



# Bearing fault detection in adjustable speed drives via self-organized operational neural networks

Sertac Kilickaya<sup>1</sup> · Levent Eren<sup>2</sup>

Received: 1 August 2024 / Accepted: 18 September 2024  
© The Author(s) 2024

## Abstract

Adjustable speed drives (ASDs) are widely used in industry for controlling electric motors in applications such as rolling mills, compressors, fans, and pumps. Condition monitoring of ASD-fed induction machines is very critical for preventing failures. Motor current signature analysis offers a non-invasive approach to assess motor condition. Application of conventional convolutional neural networks provides good results in detecting and classifying fault types for utility line-fed motors, but the accuracy drops considerably in the case of ASD-fed motors. This work introduces the use of self-organized operational neural networks to enhance the accuracy of detecting and classifying bearing faults in ASD-fed induction machines. Our approach leverages the nonlinear neurons and self-organizing capabilities of self-organized operational neural networks to better handle the non-stationary nature of ASD operations, providing more reliable fault detection and classification with minimal preprocessing and low complexity, using raw motor current data.

**Keywords** Bearing fault detection · Condition monitoring · Motor current signature analysis · Operational neural network

## 1 Introduction

Early detection and diagnosis of motor faults, which are generally classified into mechanical and electrical categories, are essential for maintaining reliable and cost-effective operation. Methods for fault detection and diagnosis (FDD) can be classified into three main types: signal-based, model-based and knowledge-based strategies. Model-based techniques create analytical models based on physical principles and system identification, but they become difficult to apply to complex systems. Signal-based methods focus on analyzing various signals, such as those related to vibration, motor current, speed, and temperature, to identify and diagnose

faults. To detect faults in electrical machinery, frequently used approaches for processing these signals include spectral estimation [3, 34], wavelet transformation and wavelet packet decomposition [6–8], time-frequency analysis [29], fast Fourier transform (FFT) [2, 10, 12], scale-invariant feature transform (SIFT) [5], and sequence analysis [27]. These methods utilize advanced signal processing techniques to extract meaningful information from the data and detect anomalies indicative of potential issues, such as bearing faults.

In contrast to classical methods, data-driven condition monitoring systems harness large volumes of data collected through advanced data acquisition and control systems. Shallow machine learning models employed for this purpose have demonstrated satisfactory performance in detecting and diagnosing faults using motor current data [32]. However, they often depend on manually crafted features and diverse classifiers, typically using limited data from motors or rotating machinery (RM). As a result, their effectiveness diminishes when applied to different types of systems, fault conditions, or larger datasets, making it challenging to develop a generic solution. Detecting faults in machines fed with ASD is a clear example of this problem, as these systems often operate at varying motor speeds.

---

Levent Eren contributed equally to this work.

---

✉ Sertac Kilickaya  
sertac.kilickaya@tuni.fi

Levent Eren  
levent.eren@ieu.edu.tr

<sup>1</sup> Faculty of Information Technology and Communication Sciences, Tampere University, Korkeakoulunkatu 7, 33720 Tampere, Finland

<sup>2</sup> Department of Electrical and Electronics Engineering, Izmir University of Economics, Sakarya Caddesi No:156, 35330 Izmir, Turkey

Numerous data-driven deep neural network (DNN) models have been proposed in the literature as solutions for FDD problem. Deep neural networks (DNNs) have the ability to automatically extract discriminative features from raw input data through the training process, which removes the necessity for manually designed statistical or transform-domain feature representations. Nevertheless, these models rely on large, well-labeled datasets to achieve effective training. In their work, Jia et al. [17] introduced a five-layer DNN model designed for fully automated intelligent FDD of rotating machinery. The model utilizes an unsupervised autoencoder for pretraining and processes vibration signal frequency spectra as input. Their method achieved a 95.8% classification accuracy on the Case Western Reserve University (CWRU) bearing dataset. A recent study by Ye et al. [36] developed an intelligent fault diagnosis method for rolling bearings that utilizes motor stator current signals, combining feature reconstruction (FR) with convolutional neural networks (CNNs) to achieve high-precision diagnosis with approximately 99% accuracy for faulty bearings. While the FR method eliminates supply frequency and its harmonics, it further increases computational complexity during preprocessing. In [37], vibration spectrum imaging (VSI) is employed to convert normalized spectral amplitudes from segmented vibratory signals into images, which are then used to train a CNN for bearing fault classification. The proposed VSI-CNN network surpasses previous methods, achieving a classification accuracy of approximately 99%. In [13], the challenges of using vibration signals for bearing fault diagnosis are discussed, highlighting their high cost and impracticality due to the need for external accelerometers. Instead, motor current signals, which are easier to measure via inverters, are used for fault diagnosis. The study introduces a deep learning method that uses raw signals from multiple motor phases, extracts features, and classifies them with CNNs. To improve accuracy, a decision-level information fusion technique combines information from all CNNs. The method's effectiveness was validated with experiments using actual bearing fault signals. Maximal overlap discrete wavelet transform (MODWT) was applied in [1] to extract features from stator current signals, converting them into a two-dimensional array. After further processing, this method identified fault patterns with an accuracy exceeding 90%. The method introduced in [11] diagnoses bearing fault progression by analyzing time-domain current signals in a semicycle and using them directly in a pattern classifier. Validated under various conditions, the approach achieves over 97% accuracy for line-fed motors and over 71% for inverter-fed motors. Efforts to improve diagnosis performance through sensor data fusion have also been made in several studies. Wan et al. [35] presented a fusion multiscale convolutional neural network (F-MSCNN), which is tailored to adapt to different speeds. The F-MSCNN utilizes raw sound and vibration

data, using a fusion layer and a multiscale convolutional layer at the start to extract a variety of features for classification. Comparative evaluations demonstrated that F-MSCNN performs well in speed generalization, with its accuracy increased by integrating sound and vibration data. Qian et al. [30] proposed a motor FDD approach using multi-feature fusion within an enhanced CNN framework. Their approach includes preprocessing current and vibration signals, implementing segmented multi-time window synchronous input, and conducting multiscale feature extraction along with time series fusion within the same time window. Experimental validation on a fault simulation platform demonstrated that integrating vibration and current signal features significantly enhances fault diagnosis accuracy and stability compared to single signal inputs. However, this method can be computationally intensive and may require significant processing power and time, potentially limiting its practical application in real-time scenarios.

To leverage computational efficiency along with enhanced performance, several studies [4, 14, 19, 28] have utilized one-dimensional (1D) CNNs with raw data or extracted features for machinery fault diagnosis under fixed conditions. Even though 1D CNNs performs well under fixed conditions, previous research [20, 21] indicates that CNNs with a uniform network architecture based on a first-order neuron model often fail to effectively address problems involving complex and highly nonlinear solution spaces. These models require a significant network depth and complexity to be effective. To overcome these limitations, Self-organized Operational Neural Networks (Self-ONNs) have been introduced, which offer a high level of heterogeneity and the ability to optimize its operators, thereby maximizing learning performance [22]. Additionally, the effectiveness of Self-ONNs has been validated in various studies focusing on motor fault diagnosis using vibration data [15, 16].

In this study, we compare the performance of state-of-the-art 1D Self-ONNs with commonly used 1D CNNs for bearing fault detection and classification using raw current data in ASD-fed machines. Figure 1 provides an overview of the implementation process for our proposed method. It details the preprocessing steps, including segmentation and min-max normalization, with an example output waveform to illustrate data transformation. The figure also outlines the architecture of the 1D Self-ONN, highlighting its three self-operational layers and two dense layers, and presents the final output with possible predictions. With the proposed pipeline, our aim is to demonstrate that 1D Self-ONNs offer competitive diagnostic performance while significantly reducing computational complexity. To offer a more robust solution, Self-ONNs integrate feature extraction and classification into a unified framework by utilizing the motor's raw current data. Its lower computational complexity makes it well-suited for real-time implementation in embedded systems as well.

