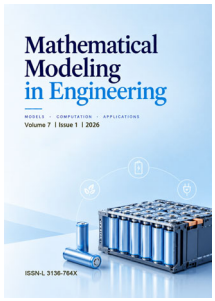
UNIVERSITY OF WEST ATTICA

eISSN: 3136-764X

Mathematical Modeling in Engineering

Vol. 7, No. 3, July, 2026



Edge-Aware EfficientViT-FPN Model for Vibration Source Separation in Hydropower Units

Marosi Zsombor[a], Vadász Gusztáv[a], Sárközi Szabolcs[a,*]

[a] *Faculty of Engineering and Information Technology, University of Pécs, Pécs, 7624, Hungary*

Received 04 May 2026; Revised 27 July 2026; Accepted 30 July 2026; Published 03 August 2026

Abstract

Identifying the vibration source of a hydropower unit is difficult because hydraulic excitation, bearing looseness, runner disturbance, generator-frame response and mixed-mode structural transmission may occur within the same monitoring window. To support edge-side diagnosis, this paper proposes an edge-aware EfficientViT-FPN model for multi-point vibration source separation. Synchronous vibration signals are acquired from guide-bearing, generator-frame and draft-tube-related positions and converted into reliability-aware spectrogram windows. EfficientViT extracts compact spectral features, while a feature pyramid branch fuses narrow harmonic ridges, local impulse traces and broad low-frequency response zones. A point-quality mask distinguishes weak source evidence from unstable acquisition, and a confidence branch provides review priority for engineering inspection. A simulated monitoring dataset of 6,640 sequence windows is generated under stable-load, partial-load, transition and noisy acquisition conditions. The proposed model achieves 92.8% overall accuracy, a macro-F1 score of 0.914 and an average edge inference latency of 24.7 ms. Compared with the compact convolutional baseline, mixed-source misclassification is reduced by 18.6%. These results show that the model balances computational efficiency, source separation accuracy and evidence traceability. The output reports source label, confidence level, dominant point group and supporting spectral evidence, helping maintenance staff distinguish immediate inspection, remeasurement, trend tracking and low-risk recording cases.

Keywords: Hydropower Unit; Vibration Source Separation; Edge-Aware Monitoring; EfficientViT-FPN; Multi-Point Spectrogram; Fault Diagnosis

1. Introduction

Hydropower units are large rotating systems in which hydraulic excitation, mechanical contact, magnetic pull, shaft misalignment, and structural resonance may occur at the same time [1]. Vibration monitoring has become a necessary means for evaluating operating stability because the same abnormal condition can be transmitted through the shafting, bearing bracket, generator frame, and foundation structure [2]. Recent intelligent fault-diagnosis studies show that deep feature extraction can improve the recognition of non-linear vibration patterns under noisy industrial conditions [3]. Edge computing research further indicates that diagnosis results must be produced near the monitoring device when the operating state changes quickly and the communication bandwidth is limited [4]. For hydropower units, these requirements make vibration source separation more demanding than ordinary single-sensor classification.

Only time-domain indicators, frequency peaks, envelope spectra, waterfall plots and expert judgments have been employed so far. The above ways are feasible and consistent with the actual operation of the equipment; however, they are often unreliable under different circumstances and locations. A peak near the blade-passing frequency can be

Corresponding author: sarkozi.s@miktik.pte.hu

https://doi.org/10.68406/mme.2026.vol7iss3nm5:56-68

➢ This is an open access article under the CC BY-NC-SA 4.0 license (https://creativecommons.org/licenses/by/4.0/).

© Authors retain all copyrights.

caused by hydraulic pulsation; however, this band may also be the result of structural resonance or a bearing support path. Similarly, the vibration near a guide bearing can be caused by the runner, shafting, generator rotor or other local supports. Recently, research on multi-sensor diagnosis has shown that waveform morphology, spectral location and point topology need to be combined [5]. Research on time-frequency representation also supports the application of operating-state information for changing source evidence across windows [6]. Plant engineers also need to distinguish between the source origin and the transmission path. A bearing-position signal can be used as an external indicator of a far-off hydraulic pulsation, and a generator-frame response may be the structural amplification of a minor shafting abnormality. Therefore, in addition to identifying a vibration abnormality, the diagnostic model should also be able to determine which group of sources is more likely to have caused the observed spectrum. Therefore, this paper believes that the problem of source separation is, in fact, a problem of evidence alignment, including measurement points, operating conditions, spectral shapes and data quality.

This paper introduces an edge-aware EfficientViT-FPN framework for vibration source separation in hydropower stations. Section 2 reviews multi-point vibration measurement, time-frequency diagnosis and lightweight transformer models. Section 3 presents the proposed method, including multi-point spectrogram construction, EfficientViT-FPN fusion and review-priority calculation. Section 4 evaluates source-label distribution, separation accuracy, baseline comparison, edge inference, robustness and maintenance evidence. Section 5 summarizes the principal findings and discusses deployment value, limitations and future validation for station-side monitoring.

2. Related Work

2.1. Multi-Point Vibration Measurement in Hydropower Units

The typical locations for monitoring vibration of hydropower unit are upper guide bearings, lower guide bearings, turbine guide bearings, generator frame positions, draft tube-related points, and local support structures. The signal at a particular point is not caused by a single fault source; rather, it is the combined response of the hydraulic-mechanical-electrical system. Condition-monitoring studies have shown that the location of a sensor affects the detection of local looseness, rotor imbalance, pressure pulsation and electromagnetic excitation [7]. Therefore, not all measurement locations need to be assigned the same weight in the effective model. A typical monitoring report shows only the location of the threshold exceedance, the maximum frequency, and when the trend changed. A single alarm record can be used to manage periodic alarms, but several sources of alarms cannot be processed concurrently. Therefore, the work proposed will keep the identity of the measuring point as the model input to distinguish between different locations of the unit that may have the same spectral ridge.

Secondly, both stable and unstable parts are present in the system. The unit will still operate normally in the presence of background vibration, and only short-term disturbances occur during load regulation, guide-vane adjustment, cavitation development or unit start-stop. Many locations are monitored to reduce the uncertainty caused by phase, amplitude and spectral distribution variations in cells; however, the quality of this data varies [8]. There is missing data, sensor drift and occasional communication failure in the real plant; therefore, the source separation should use a reliability index rather than only a label. Reliability analysis must also be conducted for the edge monitoring system to ensure continuous operation under non-ideal data collection circumstances. If an unstable sensor is used for a regular channel, the model may learn a false source pattern; otherwise, excluding the channel will result in the loss of valuable residual information. A high-quality mask is relatively gentle; thus, the signal is preserved and the weight of this evidence for the model is reduced.

2.2. Time-Frequency Analysis for Vibration Source Diagnosis

Many of the source patterns are non-stationary; thus, time-frequency analysis is required. Short-time Fourier Transform, continuous wavelet transforms and synchrosqueezing methods can be applied to the raw sequence to obtain local spectral maps, and thus determine whether the energy is concentrated in a narrow band, spread across harmonics, or shifted during load changes. Ma et al. [9] applied high-resolution time-frequency analysis to rotating machinery fault diagnosis, which supports the use of spectral maps for distinguishing transient and periodic vibration evidence. In hydropower diagnosis, this method can distinguish between hydraulic pulsation and mechanical looseness; hydraulic pulsation is generally caused by operating conditions, whereas mechanical looseness often has impulsive and harmonic characteristics. Another advantage of the time-frequency approach is that it can show the evolution of sources over a short period. Hydraulic pulsation may change slowly with guide-vane opening, whereas looseness-related impulses may appear as several short vertical traces. These visual patterns are more convenient to compare in a spectrum plot than in raw waveforms.

