Journal of Computer Science
https://dergipark.org.tr/en/pub/bbd

ISSN,e-ISSN: 2548-1304
Volume:IDAP-2023, pp:75-82, 2023
https://doi.org/10.53070/bbd.1347156

Research Paper

A New Digital Twin-Based Fault Diagnosis Approach Using Parameter Estimation and Information Entropy

İlhan AYDIN[*1], Emrullah AYDIN[2], Erhan AKIN[1]

[1,2]Bilgisayar Mühendisliği Bölümü, Fırat Üniversitesi, Elazığ, Türkiye

[2]Elektrik - Elektronik Mühendisliği Bölümü, Malatya Turgut Özal Üniversitesi, Malatya, Türkiye

(iaydin@firat.edu.tr, emrullah.aydin@ozal.edu.tr, eakin@firat.edu.tr)

Received:Aug.21,2023 Accepted:Aug.23,2023 Published:Oct.18,2023

Abstract— Induction motors are one of the important motor types used in industry. Although these motors are generally of robust construction, they are subject to failures due to ambient operating conditions. The traditional diagnostic methods are based on measuring signals such as current, vibration, temperature, and speed from an experimental setup for good and faulty motors. But finding an equivalent motor that can compare with the motor used in the industry is always difficult. Therefore, by constructing a digital twin of the real motor, signals belonging to the healthy motor can be obtained, which is equivalent to the motor in the industry. In this study, motor stator faults were tried to be diagnosed using digital twin and motor signals obtained from a real experimental setup. The faulty frequency region is determined in the spectrum by estimating the parameters related to the motor current, and the faults are determined according to the information entropy. The operation of the proposed system has been tested with data from both the digital twin and the real motor, and successful results have been obtained.

Keywords : *Induction motor, information entropy, diagnostics, stator faults, feature extraction, digital twin, classification.*

1. Introduction

Induction motors are one of the most used motor types in the industry. Wide area usage, robust structures, and low maintenance requirements are the main reasons for their use [1]. Detecting faults in these motors at an early stage prevents occurring serious faults. Faults in induction motors are generally related to shaft bearing, rotor, and stator components [2]. Stator faults account for 38% of all faults [3].

Model-based and feature extraction-based methods have been used in the literature for fault diagnosis. Model-based methods are based on the creation of a mathematical model of the motor and the comparison of the real motor with the model. Feature extraction-based methods, on the other hand, are based on the evaluation of the properties obtained for healthy and faulty motors from signals such as current, vibration, flux, torque, and voltage with a classification method. Artificial intelligence-based methods have been developed in the literature for the diagnosis of induction motor faults. Time-frequency image-based approaches using vibration signals have been proposed for the diagnosis of bearing failures [4-6]. In these approaches, the images obtained with the time-frequency image obtained by using the vibration signal and the current signal are trained with convolutional neural network models, and the faults are classified. Discrete wavelet analysis and artificial neural networks are used for the diagnosis of stator faults [7]. After applying the park vector transform to the stator current signals, the db-8 wavelet is applied. Statistical parameters of L1 and L2 norms gave effective results for stator failures. A study in which a 1-dimensional convolutional neural network and long-short-term memory architecture are integrated has been proposed for the detection of stator faults at an early stage [8]. The proposed work combines feature extraction and classification steps and integrates multiple convolutional neural network architectures to identify faults. Especially since stator faults are often confused with voltage unbalance faults, a study that makes this distinction is presented in this study. In addition, the effect of load changes is reduced in the developed method. Hilbert transform and statistical feature extraction were used to detect stator faults at an early stage [9]. Fault detection was made by giving the obtained features to the support vector machines. In order to enable the use of induction motors in electric vehicles and to detect the faults that occur in these motors, different types of faults are simulated by creating a motor model in the ANSYS environment [10]. Different machine learning and deep learning methods are used to detect the generated faults under different load conditions. It has been experimentally verified that the performance rate of the deep learning method is better than other machine learning methods. An approach using piezoelectric sensors is proposed for the detection of short-circuit faults of three-phase induction motors [11]. For feature extraction, wavelet analysis and principal component analysis were used together with the cross-correlation method. A deep learning-based approach has been proposed to detect stator winding failures under variable load conditions [12]. The proposed approach first applies the park’s vector transform to the three-phase current signal, and then the magnitude of the two park’s vector components is obtained as a signal. The recurrence plot of this signal is recorded as an image and classified by different deep neural network architectures. It has been proven that the proposed deep neural network architecture outperforms other transfer learning models.

