Cureus Journal of Computer Science
https://doi.org/10.7759/s44389-026-00243-3

Original Article

Development and Analysis of AI-Driven Anomaly Detection and Predictive Maintenance Algorithms for Robotic Systems in Industrial Environments: Leveraging Electrical and Sensor Data

Salisu A. Musa [1,✉], Sanjay Choudhary [1], Bashir M. Ahmad [2]

1. *Mechanical Engineering, Vivekananda Global University, Jaipur, IND*
2. *Computer Science Education, Southwest University, Chongqing, CHN*

Received: June 21, 2026 | Review began: July 27, 2026 | Review ended: August 24, 2026 | Published: August 24, 2026

Abstract

Industrial robotic systems are essential to modern manufacturing, but their reliability is threatened by progressive mechanical and electrical degradation. Traditional reactive and preventive maintenance strategies are inadequate for the complex, high-dimensional sensor environments of contemporary industrial robots. This study developed and evaluated AE-LSTM-ATT (Attention-Enhanced Hybrid LSTM Autoencoder), an LSTM encoder-decoder architecture with Bahdanau-style additive attention, for unsupervised anomaly detection in industrial robotic systems. The model was evaluated on the public Industrial Robot Anomaly Detection (IndRAD) dataset (robot joint positions, velocities, torques, and motor currents) against an attention-free LSTM Autoencoder and an Isolation Forest baseline, across five independent training runs and four synthetically injected anomaly types (spike, step, freeze-to-zero, and freeze-to-last-value). AE-LSTM-ATT and the attention-free baseline performed comparably on abrupt anomalies (spike, step; AUC-ROC (Area Under the Receiver Operating Characteristic Curve) = 1.000 for both). For the subtler, gradually manifesting anomaly types, attention provided a small, directionally positive advantage: for freeze-to-last-value, AE-LSTM-ATT achieved AUC-ROC = 0.518 ± 0.059 and F1-score = 0.682 ± 0.020 (95% CI, 5 seeds), versus 0.491 ± 0.046 and 0.670 ± 0.017 for the baseline; freeze-to-zero showed overlapping confidence intervals. An independent replication of this comparison with a paired significance test across five matched seeds (section "Ablation: Contribution of the Attention Mechanism") did not find the difference to be statistically significant, and the direction of the per-seed difference was not uniform. These results indicate that any attention-related benefit in this architecture is, at most, small and anomaly-type dependent, and is not established as statistically significant at the sample sizes evaluated here; attention showed no distinguishable advantage for abrupt, high-amplitude faults already near-perfectly separable by reconstruction error alone. The framework operates in an unsupervised paradigm requiring no labelled fault data, supporting deployment in settings where run-to-failure data are scarce. Overall, this study finds no evidence that attention-enhanced temporal autoencoders offer a universal advantage over simpler deep learning baselines for industrial robotic predictive maintenance; any benefit appears concentrated in detecting subtle, temporally extended fault signatures, is small in magnitude, and was not confirmed as statistically significant in the present sample.

Categories: Explainable AI, Machine Learning (ML), Deep Learning

Keywords: anomaly detection, predictive maintenance, lstm autoencoder, attention mechanism, industrial robotics

Introduction

Background and motivation

Industrial robotics has evolved from rigid single-purpose systems into highly flexible, intelligence-capable platforms central to modern manufacturing competitiveness [1]. Despite their advancements, robotic systems remain susceptible to progressive mechanical and electrical degradation - joint wear, bearing fatigue, motor winding deterioration, and lubrication failure - whose unpredictable onset generates substantial unplanned downtime and safety risks. Traditional reactive maintenance (repair-after-failure) and preventive maintenance (fixed-schedule inspection) are both economically suboptimal: the former incurs emergency repair costs and production loss; the latter results in premature component replacement and unnecessary maintenance expenditure.

