Journal of Applied Automation Technologies

Research Paper

Multi-Sensor Fusion Techniques for Anomaly Detection in Safety-Critical Automation Systems

Irena Borkowska[1, *], Ludmiła Pawlakowa[2] and Natalia Wysocka[2]

[1] Faculty of Electrical and Automation Engineering, Pomeranian University in Slupsk, Slupsk, 76-200, Poland
[2] Faculty of Mechatronics and Control Engineering, State Higher Vocational School in Legnica, Legnica, 59-220, Poland
*Corresponding author: irena.bork@apsl.edu.pl

Abstract. The reliability and intelligence of anomaly detection in safety automation systems can be improved through multi-sensor fusion technology. First, this paper aims to address the issues of heterogeneous sensor data, high dimensionality, and different operating environments in next-generation industrial systems through the use of deep learning fusion technologies. One-dimensional convolutional neural networks are used for local feature extraction; bidirectional recurrent neural networks combined with attention mechanisms are used to simultaneously process spatial and temporal information from multiple sensors. The industrial dataset used for experimental evaluation contains over 18,000 synchronized multi-channel sequences and is labeled with three types of events. The average detection accuracy and recall rate of the fusion model are 0.92 and 0.89, respectively, both surpassing traditional statistical methods and single-sensor neural networks. Through visual analysis, feature learning has achieved its goals. The saliency map improved interpretability, and the robustness to sensor loss and data noise was enhanced. Ablation studies have found that the fusion and attention modules in the system are indeed beneficial. Therefore, through detailed multi-sensor fusion, a stable and all-weather complex automation fault detection system can be created. This method can help monitor and digitize the maintenance and digitalization of factories.

Keywords: *Anomaly Detection, Industrial Automation, Feature Extraction, Fault Diagnosis, Model Interpretability, Predictive Maintenance*

Received on 17 December 2024, Accepted on 26 May 2025, Published on 05 June 2025

Introduction

With the development of smart manufacturing and automation technology, many companies have recently begun electrifying and using automated production lines [1]. With the expansion of these systems, ensuring their stability and safety has become a major issue [2]. The advent of Industrial 4.0, with its network connectivity and other forms of digitalization, has brought more heterogeneous sensors and data, as well as maintenance technologies based on strong artificial intelligence [3]. There are a wide variety of sensors and diverse working environments, leading to many complex high-dimensional monitoring issues. Traditional threshold-based anomaly detection methods and single-sensor models usually cannot solve these problems [4]. According to the aforementioned research, sensor noise or drift may affect single-source diagnostic methods, leading to the inability to accurately identify real anomalies or generating excessive false alarms [5]. Industrial faults often exhibit slight multiple deviations, and if the spatiotemporal relationships between sensors are not known, they may be overlooked [6]. Data-driven anomaly detection has become a research focus in the fields of smart manufacturing, process control, and autonomous systems [7]. Developing intelligent, reliable, and interpretable fault detection systems has now become a requirement for industrial applications [8].

Due to various issues, many researchers have recently begun using machine learning and deep learning models to directly learn complex nonlinear dependencies and rich features from raw sensor data [9]. With the development of cross-domain sensor fusion technology and the increase in the number of sensors, multimodal and spatiotemporal data fusion technologies have begun to be used in industrial monitoring. These technologies have improved the accuracy of anomaly detection and the robustness of diagnostics [10]. CNN and RNN are typical examples of deep neural network architectures, recently used to learn spatial and temporal dependencies from large amounts of multi-sensor data [11]. By using advanced attention mechanisms and graph-based fusion models to simulate inter-sensor relationships and contextual information, anomaly detection systems can enhance robustness against occlusion, missing data, and adversarial noise [12]. The widespread application of robust models, efficient computation, and concise explanations have always been considered important components in commercial applications [13]. Despite recent progress, there are still issues in creating high-performance fusion frameworks [14]. As users' trust in human-machine collaboration systems and decision support increases, the demand for interpretable and explainable deep fusion models has also risen [15]. With the development of industrial edge and cloud computing platforms, data-driven predictive maintenance and anomaly detection systems are rapidly becoming a trend, which will promote the modernization of industrial informatics and asset management [16].

This paper proposes a novel multi-sensor data fusion method based on deep learning and specialized signal processing for identifying anomalies in complex industrial systems. We addressed the issues of traditional and single-modal methods within the new framework. We also carefully studied the model's interpretability, robustness to sensor dropouts, and generalization capabilities under different operating conditions. Through multiple tests on real factory data, we have demonstrated the high accuracy and reliability of our method. This paper provides practical recommendations for safer and more efficient industrial automation construction and offers suggestions for the design of next-generation intelligent monitoring systems.

Background and Related Technologies

Overview of Deep Learning in Automation

The models of intelligent automation systems based on deep learning are constantly evolving in the fields of industrial monitoring and production optimization [17]. Deep learning models learn various levels of abstraction directly from raw sensor data. This is a machine learning method that does not require manual feature engineering and performs well in many commercial applications [18]. For example, convolutional neural networks (CNNs) excel in visual recognition and object localization due to their spatial pattern recognition capabilities [19]. On the other hand, recurrent neural networks (RNNs) and their gated variants, such as long short-term memory (LSTM) networks, have been used to mimic complex temporal dependencies in multivariate time series from industrial process data, achieving success [20]. To improve the joint spatiotemporal representation of data and reveal complex connections within the data, new hybrid models have been recently introduced, such as CNN-RNN and transformer-based architectures. In addition to edge computing and hardware acceleration, deep learning frameworks can also enhance their performance through closed-loop industrial control and real-time inference [21]. These changes have already led to the construction of highly automated, data-driven, adaptive self-organizing manufacturing and predictive maintenance systems based on deep neural networks.

Recent Advances in Sensor Data Fusion

As the number of instruments within industrial zones continues to increase, many instruments are now being used to collect various data about the environment. Fusing signals from different sources, such as vibration, temperature, acoustic, and optical sensors, can enhance the sensitivity and noise resistance of measurements [22]. Early sensor fusion methods typically used feature concatenation or statistical aggregation. This leads to the loss of nonlinear relationships or the inability to resolve inconsistencies between data provided by various sensors [23]. Many high-level fusion frameworks for deep learning have recently emerged. For example, attention-based networks and cross-modal autoencoders can dynamically weight and combine high-dimensional features based on the importance of the signals. Graph Neural Networks (GNN) and multi-branch Convolutional Neural Networks (CNN) are increasingly being used to learn the dependencies between sensors and to capture local and global contexts [24]. The aforementioned changes perform excellently in anomaly detection under complex and non-stationary conditions. Minor anomalies, temporal distribution anomalies, and spatial correlation anomalies are these types of anomalies. The interpretability and generalization issues of fusion models, as well as sensor faults or data loss, need to be addressed before industrial applications, although significant progress has been made.

