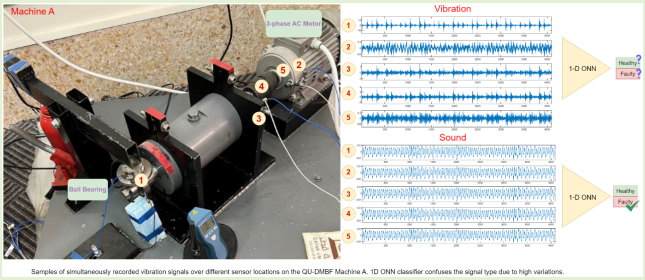
Exploring Sound Versus Vibration for Robust Fault Detection on Rotating Machinery

Serkan Kiranyaz, *Senior Member, IEEE*, Ozer Can Devecioglu, Amir Alhams, Sadok Sassi, Turker Ince, *Member, IEEE*, Onur Avci, and Moncef Gabbouj, *Fellow, IEEE*

***Abstract*—Robust and real-time detection of faults has become an ultimate objective for predictive maintenance on rotating machinery. Vibration-based deep learning (DL) methodologies have become the de facto standard for bearing fault detection as they can produce state-of-the-art detection performances under certain conditions. Despite such particular focus on the vibration signal, the utilization of sound, on the other hand, has been widely neglected. As a result, no large-scale benchmark motor fault dataset exists with both sound and vibration data. The novel and significant contributions of this study can be summarized as follows. This study presents and publically shares the Qatar University dual-machine bearing fault benchmark dataset (QU-DMBF), which encapsulates sound and vibration data from two different motors operating under 1080 working conditions. Then, we focus on the major limitations and drawbacks of vibration-based fault detection due to numerous installation and operational conditions. Finally, we propose the first DL approach for sound-based fault detection and perform comparative evaluations between the sound and vibration signals over the QU-DMBF dataset. A wide range of experimental results shows that the sound-based fault detection method is significantly more robust than its vibration-based counterpart, as it is entirely independent of the sensor location, cost-effective (requiring no sensor and sensor maintenance), and can achieve the same level of the best detection performance by its vibration-based counterpart. This study publicly shares the QU-DMBF dataset, the optimized source codes in PyTorch, and comparative evaluations with the research community.**



Samples of simultaneously recorded vibration signals over different sensor locations on the QU-DMBF Machine A. 1D ONN classifier confuses the signal type due to high variations.

***Index Terms*— Bearing fault detection, machine health monitoring, operational neural networks (ONNs).**

NOMENCLATURE

Symbol	Description
Q	Order of the Maclaurin polynomial.
$Kx \times Ky$	Size of a kernel, i.e., $\boldsymbol{w_{ik}^{l+1}}$.
$\boldsymbol{w_{ik}^{l+1}}(r,t)$	Q-dimensional array of the kernel element (r,t) from the ith neuron in layer $l+1$ to k^{h} neuron in layer l.
$w_{ik}^{l+1}(r,t,q)$	qth element of $\boldsymbol{w_{ik}^{l}}(r,t)$.
(y_k^l)	2-D output feature map of the kth neuron at the lth layer.
$\Psi\left(y_k^l, \boldsymbol{w_{ik}^{l+1}}(r,t)\right)$	Nodal operator function approximated the Maclaurin series over the elements (coefficients) of $\boldsymbol{w_{ik}^{l}}(r,t)$.
$P_i^{l+1} = \Sigma$	Pooling operator is fixed to summation for this study.
Δ_i^{l+1}	Given the cost function, E, the 2-D delta error map of the input feature map, $x_i^{l+1}(m,n)$ at layer $l+1$. Specifically, $\Delta_{\text{i}}^{l+1}(m,n) = (\partial E/\partial x_i^{l+1}(m,n))$.
Δy_k^l	Given the cost function, E, the 2-D sensitivity of the output map, $y_k^l(m,n)$. Specifically, $\Delta y_k^l(m,n) = (\partial E/\partial y_k^l(m,n))$.

Manuscript received 12 April 2024; accepted 10 May 2024. Date of publication 3 June 2024; date of current version 16 July 2024. This work was supported by Business Finland/NSF CBL project AMALIA. The associate editor coordinating the review of this article and approving it for publication was Dr. Pedro Oliveira Conceição Junior. *(Corresponding author: Moncef Gabbouj.)*

Serkan Kiranyaz is with the Department of Electrical Engineering, Qatar University, Doha 2713, Qatar (e-mail: mkiranyaz@qu.edu.qa).

Ozer Can Devecioglu and Moncef Gabbouj are with the Department of Computing Sciences, Tampere University, 33100 Tampere, Finland (e-mail: ozer.devecioglu@tuni.fi; moncef.gabbouj@tuni.fi).

Amir Alhams and Sadok Sassi are with the Department of Mechanical Engineering, Qatar University, Doha 2713, Qatar (e-mail: aa1702913@qu.edu.qa; sadok.sassi@qu.edu.qa).

Turker Ince is with the Department of Media Engineering and Technology, German International University, 10829 Berlin, Germany (e-mail: turker.ince@giu-berlin.de).

Onur Avci is with the Department of Civil and Environmental Engineering, West Virginia University, Morgantown, WV 26506 USA (e-mail: onur.avci@mail.wvu.edu).

Digital Object Identifier 10.1109/JSEN.2024.3405889

I. INTRODUCTION

ACCURATE and instantaneous fault detection for rotating machinery is crucial for motor health monitoring in several industries, such as mass production lines, manufacturing, aerospace, and energy. The most critical components

