***Journal of Dynamics, Monitoring and Diagnostics*, 2026, 5, 25-37**
https://doi.org/10.37965/jdmd.2025.888

RESEARCH ARTICLE

A Comparative Study Between SDP-CNN and Time–Frequency-CNN-Based Approaches for Fault Detection

Mario Spirto, Francesco Melluso, Armando Nicolella, Pierangelo Malfi, Chiara Cosenza, Sergio Savino, and Vincenzo Niola
University of Naples Federico II, Department of Industrial Engineering, Via Claudio 21, 80125 Naples, Italy

(Received 05 August 2025; Revised 26 August 2025; Accepted 16 September 2025; Published online 17 September 2025)

Abstract: The image-based approach is widely used in fault detection (FD) algorithms of mechanical systems. The images are derived from the vibrational signals transformed from the time to time–frequency domain, and they are used to develop a convolutional neural network (CNN) to automate the FD process. Nowadays, images are also obtained from the transformation of vibrational signals from the time domain to symmetrized dot pattern (SDP) coordinates, achieving high CNN testing accuracy. This paper shows a comparison of image-CNN approaches for FD using images obtained from time–frequency transforms and those obtained from the SDP transform as input. The comparison was conducted using experimental data from two publicly available bearing datasets, examining both the accuracy of the CNNs and the computational time required for the vibrational signal transformations. The results show that the SDP-CNN approach achieves the same accuracy as spectrogram-CNN approaches but with a significantly reduced computational time. These results support the future real-time implementation of the SDP-CNN approach for FD in mechanical systems such as bearings.

Keywords: ball bearings; CNN; fault detection; symmetrized dot pattern; time–frequency analysis

I. INTRODUCTION

Fault detection (FD) and diagnosis in rotating machinery are essential to identify anomalies, reduce downtime, and prevent unexpected failures. A typical FD process is developed from vibrational analysis based on features derived from the time, frequency, and time–frequency domain [1,2].

In recent years, image-based approaches have gained considerable attention for FD by transforming 1-D vibration signals into 2-D representations such as time–frequency spectrograms. These images enable advanced machine learning (ML) models, particularly convolutional neural networks (CNNs), to automatically learn discriminative features [3,4]. Commonly used time–frequency transforms include the short-time Fourier transform (STFT) and continuous wavelet transform (CWT), which have been successfully applied in gear and bearing FD [5–7]. Improved time–frequency transforms [8,9], deep ML models [10,11], or synchrosqueezing wavelet transform (SWT) [12] have been used to enhance the image-CNN approach. Other approaches incorporate frequency attention mechanisms within CNN architectures to enhance the focus of the ML models on the most informative frequency components while suppressing irrelevant noise [13,14]: they have shown superior performance across multiple fault types and rotating machinery components.

Recently, the symmetrized dot pattern (SDP) method has become a relevant approach for rotating machinery FD: it provides a vibrational signal transformation into a visual representation to identify faults [15]. The SDP is frequently combined with CNNs to develop an image-based FD process [16–18], demonstrating its effectiveness in systems ranging from gearboxes [19,20] and fans [21] to more recent applications in lithium battery modules [22]. The SDP images allow more accurate classification of faults than the analysis of raw signals [23] and can reveal significant differences between Gaussian and white noise, which are difficult to distinguish in traditional spectra [24]. A key advantage of this visual approach is its ability to reveal spatial relationships among primary features, which is particularly beneficial for diagnosis under variable working conditions [25]. To enhance the classical SDP-CNN approach, researchers have combined it with signal processing techniques like empirical mode decomposition and variational mode decomposition [26,27]. This fusion of methods helps in generating more informative and discriminative images for the neural network, as also demonstrated using optimized SDPs for switched reluctance motor fault diagnosis [28]. Furthermore, the SDP has proven highly valuable for the FD process under variable working conditions [29,30]. To address the domain shift problem inherent in these scenarios, advanced techniques such as adversarial gradual domain adaptation have been explored to improve model generalizability [31]. Finally, the architecture of the deep learning models processing SDP images continues to evolve. Beyond standard CNNs, novel hybrid approaches are being developed, such as combining CNNs with K-nearest neighbors for robust compound fault diagnosis [32] or utilizing convolutional probabilistic neural networks [33], further expanding the applicability and robustness of the SDP-based fault diagnosis framework.

Xu *et al.* [34] present a fast and accurate method for real-time detection of rotating stall in centrifugal fans, which converts pressure signals into images using SDP analysis and compares them with reference templates through image matching. By applying Principal Component Analysis directly to the SDP images, noise and redundancy are effectively reduced, significantly improving speed and accuracy compared to wavelet filtering. The approach achieves a minimal detection delay with computational times suitable for real-time applications, making it a

---

Corresponding author: Mario Spirto (e-mail: mario.spirto@unina.it)

reliable method for providing timely stall warnings. This approach does not involve the use of classification CNN networks or other ML techniques.

FD based on the image-CNN approach is typically carried out by deriving input images (spectrogram) from the time–frequency transform of vibrational signals. This method has demonstrated high accuracy in CNN testing across various mechanical systems and different defect types. The computational time is closely linked to the chosen time–frequency transform. In recent years, the input images have also been obtained from diagrams of vibrational signals, SDP transformed achieving high accuracy with a lower computational time. Given the high accuracy demonstrated by the image-CNN approach for FD, it is interesting to evaluate which technique offers the best trade-off between computational time and CNN accuracy, with a view toward potential real-time implementation of an image-CNN-based FD method. Many studies have investigated the SDP transform applied to signals from different sensors and systems, demonstrating the accuracy of the SDP-CNN approach for FD. Furthermore, the SDP-CNN framework has often been enhanced either by preprocessing the signals before applying the SDP transformation or by directly processing the diagrams. Nevertheless, SDP-CNN has never been compared with the conventional image-CNN approaches, which use time–frequency transform images of the acquired signals as input.

The novelty of this work lies in the direct comparison between SDP-CNN and the classical spectrogram-CNN approach. The results highlight two key points:

- SDP-CNN approach achieves accuracy comparable to spectrogram-CNN methods.
- SDP transformation requires significantly lower computational cost than other time–frequency transforms.

This demonstrates that SDP-CNN is a stronger candidate for real-time monitoring compared to other spectrogram-CNN approaches, as it provides the best trade-off between computational efficiency and predictive accuracy of the CNN.

This paper compares the image-CNN approaches between the images derived by time–frequency transformations and the images derived by the SDP transformation. The comparison was carried out both in terms of the test performance of the specifically designed and optimized CNNs, and in terms of the computational time required to obtain the transformed signals. The experimental results are obtained using two public datasets of ball bearings that include various defect types, defect sizes, and operating conditions. They have proven that the SDP-CNN approach achieves a comparable test accuracy as spectrogram-CNN approaches, but with a reduced computational time during the signal transformation from the time domain to SDP coordinates.

The main contributions of this work can be summarized as follows:

1. This study proposes a comparison between the FD based on image-CNN techniques between SDP-CNN and spectrogram-CNN.
2. The comparison is performed on experimental public datasets of ball bearings that include multiple defect types, defect sizes, and operating conditions.
3. The comparison is evaluated in terms of the test accuracy of the specifically designed and optimized CNNs and in terms of the computational time for transforming the signals.

The remainder of the paper is structured as follows: Section II outlines the workflow, the SDP coordinate transformation, and the time–frequency transformations chosen for the comparison; Section III describes the experimental datasets; Section IV presents and discusses the experimental results; and Section V summarizes the main conclusions of this work.

A. IMAGE-CNN APPROACHES

The proposed comparison is based on the performance of automated FD through CNNs. These networks use as input the images obtained from some two-dimensional transformations of the accelerometer signals. First, the vibrational signal is transformed using either SDP or a time–frequency analysis method (STFT, CWT, Kurtogram, or SWT) to highlight both time and frequency characteristics. In particular, the SDP method converts 1-D temporal signals into symmetric, visually distinct patterns in polar coordinates, while time–frequency transforms offer rich frequency-temporal insights.

Then, each transformed output diagrams are converted as a binarized image fed into optimized CNNs. Finally, the CNN outputs a classification, determining the health status of bearings. The workflow is represented in Fig. 1.

The pipeline in Fig. 1 is applied separately for each transform. It can be summarized as follows:

1. The raw signals are divided into one-second intervals.
2. They are then normalized to the range $[-1, 1]$.
3. The normalized intervals are subsequently transformed using either the SDP or a time–frequency approach.
4. The transformations produce 2D representations, which are binarized by setting a threshold.
5. The resulting images form a datastore that is split into training, validation, and test sets.
6. The three image sets are used to train, optimize, and test a CNN for FD.

