Experimental Verification of Vibration-Based Unbalance Detection in a PMSM Drive Using a CNN-Based Classification Framework

1st Dominik Łuczak
Poznan University of Technology
Poznan, Poland
dominik.luczak@put.poznan.pl
ORCID: 0000-0002-2755-5503

2nd Wiktor Nowacki
Poznan University of Technology
Poznan, Poland
wiktor.nowacki@student.put.poznan.pl
ORCID: 0009-0002-8254-8717

3rd Bartłomiej Wicher
Poznan University of Technology
Poznan, Poland
bartlomiej.wicher@put.poznan.pl
ORCID: 0000-0001-7144-0001

***Abstract*—This paper presents an experimental verification of a convolutional neural network (CNN) method for vibration-based unbalance diagnosis (also known as imbalance diagnosis) in a PMSM drive test stand. Vibration data were acquired using an industrial accelerometer and converted into grayscale image representations for CNN classification using a band-pass-filter-based signal-to-image pipeline. The study considers three variants of the moment of inertia, each measured at three sampling frequencies (1, 8, and 48 kHz) and under two operating conditions: normal and with added unbalance on a metal disk. The results confirm high classification accuracy for the training data at 900 rpm, but reduced performance at unseen speeds (800 and 1000 rpm), indicating limited generalization across operating points and the need for improved robustness of the method.**

***Index Terms*—unbalance diagnosis, imbalance diagnosis, vibration-based diagnosis, CNN, PMSM drive, moment of inertia, signal to image transformation.**

I. Introduction

Mechanical unbalance (also known as imbalance) is a common source of vibration in rotating electrical drives and may degrade efficiency, increase bearing load, and shorten machine lifetime. Reliable early detection of unbalance from measured vibration is therefore important for condition monitoring and maintenance planning.

Traditional vibration-diagnosis methods are often based on handcrafted time- or frequency-domain descriptors. Although effective in selected regimes, such descriptors may require expert tuning and can be sensitive to operating-point changes. In contrast, image-based representations combined with convolutional neural networks can learn discriminative structures directly from measured waveforms without manual feature engineering, in line with recent advances in representation learning [1]–[3].

Recent advances in deep learning for rotating machinery diagnostics have demonstrated the effectiveness of convolutional neural networks (CNNs) for image-based classifiers applied to vibration signals and their image-domain representations, including decomposition-based preprocessing, image embedding, and hybrid CNN architectures [4]–[7]. Condition-monitoring surveys provide a broader context for signal processing and feature extraction strategies [8], [9].

In contrast to many existing studies that focus on a single mechanical configuration, practical drive systems frequently operate with varying moments of inertia due to interchangeable loads or mechanical reconfiguration. Variations in inertia influence vibration propagation and may affect diagnostic performance if not properly accounted for. However, systematic experimental validation of vibration-based CNN classifiers across multiple inertia configurations remains limited in the literature, especially for PMSM drive systems operating under realistic laboratory conditions.

To address this gap, this paper presents an experimental verification of a vibration-based unbalance classification method using a CNN and grayscale signal-to-image representations derived from band-pass-filtered vibration data. The method is evaluated on a PMSM laboratory stand for three distinct moment-of-inertia variants, each measured in both balanced and imbalanced conditions. By training the CNN at one reference operating speed and testing it on unseen speeds, the study assesses classification accuracy and the extent of generalization across neighboring steady-state operating points. The underlying band-pass-filter signal-to-image concept was introduced in [10]; the present work provides its empirical verification for a PMSM drive using a professional vibration sensor. The results confirm that the proposed vibration-image + CNN pipeline can detect unbalance across differing inertia configurations, while also revealing limitations of cross-speed generalization that must be addressed before wider deployment.

II. Measurement Context and Test Cases

The study focuses on vibration-based unbalance diagnosis in a PMSM drive setup. Measurements were acquired in six conditions, formed by combining three moment-of-inertia variants (case 0, case 1, case 2) with two states in each variant: normal and unbalance (imbalanced).

This research was funded in whole by National Science Centre, Poland 2024/53/B/ST7/01308.