Although many methods have been proposed for motor fault detection, it is seen as a shortcoming that a working method has not been developed for different sizes, load conditions, and fault types. Many approaches have been proposed with a data-driven classification process according to the characteristics obtained from the good motor and the faulty motor. But finding an equivalent motor operating in the industry is always difficult. It has therefore become a necessity to create a reference virtually to which the running motor can be compared. Digital twin technology simulates real-world physical assets in the form of digital models with high-density, multidimensional, and powerful real-time features [8]. Digital twins help to achieve flexible, faster, quality, and personalized products. It also provides businesses with significant benefits such as security, diagnostics, and querying the current forecast status. The digital twin improves performance to exceed and achieve business quality in products and services, accelerating time in achieving results, and increasing hits in decision-making time. A custom-built digital twin model leverages physics-based simulation models and data-driven intelligence to provide insight into the occurrence of the fault. A study has been proposed for the detection of rotor failures using digital twin technology [13]. In order to increase the adaptability of the developed model, a model update method based on parameter sensitivity analysis has been proposed. The digital twin of the rotor system has been created and the imbalances occurring here have been determined.

Although many studies on the detection of stator faults have been suggested in the literature, there is little study on fault diagnosis with signals obtained from the digital twin. In addition, the estimation of parameters such as frequency, phase angle, and amplitude of the motor current signals and the spectrum region to be examined for fault diagnosis will be determined within the scope of this study. In this study, firstly, healthy and faulty motor signals will be taken from the motor modeled with ANSYS. Afterward, the park’s vector transformation and module of this transformation will be calculated on these signals. In addition, the supply frequency and phase angle will be calculated for the current signal obtained. The information entropy will be used to obtain information about the fault in the spectrum of the received signal. Then, the faults are determined according to the change in the spectrum by using a support vector machine classifier.

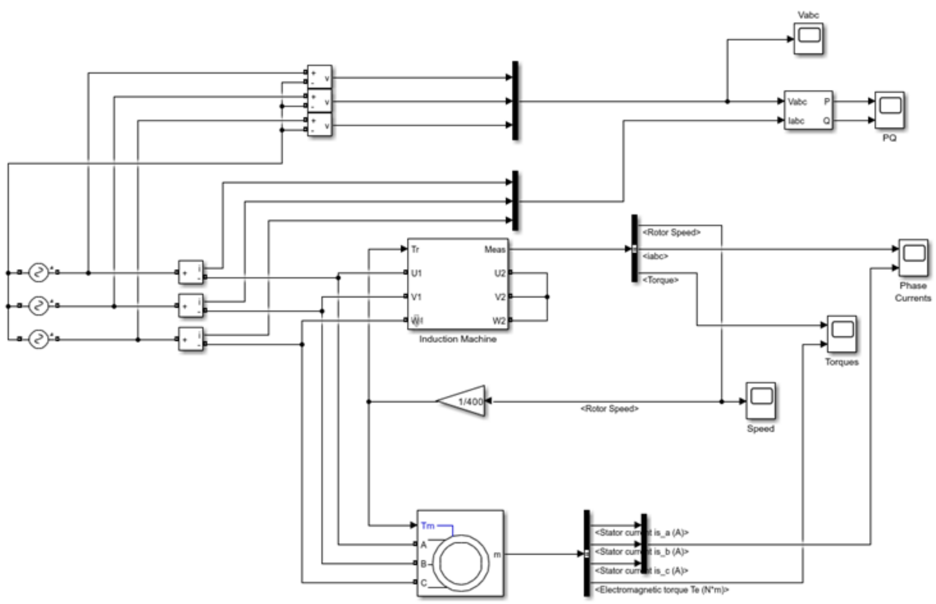


Figure 1. 500W Induction motor model obtained using the voltage approach behind the reactance