The structure of the paper is as follows: Section 2 examines 1D Self-ONNs, comparing them to traditional 1D CNNs and ONNs. Section 3 discusses the experimental setup and outcomes using a benchmark ASD-fed motor current dataset, evaluating the proposed method’s performance against 1D CNNs. Section 4 concludes the paper and suggests areas for future research.

## 2 1D self-organized operational neural networks

This section introduces the concept of generative neurons and their integration into 1D Self-ONNs. The classical linear neuron model forms the traditional CNN architecture, which also includes constraints such as limited connections and weight sharing at the kernel level. These constraints give rise to the convolution equations employed in CNNs.

The output of the  $k$ -th neuron in the  $l$ -th layer of a 1D CNN can be expressed as follows:

$$x_k^l = b_k^l + \sum_{i=0}^{N_{l-1}} x_{ik}^l \tag{1}$$

where  $b_k^l$  represents the bias of the corresponding neuron and  $x_{ik}^l$  can be written as:

$$x_{ik}^l = \text{Conv1D}(w_{ik}, y_i^{(l-1)}) \tag{2}$$

In this equation,  $w_{ik} \in \mathbb{R}^K$  represents the kernel that connects the  $i^{\text{th}}$  neuron of the  $(l - 1)^{\text{th}}$  layer to the  $k^{\text{th}}$  neuron of the  $l^{\text{th}}$  layer. On the other hand,  $x_{ik}^l \in \mathbb{R}^M$  is the input map, and  $y_i^{(l-1)} \in \mathbb{R}^M$  is the output of the  $i^{\text{th}}$  neuron in the  $(l - 1)^{\text{th}}$  layer.

The convolution operation, as outlined in Eq. 2, is expanded in the following form:

$$x_{ik}^l(m) = \sum_{r=0}^{K-1} w_{ik}^l(r) y_i^{(l-1)}(m+r) \tag{3}$$

The kernel  $w_{ik}$  and the shifted versions of the  $i$ -th neuron’s output  $y_i^{(l-1)}$  in the  $(l - 1)$ -th layer are multiplied element-wise and summed over the kernel’s length to produce an  $M$ -dimensional input vector  $x_{ik}^l$ .

On the other hand, in ONNs, Eq. 3 can be generalized, yielding a concise representation for the output of an operational neuron:

$$\bar{x}_{ik}^l(m) = P_k^l \left( \psi_k^l(w_{ik}^l(r), y_i^{(l-1)}(m+r)) \right)_{r=0}^{K-1} \tag{4}$$

where  $\psi_k^l(\cdot) : \mathbb{R}^{M \times K} \rightarrow \mathbb{R}^{M \times K}$  and  $P_k^l(\cdot) : \mathbb{R}^K \rightarrow \mathbb{R}^1$  are called nodal and pool operators, respectively, assigned to the  $k$ -th neuron of the  $l$ -th layer.

The Greedy Iterative Search (GIS) algorithm is often used iteratively to explore a potential set of operators, aiming to identify an optimal combination of pool  $P$  and nodal  $\psi$  operators for an ONN. Subsequently, these optimal operators are allocated for all neurons within the respective hidden layer such that the final ONN configuration is formed. However, the conventional ONN architecture suffers from several limitations [22]. Firstly, it limits heterogeneity by using the same operator set for all neurons within a layer. Secondly, the process of manually crafting a selection of potential operators and seeking the optimal one for each neuron introduces considerable overhead. Lastly, the inability to express the appropriate operator with well-defined functions limits adaptability and customization to suit the specific learning problem.

To address these issues, Self-ONNs with generative neurons were introduced [22]. Unlike conventional CNN and ONN architectures, Self-ONNs with generative neurons offer flexibility by generating nodal operators during training without predefined sets or prior search processes. This implementation eliminates the need for a single nodal operator across all neurons in a hidden layer, as each neuron in Self-ONNs can generate various nodal operators. Figure 2 illustrates the 1D kernels of CNN, ONN, and Self-ONN with generative neurons, highlighting that while CNN and ONN architectures feature fixed nodal operators for convolutional and operational neurons, Self-ONNs are able to produce any nodal operator  $\Psi$  for each kernel element as the training progresses.

The nodal operators within Self-ONNs are derived through the application of Taylor series function approximation, where for a function  $f(x)$  with infinitely differentiable properties, the Taylor series is represented around a given point  $a$  as follows:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x - a)^n \tag{5}$$

Then, we can approximate Eq. 5 up to the  $Q$ -th order, and express the Taylor polynomial as:

$$f(x)^{(Q,a)} = \sum_{n=0}^Q \frac{f^{(n)}(a)}{n!} x^n \tag{6}$$

This equation facilitates the approximation of any function  $f(x)$  around a given point  $a$ . During backpropagation, the coefficients  $\frac{f^{(n)}}{n!}$  are optimized iteratively to customize the nodal operator for each kernel element. For instance, if neuron outputs are bounded by the tanh activation function, the