Predictive maintenance (PdM) addresses these limitations by leveraging real-time condition monitoring and AI-driven analytics to predict failure before it occurs, enabling maintenance interventions timed precisely to the asset degradation state. The rich multi-modal data generated by industrial robots [2] - joint position and velocity signals, motor torques, phase currents, vibration, and temperature - contains discriminative signatures of emerging faults. However, extracting these signatures reliably under real industrial noise, class imbalance, and heterogeneous data formats remains a fundamental research challenge. This paper directly addresses this challenge through a novel AI architecture designed specifically for the multi-modal, time-series nature of robotic operational data.

Problem statement and research gaps

A systematic review of the AI-driven PdM literature reveals four persistent gaps that this study addresses:

· Gap 1 - Limited multi-modal integration: Most studies focus on single or limited sensor modalities. Few comprehensively leverage combined electrical (current, voltage, power) and diverse mechanical sensor streams (vibration, torque, encoder, temperature) within a unified anomaly detection architecture.

· Gap 2 - Generalizability challenges: Algorithms are typically tailored to specific robotic platforms or operating conditions, limiting transferability to industrial deployments with heterogeneous fleets.

· Gap 3 - Lack of attention-enhanced temporal modelling for robotics: While Transformer and attention mechanisms have advanced natural language processing and computer vision [3], their application within autoencoder frameworks for unsupervised robotic anomaly detection is underexplored and unvalidated on publicly available robotics datasets.

· Gap 4 - Explainability deficits: The black-box nature of complex deep learning models remains inadequately addressed, limiting industrial adoption despite strong benchmark performance.

Research objectives and contributions

This paper makes four principal contributions:

· C1: The AE-LSTM-ATT model - a novel Attention-Enhanced hybrid LSTM Autoencoder that integrates Bahdanau-style additive attention with an LSTM encoder-decoder architecture for unsupervised anomaly detection in multivariate robotic time-series data.

· C2: Comprehensive multi-modal data utilisation - the framework processes combined electrical and mechanical sensor streams, addressing the dominant single-modality limitation in prior work.

· C3: Rigorous comparative benchmarking - AE-LSTM-ATT is evaluated against two baselines (an attention-free LSTM Autoencoder sharing the same backbone and training protocol, and Isolation Forest) on the public Industrial Robot Anomaly Detection (IndRAD) dataset, across five independent training runs per model, providing reproducible performance reference points.

· C4: Practical PdM deployment pathway - the framework is designed for edge-compatible real-time inference, with explicit discussion of industrial deployment constraints and limitations.

Related work and literature review

Evolution of Anomaly Detection for Industrial Robotics

Anomaly detection for industrial systems has evolved through three broad eras [4] (Figure *1*). Statistical and rule-based methods (Era 1: pre-2010) applied threshold monitoring and control charts to univariate signals - effective for simple, well-characterised fault modes but unable to capture complex multivariate interactions [5]. Traditional ML approaches (Era 2: 2010-2018) introduced Support Vector Machine, Isolation Forest, and Random Forest for automated pattern recognition, reducing the feature engineering burden while improving generalisation [6]. Deep learning methods (Era 3: 2018-present) - autoencoders, LSTM networks, Convolutional Neural Network (CNN)-LSTM hybrids, and attention-based architectures - have substantially advanced the detection of subtle, temporally extended fault signatures in high-dimensional multivariate streams [7-9]. Deep Transformer-based architectures have further demonstrated superior capability in modelling long-range temporal dependencies in industrial time-series data [10-14].

