

Video-Based Vibration Analysis for Predictive Maintenance: A Motion Magnification and Random Forest Approach

Walid Gomaa^{1,2}^a, Abdelrahman Wael Ammar¹^b, Ismael Abbo¹^c, Mohamed Galal Nassef¹^d,
Tetsuji Ogawa³^e and Mohab Hossam¹^f

¹Faculty of Engineering, Egypt-Japan University of Science and Technology, Alexandria, Egypt

²Faculty of Engineering, Alexandria University, Alexandria, Egypt

³Graduate School of Fundamental Science and Engineering, Waseda University, Tokyo, Japan

Keywords: Condition Monitoring, Machinery, Video Motion Magnification, Machine Learning, Non-Contact Approach.

Abstract: Condition monitoring of high-speed machinery is critical to prevent unexpected breakdowns that could lead to injuries and cost billions. Traditional contact-based vibration sensors face limitations including measurement perturbations, point-specific data coverage, and installation constraints. This paper presents a novel non-contact machinery fault detection framework combining Eulerian video motion magnification with machine learning classification. The methodology comprises two integrated components. Primarily, a video-based vibration analysis pipeline utilizing Eulerian motion magnification with dense optical flow, which accomplish comprehensive signal processing for feature extraction using Fast Fourier transform. Then, a Random Forest classifier trained on video-derived temporal and frequency domain features. The system was validated based on ground-truth data from the Gunt PT500 machinery diagnosis and the Gunt TM170 balancing apparatus under four operational conditions: normal operation, outer ring bearing fault, and two imbalance severities (10g and 37g). Hence, experimental results demonstrate exceptional performance with 96.7% overall accuracy and a macro-averaged F1-score of 0.965 in discriminating fault conditions using solely video-derived features. The video processing allowed to identify distinct vibration signatures, from imbalance conditions showing amplitude variations to proportional fault severity ultimately offering a cost-effective solution for industrial condition monitoring applications.

1 INTRODUCTION

Modern manufacturing depends heavily on complex machinery, making reliable condition monitoring a cornerstone of operational efficiency. Unplanned equipment failures can halt production, incur safety risks, and generate substantial costs—estimated at \$28.6 billion per year in Australia alone (Li et al., 2025; de Koning et al., 2024). Early detection of issues such as overheating, fluid leaks, or drive-belt slippage is therefore critical to avert these disruptions.

A range of diagnostic techniques has emerged, including particle debris analysis, acoustic emission

monitoring, and vibration signal inspection (Zhong et al., 2025; Xu and Lu, 2025). Among these, vibration and acoustic methods are especially attractive because they permit online health assessment without machine disassembly. However, these contact sensors suffer from alignment errors and environmental noise, which can degrade accuracy by up to 15% (Kalaiselvi et al., 2018; Kiranyaz et al., 2024). These limitations have spurred interest in non-contact alternatives.

Video Motion Magnification (VMM) acts like a “motion microscope,” amplifying minute movements in video sequences to reveal hidden dynamics (Lima et al., 2025). Its applications span medical diagnostics (Nieto, 2025), structural health and condition monitoring (Lado-Roigé and Pérez, 2023; Yang and Jiang, 2024; Zhang et al., 2025). Initial VMM implementations relied on Lagrangian tracking which are computationally heavy with optical-flow methods sensitive to occlusion and complex motion (Śmieja et al., 2021). Eulerian techniques then emerged, de-

^a <https://orcid.org/0000-0002-8518-8908>

^b <https://orcid.org/0009-0005-8695-9144>

^c <https://orcid.org/0009-0003-4834-8962>

^d <https://orcid.org/0000-0002-3192-7154>

^e <https://orcid.org/0000-0002-7316-2073>

^f <https://orcid.org/0000-0003-2075-2050>

tecting pixel-level intensity changes for faster processing, though prone to blurring and noise artifacts that could reduce magnification fidelity by roughly 20% (Wu et al., 2012; Śmieja et al., 2021). More recently, data-driven models using Convolutional Neural Networks (CNNs) or Swin Transformers have learned optimized filters, delivering sharper results and up to 30% better noise suppression, yet often depend on synthetic training sets that may not capture real-world variability (Lado-Roigé et al., 2023; Lado-Roigé and Pérez, 2023; Oh et al., 2018).

In this work, we build upon Eulerian VMM by integrating enhanced optical-flow estimation and tailored feature extraction to strengthen motion amplification. Coupling these video-derived features with a Random Forest classifier trained on ground-truth vibration measurements from a rotating-machinery test rig, we introduce a machine learning-driven, non-contact vibration analysis framework. Our approach aims to overcome the drawbacks of traditional contact sensors and synthetic-data bias, offering a robust solution for continuous condition monitoring in industrial environments.