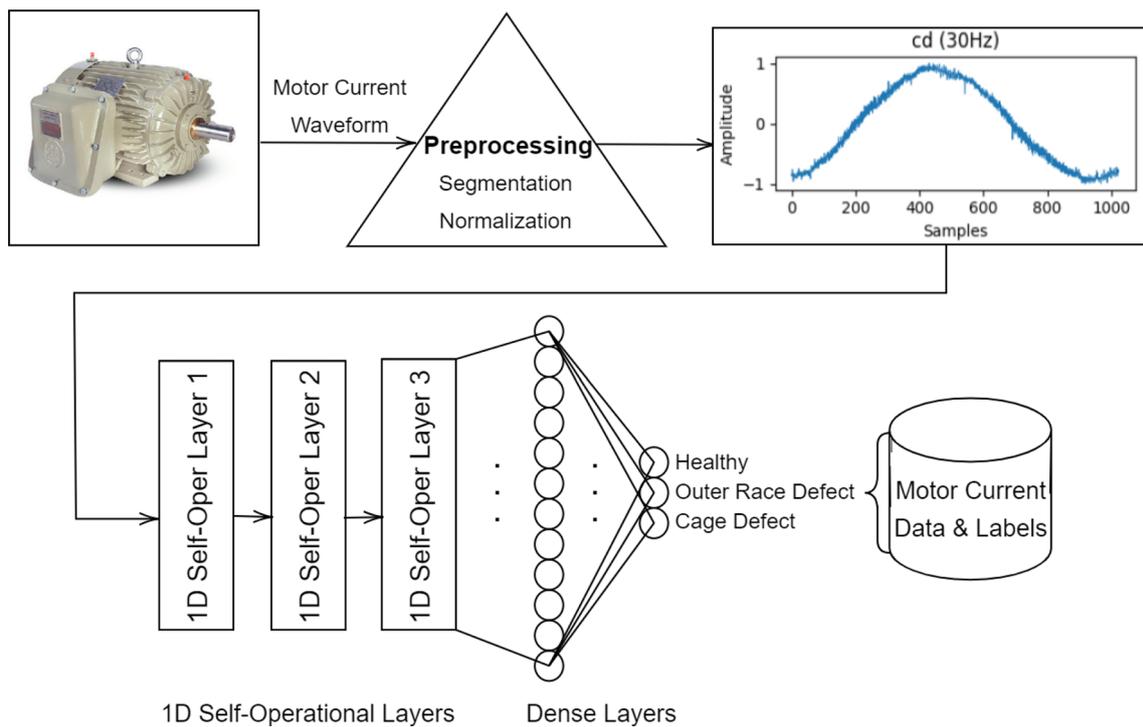


Fig. 1 Overall system diagram

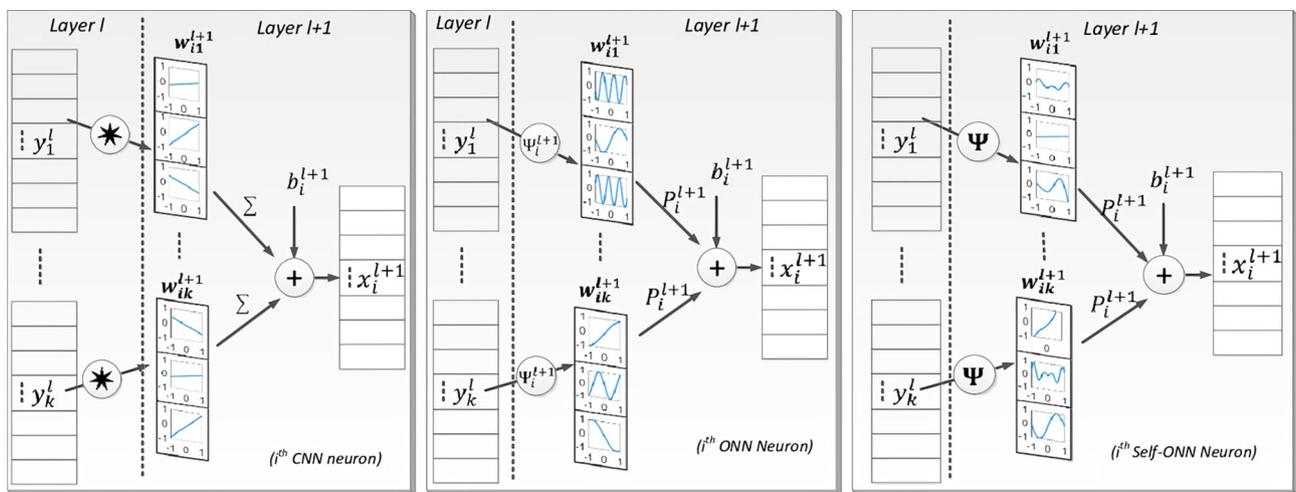


Fig. 2 1D nodal operations of the  $i$ -th neuron of CNN (left), ONN (middle), and Self-ONN (right) [15]

$Q$ -th order Maclaurin series allows the generation of diverse transformations near the midpoint 0. This principle underlies the concept of generative neurons in Self-ONNs. The nodal transformation of a generative neuron can then be summarized in the following form:

$$\tilde{\psi}_k^l(w_{ik}^{l(Q)}(r), y_i^{(l-1)}(m+r)) = \sum_{q=1}^Q w_{ik}^{l(Q)}(r, q)(y_i^{(l-1)}(m+r))^q \quad (7)$$

The  $K \times 1$  kernel vector  $w_{ik}^l$  in 1D CNN topology is replaced by a  $K \times Q$  matrix  $w_{ik}^{l(Q)}$  where  $Q$  is the degree

of the Taylor polynomial in Self-ONNs. This matrix  $w_{ik}^{l(Q)}$  is created by substituting each element  $w_{ik}^l(r)$  with a  $Q$ -dimensional vector  $w_{ik}^{l(Q)}(r) = [w_{ik}^{l(Q)}(r, 0), w_{ik}^{l(Q)}(r, 1), \dots, w_{ik}^{l(Q)}(r, Q-1)]$ .

Therefore, the operator  $\tilde{\psi}_k^l$  varies for each individual output  $y_i^{(l-1)}$ , leading to  $Q$  times the number of parameters present in the CNN model. Finally, the input map of the generative neuron  $\tilde{x}_{ik}^l$  is determined as follows:

$$\tilde{x}_{ik}^l(m) = P_k^l \left( \psi_k^l \left( w_{ik}^l(r), y_i^{(l-1)}(m+r) \right) \right)_{r=0}^{K-1} \quad (8)$$

To sum up, the Self-ONN model differs from traditional CNNs in that it employs a unique approach to nonlinearity. Basically, each layer of the Self-ONN utilizes several powers of activations, as determined by the hyperparameter  $Q$ , to create a more flexible neuron model. This design enhances the network's ability to learn from challenging datasets by capturing nonlinear relationships more effectively. In addition, Self-ONNs offer several advantages over CNNs and ONNs. Firstly, they eliminate the need to search for optimal operators for each neuron connection by enabling self-organization of network operators through generative neurons during training. Secondly, they allow for greater heterogeneity by not restricting each kernel connection to a single nodal operator, unlike ONNs. Lastly, Self-ONN layers offer greater parallelization efficiency compared to ONNs. Self-ONNs utilize the standard backpropagation (BP) algorithm for learning, similar to CNNs. The network weights and biases are updated by calculating the gradient of the loss function with respect to each parameter during training. Detailed forward and backpropagation formulations for Self-ONN neurons are described in [22] and [25].

### 3 Test and evaluation

In testing, motor current signals are collected for healthy bearings and bearings with two different fault types (outer race and cage defect) and three different input frequencies (60 Hz, 45 Hz and 30 Hz) are used to evaluate the performance of 1D Self-ONNs. Current waveforms are captured from the ASD test setup in Fig. 3 [33]. The setup incorporates an inverter that generates pulse width modulation (PWM) voltage signals with fundamental frequencies between 2 and 60 Hz, alongside a carrier frequency fixed at 9.2 kHz. It is paired with a 745.7 W, 3450 rpm, 208 V, 60 Hz, 3-phase, 2-pole induction motor. The motor employs ORS 6203-ZZ-C3 bearings on both ends, each with eight balls. Specifically, the bearings on the shaft-end undergo tests for outer race defect (OD) and cage defect (CD), against a healthy bearing.

We chose to focus on outer race defects and cage defects as they represent two of the most prevalent types of bearing faults in induction motors. Outer-race defects are particularly critical due to their direct impact on the bearing's interaction with the load, which can result in severe operational problems. On the other hand, cage defects can disrupt the even distribution of load across the bearing balls, accelerating wear and potentially leading to failure.

For the cage defect, the bearing cage was intentionally deformed by pressing a center punch between two adjacent balls, disrupting the usual rotation of the cage. To simulate an outer race defect, a radial load was applied using a belt-driven mechanism. In this experiment, a single 1/32 inch diameter hole was drilled into the outer race to generate a consistent

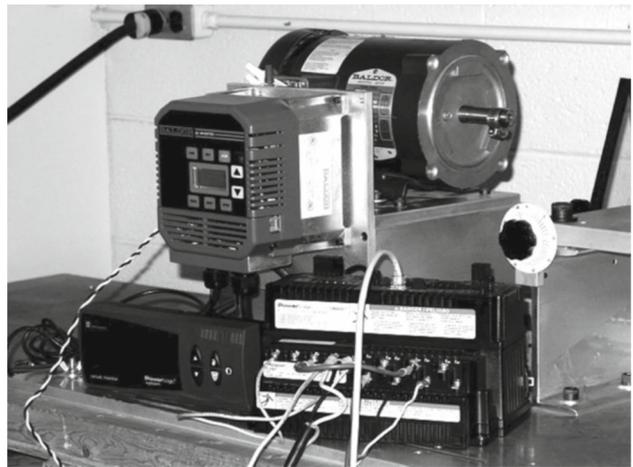


Fig. 3 ASD test setup



Fig. 4 Cage defect

defect for testing purposes. The bearings with cage defect and outer race defects (depicted with 2 holes) are shown in Figs. 4 and 5, respectively. These illustrations are for reference purposes; in the actual test, the outer race had only one 1/32 inch diameter hole.