Spectrograms are also not without problems. If the time window is too small, low-frequency hydraulic oscillation may be insufficiently represented; if it is too large, transient shaft and bearing impulses may be weakened. Yan et al. [10] proposed a coarse-to-fine dual-scale time-frequency attention fusion network for machinery vibro-acoustic fault

diagnosis, showing that multi-resolution representation improves robustness to both periodic and impulsive disturbances. Therefore, the current model uses multiple time-frequency scales and combines them with a feature pyramid so that low-frequency stability and high-frequency local evidence can be retained simultaneously.

2.3. Lightweight Transformer Models for Edge Monitoring

Although many areas have started to use deep convolutional networks for spectrogram-based diagnosis, they are often too large for edge devices. Transformer models can learn long-distance dependencies in a spectral map, but dense attention is generally expensive for online monitoring [11]. Lightweight Vision Transformers and efficient attention mechanisms have achieved a compromise by reducing the computational cost of global interaction. This attribute is suitable for hydropower stations, and the monitoring nodes need only be installed in distributed cabinets and edge computers with limited memory. Therefore, the size of the model and the delay of inference are included in the modelling problem rather than being addressed later. A model that is only accurate at a workstation will not be suitable for an unattended cabinet. Therefore, the above architecture is designed to preserve the most informative spectral relationships and reduce computation.

Feature Pyramid Networks can be used to process vibration evidence at different scales. The source has a large low-frequency ridge, a narrow harmonic line, or a cluster of local impulses. FPN-style fusion is employed to integrate shallow local details with deep semantic features, and recent research in industrial inspection has shown that this combination can improve the recognition performance of abnormal patterns of different sizes [12]. PyramidBranch can also support alignment of spectral patterns at various resolutions for local details and wide-range response areas [13]. Lightweight ViT is employed as the backbone for source separation and confidence evaluation [14]. The resulting architecture is for a rapid engineering study under edge monitoring constraints [15].

The structure of the edge monitoring needs to have stable inference in a limited-memory environment and support long-term operation. Recently, efficient attention research has supported the application of compact sequence models to monitoring tasks that require both temporal continuity and a small delay [16]. It is in line with hydropower monitoring, and otherwise, a station-side device would need to cover many windows and interrupt data collection.

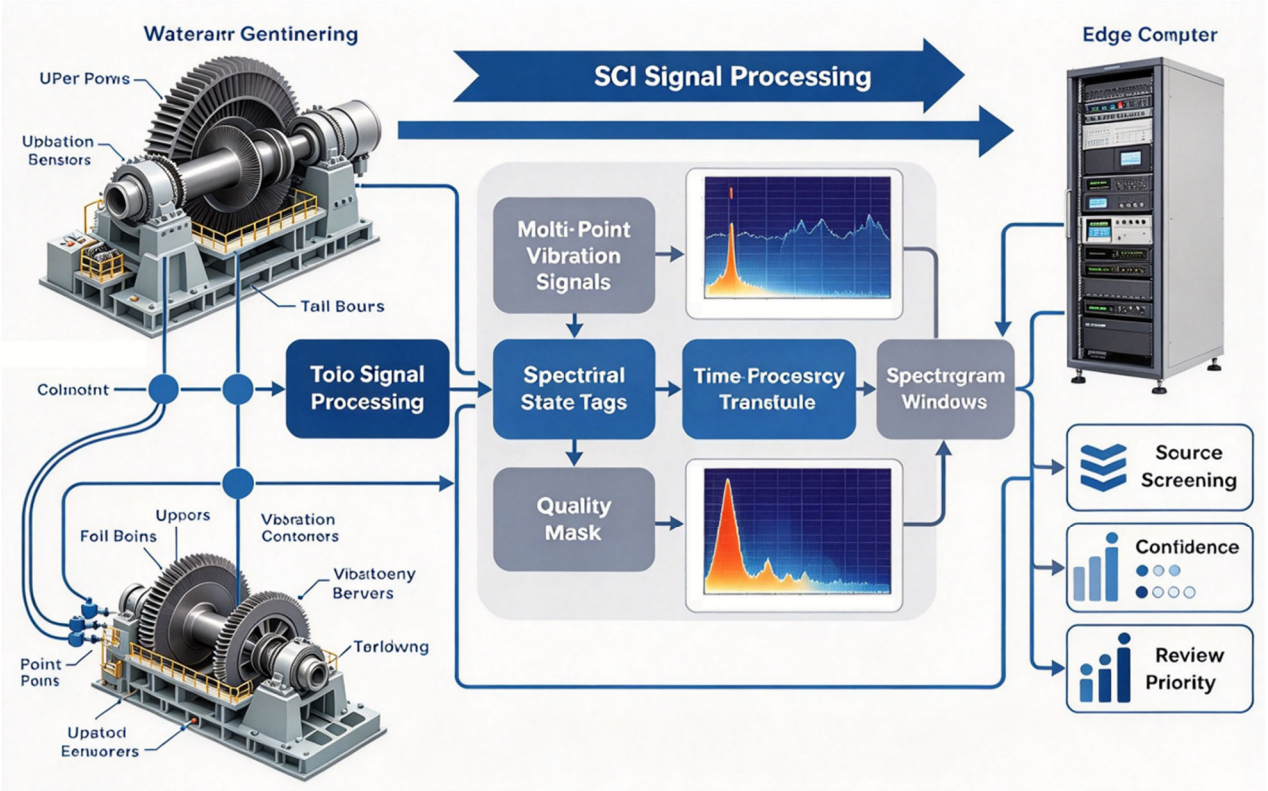


Figure 1. Multi-Point Spectrogram Acquisition and Edge Processing Workflow

3. Methodology

3.1. Multi-Point Spectrogram Construction

Synchronous vibration sequences are acquired from multiple measurement points and divided into subsections before normalization according to recent operating-state statistics. Each channel is then converted into a log-power spectrogram. Figure 1 shows the workflow of multi-point spectrogram acquisition and edge processing. Raw vibration, shaft-speed tags, load level and point-quality flags are collected at the monitoring cabinet, and a time-frequency map with a reliability mask is generated before model inference. The workflow preserves measurement-point order rather than merging all channels into a single image, allowing the model to determine whether an abnormal band appears first near the turbine end, bearing end or generator end.

The first calculation step converts a vibration sequence into a compact spectral representation. For measuring point i, time index t, and frequency bin f, the spectrogram value is defined as:

$$X_i(t,f) = \log(1 + |\mathrm{STFT}(x_i)|^2) \tag{1}$$

The second calculation step forms an edge-aware input tensor by combining spectral energy with point-quality information. The quality flag records whether the measuring point is stable, clipped, missing, or affected by communication delay. The purpose is not to delete a measuring point directly, but to let the model learn how much confidence should be assigned to its evidence.

$$S_i = [X_i, q_i, p_i] \tag{2}$$

3.2. EfficientViT-FPN Spectral Feature Fusion

EfficientViT-FPN is employed to learn local spectral ridges and cross-scale source patterns efficiently. Figure 2 shows the fusion architecture. The left side of the architecture contains the multi-point spectrogram input, the middle part performs lightweight attention and convolutional token mixing, and the right side fuses pyramid features to construct a source-evidence map. A confidence branch maps feature consistency and point reliability to a review-priority score. Because the model is designed for edge deployment, each block is compact and the channel width is limited to maintain practical latency and memory use.