For clarity, the variants correspond to increasing total inertia of the drive setup: case 0 denotes the baseline (lowest-inertia) configuration, case 1 denotes an intermediate-inertia configuration, and case 2 denotes the highest-inertia configuration.

Imbalance was introduced by adding mass on a metal disk, while the normal condition corresponds to operation without the added imbalance mass.

Vibration was measured using a Digiducer 333D05 sensor mounted on the laboratory stand. The resulting waveforms were used as the only diagnostic input to the CNN classifier.

III. Vibration-Based Classification Method

The classification pipeline converts vibration time series into compact greyscale images used for CNN classification.

A. Signal-to-Image Encoding

The signal-to-image processing chain, illustrated in Fig. 1, employs a sliding window of 841 samples at a sampling frequency of 8 kHz. This window length corresponds to slightly more than one mechanical revolution at the reference speed of 900 rpm. Consecutive windows are shifted by 421 samples (50% overlap), providing dense temporal coverage while limiting redundancy.

B. Band-Pass Filter Signal Mapping

Within each analysis window, the vibration signal is processed using a bank of band-pass filters as depicted in Fig. 1. The filter outputs are arranged into a two-dimensional grayscale images of size 29×29 which all gathered together serves as the CNN input. Image size depends on the number of filters used, in this case 9, and is calculated as 29×3 to have rectangular images. in other words input images consist of 9 sub images of size 29×29 pixels at 8 kHz sampling frequency. The grid size is selected to capture band-limited vibration characteristics associated with imbalance while maintaining a compact image representation suitable for efficient convolutional processing.

Figure 1 illustrates the full processing chain with band-pass filtering, channel mapping, and CNN-based classification.

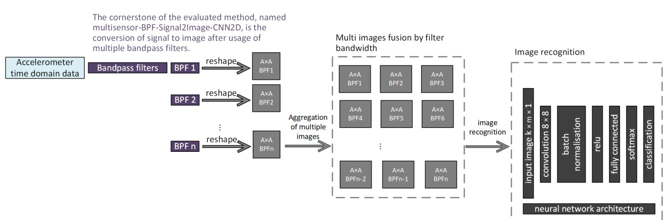


Fig. 1. Signal-to-image method with band-pass filtering used in this work for vibration-based imbalance classification.

Representative examples of the generated band-pass-filter image tiles for channel 1 are shown in Figs. 15–17. Each grid is organized by operating condition (imbalance in the top row, normal in the bottom row) and by inertia variant (case 0–2 in columns). The visual patterns confirm that the BPF mapping preserves class-dependent texture differences while maintaining consistent structure across inertia variants.

At 1 kHz and 8 kHz, the grids exhibit richer multi-band texture over a larger image area, while at 48 kHz the most discriminative structure is concentrated in the upper region of each tile. This frequency-dependent behavior is consistent with the selected filter-bank mapping and supports using sampling-frequency-specific data in CNN evaluation.

The resulting grayscale representation combines band-limited vibration information in a compact two-dimensional structure suitable for convolutional feature extraction.

IV. Experimental Setup

Experiments were conducted on a PMSM laboratory stand configured with three different moments of inertia (case 0, case 1, and case 2). For each configuration, two operating states were measured: a balanced reference condition and an imbalanced condition created by adding mass to a shaft-mounted metal disk. Vibration signals were acquired using a Digiducer 333D05 accelerometer and sampled at 1 kHz, 8 kHz, and 48 kHz for a duration of 20 s per measurement, yielding 20,000, 160,000, and 960,000 samples, respectively.

A. Additional Case Visualizations

Figures 3–5 present additional samples for cases 0–2 in normal and imbalance conditions.