2. Digital Twin Model of Induction Motor

The design of the motor, which is healthy and has different degrees of stator faults, was created in the ANSYS program and the motor model was created in the Matlab/Simulink [14]. Using the 500W induction motor model, which is obtained using the voltage approach behind the modeled motor reactance, the short-circuit resistance due to the loss of insulation in the stator winding occurs as inter-turn short-circuits that rapidly transform into phase-to-phase or phase-to-earth faults if the fault is not detected and isolated quickly. By changing the fault rate, the simulation was run with a frequency of 10kHz, and current changes were obtained. The variable error rate is taken as 0.05% - 5% - 10%. It was obtained by providing the error resistance to be 0.20 – 1000 frequency 10kHz. The Simulink model used is given in Figure 1. Fault conditions are created by changing the motor parameters in Figure 1 as shown in Figure 2.

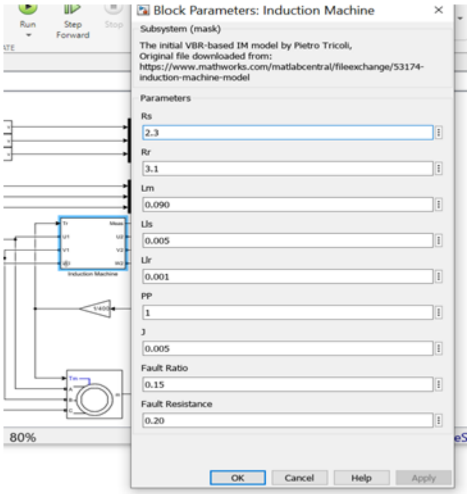


Figure 2. Fault rate and fault resistance change screen

After the motor model is created, healthy and faulty signals will be obtained from the motor. The severity of the fault will be determined by running the methods on the signals obtained.

3. The Proposed Entropy-Based Diagnosis Model

A new feature extraction approach has been proposed for the detection of stator faults in the induction motor from the current signals obtained. This approach basically consists of two stages. The first step is to obtain a new feature signal by park's vector conversion from the three-phase current signals. Then, the mutual information is estimated from the obtained feature signal. In the mutual information method according to the frequency, the interval to be analyzed is determined. The obtained signals are then distinguished as healthy and faulty according to the support vector machine classifier. The proposed method has been tested under different operating speeds and successful results have been obtained. The block diagram of the proposed method is shown in Figure 3.

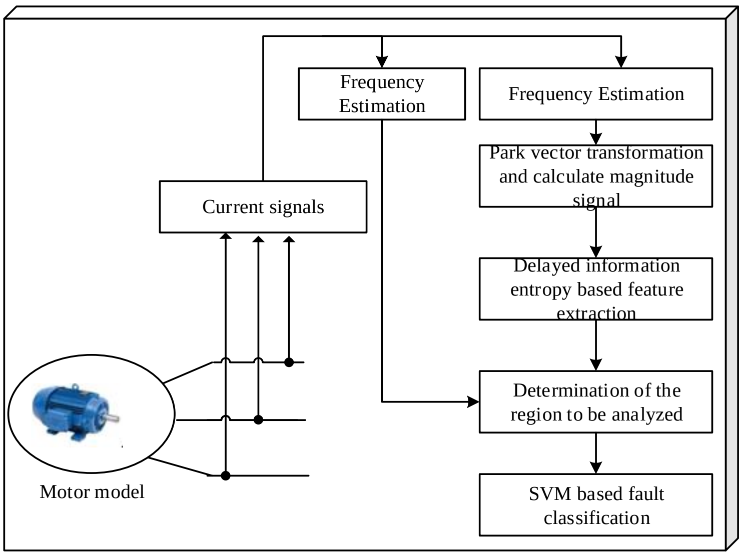


Figure 3. Information entropy-based stator fault diagnosis

With the system given in Figure 3, the three-phase current signals are first received from the motor. Then, the feature signals are obtained by parking vector transformation. The supply frequency of the motor is also determined by using three-phase current signals. Then, the signal obtained by taking the module of the two parking vector components is given to the mutual information method. According to the frequency information obtained in the frequency estimation, the region of the delay component affected by the fault is determined and the signals in this region are given to the support vector machine and the classification process is performed.