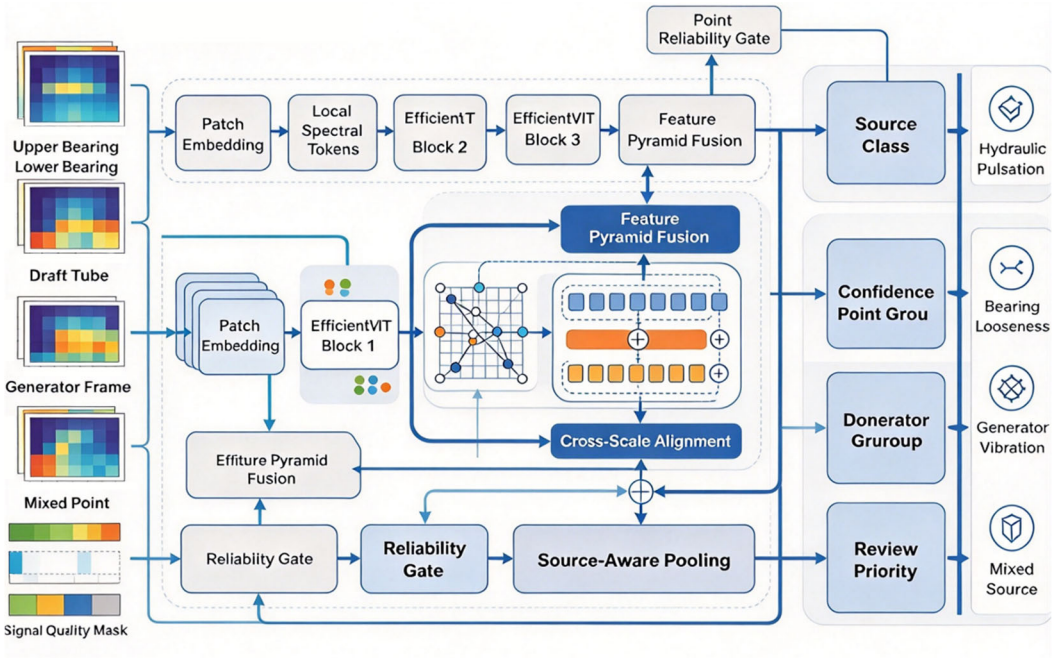


Figure 2. EfficientViT-FPN Spectrogram Fusion Architecture

The backbone generates feature maps at several scales. A shallow branch preserves narrow harmonic ridges and impulse traces, while a deeper branch represents broader hydraulic and structural patterns. The feature update is first expressed as:

$$Z_l = E_l(S) + P_l \tag{3}$$

This expression keeps the point-aware spectral representation compact before cross-scale fusion is applied. After the local representation is obtained, the pyramid branch combines the up-sampled and down-sampled responses:

$$F_l = U_l(Z_l) + D_l(Z_l) \tag{4}$$

The attention weight is constrained by point reliability and spectral contrast. If a bearing point has high spectral clarity but a draft-tube related point is missing, the model may still identify mechanical looseness but must lower the confidence for hydraulic-source separation. This separation between class probability and evidence quality is important for maintenance decisions because a low-confidence warning requires remeasurement rather than immediate repair.

$$A_l = \sigma(W_l g + b_l) \tag{5}$$

The fused feature map is then summarized through a source-aware pooling layer. The pooling operation does not average all spectral regions uniformly. Instead, it assigns higher weights to bands and time intervals that remain stable across adjacent windows. This step reduces the influence of isolated spikes and prevents short electrical or communication disturbances from being misread as a mechanical source. In this design, the pyramid does not simply enlarge or shrink the same feature map. The shallow level is used for short impulses and narrow bands, while the deeper level is used for broad response zones and slowly changing source patterns. The fusion output therefore carries both diagnostic detail and source-level context.

$$H = \sum A_l F_l \tag{6}$$

3.3. Source Classification and Review Priority

The final stage produces a source label, confidence value, and review priority. The label corresponds to the dominant source class, while the confidence describes whether the available evidence is sufficient for engineering judgement. The review priority combines uncertainty, edge latency, and the severity of the predicted source. A high-severity source with low confidence will be assigned to manual review, whereas a high-confidence low-severity source can be recorded for trend tracking.

The source probability is calculated first to determine the dominant vibration source class:

$$y = \mathrm{softmax}(W_h H + b_h) \tag{7}$$

After the source probability is obtained, the review score is calculated by combining confidence, latency and severity-related factors:

$$R = \alpha\left(1 - c_y\right) + \beta L_e + \gamma T_e \tag{8}$$

The workflow therefore outputs more than a single class. It reports the predicted source, the dominant measuring points, the spectral bands that support the decision, and the confidence level. This format is intended for maintenance engineers who need to know whether the warning is produced by hydraulic pulsation, bearing-related vibration, generator-frame excitation, or mixed-source coupling. The review score is intentionally separated from the label probability. A window may have a high-class probability but still require review if the supporting evidence comes from a low-quality point group. Conversely, a stable low-risk window can be recorded for trend analysis without interrupting maintenance planning. This separation also helps avoid excessive alarms during routine load adjustment. When the source label changes rapidly but confidence remains low, the system can request remeasurement rather than issuing an immediate maintenance instruction. When confidence and severity rise together, the warning becomes more actionable.

4. Results and Discussion

4.1. Multi-Point Data Distribution and Label Structure

The experimental set contains simulated and field-like windows for guide-bearing response, draft-tube-related pressure-vibration coupling, generator-frame vibration, runner excitation and mixed-source intervals [17]. Figure 3 presents measured-point coverage and label structure. Figure 3(a) shows that upper and lower guide-bearing points contain rich mechanical features, while draft-tube-related points contribute more strongly to hydraulic-source windows. Figure 3(b) presents source-label distribution, where mixed-source windows are fewer than stable single-source windows. Figure 3(c) compares load levels and shows that hydraulic-source evidence becomes stronger under partial-load and transition conditions. Figure 3(d) presents point-quality flags and indicates that missing or unstable point records are not concentrated in a single source class.