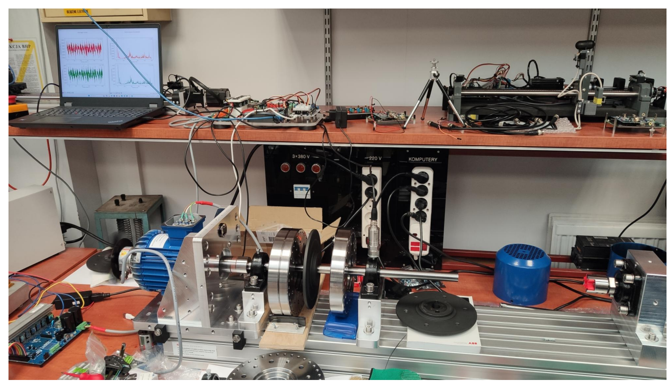
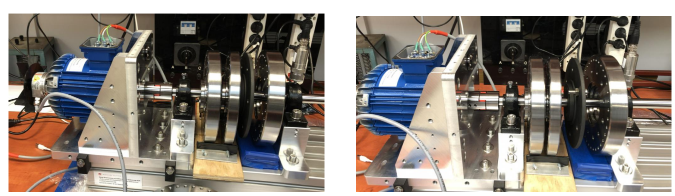


Fig. 2. Laboratory setup used for experimental verification, including the inverter platform, PMSM drive, Digiducer 333D05 vibration sensor, and workstation used for supervision and data acquisition.

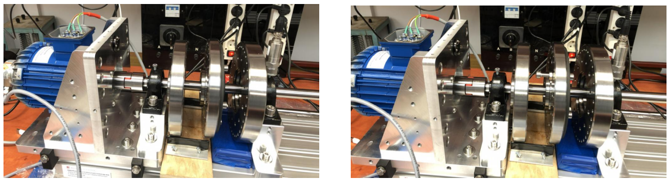


(a) Case 0 normal (b) Case 0 imbalance

Fig. 3. Case 0 visualizations for normal and imbalance conditions.

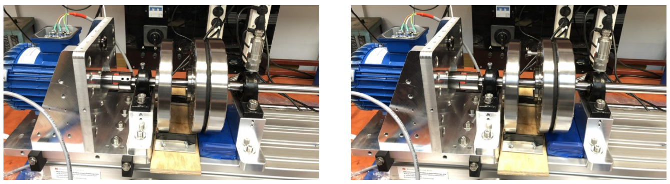
B. Data Volume

After sliding-window segmentation and signal-to-image conversion, the number of grayscale images per condition depends on the selected sampling frequency (1 kHz, 8 kHz, or 48 kHz), while the acquisition duration is fixed at 20 s per



(a) Case 1 normal (b) Case 1 imbalance

Fig. 4. Case 1 visualizations for normal and imbalance conditions.



(a) Case 2 normal (b) Case 2 imbalance

Fig. 5. Case 2 visualizations for normal and imbalance conditions.

measurement. The dataset includes all six conditions (case 0, case 1, case 2, each in normal and imbalance condition).

C. Hardware Platform

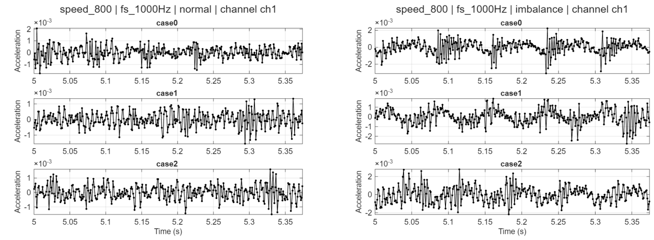
The laboratory setup used for data acquisition is shown in Fig. 2. The rig is based on an STM32 NUCLEO-F303 controller with STEVAL-IHM023V3 and X-NUCLEO-IHM09M2 motor-control expansion boards. The power stage supplies the blue PMSM drive visible in the photograph. Vibration data were captured with a Digiducer 333D05 sensor.

The project was executed using a Lenovo ThinkPad P16 Gen 2 workstation with a QHD display, equipped with an Intel Core i7-14700HX processor featuring 20 cores (8 P-cores and 12 E-cores, 28 threads), 128 GB RAM, and two 1 TB storage drives. The system incorporated an NVIDIA RTX 3500 GPU with 5120 CUDA cores and 12 GB GDDR6 memory, providing hardware-accelerated computation, and operated on Windows 11 Pro.