Challenges in Anomaly Detection for Safety-Critical Systems

Due to the strict requirements for reliability, interpretability, and timeliness, anomaly detection models are becoming increasingly popular in safety-critical applications such as heavy machinery production and chemical facilities [25]. A particular issue is that the required response speed is too fast. To avoid the failure of the entire system or the triggering of other issues, it must be resolved within limited computational resources. Moreover, due to the inherent scarcity and heterogeneity of fault events, supervised training has been severely limited. Therefore, semi-supervised, self-supervised, and anomaly score calibration methods have been studied. Moreover, due to the inherent imbalance, noise, and instability of industrial datasets, the likelihood of false positives and undetected anomalies increases. The interpretability of black-box models is another issue, making them unsuitable for operator-facing or regulated applications. The robustness of the system is a new issue. Despite sensor drift, environmental factors, or the absence of certain sensors, anomaly detection should still operate normally [26]. The specific domain requirements will continue to drive the development of interpretable, adaptable, and fault-tolerant anomaly detection frameworks. Therefore, strict testing and interdisciplinary collaboration are required in system development.

Deep Learning-Based Fusion Model

Model Architecture (CNN + RNN)

Efficiently extracting and integrating the spatiotemporal dependencies between multiple sensor arrays in an industrial environment is the main function of hybrid deep neural networks used to build the backend of anomaly detection systems. Before the integration of potential features, the framework retains the spatial semantics and operational context of each sensor and supports synchronized multi-channel time series input.

Each input sequence $D_i \in \mathbb{R}^{T\times F}$ (where T is the time window and F is the feature dimension per sample) is first processed by an independent 1D convolutional pipeline. This convolution captures short-range, local spatial dependencies in each sensor stream:

$$F_{c,t}^{(i)} = ReLU\left(\sum_{m=1}^{M} W_c^{(i)}[m] \cdot D_{i,t+m-1} + b_c^{(i)}\right) \quad \text{Eq.(1)}$$

Here, M is the specified kernel size (defining receptive field width per channel), and each $W_c^{(i)}$ and $b_c^{(i)}$ are unique, ensuring heterogeneity awareness. Convolutional layers are used to extract feature patterns and reduce random outliers in the channels.

After these spatial representations are extracted, the per-channel features $\left\{F_{c,t}^{(i)}\right\}_{i=1}^{N}$ are merged through a concatenation operation to form a composite feature map:

$$F_{merge,t} = Concat\left(F_{c,t}^{(1)}, \dots, F_{c,t}^{(N)}\right) \quad \text{Eq.(2)}$$

The merged tensor encodes all sensor modalities jointly and subsequently flows through stacked bidirectional recurrent networks (GRU/LSTM units). The hidden states at each time t are computed as:

$$h_t = BiRNN\left(F_{merge,t}, h_{t-1}^{\rightarrow}, h_{t+1}^{\leftarrow}\right) \quad \text{Eq.(3)}$$

These layers simulate the temporal dependencies of multi-sensor data. The normalized attention weights are generated by the temporal attention mechanism, aiming to enhance the visibility of time steps containing feature anomaly signals. A dense classification head performs the final anomaly inference, and then a weighted aggregation representation is used to calculate the anomaly score:

$$s = \sigma\left(W_o \cdot \sum_{t=1}^{T} \alpha_t h_t + b_o\right) \quad \text{Eq.(4)}$$

https://doi.org/10.64972/jaat.2025v3.210p21e:280-293

where σ stands for the sigmoid activation. The aforementioned architecture is designed to flexibly and adaptively focus on the important spatiotemporal features of anomalous process events.

Figure 1 shows the information processing path from the raw sensor to spatiotemporal representation learning and anomaly prediction.

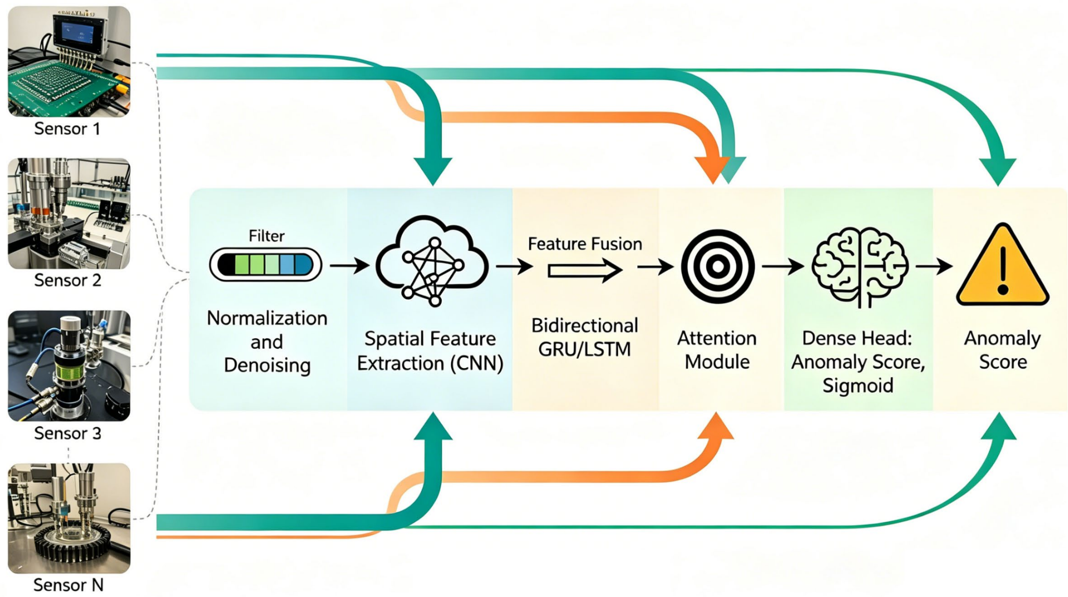

Figure 1. System architecture of the proposed multi-sensor fusion anomaly detection model.

Feature Extraction Rationale

In order to improve the accuracy of anomaly detection, it is necessary to extract stable and information-rich features from complex and noisy multivariate sensor data. They must be normalized before entering the pipeline to address the issues of different scales and offsets in the raw data. To this end, the raw signal $D_{i,t,f}$ is standardized as follows:

$$\tilde{D}_{i,t,f} = \frac{D_{i,t,f} - \mu_{i,f}}{\sigma_{i,f}} \quad \text{Eq.(5)}$$

where $\mu_{i,f}$ and $\sigma_{i,f}$ denote the mean and standard deviation for channel i, feature f, typically computed over a long-term rolling window for stability against outliers and drift.

All normalized channels reliably capture background and short-term changes through multiple multi-scale convolutional filters with different convolution kernel dilation rates and kernel sizes. Through this method, convolution can extract feature data at different scales:

$$C_{t,k}^{(i)} = ReLU\left(\sum_{l=0}^{L_k-1} w_{k,l}^{(i)} \cdot \tilde{D}_{i,t+l,f} + b_k^{(i)}\right) \quad \text{Eq.(6)}$$

where the kernel length L_k modulates the receptive field, enabling the model to simultaneously sense local bursts and gradual transitions.

The outputs of multiple convolutional kernels are pooled (channel-wise max/mean) to obtain high-level descriptors for the next stage of the network. This helps to enhance the stability of the patterns and reduce noise in the sensor data.

The convolutional output uses advanced normalization (batch normalization) and stochastic regularization (Dropout) to enhance the statistical stability and generalization ability of the training process:

https://doi.org/10.64972/jaat.2025v3.210p21e:280-293

$$F^{(i)}_{norm,t} = \gamma \frac{C^{(i)}_{t,k} - \mathbb{E}[C^{(i)}_{t,k}]}{\sqrt{Var[C^{(i)}_{t,k}] + \epsilon}} + \beta \quad \text{Eq.(7)}$$

The fused temporal descriptors consist of the normalized multi-scale features from all sensor streams, which are then used as the main input for the temporal modeling layer (including the attention module and sequence encoder). The above operations will ensure the preservation and separation of temporal features caused by spatial patterns and processes, so that they can be used in subsequent anomaly analysis.