3.1. Park's Vector Transform

The park's vector transform allows the characteristics of the phase currents to be expressed with two components in three-phase induction motors. The park vector components obtained by using three-phase currents are shown in (1).

$$I_d = \sqrt{\frac{2}{3}} i_A - \frac{1}{\sqrt{6}} i_B - \frac{1}{\sqrt{6}} i_C$$
$$I_q = \frac{1}{\sqrt{2}} i_B - \frac{1}{\sqrt{2}} i_C \qquad (1)$$

In (1), i_A, i_B, $and_{,} i_C$ represent the phase currents. Under ideal conditions, the three-phase currents form the parking vector with the following components.

$$I_d = \frac{\sqrt{6}}{2} i_M \sin \omega t$$
$$I_q = \frac{\sqrt{6}}{2} i_M \sin(\omega t - \frac{\pi}{2}) \qquad (2)$$

In this equation, $i_{M,}$ ω, and t represent the maximum value of the phase supply current, supply frequency, and time variable, respectively. The park's transformation on the three-phase current signal is shown in Figure 4.

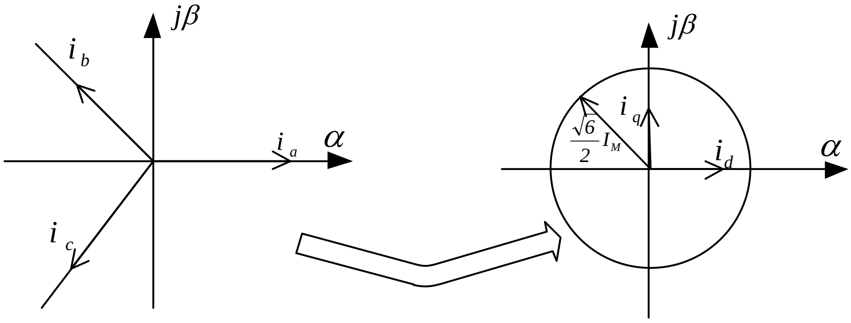


Figure 4. Park's vector transformation

The park's vector transform creates a circular pattern at the center of the two components. By monitoring the deviations in the patterns obtained with this transformation, faulty conditions are detected. Park's vector transform is used to determine voltage unbalance, broken rotor bar faults, winding faults, and stator open phase faults. While the two-dimensional pattern formed by the Park's vector components in a healthy motor is in the form of a circle, it turns into an ellipse in the case of faults. After the park's vector components are obtained, the modulus of the two components is obtained according to (3).

$$I_M = \sqrt{I_d^2 + I_q^2} \qquad (3)$$

3.2. Mutual Information

In recent years, information theory has been successfully applied to different pattern recognition problems. The performances of the methods in information theory in this area are their ability to measure the relationships between system variables and their performance in classification and pattern recognition. The two main measurements used in this field are entropy and mutual information. Entropy is a method used to measure the uncertainty in a system and has also been used in fault diagnosis. On the other hand, the mutual information method is a common knowledge that measures the independence between two variables.

Mutual information can measure the linear dependence of two variables. Estimating information from observations of one random event in another is a metric derived from Shannon's information theory. It can be considered as a nonlinear equivalent of the correlation function. Usually, mutual information is measured between signals such as X and Y received from two different systems. However, in this study, there is only a one-time series obtained from the parking vector transformation. Therefore, the system used is based on the automatic mutual information method.

$$M(X_t, X_{t+\tau}) = \sum_{x_n, x_{n+\tau}} P(x_n, x_{n+\tau}) \ln\left(\frac{P(x_n, x_{n+\tau})}{P(x_n)P(x_{n+\tau})}\right) \quad (4)$$

In (4), the value of $P(x_{n+\tau})$ is the common probability density for the signal $x_{n+\tau}$and x_n. $P(x_n)$ represents the individual probability density for x_n and $P(x_{n+\tau})$ represents the individual probability distribution for $x_{n+\tau}$. Here x_n and $x_{n+\tau}$ are signal samples of the same series delayed from each other by τ time delay.

In this study, the time series is taken as the module of two parking vector transformations, and the mutual information value of the original signal and the time-delayed signal is calculated. Since we have limited data, mutual information is obtained as a sample estimator. The park's vector transform module exhibits different distortions due to stator failures in a certain period of time. The resulting feature signal is given to a support vector machine classifier and faults are determined.