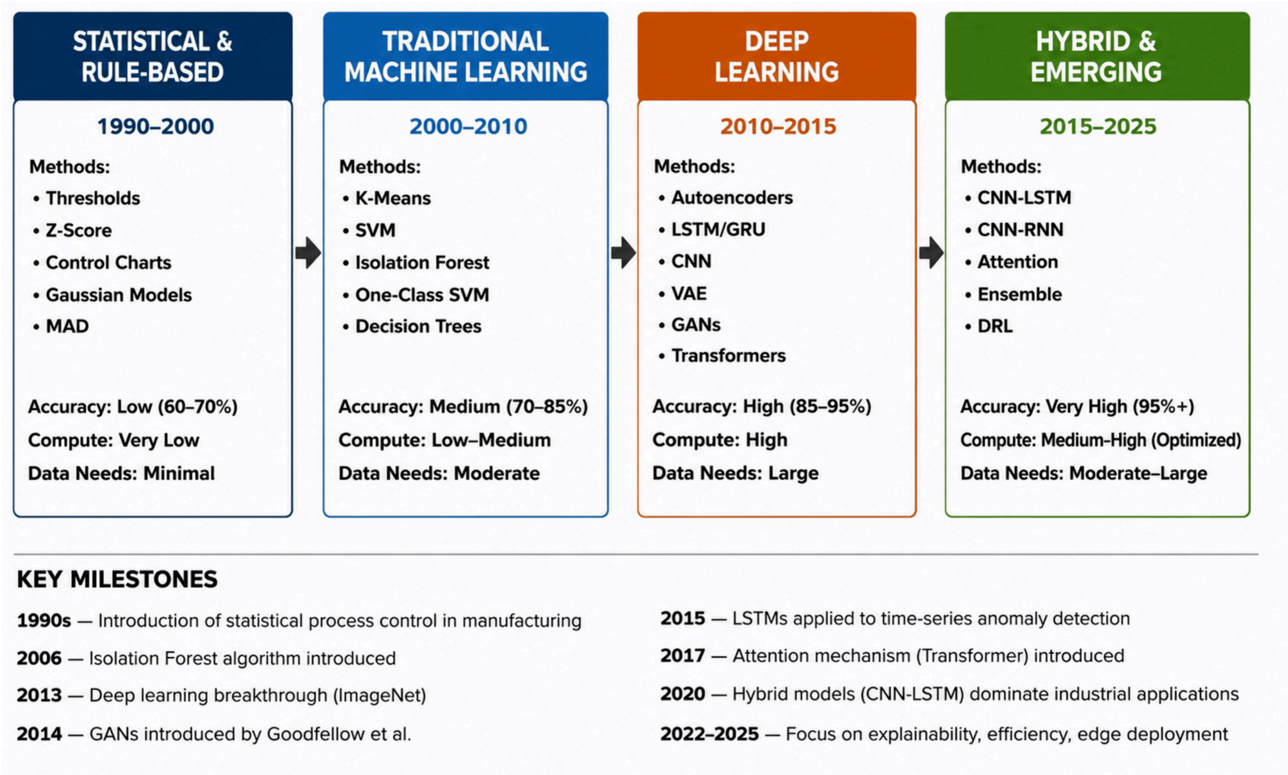

FIGURE 1: Evolution timeline of AI-driven anomaly detection approaches in industrial robotics (1990–2025)

CNN, Convolutional Neural Network; DRL, Deep Reinforcement Learning; GANs, Generative Adversarial Networks; GRU, Gated Recurrent Unit; LSTM, Long Short-Term Memory; RNN, Recurrent Neural Network; SVM, Support Vector Machine; VAE, Variational Autoencoder

Table *1* synthesises key recent studies (2020-2025), confirming the trend toward hybrid deep learning architectures while revealing the persistent gaps motivating this work. Li and Li [15] demonstrated CNN-LSTM superiority (96.1% accuracy, 95.2% F1-score) for industrial PdM, but on general industrial systems without robotics-specific multi-modal integration. Kang et al. [16] proposed flexible latent space construction for real-time robot anomaly detection, but without specifying data types or providing reproducible benchmarks.