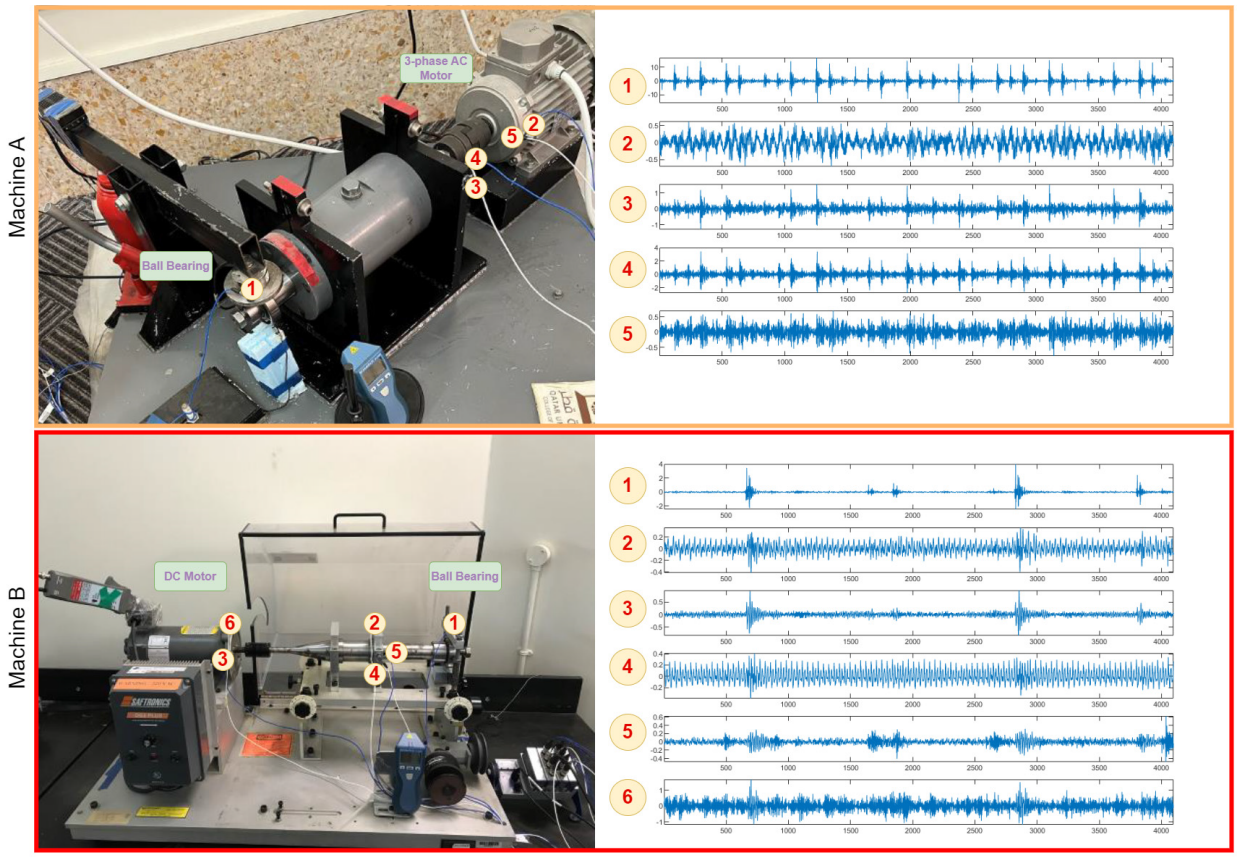


Fig. 1. Samples of *simultaneously* recorded vibration signals over different sensor locations on the QU-DMBF [62].

of rotating machinery are its bearings, as they tend to fail in time, which will eventually cause unexpected downtime, high maintenance costs, and even catastrophic accidents if not detected and prevented in advance.

Numerous methodologies have focused on the vibration signals to detect and identify the bearing faults. They can be classified as model-based methods [1], [2], [3], [4], signal-processing approaches [5], [6], [7], [8], [9], [10], [11], [12], conventional machine learning (ML), and recent deep learning (DL) methods [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28]. Especially during the last decade, DL-based methods based on the vibration signal have increased tremendously. This is an expected outcome since the vibration signal can quickly reveal the track changes over the mechanical behavior of a bearing. Thus, it has become a de facto standard in this domain. However, acquiring a reliable and high-quality vibration signal requires good-quality sensors, which further demands periodic maintenance. This may induce significant costs and pose particular risks of malfunctioning over time. A fundamental problem is that the vibration signal is susceptible to the sensor location on the machinery. Take, for example, sample vibration signals acquired from the two machines of the Qatar University dual-machine bearing fault benchmark dataset (QU-DMBF) dataset [28], as shown in Fig. 1. Despite some of the sensors being in close vicinity, the signals from them are entirely different from one another, e.g., see the signals of sensors #2 and #5 from Machine A and sensors #6 and #3 as well as #2, #5, and #4 from Machine B. Therefore, any slight mispositioning of the sensor may result in poor detection performance by the DL method. Moreover, installing such wired sensorial equipment to nearby locations of rotating bearings will lead to specific challenges and operational drawbacks. Finally, vibration signals can easily get corrupted by several background noises, such as ambient vibrations, electrical interference of the motor, and sensor variations. Most aforementioned studies ignored such variations and evaluated their proposed method assuming only one or few working conditions over the early benchmark datasets with limited vibration data and a fixed sensor location. In particular, assuming sufficient fault data for all working conditions to train the fault detector may not be feasible in practice.

Conversely, the sound signal has none of these setbacks. For instance, it does not require an additional sensor. Thus, there is no need for any maintenance and installation since a mobile phone can easily acquire sound data from any location, which does not affect the sound signal pattern. Moreover, it is immune from any electrical or sensorial noise. Despite such crucial advantages, during the last two decades, only a few sound-based fault detection methods have been proposed in the literature [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], which utilize conventional ML methods such as K-nearest neighbor (KNN)

and support vector machines (SVM) classifiers over the manually extracted/selected features. Two possible reasons can be listed: 1) the lack of a benchmark dataset providing a large volume of both vibration and sound data over several working conditions for different motors and sensor locations and 2) an early study [36] has reported certain limitations of sound signals for identifying defects in gearboxes. It was concluded that the application of sound to gear defect detection needs to be improved, particularly for fault identification. This may have discouraged the researchers from pursuing this direction and drawn the focus instead on the vibration-based approaches. As a result, only a few DL-based methods over the sound signal have ever been proposed [37], [38], [39], and no comparative evaluations against the vibration counterpart have yet to be reported. Among them, [37] is not a DL-based method. The authors employed a stacked sparse autoencoder to extract the fault features first, and a simple softmax regression was adopted for fault classification. This is also true for [38], where a simple LSTM network has been trained over only 100 sound samples and tested over 90. With such a minimal dataset, it is obviously not reasonable to train a deep network. Finally, [39] is the only DL method, but it was again tested on a minimal dataset. The authors used 80% of the data randomly selected for training and the rest for testing. This not only biases the test results but is also infeasible in practice.

To address all the aforementioned drawbacks, this study proposes the first DL-based fault detection method over the raw sound signal and performs a wide range of experiments for comparative evaluations against its vibration-based counterpart. The QU-DMBF dataset compiled for this purpose is the most extensive sound/vibration dataset, with 13.5 h of data acquired from two motors (DC and three-phase AC) with 1080 working conditions. For simplicity and unbiased comparisons, we do not use any domain adaptation (DA) methodology [27], [28], [32], and hence, for both signal types, we evaluate their performance over the "unseen" working conditions using a single network model trained over the same data and with an identical experimental setup. In this way, we can test and compare their *robustness* for detecting the "unseen" fault cases, especially when the sensor locations and fault severities for testing differ from the ones used for training.

Derived from generalized operational perceptrons (GOPs) [48], [49], [50] and native operational neural networks (ONNs) [47], [48], self-organized ONNs (Self-ONNs) are the latest heterogeneous nonlinear network models with generative [44], [45], [46] and super neurons [58] that can perform optimal nonlinear operations for each kernel element. Therefore, in many mechanical faults [25], [26], [27], [28], [29] or health/structural anomaly detection [51], [52], [53], [54] tasks, along with the numerous data mining and pattern recognition problems [55], [56], [57], [58], [59], [60], [61], they have outperformed their linear counterparts, the conventional CNN models, significantly even with reduced network complexity and depth. In particular, the studies [25], [26], [27], [28] proposed state-of-the-art fault detection methodologies on rotating machinery using 1-D Self-ONNs. In this study, we aim to leverage this superiority further to achieve an elegant computational efficiency for sound-based fault detection. So, a seven-layer 1-D Self-ONN with only 80 generative neurons is used for both sound and vibration-based fault detection, and this demonstrates that even with such a compact and lightweight network model, robust and highly accurate sound-based fault detection performance can be achieved.

The rest of the article is organized as follows. Section II presents the exploration methodology with the proposed network model and the QU-DMBF benchmark dataset. Section III presents the fault detection results with detailed comparative evaluations over the QU-DMBF dataset. Finally, Section IV concludes the article and summarizes this article's novel and significant contributions.