Below there are a description of the used transformations and the CNNs.

B. SDP TRANSFORMATIONS

The SDP technique transforms a signal into a normalized two-dimensional representation by plotting amplitude values in a polar coordinate system, resulting in a symmetrical diagram. This visualization method highlights variations in signal amplitude and frequency, aiding in the diagnosis of faults in rotating systems such as bearings. The differences between various signals are reflected in the distinct shapes of the petals forming the “snowflake” pattern. The “snowflake diagram” refers to symmetric image generated by SDP, that is, scatter representation in polar coordinates, where each data point is mirrored to form radial symmetry. This creates a distinctive structure whose geometric differences directly correspond to variations in underlying signal characteristics, making it easier to visually distinguish between fault types or detect subtle anomalies. This diagram is constructed using k planes of symmetry, which are created by circular repetition of the petals through specific angles γ:

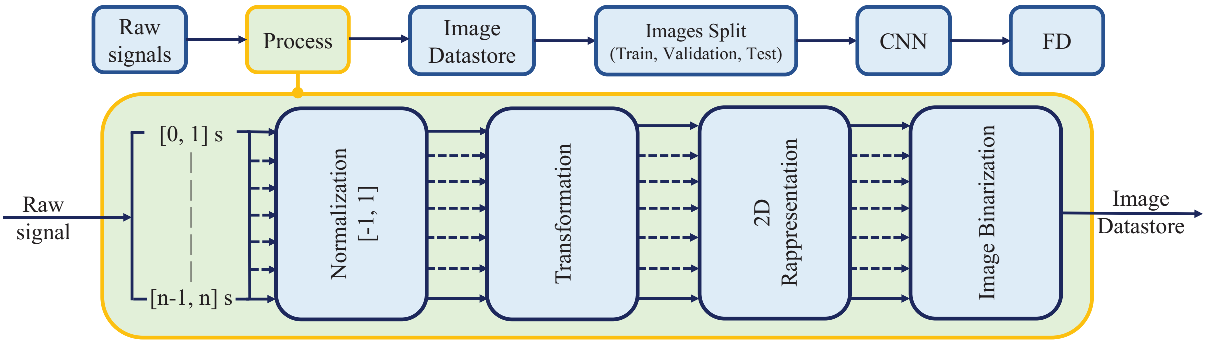

Fig. 1. Image-CNN approaches workflow.

$$\gamma = \frac{360}{k} \tag{1}$$

Not overlapping the petals is a necessary condition for plotting the snowflake.

It is possible to transform a signal $y = y_1, \ldots, y_i, \ldots, y_n$ in polar coordinates using the SDP transform by the formulas:

$$\rho_i = \frac{y_i - max(y)}{max(y) - min(y)} \tag{2}$$

$$\alpha_i = \gamma + \frac{y_{i+h} - min(y)}{max(y) - min(y)}\lambda \tag{3}$$

$$\beta_i = \gamma - \frac{y_{i+h} - min(y)}{max(y) - min(y)}\lambda \tag{4}$$

where $min(y)$ and $max(y)$ are the minimum and the maximum value of y, respectively, ρ_i is the radius of the i-th point, α_i is the clockwise deflection angle of the i-th point along the symmetry plane, β_i is the anticlockwise deflection angle of the i-th point along the symmetry plane, h is the delay coefficient, γ is the rotation angle of the symmetry plane, and λ is the gain of the deflection angle.

C. TIME–FREQUENCY TRANSFORMATIONS

The most common time–frequency transforms are briefly described below. The diagrams of these transformations are used for image generation. These images will then be used as input for automatic FD via CNN.

The STFT [37] is a time–frequency transformation that allows to divide the time-signal x into time-segment which the discrete Fourier transform (DFT) is calculated. The STFT $X(\tau,\omega)$ of a time-domain signal $x(t)$ can be expressed as:

$$X(\tau,\omega) = \int_{-\infty}^{\infty} x(t) \cdot \omega(t-\tau) \cdot e^{-i\omega t} dt \tag{5}$$

where $\omega(t-\tau)$ is the window function centered in τ that defined the window of the $x(t)$ where the DFT must be calculated, and ω is the angular frequency.

The Kurtogram [38] is based on the Spectral Kurtosis SK a statistical measure that quantifies the impulsivity of a time-signal $x(t)$ as a function of frequency f:

$$SK(f) = \frac{E[|X(f)|^4]}{([|X(f)|^2])^2} - 2 \tag{6}$$

where $X(f)$ is the Fourier transform of $x(t)$ and $E[\cdot]$ is the expected value.

The CWT [39] is a time–frequency transform of a time-signal $x(t)$ defined as:

$$X_w(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(t) \cdot \overline{\psi}\left(\frac{t-b}{a}\right) dt \tag{7}$$

where $\psi(t)$ is the mother wavelet and $\overline{\psi}$ is the complex conjugates, $a > 0$ is the scale factor that expands or compresses the wavelet function, and $b \in \mathbb{R}$ is the translation factor that moving the wavelet function along the time.

The SWT [40] is a time–frequency transformation of a time-signal $x(t)$ that improve the resolution of CWT concentrating energy around the estimated instantaneous frequencies. After the calculation of the CWT of $x(t)$ as shown in Eq. 7, the instantaneous frequency ω_s is estimated as:

$$\omega_s(a,b) = -i \cdot \frac{\partial_b X_w(a,b)}{X_w(a,b)} \tag{8}$$

where $\omega_s(a,b)$ returns the local instantaneous frequency ω_s at each point (a,b). The SWT of $x(t)$ move the energy from (a,b) to (ω_s,b) to improve the time–frequency representation:

$$T_s(\omega_s,b) = \int X_w(a,b) \cdot a^{-3/2} \delta(\omega_s(a,b) - \omega) da \tag{9}$$

where δ is the Dirac function.

D. IMAGE BINARIZATION, CNN DEVELOPMENT, AND OPTIMIZATION

Once the 2D representation of the signal was obtained, the resulting image was binarized:

- If the signal was transformed using SDP, a threshold was set to highlight all the dot patterns.
- If the signal was transformed using a time–frequency approach, a threshold was selected to emphasize the distinctive features of each defect while suppressing noise.

The obtained binarized images were used as input to train and optimize CNNs for FD. A CNN is a specialized form of artificial neural network designed for image recognition tasks [41]. It processes single-channel binary images through a deep convolutional framework, where the architecture depth and hyperparameters are fine-tuned using Bayesian optimization [42]. Each convolutional block contains the following sequence:

- A convolutional layer equipped with multiple filters.
- Batch normalization (BN) stabilizes and accelerates training by normalizing layer inputs.
- Rectified Linear Unit (ReLU) activation function to introduce nonlinearity.
- Max pooling to down-sample the spatial dimensions, reducing computational complexity.

The output from the final convolutional block is flattened into a one-dimensional vector and passed through a fully connected neural network (FNN). This is followed by a ReLU activation and a Softmax layer to produce the final classification probabilities.

The network is trained using the cross-entropy loss function, optimized via stochastic gradient descent with momentum (SGDM). Bayesian optimization, employing the expected improvement criterion, is utilized to tuning the network hyperparameters.

The final architecture of the optimized CNNs is structured as reported in [30].

II. EXPERIMENTAL DATASETS

The comparison was conducted using two different datasets: the Case Western Reserve University (CWRU) [35] and the Hanoi University of Science and Technology (HUST) [36] dataset. These datasets were selected to cover different working conditions, defect types and sizes, as well as ball bearings of different sizes.

A. DATASET 1: CWRU

The CWRU dataset was provided by the Case Western Reserve University Bearing Data Center. This dataset is widely used as a benchmark for bearing fault diagnosis. This dataset is widely used as a benchmark for bearing fault diagnosis. The test rig includes a 3 hp Reliance electric motor, torque sensor, encoder, dynamometer, and two test bearings, one at the drive end (DE) and one at the fan end (FE). The faults are artificially introduced via electro-discharge machining in three bearing locations: inner race (IR), ball (B), and outer race (OR). The fault diameters ranged from 7 mils to 40 mils, and OR faults were further categorized by their placement relative to the load zone (at 3 o'clock, 6 o'clock, or 12 o'clock). Vibration data is collected from accelerometers at the DE, FE, and base (BA) at 12 or 48 kHz. Experimental conditions spanned 0 to 3 hp motor loads, corresponding to speeds between 1797 RPM and 1720 RPM, ensuring a range of operating scenarios for evaluating fault diagnosis techniques. Tests are run at constant speed and load.