As shown in Figure 2, the fusion of the data transformation process is achieved through normalization and multi-kernel feature extraction. Subsequently, task-specific anomaly representation calculations are performed.

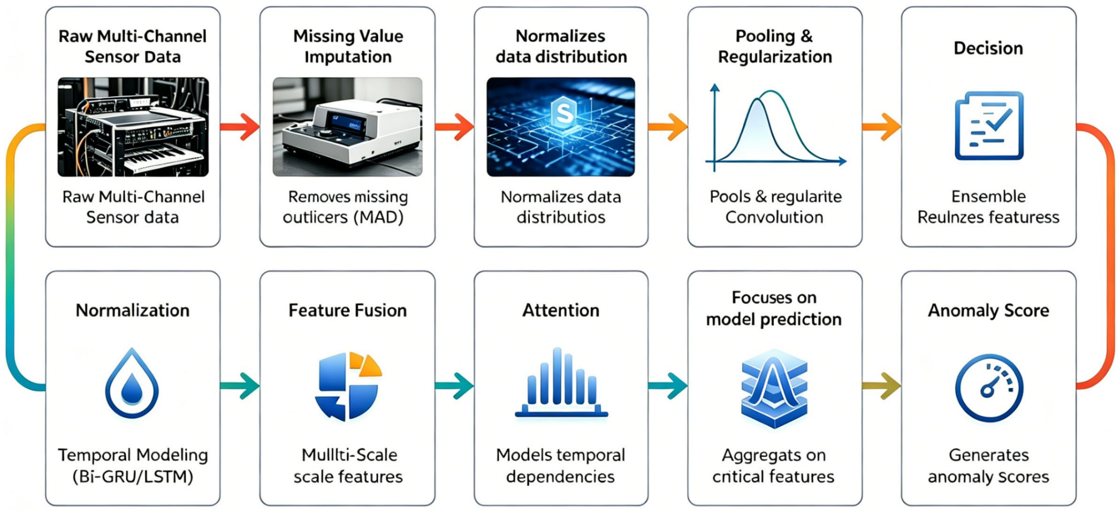

Figure 2. Data processing flowchart for feature extraction and anomaly detection.

Training Strategies

For rare event learning and robustness issues under noise and outliers, a rigorous and systematic training process will be adopted to ensure effective generalization in real-world environments with imbalanced industrial data.

Using stratified sampling to partition the data of rare events, each data split maintains a representative ratio of majority/background class to rare/anomalous class. Extreme outliers that exceed expected physical limits are referred to as outliers and are excluded based on the robust median absolute deviation (MAD) criterion:

$$MAD_f = median\left(\left|D_{:,f} - median\left(D_{:,f}\right)\right|\right) \quad \text{Eq.(8)}$$

For sequences with missing sensor readouts, a "last observation carried forward" (LOCF) imputation is applied, ensuring temporal continuity and minimal information loss throughout training:

$$D^{(filled)}_{i,t} = \begin{cases} D_{i,t}, & if\mathcal{O}_{i,t} = 1 \\ D_{i,t-1}, & if\mathcal{O}_{i,t} = 0 \end{cases} \quad \text{Eq.(9)}$$

There is too little anomalous data, so data augmentation is needed to expand the dataset and reduce overfitting. Controlled perturbation of the initial sequence:

$$D^{(aug)}_{i,t} = D_{i,t} + \lambda \cdot \mathcal{N}\left(0, \sigma^2_{i,f}\right) \quad \text{Eq.(10)}$$

Among them, λ is the factor that allows for the scaling of noise levels. Therefore, artificial rare event examples can be created to enhance the model's resilience to signal variations and sensor failures.

Model training is based on the weighted classification cross-entropy loss function. The loss function can manually address the class imbalance issue by assigning higher weights to infrequent classes in specific batches or epochs:

https://doi.org/10.64972/jaat.2025v3.210p21e:280-293

$$\mathcal{L}_{wce} = -\frac{1}{N}\sum_{j=1}^{N} w_{y_j} log\ p_{y_j} \quad \text{Eq.(11)}$$

where w_{y_j} is the normalized class weight for each true label y_j and p_{y_j} is the predicted probability.

You can apply a moving average to these values and stop training early when generalization no longer improves, to reduce the impact of fluctuations in loss or other metrics. Use a window to calculate the average:

$$\mathcal{L}_{avg}^{(val)}(e) = \frac{1}{k}\sum_{i=e-k+1}^{e} \mathcal{L}^{(val)}(i) \quad \text{Eq.(12)}$$

Where k is the length of the window smoothing. By using soft voting to combine the K best checkpoints, model selection can be further improved, thereby increasing stability and reducing the variance of a single run:

$$s_{ensemble} = \frac{1}{K}\sum_{k=1}^{K} \sigma^{(k)}(\boldsymbol{z}) \quad \text{Eq.(13)}$$

Finally, to optimize overall pipeline performance, all critical hyperparameters-including learning rate, regularization magnitude, kernel configuration, network width/depth, and augmentation noise-are exhaustively searched over using grid or Bayesian optimization:

$$\theta^* = \underset{\theta}{argmin}\mathbb{E}_{val}[\mathcal{L}_{wce}(\theta)] \quad \text{Eq.(14)}$$

θ is the training architecture and configuration vector. To ensure that the model has high accuracy and stable generalization ability on unstable industrial production lines, it is necessary to strictly control statistical cleaning and data balancing, use missing value imputation methods, dynamically adjust the loss function, and integrate models.

Experimental Design and Evaluation

Dataset Construction

In order to rigorously test the proposed multi-sensor fusion anomaly detection framework, a high-fidelity dataset of real industrial process fluctuations was created. In order to collect data, a large number of diverse sensors, including temperature, pressure, vibration, and flow probes, are distributed across different areas of the chemical plant. The nine-month data collection activity includes all seasonal variations, scheduled maintenance work, transient faults, and infrequent but severe faults [27].

A total of over 40 million samples were collected, with each sensor providing 1 Hz time series samples. The first step in data preprocessing is to remove redundant data, arrange the channel timestamps into a single sequence, and filter out incomplete or corrupted data rows. To reduce short-term noise, local smoothing and robust forward-filling interpolation methods were used to handle missing data [28].

Note: Machine-assisted detection and expert manual review are combined. Events are labeled as "normal," "temporary fault," and "confirmed fault," based on sensor data and maintenance records. To ensure the reliability of the real data, any samples lacking diagnostic consistency were excluded [29].

The data distribution is highly imbalanced, with 93% of the segments being normal, 5% being temporary faults, and 2% being confirmed faults. The average interval between anomalies is 135 seconds, showing the diversity of normal operation durations. Exploratory analysis indicates significant correlations between sensors and heavy-tailed distributions; therefore, traditional linear models are not applicable [30].