II. Exploration Methodology and QU-DMBF Dataset

A. 1-D Self-Organized ONNs

Self-ONNs [26], [28], [51], [52], [53], [54], [55], [56], [57], [58], [59], [60], [61] were proposed as the superset of the CNNs with a controllable parameter, Q, which determines the level of the nonlinearity (degree of the polynomials) of each kernel transformation. When $Q = 1$ for all neurons in the network, a Self-ONN will reduce to a CNN. Due to the "on-the-fly" generation of the nonlinear nodal operator, the network can create the best possible basis functions to achieve the highest learning performance. So, with such optimized nonlinearity and heterogeneity, a Self-ONN can easily surpass an equivalent or even a significantly deeper and more complex CNN.

Each generative neuron can have an arbitrary nodal function, Ψ, for each kernel element of each connection. This incredible flexibility permits the formation of any nodal operator function. For the formation of "any" nodal function, we use the Taylor series approximation as stated in

$$\Psi(\mathbf{w}, \mathbf{y}) = w_0 + w_1 y + w_2 y^2 + \cdots + w_Q y^Q. \quad (1)$$

Any other function approximation technique (e.g., Fourier series or DWT) can also be used, but such techniques will also require the computation of certain basis (nonlinear) computationally demanding functions. However, as formulated in [44] and [45], the generative neurons can be turned into purely a set of convolutions, yielding a great computational advantage. Let $x_{ik}^l \in \mathbb{R}^M$ be the contribution of the ith neuron at the $(l-1)^{\text{th}}$ layer to the input map of the l^{th} layer. Therefore, it can be expressed as

$$\tilde{x}_{ik}^l(m) = \sum_{r=0}^{K-1} \sum_{q=1}^{Q} w_{ik}^{l(Q)}(r, q) \left(y_i^{l-1}(m+r)\right)^q \quad (2)$$

where $y_i^{l-1} \in \mathbb{R}^M$ is the output map of the ith neuron at the $(l-1)^{\text{th}}$ layer, $w_{ik}^{l(Q)}$ is a learnable kernel of the network, which is a $K \times Q$ matrix, i.e., $w_{ik}^{l(Q)} \in \mathbb{R}^{K \times Q}$, formed as, $w_{ik}^{l(Q)}(r) = [w_{ik}^{l(Q)}(r, 1), w_{ik}^{l(Q)}(r, 2), \ldots, w_{ik}^{l(Q)}(Q)]$. By the commutativity of the summation operations in (2), one can alternatively express

$$\tilde{x}_{ik}^l(m) = \sum_{q=1}^{Q} \sum_{r=0}^{K-1} w_{ik}^{l(Q)}(r, q-1)\, y_i^{l-1}(m+r)^q. \quad (3)$$

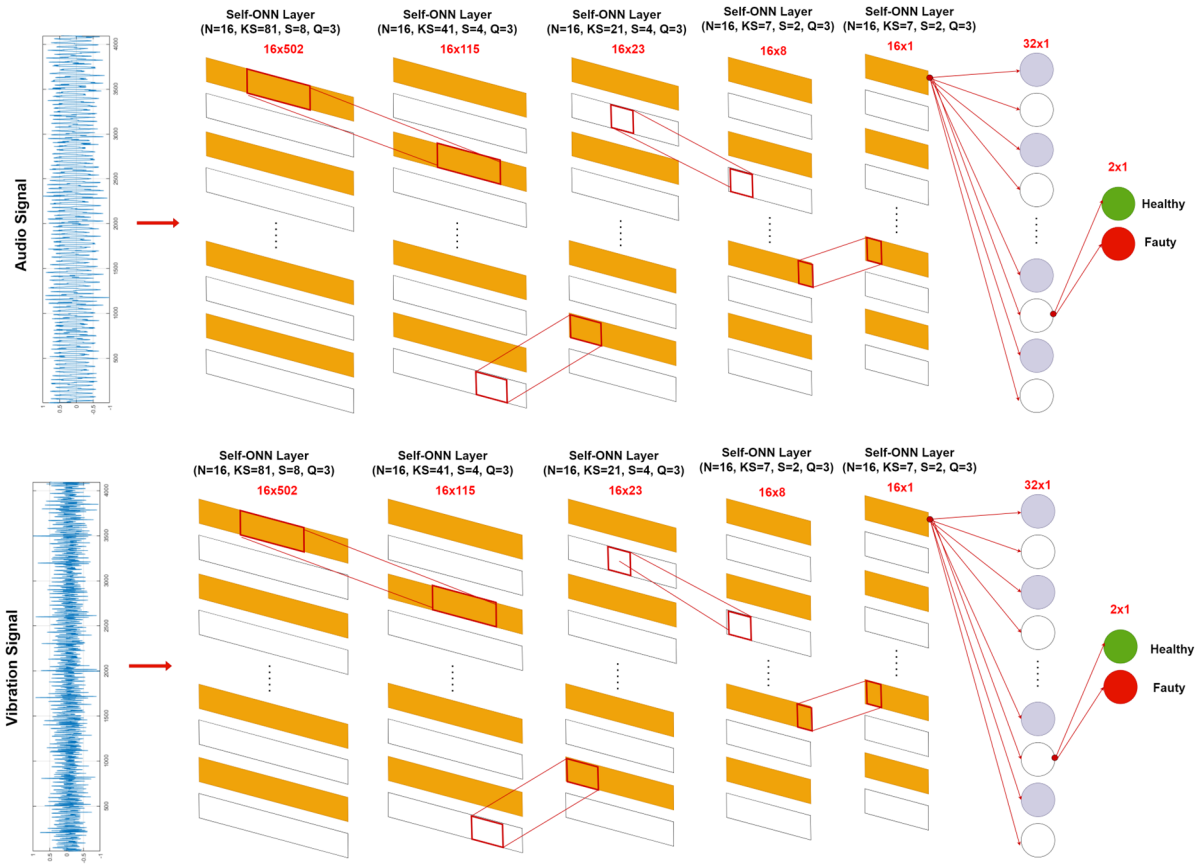


Fig. 2. Identical seven-layer Self-ONN models for detecting sound (top) and vibration-based (bottom) fault. The number of neurons (N), 1-D kernel size (KS), stride (S), and Q values are also presented for each layer.

One can simplify this expression as follows:

$$\tilde{x_{ik}^{l}} = \sum_{q=1}^{Q} \text{Conv1D}\left(w_{ik}^{l(Q)}, \left(y_i^{l-1}\right)^q\right). \tag{4}$$

Equation (4) simply shows that the kernel operations can indeed be executed by performing a Q number of 1-D convolutions. Finally, the input map of this neuron can be composed as

$$x_k^l = b_k^l + \sum_{i=0}^{N_{l-1}} x_{ik}^l \tag{5}$$

where b_k^l is the bias associated with this neuron. Passing the input map through the activation function will generate the output map, $y_k^l = f(x_k^l)$, which will then contribute to the input maps of the neurons on the next layer and so on. For a parallel processing implementation, in [44], [46], and [52], the raw-vectorized formulations of both forward propagation and back-propagation (BP) are presented.