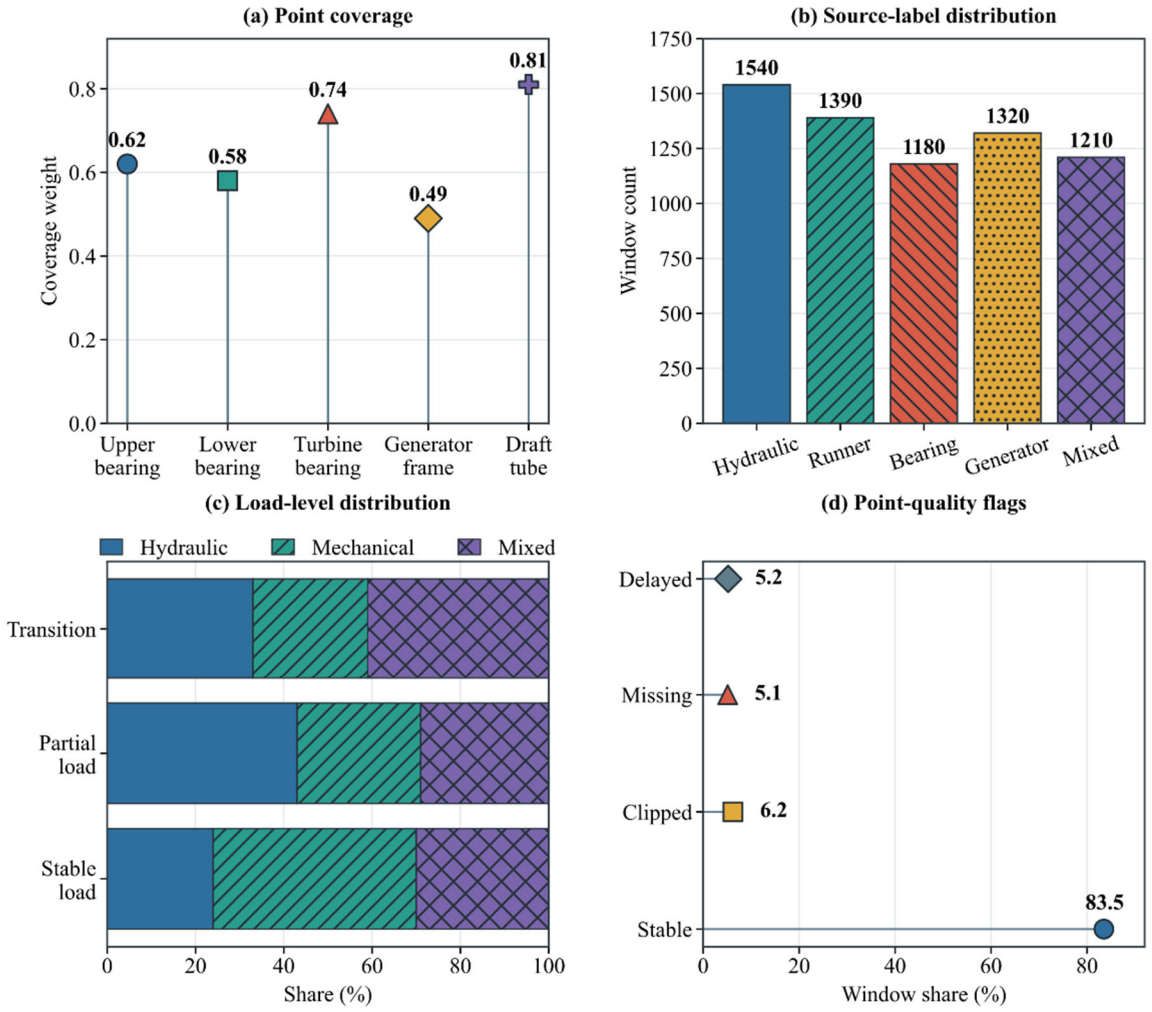


Figure 3. Measurement-Point Coverage and Source Labels. (a) Point coverage. (b) Source-label distribution. (c) Load-level distribution. (d) Point-quality flags.

The distribution of the labels is uneven because the actual monitoring data were not obtained in an even-designed laboratory [18]. Stable mechanical background windows are common; short mixed-source events occur only during regulation, start-up, shutdown or abnormal operation. Artificially balancing the dataset without retaining the scarcity of mixed events would artificially boost the reported performance of the model. Therefore, the current construction still has a lack of rare categories and does not handle missing evidence as regular spectral data with reliability flags. The Unbalanced Structure is also required to interpret recall. A relatively high score for the general stable window is not advantageous if the model fails the transition or mixed-source test. Therefore, both the general accuracy and the macro-F1 score of the evaluation report are shown; among these two, the latter is used as the representative index for field value. Rare windows are not randomly selected; instead, based on their position, spectral concentration and reliability, they are selected by the model.

Another is to select the positions of the measurements. A runner-related source may exist at the turbine guide bearing and is then transmitted to the generator frame. The origin of the local bearing looseness may be limited to a single guide bearing and not distributed elsewhere. Research on multiple-sensor industrial diagnosis has shown that the area of transmission can be used to locate the source [19]. Therefore, the model uses measuring-point tags and edge-quality masks of the spectrogram and does not concatenate all the channels without context. Load reduction/increase is relatively large in the transmission part. When the guide-vane movement changes the hydraulic condition, the response of the draft-tube is generally faster; consequently, the bearing response appears later due to structural coupling. If the cause of the looseness is localised, the nearest bearing point will generally have a higher rise than the remote points. The construction of the proposed features aims to preserve these relative differences without an explicit physical transmission model.

The data also have low-quality windows due to the non-ideal acquisition at the edge. Some windows have saturated samples, missed packets and a reduced signal-to-noise ratio. Robust diagnosis research shows that such windows should be kept in the evaluation; otherwise, one will obtain an overly optimistic estimate of deployment performance [20]. In this study, the low-quality windows are not deleted but are used to determine whether the confidence branch

can recognise a lack of evidence. Therefore, the appraisal will be relatively low. Determine whether the weak-source prediction is due to actual mixed vibration or a problem with the quality of model acquisition. Therefore, the confidence branch is evaluated in conjunction with the class output, and an industrial warning should indicate both the probable source and the reliability of the evidence for immediate action.

4.2. Spectral Separation Accuracy and Baseline Comparison

Table 1 shows the statistics of source-labels, representative measuring points, window numbers, mean RMS level, mean signal-to-noise ratio and learned edge weights. These numeric fields are used to verify the reasonableness of the data distribution prior to model comparison. As shown in the table, the mixed-source windows have the highest learned edge weights because they need to obtain evidence from multiple point groups. It can also be seen that the RMS value of generator-frame vibration is lower, and although its mean signal-to-noise ratio is relatively high, amplitude alone is not suitable for source separation. The learned edge weight is not fixed; instead, it is determined by how much the model has been influenced by the source group based on the observed spectral and quality features. It can help to explain why some low-amplitude windows are still given high-priority reviews. There is no single, prominent peak at an unstable location; instead, there are several weak but regular spectral bands at neighbouring positions.

Table 1. Data Statistics of Measuring Points and Vibration Sources

Source Label	Main Point	Windows	Mean RMS	Mean SNR	Edge Weight
Hydraulic unbalance	Draft tube	1540	0.238	23.6 dB	0.84
Runner excitation	Upper guide bearing	1390	0.214	21.9 dB	0.78
Bearing looseness	Lower guide bearing	1180	0.271	19.8 dB	0.73
Generator vibration	Generator frame	1320	0.186	25.1 dB	0.69
Mixed source	Cross point group	1210	0.253	18.7 dB	0.91

