of accelerometer sensors on the motor to gather vibration signals, a process that demands expertise. On the other hand, diagnostics using motor acoustic data does not mandate a direct connection to the motor, making the collection of acoustic data both economical and effortless [4]. Nevertheless, the scarcity of publicly available acoustic data has restricted studies in this field.

The integration of machine learning and deep learning models in motor faults detection offers advantages such as the optimization of maintenance processes in industrial facilities and the early detection of malfunctions. However, these models necessitate substantial amounts of data. While an increase in data enhances the accuracy of fault detection models [5], it also prolongs the training process. Therefore, the utilization of pre-trained neural networks in fault detection aids in achieving optimal performance with reduced data requirements and training time. In the study [6], irreversible-demagnetization fault and bearing faults in permanent magnet synchronous motor were classified with an accuracy rate of 96.65% using VGG16. In the study [7], bearing faults were classified with AlexNet, VGG19, GoogLeNet and ResNet50. The test accuracy rate in fault classification using VGG-19 is 85.46% for the SGD optimizer, 98.22% for the Adam optimizer and 97.92% for the Adamax optimizer. In the study [8], the study focused on classifying bearing faults using ResNet50, VGG16, and VGG19. The highest accuracy rate of 99.92% was achieved with VGG-19.

In this study, we classified induction motor faults using pre-trained VGG16 and VGG19 models. This approach was adopted to address the challenge of limited publicly available data, which often hinders the use of engine sound signals for engine fault detection. Time domain sound data belonging to healthy and 7 fault classes in the UOEMD-VAFCVS dataset [9] were converted to spectrograms using Short Time Fourier Transform (STFT). Spectrograms were resized to match the input size expected by the VGG16 and VGG19 models, which is 224×224 pixels. Fault classification using VGG16 achieved an accuracy rate of 91.52%, while VGG19 achieved an accuracy rate of 92.11%.

## 2. Short Time Fourier Transform (STFT)

Short-Time Fourier Transform (STFT) is a spectrum analysis technique used to analyze the frequency components of signals that change over time. It works by dividing the signal into small time segments and applying the Fourier transform to each segment, allowing us to observe how the signal's frequency content changes over time.

STFT is valuable because it provides a way to visualize the time-varying frequency components of a signal in the time-frequency plane. This helps us better understand the signal's characteristics in both the time and frequency domains. STFT is commonly used in fields such as image and audio processing to create spectrograms, which provide a detailed representation of a signal's frequency and time components. Where x(t) is input signal, g(t) is window function, ω is the angular frequency parameter and τ is shifting parameter, STFT of the input signal is given in Equation 1.

When using STFT, the process typically involves applying a windowing function with a specific amount of overlap. This adjustment shifts the starting point of each window by a certain amount, typically a fraction of the window size, to create the next window. Overlap plays a crucial role in balancing the time and frequency resolutions of STFT. A higher overlap results in increased time resolution because the windows are applied more frequently, capturing more detail in the time domain.

$$\mathcal{F}(\tau,\omega) = \int_{-\infty}^{+\infty} x(t)\, \mathrm{g}(\mathrm{t}-\tau) e^{-j\omega t} dt \tag{1}$$

## 3. Transfer Learning

Transfer learning is the transfer of information learned in one task to another task. Classifying with pre-trained models is achieved with transfer learning approach. A pre-trained model is an artificial neural network model that has been pre-trained on a large dataset, and tuned to be able to solve a general task or tasks. Pre-trained models are often used to solve complex problems such as visual perception, natural language processing or audio processing. These models are typically trained for a general task on a large dataset and then can be tuned on a smaller dataset to solve a specific task. This can often be useful to achieve good performance when working with smaller datasets.

### 3.1 VGG16

VGG16 is a deep convolutional neural network (CNN) model developed by researchers from the University of Oxford [10]. VGG16 is a model trained on the ImageNet dataset and achieves high accuracy in the image classification task. The model is particularly notable for its depth (number of layers) and simplicity. VGG16 contains 13 convolutional layers, 5 maxpooling layers and 3 fully connected layers (FC). Convolutional layers are used to learn features in images, while fully connected layers are used to perform classification. Each convolution layer has a 3×3 filter size with a stride of 1 and each maxpooling layer has a filter size of 2×2 and a stride of 2. Figure 1 shows the architecture of VGG16.