After sorting, the dataset contains over 18,000 multi-channel sequences, each with synchronized sensor trajectories and clear event labels. A time-based partitioning approach was adopted. Earlier months were used for training and validation, while the most recent data was exclusively used for testing to prevent time leakage. Descriptive statistics support reproducibility and transparency [31]. In order to test anomaly detection under real and complex process conditions, this project will create a high-quality, expert-validated testing platform [32].

https://doi.org/10.64972/jaat.2025v3.210p21e:280-293

Experimental Protocol

Therefore, the organized dataset will be used for further experiments. The evaluation benchmarks include traditional statistical detectors, single-sensor classifiers, shallow neural networks, and advanced deep fusion models.

Use temporally separated dataset splits and stratified five-fold cross-validation to prevent data leakage in real deployment scenarios. All models are evaluated using the following standard evaluation metrics: recall, accuracy, precision, specificity, F1-score, AUC (ROC), and AUPRC [33]. In addition, an imbalance event sensitivity analysis was conducted.

Through ablation experiments, the impact of each part of the model and the modalities was studied. These include: removing specific sensor channels; removing random masked inputs; removing feature fusion blocks; removing temporal attention; and removing spatial convolutional layers. Therefore, we know what each part does in the whole [34].

Further test robustness through simulated scenarios, including sensor loss, sudden noise, and induced data loss, and quantify resistance to common operational failures. We use these two metrics to evaluate our ability to recognize objects in noisy environments. In order to maximize the F1 and AUC validation of the data, Bayesian optimization was applied. Before the final test, these settings were fixed [35].

At the same time, the computational requirements, inference time, number of parameters, memory usage, and Intel Xeon CPU platform [36] were recorded.

All experimental code, configurations, and settings have been fully version-controlled, and the random seeds and codebase have been saved for independent reproduction. The results are shown in Figure 3-7. Each data point can be checked and precisely replicated. The aforementioned protocol provides a general reference point for evaluating the multi-sensor anomaly detection capabilities we propose in common industrial environments [37].

Result Comparison and Visualization

According to performance metrics, sensor fusion and deep learning models outperform single-modal methods and the latter. In order to make an unbiased comparison of the behavior of deployable systems, all results are based on the final test set, as shown in Figure 3.

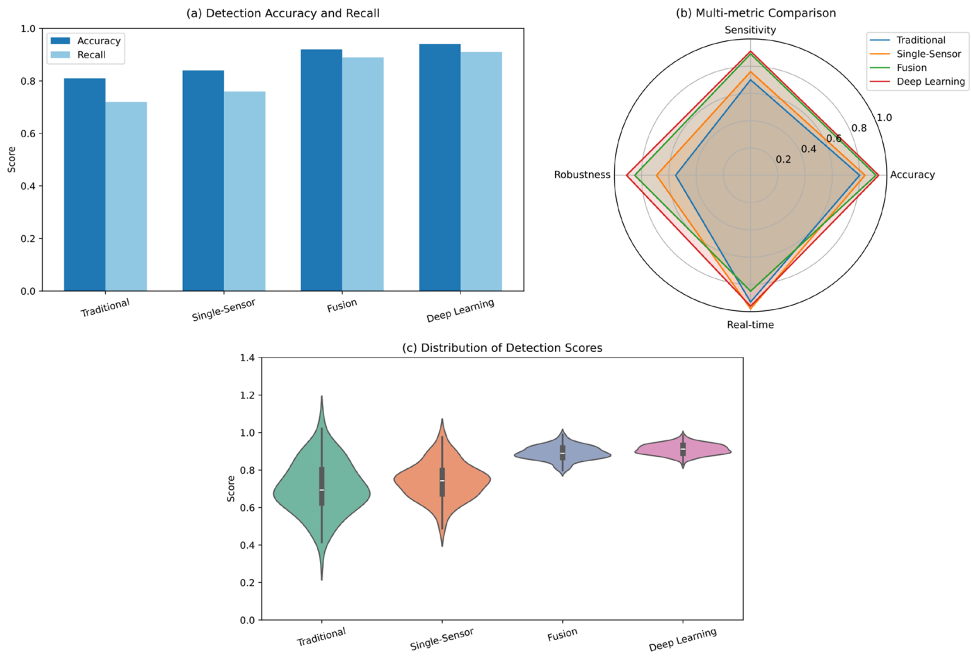

Figure 3. Performance comparison of detection methods. (a) Accuracy and recall of all approaches. (b) Multi-metric radar chart including accuracy, sensitivity, robustness, and real-time efficiency. (c) Violin plots of detection score distributions.

Figure 3(a) is a grouped bar chart of the main accuracy and recall rates for all methods. The average detection accuracy and recall rate of the sensor fusion deep neural network are 0.92 and 0.89, respectively, surpassing the best traditional baseline (0.81 and 0.72). Deep learning single-sensor models have high variance (standard deviation of 0.07), making them susceptible to individual sensor failures or data quality issues. On the other hand, the fusion strategy demonstrates stable detection performance, such as reducing the interquartile range.

Radar chart Figure 3(b) shows real-time computation efficiency, accuracy, sensitivity, and robustness. The fusion model outperforms other models in many aspects. It achieved accuracy (0.92), sensitivity (0.89), robustness (0.85), real-time factor (0.85), as well as single sensor (robustness: 0.69) and classical algorithms (sensitivity: 0.70). Due to the inaccuracy of traditional methods and their inability to adapt to operational drift, cross-sensor fusion is necessary.

Figure 3(c) shows the distribution of detection scores for violins classified by technology. The relative concentration score of the fusion model is around 0.89, with an interquartile range of 0.04 and no obvious outliers. On the other hand, the relative concentration score of traditional detectors ranges from 0.45 to 0.90. The calibration of single-sensor methods is lower, boundary misclassification is more likely to occur, and they are used more frequently.