To gain an understanding of how various hyperparameters impact the models' performance, we have run preliminary experiments using a variety of model configurations. This involves evaluating multiple training hyperparameters and network models and assessing their performances across the validation set. The best network configuration is used to evaluate fault detection performances over sound and vibration datasets. Fig. 2 shows the best 1-D Self-ONN sound and vibration-based fault detection models. Each model contains five operational layers with a total of 80 generative neurons, one dense layer with 32 generative perceptrons, and an output layer with two generative perceptrons. For each layer, $Q = 3$ and tanh activation functions are used. The kernel sizes at the consecutive operational layers are set as 81, 41, 21, 7, and 7, respectively.

Both vibration and sound data are framed into one-second segments without any overlapping. The sampling frequency of both signals is 4096 Hz; thus, each segment has $m = 4096$ samples. For a strict magnitude invariance during the training (BP) and forward propagation on the 1-D Self-ONN, each segment is further normalized as follows:

$$X_N(i) = \frac{2(X(\text{i}) - X_{\min})}{X_{\max} - X_{\min}} - 1 \tag{6}$$

where $X(i)$ is the ith original sample amplitude in the segment, $X_N(i)$ is the ith sample amplitude of the normalized segment, and $X_{\min}$ and $X_{\max}$ are the minimum and maximum amplitudes within the segment, respectively. This will scale the segment linearly in the range of $[-1\ 1]$, where $X_{\min} \to -1$ and $X_{\max} \to 1$. The proposed 1-D Self-ONN model receives the normalized segment samples as input channel and produces the binary output (class) vector.

Nomenclature section presents the formula abbreviations and mathematical symbols used in this article. The details of

the BP training for Self-ONNs[1] can be obtained from [44], [45], and [46].

B. Qatar University Dual-Machine Bearing Fault Benchmark Dataset

Researchers at Qatar University created the benchmark dataset for this study using two distinct electric machines (Machine A and Machine B). Fig. 1 depicts the experimental setup, including the positioning of the sensors and the installation of two machines. A three-phase AC motor, two double-row bearings, and a shaft with a maximum speed of 2840 r/min make up Machine-A's configuration. Two different loads (0.18 and 0.23 kN) were applied using an arm that is attached to a hydraulic jack from one end and to a plate from the other end. This plate is used to apply the load on the outer race of the bearing. Mounted on the bearing housing are PCB accelerometers (352C33 high-sensitivity Quartz ICP). The machine measures 100 × 100 × 40 cm and weighs 180 kg. The faults are deliberately inserted in the raceways of the outer and inner rings by electric discharge machining (EDM). It is a metal fabrication process that removes material from a workpiece by using electrical discharges between two electrodes without physical contact, offering high precision for conductive materials.

The following is the outline of Machine-A's working conditions.

1) There are 19 distinct bearing configurations: one healthy case and 18 faulty cases that consist of nine outer ring defects and nine inner ring defects. The range of defect sizes varies from 0.35 to 2.35 mm.
2) There are five distinct locations for accelerometers: two radial directions and three horizontal locations.
3) There are two distinct force levels: 0.18 and 0.23 kN.
4) Three distinct RPM ranges: 480, 680, and 1010 r/min.

For a healthy bearing, we recorded data for 270 s of operation, and for a damaged bearing case, we recorded 30 s; $30 \times 18 \times 5 \times 2 \times 3 = 16\,200$ s (4.5 h) of defect data, and $270 \times 5 \times 2 \times 3 = 8100$ s (2.25 h) of healthy data are the totals that come from this.

Machine B, on the other hand, has a DC motor, two single-row bearings, and a shaft that rotates at a varying speed of around 2000 r/min (max. 2500 r/min). A constant load of 0.18 kN was applied to the bearing by tightening two bolts at the two ends of a rectangular thick plate to let it press on the outer race of the bearing. PCB accelerometers are mounted on the bearing housing (353B33 high-sensitivity Quartz ICP). The machine is 59 kg in weight and has dimensions of 100 × 63 × 53 cm. The following are the different working conditions for Machine B.

1) There are nine bearing configurations with an outer ring defect, nine with an inner ring defect, and one healthy bearing configuration. The range of fault sizes varies from 0.35 to 2.35 mm.
2) There are six distinct locations for accelerometers.
3) A constant force (load) of 0.18 kN.
4) Five distinct RPM ranges: 240, 360, 480, 700, and 1020.

For each working condition of a healthy bearing, 270 s of vibration/sound data are recorded. Thus, the total duration of the vibration data for a healthy bearing is $270 \times 6 \times 1 \times 5 = 8100$ s (2.25 h). Similarly, 30 s of sound and vibration data are recorded for every working condition of a faulty bearing. The faulty to healthy data ratio is 2:1, and the total time is $30 \times 18 \times 6 \times 1 \times 5 = 16\,200$ s (4.5 h). Consequently, the entire duration of the dataset on machine B is 24 300 s (6.75 h). For all working conditions in both machines, the sound was recorded simultaneously with the same sampling frequency as the vibration data.

C. Experimental Setup

A recent study [28] has shown that the most reliable vibration data for fault detection is acquired from the closest accelerometer to the bearings, i.e., sensor #1 for both machines, as shown in Fig. 1. So, we have selected a part of the data of this accelerometer for training both fault detectors and use the rest of the data and the data of other sensors for testing. In particular, the sound and vibration data of sensor #1 acquired from the two most minor defects (0.35 and 0.5 mm) are used for training, corresponding to 124 fault segments. This corresponds to only 4.44% and 3.37% of the Machine-A and Machine-B fault data used for training, respectively. The exact number of healthy segments is also used to compose an isolated training data partition to evaluate the robustness of both fault detectors against the variations in sensor locations, fault severities, speed, and load over both machines.

The Adam optimizer is used for the BP training with the initial learning factor, $\varepsilon = 10^{-4}$, and the mean-squared error (MSE) as the loss function. Twenty percent of the training data is spared as validation to select the best Self-ONN model for testing. We implemented both fault detector networks using the FastONN library [46] based on PyTorch.

Commonly used performance metrics, precision (P), recall (R), $F1$-score ($F1$), and accuracy (Acc) are computed to evaluate the fault detection performances. The calculation of true positives (TP), false negatives (FN), and false positives (FP) are obtained per vibration/sound segment classification in the test set. Accordingly, these performance metrics can be expressed as follows:

$$P = \frac{\text{TP}}{\text{TP}+\text{FP}}, \quad R = \frac{\text{TP}}{\text{TP}+\text{FN}}$$
$$F1 = \frac{2PR}{P+R}, \quad \text{Acc} = \frac{\text{TP}+\text{TN}}{\text{TP}+\text{TN}+\text{FP}+\text{FN}}. \tag{7}$$

III. Experimental Results

In Section III-A, we shall perform an extended set of comparative evaluations between sound and vibration-based fault detection. In Section III-B, the computational complexity analysis of the proposed network model will be examined in-depth.