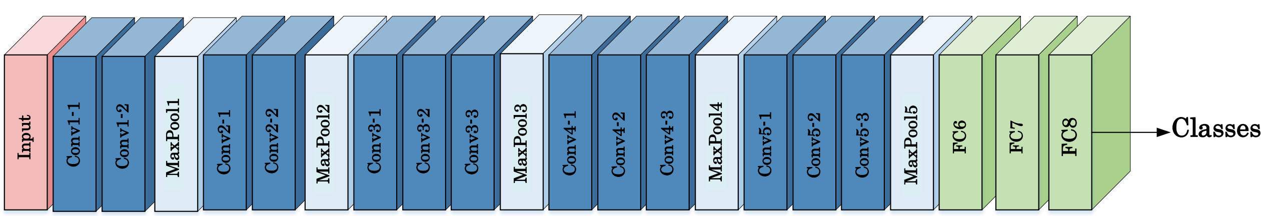


Fig. 1: VGG16 architecture.

### 3.2 VGG19

VGG19 is an expansion of the VGG16 model and has more layers. VGG19 was trained on the ImageNet dataset and achieved high accuracy in the image classification task. VGG19 contains 16 convolutional layers, 5 maxpooling layers and 3 fully connected layers. Each convolution layer has a 3×3 filter size with a stride of 1 and each maxpooling layer has a filter size of 2×2 and a stride of 2. Figure 2 shows the architecture of VGG19.

The basic idea of VGG19 is that as the depth (number of layers) increases, the model can learn more complex features. However, adding more layers can make the training process of the model more difficult and require more computing power. VGG19 achieves this balance, effectively learning complex features while making training manageable. VGG19 has been an important milestone in the development of deep learning models, especially in the field of image processing. The model has shown that the complexity of deep neural network architectures can increase and deeper models can perform better.

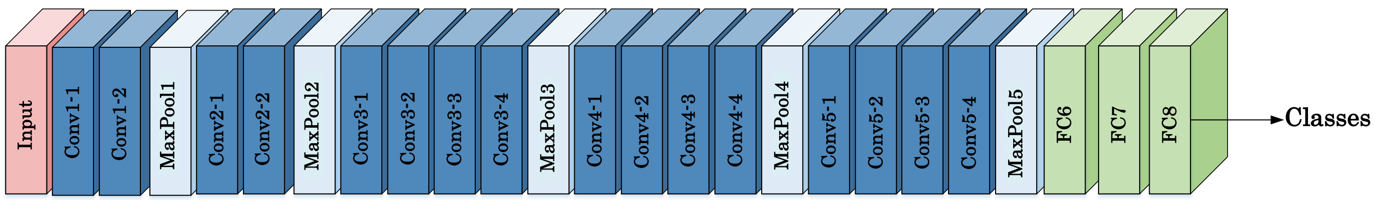



Fig. 2: VGG19 architecture.

## 4. Materials and Methods

In this study, UOEMD-VAFCVS public electricity motor fault dataset [9] is used. Figure 3 shows UOEMD-VAFCVS dataset test rig setup [11]. The maximum output power of the motor at 60 Hz frequency is 3.00 Hp and the maximum speed is 3600 rpm. The microphone (PCB, type 130F20) is positioned within 2 cm of the left bearing housing and is supported by a separate stand designed to isolate vibrations.