The remaining sections are organized as follows. We present related works in section 2, followed by a detailed scaffold of our methodology in section 3. Then we present, illustrate and discuss our results in section 4. Finally, we acknowledge perspectives in section 5.

2 RELATED WORKS

According to the review (Śmieja et al., 2021), VMM techniques are commonly divided into two main categories: Eulerian and Lagrangian approaches. Parallel in fluid dynamics, Eulerian methods observe changes within a stationary volume here, by decomposing video sequences into motion-related representations, which are then mathematically enhanced and reconstructed into magnified output frames. On the other hand, Lagrangian methods trace elements as they move here, focus on tracking individual pixel or feature trajectories across time. Hence, Eulerian approaches are particularly effective for detecting subtle motions but can introduce blur when applied to larger movements.

Eulerian VMM involves three stages: spatial decomposition, motion extraction and amplification, and noise suppression. The VMM problem can be formulated following (Oh et al., 2018; Lado-Roigé et al., 2023): given an image signal $I(x;t)$, where x denotes spatial location and t time, and a displacement function $\delta(t)$ representing temporal motion, the

original signal function derived from the original image $f(x)$ is expressed in Equation 1.

$$I(x;t) = f(x + \delta(t)) \quad I(x;0) = f(x) \quad (1)$$

The goal is to synthesize a magnified version of the signal expressed in Equation 2.

$$\hat{I}(x;t) = f(x + (1 + \alpha) \cdot \delta(t)) \quad (2)$$

Here, α denotes the amplification factor. In practice, only motions within certain frequency ranges are meaningful, and thus a temporal band-pass filter $T(\cdot)$, is applied to isolate those components of $\delta(t)$.

Before learning-based advances, most VMM methods relied on multi-frame temporal filtering to extract and enhance relevant motion signals (Wu et al., 2012; Wadhwa et al., 2013; Wadhwa et al., 2014). Notably, (Wadhwa et al., 2013) proposed an Eulerian magnification approach using local phase variations which extracts motion through a complex steerable pyramid and multi-scale orientation filters.

Oh et al. shifted to a CNN-driven pipeline that learns motion features end-to-end, removing the need for explicit temporal filtering (Oh et al., 2018). Due to scarce real videos, they trained on synthetically generated sequences with controlled sub-pixel displacements, enabling adaptive, data-driven magnification. These allowed the application of VMM techniques in areas such as Structural Health Monitoring (SHM) and machinery fault diagnosis.

Hence, Yang et al. augmented phase-based VMM (PVMM) with a gradient-domain filter (GDGIF) (Yang and Jiang, 2024). In SHM experiments against the original PVMM method (Wadhwa et al., 2013), their approach raised the SSIM to 0.9222 and the PSNR to 35.2236dB. However, these filters require manual tuning for the relevant frequency band, making the process laborious. To address this drawback, the work of (Lado-Roigé et al., 2023) followed the LB-VMM framework targeting micro-displacement bands with minimal prior information. Under an 20% damage condition, frequency estimates deviated only by 0.01% from accelerometer readings.

Likewise, machinery fault diagnosis has embraced VMM. (Zhao et al., 2023) proposed a luminance-based magnification framework to detect rotor unbalance, loose bearings, and misalignment by amplifying one to two times rotational frequencies. A rotor unbalance appeared as axial shaking, loosened bolts as vertical displacement, and misalignment as axial motion. Separately, a Swin Transformer-based method used Residual Swin Transformer Blocks to enhance magnification (Lado-Roigé and Pérez, 2023). Though fine-tuned on real videos and yielding a 9.63% boost in MUSIQ (Ke et al., 2021) scores, its initial training relied on synthetic datasets like PASCAL VOC

and MS COCO (Lin et al., 2014; Everingham et al., 2010). Similar caveats apply to LB-VMM and the band-passed acceleration methods (Oh et al., 2018; Lado-Roigé et al., 2023; Zang et al., 2025). While synthetic data simplifies controlled testing, it cannot fully replicate real-world lighting, texture, and motion artifacts.

3 METHODOLOGY

This study proposes a video-based approach for machinery fault detection and condition monitoring, consisting of two main components: (1) a non-contact video motion magnification pipeline for extracting vibration signals, and (2) a Random Forest classifier using features derived from the magnified video data.

3.1 Dataset Collection

All extracted features are derived from the working conditions of two rotating machines. The first is the Machinery Diagnosis System, Gunt PT500, which supports vibration measurement and field balancing tasks (Song et al., 2025; Gunt, 2025a). Three operating conditions were tested: a normal condition with healthy components and no added mass, a faulty condition with outer ring bearing damage, and a weighted condition where a 10g mass was added at a 90° angle on the flywheel, as shown in Figure 1a.