During the evaluation of defective bearings, the PWM outputs from the inverter were captured at 30, 45, and 60 Hz, resulting in no-load speeds of 1797, 2697, and 3596 rpm, respectively. Data is collected using a SquareD CM4000 series Circuit Monitor, which can sample 3-phase voltages and currents up to 8 channels, at a rate of 30,720 Hz per channel. The Circuit Monitor has built-in memory for storing waveform data, which can be transferred to a PC through a serial or Ethernet connection. To avoid any issues with aliasing, especially since the PWM inverter has a fixed frequency of 9.2 kHz, the stator current data is sampled at the highest rate permitted by the Circuit Monitor, which is 30,720 Hz. Figure 6 illustrates the normalized current waveforms of



**Fig. 5** Outer-race defect

the ASD system for the given input frequencies, and Fig. 7 depicts their fast Fourier transform (FFT).

The relationship between bearing fault characteristic frequencies and bearing geometry has been thoroughly documented in numerous studies. Mathematical expressions that connect these fault frequencies to parameters such as the number of rolling elements, shaft rotational speed, pitch diameter, and contact angle are also detailed in [15]. For this study, outer race and cage defect characteristic frequencies at different ASD system frequencies are calculated and provided in Tables 1 and 2.

Bearing faults generate mechanical vibrations at the given fundamental frequencies and these vibrations cause air gap eccentricity leading to irregularities in the air gap flux density. The fluctuations in flux density alter the machine's inductances, resulting in distortions in the stator current at the vibrational harmonics. For line-driven motors, the characteristic current frequencies,  $f_{CF}$ , caused by these vibration frequencies are calculated using the following equation:

$$f_{CF} = |f_s \pm kf_v| \quad (9)$$

where  $f_s$  is the power supply frequency in Hertz,  $f_v$  is the vibration frequency in Hertz, and  $k$  is the vibration modulation index. For the ASD-fed machines, the motor is powered by a PWM voltage waveform. The power supply for the ASD, denoted as  $f_{s,ASD}$ , primarily consists of  $n$  odd harmonics of the fundamental drive frequency  $f_d$ , and it can be expressed as:

$$f_{s,ASD} = nf_d \quad (10)$$

Therefore, the characteristic current frequency of the bearing faults in the ASD can be expressed as:

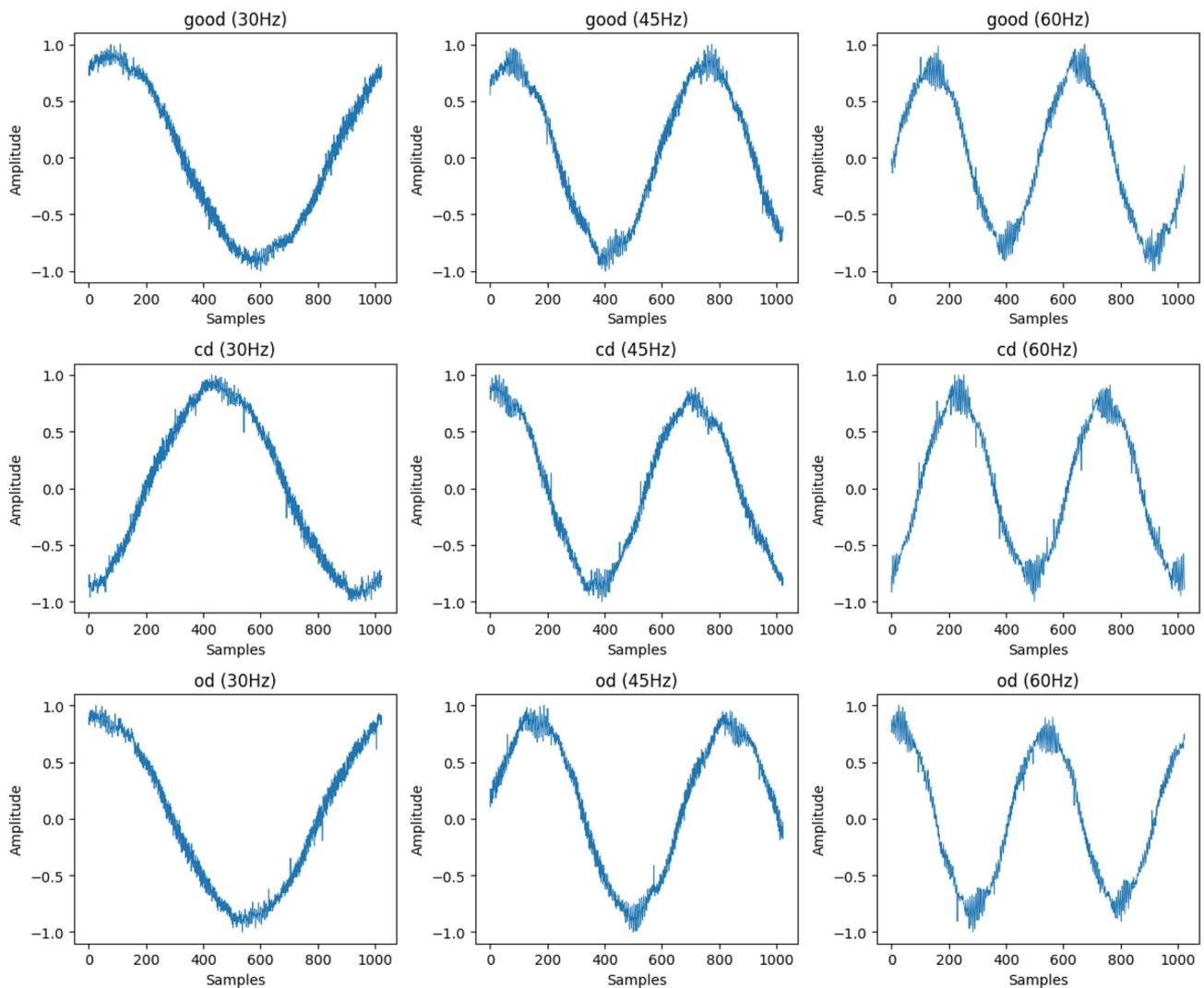
$$f_{CF,ASD} = |nf_d \pm kf_v| \quad (11)$$

where  $f_d$  is the fundamental supply frequency in Hertz,  $f_v$  is the vibration frequency in Hertz,  $n$  is the PWM harmonic index of the fundamental supply, and  $k$  is the vibration modulation index. Consequently, an ASD system with a bearing fault produces current frequency components with  $n$  times as many harmonics as those found in a line-driven motor with a similar bearing defect.

Our study explores the impact of varying the polynomial degree  $q$  in the 1D self-organized operational layers (Oper1D), testing  $q$  values of 1 (i.e., Conv1D), 3, and 5 to analyze how higher polynomial terms affect the model's ability to capture complex patterns in the motor current signature for each case.