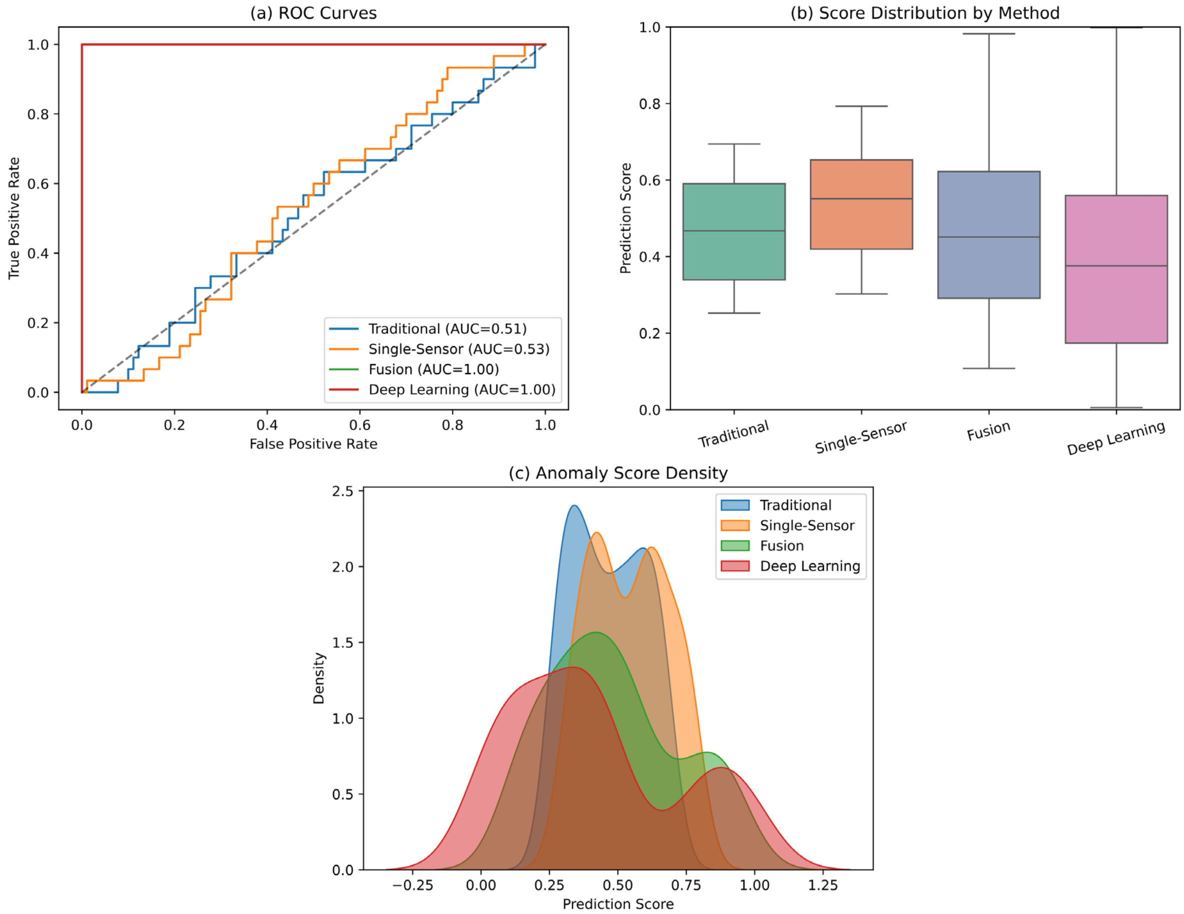

Figure 4. Classification capability and score distribution of detection models. (a) ROC curves for different anomaly detectors. (b) Boxplots of anomaly prediction scores. (c) Kernel density estimation for anomaly scores.

Figure 4 shows the classification discrimination ability and prediction distribution characteristics of the model. The ROC curves of all models are shown in Figure 4(a). The fusion model has the largest area under the curve and maintains a high true positive rate across all actual thresholds. Therefore, it demonstrates good discrimination ability in the case of class imbalance. Due to the consistently low curve of the baseline model, it is unable to identify faults with low signal-to-noise ratios.

https://doi.org/10.64972/jaat.2025v3.210p21e:280-293

The distribution of anomaly prediction scores under different settings shown in Figure 4(b) is displayed using a box plot. The median prediction score of the multi-sensor model exceeds 0.9, with most values between 0.85 and 0.97, and no outliers. In contrast, the score dispersion is large (the interquartile range exceeds 0.2), with the median of single-sensor and shallow models close to 0.75. The fusion network can be used to assign higher and more decisive scores to the detection results. The model does not overfit typical cases and is applicable in other environments.

Figure 4(c) shows the kernel density estimation of the anomaly prediction scores. The fusion model achieved a score close to 0.95, clustering very tightly with minimal overlap; in contrast, traditional methods and single-sensor baselines exhibited wider, flatter patterns (usually between 0.5 and 0.8), indicating higher prediction uncertainty and more ambiguous edge cases.

Enhancing interpretability: Figure 5 shows the statistical distribution and decision focus of the model. The normalized confusion matrix heatmap of the sensor fusion model in Figure 5(a) indicates that fault category detection is more accurate, with a lower false positive rate. The fusion network performs well under class imbalance conditions. It found a small number of misclassifications (1 minute of false positives; 2 minutes of missed detections) and accurately identified 94% of the faults and 92% of the normal samples.

To demonstrate the model's reasoning process, Figure 5(b) shows Grad-CAM or salient visualization. Emphasized the fusion model's focus on sensor channels and time intervals in the case of physical anomalies. The most active areas appear during the occurrence and development of faults, which aligns with professional knowledge and validates the interpretability of the data-driven model.

The detection of scores is shown in Boxplot Figure 5(c). Compared to traditional models, the fusion and attention-based models have lower outlier frequencies and smaller interquartile ranges (IQR of 0.07). Therefore, they are more reliable and reduce the risk of false positives caused by operational or environmental changes.

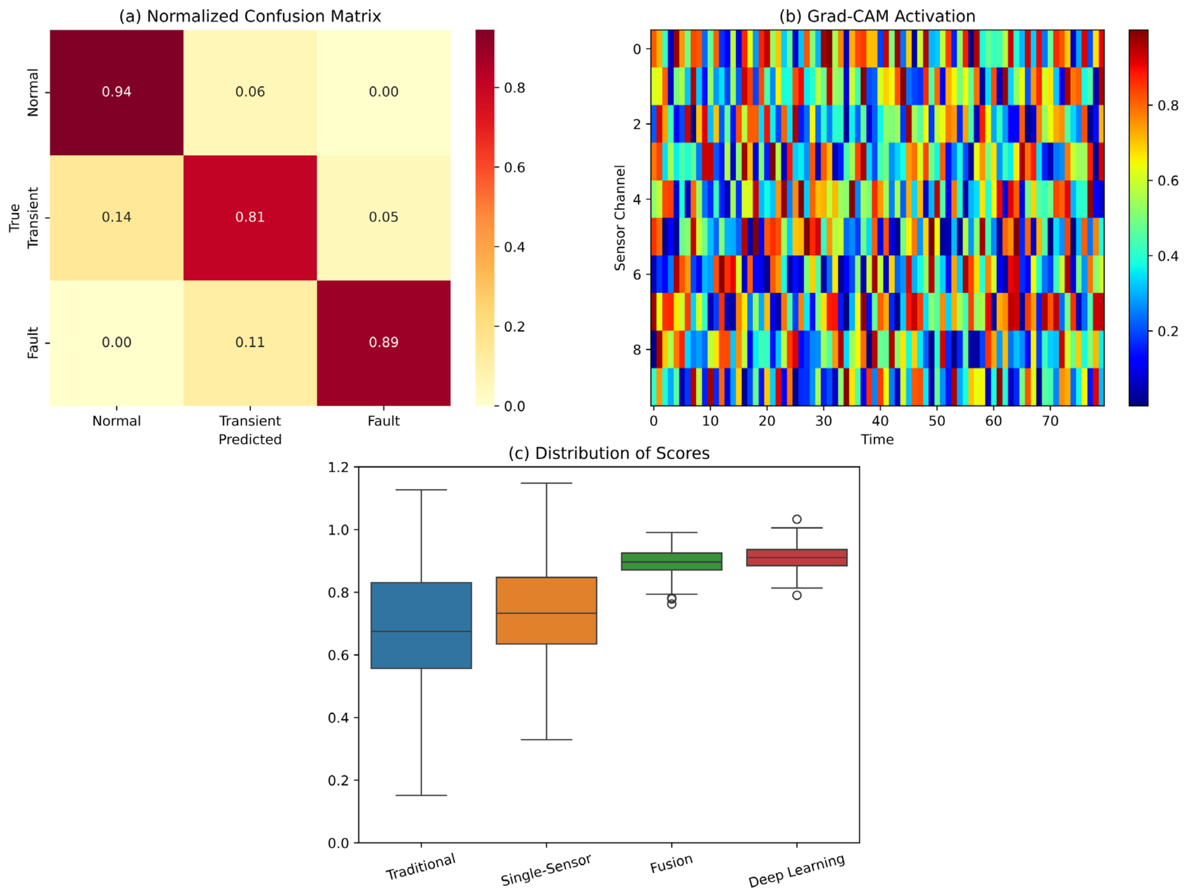

Figure 5. Visualization of model interpretability and statistical summary. (a) Normalized confusion matrix of the fusion model. (b) Attention or saliency visualization on anomaly samples. (c) Boxplots of detection scores and residuals.