The second setup is the Balancing Apparatus Gunt TM170, used to illustrate imbalance scenarios (Gunt, 2025b). It includes weighted condition where 37g is placed across different slots on the balancing foundation (Figure 1b). These configurations serve as the basis for video collection and subsequent feature extraction through motion magnification.

3.2 Our Proposed Framework

The overall methodology, as depicted in Figure 2 is structured as an integrated pipeline comprising sequential processing stages that work in tandem.

3.2.1 Motion Magnification

The primary component of our methodology introduces an approach to vibration monitoring that eliminates the need for physical sensor contact with machinery. Here, the computer vision-based system leverages Eulerian Motion Magnification as the cornerstone step, combined with dense optical flow extraction to capture vibration signatures directly from video recordings.

The process begins with an interactive region of interest (ROI) selection corresponding to the vibrating machine parts. The mathematical representation of ROI is defined in Equation 3.

$$I_{ROI}(x, y, t) = I(x_0 + x, y_0 + y, t) \quad (3)$$

Where (x_0, y_0) represents the top-left corner of the selected ROI, and (x, y) are the relative coordinates within the cropped region of the initial frames from Equation 1.

Given an input video sequence $I(x, t)$ as presented in Equation 1, the Eulerian magnification process is applied on a region of interest (ROI) of the frames to obtain an output similar to Equation 2. Then, the dense optical flow computation quantifies pixel-wise motion vectors between consecutive frames within the ROI as temporal displacement fields vibration patterns. The motion magnitude is computed as the Euclidean norm of the flow vectors, and the spatial averaging across the ROI produces the final motion time-series as presented in Equation 4.

$$S(t) = \frac{1}{|ROI|} \sum_{(x,y) \in ROI} \sqrt{u(x, y, t)^2 + v(x, y, t)^2} \quad (4)$$

Where u, v are the optical flow vectors, and x, y are the relative spatial coordinates.

3.2.2 Feature Extraction

The extracted motion timeseries undergoes processing to generate discriminative features for fault classification. The preprocessing pipeline includes DC component removal by subtracting the signal mean to eliminate any constant offset as developed in Equation 5.

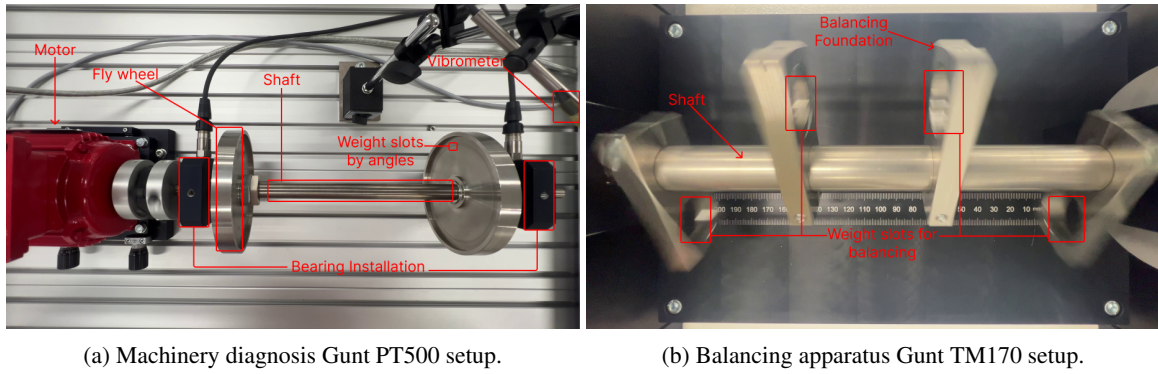
$$x_i = S(t_i) - \frac{1}{N} \sum_{j=1}^N S(t_j) \quad (5)$$

With $S(t_i)$ representing the raw motion timeseries samples from Equation 4, x_i are the DC-removed signal samples, N is the total number of samples. Statistical measures are computed from the DC-removed motion signal to capture the temporal characteristics of each working condition. Particularly, the key time domain features include the mean, variance, skewness and zero-crossing rate.

As each operating condition is characterized by distinct vibrational signatures, this pipeline aims to generate comprehensive datasets for each working condition of the captured scene.

3.2.3 Frequency Analysis

Comprehensive frequency domain analysis is performed using Fast Fourier Transform (FFT) to identify dominant vibrational frequencies and extract



(a) Machinery diagnosis Gunt PT500 setup.

(b) Balancing apparatus Gunt TM170 setup.