The Self-ONN architecture is designed with specific training hyperparameters, including a batch size of 16 and 200 epochs, optimized using the Adam optimizer with a learning rate of 0.0005. The input to the network is the motor current waveform, with varying time-domain samples: 1024 for 30 Hz, 682 for 45 Hz, and 512 for 60 Hz input frequencies. Before being processed by the network, each frame undergoes a normalization process to scale the data within a consistent range, typically  $[-1, 1]$ . The architecture includes three operational layers, each consisting of a custom 1D self-operational layer, a hyperbolic tangent (tanh) activation function and a max-pooling layer. The max-pooling layer, tailored to the input frequency, uses a window of 2 data points for 60 Hz and a window of 4 data points for 30 Hz and 45 Hz, reducing data dimensionality while accommodating the differences in input size. After the final self-operational layer in the Self-ONN architecture, the data is flattened into a 1D vector and then processed by two dense layers. The first dense layer has 12 neurons and uses tanh activation function. The final output layer with 3 neurons employs a Softmax activation function, which converts the processed data into probabilities for three distinct classes: healthy, outer race defect, and cage defect. The proposed 1D Self-ONN model for 60 Hz input frequency is illustrated in Fig. 8.

For all input frequencies, we employed the sliding window method as a preprocessing step to augment the data for training. This technique involves segmenting the continuous motor current waveform into shorter, fixed-length segments, which are more manageable for subsequent processing and analysis. For 60 Hz input frequency, the frame length is set to 512 data points and the hop length is 256 points, meaning that we apply 50% overlapping. Post-segmentation, each frame undergoes a normalization process to scale the data within a consistent range, typically  $[-1, 1]$ . Each normalized frame is treated as an independent data sample for the neural network. Training and evaluation are conducted within a 5-fold stratified cross-validation setup to ensure that each fold is representative of the overall dataset, maintaining the proportion of each class label. This method enhances the generalizability of the model by validating it across multiple, independent



**Fig. 6** Normalized current waveforms of the ASD system for the given input frequencies

subsets of the data. Within each fold, the model trains on the training subset, iteratively adjusting weights and biases to minimize the cross-entropy loss. After training, the model’s performance is assessed using the validation subset specific to that fold. This process is repeated for each of the five folds, with the model being reinitialized each time to ensure the learning is specific to the data in the fold and not influenced by previous data. Finally, the results from all folds are aggregated to provide a robust estimate of the model’s performance across the entire dataset.

For the 60 Hz input frequency, the performance metrics of recall, precision and F1-score across different models are given in Table 3. These metrics evaluate the classifier’s ability to differentiate specific events from non-events. Precision represents the ratio of correctly identified events to all detected events; Recall indicates the proportion of correctly classified events among all events, while the F1-score signi-

fies the harmonic mean of the model’s Precision and Recall. These metrics are calculated based on the counts of false negatives (FN), false positives (FP), true negatives (TN), and true positives (TP) as follows:

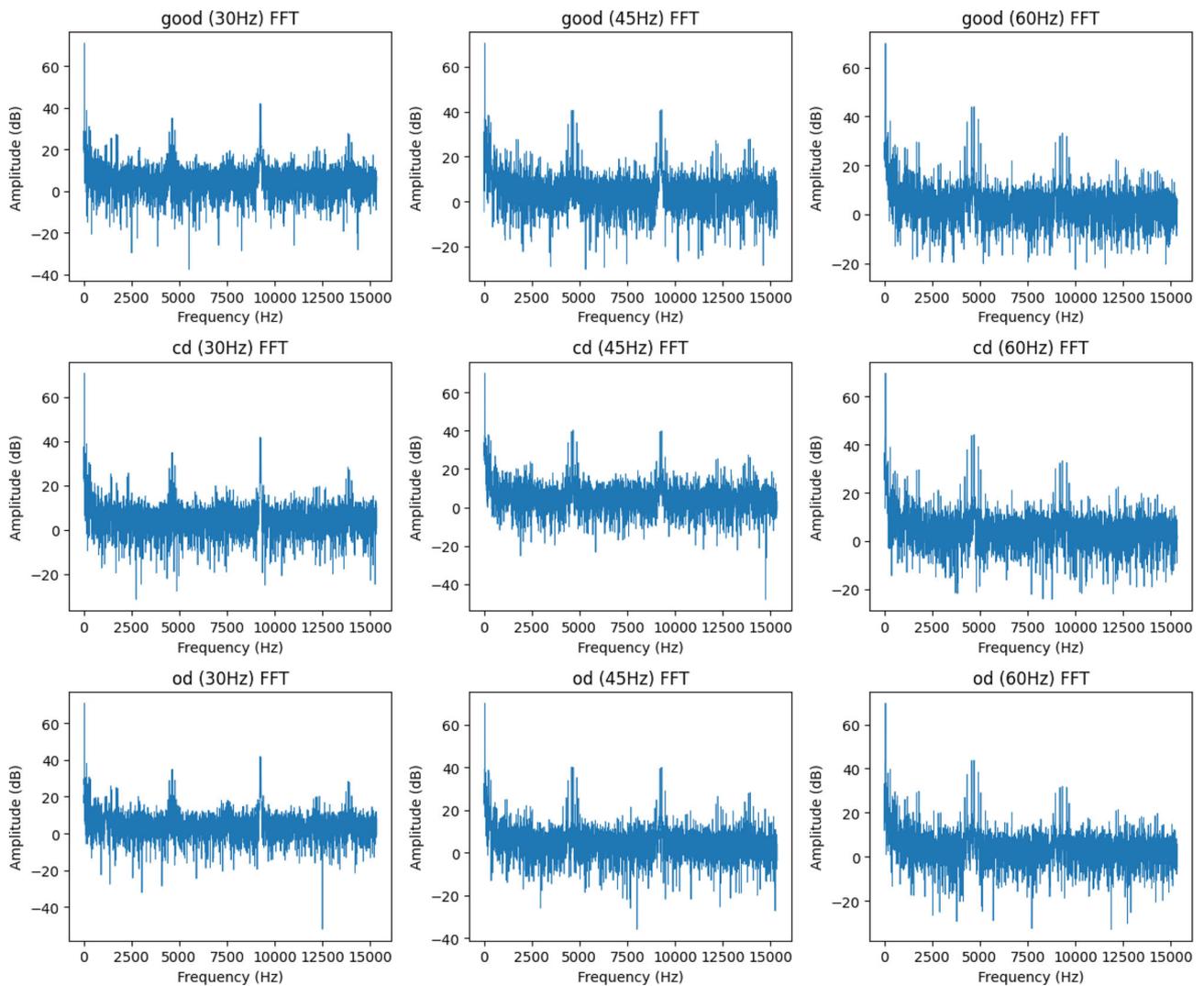
$$\text{Recall} = \frac{TP}{TP + FN} \tag{12}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{13}$$

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{14}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{15}$$

The confusion matrices for 1D CNN ( $q=1$ ) and Self-ONN ( $q=3$ ) are also provided in Table 4. 1D Self-ONN with a polynomial degree of  $q = 3$  exhibited superior performance, achieving precision, recall, and F1 scores of 0.90, 0.89, and



**Fig. 7** Frequency-domain (FFT) plots for each fault case and input frequency

**Table 1** Outer-race defect characteristic frequencies at different ASD system frequencies

System frequency (Hz)	Rotor speed (rpm)	$f_{OD}$ (Hz)
30	1797	95.8
45	2697	143.8
60	3596	191.8

**Table 2** Cage defect characteristic frequencies at different ASD system frequencies

System frequency (Hz)	Rotor speed (rpm)	$f_{CD}$ (Hz)
30	1797	12.0
45	2697	18.0
60	3596	24.0

0.89, respectively. Furthermore, the 1D CNN ( $q=1$ ) model achieved an overall accuracy of 87.31%, while the 1D Self-ONN ( $q=3$ ) achieved a higher accuracy of 89.46%. This underscores that a moderate level of model complexity, as seen with  $q = 3$ , is optimal for capturing the essential features in the motor current data without overfitting, making it the most effective model configuration in our analysis for 60 Hz case.