3.3. Frequency Estimation and Determination of the Mutual Information Region and Fault Classification

In order to estimate the motor operating condition, it is necessary to determine the angular frequency. For this purpose, the angular frequency is first found with a simple search method. The angular frequency is expressed by the following equation.

$$\omega = 2 * \pi * f \quad (5)$$

In (5), *f* represents the supply frequency of the motor. Normally for a line-fed motor, this frequency is 50 Hz. However, in the case of a power supply with a driver, this frequency is constantly changing. Therefore, it has to be estimated. For its estimation, the value that maximizes the following function must be estimated.

$$\arg\max_{\omega} \frac{1}{3} \sum_{k=0}^{2} x_k P_G(\omega) x_k \quad (6)$$

In (6), $P_G(\omega)$ is calculated by using (7).

$$P_G(\omega) = G(\omega)(G(\omega)^T G(\omega))^{-1} G(\omega)^T \quad (7)$$

In (7), $G(\omega)$ is a matrix of Nx2 size and is defined as follows.

$$G(\omega_0) = \left[g(\omega_0) \quad g(\omega_0)^* \right] \quad (8)$$

The (.)* in (8) represents the complex conjugate and $g(\omega_0)$ is given according to (9).

$$g(\omega_0) = \left[1 \quad e^{j\omega_0} \quad \ldots \quad e^{j\omega_0 (N-1)} \right]^T \quad (9)$$

After estimating the supply frequency of the motor, the region to be examined in the mutual information method should be determined. For this purpose, the time delay to be examined is estimated according to the following equation.

$$location = F_S x(1/2xf) \quad (10)$$

After the fault region is obtained from mutual information, this region is recorded as a dataset for healthy and each faulty condition. The obtained dataset is given to support vector machine classifier and faults are classified.

4. Experimental Results

A tuned resistor is connected to one of the motor phases to generate stator faults. The degree of faults has been changed to obtain different faults. Then, starting from 20 Hz to 50 Hz, the data was obtained by running the motor at different supply frequencies with 10 Hz intervals. In Figure 5, phase signals are given for a healthy condition and two stator-related fault conditions.

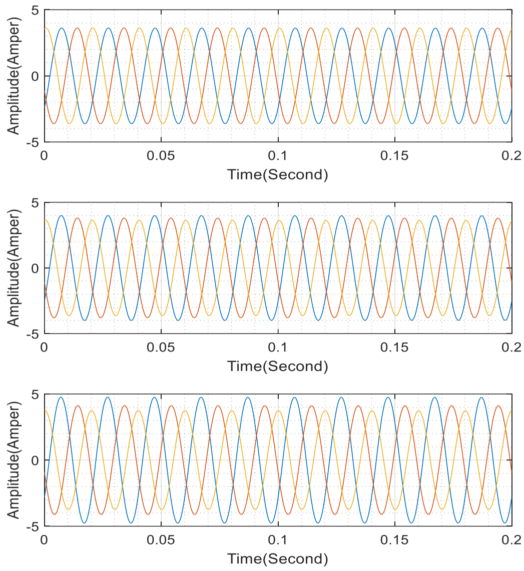


Figure 5. Current signals for solid and two fault states (a) Solid, 5% and 10% of fault resistance altered fault

As can be seen in Figure 5, it is very difficult to detect faults using direct current signals. Therefore, feature extraction was performed with preprocessing. In Figure 6, the feature extraction results for the healthy and two-faulty conditions are given.