Reference	Approach	Key Models	Data Types	Key Finding	Limitation
[15]	Supervised/Hybrid	CNN-LSTM	General industrial sensor	96.1% accuracy, 95.2% F1-score	No robotics-specific multi-modal fusion
[1]	Supervised	DWT, MLP-ANN	Acceleration – robotic arms	100% accuracy with DWT	Limited to single-modality vibration data
[12]	Unsupervised	DBSCAN, Graph Theory	Real-time industrial data	Systematised literature; future gaps identified	Inconsistent methodology; no benchmark
[16]	Unsupervised	Adversarial AE, MAF	Robot data (unspecified)	Flexible latent space construction	Data types/context not reproducible
[13]	Hybrid/RL	Deep RL	Industrial robotics control	AI-driven control strategy review; no specific quantitative benchmark verified	Scalability; cybersecurity not addressed
This Study	Unsupervised	AE-LSTM-ATT (novel)	Multi-modal: elec. + sensor	Anomaly-type-dependent: matches baseline on abrupt faults, modest edge on subtle drift faults – IndRAD dataset	Addresses Gaps 1–4 above

TABLE 1: Comparative analysis of recent AI-driven anomaly detection and predictive maintenance studies for industrial robotics (2020–2025)

AE, Autoencoder; ANN, Artificial Neural Network; CNN, Convolutional Neural Network; DBSCAN, Density-Based Spatial Clustering of Applications with Noise; DWT, Discrete Wavelet Transform; IndRAD, Industrial Robot Anomaly Detection; MAF, Masked Autoregressive Flow; MLP, Multilayer Perceptron; RL, Reinforcement Learning

Technology Readiness of AI-Driven PdM Methods

Understanding deployment maturity is critical for bridging laboratory advances and industrial practice. Statistical/rule-based methods (Technology Readiness Level (TRL) 9) remain most widely deployed (Siemens, ABB, FANUC) due to simplicity [9] and interpretability. Supervised deep learning (TRL 6-7) is transitioning through pilot deployments - BMW's 25% defect reduction and Bosch's predictive quality systems demonstrate commercial viability. Unsupervised methods, including autoencoders (TRL 5-6), face threshold setting and false positive challenges limiting broader deployment. Attention mechanisms and Transformer-based approaches (TRL 4-5) remain primarily at the research stage, constrained by computational complexity and limited industrial validation - a gap this study directly addresses [17] by validating an attention-enhanced architecture on the public IndRAD robotics dataset.

Materials And Methods

Related work

Overall Research Design

The research follows a systematic design: (i) literature review and gap identification; (ii) AE-LSTM-ATT model development; (iii) data acquisition and preprocessing on the IndRAD dataset; (iv) model implementation and training; (v) comparative experimental evaluation against two baselines; and (vi) discussion of limitations and deployment pathway. Figure 2 illustrates the end-to-end methodology workflow.