Figure 1: Setups for machinery diagnosis and balancing apparatus.

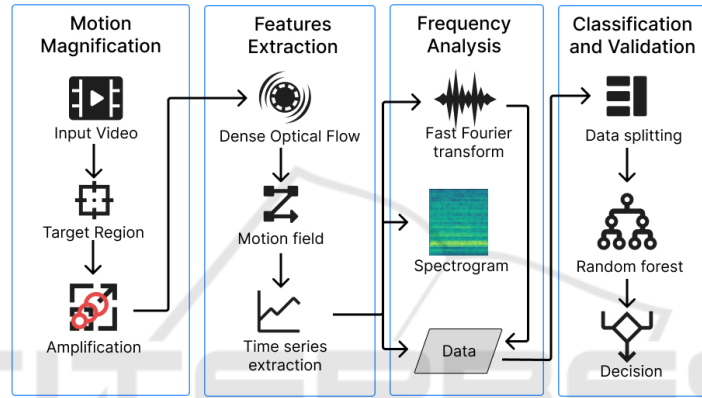


Figure 2: Video-based vibration analysis framework.

spectral characteristics relevant to the condition monitored. The FFT computation transforms the discrete time-domain signal to frequency domain, with an implementation of the Discrete Fourier Transform (DFT).

From the frequency spectrum, multiple discriminative features are extracted for analysis. These include the total spectral energy (Equation 6), the maximum amplitude and its corresponding frequency (Equations 7 and 8), the spectral centroid (Equation 9), spectral bandwidth (Equation 10), and the spectral entropy (Equation 11).

$$\text{Total Energy} = \sum_{k=0}^{N/2-1} A(k)^2 \quad (6)$$

$$\text{Max Amplitude} = \max_k A(k) \quad (7)$$

$$HF = f_k, \text{ where } k = \arg \max_j A(j) \quad (8)$$

$$\text{Spectral Centroid} = \frac{\sum_{k=0}^{N/2-1} f_k \cdot A(k)}{\sum_{k=0}^{N/2-1} A(k)} \quad (9)$$

$$SB = \sqrt{\frac{\sum_{k=0}^{N/2-1} (f_k - f_c)^2 \cdot A(k)}{\sum_{k=0}^{N/2-1} A(k)}} \quad (10)$$

$$\text{Spectral Entropy} = - \sum_{k=0}^{N/2-1} P_k \log_2(P_k) \quad (11)$$

Here, $A(k)$ represents the normalized amplitude at frequency bin k , f_k is the corresponding frequency, f_c is the spectral centroid, and $P_k = A(k) / \sum_{j=0}^{N/2-1} A(j)$ is the normalized probability density for entropy calculation.

3.2.4 Classification and Validation

The final component of our framework applies Random Forest-based machine learning classification to the comprehensive features extracted from the video-based vibration analysis. Here, an important step involves a preprocessing pipeline to ensure data quality through systematic process such as: data structuring and, 80-20 train-test split using stratified sampling.

Furthermore, the performance evaluation encompasses multiple metrics for comprehensive assess-

ment including the accuracy, and the F1-score computed based on obtained true positives, true negatives, false positives, and false negatives predictions.

In the following section, we will present the results of our framework in ground truth test rigs.

4 RESULTS AND DISCUSSION

In this section, we present the configuration in which we captured the machinery scene and the empirical results we obtained from our proposed framework depicted in Figure 2.

4.1 Experimental Setup

The baseline established in subsection 3.1, we recorded a set of 1 minute videos of both the Gunt PT500 and Gunt TM170 (Figure 1) at 4K@60FPS resolution for each of their operating conditions running at 250RPM. More precisely, for the Gunt PT500, the camera has been installed from above at 50cm height of the base plate supporting the machine; and from the front at 91.5cm away and elevated at 120cm height. Regarding the Gunt TM170, we recorded from the front view at 11cm away with 12cm height then 16cm away with 20cm height.

Furthermore, we fed the captured scenes into our proposed framework which runs on a Linux distribution in a DELL Intel Core i5-2410M 2.3GHz 8GB RAM to extract and process the features relevant for frequency analysis.

4.2 Dataset Description

The dataset extraction of our proposed framework maximizes data utilization and create sufficient training samples using a sliding window approach implemented for temporal segmentation. More particularly, each video is systematically divided into overlapping clips of 20s with a 5s overlap interval. Thus, this segmentation strategy produces nine clips per initial video, where the first clip encompasses the range 0 to 20s, the second one covers 5s to 25s, and so on. We obtained a set of 272 compiled clips for each of the operational condition. Hence, this overlapping segmentation approach served multiple purposes: it increased the available training data from the limited original recordings, with temporal continuity between adjacent clips, and ensured that transient vibrational events were captured in multiple segments.