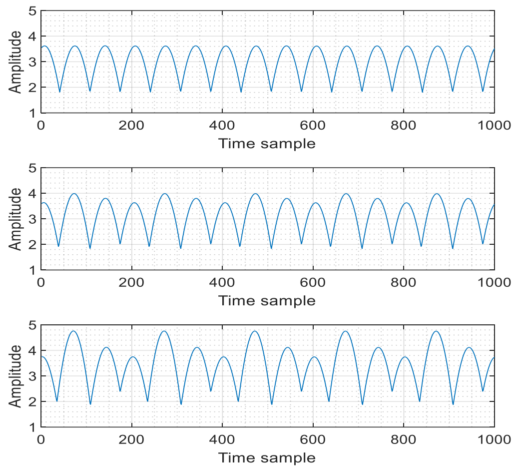
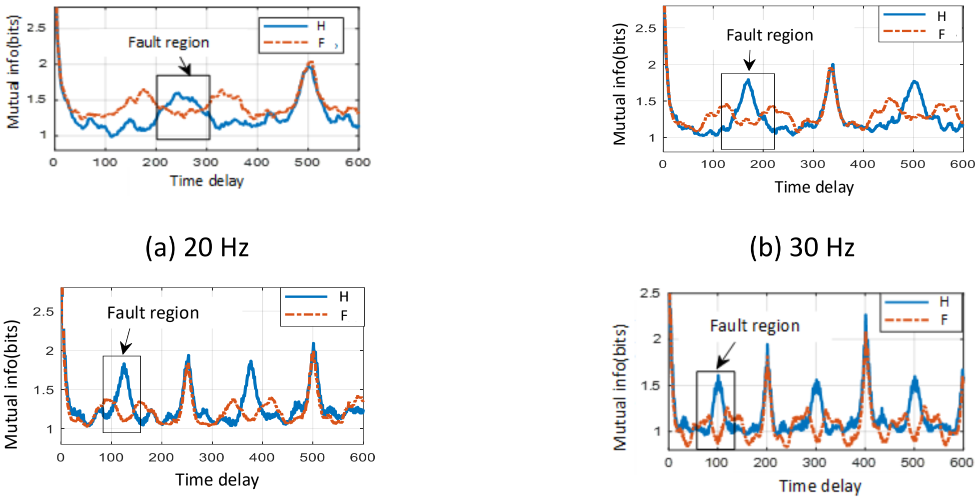


Figure 6. Feature signals for healthy and two fault states (a) Solid, 5% and 10% of fault resistance altered fault.

As can be seen in Figure 6, both the amplitude changes and local peaks occur on the sides of the sinus signal in faulty situations. In Figure 7, mutual information results are given for the fault created.

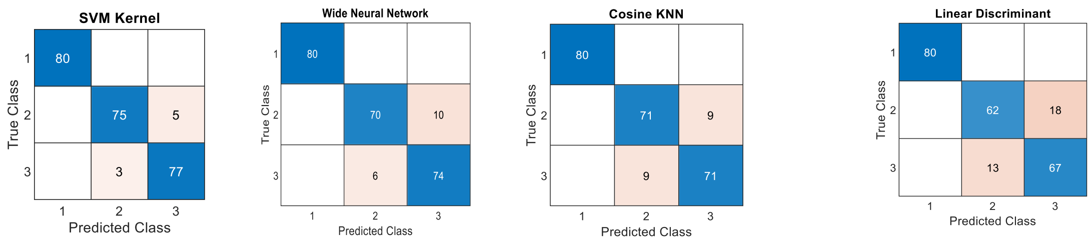
The results of the proposed method at four different operating frequencies are given in Figure 7. The equation in the location formula given before was used to determine the area related to the fault. In fact, the basic operation is to determine a period data field. At the midpoint of this one-period data, and around it, the mutual information represents a significant variation. For example, in feeding with 20 Hz, this value becomes a peak around place=(10000/(2*20))=250. This peak occurs at approximately 166 time delays for 30 Hz, delays for 40 Hz, and 100-time delays for 50 Hz. By examining the amplitudes at the 50-unit delay taken from the right and left sides of the peak, it can be determined whether a fault has occurred. Matlab classification Learner is used for fault classification. Apart from support vector machines, the performance of other classifiers is also evaluated. The complexity matrix of the best-performing classifiers is given in Figure 8.



(a) 20 Hz (b) 30 Hz

(c) 40 Hz (d) 50 Hz

Figure 7. Mutual information of healthy and faulty conditions for different operating frequencies



(a) SVM (b) WideNN (c) Cosine KNN (d) Linear Discriminant

Figure 8. Fault classification with different machine learning approaches.

The performance of each machine learning was calculated according to the confusion matrices given in Figure 8. For this purpose, accuracy, precision, recall, and F1 measurements were used. In Table 1, performance criteria are given according to these confusion matrices.

Table 1. Comparisons of performance metrics

The method	F1	Precision	Recall	Accuracy
SVM	96,68	96,69	96,67	96,67
WideNN	93,37	93,40	93,33	93,33
Cosine KNN	92,50	92,50	92,50	92,50
Linear Discriminant	87,12	87,08	87,16	87,08

As seen in Table 1, support vector machines gave better classification performance in all metrics. Support vector machine gives better generalization performance.