D. Representative Speed-Based Samples

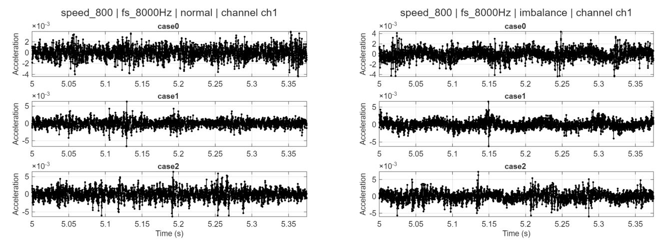
Figures 6–14 show representative normal and imbalance image samples for all measured speed and sampling-frequency combinations used in this study. Plots start at 5 s and cover a 70 ms interval, which corresponds to slightly more than one mechanical revolution at the reference speed of 900 rpm. The images confirm that the imbalance condition produces clear changes in vibration patterns compared with the normal condition, and that these changes are preserved across different operating speeds and sampling frequencies. The observed differences in amplitude and waveform shape between normal and imbalance conditions provide discriminative features that the CNN can exploit for accurate classification. The consistency of these patterns across the tested speeds supports the generalization capability of the proposed method.

Tables I and II summarize the acquisition parameters and PMSM characteristics. Each sample vibration segment is mapped to one grayscale image of size 29×29 pixels at 8 kHz



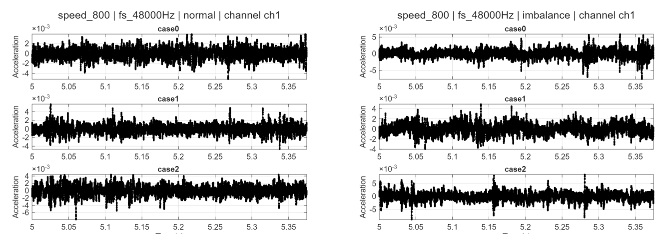
(a) Normal (b) Imbalance

Fig. 6. Speed 800, sampling frequency 1000 Hz.



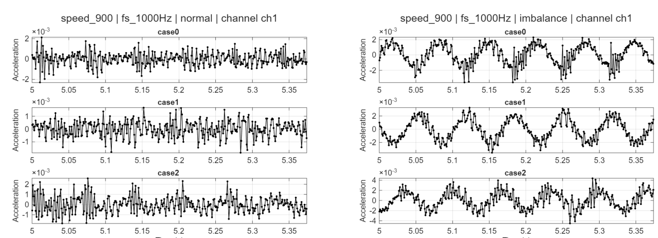
(a) Normal (b) Imbalance

Fig. 7. Speed 800, sampling frequency 8000 Hz.



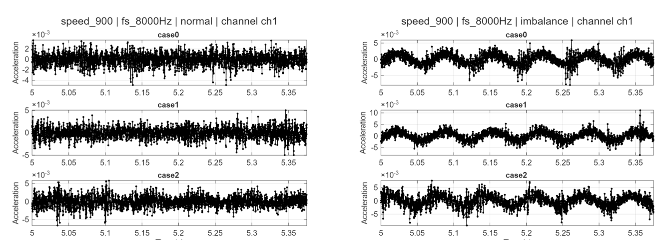
(a) Normal (b) Imbalance

Fig. 8. Speed 800, sampling frequency 48000 Hz.



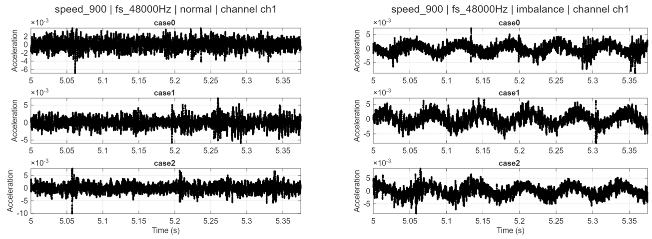
(a) Normal (b) Imbalance

Fig. 9. Speed 900, sampling frequency 1000 Hz.