A. Results

Once the fault detectors of both sound and vibration data are trained over the training data (a fraction of healthy and

[1]The optimized PyTorch implementation of Self-ONNs with generative and super neurons is publicly shared in https://github.com/MertDuman/sonn

TABLE I
CONFUSION MATRICES FOR AUDIO (LEFT) AND VIBRATION (RIGHT) DATASETS FOR MACHINE-A USING 1-D SELF-ONNS

		Output	
TRUTH		Healthy	Faulty
	Healthy	942	6
	Faulty	16	2582

		Output	
TRUTH		Healthy	Faulty
	Healthy	942	6
	Faulty	3	2595

TABLE II
CONFUSION MATRICES FOR AUDIO (LEFT) AND VIBRATION (RIGHT) DATASETS FOR MACHINE-B USING 1-D SELF-ONNS

		Output	
TRUTH		Healthy	Faulty
	Healthy	746	44
	Faulty	42	2102

		Output	
TRUTH		Healthy	Faulty
	Healthy	786	4
	Faulty	90	2054

TABLE III
FAULT DETECTION PERFORMANCES PER SENSOR OVER SOUND AND VIBRATION DATA FOR MACHINE-A

Test Sensor #	Accuracy	Precision	Recall	F1-Score
Sound @ Sensors 1-5	99.38	**99.77**	99.38	99.58
Vibration @ Sensor #1	**99.75**	**99.77**	**99.88**	**99.83**
Vibration @ Sensor #2	62.15	85.93	57.81	69.12
Vibration @ Sensor #3	68.53	99.40	57.39	72.77
Vibration @ Sensor #4	89.34	93.74	91.57	92.64
Vibration @ Sensor #5	48.36	99.61	29.64	45.68

TABLE IV
FAULT DETECTION PERFORMANCES PER SENSOR OVER SOUND AND VIBRATION DATA FOR MACHINE-B

Test Sensor #	Accuracy	Precision	Recall	F1-Score
Sound @ Sensors 1-6	**97.07**	97.95	**98.04**	**98.00**
Vibration @ Sensor #1	96.80	**99.81**	95.80	97.76
Vibration @ Sensor #2	58.73	99.89	43.56	60.67
Vibration @ Sensor #3	74.16	99.86	64.74	78.55
Vibration @ Sensor #4	53.31	99.49	36.29	53.18
Vibration @ Sensor #5	56.85	99.66	41.09	58.19
Vibration @ Sensor #6	44.96	72.63	39.60	51.25

faulty data acquired by sensor #1), they are tested over the test data of each machine in the QU-DMBF dataset. The confusion matrices for machines A and B are presented in Tables I and II, respectively. Accordingly, the average fault detection performances over each sensor are individually computed using the aforementioned standard metrics and presented in Tables III and IV, respectively.

As mentioned earlier, the studies [25], [26], [27], [28] have already demonstrated the superiority of the Self-ONNs over the conventional CNNs. In this study, to perform comparative evaluations between the two network models for audio and vibration-based fault detection, we used a CNN model with the same configuration as the proposed 1-D Self-ONN, trained and tested over identical data partitions. Fig. 3 shows the

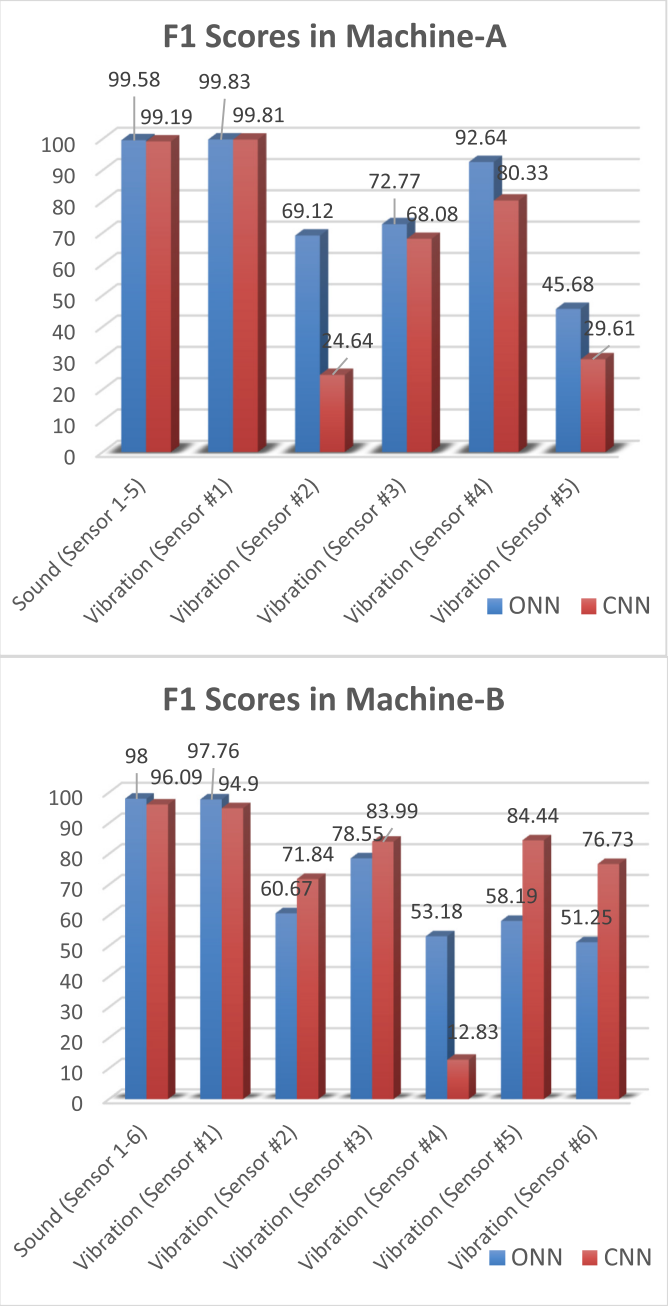


Fig. 3. Over the audio and vibration (Sensors #1–#6) data partitions of QU-DMBF dataset, $F1$ scores achieved by Self-ONNs (blue) and CNNs (red) in Machine-A (top) and Machine-B (bottom).

average $F1$ plots over the audio and vibration (per sensor) data partitions of the QU-DMBF dataset achieved by Self-ONNs (blue) and CNNs (red) in Machine-A (top) and Machine-B (bottom).

B. Computational Complexity Analysis

This section presents the computational complexity analysis by measuring the inference time, network size, and total number of parameters (PARs) of the proposed Self-ONN configuration. Comprehensive formulations of the PARs for Self-ONNs can be found in [46]. All trials used a 2.2 GHz Intel Core i7 PC equipped with an NVIDIA GeForce RTX 3080 graphics card and 16 GB of RAM. PyTorch[2] and the FastONN library [46] were utilized for the fault detectors. The proposed seven-layer Self-ONN model contains 377K parameters in total. For a single CPU implementation, the FP time for classifying a 1-s segment (sound or vibration) is 4 ms. With a single CPU, this demonstrates that the proposed sound-based fault detector can work 250 times faster than the real-time requirements. This indicates that the proposed approach can be used for a real-time motor health monitoring implementation, even as a typical application on a mobile phone. This will not only void the need for a sensor, but it also enables instantaneous health monitoring capabilities for any rotating machinery in any industry.