To compare 1D CNN and Self-ONN architectures with similar computational complexity, we introduce a variant of the 1D CNN model denoted as "1D CNN (\*2)" in Table 3. This modified CNN model features double the number of filters in its initial two convolutional layers and has a similar number of trainable parameters with 1D Self-ONN ( $q = 3$ ) to ensure a fair comparison. Despite this increase in complexity, the 1D CNN (\*2) model fails to surpass the classification performance achieved by the 1D Self-ONN with  $q = 3$ .

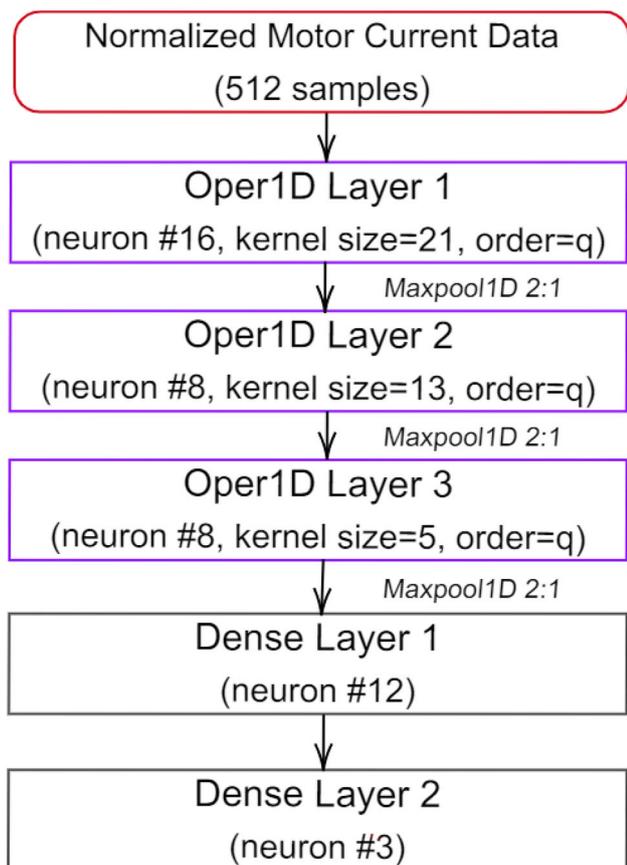


Fig. 8 1D Self-ONN architecture for 60 Hz input frequency

Table 3 Performance metrics of 1D CNN and Self-ONN models at 60 Hz input frequency

Model	Precision	Recall	F1-score
1D CNN	0.87	0.87	0.87
1D CNN (*2)	0.85	0.85	0.85
1D Self-ONN (q=3)	0.90	0.89	0.89
1D Self-ONN (q=5)	0.87	0.87	0.87

Table 4 Confusion matrices for 1D CNN (in parentheses) and 1D Self-ONN (q=3) for 60 Hz input frequency

Ground truth	Prediction		
	Healthy	CD	OD
Healthy	141 (131)	3 (1)	11 (23)
CD	4 (1)	148 (147)	3 (7)
OD	20 (19)	8 (8)	127 (128)

Table 5 Performance metrics of 1D CNN and Self-ONN models at 30 Hz input frequency

Model	Precision	Recall	F1-score
1D CNN	0.86	0.86	0.86
1D CNN (*2)	0.86	0.86	0.86
1D Self-ONN (q=3)	0.89	0.89	0.89
1D Self-ONN (q=5)	0.86	0.87	0.87

Table 6 Confusion matrices for 1D CNN (in parentheses) and 1D Self-ONN (q=3) for 30 Hz input frequency

Ground truth	Prediction		
	Healthy	CD	OD
Healthy	63 (62)	11 (12)	1 (1)
CD	9 (6)	64 (64)	2 (5)
OD	0 (1)	2 (7)	73 (67)

For the 30 Hz input frequency, the model undergoes training with the same configuration as previously described, but with a minor modification in the preprocessing step. This adjustment involves changing the frame size to 1024 and the hop length to 512 to accommodate the decreased input frequency.

For the 30 Hz input frequency, the same performance metrics for each model are detailed in Tables 5, and 6 shows the confusion matrices. At this frequency, the 1D CNN achieved an accuracy of 85.78%, while the 1D Self-ONN ( $q = 3$ ) performed better with an overall accuracy of 88.89%. The 1D Self-ONN model with  $q = 3$  again achieves the highest scores across precision, recall, and F1 score, each standing at 0.89. It is evident that the model’s performance is sensitive to the degree of the polynomial transformation applied within the Oper1D layers. Specifically, the model with  $q = 3$  consistently outperforms the other configurations, maintaining a higher accuracy on average across all folds.

For comparison, we again introduce a variant of the 1D CNN model denoted as "1D CNN (\*2)" with similar computational complexity as 1D Self-ONN ( $q = 3$ ) in Table 5. Even though we double the number of filters in the first two convolutional layers, the 1D CNN (\*2) model again fails to surpass the classification performance achieved by the 1D Self-ONN with  $q = 3$  for the 30 Hz input frequency.

**Table 7** Performance metrics of 1D CNN and Self-ONN models at 45 Hz input frequency

Model	Precision	Recall	F1-score
1D CNN	0.95	0.94	0.95
1D CNN (*2)	0.96	0.96	0.96
1D Self-ONN (q=3)	0.99	0.99	0.99
1D Self-ONN (q=5)	0.97	0.97	0.97

**Table 8** Confusion matrices for 1D CNN (in parentheses) and 1D Self-ONN (q=3) for 45 Hz input frequency

Ground truth	Prediction		
	Healthy	CD	OD
Healthy	113 (106)	2 (9)	0 (0)
CD	2 (5)	113 (109)	0 (1)
OD	0 (2)	0 (2)	115 (111)

Finally, for the 45 Hz input frequency, the model was trained with an adjusted input window size of 682 and a hop length of 341 samples to capture one full cycle of the signal. This configuration was applied to the same Self-ONN model architecture used for 30 Hz, with max-pooling window sizes of 4 data points in each operational layer, as before. The Self-ONN with  $q=3$  demonstrated exceptional performance, achieving average precision, recall, and F1-scores of 0.99 across all classes. In contrast, the 1D CNN models, including the enhanced version with doubled filters (1D CNN (\*2)) in the first two 1D convolutional layers, achieved a maximum F1-score of 0.96, which is lower than the performance of the Self-ONN. Additionally, from Table 8, we can see that the 1D CNN achieves an accuracy of 94.49%, whereas the 1D Self-ONN ( $q = 3$ ) achieves a notably higher accuracy of 98.84%. Overall, 1D Self-ONNs with  $q = 3$  have demonstrated potential for outperforming traditional 1D CNNs in terms of recall, precision and F1-score at all input frequencies for bearing fault diagnosis from motor current data.

Table 9 offers a comparison of the trainable parameter count and total MACs (Multiply-Accumulate operations) for each neural network employed in this study. To compute the trainable parameters, in the generative neuron, each kernel connection involves  $Q$  times the usual number of parameters. Thus, total number of trainable parameters for the  $k$ -th neuron in the  $l$ -th layer, represented as  $n_k^l$ , is computed using the formula below:

$$n_k^l = N_{l-1} \cdot K_k^l \cdot Q_k^l \quad (16)$$

In Eq. 16,  $N_{l-1}$  represents the neuron count in layer  $l - 1$ ,  $K_k^l$  denotes the kernel size utilized within the neuron, and  $Q_k^l$  indicates the chosen approximation order for this neuron. Regarding the number of Multiply-Accumulate operations,

it is worth noting that to generate a single element in the output  $\bar{x}_{ik}^l$ , the process requires  $K_k^l \times Q_k^l$  MAC operations for each output map  $y_i^{l-1}$  from the previous layer. Extending the logic, we arrive at the below generalization:

$$\text{MAC}_k^l = N_{l-1} \cdot |\bar{x}_{ik}^l| \cdot K_k^l \cdot Q_k^l \quad (17)$$

where the notation  $|\cdot|$  denotes the use of the cardinality operator. For ease of notation, the equation does not include the bias term and the computational expense related to Hadamard exponentiation.