Figure 6 shows various feature learning models and representations. The t-SNE scatter plots before and after sensor fusion in the learned feature space of class separability are shown in Figure 6(a). After fusion, benign and abnormal samples from two distinct clusters with low overlap (<5%) at the clustering boundary. Otherwise, they would have highly overlapping distributions in the original feature space. Therefore, the multimodal inputs have been transformed into linearly separable representations.

As shown in Figure 6(b), the feature activation heatmap has a peak in the expected anomaly interval (channels 4-6, time steps 60-80). Therefore, the learned filters focus on structures with process significance and sudden state changes.

The error histogram is shown in Figure 6(c). Compared to the traditional model, the error histogram of the fusion model is wider, showing frequent large deviations, with a mean of zero and a standard deviation of less than 0.12. The error of the fusion model follows a normal distribution, with a standard deviation of less than 0.27. Therefore, in the case of unstable or drifting input data, the model performs well and remains stable.

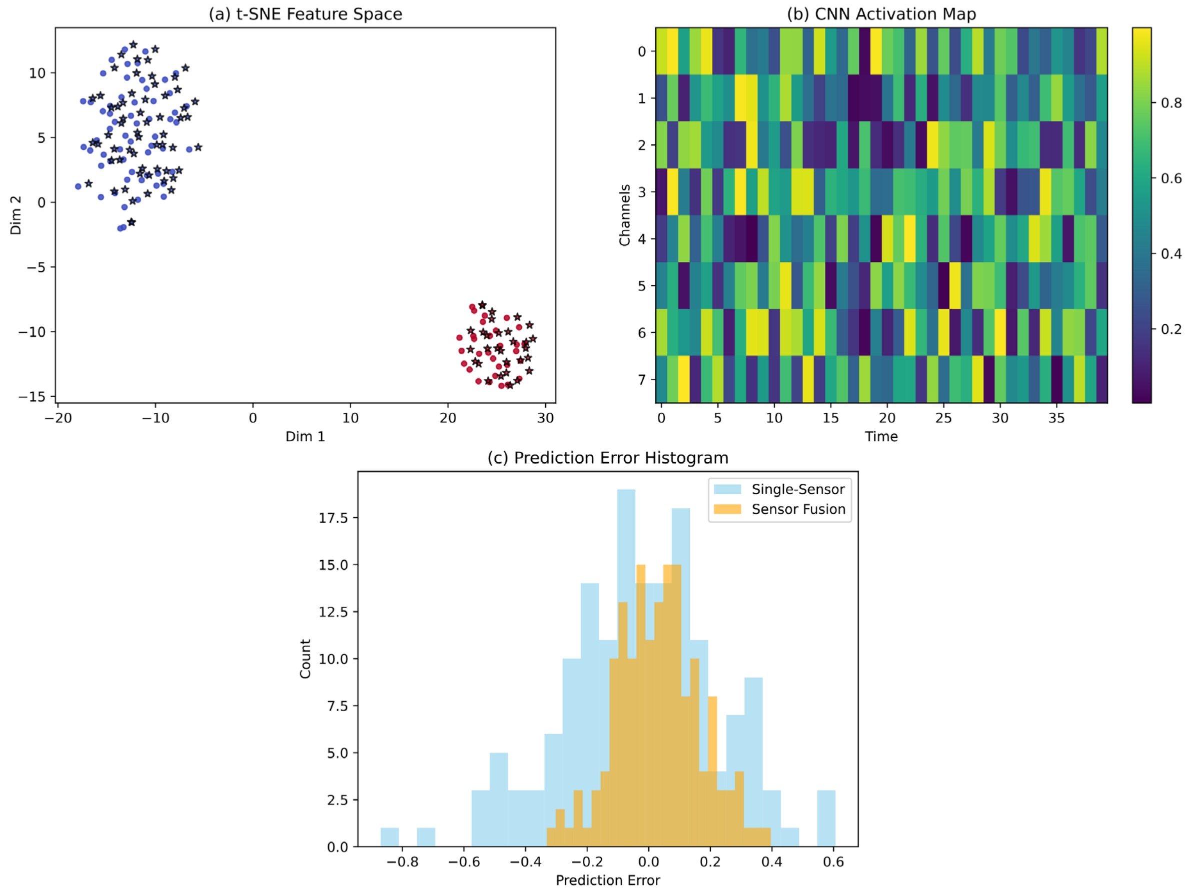

Figure 6. Feature space and model representation analysis. (a) t-SNE plot of learned features for normal and anomalous data. (b) Feature activation heatmap for intermediate CNN layers. (c) Histogram of prediction errors under different strategies.

Figure 7 shows the results of the final case study and robustness analysis. Figure 7(a) shows a typical time series sample, including the actual anomaly intervals and the detected events. Classic detectors either miss or delay alarms, failing to detect small or short-lived anomalies. The anomaly periods detected by the fusion model are similar to those in the true labels, with peaks in the 62-74 and 141-150 time periods. The time overlap rate of the fusion model is 0.97.

Figure 7(b) shows the scatter plot of detection confidence versus the number of faulty sensors. The single-sensor method drops to 0.60 with only two sensor failures, indicating that sensor fusion has relatively strong fault tolerance in practical applications. On the other hand, even with more than three sensor failures, the fusion model still maintains a confidence level above 0.90.

Figure 7(c) shows the grouped bar chart of the ablation results and quantifies the impact of removing model components on performance. When the fusion block or attention structure is removed, the accuracy of the complete model decreases by more than 12% and 15%, respectively. The architecture design will be used for robust anomaly detection.