IV. Discussion

Several important observations can be made regarding the fault detection results presented in Nomenclature section and Table I. First, sound and vibration-based fault detectors obtain almost identical fault detection performances, but only when the vibration detector is tested on the same sensor data used for its training (sensor #1). In this case, the proposed 1-D Self-ONN model has achieved remarkable performance levels between 97.6% and 99.83% $F1$ scores on machines A and B, respectively. However, the performance of vibration-based fault detection significantly deteriorates when the sensor location is altered, even for a slight change. Performance levels around 50% (highlighted in red in both tables) can be observed on both machines, indicating a detection failure since 50% accuracy (or precision, recall, or $F1$) is a bottom-line performance level for a binary classification problem. It is worth noting that there is a significant performance gap even over the results obtained from the data of the sensors, which are pretty close to each other, e.g., see the accuracies obtained for sensor pairs #2–#5 and #3–#4 on Machine A. Such high variations from nearby sensors may also indicate that the mounting of the sensors to the surface of the machines differ from each other, and this, in turn, alters the acquired vibration signals significantly, as witnessed in Fig. 1. Besides the challenges of installing accelerometers over rotating machinery, this further indicates how sensitive fault detection can be concerning the sensor locations and mounting. On the other hand, sound-based detection has none of such drawbacks, sensitivities, or limitations, as the results on both machines show that it can consistently achieve a high detection performance with such a lightweight network model trained over a minority of the fault data (<5% of the fault data).

Finally, in accordance with the findings in the studies [25], [26], [27], [28], our comparative evaluations over both sound and vibration data presented in the plots of Fig. 3 have demonstrated the superiority of the Self-ONNs over the conventional CNNs under the fair evaluation setup.

V. Conclusion

With the recent advances in DL models, vibration-based fault detection has become the de facto standard for rotating machinery. This study explores sound versus vibration in a DL

[2]Trademarked.

method for fault detection in terms of detection performance and robustness over the largest benchmark dataset, QU-DMBF, ever composed with 1080 working conditions. The significant and novel contributions of this study can be summarized as follows.

1) This pioneering study proposes a DL approach for sound-based bearing fault detection and compares it against its vibration-based counterpart.
2) To achieve the objectives of this study, the QU-DMBF benchmark dataset is compiled to obtain sound and vibration data simultaneously for 1080 working conditions. The QU-DMBF dataset, our results, and the optimized PyTorch implementation of the proposed sound-based detection approach are now publicly shared with the research community [62].
3) The proposed sound-based fault detector, based on a compact model of 1-D Self-ONN, achieves state-of-the-art detection performance levels despite only a minority of the sound data being used for training a compact Self-ONN model and tested over a data partition with a large number of "unseen" working conditions.
4) An extended set of comparative evaluations has demonstrated that the sound-based detection performance can match the *best* vibration-based counterpart, observed only when the detector model is trained and tested over the same sensor data.
5) Another important observation of this study is that vibration-based fault detectors are highly sensitive to sensor location. A slight relocation of the sensor can cause a significant deterioration in the detection performance. This is a crucial advantage of sound-based detectors, which are invariant to such changes.

In brief, the proposed method makes motor health monitoring significantly more robust, practical, cheaper, and accessible. Thus, it has the potential to make a crucial impact on the other related fields of health monitoring and predictive maintenance, e.g., mechanical fault detection on vehicles or other engines used for transportation.

Future research will focus on sound-based fault identification and localization to determine the exact nature of each fault, predict its severity, and identify its location for full-scale motor health monitoring.

REFERENCES

[1] Z. Gao, C. Cecati, and S. X. Ding, "A survey of fault diagnosis and fault-tolerant techniques—Part I: Fault diagnosis with model-based and signal-based approaches," *IEEE Trans. Ind. Electron.*, vol. 62, no. 6, pp. 3757–3767, Jun. 2015.

[2] X. Dai and Z. Gao, "From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis," *IEEE Trans. Ind. Informat.*, vol. 9, no. 4, pp. 2226–2238, Nov. 2013.

[3] F. Filippetti, A. Bellini, and G.-A. Capolino, "Condition monitoring and diagnosis of rotor faults in induction machines: State of art and future perspectives," in *Proc. IEEE Workshop Electr. Mach. Des., Control Diagnosis (WEMDCD)*, Mar. 2013, pp. 196–209.

[4] S. Sassi, B. Badri, and M. Thomas, "A numerical model to predict damaged bearing vibrations," *J. Vib. Control*, vol. 13, no. 11, pp. 1603–1628, 2007.

[5] R. R. Schoen, T. G. Habetler, F. Kamran, and R. G. Bartfield, "Motor bearing damage detection using stator current monitoring," *IEEE Trans. Ind. Appl.*, vol. 31, no. 6, pp. 1274–1279, Dec. 1995.

[6] M. E. H. Benbouzid, "A review of induction motors signature analysis as a medium for faults detection," *IEEE Trans. Ind. Electron.*, vol. 47, no. 5, pp. 984–993, Oct. 2000.

[7] J. Pons-Llinares, J. A. Antonino-Daviu, M. Riera-Guasp, S. B. Lee, T.-J. Kang, and C. Yang, "Advanced induction motor rotor fault diagnosis via continuous and discrete time–frequency tools," *IEEE Trans. Ind. Electron.*, vol. 62, no. 3, pp. 1791–1802, Mar. 2015.

[8] L. Eren and M. J. Devaney, "Bearing damage detection via wavelet packet decomposition of the stator current," *IEEE Trans. Instrum. Meas.*, vol. 53, no. 2, pp. 431–436, Apr. 2004.

[9] S. Nandi, H. A. Toliyat, and X. Li, "Condition monitoring and fault diagnosis of electrical motors—A review," *IEEE Trans. Energy Convers.*, vol. 20, no. 4, pp. 719–729, Dec. 2005.

[10] V. T. Do and U.-P. Chong, "Signal model-based fault detection and diagnosis for induction motors using features of vibration signal in two-dimension domain," *Strojniaki vestnik J. Mech. Eng.*, vol. 57, no. 9, pp. 655–666, Sep. 2011, doi: 10.5545/sv-jme.2010.162.

[11] Z. Qiao, Y. He, C. Liao, and R. Zhu, "Noise-boosted weak signal detection in fractional nonlinear systems enhanced by increasing potential-well width and its application to mechanical fault diagnosis," *Chaos, Solitons Fractals*, vol. 175, Oct. 2023, Art. no. 113960, doi: 10.1016/j.chaos.2023.113960.

[12] A. Bellini, F. Filippetti, C. Tassoni, and G.-A. Capolino, "Advances in diagnostic techniques for induction machines," *IEEE Trans. Ind. Electron.*, vol. 55, no. 12, pp. 4109–4126, Dec. 2008.

[13] X. Lou and K. A. Loparo, "Bearing fault diagnosis based on wavelet transform and fuzzy inference," *Mech. Syst. Signal Process.*, vol. 18, no. 5, pp. 1077–1095, Sep. 2004.

[14] W. Sun, S. Shao, R. Zhao, R. Yan, X. Zhang, and X. Chen, "A sparse auto-encoder-based deep neural network approach for induction motor faults classification," *Measurement*, vol. 89, pp. 171–178, Jul. 2016.

[15] M. Xia, T. Li, L. Liu, L. Xu, and C. W. Silva, "Intelligent fault diagnosis approach with unsupervised feature learning by stacked denoising autoencoder," *IET Sci., Meas. Technol.*, vol. 11, no. 6, pp. 687–695, Sep. 2017.

[16] F. Jia, Y. Lei, L. Guo, J. Lin, and S. Xing, "A neural network constructed by deep learning technique and its application to intelligent fault diagnosis of machines," *Neurocomputing*, vol. 272, pp. 619–628, Jan. 2018.

[17] Z. Chen, C. Li, and R.-V. Sanchez, "Gearbox fault identification and classification with convolutional neural networks," *Shock Vib.*, vol. 2015, pp. 1–10, 2015, Art. no. 390134, doi: 10.1155/2015/390134.