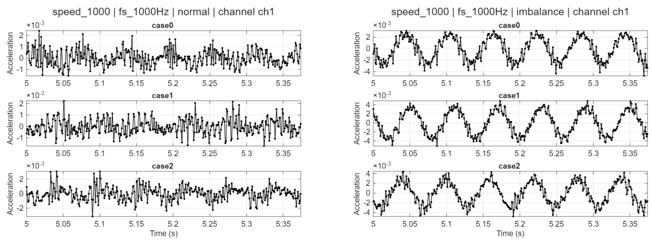
(a) Normal (b) Imbalance

Fig. 10. Speed 900, sampling frequency 8000 Hz.



(a) Normal (b) Imbalance

Fig. 11. Speed 900, sampling frequency 48000 Hz.



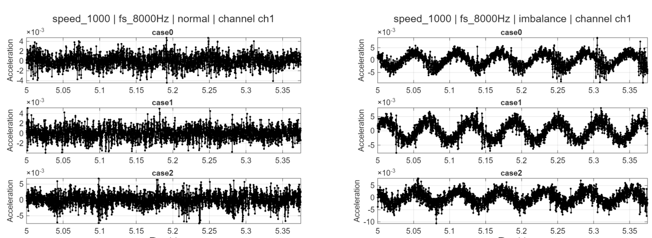
(a) Normal (b) Imbalance

Fig. 12. Speed 1000, sampling frequency 1000 Hz.

sampling frequency. The dataset includes case 0, case 1, and case 2, each in normal and imbalance conditions.

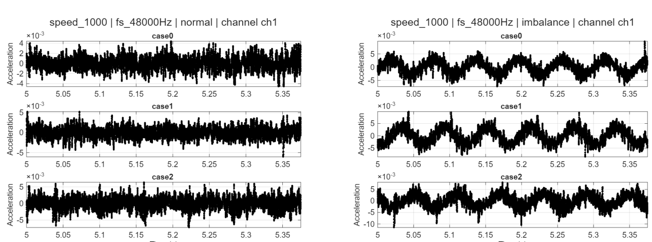
V. Results and Discussion

The CNN achieved very high training and validation accuracy at the reference speed of 900 rpm, but the test results at the unseen 800 rpm and 1000 rpm operating points were clearly less favorable. Representative confusion matrices are shown in Figs. 18–20. These plots indicate that the proposed



(a) Normal (b) Imbalance

Fig. 13. Speed 1000, sampling frequency 8000 Hz.



(a) Normal (b) Imbalance

Fig. 14. Speed 1000, sampling frequency 48000 Hz.

TABLE I
Verification Summary

Item	Value	Description
Data length	20,000 / 160,000 / 960,000	20 s at 1 / 8 / 48 kHz
Images per condition	399	one class at 1 kHz and one case
Images per full set	2,394	normal and imbalance classes at 1 kHz (single speed) for case 0–2
Train/validation set	speed 900	all cases/states, 0.8/0.2 split
Test set	speeds 800 and 1000	all cases/states, no split
Max epochs	10	training configuration
Initial learning rate	1e-4	training configuration

TABLE II
PMSM Parameters

Parameter	Value
Rated power	0.75 kW
Rated speed	1500 rpm (≈157.08 rad/s)

representation is highly discriminative for the training distribution, yet its cross-speed generalization remains incomplete.

Figures 18, 19, and 20 present training and test confusion matrices for three sampling frequencies. Training confusion matrices are dominated by the diagonal, indicating clear class separation at 900 rpm. In contrast, test results show increased off-diagonal errors, particularly imbalance samples misclassified as normal—most pronounced at 8 and 48 kHz, where performance approaches random balance. This indicates a clear train–test mismatch rather than robust generalization.

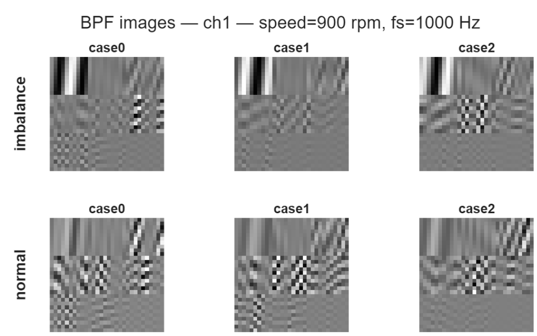


Fig. 15. Band-pass-filter image grid for channel 1 at the sampling frequency of 1000 Hz (rows: imbalance/normal, columns: case 0–2).