From Tables 3, 5, 7, and 9, we conclude that the 1D Self-ONN ( $q = 3$ ) model, with comparable trainable parameters and MAC operations to the 1D CNN, outperforms it in accuracy, precision, and F1-score. Additionally, 1D Self-ONN layers can be efficiently parallelized on Graphics Processing Units (GPUs) [22].

To further highlight the computational advantage of our approach, we compared the computational complexity of our method with various signal processing-based techniques in Table 10. The number of basic operations required by each algorithm was evaluated, taking into account the varying lengths of data samples used. The results, summarized based on findings from studies [26] and [1], demonstrate the relative efficiency and effectiveness of our method.

We also evaluated the training durations and inference speeds for both the 1D CNN and Self-ONN models. All models were run on a single NVIDIA RTX 3070 GPU to ensure consistency. Inference speeds were computed by averaging over 100 runs, while the total training duration represents the cumulative time across all 5 folds. The results, detailed in Table 11, illustrate the trade-offs between training time and inference speed across different input frequencies.

At 60 Hz, 1D CNNs trained in 87 to 88 s with inference speeds of 738 to 758 microseconds, while 1D Self-ONNs required 150 to 213 s for training and had inference speeds of 903 to 1127 microseconds. At 45 Hz and 30 Hz, CNNs were also faster in both training and inference compared to Self-ONNs. Despite these differences, the performance gap in terms of training and inference duration between CNNs and Self-ONNs was relatively modest. Although Self-ONNs require longer training times, they offer significant advantages in parallel processing. By avoiding the need for deeper networks and utilizing higher-order activations within a Self-ONN layer, Self-ONNs can be run faster with proper parallel computation.

Table 12 presents a comparative analysis of the proposed method against other machine learning techniques from the literature for fault diagnosis in ASD-fed machines. In this table, OR, IR, BF, and CD stand for outer race fault, inner race fault, ball fault, and cage defect, respectively. The results presented here lead to several key conclusions. Firstly, vibration-based fault diagnosis in ASD-fed machines gener-

**Table 9** Computational complexity comparison

Input freq.	Model	Trainable parameters	Total MACs
60 Hz	1D CNN (q=1)	8,547	645,156
	1D CNN (*2) (q=1)	14,219	2,136,100
	1D Self-ONN (q=3)	13,187	1,923,108
	1D Self-ONN (q=5)	17,827	3,201,060
30 Hz	1D CNN (q=1)	3,939	792,100
	1D CNN (*2) (q=1)	9,611	2,434,596
	1D Self-ONN (q=3)	8,579	2,373,156
	1D Self-ONN (q=5)	13,219	3,954,212

**Table 10** Computational complexity comparison with signal processing-based techniques

Reference	Method	Operations (approx.)
[26]	Sub-nyquist strategy with reduced data length	$\sim 10^5$
This study	1D Self-ONN	$\sim 10^6$
[2]	Empirical demodulation and FFT	$\sim 10^6$
[24]	Spectrum synch	$\sim 10^7$
[12]	FFT	$\sim 10^7$
[23]	Spectral kurtosis	$\sim 10^8$
[1]	MODWT and image edge detection	$\sim 10^8$
[34]	Subspace spectral estimation	$\sim 10^{11}$

**Table 11** Total training duration and inference speed for different models and input frequencies

Input freq.	Model	Training duration (s)	Inference speed ( $\mu$ s)
60 Hz	1D CNN (q=1)	87	738
	1D CNN (*2) (q=1)	88	758
	1D Self-ONN (q=3)	150	903
	1D Self-ONN (q=5)	213	1127
45 Hz	1D CNN (q=1)	64	728
	1D CNN (*2) (q=1)	66	788
	1D Self-ONN (q=3)	113	868
	1D Self-ONN (q=5)	160	1087
30 Hz	1D CNN (q=1)	44	908
	1D CNN (*2) (q=1)	46	916
	1D Self-ONN (q=3)	78	1047
	1D Self-ONN (q=5)	112	1197

ally offers superior accuracy and earlier detection compared to the methods that use current data. However, the effectiveness of vibration sensors is heavily dependent on their mounting location, and they tend to be more expensive [13]. When comparing machine learning approaches that utilize current data, many demonstrate high diagnostic accuracy but often require extensive preprocessing. Additionally, these

methods frequently use test bearings with larger defects than those used in our study. The strength of our approach lies in its minimal preprocessing and lower computational complexity, making it suitable for direct implementation on embedded devices while maintaining competitive fault diagnosis accuracy.

**Table 12** Comparison of ML-based methods for fault diagnosis in ASD-fed machines

Reference	Data	Load (%)	Fault type	Fault severity (mm)	Accuracy (%)
[37]	Vibration	–	OR IR BF	–	98.26
[11]	Current	0 and 100	distributed	–	> 71
[9]	Vibration	Variable	OR IR BF	–	> 95
[18]	Current	0 and 100	OR	1.60	99.84
[1]	Current	0 and 100	OR BF Distributed	1.58	94.68–97.74
[31]	Current	Variable	OR	–	96.1
1D CNN	Current	0	OR CD	0.79	85.78 (30 Hz) 94.49 (45 Hz) 87.31 (60 Hz)
1D Self-ONN ( $q=3$ )	Current	0	OR CD	0.79	88.89 (30 Hz) 98.84 (45 Hz) 89.46 (60 Hz)

## 4 Conclusion

In the field of motor fault diagnostics, particularly pertaining to bearing anomalies, the analysis of motor current signatures has emerged as a vital tool for predictive maintenance and reliability assurance. Our research contributes to this domain by utilizing 1D Self-ONNs to improve the detection of bearing fault characteristics directly from raw motor current waveforms, while keeping the model complexity low. A key part of our approach is fine-tuning the polynomial degree  $q$  in the 1D operational layers of Self-ONNs. Empirical results indicate that a moderate complexity setting ( $q = 3$ ) in the Self-ONN surpasses the traditional 1D Convolutional Neural Network (CNN) in identifying bearing defects in ASD-fed machines, as reflected by superior precision, recall, and F1-score metrics across distinct input frequencies. At 30 Hz, the Self-ONN with  $q = 3$  achieved an average F1-score across all classes (healthy, outer race defect and cage defect) of 0.89, compared to 0.86 for the CNN. At 60 Hz, the Self-ONN with  $q = 3$  also outperformed the CNN with an average F1-score of 0.89, while the CNN's average F1-score was 0.87. Similarly, at 45 Hz, the Self-ONN with  $q = 3$  achieved a near-perfect average F1-score of 0.99, significantly surpassing the CNN's average F1-score of 0.95. These results clearly demonstrate that the Self-ONN with  $q = 3$  provides a consistently higher F1-score than the CNN across all tested frequencies, highlighting its enhanced capability in extracting relevant features from the current signatures. The implementation of such a model may facilitate a significant step forward in minimizing unscheduled down-

time and extending the lifespan of industrial machinery. In future work, we plan to apply a similar domain adaptive learning scheme as in [16] for bearing fault diagnosis using raw motor current data, aiming to enhance diagnostic accuracy and reliability under varying speeds of ASD-fed machines.