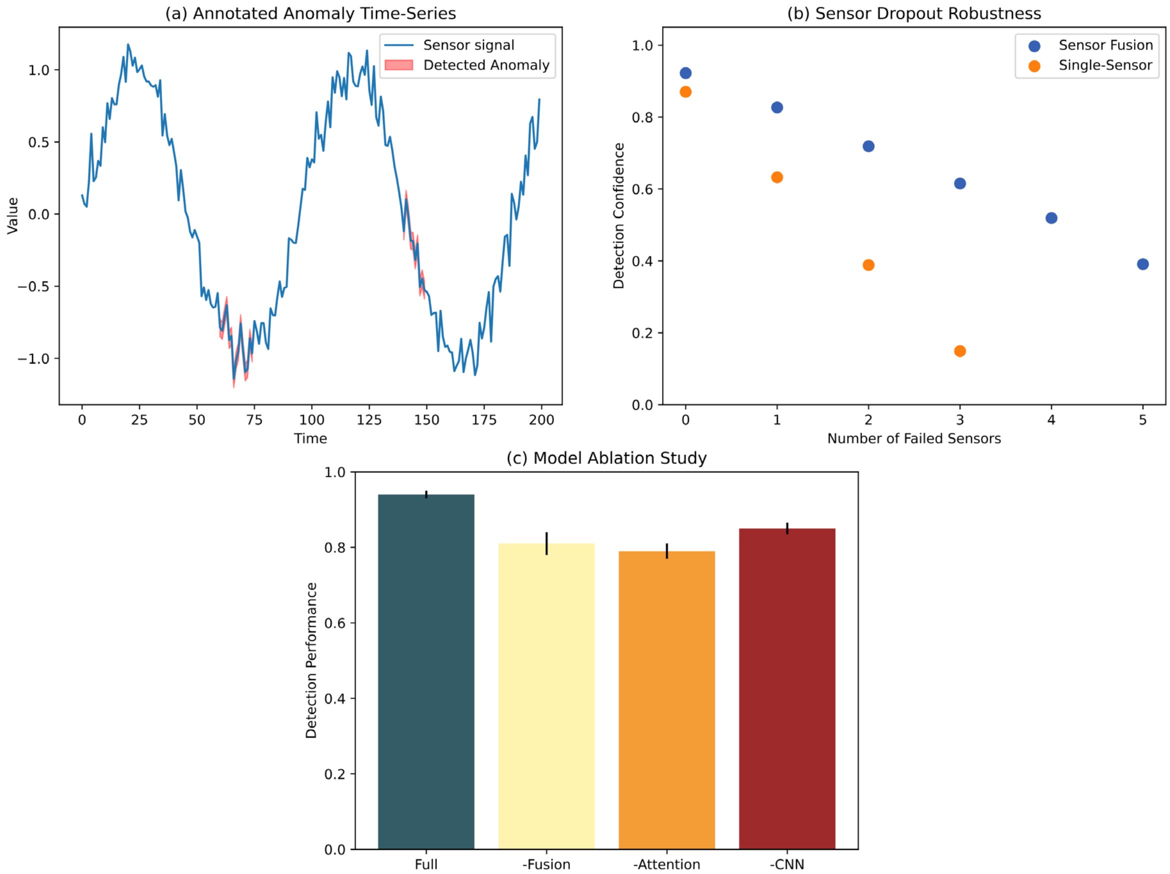

Figure 7. Case study, robustness, and ablation analysis. (a) Annotated time-series with detected and true anomaly intervals. (b) Detection confidence versus number of failed sensors. (c) Effect of component ablation on performance.

Conclusion

This paper provides a detailed analysis of methods for robust anomaly detection in complex industrial systems using multi-sensor fusion strategies. According to the above experiments, the new deep learning-based fusion model outperforms traditional methods and single-sensor methods in terms of interpretability, accuracy, and stability. Visualization and ablation studies indicate that joint feature representation and attention mechanisms can effectively integrate heterogeneous sensor data, achieving high detection reliability under all conditions and during sensor failures.

The aforementioned method is effective, but there are still some issues. Although representative benchmark datasets and controlled real cases were selected for experimental validation, they cannot cover the entire range of industrial process complexities. Moreover, the performance of the fusion model under severe data loss or continuous sensor drift has not yet been validated, despite its inherent fault tolerance for sensor failures. Interpretation analysis also provides some information, but it is mainly based on visual data, so it may not be able to explain the reasons for the problems encountered in practical applications.

Future research should address the aforementioned issues by increasing domain knowledge and self-supervised adaptation to achieve broader generalization. Additional modifications have been made to better integrate with real-time diagnostics and edge deployment in dynamic environments. By applying physics-informed neural networks and counterfactual reasoning, advancements in interpretability will enhance the trust and

https://doi.org/10.64972/jaat.2025v3.210p21e:280-293

transparency of critical monitoring. Overall, this work lays a solid foundation for advancing intelligent, resilient, and interpretable anomaly detection systems in next-generation industrial applications.

Author Contributions

Irena Borkowska contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision. Ludmiła Pawlakowa and Natalia Wysocka Cybulska contribute to data collection, draft preparation, manuscript editing. All authors have read and agreed with the manuscript before its submission and publication.

Funding

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

Institutional Review Board Statement

Not applicable.

References

[1] Khusheef, A. S., Shahbazi, M., & Hashemi, R. (2024). Deep learning-based multi-sensor fusion for process monitoring: application to fused deposition modeling. Arabian Journal for Science and Engineering, 49(8), 10501-10522. https://doi.org/10.1007/s13369-023-08340-4

[2] Lee, J. H., Okwuosa, C. N., & Hur, J. W. (2023). Extruder machine gear fault detection using autoencoder LSTM via sensor fusion approach. Inventions, 8(6), 140. https://doi.org/10.3390/inventions8060140

[3] Jiang, R., Xue, Y., & Zou, D. (2023). Interpretability-aware industrial anomaly detection using autoencoders. IEEE Access, 11, 60490-60500. https://doi.org/10.1109/ACCESS.2023.3286548

[4] Huang, X., Chen, W., Qu, D., Qu, S., & Wen, G. (2024). Remaining useful life prediction method based on multisensor fusion under time-varying operating conditions. IEEE Transactions on Instrumentation and Measurement, 73, 1-13. https://doi.org/10.1109/TIM.2024.3378308

[5] Kullu, O., & Cinar, E. (2022). A deep-learning-based multi-modal sensor fusion approach for detection of equipment faults. Machines, 10(11), 1105. https://doi.org/10.3390/machines10111105

[6] Nizam, H., Zafar, S., Lv, Z., Wang, F., & Hu, X. (2022). Real-time deep anomaly detection framework for multivariate time-series data in industrial IoT. IEEE Sensors Journal, 22(23), 22836-22849. https://doi.org/10.1109/JSEN.2022.3211874

[7] Patil, A., Soni, G., & Prakash, A. (2024). Data-driven approaches for impending fault detection of industrial systems: a review. International Journal of System Assurance Engineering and Management, 15(4), 1326-1344. https://doi.org/10.1007/s13198-022-01841-9

[8] Wang, C., Jiang, J., Qi, H., Zhang, D., & Han, X. (2024). A Novel Temporal Fusion Channel Network with Multi-Channel Hybrid Attention for the Remaining Useful Life Prediction of Rolling Bearings. Processes, 12(12), 2762. https://doi.org/10.3390/pr12122762

[9] Lai, T. T., Tran, T. P., Cho, J., & Yoo, M. (2023). Noise-tolerant data reconstruction based on convolutional autoencoder for wireless sensor network. Applied Sciences, 13(18), 10090. https://doi.org/10.3390/app131810090

[10] Pan, H., Sun, W., Sun, Q., & Gao, H. (2021). Deep learning based data fusion for sensor fault diagnosis and tolerance in autonomous vehicles. Chinese journal of mechanical engineering, 34(1), 72. https://doi.org/10.1186/s10033-021-00568-1

[11] Lin, X., Chao, S., Yan, D., Guo, L., Liu, Y., & Li, L. (2023). Multi-sensor data fusion method based on self-attention mechanism. Applied Sciences, 13(21), 11992. https://doi.org/10.3390/app132111992

[12] Mansouri, T., & Vadera, S. (2022). A deep explainable model for fault prediction using IoT sensors. IEEE Access, 10, 66933-66942. https://doi.org/10.1109/ACCESS.2022.3184693