A. Sampling-Frequency-Specific CNN Results

Additional CNN results obtained for the band-pass-filter image representation on channel 1 are shown in Figs. 18–20. These plots provide a frequency-specific view of classification quality for sampling frequencies of 1 kHz, 8 kHz, and 48 kHz. For all three sampling frequencies, the training confusion matrices remain strongly diagonal, but the test matrices exhibit

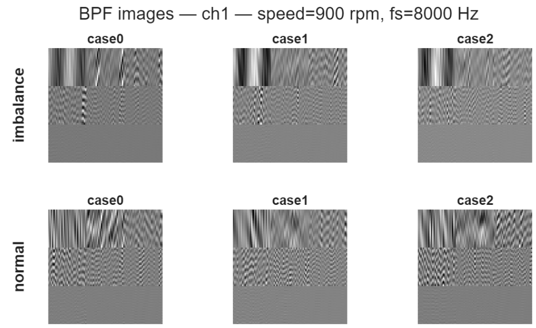


Fig. 16. Band-pass-filter image grid for channel 1 at the sampling frequency of 8000 Hz (rows: imbalance/normal, columns: case 0–2).

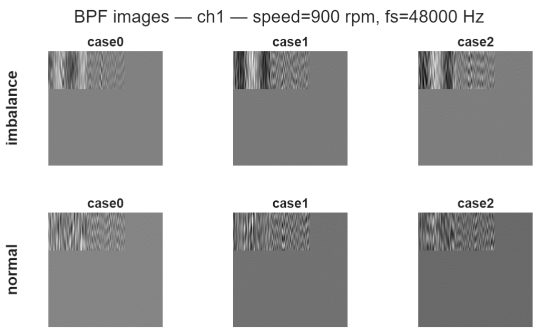
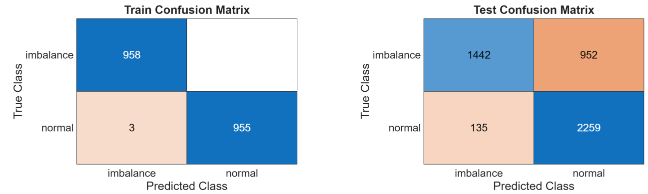


Fig. 17. Band-pass-filter image grid for channel 1 at the sampling frequency of 48000 Hz (rows: imbalance/normal, columns: case 0–2).

a noticeable loss of separability when the model is evaluated at unseen speeds.

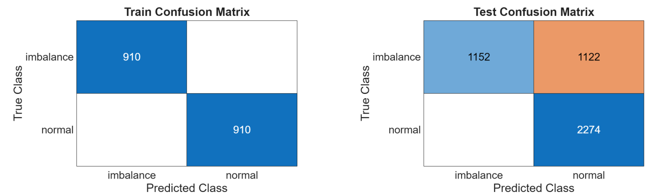
At 1 kHz, off-diagonal errors remain, but the diagonal structure is still clearer than at higher rates. At 8 and 48 kHz, imbalance false negatives increase, indicating that added high-frequency content does not improve robustness; instead, it raises computational cost without enhancing cross-speed performance. An important practical observation is that increasing the sampling frequency did not provide a clear improvement in classification performance, while it increased computational load (more samples to process, larger intermediate arrays, and longer training/inference time). For fair comparison across sampling frequencies, image generation was performed for the same time interval (70 ms). Because the number of samples in 70 ms depends on the sampling frequency, the generated image size changed accordingly (e.g., 70 samples at 1 kHz, 560 samples at 8 kHz, and 3360 samples at 48 kHz before mapping), which directly affects computation cost.