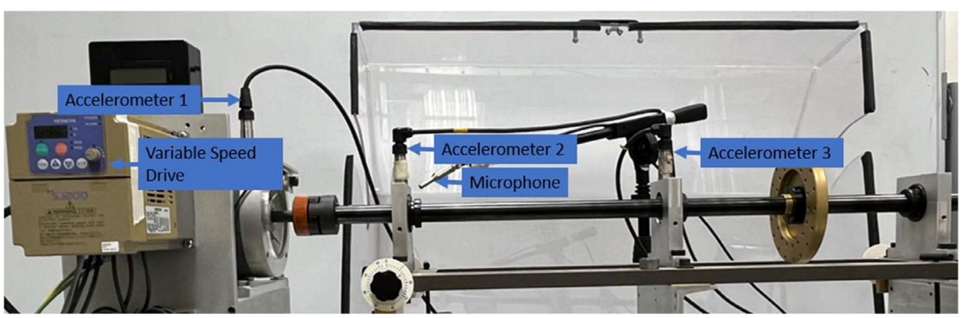



Fig. 3: UOEMD-VAFCVS dataset test rig setup [11]

The induction motor dataset's sampling frequency is 42 kHz. Time domain acoustic signals in this dataset are converted into spectrograms by applying STFT. In this study, Hanning windowing method is applied. Acoustic signals are divided into segments of 512 data points each. Overlap value is set to 460. Figure 4 depicts spectrogram images corresponding to each class. Spectrogram images are randomly divided, with 80% allocated to the train folder (20480 spectrograms), 10% to the validation folder (2560 spectrograms) and 10% to the test folder (2560 spectrograms). The spectrograms in the train and validation folders are utilized to train the final classification layers of the VGG16 and VGG19 pre-trained models, using the dataset at hand. The performance of the re-trained and updated VGG16 and VGG19 models is evaluated using the test dataset. Figure 5 summarizes the steps described.