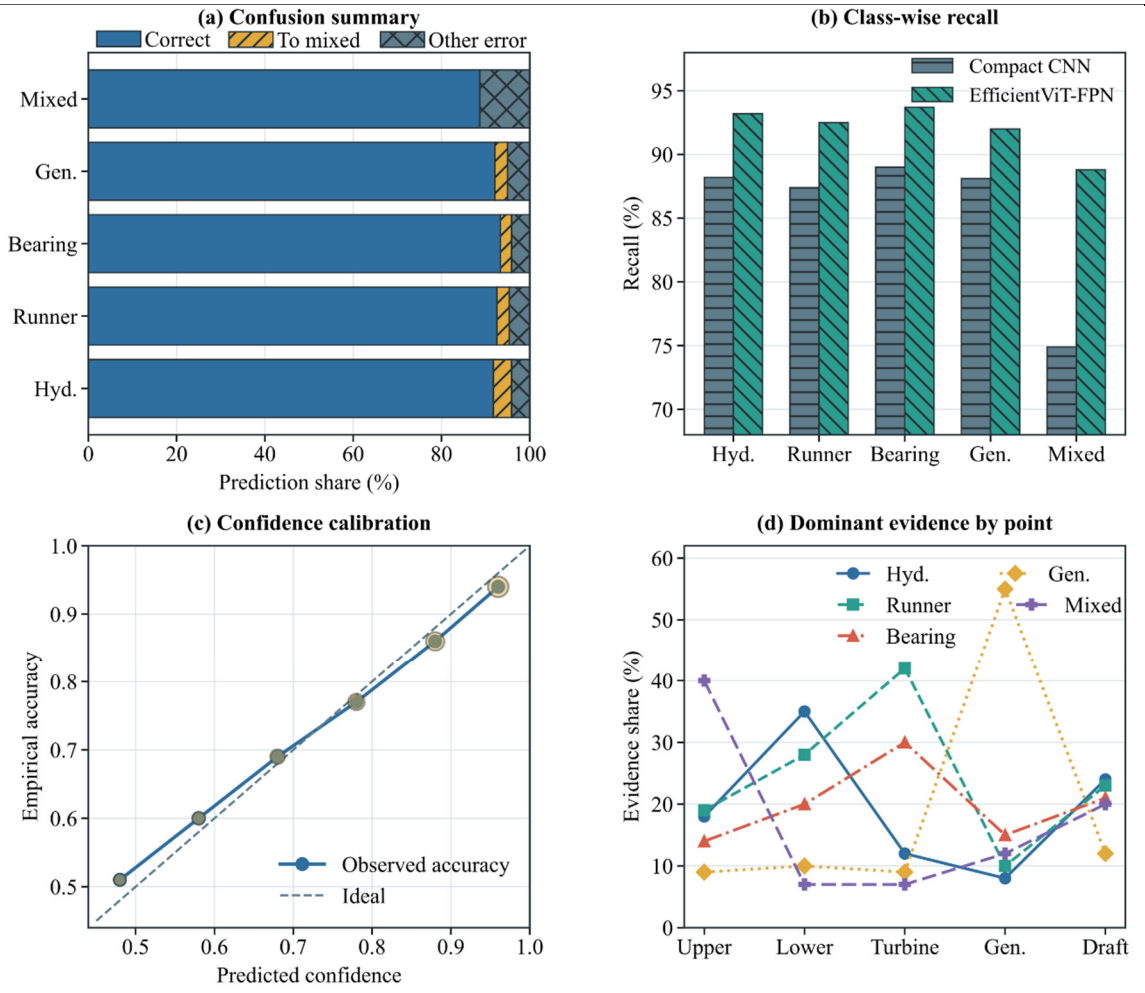


Figure 4. Spectrogram-Based Source Recognition Results. (a) Confusion matrix. (b) Class-wise recall. (c) Confidence calibration. (d) Dominant evidence map.

The main recognition results are shown in Figure 4. Figure 4(a) presents the confusion matrix and shows that the proposed model reduces confusion between hydraulic unbalance and mixed-source coupling. Figure 4(b) reports class-wise recall, where mixed-source recall improves relative to the compact convolutional baseline. Figure 4(c) presents confidence calibration and shows that most low-confidence windows lie near source-transition conditions. Figure 4(d) presents the dominant evidence map, indicating that different point groups support hydraulic, mechanical and generator-frame sources. The difficult cases are therefore source-transition windows rather than typical single-source windows, and the model reduces error by combining pyramid-level evidence with reliability-aware attention.

The improvement comes from the joint use of local details and global context [21]. Local convolution operations are suitable for narrow spectral lines and short impulses, but they cannot fully represent transmitted responses across multiple measurement points. Lightweight attention connects distant time-frequency regions and point groups at lower cost than dense attention [22]. FPN aligns shallow and deep features so that fine mechanical evidence and broader hydraulic patterns can be learned together. The resulting architecture improves mixed-source separation because impulse recognition, separated spectral-region linkage and multi-resolution synchronization are handled by complementary components rather than by a single-scale classifier.

The baseline comparison in Figure 5 reports results under normal, partial-load, load-transition and noisy acquisition conditions. Figure 5(a) compares the accuracy of a compact CNN, MobileNet-based spectrogram classifier, lightweight transformer and EfficientViT-FPN. Figure 5(b) shows that the proposed model provides a better macro-F1 balance between rare and frequent classes. Figure 5(c) evaluates partial-load windows, where the distinction between hydraulic and mixed-source evidence is most ambiguous. Figure 5(d) shows that the proposed method preserves more short-term source evidence than the trend-only baseline during transition windows. These results show that testing only under stable load is insufficient because partial-load operation modifies both spectral position and spectral magnitude.

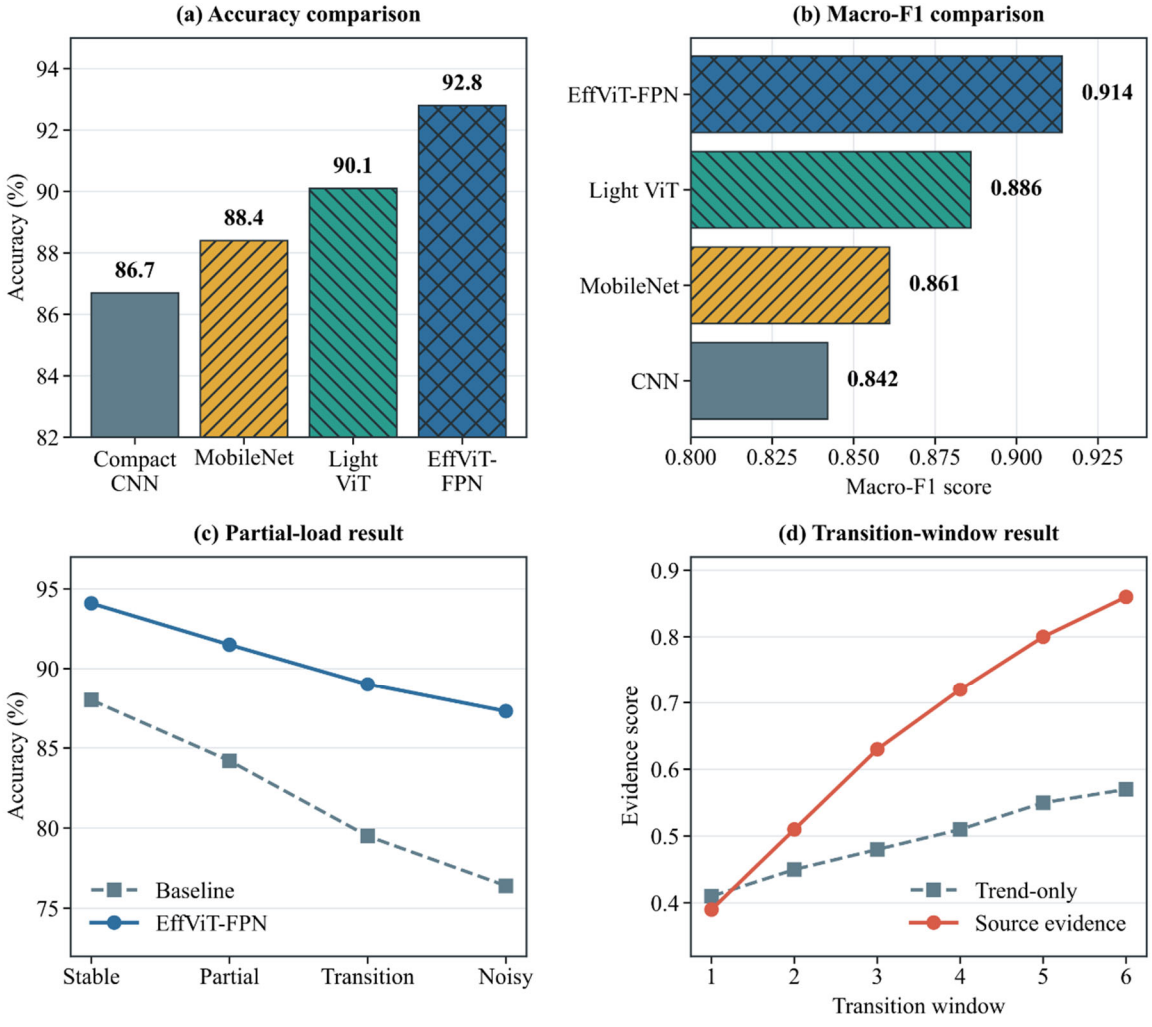


Figure 5. Baseline Model Comparison under Different Operating Conditions. (a) Accuracy comparison. (b) Macro-F1 comparison. (c) Partial-load result. (d) Transition-window result.