[18] O. Janssens et al., "Convolutional neural network based fault detection for rotating machinery," *J. Sound Vibrat.*, vol. 377, pp. 331–345, Sep. 2016.

[19] X. Guo, L. Chen, and C. Shen, "Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis," *Measurement*, vol. 93, pp. 490–502, Nov. 2016.

[20] S. Shao, P. Wang, and R. Yan, "Generative adversarial networks for data augmentation in machine fault diagnosis," *Comput. Ind.*, vol. 106, pp. 85–93, Apr. 2019.

[21] O. Abdeljaber, S. Sassi, O. Avci, S. Kiranyaz, A. A. Ibrahim, and M. Gabbouj, "Fault detection and severity identification of ball bearings by online condition monitoring," *IEEE Trans. Ind. Electron.*, vol. 66, no. 10, pp. 8136–8147, Oct. 2019, doi: 10.1109/TIE.2018.2886789.

[22] S. Liu, J. Xie, C. Shen, X. Shang, D. Wang, and Z. Zhu, "Bearing fault diagnosis based on improved convolutional deep belief network," *Appl. Sci.*, vol. 10, no. 18, p. 6359, Sep. 2020.

[23] S. Zhang, S. Zhang, B. Wang, and T. G. Habetler, "Deep learning algorithms for bearing fault diagnostics—A comprehensive review," *IEEE Access*, vol. 8, pp. 29857–29881, 2020, doi: 10.1109/ACCESS.2020.2972859.

[24] O. Avci, O. Abdeljaber, S. Kiranyaz, S. Sassi, A. Ibrahim, and M. Gabbouj, "One-dimensional convolutional neural networks for real-time damage detection of rotating machinery," in *Proc. Rotating Machinery, Opt. Methods Scanning LDV Methods*, vol. 6, 2022, pp. 73–83.

[25] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-time motor fault detection by 1-D convolutional neural networks," *IEEE Trans. Ind. Electron.*, vol. 63, no. 11, pp. 7067–7075, Nov. 2016.

[26] T. Ince et al., "Early bearing fault diagnosis of rotating machinery by 1D self-organized operational neural networks," *IEEE Access*, vol. 9, pp. 139260–139270, 2021.

[27] T. Ince, S. Kilickaya, L. Eren, O. C. Devecioglu, S. Kiranyaz, and M. Gabbouj, "Improved domain adaptation approach for bearing fault diagnosis," in *Proc. 48th Annu. Conf. IEEE Ind. Electron. Soc.*, Oct. 2022, pp. 1–6.

[28] S. Kiranyaz et al., "Zero-shot motor health monitoring by blind domain transition," 2022, *arXiv:2212.06154*.

[29] O. Can Devecioglu et al., "Sound-to-vibration transformation for sensorless motor health monitoring," 2023, *arXiv:2305.07960*.

[30] T. J. Holroyd and N. Randall, "Use of acoustic emission for machine condition monitoring," *NDT E Int.*, vol. 27, no. 4, p. 210, Jan. 1994.

[31] D. Mba, A. Cooke, D. Roby, and G. Hewitt, "Opportunities offered by acoustic emission for shaft-steal rubbing in power generation turbines: A case study," in *Proc. Int. Conf. Condition Monitor*, 2003, pp. 2–4.

[32] Y. Kim, A. Tan, J. Mathew, and B. Yang, "Experimental study on incipient fault detection of low speed rolling element bearings: Time domain statistical parameters," in *Proc. 12th Asia–Pacific Vibrat. Conf.*, 2007, pp. 6–9.

[33] M. Altaf et al., "Automatic and efficient fault detection in rotating machinery using sound signals," *Acoust. Aust.*, vol. 47, no. 2, pp. 125–139, Aug. 2019.

[34] E. Kim, C. A. Tan, J. Mathew, V. Kosse, and B.-S. Yang, "A comparative study on the application of acoustic emission technique and acceleration measurements for low speed condition monitoring," in *Proc. 12th Asia Pacific Vibrat. Conf.*, 2007, pp. 1–11.

[35] D. Mba, "The detection of shaft-seal rubbing in large-scale turbines using acoustic emission," in *Proc. 14th Int. Congr. Condition Monitor. Diagnostic Eng. Manage.*, 2001, pp. 21–28.

[36] C. K. Tan and D. Mba, "Limitation of acoustic emission for identifying seeded defects in gearboxes," *J. Nondestruct. Eval.*, vol. 24, no. 1, pp. 11–28, Mar. 2005.

[37] J. Lu, K. Wang, C. Chen, and W. Ji, "A deep learning method for rolling bearing fault diagnosis based on attention mechanism and Graham angle field," *Sensors*, vol. 23, no. 12, p. 5487, Jun. 2023, doi: 10.3390/s23125487.

[38] H. Nakamura, K. Asano, S. Usuda, and Y. Mizuno, "A diagnosis method of bearing and stator fault in motor using rotating sound based on deep learning," *Energies*, vol. 14, no. 5, p. 1319, Mar. 2021.

[39] D. Zhang, E. Stewart, M. Entezami, C. Roberts, and D. Yu, "Intelligent acoustic-based fault diagnosis of roller bearings using a deep graph convolutional network," *Measurement*, vol. 156, May 2020, Art. no. 107585.

[40] T. Tran and J. Lundgren, "Drill fault diagnosis based on the scalogram and mel spectrogram of sound signals using artificial intelligence," *IEEE Access*, vol. 8, pp. 203655–203666, 2020.

[41] S. K. Gundewar and P. V. Kane, "Rolling element bearing fault diagnosis using supervised learning methods- artificial neural network and discriminant classifier," *Int. J. Syst. Assurance Eng. Manage.*, vol. 13, no. 6, pp. 2876–2894, Dec. 2022.

[42] K. Vernekar, H. Kumar, and K. V. Gangadharan, "Engine gearbox fault diagnosis using machine learning approach," *J. Quality Maintenance Eng.*, vol. 24, no. 3, pp. 345–357, Aug. 2018.

[43] F. Mohd-Yasin, C. E. Korman, and D. J. Nagel, "Measurement of noise characteristics of MEMS accelerometers," *Solid-State Electron.*, vol. 47, no. 2, pp. 357–360, Feb. 2003.

[44] S. Kiranyaz, J. Malik, H. B. Abdallah, T. Ince, A. Iosifidis, and M. Gabbouj, "Self-organized operational neural networks with generative neurons," *Neural Netw.*, vol. 140, pp. 294–308, Aug. 2021.

[45] J. Malik, S. Kiranyaz, and M. Gabbouj, "Self-organized operational neural networks for severe image restoration problems," in *Neural Networks*, vol. 135. Amsterdam, The Netherlands: Elsevier, Jan. 2021, pp. 201–211.

[46] J. Malik, S. Kiranyaz, and M. Gabbouj, "FastONN - Python based open-source GPU implementation for operational neural networks," 2020, *arXiv:2006.02267*.

[47] S. Kiranyaz, T. Ince, A. Iosifidis, and M. Gabbouj, *Operational Neural Networks*. Cham, Switzerland: Springer, Mar. 2020.