B. DATASET 2: HUST

The HUST bearing dataset comprises 99 vibration recordings across five types of bearings (6204, 6205, 6206, 6207, and 6208) and includes faults in the IR, OR, and B as well as combined fault scenarios. Faults are introduced via wire-cut micro-cracks about 0.2 mm and simulating early-stage damage. Data were collected at loads 0, 200, and 400 W and include both steady-state and run-up (start-up transient) conditions. Each recording lasts 10 seconds and is sampled at 51.2 kHz. The acquisition setup features a 750 W induction motor, a multistep shaft, a powder brake to impose load, and a torque transducer plus dynamometer for monitoring. A PCB 325C33 accelerometer captures radial vibrations. The rotational speed data are captured by a dynamometer, which monitors speed evolution throughout both steady and transient phases of each test.

III. RESULTS AND DISCUSSION

The results obtained separately for each dataset are presented and discussed below. The only signals preprocessing was the normalization of each one-second segment in the range between −1 and 1.

To ensure fairness in the comparison among the different transformations, as detailed below, the following choices were made:

- All execution times were measured on the same hardware and within the same MATLAB environment.
- All signals underwent the same preprocessing.
- All signals were segmented into 1-second intervals for the transformations.
- All images were binarized in the same way.
- For the time–frequency methods, typical parameter settings were adopted to avoid favoring specific transformations.

A. DATASET 1 RESULTS

The signals sampled at 12 kHz by the DE accelerometer and all fault types with diameters of 0.007, 0.014, and 0.021 mm were considered. The three OR location faults were included in one class. Figure 2 shows one second of raw signals.

Figure 2 shows the raw vibrational signals in N, B, IR, and OR conditions. The N conditions presents a stationary noise without evident impulses, while it is possible to see high-frequency components in the defects conditions,

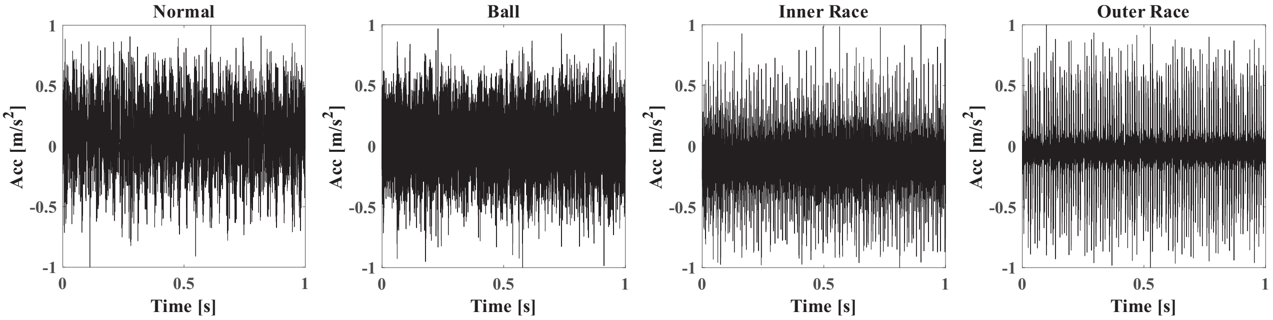

Fig. 2. Raw signals for dataset 1.

especially in IR and OR signals. These characteristics highlight a signal modulation due to the periodic impact between the defect and the rolling elements. The different morphology of the signals suggests the possibility of discriminating between different conditions.

It is necessary to choose the SDP parameters k, h, and λ appropriately to transform signals from the time domain into SDP coordinates:

- k was chosen equal to 6 from which γ was obtained equal to 60 from Eq. 1. This choice was made to obtain a better definition of the center of the diagram.
- h and λ were set as shown in [43].

Each snowflake diagram was obtained by choosing a one-second acquired signal in the time domain. The calibration process of parameters h and λ was performed for signals acquired in the N case and is shown in Fig. 3.

In Fig. 3, it is possible to see the influence of h and λ on the snowflake diagram. The conditions h and λ to be properly chosen is that:

- There is no overlap between the petals.
- The petals must be sufficiently open to be clearly defined.

To fulfill both conditions, $h = 14$ and $\lambda = 30$ were chosen.

Figure 4 shows the SDP transformation for one second of raw signals.

In Fig. 4, the N conditions exhibit a smooth, continuous spiral with symmetrically distributed points, reflecting stationary noise without impulsive components, and the point density is uniform, indicating the absence of significant transient events. In this case, the petals indicate a signal dominated by stationary noise, free from impulsive components, and therefore associated with a regular operating condition of the bearing.

The B conditions show more concentrated petals, suggesting the presence of pulses generated by the rotating elements: the radial symmetry remains, but the structure is significantly more segmented and less continuous than in the N case. Physically, the distribution of petals corresponds to the generation of impulses each time the damaged ball meets the inner or outer race, while still preserving a certain radial symmetry. For IR, the petals are narrower and more sharply defined, while in the OR case, the petals appear broader and less defined, with greater radial dispersion at the petal tips. These visual transformations align with observations in the raw signals, where high-frequency impulsive components appear in defective cases. In the SDP domain, these manifest as petal-shaped structures. In the case of an IR fault, the petals appear narrower and more sharply defined: this occurs because the inner race, being connected to the rotating shaft, generates repetitive high-frequency impacts each time the balls pass over the damaged area, with high regularity. Conversely, in the case of an OR fault, the petals appear broader, less defined, and characterized by greater radial dispersion, since the outer race is fixed and is excited only when the balls roll over the damaged zone, producing fewer regular impulses with higher angular variability. The morphological differences

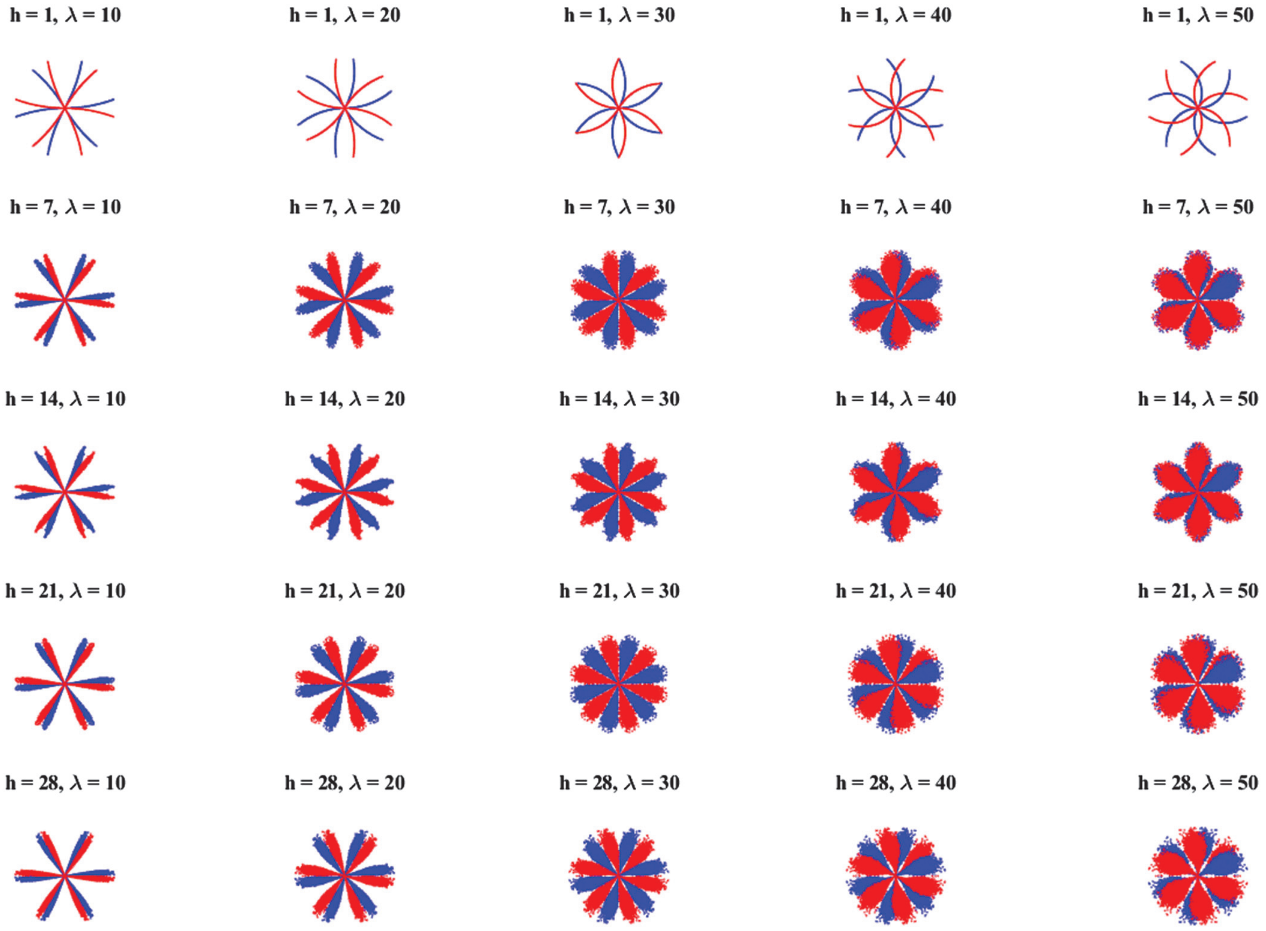

Fig. 3. SDP parameters tuning.