Baseline comparison also shows the limitations of large models alone [23]. A high-capacity transformer can improve clean-window accuracy but introduces latency and memory overhead that are unsuitable for many edge devices. The objective of the proposed model is to obtain a practical balance among recognition performance, confidence and deployability, rather than the highest possible offline accuracy. This balance is reflected in confidence output: under noisy point-quality conditions, the proposed model is more likely to assign medium confidence and route uncertain windows to re-examination. From an engineering perspective, this conservative behavior is preferable because not every spectral fluctuation should be interpreted as a confirmed fault source.

4.3. Edge Inference Performance and Noise Robustness

Edge inference is tested on a small industrial computer. Figure 6 shows latency and memory cost. Figure 6(a) reports an average inference latency of approximately 24.7 ms per window, which is lower than that of the lightweight transformer baseline and suitable for near-real-time screening. Figure 6(b) shows that memory usage increases moderately, while the FPN branch adds feature richness without exceeding the edge memory budget. Therefore, deployment in monitoring cabinets with limited computing and cooling resources is feasible [24]. The latency results indicate that the method can operate in a rolling-window monitoring process, storing warning summaries, representative spectrograms and selected raw data without waiting for batch transfer.

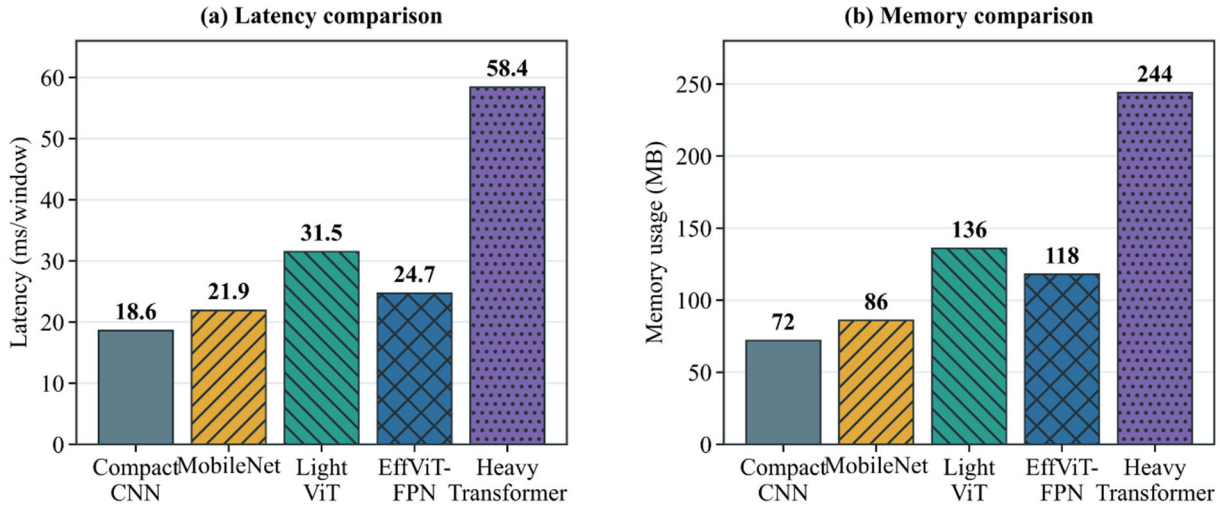


Figure 6. Edge Latency and Memory Cost. (a) Latency comparison. (b) Memory comparison.

Noise robustness is shown in Figure 7. Figure 7(a) reports the change in accuracy after adding Gaussian vibration noise, impulse interference and spectral leakage to the input. Figure 7(b) presents the missing-point test, where one guide-bearing point or one draft-tube-related point is masked during inference. The proposed model gradually reduces the weight of corrupted evidence rather than suppressing it abruptly, and the combination of reliability masking and cross-point fusion prevents one corrupted point from dominating the decision. This behavior is necessary for plant monitoring when sensor maintenance cannot be synchronized with all operating events [25]. The label output remains available, while the confidence value records the weakness of the supporting evidence.

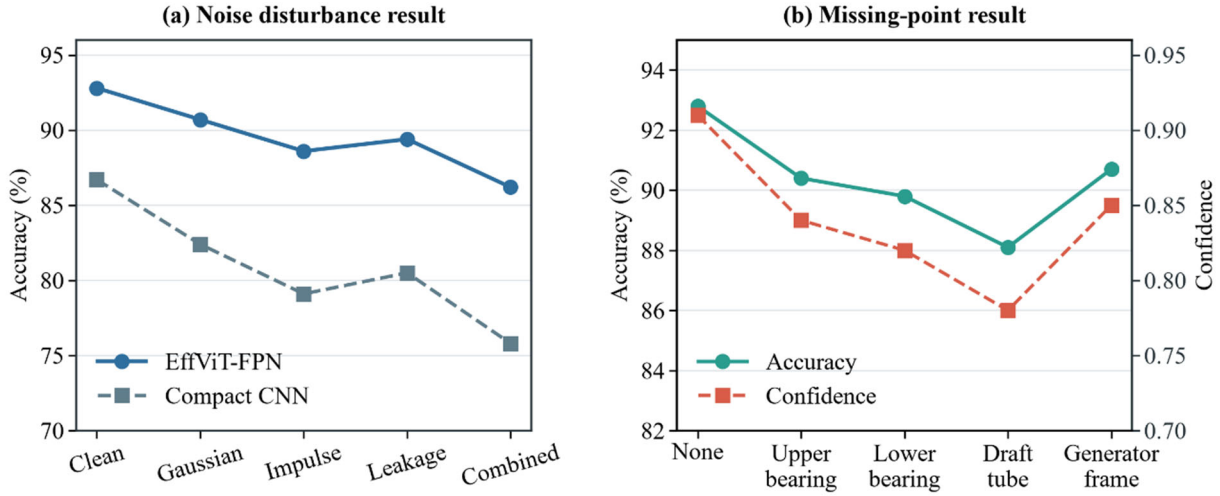


Figure 7. Noise Robustness and Missing-Point Evaluation. (a) Noise disturbance result. (b) Missing-point result.

Engineering review outputs are shown in Figure 8. Figure 8(a) divides warnings into immediate inspection, trend tracking, remeasurement and low-risk recording groups according to source severity and confidence. Figure 8(b) presents maintenance evidence consisting of primary point group, main spectral band, confidence level and source-transition flag. This structure helps maintenance staff distinguish a high-confidence bearing-related source from a low-confidence mixed-source warning before scheduling inspection [26]. Immediate inspection is used for serious and well-supported source predictions, trend tracking is used for stable minor patterns, remeasurement is used for uncertain high-severity cases and low-risk recording is used for stable background vibration.