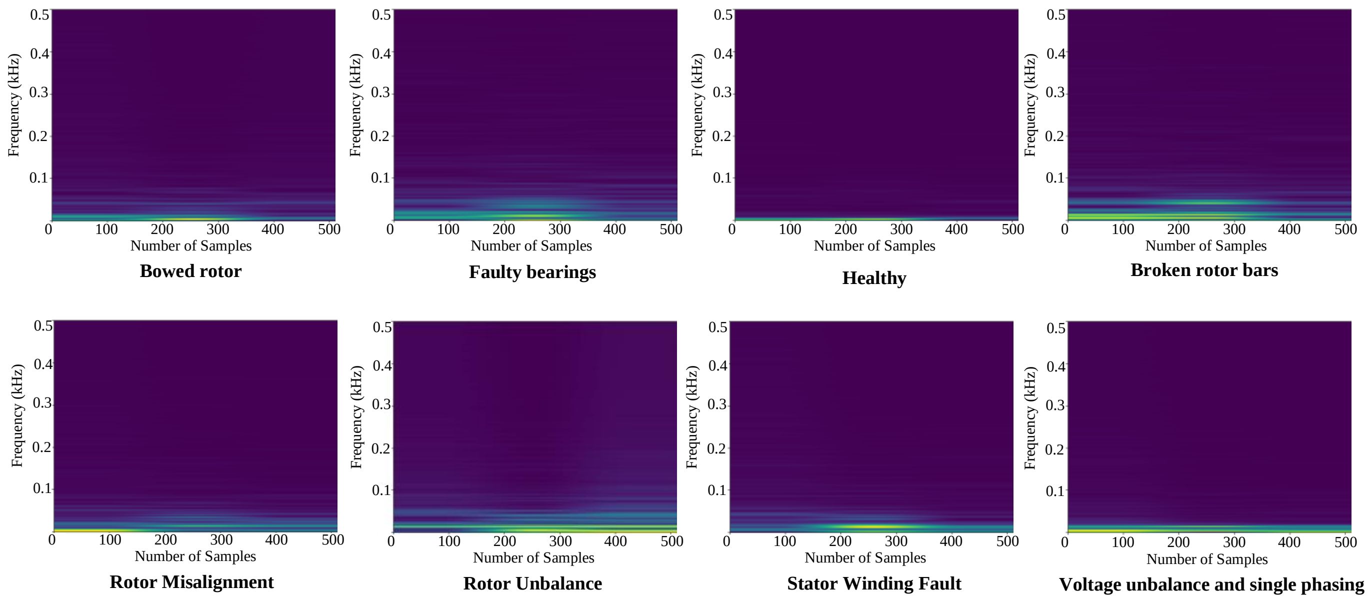



Fig. 4: Spectrogram images for each class

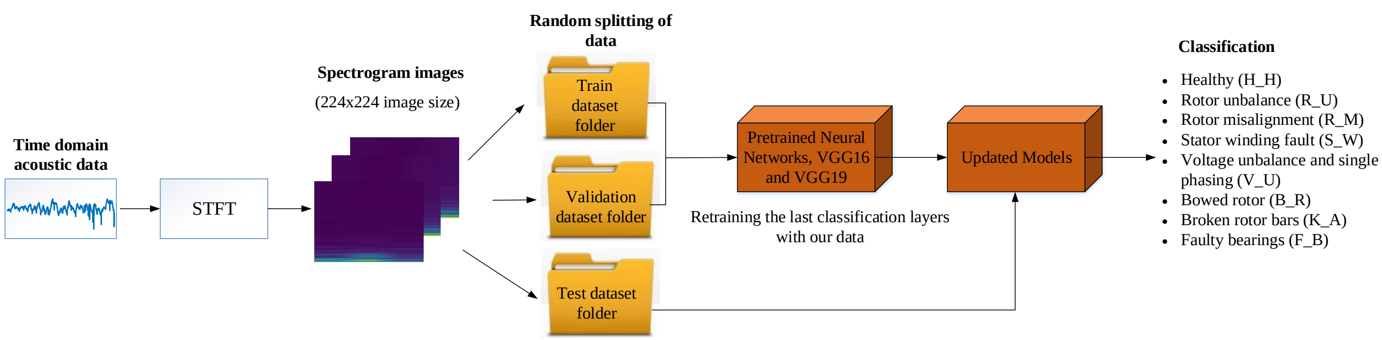


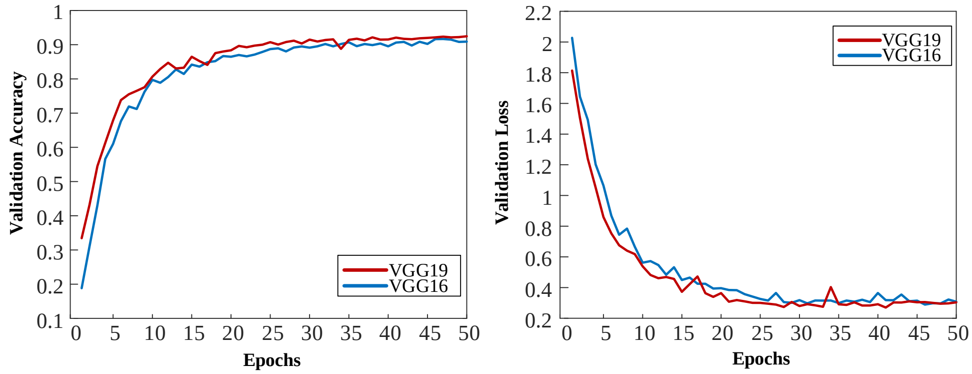

Fig. 5: Proposed method

## 5. Experimental Results

The study is carried out using Google Colab Notebook. The model used the "Categorical_crossentropy" loss function, Adagrad optimizer, and a batch size of 32, with results obtained after training for 50 epochs.