Representative vibration waveforms exhibit temporal patterns used for grayscale image generation. Compared with the normal condition, imbalance cases exhibit clear changes in waveform amplitude and shape that are exploited by the CNN for accurate classification. The method was examined for all three inertia variants (case 0, case 1, case 2), each with and without imbalance, and produced consistent discrimination



(a) Training (b) Test

Fig. 18. Band-pass-filter CNN confusion matrices for channel 1 at the sampling frequency of 1000 Hz.



(a) Training (b) Test

Fig. 19. Band-pass-filter CNN confusion matrices for channel 1 at the sampling frequency of 8000 Hz.

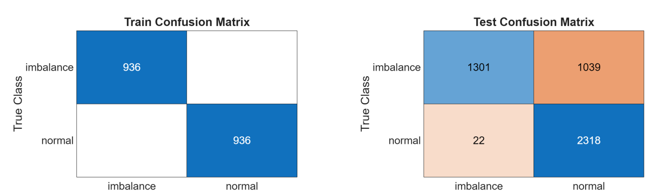
performance across tested operating speeds.

B. Training Behaviour

The training process exhibited stable convergence for all evaluated sampling frequencies. Both training and validation accuracy increased rapidly during the initial epochs and reached plateau values close to 100%, while the loss function decreased monotonically. These observations are consistent with the confusion matrix results and confirm that the proposed signal-to-image representation provides discriminative features for CNN-based imbalance classification. An additional training trace for the band-pass-filter CNN using channel 1 at the sampling frequency of 1 kHz is shown in Fig. 21. The plot confirms a monotonic reduction of the loss function together with a rapid increase in training and validation accuracy toward 100%, which is consistent with the confusion-matrix results reported above.

Near-perfect training performance with weaker test results indicates that convergence does not guarantee generalization across operating points; although optimization is stable, the learned representation transfers imperfectly to unseen speeds.

Training was performed in MATLAB Deep Learning Toolbox using the stochastic gradient descent with momentum



(a) Training (b) Test

Fig. 20. Band-pass-filter CNN confusion matrices for channel 1 at the sampling frequency of 48000 Hz.

optimizer. Execution was forced on GPU, and the built-in training-progress plot was used for monitoring.

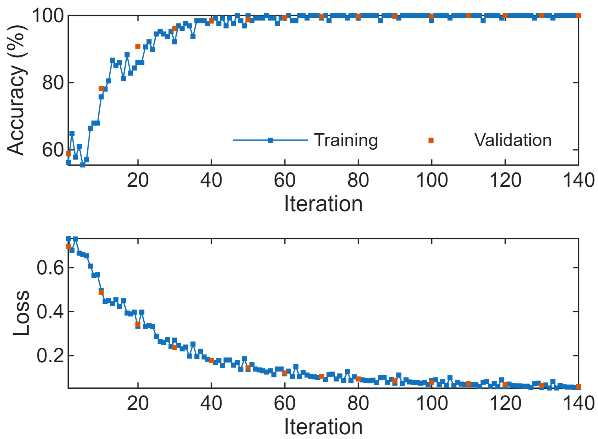


Fig. 21. Training progress of the band-pass-filter CNN for channel 1 at the sampling frequency of 1000 Hz.

C. Scope, Limitations, and Practical Implications

CNN-based image representations generally outperform classical feature-based classifiers (e.g., decision trees, SVM, k-NN) in fault diagnosis, as also indicated in [11], likely due to their ability to capture spatial vibration patterns. However, no direct benchmark on the same PMSM dataset is provided here, so this advantage remains indicative rather than experimentally confirmed; a like-for-like comparison will be included in future work. The evaluation is limited to three steady-state speeds (train: 900 rpm; test: 800 and 1000 rpm), revealing a train–test gap but not ensuring robustness over a wider operating range. Transient regimes were not considered and are expected to introduce additional non-stationarity. Only a single imbalance magnitude was analyzed, leaving sensitivity to incipient faults unverified. Future studies will include multiple severity levels to determine detection limits and confidence scaling. A sampling rate of 1 kHz offers the best cost–performance trade-off; higher rates (8 and 48 kHz) increase computational load without improving accuracy for the 70 ms window. Finally, results are based on a controlled laboratory setup. Variability in industrial environments may affect performance, requiring further validation and possibly domain adaptation for practical deployment.