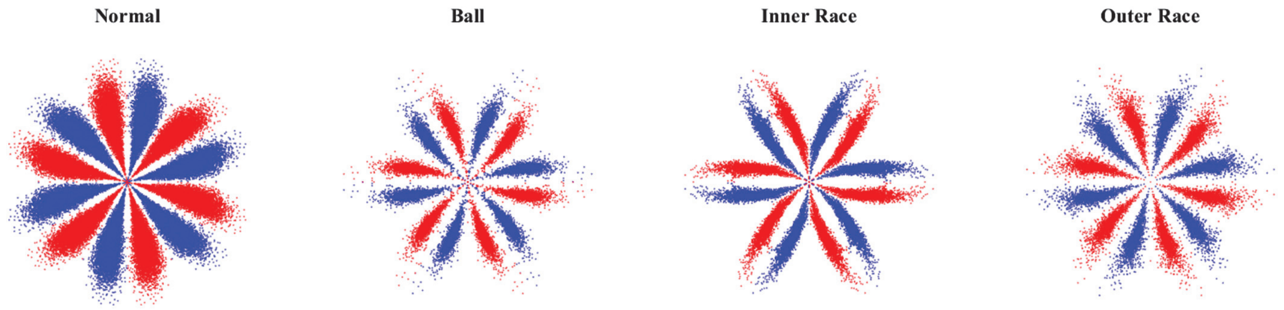

Fig. 4. SDP snowflakes for dataset 1.

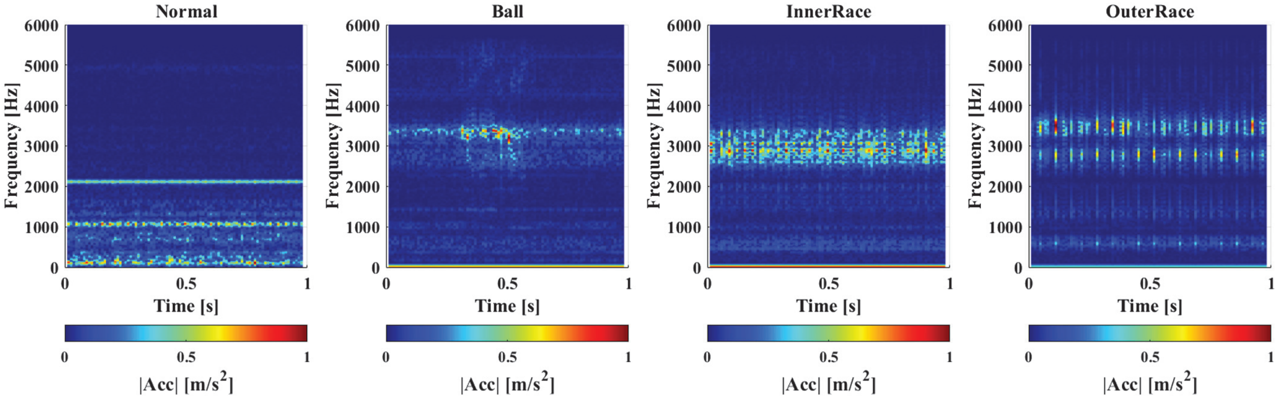

Fig. 5. STFT diagrams for dataset 1.

between IR and OR patterns suggest that SDP is sensitive not only to the presence of faults but also to their location within the bearing, enabling potential automatic discrimination based on characteristic patterns.

The diagrams of the time–frequency transforms described in Section B are commented and given below to make the comparison between the SDP and the time–frequency diagrams.

The STFTs of the raw signals (Fig. 2) are reported in Fig. 5. Each FFT transform was performed on 512 samples using a Hamming window of 256 samples with an overlap of 128 samples.

In Figure 5, the spectrum of the N case shows low noise and an energy concentration in the lower frequency band (<1500 Hz), likely associated with harmonic components of shaft rotation. In the B case, it is possible that distinct impulsive components distributed over time are observed, with significant energy around 3000 Hz: this reflects the transient nature of impacts caused by a localized fault on the ball, which generate intermittent and broader frequency distributions. For the IR conditions, the spectrum shows an energy band around 3000 Hz, with a dense pattern throughout the entire time interval. This behavior is consistent with the cyclic nature of impacts due to a fault on the inner race, which recurs with every rotor revolution, causing high-energy signals. Finally, OR shows a structure of periodic horizontal bands, suggesting the presence of regular impulses corresponding to the fixed (outer) race fault and the rotor rotation. The multiple bands indicate the generation of harmonics due to repetitive impacts at a constant frequency.

Figure 6 shows the Kurtograms.

In the N case of Fig. 6, the spectral kurtosis is low and non-localized, indicating the absence of significant transients. This is consistent with regular vibration behavior typical of a healthy bearing, while the B case shows high kurtosis at short windows and in frequencies ranging around 2500–4000 Hz, suggesting that transient impacts generated by the ball defect are localized and high frequency. The IR shows transient activity around 3500–5000 Hz and intermediate window lengths, consistent with the regular, high-energy impacts that occur cyclically. The OR case exhibits a distribution of kurtosis around 1000–3000 Hz and short windows, consistent with an impulsive structure that repeats regularly.

Figure 7 shows the CWTs. The chosen mother wavelet was the Morse wavelet.

The CWT of the N case in Fig. 7 shows well distributed around 2000 and 4000 Hz with low temporal variability, typically of a regular vibration signal due to shaft rotational frequencies. B defect shows energy concentrated in the high-frequency band (4500–5000 Hz), while IR shows repeatable events in the same frequency band. The OR defect shows very high energy in the upper bands and around 3500 Hz.

In Fig. 8, the SWTs are reported. The chosen mother wavelet was the Morse wavelet with 10 voices per octave.

Fig. 8 shows how the SWT improved the frequency localization of CWT. The N case shows harmonic distribution around 2000 and 5000 Hz with periodic vibrations. The B case shows components around 4000 and 5000 Hz, the IR case highlights harmonics around 4500 Hz, and OR shows components mainly above 5000 Hz.

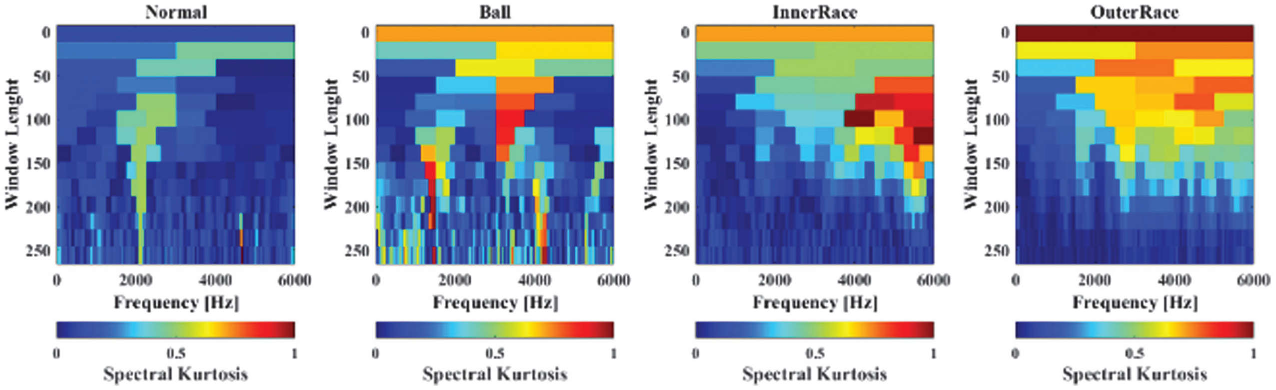

Fig. 6. Kurtogram for dataset 1.