5. Conclusion

In this study, signals obtained from the digital twin model of a large powerful motor are used to detect faults in an induction motor. The digital twin model is of great importance, especially in terms of creating a reference model for an engine operating in industrial environments. In addition, different operating frequencies were determined by making parameter estimations from the signals obtained. The area related to the failure in the information entropy of the signal obtained according to the operating frequency was determined and the severity of the failure was classified with the support vector machines. Compared to other machine learning methods, a 3%

performance increase has been achieved. In future studies, the results of the real engine and the digital model will be compared and only a solid model of an engine in the industry will be created to detect the faults.

Acknowledgments

This work was supported by the TUBITAK (The Scientific and Technological Research Council of Turkey) under Grant No: 122E412.

6. References

[1] Benbouzid, Mohamed (2021) Signal processing for fault detection and diagnosis in electric machines and systems. Institution of Engineering and Technology, London.

[2] Habbouche, H., Amirat, Y., Benkedjouh, T., and Benbouzid, M (2021) Bearing fault event-triggered diagnosis using a variational mode decomposition-based machine learning approach. *IEEE Transactions on Energy Conversion*, 37(1), 466-474.

[3] Almounajjed, A., Sahoo, A. K., Kumar, M. K., and Assaf, T (2022) Fault diagnosis and investigation techniques for induction motor. *International Journal of Ambient Energy*, 43(1), 6341-6361.

[4] Aydin, I., Kaner, S (2020) A New Hybrid Diagnosis of Bearing Faults Based on Time-Frequency Images and Sparse Representation. *Traitement du Signal*, 37(6).

[5] Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P., and Gao, R. X (2019) Deep learning and its applications to machine health monitoring. *Mechanical Systems and Signal Processing*, 115, 213-237.

[6] Zhang, S., Zhang, S., Wang, B., and Habetler, T. G (2020) Deep learning algorithms for bearing fault diagnostics—A comprehensive review. *IEEE Access*, 8, 29857-29881.

[7] Almounajjed, A., Sahoo, A. K., Kumar, M. K., and Subudhi, S. K (2023) Stator Fault Diagnosis of Induction Motor Based on Discrete Wavelet Analysis and Neural Network Technique. *Chinese Journal of Electrical Engineering*, 9(1), 142-157.

[8] Husari, F., Seshadrinath, J (2021) Early stator fault detection and condition identification in induction motor using novel deep network. *IEEE Transactions on Artificial Intelligence*, 3(5), 809-818.

[9] Okwuosa, C. N., Hur, J. W (2023) An intelligent hybrid feature selection approach for SCIM inter-turn fault classification at minor load conditions using supervised learning. *IEEE Access*.

[10] Aishwarya, M., Brisilla, R. M (2023) Design and Fault Diagnosis of Induction Motor Using ML-Based Algorithms for EV Application. *IEEE Access*, 11, 34186-34197.

[11] Lucas, G. B., De Castro, B. A., Ardila-Rey, J. A., Glowacz, A., Leão, J. V. F., and Andreoli, A. L (2023) A Novel Approach Applied to Transient Short-Circuit Diagnosis in TIMs by Piezoelectric Sensors, PCA, and Wavelet Transform. *IEEE Sensors Journal*, 23(8), 8899-8908.

[12] Das, A. K., Das, S., Pradhan, A. K., Chatterjee, B., and Dalai, S (2023) RPCNNet: A Deep Learning Approach to Sense Minor Stator Winding Inter-Turn Fault Severity in Induction Motor under Variable Load Condition. *IEEE Sensors Journal*, 23(4), 3965-3972.

[13] Wang, J., Ye, L., Gao, R. X., Li, C., and Zhang, L (2019) Digital Twin for rotating machinery fault diagnosis in smart manufacturing. *International Journal of Production Research*, 57(12), 3920-3934.

[14] Şahin, İ. L. K. E. R., Bayazıt, G. H., and Keysan, O. (2020) A simulink model for the induction machine with an inter-turn short circuit fault. IEEE International Conference on Electrical Machines (ICEM) pp. Gothenburg, Sweden 1273-1279.