VI. Conclusion

This paper presented an experimental verification of a CNN-based method for vibration-based imbalance diagnosis using grayscale signal-to-image representations. The approach was validated on vibration data acquired from a PMSM drive operating with three different moments of inertia, each evaluated under balanced and imbalanced conditions. The results confirm that the band-pass-filter image representation combined with CNN classification is highly effective for the trained operating point and remains promising for cross-speed diagnosis, but the confusion matrices also reveal a substantial reduction in test performance at unseen speeds, especially for 8 kHz and 48 kHz. Future work will therefore focus on direct comparison with classical classifiers, additional imbalance magnitudes, broader steady-state and transient speed regimes, validation on more realistic industrial environments, and deployment on embedded platforms.

Data Availability

The dataset used in this study will be publicly available on Zenodo [12] after the publication of this work.

Author Contributions

D. Łuczak: conceptualization and methodology, principal investigation, experimental design, data curation, model training, and manuscript writing. W. Nowacki: data collection and laboratory stand setup. B. Wicher: laboratory stand setup and commissioning.

Acknowledgment

This research was funded in whole by National Science Centre, Poland, grant no. 2024/53/B/ST7/01308. The authors gratefully acknowledge the financial support provided the project entitled *"Step Towards Self-Aware Electric Drives"*, which made this study possible.

References

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[2] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, 2013, doi: 10.1109/TPAMI.2013.50.

[3] R. B. Randall and J. Antoni, "Rolling element bearing diagnostics—A tutorial," *Mech. Syst. Signal Process.*, vol. 25, no. 2, pp. 485–520, 2011, doi: 10.1016/j.ymssp.2010.07.017.

[4] F. Yang, X. Tian, L. Ma, *et al.*, "An optimized variational mode decomposition and symmetrized dot pattern image characteristic information fusion-based enhanced CNN ball screw vibration intelligent fault diagnosis approach," *Measurement*, vol. 233, p. 114382, 2024, doi: 10.1016/j.measurement.2024.114382.

[5] Y. Kim and H. Yoon, "Integrating vibration signal analysis and image embedding for enhanced bearing fault diagnosis in manufacturing," *Scientific Reports*, vol. 15, 2025, doi: 10.1038/s41598-025-94351-0.

[6] F. Cong, Q. Bo, Y. Fan, *et al.*, "A novel multiscale petal pattern and CNN-BiLSTM structure for acoustic signal-based rolling bearing fault diagnosis," *Journal of Vibration and Acoustics*, vol. 147, no. 4, 2025, doi: 10.1115/1.4069853.

[7] J. Liu, R. Hao, and F. Deng, "Gearbox fault diagnosis method based on RIME-VMD and CNN-BiLSTM," *Journal of Vibration and Control*, 2026, doi: 10.1177/10775463261438471.

[8] I. Howard, "A review of rolling element bearing vibration detection, diagnosis and prognosis," *Mech. Syst. Signal Process.*, vol. 47, no. 1–2, pp. 1–29, 2014, doi: 10.1016/j.ymssp.2014.02.005.

[9] R. B. Randall, *Vibration-Based Condition Monitoring*, John Wiley & Sons, Chichester, UK, 2011, doi: 10.1002/9780470977668.

[10] D. Łuczak, "Data-driven rotary machine fault diagnosis using multisensor vibration data with band-pass filtering and convolutional neural network for signal-to-image recognition," *Electronics*, vol. 13, no. 15, p. 2940, 2024, doi: 10.3390/electronics13152940.

[11] D. Łuczak, S. Brock, and K. Siembab, "Fault detection and localization of a three-phase inverter with PMSM load using a convolutional neural network," *Actuators*, vol. 12, no. 3, 2023, doi: 10.3390/act12030125.

[12] D. Łuczak, W. Nowacki, and B. Wicher, "PMSM Vibration Dataset for Unbalance Diagnosis at Rotational Speeds 800, 900, and 1000 rpm" Zenodo, Apr. 24, 2026, doi: 10.5281/zenodo.19748543.