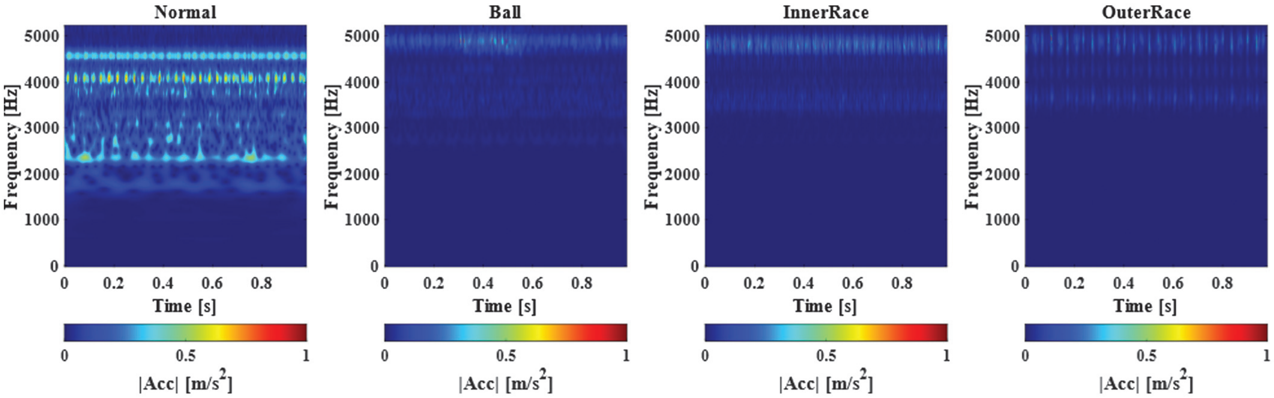

Fig. 7. CWT diagram for dataset 1.

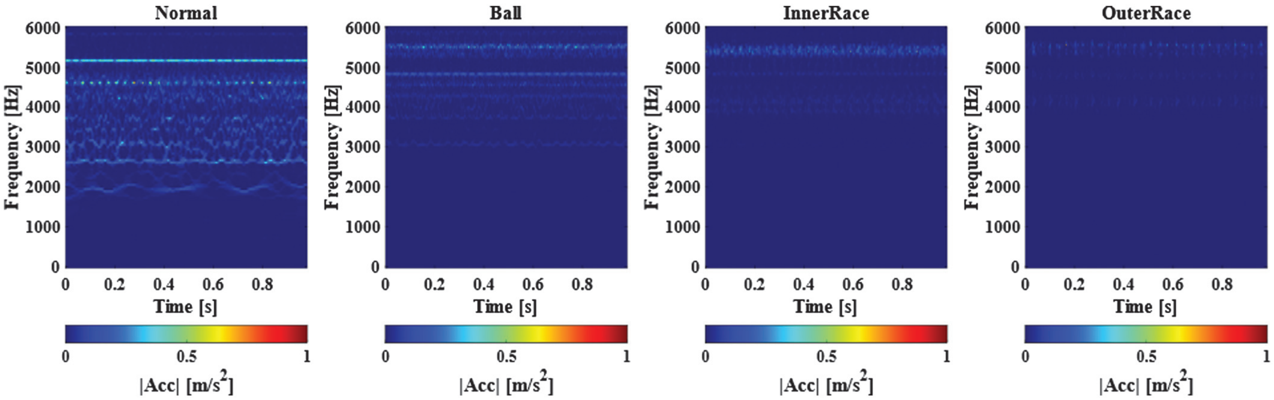

Fig. 8. SWT diagram for dataset 1.

Comparing the various techniques, the SDP demonstrates a high visual ability to distinguish bearing states but is limited to qualitative aspects and requires ML approaches for automated diagnostics. Time–frequency transforms show a dynamic spectral content, allowing the emergence of fault-related bands and pulses to be observed.

The images obtained from the transformed signals were binarized and used as input to develop five CNNs that predict whether the bearing is healthy or has one of the previously listed defects. The images were split into 50 % for training, 20 % for validation, and 30 % for testing using a stratified partitioning strategy. This division ensures that the number of samples between classes (N, IR, OR, and B) is always balanced between the three subsets. The developed CNN was optimized by tuning several hyperparameters, using a Bayesian optimization process to identify the most effective configuration. Table I reports all the hyperparameters considered, together with the corresponding search ranges defined for the optimization procedure.

It is worth noting that the training and optimization times of the various networks are all comparable. The used PC is characterized by a processor 13th Gen Intel Core

Table I. Hyperparameter range

Hyperparameter	**Range**
Convolutional layers	1 ÷ 5
Stride of convolutional layers	2 ÷ 5
Filters size of convolutional layers	2 ÷ 5
Filters numbers of convolutional layers	2 ÷ 50
Stride of pooling	2 ÷ 5
Pooling size	1 ÷ 5
Fully connected layers number	1 ÷ 5
Neurons numbers of fully connected layers	1 ÷ 1000
Initialization learning rate	$10^{-6} \div 10^{-1}$
Momentum	0.01 ÷ 0.99
Mini batch size	2 ÷ 64
L2 regularization	$10^{-6} \div 10^{-1}$

i9-13900 of 2.00 GHz, 32 GB of RAM, and a 12 GB NVIDIA GeForce RTX 3060.

The architectures of the optimized CNNs for dataset 1 are reported in Table II.

Table III shows the performance metrics obtained during the test phase of each CNN and the computational time to perform the signal transformations. The metrics used are accuracy (the percentage of correctly classified samples over the total), precision (the percentage of samples identified as faulty that are actually faulty), recall (the percentage of actual faults correctly detected by the network), and F1-score (the harmonic mean of precision and recall, which measures the balance between false alarms and missed faults).

The results in Table III clarify that the use of the time–frequency transform-derived images achieve perfect performance across all metrics, suggesting comprehensive and discriminative information for distinguishing between all classes.

The CNN based on SDP achieves a lower performance of 0.43% for the accuracy, 2.33% for the precision, and 1.18% for the F1-score; a recall of 100% suggest no false negatives. Instead, the SDP is the faster transformation thanks to its extremely lightweight computational nature without the spectral processing. The STFT offers an excellent compromise between computation time and information yield, making it practical for diagnostic pipelines, while the other time–frequency transformations require a significantly higher computational time. Indeed, Table III confirms that the lowest computational time is relative to the SDP transform and comparing it with the STFT (the second fastest techniques) the SDP is 3.78 times lower than the STFT.

Table IV shows the mean, median, and standard deviation of the execution time for each transformation.

Table II. Optimized CNNs for dataset 1

SDP	**SWT**	**Kurtogram**	**CWT**	**SWT**
IML	IML	IML	IML	IML
C2DL(2,48,2)	C2DL(4,11,2)	C2DL(2,19,2)	C2DL(3,44,4)	C2DL(2,16,2)
BNL	BNL	BNL	BNL	BNL
RL	RL	RL	RL	RL
MP2DL(3,3)	MP2DL(4,2)	MP2DL(3,3)	MP2DL(2,2)	MP2DL(3,2)
FCL(192)	C2DL(2,46,2)	C2DL(2,19,2)	FCL(856)	C2DL(2,16,2)
RL	BNL	BNL	RL	BNL
FCL(658)	RL	RL	FCL(4)	RL
RL	MP2DL(2,2)	MP2DL(2,2)	SL	MP2DL(2,2)
FCL(4)	FCL(139)	C2DL(2,39,2)		C2DL(2,32,2)
SL	RL	BNL		BNL
/	FCL(4)	RL		RL
/	SL	MP2DL(2,2)		MP2DL(2,2)
/	/	FCL(366)		FCL(486)
/	/	RL		RL
/	/	FCL(774)		FCL(597)
/	/	RL		RL
/	/	FCL(4)		FCL(4)
/	/	SL		SL

Table III. Dataset 1 results

Method	**Accuracy [%]**	**Precision [%]**	**Recall [%]**	**F1-score [%]**	**Computational time [s]**
SDP	99.57	97.67	100	98.82	0.18
STFT	100	100	100	100	0.60
Kurtogram	100	100	100	100	18.80
CWT	100	100	100	100	11.58
SWT	100	100	100	100	76.58

Table IV. Dataset 1 stats

Method	Mean [s]	Std [s]	Median [s]
SDP	$2.285 * 10^{-4}$	$3.084 * 10^{-4}$	$2.492 * 10^{-4}$
STFT	$7.803 * 10^{-4}$	$1.738 * 10^{-4}$	$7.522 * 10^{-4}$
Kurtogram	0.028	$6.968 * 10^{-4}$	0.028
CWT	0.017	0.001	0.017
SWT	0.114	0.004	0.116

Table V. Dataset 1 paired t-tests results

Method	h	p-Value	Test statistics	Std
SDP-STFT	1	0	−82.752	$1.713 * 10^{-4}$
SDP-Kurtogram	1	0	$-1.018 * 10^{3}$	$6.985 * 10^{-4}$
SDP-CWT	1	0	−362.201	0.001
SDP-SWT	1	0	−823.325	0.004

The results in Table IV confirm that the SDP method is significantly faster than the others, with average times about 10^{-4} s. In contrast, Kurtogram, CWT, and SWT exhibit considerably higher computational costs.