4.3 Vibration Signal Analysis

The video processing pipeline successfully extracted meaningful vibration signatures for each of the four monitored test conditions: normal operation, 10g imbalance, 37g imbalance, and outer ring bearing fault.

The extracted time-domain signals reveal clear, condition-dependent signatures. In the healthy state (Fig. 3a), oscillation amplitudes remain within ± 0.002 , reflecting a stable baseline with only sporadic spikes. Introducing a 10g imbalance (Fig. 3b) elevates peak excursions to nearly ± 0.4 , indicating emergent impulsive events and increased variability. Under 37g loading (Fig. 3c), the system exhibits the most severe response, with swings up to ± 5 and sustained high-magnitude bursts, underscoring the method's sensitivity to imbalance severity. The outer-ring fault (Fig. 3d) produces moderate amplitudes ($\approx \pm 0.7$) but with pronounced periodic impacts characteristic of bearing defects, distinguishing it from pure imbalance cases. Together, these patterns confirm that our pipeline reliably captures both subtle and extreme vibration behaviors directly from video data.

The FFT results further validate condition separability. The normal spectrum (Fig. 4a) shows low-level peaks near 8Hz, 23.5Hz, and 27Hz (amplitudes $\approx 2.5 \times 10^{-5}$), representing nominal rotational harmonics. With a 10g mass (Fig. 4b), energy concentrates between 20–30Hz, peaking at 28–29Hz (≈ 0.012), signaling elevated high-frequency content due to imbalance. The 37g case (Fig. 4c) displays a broader spectral footprint, with significant components at 3Hz, 19Hz, and 20Hz (0.1–0.12), reflecting combined low- and high-frequency excitation. Faulted bearing (Fig. 4d) presents distinct peaks at 21Hz, 27Hz, and 28.5Hz (0.017–0.018), yielding a spectral profile that diverges from both healthy and imbalance conditions. The four-order-of-magnitude span in amplitude (2.5×10^{-5} –0.12) demonstrates that video-derived spectra quantitatively capture condition-specific dynamics suitable for reliable classification.

Spectrograms analysis add temporal resolution to these findings. Healthy state (Fig. 5) maintain uniform energy across the band at -40 to -20 dB. The 10g imbalance (Fig. 6) shows band-limited energy surges of ≈ -60 dB at rotational harmonics, while the 37g condition (Fig. 7) exhibits stronger bursts between 10–20Hz, particularly at 7–15s approaching -20 dB. The bearing fault (Fig. 8) is marked by a high-energy event at 6–8s spanning 15–25Hz greater than -60 dB, a different signature from imbalance cases.