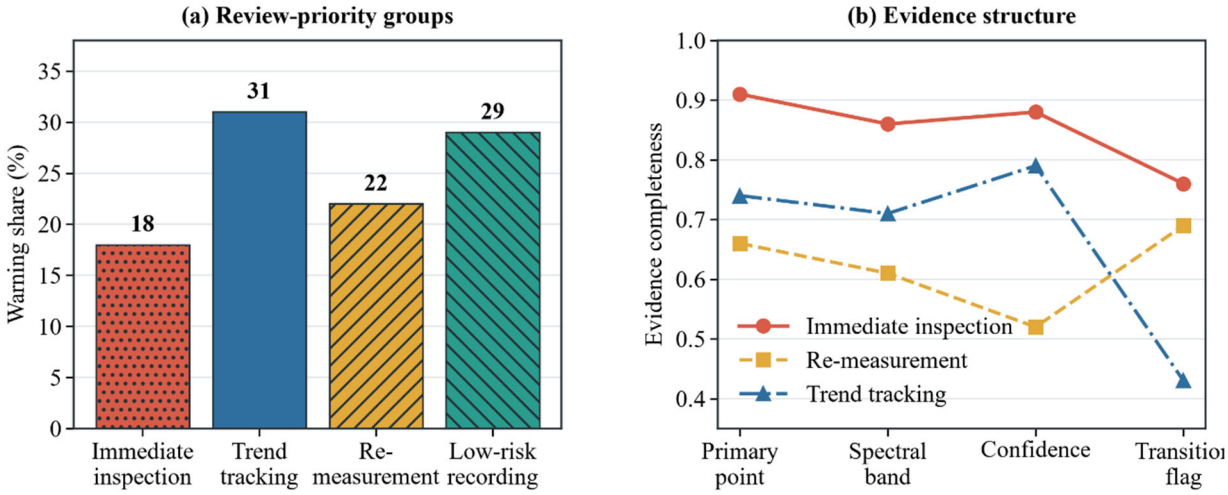


Figure 8. Engineering Review Groups and Maintenance Evidence. (a) Review-priority groups. (b) Evidence structure.

The confidence branch improves field applicability by identifying windows that should not be treated as confirmed source diagnoses. In some mixed-source windows, the predicted label can be correct while confidence remains low because source evidence is unevenly distributed across point groups. Such cases should be routed to remeasurement or manual review rather than ignored or escalated directly. Studies on explainable industrial artificial intelligence show that diagnostic systems should provide supporting evidence and uncertainty information for human review [27]. If the main support for a warning comes from a low-quality point, acquisition quality should be checked first. If several point groups show stable spectral evidence, the warning is more likely to reflect an actual unit condition.

The results show that edge-aware Design reduces the amount of transmitted data. Only spectrogram summaries, source labels, confidence scores and evidence descriptors need to be uploaded for the regular monitoring; raw waveforms can be stored locally for a short time and then displayed. In line with recent studies on edge intelligence, local preprocessing and selective transmission have been proposed for industrial time-series monitoring [28]. First, the current experiment is conducted using artificial field data and regulated disturbance conditions [29]. The future deployment will have longer station-specific records and actual maintenance results [30]. Station-level validation needs to perform confidence calibration before threshold adjustment [31]. Another is that the new workflow can be introduced gradually. First, use the model to evaluate a station and, after confirming it is suitable, add its reliability data to the maintenance timetable. A model can also be trained with station-specific data without changing the main data interface. Therefore, a staged application will be used instead of immediate expert diagnosis. The remaining difference between the experimental evaluation and field deployment is mainly due to prolonged environmental changes. Fluctuations in the water table, seasonal variations of sediment cover, base erosion, etc., will also change the background spectrum. Therefore, the deployable version should have a small validation set for each station and compare the new warning with recently verified events before modifying the threshold. It will not change the architecture, but it will alter the control mode of the model after installation.

5. Conclusion

This study developed an edge-aware EfficientViT-FPN model for vibration source separation in hydropower units, and the overall conclusion is that multi-point spectrogram evidence, point-quality confidence and edge-deployable feature fusion should be evaluated together rather than through isolated single-sensor classification. First, the constructed monitoring dataset shows that the model achieves 92.8% accuracy, a macro-F1 score of 0.914 and an edge inference latency of 24.7 ms, while reducing mixed-source misclassification by 18.6% compared with the compact convolutional baseline. Second, the comparative results show that local spectral ridges, cross-point relationships and multi-resolution evidence improve recognition under partial-load, transition and noisy acquisition conditions. Third, the main limitation is that the current study uses simulated field-like records and simplified source labels, so long-term station records, verified maintenance events and operating logs are required for deployment validation.

The proposed output separates source label, confidence level, dominant point group, spectral evidence and review priority. This structure allows immediate inspection for high-confidence severe sources, trend tracking for stable minor patterns, remeasurement for uncertain high-severity windows and low-risk recording for background vibration. The confidence branch is especially useful for mixed-source intervals and missing-point cases because it prevents weak evidence from being treated as a confirmed fault. The model therefore supports maintenance decision making without removing expert review, and each warning remains linked to point-quality information, spectral bands and review evidence for later audit.

Future work should incorporate long-term station data, verified maintenance records and operating logs across different unit capacities and sensor layouts. Field deployment should also consider water-level variation, long-term sensor drift, sediment effects, operator feedback and station-specific background spectra. A verified field dataset can be used to establish review thresholds for different units and operating hours. These extensions will help convert the controlled evaluation into a practical hydropower condition-monitoring tool while preserving traceable evidence for maintenance scheduling and acquisition-quality review.

6. Declarations

6.1. Author Contributions

Marosi Zsombor contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision, project administration, and funding acquisition. Vadász Gusztáv contributes to validation, analysis, investigation, data collection, draft preparation. Sárközi Szabolcs contributes to conceptualization, validation, analysis. All authors have read and agreed with the manuscript before its submission and publication.

6.2. Funding and Acknowledgments

This research received no specific financial support from any funding agency.

6.3. Institutional Review Board Statement

Not applicable.

6.4. Informed Consent Statement

Not applicable.

6.5. Data Availability

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

6.7. Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors utilized Grammarly for linguistic polishing, logical organization and draft revision. The generative AI tool was not involved in research design, data collection, data analysis, or the equation of core conclusions. All outputs generated by AI were reviewed, revised and verified manually by the authors. The authors take full responsibility for all content of this manuscript.

7. References

[1] García Márquez, F. P., Peinado Gonzalo, A., & Papaelias, M. (2026). A comprehensive review of condition monitoring systems for hydropower stations: Technologies, applications, and future trends. Electric Power Systems Research, 251, 112339. https://doi.org/10.1016/j.epsr.2025.112339

[2] Lu, Z., Tao, R., Xiao, R., & Li, P. (2024). Forecasting the hydropower unit vibration based on adaptive variational mode decomposition and neural network. Applied Soft Computing, 150, 111040. https://doi.org/10.1016/j.asoc.2023.111040

[3] Dao, F., Zeng, Y., Zou, Y., & Qian, J. (2024). Fault diagnosis method for hydropower unit via the incorporation of chaotic quadratic interpolation optimized deep learning model. Measurement, 237, 115199. https://doi.org/10.1016/j.measurement.2024.115199

[4] Dao, F., Zeng, Y., Zou, Y., & Qian, J. (2024). Wear fault diagnosis in hydro-turbine via the incorporation of the IWSO algorithm optimized CNN-LSTM neural network. Scientific Reports, 14, 25278. https://doi.org/10.1038/s41598-024-77251-7