[48] S. Kiranyaz, J. Malik, H. B. Abdallah, T. Ince, A. Iosifidis, and M. Gabbouj, "Exploiting heterogeneity in operational neural networks by synaptic plasticity," in *Neural Computing and Applications*. Cham, Switzerland: Springer, Jan. 2021, pp. 1–19.

[49] S. Kiranyaz, T. Ince, A. Iosifidis, and M. Gabbouj, "Progressive operational perceptrons," *Neurocomputing*, vol. 224, pp. 142–154, Feb. 2017.

[50] D. T. Tran, S. Kiranyaz, M. Gabbouj, and A. Iosifidis, "Progressive operational perceptron with memory," *Neurocomputing*, vol. 379, pp. 172–181, Oct. 2020.

[51] O. C. Devecioglu, J. Malik, T. Ince, S. Kiranyaz, E. Atalay, and M. Gabbouj, "Real-time glaucoma detection from digital fundus images using self-ONNs," *IEEE Access*, vol. 9, pp. 140031–140041, 2021.

[52] J. Malik, O. C. Devecioglu, S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ECG classification by 1D self-operational neural networks," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 5, pp. 1788–1801, May 2022, doi: 10.1109/TBME.2021.3135622.

[53] M. Gabbouj et al., "Robust peak detection for Holter ECGs by self-organized operational neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 11, pp. 9363–9374, Nov. 2023, doi: 10.1109/TNNLS.2022.3158867.

[54] M. U. Zahid, S. Kiranyaz, and M. Gabbouj, "Global ECG classification by self-operational neural networks with feature injection," *IEEE Trans. Biomed. Eng.*, vol. 70, no. 1, pp. 205–215, Jan. 2023, doi: 10.1109/TBME.2022.3187874.

[55] T. Ince, S. Kiranyaz, O. Can Devecioglu, M. Salman Khan, M. Chowdhury, and M. Gabbouj, "Blind restoration of real-world audio by 1D operational GANs," 2022, *arXiv:2212.14618*.

[56] M. Soltanian, J. Malik, J. Raitoharju, A. Iosifidis, S. Kiranyaz, and M. Gabbouj, "Speech command recognition in computationally constrained environments with a quadratic self-organized operational layer," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2021, pp. 1–6.

[57] A. Rahman et al., "Robust biometric system using session invariant multimodal EEG and keystroke dynamics by the ensemble of self-ONNs," in *Computers in Biology and Medicine*, vol. 142. Amsterdam, The Netherlands: Elsevier, Mar. 2022.

[58] S. Kiranyaz et al., "Super neurons," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 8, no. 1, pp. 206–228, Feb. 2024, doi: 10.1109/TETCI.2023.3314658.

[59] J. Malik, S. Kiranyaz, M. Yamac, and M. Gabbouj, "BM3D vs 2-layer ONN," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2021, pp. 1994–1998.

[60] S. Kiranyaz et al., "Blind ECG restoration by operational cycle-GANs," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 12, pp. 3572–3581, Dec. 2022.

[61] J. Malik, S. Kiranyaz, and M. Gabbouj, "Operational vs convolutional neural networks for image denoising," 2020, *arXiv:2009.00612*.

[62] *Exploring-Sound-vs-Vibration-for-Robust-Fault-Detection-on-Rotating-Machinery, Version 1.0, Source Code*. Accessed: May 24, 2024. [Online]. Available: https://github.com/OzerCanDevecioglu/Exploring-Sound-vs-Vibration-for-Robust-Fault-Detection-on-Rotating-Machinery

[63] *Zero Shot Bearing Fault Detection by Blind Domain Transition, Version 1.0, Source Code*. Accessed: May 24, 2024. [Online]. Available: https://github.com/OzerCanDevecioglu/Zero-Shot-Bearing-Fault-Detection-by-Blind-Domain-Transition

Serkan Kiranyaz (Senior Member, IEEE) works as a Professor at Qatar University, Doha, Qatar. He published two books, seven book chapters, ten patents/applications, more than 100 journal articles in several IEEE TRANSACTIONS and other high-impact journals, and more than 120 papers in international conferences. He made significant contributions to bio-signal analysis, classification and segmentation, computer vision with applications to recognition, classification, multimedia retrieval, evolving systems and evolutionary machine learning, swarm intelligence, and evolutionary optimization. His principal research interests include machine learning and signal processing.

Ozer Can Devecioglu was born in Turkey, in 1996. He received the B.S. and M.S. degrees from Izmir University of Economics, Izmir, Turkey, in 2019 and 2022, respectively.

He is currently working as a Researcher at Tampere University, Tampere, Finland. His research interests include machine learning, artificial neural networks, image processing, and signal processing.

Amir Alhams received the bachelor's degree in mechanical engineering from Qatar University, Doha, Qatar, in 2021, where he is currently pursuing the master's degree in mechanical engineering.

His research interests include semi-active vibration control of vehicles, condition monitoring, and diagnosis of rotating machine faults using artificial intelligence.

Sadok Sassi received the Ph.D. degree in applied mechanics from the Polytechnic of Montreal, Montreal, QC, Canada, in 1994.

He is a Professor of Mechanical Engineering at Qatar University, Doha, Qatar. His current research interests include mechanical vibration, stability of structures, condition monitoring of rotating machinery, troubleshooting and diagnosis of damaged machine components, numerical simulation and experimental testing of rotating mechanical systems, and development of smart-material-based intelligent dampers.

Turker Ince (Member, IEEE) received the B.S. degree from Bilkent University, Ankara, Turkey, in 1994, the M.S. degree from Middle East Technical University, Ankara, Turkey, in 1996, and the Ph.D. degree from the University of Massachusetts Amherst (UMass-Amherst), Amherst, MA, USA, in 2001, all in electrical engineering.

From 1996 to 2001, he was a Research Assistant at the Microwave Remote Sensing Laboratory, UMass-Amherst. He worked as a Design Engineer at Aware Inc., Boston, MA, USA, from 2001 to 2004, and at Texas Instruments Inc., Dallas, TX, USA, from 2004 to 2006. In 2006, he joined the Faculty of Engineering at Izmir University of Economics, Turkey, and joined German International University, Berlin, Germany, in 2023, where he is currently a Professor with the Department of Media Engineering and Technology. His research interests are in the areas of machine learning, signal processing, biomedical engineering, predictive analytics, and remote sensing.

Onur Avci received the M.S. and Ph.D. degrees from Virginia Tech, Blacksburg, VA, USA, in 2002 and 2005, respectively.

After his graduation, he worked at major structural engineering firms as a Professional Engineer. He is a Faculty Member at the Department of Civil and Environmental Engineering, West Virginia University, Morgantown, WV, USA. During his industry design experience, he was involved in the analysis, design, renovation, and demolition of more than 5 million meter square of structural space. During his academic experience, he has attracted more than U.S. $4.2 million in research grants. In his research, he focuses on structural dynamics and machine learning applications in structural engineering.

Moncef Gabbouj is a Professor at the Department of Computing Sciences, Tampere University, Tampere, Finland. He was an Academy of Finland Professor. His research interests include big data analytics, multimedia analysis, artificial intelligence, machine learning, pattern recognition, nonlinear signal processing, video processing, and coding.

Dr. Gabbouj is a Fellow of the Asia-Pacific Artificial Intelligence Association. He is a member of the Academia Europaea, the Finnish Academy of Science and Letters, and the Finnish Academy of Engineering Sciences.