Table V show the paired t-test results.

Table V results highlight that all comparisons indicate a statistically significant difference between the execution times of SDP and the other methods, with p-values equal to zero. The negative test statistics further confirm that SDP consistently achieves lower computational costs compared to STFT, Kurtogram, CWT, and SWT.

B. DATASET 2 RESULTS

The bearings ID 6205, 6206, 6207, and 6208 and all faults at all operating conditions were considered. Figure 9 shows one second of raw signals.

The SDP tuning process is performed as previously seen for dataset 1. Also in this case, $h = 14$ and $\lambda = 30$ were chosen to obtain the snowflake in Fig. 10.

Below are the STFTs (Fig. 11), Kurtograms (Fig. 12), CWTs (Fig. 13), and SWTs (Fig. 14).

The CNNs were trained by always dividing the binarized images into 50% train, 20% validation, and 30% for the test and always optimized according to the hyperparameters shown in Table I. The architectures of the optimized CNNs for dataset 2 are reported in Table VI, while in Table V the CNN test phasing results and the computational time to perform the signal transformations are reported.

Table VII highlights a performance drop for dataset 2 compared to dataset 1 across all image-CNN approaches, except for the CWT-CNN method. This decline can be attributed to two factors:

- Dataset 2 includes a greater number of operating conditions than dataset 1.
- Dataset 2 involves five ball bearings of different sizes, whereas dataset 1 contains only a single type of bearing.

Observing Table VII, it is possible to see that the CWT confirms its effectiveness in more complex datasets, while Kurtogram and SWT detect faults perfectly but with a low false positive rate. The STFT is the least effective, suggesting less than perfect effectiveness with variable datasets. The SDP confirms the good performance seen for dataset 1, albeit with a slight decrease in accuracy. The SDP is again the transform with the lowest computational time at only

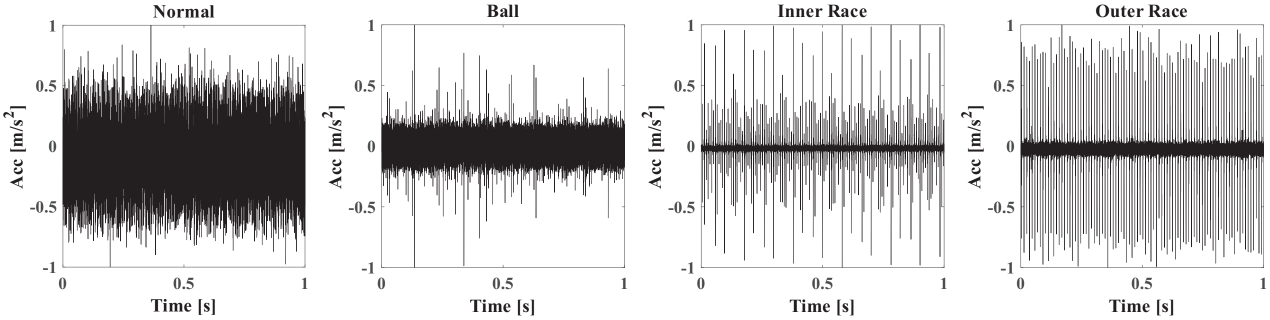

Fig. 9. Raw signals for dataset 2.

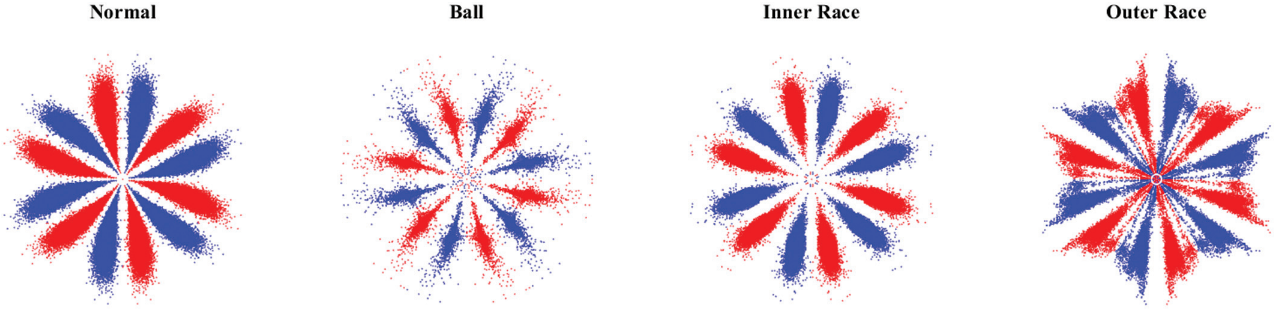

Fig. 10. SDP snowflake for dataset 2.

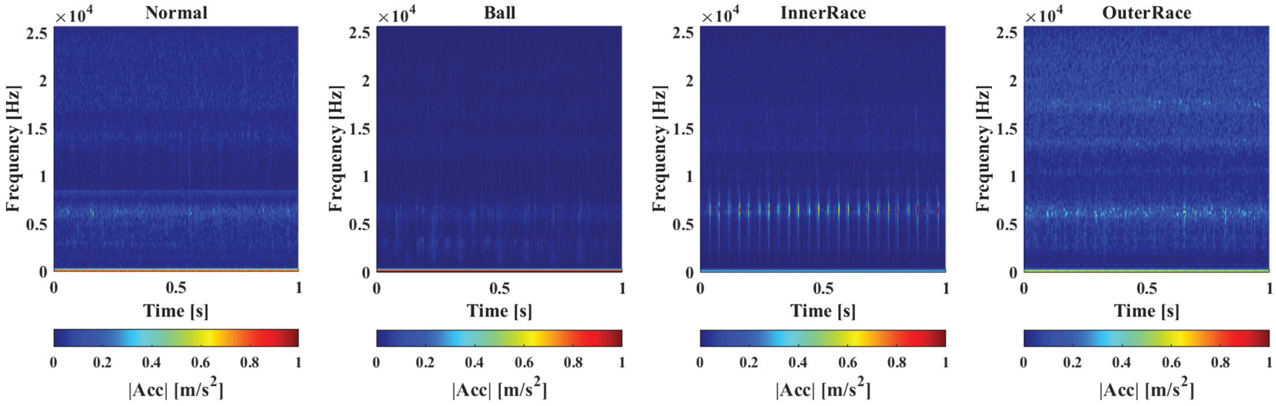

Fig. 11. STFT diagrams for dataset 2.

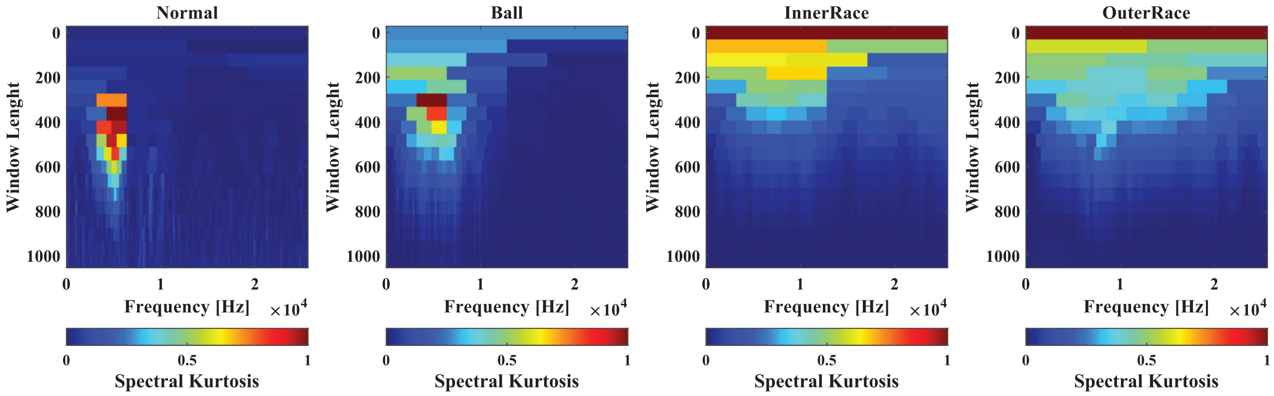

Fig. 12. Kurtogram for dataset 2.