[5] Chen, Y., Wang, J., Chen, W., Dao, F., Zeng, Y., & Lv, S. (2024). Fault diagnosis in hydropower units based on chaotic Kepler optimization algorithm-enhanced BiLSTM model. Energy Reports, 12, 5163-5176. https://doi.org/10.1016/j.egyr.2024.11.008

[6] Liu, T., Kong, F., Yang, L., & Guo, Z. (2024). Operational risk assessment of hydropower units based on PSSCA-VMD-CNN-GBiLSTM and multi-feature fusion. Computers and Electrical Engineering, 118, 109412. https://doi.org/10.1016/j.compeleceng.2024.109412

[7] Lang, X., Nilsson, H., & Mao, W. (2024). Analysis of hydropower plant guide bearing vibrations by machine learning based identification of steady operations. Renewable Energy, 236, 121463. https://doi.org/10.1016/j.renene.2024.121463

[8] Luo, Y., Feng, C., Zhang, Y., Wang, J., & Zhang, Y. (2025). Vibration fault diagnosis of hydroelectric generator sets based on multi-dimensional feature extraction and CNN-BiLSTM. Measurement, 264, 120034. https://doi.org/10.1016/j.measurement.2025.120034

[9] Liu, Y., Xu, Y., Liu, J., Li, S., Cao, H., & Chen, J. (2024). Operational risk assessment for hydroelectric generating units using multi-head spatio-temporal attention and adaptive fuzzy clustering. Measurement Science and Technology, 35(2), 025011. https://doi.org/10.1088/1361-6501/ad076b

[10] Matania, O., Dattner, I., Bortman, J., Kenett, R. S., & Parmet, Y. (2024). A systematic literature review of deep learning for vibration-based fault diagnosis of critical rotating machinery: Limitations and challenges. Journal of Sound and Vibration, 590, 118562. https://doi.org/10.1016/j.jsv.2024.118562

[11] Kibrete, F., Woldemichael, D. E., & Gebremedhen, H. S. (2024). Multi-sensor data fusion in intelligent fault diagnosis of rotating machines: A comprehensive review. Measurement, 232, 114658. https://doi.org/10.1016/j.measurement.2024.114658

[12] Xiao, X., Li, C., He, H., Huang, J., & Yu, T. (2024). Rotating machinery fault diagnosis method based on multi-level fusion framework of multi-sensor information. Information Fusion, 112, 102621. https://doi.org/10.1016/j.inffus.2024.102621

[13] Zhang, X., Wang, H., Wang, C., Liu, M., & Xu, G. (2024). Time-segment-wise feature fusion transformer for multi-modal fault diagnosis. Engineering Applications of Artificial Intelligence, 138, 109358. https://doi.org/10.1016/j.engappai.2024.109358

[14] Liu, Y., Yu, Z., & Xie, M. (2024). Cascading time-frequency transformer and spatio-temporal graph attention network for rotating machinery fault diagnosis. IEEE Transactions on Instrumentation and Measurement, 73, 1-10. https://doi.org/10.1109/TIM.2024.3453312

[15] Xiao, Y., Shao, H., Wang, J., Yan, S., & Liu, B. (2024). Bayesian variational transformer: A generalizable model for rotating machinery fault diagnosis. Mechanical Systems and Signal Processing, 207, 110936. https://doi.org/10.1016/j.ymssp.2023.110936

[16] Li, Y., Zhou, Z., Sun, C., Chen, X., & Yan, R. (2024). Variational attention-based interpretable transformer network for rotary machine fault diagnosis. IEEE Transactions on Neural Networks and Learning Systems, 35(5), 6180-6193. https://doi.org/10.1109/TNNLS.2022.3202234

[17] Su, H., Xiang, L., & Hu, A. (2024). Application of deep learning to fault diagnosis of rotating machineries. Measurement Science and Technology, 35(4), 042003. https://doi.org/10.1088/1361-6501/ad1e20

[18] Cação, J., Santos, J., & Antunes, M. (2025). Explainable AI for industrial fault diagnosis: A systematic review. Journal of Industrial Information Integration, 47, 100905. https://doi.org/10.1016/j.jii.2025.100905

[19] Tang, S., Ren, Z., Zheng, L., & Wang, J. (2026). Explainable artificial intelligence in rotating machinery fault diagnosis: A comprehensive review and emerging trends. Sensors, 26(17), 5379. https://doi.org/10.3390/s26175379

[20] Shaalan, M., Ge, M., Hussain, M., & Rekik, Y. (2024). An explainable predictive maintenance strategy for multi-fault diagnosis of rotating machines using multi-sensor data fusion. Decision Analytics Journal, 10, 100425. https://doi.org/10.1016/j.dajour.2024.100425

[21] Feisa, T. T., Gebremedhen, H. S., Kibrete, F., Woldemichael, D. E., & Enyew, G. G. (2025). Artificial intelligence in fault diagnosis of industrial machinery: A comprehensive review. Structural Control and Health Monitoring, 2025, 4640227. https://doi.org/10.1155/stc/4640227

[22] Meitz, L., Senge, J., Wagenhals, T., Schöler, T., Hähner, J., Edinger, J., & Krupitzer, C. (2025). A literature review framework and open research challenges for predictive maintenance in Industry 4.0. Computers & Industrial Engineering, 206, 111193. https://doi.org/10.1016/j.cie.2025.111193

[23] Pierleoni, P., Belli, A., Palma, L., & Sabbatini, L. (2024). Resource-efficient Edge AI solution for predictive maintenance. Procedia Computer Science, 232, 348-357. https://doi.org/10.1016/j.procs.2024.01.034

[24] Lu, W., Wang, W., Qin, X., & Cai, Z. (2024). A lightweight kernel density estimation and adaptive synthetic sampling method for fault diagnosis of rotating machinery with imbalanced data. Applied Sciences, 14(24), 11910. https://doi.org/10.3390/app142411910

[25] Mhlanga, D. (2025). Enabling intelligent industrial automation: A review of machine learning applications with digital twin and Edge AI integration. Automation, 6(3), 37. https://doi.org/10.3390/automation6030037

[26] Hamasha, M. M., Albedoor, Q., & Hamasha, S. (2025). A comprehensive framework for IoT-driven predictive maintenance: Leveraging AI and edge computing for enhanced equipment reliability. Journal of Applied Engineering Science, 23(3), 471-486. https://doi.org/10.5937/jaes0-57002

[27] Cação, J., Santos, J., & Antunes, M. (2025). Explainable artificial intelligence based intelligent fault diagnosis: A systematic review from applications to insights. Reliability Engineering & System Safety, 264, 111935. https://doi.org/10.1016/j.ress.2025.111935

[28] Rony, M. R. A., Ali, M. H., & Chowdhury, M. S. (2026). Deep learning and few-shot learning-enabled machine health monitoring for rotating machinery: A critical review of vibration-based fault diagnosis, prognostics, and industrial deployment. Quality and Reliability Engineering International. https://doi.org/10.1002/qre.70302

[29] Coussement, K., Abedin, M. Z., Kraus, M., Maldonado, S., & Topuz, K. (2024). Explainable AI for enhanced decision-making. Decision Support Systems, 184, 114276. https://doi.org/10.1016/j.dss.2024.114276

[30] Gupta, S., Kumar, A., & Maiti, J. (2024). A critical review on system architecture, techniques, trends and challenges in intelligent predictive maintenance. Safety Science, 177, 106590. https://doi.org/10.1016/j.ssci.2024.106590

[31] Lash, M. T. (2024). Explainable machine learning for decision support systems: Requirements, methods and evaluation. Decision Support Systems, 187, 114304. https://doi.org/10.1016/j.dss.2024.114304