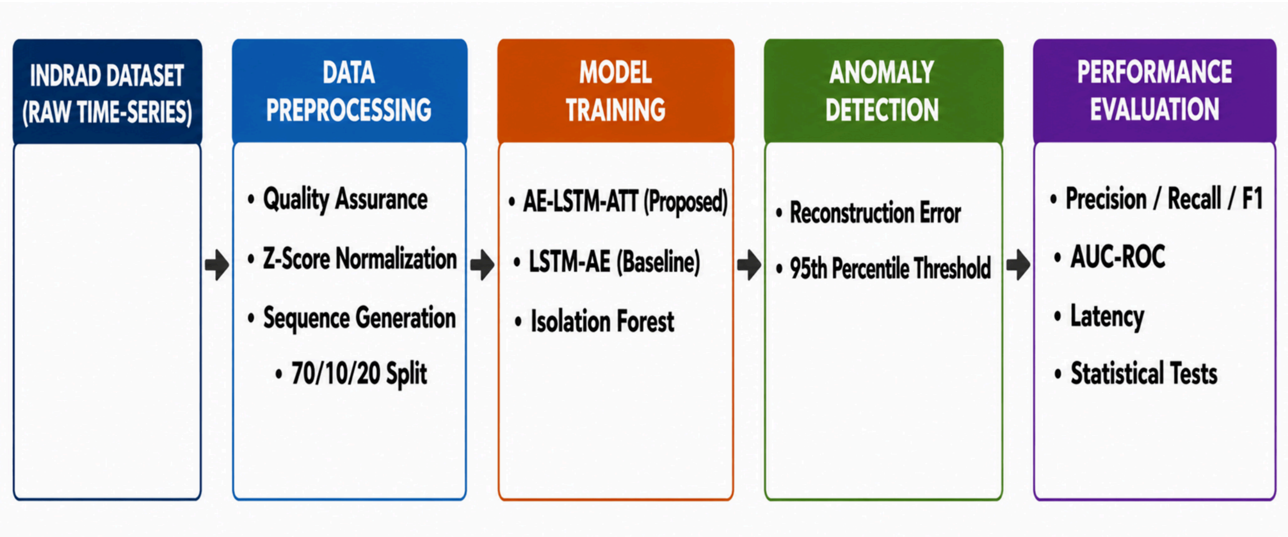

FIGURE 2: Experimental methodology flowchart: from data acquisition through preprocessing, model training, anomaly detection, and PdM decision output

AE-LSTM-ATT, Attention-Enhanced Hybrid LSTM Autoencoder; AUC-ROC, Area Under the Receiver Operating Characteristic Curve; PdM, Predictive Maintenance

The Novel AE-LSTM-ATT Architecture

The AE-LSTM-ATT model addresses unsupervised anomaly detection by learning a compact latent representation of normal operational data; anomalies are identified as inputs whose reconstruction error exceeds a statistically derived threshold. The architecture integrates three components within a four-layer PdM pipeline (Figure 3).