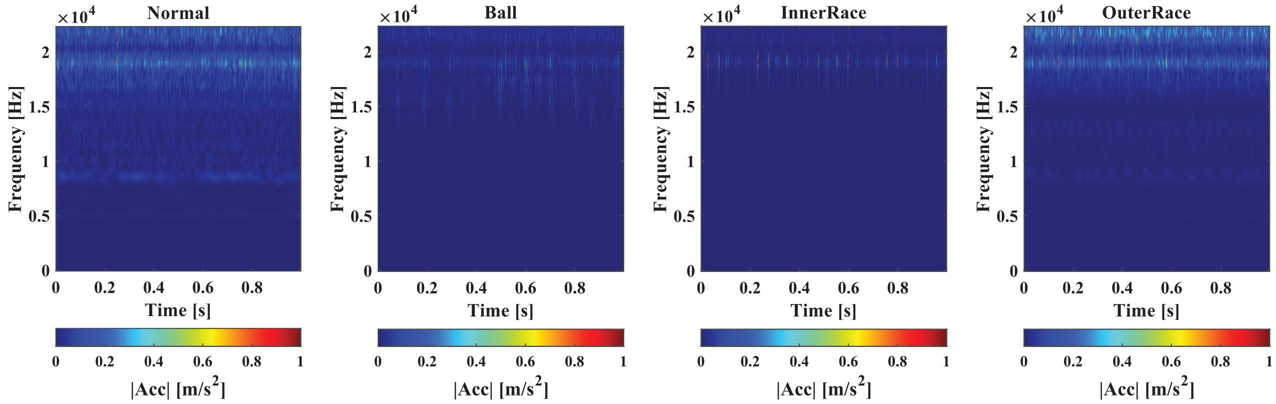

Fig. 13. CWT diagram for dataset 2.

0.26, while the STFT is still second at 4.54 times higher than the SDP.

Tables VIII and IX show the comparison between the computational time statistics for each transform and the paired t-tests results, respectively.

The results in Tables VIII and IX confirm that the SDP is the faster transformation.

IV. CONCLUSIONS

In this work, image-based approaches for the FD of ball bearings were compared. Time–frequency and SDP transforms of vibrational signals were computed, and the images were derived to be used as input for CNNs. Experimental results were obtained on two public datasets of ball bearings that included different defect types, defect sizes, and operating conditions. All techniques achieve high test accuracy and low false positive and negative rates. Furthermore, the techniques were also compared in terms of the computational time required to perform the transforms, demonstrating that SDP is always the most advantageous.

The results highlight that the SDP-CNN achieves accuracy comparable to spectrogram-CNN methods, and the SDP transformation requires significantly lower computational

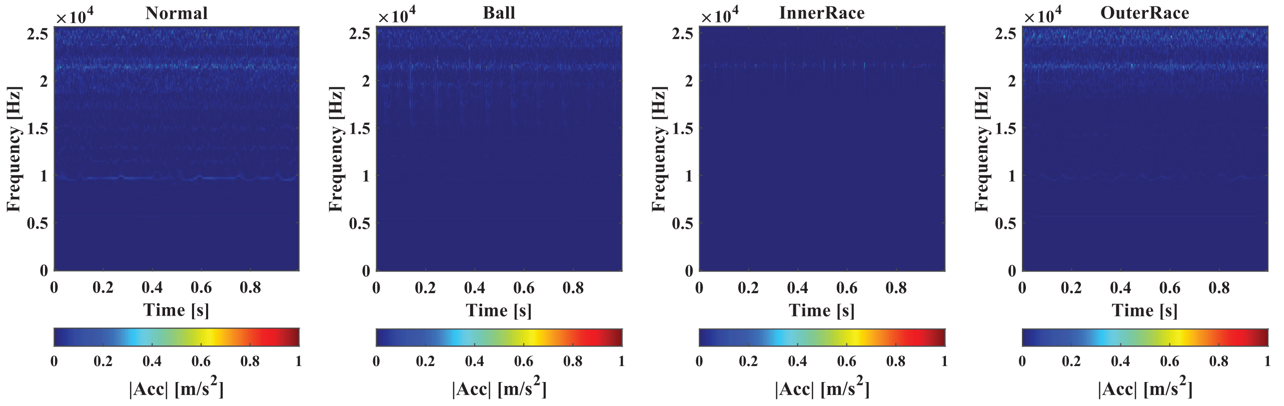

Fig. 14. SWT diagram for dataset 2.

Table VI. Optimized CNNs for dataset 2

SDP	SWT	Kurtogram	CWT	SWT
IML	IML	IML	IML	IML
C2DL(3,22,3)	C2DL(3,32,2)	C2DL(2,27,2)	C2DL(3,8,2)	C2DL(2,45,2)
BNL	BNL	BNL	BNL	BNL
RL	RL	RL	RL	RL
MP2DL(3,3)	MP2DL(4,4)	MP2DL(2,3)	MP2DL(3,3)	MP2DL(4,2)
FCL(99)	FCL(507)	FCL(776)	FCL(559)	FCL(474)
RL	RL	RL	RL	RL
FCL(339)	FCL(810)	FCL(265)	FCL(358)	FCL(919)
RL	RL	RL	RL	RL
FCL(354)	FCL(4)	FCL(529)	FCL(156)	FCL(824)
RL	SL	RL	RL	RL
FCL(4)	/	FCL(4)	FCL(4)	FCL(4)
SL	/	SL	SL	SL

Table VII. Dataset 2 results

Method	Accuracy [%]	Precision [%]	Recall [%]	F1-score [%]	Computational time [s]
SDP	98.18	96.77	96.77	96.77	0.26
STFT	96.77	92.50	97.37	94.87	1.18
Kurtogram	98.55	95.12	100	97.50	71.63
CWT	100	100	100	100	49.90
SWT	98.39	94.59	100	97.22	297.76

Table VIII. Dataset 2 stats

Method	Mean [s]	Std [s]	Median [s]
SDP	$4.402 * 10^{-4}$	$4.610 * 10^{-5}$	$4.298 * 10^{-6}$
STFT	0.002	$3.394 * 10^{-4}$	0.002
Kurtogram	0.146	0.031	0.151
CWT	0.106	0.005	0.105
SWT	0.652	0.022	0.648

Table IX. Dataset 2 paired t-tests

Method	h	p-Value	Test statistics	Std [s]
SDP-STFT	1	0	−103.787	$3.214 * 10^{-4}$
SDP-Kurtogram	1	0	−676.329	0.005
SDP-CWT	1	0	−425.156	0.005
SDP-SWT	1	0	−627.934	0.022

cost than other time–frequency transforms. In addition, CNNs trained with SDP images have consistently shown accuracy above 98% in the testing phase. The latter result necessitates finding a trade-off between computational time and accuracy since the best results in the test phase were achieved by the CWT for both datasets, but at the expense of a much higher computational time to obtain the CNN input.

The low computational time and high accuracy promote the SDP for a future study for real-time implementation of the technique over other time–frequency transforms.

In future research, it will be essential to thoroughly evaluate the performance of the SDP-CNN approach in comparison with spectrogram-CNN methods under more realistic and challenging conditions. This includes testing its robustness in dynamic operating environments where system parameters (such as load and speed) fluctuate, assessing its resilience to high levels of noise that may be present in practical applications, and extending the analysis to other types of mechanical systems, such as complex gear assemblies and other rotating machinery. Such comprehensive validation will provide a deeper understanding of the generalizability, reliability, and practical applicability of the SDP-CNN approach for real-world FD and monitoring tasks.

Nomenclature

IML	Image input layer
C2DL	Convolutional 2D layer
BNL	Batch normalization layer
RL	Relu layer
MP2DL	Max pooling 2D layer
FCL	Fully connected layer
SL	Softmax layer
C2DL(x,y,z)	C2DL(filters size, filters number, stride)
MP2DL(x,y)	MP2DL(pooling size, stride)
FCL(x)	FCL(numbers of output neurons)

FUNDING

This research received no external funding.

DATA AVAILABILITY STATEMENT

The raw data used in this research were public datasets described in [35,36]. The processed data are not publicly available but may be obtained from the corresponding author upon reasonable request.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

REFERENCES

[1] X. Chen *et al.*, “Large models for machine monitoring and fault diagnostics: Opportunities, challenges, and future direction,” *J. Dyn. Monit. Diagn*, vol. 4, no. 2, pp. 76–90, 2025.

[2] R. Yan *et al.*, “Transfer learning for prognostics and health management: Advances, challenges, and opportunities,” *J. Dyn. Monit. Diagn*, vol. 3, no. 2, pp. 60–82, 2024.

[3] O. Janssens *et al.*, “Convolutional neural network based fault detection for rotating machinery,” *J. Sound Vib.*, vol. 377, pp. 331–345, 2016.

[4] H. O. A. Ahmed and A. K. Nandi, “Vibration image representations for fault diagnosis of rotating machines: A review,” *Mach.*, vol. 10, no. 12, p. 1113, 2022.