[13] Wu, Y., Dai, H. N., & Tang, H. (2021). Graph neural networks for anomaly detection in industrial Internet of Things. IEEE Internet of Things Journal, 9(12), 9214-9231. https://doi.org/10.1109/JIOT.2021.3094295

[14] Xu, X., Li, X., Ming, W., & Chen, M. (2022). A novel multi-scale CNN and attention mechanism method with multi-sensor signal for remaining useful life prediction. Computers & Industrial Engineering, 169, 108204. https://doi.org/10.1016/j.cie.2022.108204

[15] Lu, H., Du, M., Qian, K., He, X., & Wang, K. (2021). GAN-based data augmentation strategy for sensor anomaly detection in industrial robots. IEEE Sensors Journal, 22(18), 17464-17474. https://doi.org/10.1109/JSEN.2021.3069452
[16] Zhang, L., Bai, W., Xie, X., Chen, L., & Dong, P. (2023). TMANomaly: Time-series mutual adversarial networks for industrial anomaly detection. IEEE Transactions on Industrial Informatics, 20(2), 2263-2271. https://doi.org/10.1109/TII.2023.3288226
[17] Liu, Z., & Hui, J. (2024). Advancing predictive maintenance: a deep learning approach to sensor and event-log data fusion. Sensor Review, 44(5), 563-574. https://doi.org/10.1108/SR-03-2024-0183
[18] Zideh, M. J., Chatterjee, P., & Srivastava, A. K. (2023). Physics-informed machine learning for data anomaly detection, classification, localization, and mitigation: A review, challenges, and path forward. IEEE Access, 12, 4597-4617. https://doi.org/10.1109/ACCESS.2023.3347989
[19] Wang, Q., Wang, K., Li, Q., Yang, Z., Jin, G., & Wang, H. (2020). MBNN: A multi-branch neural network capable of utilizing industrial sample unbalance for fast inference. IEEE Sensors Journal, 21(2), 1809-1819. https://doi.org/10.1109/JSEN.2020.3017686
[20] Xia, Y., Chen, Y., Zhao, Y., Kuang, L., Liu, X., Hu, J., & Liu, Z. (2024). FCLLM-DT: Enpowering federated continual learning with large language models for digital-twin-based industrial IoT. IEEE Internet of Things Journal, 12(6), 6070-6081. https://doi.org/10.1109/JIOT.2024.3510553
[21] Foshati, A., & Ejlali, A. (2024). Enhancing sensor fault tolerance in automotive systems with cost-effective cyber redundancy. IEEE Transactions on Intelligent Vehicles, 9(4), 4794-4803. https://doi.org/10.1109/TIV.2024.3379928
[22] Yan, W., Wang, J., Lu, S., Zhou, M., & Peng, X. (2023). A review of real-time fault diagnosis methods for industrial smart manufacturing. Processes, 11(2), 369. https://doi.org/10.3390/pr11020369
[23] Wang, W., Wang, Z., Zhou, Z., Deng, H., Zhao, W., Wang, C., & Guo, Y. (2021). Anomaly detection of industrial control systems based on transfer learning. Tsinghua Science and Technology, 26(6), 821-832. https://doi.org/10.26599/TST.2020.9010041
[24] Qiu, S., Cui, X., Ping, Z., Shan, N., Li, Z., Bao, X., & Xu, X. (2023). Deep learning techniques in intelligent fault diagnosis and prognosis for industrial systems: A review. Sensors, 23(3), 1305. https://doi.org/10.3390/s23031305
[25] Xu, P., Gao, Q., Zhang, Z., & Zhao, K. (2024). Multi-source data based anomaly detection through temporal and spatial characteristics. Expert Systems with Applications, 237, 121675. https://doi.org/10.1016/j.eswa.2023.121675
[26] Wang, X., Garg, S., Lin, H., Hu, J., Kaddoum, G., Piran, M. J., & Hossain, M. S. (2021). Toward accurate anomaly detection in industrial internet of things using hierarchical federated learning. IEEE Internet of Things Journal, 9(10), 7110-7119. https://doi.org/10.1109/JIOT.2021.3074382
[27] Shen, J., Morrison, M., Miao, H., & Gu, F. (2024, October). Harnessing deep learning for fault detection in Industry 4.0: A multimodal approach. In 2024 IEEE 6th International Conference on Cognitive Machine Intelligence (CogMI) (pp. 288-294). IEEE. https://doi.org/10.1109/CogMI62246.2024.00045
[28] Huang, P., Wang, Q., Chen, H., & Lu, G. (2023). Gas sensor array fault diagnosis based on multi-dimensional fusion, an attention mechanism, and multi-task learning. Sensors, 23(18), 7836. https://doi.org/10.3390/s23187836
[29] Neuer, M. J. (2020, October). Quantifying uncertainty in physics-informed variational autoencoders for anomaly detection. In Cybersecurity workshop by European Steel Technology Platform (pp. 28-38). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-69367-1_3
[30] Alharbi, B., Liang, Z., Aljindan, J. M., Agnia, A. K., & Zhang, X. (2022). Explainable and interpretable anomaly detection models for production data. SPE Journal, 27(01), 349-363. https://doi.org/10.2118/208586-PA
[31] Latif, S., Driss, M., Boulila, W., Huma, Z. E., Jamal, S. S., Idrees, Z., & Ahmad, J. (2021). Deep learning for the industrial internet of things (iiot): A comprehensive survey of techniques, implementation frameworks, potential applications, and future directions. Sensors, 21(22), 7518. https://doi.org/10.3390/s21227518
[32] Zhou, X., Hu, Y., Liang, W., Ma, J., & Jin, Q. (2020). Variational LSTM enhanced anomaly detection for industrial big data. IEEE Transactions on Industrial Informatics, 17(5), 3469-3477. https://doi.org/10.1109/TII.2020.3022432
[33] Sathupadi, K., Achar, S., Bhaskaran, S. V., Faruqui, N., Abdullah-Al-Wadud, M., & Uddin, J. (2024). Edge-cloud synergy for AI-enhanced sensor network data: A real-time predictive maintenance framework. Sensors, 24(24), 7918. https://doi.org/10.3390/s24247918

[34] Dehlaghi-Ghadim, A., Moghadam, M. H., Balador, A., & Hansson, H. (2023). Anomaly detection dataset for industrial control systems. IEEE Access, 11, 107982-107996. https://doi.org/10.1109/ACCESS.2023.3320928
[35] Kong, X., & Ge, Z. (2021). Deep learning of latent variable models for industrial process monitoring. IEEE Transactions on Industrial Informatics, 18(10), 6778-6788. https://doi.org/10.1109/TII.2021.3134251
[36] Yuan, L., Andrews, J., Mu, H., Vakil, A., Ewing, R., Blasch, E., & Li, J. (2022). Interpretable passive multi-modal sensor fusion for human identification and activity recognition. Sensors, 22(15), 5787. https://doi.org/10.3390/s22155787
[37] Barraza, J. F., Droguett, E. L., & Martins, M. R. (2024). FS-SCF network: Neural network interpretability based on counterfactual generation and feature selection for fault diagnosis. Expert Systems with Applications, 237, 121670. https://doi.org/10.1016/j.eswa.2023.121670