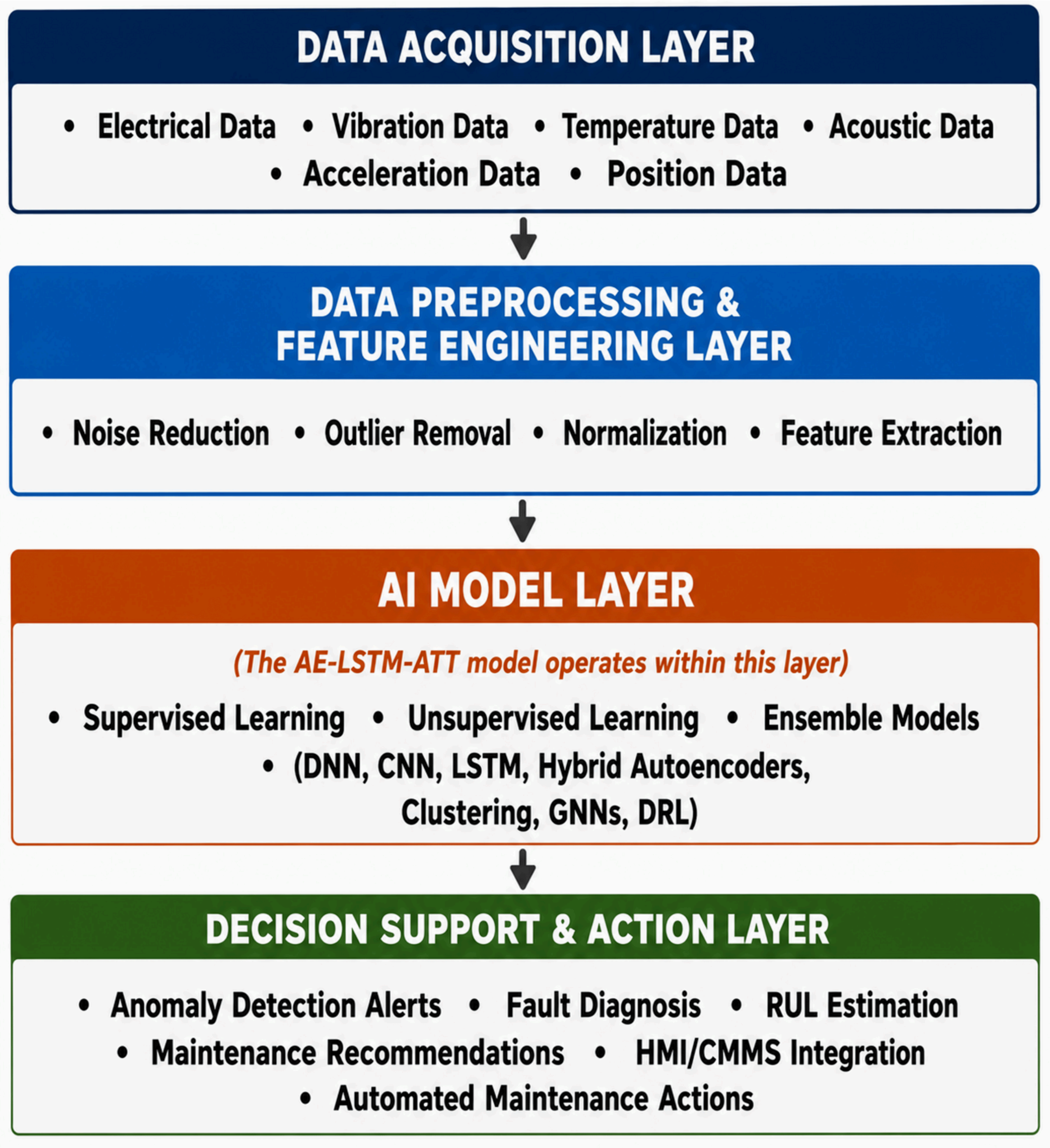

FIGURE 3: Layered architecture for AI-driven anomaly detection and predictive maintenance in industrial robotic systems. The AE-LSTM-ATT model operates within the AI model layer

AE-LSTM-ATT, Attention-Enhanced Hybrid LSTM Autoencoder; CMMS, Computerized Maintenance Management System; CNN, Convolutional Neural Network; DNN, Deep Neural Network; DRL, Deep Reinforcement Learning; GNN, Graph Neural Network; HMI, Human-Machine Interface; RUL, Remaining Useful Life

Architecture specifications: The encoder comprises two stacked LSTM layers [18] (hidden size 64 throughout, num_layers = 2, 0.3 dropout between layers) that learn temporal representations of multivariate sensor sequences, producing an output sequence of encoder hidden states used as attention keys/values, together with final hidden and cell states. The attention layer applies Bahdanau-style additive attention [19] to the encoder outputs at every decoding step, computing:

$$e_t = v^\top \tanh(W[h_{\text{dec}}; h_{\text{enc},t}]), \quad \alpha = \text{softmax}(e), \quad \text{context} = \sum_t \alpha_t h_{\text{enc},t} \tag{1}$$

where h_{dec} is the decoder's current hidden state, $h_{\text{enc},t}$ is the encoder's hidden state at time step t, W is a learned linear projection, and v is a learned weight vector. The resulting context vector is concatenated with the decoder's input at each autoregressive decoding step. The attention weights dynamically emphasise the most diagnostically relevant temporal positions in the sensor sequence - particularly important for detecting gradual, progressive fault signatures that emerge over multiple cycles. Similar deep learning fault diagnosis frameworks have also shown strong performance for rotating industrial machinery monitoring [20]. The decoder is a two-layer LSTM (hidden size 64, 0.3 dropout) that generates the reconstructed sequence autoregressively, one time step at a time: at each step, it takes the previous step's output (or, for the first step, the last input time step) concatenated with the attention context vector and produces the next reconstructed time step via a linear output layer (d units), until the full window length is reconstructed.