[5] I. Lupea and M. Lupea, “Continuous wavelet transform and CNN for fault detection in a helical gearbox,” *Appl. Sci.*, vol. 15, no. 2, p. 950, 2025.

[6] L. H. Wang *et al.*, “Motor fault diagnosis based on short-time Fourier transform and convolutional neural network,” *Chin. J. Mech. Eng.*, vol. 30, pp. 1357–1368, 2017.

[7] T. D. Nguyen *et al.*, “A distinguished deep learning method for gear fault classification using time–frequency representation,” *Discover Appl. Sci.*, vol. 6, no. 7, pp. 340, 2024.

[8] E. Piedad *et al.*, Deep learning-based machine condition diagnosis using short-time fourier transformation variants. In 2024 International Conference on Diagnostics in Electrical Engineering (Diagnostika) (pp. 1-4). IEEE. Sep. 2024.

[9] Q. Xiao *et al.*, “Improved variational mode decomposition and CNN for intelligent rotating machinery fault diagnosis,” *Entropy (Basel)*, vol. 24, no. 7, p. 908, 2022.

[10] G. Fu, Q. Wei, and Y. Yang, “Bearing fault diagnosis with parallel CNN and LSTM,” *Math. Biosci. Eng: MBE*, vol. 21, no. 2, pp. 2385–2406, 2024

[11] Y. Xin *et al.*, “An intelligent fault diagnosis method of rotating machinery based on deep neural networks and time-frequency analysis,” *J. Vibroeng.*, vol. 20, no, 6, pp. 2321–2335, 2018.

[12] Z. Zhai *et al.*, “Rolling bearing fault diagnosis based on a synchrosqueezing wavelet transform and a transfer residual convolutional neural network,” *Sensors*, vol. 25, no. 2, p. 325, 2025

[13] H. Zhou *et al.*, “A rolling bearing fault diagnosis method based on a convolutional neural network with frequency attention mechanism. *Struct. Health Monit*, vol. 23, no. 4, pp. 2475–2495, 2024

[14] X. Pan *et al.*, “Real-time intelligent diagnosis of co-frequency vibration faults in rotating machinery based on lightweight-convolutional neural networks,” *Chin. J. Mech. Eng.*, vol. 37, no. 1, p. 41, 2024.

[15] C. A. Pickover, “On the use of symmetrized dot patterns for the visual characterization of speech waveforms and other sampled data,” *J Acoust Soc Am*, vol. 80, no. 3, pp. 955–960, 1986

[16] T. E. Wu, S. H. Huang, and C. H. Lai, “Helical gear defect detection system based on symmetrized dot pat-tern and convolutional neural network,” IEEE Access 2024.

[17] X. Zhu *et al.*, “Rotor fault diagnosis using a convolutional neural network with symmetrized dot pattern images,” *Measurement*, vol. 138, pp. 526–535, 2019

[18] S. C. Yen *et al.*, Convolutional Probabilistic Neural Network Approach for Fault Diagnosis in Wind Turbine Gearbox with Symmetrized Dot Patter. In 2024 10th International Conference on Applied System Innovation (ICASI) (pp. 01-03). IEEE. April. 2024

[19] S. D. Lu *et al.*, A Novel Hybrid CNN–KNN (CKNN) Approach for Robust Compound Fault Diagnosis in Wind Turbine Gearboxes Using Symmetrized Dot Pattern Analysis. In 2025 IEEE Industry Applications Society Annual Meeting (IAS) (pp. 1-8). IEEE. June. 2025.

[20] C. Zhang, W. Wang, and H. Li, “Tool wear prediction method based on symmetrized dot pattern and multi-covariance Gaussian process regression,” *Measurement*, vol. 189, p. 110466, 2022.

[21] X. Xu *et al.*, “Fan fault diagnosis based on symmetrized dot pattern analysis and image matching,” *J. Sound Vib.*, vol. 374, pp. 297–311, 2016.

[22] M. H. Wang, J. X. Hong, and S. D. Lu, “Fault diagnosis of lithium battery modules via symmetrized dot pattern and convolutional neural networks,” *Sensors*, vol. 25, no. 1, p. 94, 2024.

[23] X. Liu *et al.*, “Fault diagnosis of rolling bearings based on the improved symmetrized dot pattern enhanced convolutional

neural networks," *J. Vib. Eng. Technol.*, vol. 12, no. 2, pp. 1897–1908, 2024.

[24] J. D. Wu, W. J. Luo, and K. C. Yao, "Acoustic signal classification using symmetrized dot pattern and convolutional neural network," *Mach.*, vol. 10, no. 2, p. 90, 2022.

[25] W. Jia *et al.*, "A zero-shot bearing fault diagnosis framework utilizing spatial relationships among primary features under variable working conditions," *Mech. Syst. Sig. Process.*, vol. 238, p. 113103, 2025

[26] Y. Sun *et al.*, "Fault diagnosis of rolling bearing based on empirical mode decomposition and improved manhattan distance in symmetrized dot pattern image," *Mech. Syst. Sig. Process.*, vol. 159, p. 107817, 2021

[27] F. Yang *et al.*, "An optimized variational mode decomposition and symmetrized dot pattern image characteristic information fusion-based enhanced CNN ball screw vibration intelligent fault diagnosis ap-proach," *Measurement*, vol. 229, p. 114382, 2024

[28] J. Song *et al.*, "Multisource deep feature fusion of optimized symmetrized dot patterns for SRM fault diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–13, 2024

[29] J. Liu *et al.*, "Incipient fault detection of planetary gearbox under steady and varying condition," *Expert Syst. Appl.*, vol. 233, p. 121003, 2023

[30] M. Spirto *et al.*, "Enhancing SDP-CNN for gear fault detection under variable working conditions via multi-order tracking filtering," *J. Dyn. Monit. Diagn.*, 2025. https://doi.org/10.37965/jdmd.2025.813

[31] J. Hao *et al.*, "An adversarial gradual domain adaptation approach for fault diagnosis via intermediate domain generation," *Nondestruct. Test. Eval.*, pp. 1–21, 2025

[32] S. D. Lu *et al.*, A Novel Hybrid CNN–KNN (CKNN) Approach for Robust Compound Fault Diagnosis in Wind Turbine Gearboxes Using Symmetrized Dot Pattern Analysis. In 2025 IEEE Industry Applications Society Annual Meeting (IAS) (pp. 1-8). IEEE. June. 2025

[33] S. C. Yen *et al.*, Convolutional Probabilistic Neural Network Approach for Fault Diagnosis in Wind Turbine Gearbox with Symmetrized Dot Patter. In 2024 10th International Conference on Applied System Innovation (ICASI) (pp. 01-03). IEEE. April. 2024.

[34] X. Xu, M. Qi, and H. Liu, "Real-time stall detection of centrifugal fan based on symmetrized dot pattern analysis and image matching," *Measurement*, vol. 146, pp. 437–446, 2019

[35] K. Loparo, Case western reserve university bearing data center. Bearings Vibration Data Sets, Case Western Reserve University, 22-28. 2012

[36] N. D. Thuan and H. S. Hong, "HUST bearing: A practical dataset for ball bearing fault diagnosis," *BMC Res Notes*, vol. 16, no. 1, p. 138, 2023

[37] K. Wang *et al.*, 2023"An improved second-order multisynchrosqueezing transform for the analysis of non-stationary signals," *J. Dyn. Monit. Diagn.*, vol. 2, no. 3, pp. 183–189.

[38] J. Antoni and R. B. Randall, "The spectral kurtosis: Application to the vibratory surveillance and diagnostics of rotating machines," *Mech. Syst. Sig. Process.*, vol. 20, no. 2, pp. 308–331, 2006

[39] D. Helm and M. Timusk, "Wavelet denoising applied to hardware redundant systems for rolling element bearing fault detection," *J. Dyn. Monit. Diagn.*, vol. 2, no. 2, pp. 102–114, 2023.

[40] D. Zhang and Z. Feng, "Wind turbine planetary gearbox fault diagnosis via proportion-extracting synchrosqueezing chirplet transform," *J. Dyn. Monit. Diagn.*, vol. 2, no. 3, pp. 177–182, 2023.

[41] G. J. Jiang *et al.*, "Rolling bearing fault diagnosis based on convolutional capsule network," *J. Dyn. Monit. Diagn*, vol. 2, no. 4, pp. 275–289, 2023

[42] Z. Fan *et al.*, "New sensing technologies for monitoring machinery, structures, and manufacturing processes," *J. Dyn. Monit. Diagn.*, vol. 2, no. 2, pp. 69–88, 2023.

[43] J. González *et al.*, "Symmetrized dot pattern analysis for the unsteady vibration state in a Sirocco fan unit," *Appl. Acoust.*, vol. 152, pp. 1–12, 2019.