Collectively, analysis of 272 video clips yielded

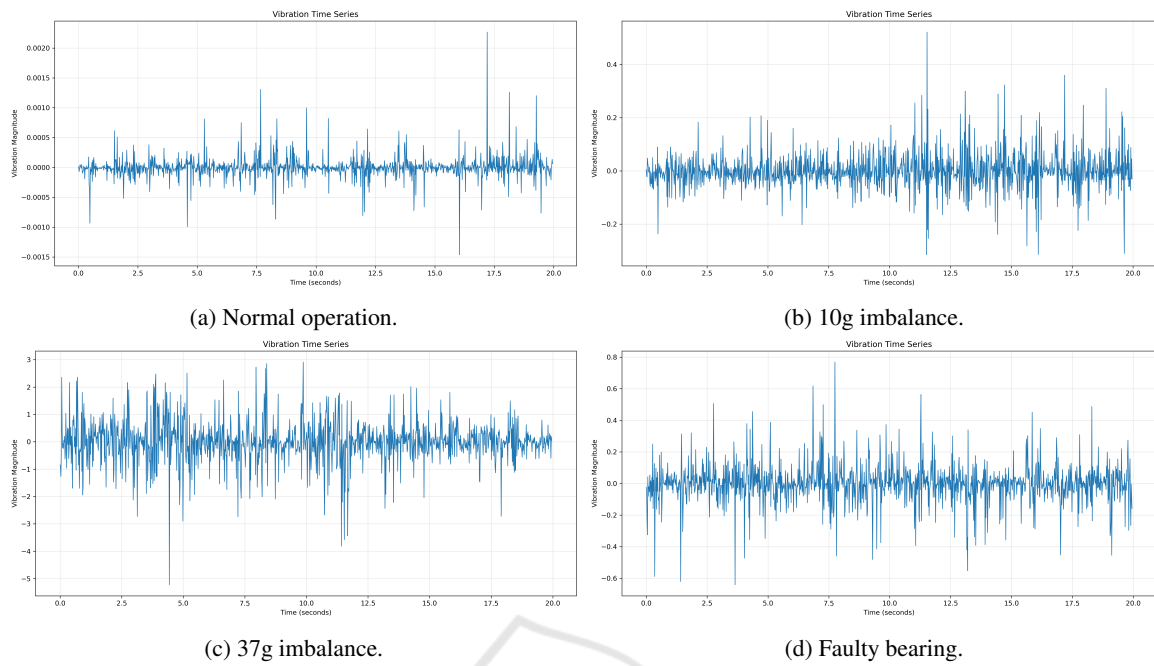


Figure 3: Time-domain vibration signals extracted from video recordings under different conditions.

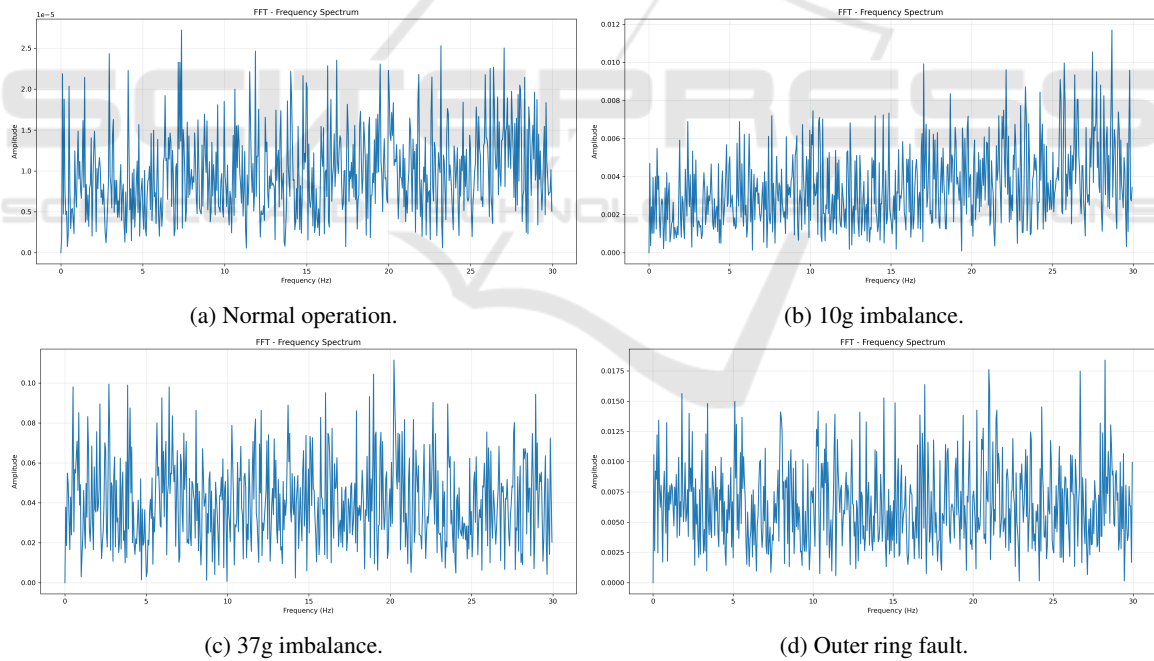


Figure 4: Frequency spectra from FFT analysis of vibration signals under different conditions.

14 discriminative frequency features. Their consistency across time, frequency, and time–frequency domains underpins the subsequent Random Forest classifier’s strong performance.

4.4 Classification Performance

The Random Forest classifier demonstrated excellent performance in distinguishing between the four machinery conditions using the video-derived features. We configured its parameters with 200 estimators at a maximum depth of 10. To ensure robust model eval-

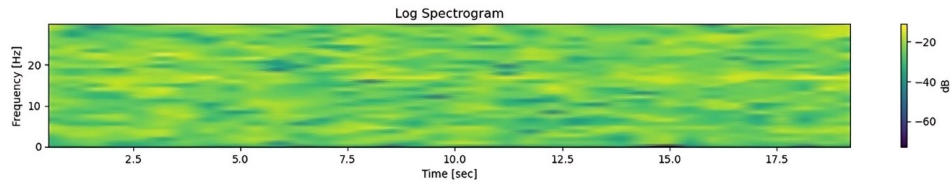


Figure 5: Spectrogram analysis under normal operating condition.