**Author Contributions** S.K. wrote the main manuscript text and validated the proposed method, while L.E. provided supervision and assisted with data curation. All authors reviewed the manuscript.

**Funding** Open access funding provided by Tampere University (including Tampere University Hospital).

**Data Availability** No datasets were generated or analysed during the current study.

## Declarations

**Conflict of interest** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Aviña-Corral V, de Jesus Rangel-Magdaleno J, Peregrina-Barreto H et al (2022) Bearing fault detection in asd-powered induction machine using modwt and image edge detection. *IEEE Access* 10:24181–24193
2. Batista FB, Lamim Filho PCM, Pederiva R et al (2016) An empirical demodulation for electrical fault detection in induction motors. *IEEE Trans Instrum Meas* 65(3):559–569
3. Benbouzid MEH (2000) A review of induction motors signature analysis as a medium for faults detection. *IEEE Trans Industr Electron* 47(5):984–993
4. Chen CC, Liu Z, Yang G et al (2020) An improved fault diagnosis using 1d-convolutional neural network model. *Electronics* 10(1):59
5. Chong UP et al (2011) Signal model-based fault detection and diagnosis for induction motors using features of vibration signal in two-dimension domain. *Strojniški vestnik* 57(9):655–666
6. Eren L, Devaney MJ (2003) Motor current analysis via wavelet transform with spectral post-processing for bearing fault detection. In: *Proceedings of the 20th IEEE instrumentation technology conference (Cat. No. 03CH37412)*, IEEE, pp 411–414
7. Eren L, Devaney MJ (2004) Bearing damage detection via wavelet packet decomposition of the stator current. *IEEE Trans Instrum Meas* 53(2):431–436
8. Eren L, Teotrakool K, Devaney MJ (2007) Bearing fault detection via wavelet packet decomposition with spectral post processing. In: *2007 IEEE instrumentation and measurement technology conference IMTC 2007*, IEEE, pp 1–4
9. Ewert P, Orłowska-Kowalska T, Jankowska K (2021) Effectiveness analysis of pmsm motor rolling bearing fault detectors based on vibration analysis and shallow neural networks. *Energies* 14(3):712
10. Filippetti F, Bellini A, Capolino GA (2013) Condition monitoring and diagnosis of rotor faults in induction machines: State of art and future perspectives. In: *2013 IEEE workshop on electrical machines design. IEEE, control and diagnosis (WEMDCD)*, pp 196–209
11. Fontes Godoy W, Morinigo-Sotelo D, Duque-Perez O et al (2020) Estimation of bearing fault severity in line-connected and inverter-fed three-phase induction motors. *Energies* 13(13):3481
12. Gyftakis KN, Antonino-Daviu JA, Garcia-Hernandez R et al (2015) Comparative experimental investigation of broken bar fault detectability in induction motors. *IEEE Trans Ind Appl* 52(2):1452–1459
13. Hoang DT, Kang HJ (2019) A motor current signal-based bearing fault diagnosis using deep learning and information fusion. *IEEE Trans Instrum Meas* 69(6):3325–3333
14. Ince T, Kiranyaz S, Eren L et al (2016) Real-time motor fault detection by 1-d convolutional neural networks. *IEEE Trans Industr Electron* 63(11):7067–7075
15. Ince T, Malik J, Devocioglu OC et al (2021) Early bearing fault diagnosis of rotating machinery by 1d self-organized operational neural networks. *Ieee Access* 9:139260–139270
16. Ince T, Kilickaya S, Eren L, et al (2022) Improved domain adaptation approach for bearing fault diagnosis. In: *IECON 2022—48th annual conference of the IEEE industrial electronics society, IEEE*, pp 1–6
17. Jia F, Lei Y, Lin J et al (2016) Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. *Mech Syst Signal Process* 72:303–315
18. Jiménez-Guarneros M, Morales-Perez C, de Jesus Rangel-Magdaleno J (2021) Diagnostic of combined mechanical and electrical faults in asd-powered induction motor using modwt and a lightweight 1-d cnn. *IEEE Trans Industr Inf* 18(7):4688–4697
19. Kılıçkaya S (2022) Microcontroller-based real-time motor bearing fault detection and diagnosis using 1d convolutional neural networks. Master's thesis, İzmir Ekonomi Üniversitesi
20. Kiranyaz S, Ince T, Iosifidis A et al (2017) Progressive operational perceptrons. *Neurocomputing* 224:142–154
21. Kiranyaz S, Ince T, Iosifidis A et al (2020) Operational neural networks. *Neural Comput Appl* 32(11):6645–6668
22. Kiranyaz S, Malik J, Abdallah HB et al (2021) Self-organized operational neural networks with generative neurons. *Neural Netw* 140:294–308
23. Leite VC, da Silva JGB, Veloso GFC et al (2014) Detection of localized bearing faults in induction machines by spectral kurtosis and envelope analysis of stator current. *IEEE Trans Industr Electron* 62(3):1855–1865
24. Li DZ, Wang W, Ismail F (2015) A spectrum synch technique for induction motor health condition monitoring. *IEEE Trans Energy Convers* 30(4):1348–1355
25. Malik J, Kiranyaz S, Gabbouj M (2021) Self-organized operational neural networks for severe image restoration problems. *Neural Netw* 135:201–211
26. Naha A, Samanta AK, Routray A et al (2017) Low complexity motor current signature analysis using sub-nyquist strategy with reduced data length. *IEEE Trans Instrum Meas* 66(12):3249–3259
27. Nandi S, Toliyat HA, Li X (2005) Condition monitoring and fault diagnosis of electrical motors—a review. *IEEE Trans Energy Convers* 20(4):719–729
28. Ozcan IH, Devocioglu OC, Ince T et al (2022) Enhanced bearing fault detection using multichannel, multilevel 1d cnn classifier. *Electr Eng* 104(2):435–447
29. Pons-Llinares J, Antonino-Daviu JA, Riera-Guasp M et al (2014) Advanced induction motor rotor fault diagnosis via continuous and discrete time-frequency tools. *IEEE Trans Industr Electron* 62(3):1791–1802
30. Qian L, Li B, Chen L (2022) Cnn-based feature fusion motor fault diagnosis. *Electronics* 11(17):2746
31. Skyllvik AJ, Robbersmyr KG, Van Khang H (2019) Data-driven fault diagnosis of induction motors using a stacked autoencoder network. In: *2019 22nd International Conference on Electrical Machines and Systems (ICEMS)*, IEEE, pp 1–6
32. Teotrakool K, Devaney MJ, Eren L (2008) Bearing fault detection in adjustable speed drives via a support vector machine with feature selection using a genetic algorithm. In: *2008 IEEE Instrumentation and Measurement Technology Conference*, IEEE, pp 1129–1133
33. Teotrakool K, Devaney MJ, Eren L (2009) Adjustable-speed drive bearing-fault detection via wavelet packet decomposition. *IEEE Trans Instrum Meas* 58(8):2747–2754
34. Trachi Y, Elbouchikhi E, Choqueuse V et al (2016) Induction machines fault detection based on subspace spectral estimation. *IEEE Trans Industr Electron* 63(9):5641–5651
35. Wan H, Gu X, Yang S et al (2023) A sound and vibration fusion method for fault diagnosis of rolling bearings under speed-varying conditions. *Sensors* 23(6):3130
36. Ye X, Li G (2024) An intelligent fault diagnosis method for rolling bearing using motor stator current signals. *Meas Sci Technol* 35(8):086131
37. Youcef Khodja A, Guersi N, Saadi MN et al (2020) Rolling element bearing fault diagnosis for rotating machinery using vibration spectrum imaging and convolutional neural networks. *Int J Adv Manuf Technol* 106:1737–1751