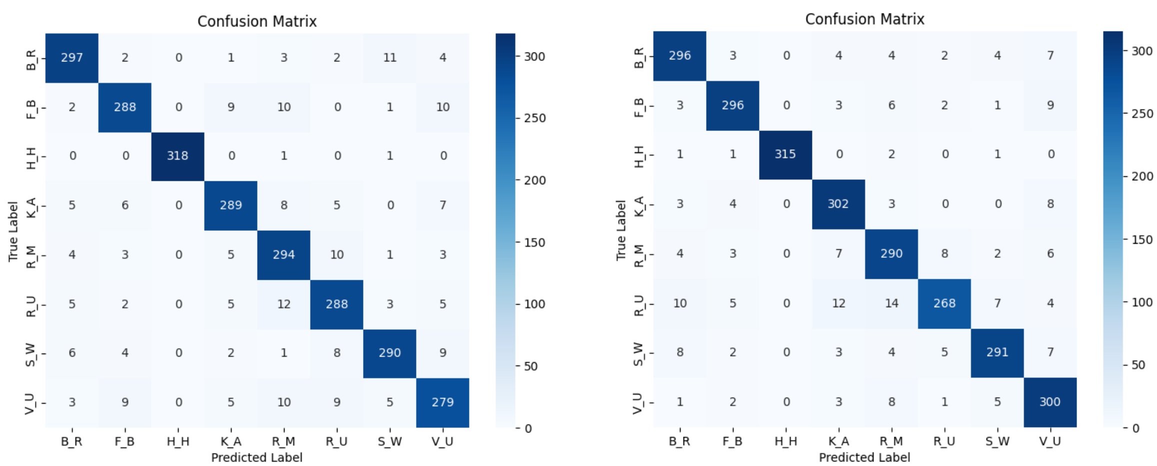
In Figure 6, validation accuracy and loss graphs for the VGG16 and VGG19 models are presented. The highest accuracy rate achieved by the VGG16 model over 50 epochs is 91.6796%, and for VGG19, it is 92.4609%.

In Figure 7, confusion matrices are given. Both models showed the highest accuracy when classifying data belonging to the "Healthy" class. VGG16 showed the lowest performance when classifying data belonging to the "Voltage unbalance and single phasing" class. It misclassified 41 of the 320 data belonging to this class. VGG19 showed the lowest performance when classifying data belonging to the "Rotor Unbalance" class. It misclassified 52 of the 320 data belonging to this class.




(a) (b)

Fig. 6: (a) Validation accuracy and (b) validation loss graphs for VGG16 and VGG19




(a) (b)

Fig. 7: Confusion matrix (a) VGG16 (b) VGG19

Spectrogram data were classified using a custom 2D-CNN model to investigate the impact of using pre-trained models on the accuracy of engine fault detection using acoustic data. The custom 2D-CNN model consists of 4 convolutional layers, 4 max-pooling layers, and 3 dense layers. The performance of models trained on a large dataset is reflected in the accuracy rate for detecting motor faults. Table 2 presents the accuracy rates of the VGG16, VGG19, and custom 2D-CNN models. To evaluate the performance of the VGG16 and VGG19 model in detail, Accuracy, Precision, Sensitivity, Specificity and F1-Score values are given in Table 2 for each class and overall.

Table 1: Accuracy rates comparision

| Model | Accuracy (%) |
|---|---|
| VGG16 | 91.52 |
| VGG19 | 92.11 |
| Custom 2D-CNN | 83.36 |

Table 2: Performace Values of VGG16 and VGG19

| Pre-trained Models | Classes | Accuracy (%) | Precision (%) | Sensitivity (%) | Specificity (%) | F1-Score |
|---|---|---|---|---|---|---|
| **VGG16** | B_R | 98.1250 | 92.2360 | 92.8125 | 98.8839 | 92.5234 |
| | F_B | 97.7344 | 91.7197 | 90.0000 | 98.8393 | 90.8517 |
| | H_H | 99.9219 | 100.0000 | 99.3750 | 100.0000 | 99.6865 |
| | K_A | 97.7344 | 91.4557 | 90.3125 | 98.7946 | 90.8805 |
| | R_M | 97.2266 | 86.7257 | 91.8750 | 97.9911 | 89.2261 |
| | R_U | 97.4219 | 89.4410 | 90.0000 | 98.4821 | 89.7196 |
| | S_W | 97.9688 | 92.9487 | 90.6250 | 99.0179 | 91.7722 |
| | V_U | 96.9141 | 88.0126 | 87.1875 | 98.3036 | 87.5981 |
| | ***Overall*** | **91.52** | **91.5674** | **91.5234** | **98.7891** | **91.5323** |
| **VGG19** | B_R | 97.8906 | 90.7975 | 92.5000 | 98.6607 | 91.6409 |
| | F_B | 98.2813 | 93.6709 | 92.5000 | 99.1071 | 93.0818 |
| | H_H | 99.8047 | 100.0000 | 98.4375 | 100.0000 | 99.2126 |
| | K_A | 98.0469 | 90.4192 | 94.3750 | 98.5714 | 92.3547 |
| | R_M | 97.2266 | 87.6133 | 90.6250 | 98.1696 | 89.0937 |
| | R_U | 97.2656 | 93.7063 | 83.7500 | 99.1964 | 88.4488 |
| | S_W | 98.0859 | 93.5691 | 90.9375 | 99.1071 | 92.2345 |
| | V_U | 97.6172 | 87.9765 | 93.7500 | 98.1696 | 90.7716 |
| | ***Overall*** | **92.11** | **92.2191** | **92.1094** | **98.8728** | **92.1048** |