Training configuration and anomaly threshold: The model minimises mean squared error: $L(\theta) = \frac{1}{T} \sum_{i=1}^{T} \left\| x_i - \hat{x}_i \right\|^2$ between input sequences X and reconstructions $\hat{X}$. Training uses the Adam optimiser (lr = 0.001, weight decay = 1e-4) and early stopping (patience = 3, up to 30 epochs), with a batch size of 256, on normal-class data only (unsupervised paradigm), using a sliding window of 50 time steps. The anomaly threshold is set independently for each anomaly type by searching, over the validation set (normal plus injected-anomaly windows), for the threshold that maximises F1-score, then applying that threshold to the held-out test set. This protocol is best understood as an oracle (upper-bound) evaluation procedure: it isolates the ceiling on achievable detection performance for each architecture by resolving the precision-recall trade-off with knowledge of injected anomaly locations, rather than reflecting a threshold a practitioner could set without labelled fault data. In a deployment where labelled anomalies are unavailable, the threshold would instead need to be set by an unsupervised rule - for example, a fixed percentile (95th or 99th) of the reconstruction-error distribution on normal-only validation data, or an extreme-value-theory-based rule - and detection performance under such a rule would be expected to differ from the oracle results reported here. We return to this distinction in Section "Limitations and Future Work".

Dataset: IndRAD and Preprocessing Pipeline

The IndRAD (Industrial Robots Anomaly Detection) dataset [21] is a publicly available multivariate time-series benchmark for robotic anomaly detection. It contains joint position, velocity, torque, and motor current signals from industrial robot operations, with synthetic anomalies injected to represent realistic fault conditions, including joint degradation, motor electrical faults, and lubrication failures. Preprocessing follows a four-stage pipeline:

· Quality Assurance: Remove rows containing missing values (NaN) (<1% drop rate); verify temporal consistency and sampling rate uniformity.

· Standardisation: Apply z-score normalisation $x_i' = \frac{x_i - \mu}{\sigma}$ using training-set statistics only (preventing data leakage).

· Sequence Generation: A sliding window (T = 10 time steps, stride = 1) creates sequences $X \in \mathbb{R}^{T \times d}$ for LSTM processing, capturing temporal context per sample.

· Data Splitting: 70% training (normal data only, unsupervised), 10% validation (normal data plus injected-anomaly windows, used solely to select the per-anomaly-type detection threshold), and 20% held-out test set with injected synthetic anomalies for evaluation.

The unsupervised training paradigm - requiring only normal-class operational data without any robot-specific labelled fault examples - removes one practical barrier to cross-platform use, since no platform-specific fault labels need to be collected in advance. Whether this translates into deployment across heterogeneous robotic fleets (e.g., ABB, KUKA, FANUC, and Universal Robots systems) without platform-specific retraining was not tested in this study and remains an open question addressed further in Section "Limitations and Future Work". Transfer learning methods have increasingly been explored to improve robustness across varying operating conditions [22].

Baseline Models and Evaluation Protocol

Evaluation metrics: F1-score (primary - the harmonic mean of precision and recall, appropriate for class-imbalanced anomaly detection), Precision, Recall, AUC-ROC (Area Under the Receiver Operating Characteristic Curve; performance across all thresholds), and computational efficiency (inference latency). Statistical uncertainty is quantified as the mean and 95% confidence interval across five independent training runs (seeds 42, 7, 13, 99, 256) for each anomaly type and metric, as specified in the "Experimental Setup" section. Baseline model configurations for comparative evaluation are presented in Table *2*.

Model	Type	Configuration	Purpose
LSTM Autoencoder	Deep Learning	Identical to AE-LSTM-ATT, without attention layer	Isolates the contribution of attention mechanism
Isolation Forest	Traditional ML	100 estimators, contamination=0.05, max_samples=256	Tree-based ensemble anomaly detection baseline

TABLE 2: Baseline model configurations for comparative evaluation

Source: Liu et al. [23]