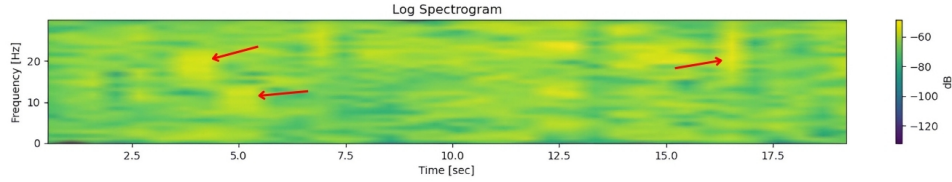


Figure 6: Spectrogram analysis for 10g imbalance showing time-varying frequency patterns.

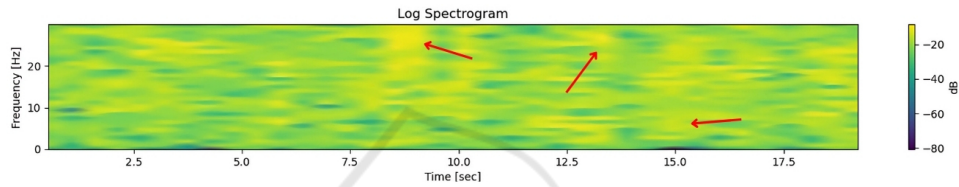


Figure 7: Spectrogram analysis for 37g imbalance with time-varying frequency patterns.

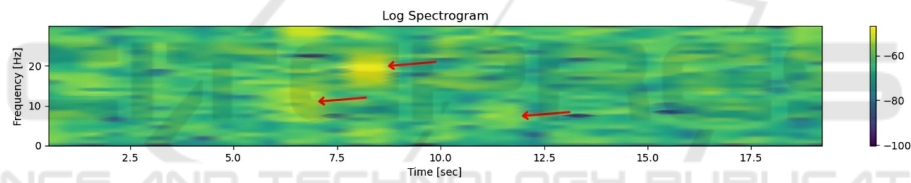


Figure 8: Spectrogram analysis for faulty condition showing time-varying frequency patterns.

uation, the dataset was partitioned using an 80 – 20 split, allocating 80% of the clips (218 clips) for training purposes and reserving the remaining 20% (54 clips) for testing. Hence, the model achieved an overall accuracy of 96.7% on the test set, demonstrating robust fault detection capabilities using solely video-derived vibration features.

Moreover, the macro-averaged F1-score of 0.965 indicates balanced performance across all fault categories. Figure 9 presents the resulting confusion matrix, where each cell represents the number of classified video clips from the 54 clips test set, revealing detailed classification patterns and potential areas for improvement.

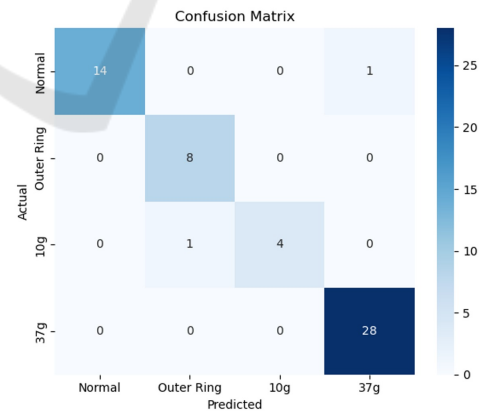


Figure 9: Classification performance of all conditions.

5 CONCLUSIONS

This work confirms that video-based motion magnification can serve as a reliable, non-contact approach for machinery fault diagnosis. We introduced

a feature extraction pipeline capturing both temporal and frequency characteristics of vibrations directly from video frames, and validated it, based on ground-truth measurements of rotating equipment unlike prior synthetic-data studies.

Experimental tests on the Gunt PT500 and TM170

under four operating conditions demonstrated that the extracted features effectively discriminate normal, faulty, and imbalance states. A Random Forest classifier trained on these features achieved 96.7% accuracy and a macro-averaged F1-score of 0.965, underscoring the method's robustness. Nevertheless, future work will extend this framework to diverse machinery types and fault scenarios, and explore real-time processing for continuous industrial monitoring.

ACKNOWLEDGEMENTS

This work is funded by the Japan International Cooperation Agency (JICA) Program, Egypt – Grant Number (TB1-25-4: "Contactless Predictive Maintenance for Rotating Equipment: An AI and Vision-Based Approach").