## 6. Conclusion

In this study, motor acoustic signals were converted to spectrograms using Short-Time Fourier Transform (STFT) to classify faults in induction motors. The spectrograms were classified with the VGG16 and VGG19 models, which were pre-trained on images, into 8 classes (Bowed rotor, Faulty bearings, Healthy, Broken rotor bars, Rotor Misalignment, Rotor Unbalance, Stator Winding Fault, Voltage unbalance and single phasing). The accuracy rates of VGG16 and VGG19 are 91.52% and 92.11% respectively. This study focuses on the need for more extensive use of motor acoustic signals in fault detection. The challenge of limited motor acoustic data can be overcome with pre-trained models.

## References


[1] Y. Zhou, Q. Shang and C. Guan, "Three-Phase Asynchronous Motor Fault Diagnosis Using Attention Mechanism and Hybrid CNN-MLP By Multi-Sensor Information," in *IEEE Access*, vol. 11, pp. 98402-98414, 2023, doi: 10.1109/ACCESS.2023.3307770.

[2] M. Gana, H. Achour and M. Laghrouche, “Enhanced motor fault detection system based on a dual-signature image classification method using CNN,” *Engineering Research Express*, 5(1), 015009, 2023.

[3] J. G. Jang, C. M. Noh, S. S. Kim, S. C. Shin, S. S. Lee and J. C. Lee, “Vibration data feature extraction and deep learning-based preprocessing method for highly accurate motor fault diagnosis,” *Journal of Computational Design and Engineering*, 10(1), 204-220, 2023.

[4] A. Glowacz, M. Sulowicz, J. Kozik, K. Piech, W. Glowacz, Z. Li, ... and H. Lui, “Fault diagnosis of electrical faults of three-phasei nduction motors using acoustic analysis,” *Bulletin of the Polish Academy of Sciences. Technical Sciences*, 72(1), 2024.
[5] D. Neupane, and J. Seok, “Bearing fault detection and diagnosis using case western reserve university dataset with deep learning approaches: A review,” *Ieee Access*, 8, 93155-93178, 2020.
[6] Z. Ullah, B. A. Lodhi and J. Hur, “Detection and identification of demagnetization and bearing faults in PMSM using transfer learning-based VGG,” *Energies*, 13(15), 3834, 2020.
[7] C. Grover and N. Turk, “A novel fault diagnostic system for rolling element bearings using deep transfer learning on bispectrum contour maps,” *Engineering Science and Technology, an International Journal*, 31, 101049, 2022.
[8] Y. Yoo and S. Jeong, “Vibration analysis process based on spectrogram using gradient class activation map with selection process of CNN model and feature layer,” *Displays*, 73, 102233, 2022.
[9] M. Sehri, P. Dumond, “University of Ottawa Electric Motor Dataset – Vibration and Acoustic Faults under Constant and Variable Speed Conditions (UOEMD-VAFCVS)”, Mendeley Data, V1, 2023, doi: 10.17632/msxs4vj48g.1.
[10] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” arXiv 1409.1556, 2014.
[11] M. Sehri and P. Dumond, “University of Ottawa Constant and Variable Speed Electric Motor Vibration and Acoustic Fault Signature Dataset,” *Data in Brief*, 110144, 2024.