REFERENCES

- de Koning, M., Machado, T., Ahonen, A., Strokina, N., Dianatfar, M., De Rosa, F., Minav, T., and Ghabcheloo, R. (2024). A comprehensive approach to safety for highly automated off-road machinery under regulation 2023/1230. *Safety Science*, 175:106517.
- Everingham, M., Van Gool, L., Williams, C. K., Winn, J., and Zisserman, A. (2010). The pascal visual object classes (voc) challenge. *International journal of computer vision*, 88:303–338.
- Gunt (2025a). Pt 500 machinery diagnostic system, base unit.
- Gunt (2025b). Tm 170 balancing apparatus.
- Kalaiselvi, S., Sujatha, L., and Sundar, R. (2018). Fabrication of mems accelerometer for vibration sensing in gas turbine. In *2018 IEEE SENSORS*, pages 1–4.
- Ke, J., Wang, Q., Wang, Y., Milanfar, P., and Yang, F. (2021). Musiq: Multi-scale image quality transformer. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 5148–5157.
- Kiranyaz, S., Devocioglu, O. C., Alhams, A., Sassi, S., Ince, T., Avci, O., and Gabbouj, M. (2024). Exploring sound vs vibration for robust fault detection on rotating machinery. *IEEE Sensors Journal*.
- Lado-Roigé, R., Font-Moré, J., and Pérez, M. A. (2023). Learning-based video motion magnification approach for vibration-based damage detection. *Measurement: Journal of the International Measurement Confederation*, 206.
- Lado-Roigé, R. and Pérez, M. A. (2023). Stb-vmm: Swin transformer based video motion magnification. *Knowledge-Based Systems*, 269:110493.
- Li, H., Sun, Y., Adumene, S., Goleiji, E., and Yazdi, M. (2025). Machinery safety improvement in manufacturing-oriented facilities: a strategic framework. *The International Journal of Advanced Manufacturing Technology*, pages 1–30.
- Lima, J. A., Miosso, C. J., and Farias, M. C. (2025). Syn-flowmap: A synchronized optical flow remapping for video motion magnification. *Signal Processing: Image Communication*, 130:117203.
- Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., and Zitnick, C. L. (2014). Microsoft coco: Common objects in context. In *Computer vision—ECCV 2014: 13th European conference, zurich, Switzerland, September 6–12, 2014, proceedings, part v 13*, pages 740–755. Springer.
- Nieto, J. J. (2025). Digital twins, history, metrics and future directions. *Mechatronics Technology*, 2(1).
- Oh, T.-H., Jaroensri, R., Kim, C., Elgharib, M., Durand, F., Freeman, W. T., and Matusik, W. (2018). Learning-based video motion magnification. In *Proceedings of the European conference on computer vision (ECCV)*, pages 633–648.
- Song, Y., Zhuang, Y., Wang, D., Li, Y., and Zhang, Y. (2025). Fault diagnosis in rolling bearings using multi-gaussian attention and covariance loss for single domain generalization. *IEEE Transactions on Instrumentation and Measurement*.
- Wadhwa, N., Rubinstein, M., Durand, F., and Freeman, W. T. (2013). Phase-based video motion processing. *ACM Transactions on Graphics (ToG)*, 32(4):1–10.
- Wadhwa, N., Rubinstein, M., Durand, F., and Freeman, W. T. (2014). Riesz pyramids for fast phase-based video magnification. In *2014 IEEE International Conference on Computational Photography (ICCP)*, pages 1–10. IEEE.
- Wu, H.-Y., Rubinstein, M., Shih, E., Guttag, J., Durand, F., and Freeman, W. (2012). Eulerian video magnification for revealing subtle changes in the world. *ACM transactions on graphics (TOG)*, 31(4):1–8.
- Xu, X. and Lu, Y. (2025). Acoustic modal analysis method for helicopter acoustic scattering. *AIAA Journal*, pages 1–21.
- Yang, Y. and Jiang, Q. (2024). A novel phase-based video motion magnification method for non-contact measurement of micro-amplitude vibration. *Mechanical Systems and Signal Processing*, 215.
- Zang, Z., Yang, X., Zhang, G., Li, S., and Chen, J. (2025). Video-based subtle vibration measurement in the presence of large motions. *Measurement*, 240:115559.
- Zhang, G., Hou, J., Wan, C., Li, J., Xie, L., and Xue, S. (2025). Non-contact vision-based response reconstruction and reinforcement learning guided evolutionary algorithm for substructural condition assessment. *Mechanical Systems and Signal Processing*, 224:112017.
- Zhao, H., Zhang, X., Jiang, D., and Gu, J. (2023). Research on rotating machinery fault diagnosis based on an improved eulerian video motion magnification. *Sensors* 2023, Vol. 23, Page 9582, 23:9582.
- Zhong, S., Jin, G., Chen, Y., Ye, T., and Zhou, T. (2025). Review of vibration analysis and structural optimization research for rotating blades. *Journal of Marine Science and Application*, 24(1):120–136.
- Śmieja, M., Mamala, J., Prażnowski, K., Ciepliński, T., and Łukasz Szumilas (2021). Motion magnification of vibration image in estimation of technical object condition-review. *Sensors*, 21.