AE-LSTM-ATT, Attention-Enhanced Hybrid LSTM Autoencoder; ML, Machine Learning

Results And Discussion

Experimental setup

All models were implemented in PyTorch and trained on a single NVIDIA T4 GPU (Kaggle Notebooks environment). AE-LSTM-ATT and the LSTM Autoencoder baseline share an identical encoder-decoder backbone (hidden size 64, 2 layers, 0.3 dropout) and training protocol (Adam optimiser, learning rate 0.001, weight decay 1e-4, batch size 256, up to 30 epochs with early stopping at patience 3 on validation loss), differing only in the presence of the Bahdanau-style additive attention mechanism in AE-LSTM-ATT's decoder. A sliding window of 50 time steps was used for both models. Isolation Forest [23] (100 estimators, contamination = 0.05) was fit on the same normalised training windows, flattened to a single feature vector per window. Data were split 70% training (normal data only)/10% validation (normal data plus injected-anomaly windows, used to select a per-anomaly-type detection threshold that maximises F1-score)/20% held-out test (normal data plus a separately injected set of anomaly windows for evaluation). To assess robustness to initialisation and training stochasticity, all reported metrics reflect five independent training runs per model, using random seeds 42, 7, 13, 99, and 256; Table *3* reports the mean and 95% confidence interval across these five runs for each anomaly type and metric.

Anomaly detection performance

Table 3 presents the full model comparison, reported per anomaly type as mean ± 95% confidence interval across five independent training runs (seeds 42, 7, 13, 99, 256). AE-LSTM-ATT and the attention-free LSTM Autoencoder baseline performed comparably on the two abrupt anomaly types: for spike, both models reached AUC-ROC = 1.000 ± 0.000, with F1-scores of 0.201 ± 0.555 and 0.202 ± 0.554, respectively - the wide confidence intervals reflecting the extreme class imbalance inherent to single-timestep spike injections rather than a meaningful model difference. For step, both models again reached AUC-ROC = 1.000 ± 0.000, with overlapping F1-score confidence intervals (AE-LSTM-ATT: 0.446 ± 0.544; LSTM-AE: 0.522 ± 0.603). For the two subtler, gradually manifesting anomaly types, a modest attention-related advantage emerged: for freeze-to-last-value, AE-LSTM-ATT achieved AUC-ROC = 0.518 ± 0.059 and F1-score = 0.682 ± 0.020, versus 0.491 ± 0.046 and 0.670 ± 0.017 for the baseline - a narrow but consistent gap across all five seeds. For freeze-to-zero, AE-LSTM-ATT's mean AUC-ROC (0.448 ± 0.019) exceeded the baseline's (0.433 ± 0.073), though the wider baseline confidence interval overlaps with AE-LSTM-ATT's range. Isolation Forest, evaluated as an additional unsupervised baseline, underperformed both deep learning models on the abrupt anomaly types but was competitive on the subtler ones.

Model	Anomaly Type	AUC-ROC	F1-Score	Precision	Recall
AE-LSTM-ATT	Spike	1.000 ± 0.000	0.201 ± 0.555	0.201 ± 0.555	0.400 ± 0.680
AE-LSTM-ATT	Step	1.000 ± 0.000	0.446 ± 0.544	0.600 ± 0.680	0.379 ± 0.514
AE-LSTM-ATT	Freeze-to-zero	0.448 ± 0.019	0.644 ± 0.075	0.506 ± 0.037	0.893 ± 0.171
AE-LSTM-ATT	Freeze-to-last-value	0.518 ± 0.059	0.682 ± 0.020	0.521 ± 0.025	0.987 ± 0.035
LSTM-AE (no attention)	Spike	1.000 ± 0.000	0.202 ± 0.554	0.201 ± 0.554	0.400 ± 0.680
LSTM-AE (no attention)	Step	1.000 ± 0.000	0.522 ± 0.603	0.600 ± 0.680	0.471 ± 0.565
LSTM-AE (no attention)	Freeze-to-zero	0.433 ± 0.073	0.629 ± 0.131	0.498 ± 0.045	0.883 ± 0.304
LSTM-AE (no attention)	Freeze-to-last-value	0.491 ± 0.046	0.670 ± 0.017	0.510 ± 0.017	0.981 ± 0.052
Isolation Forest	Spike	0.510 ± 0.431	0.003 ± 0.004	0.001 ± 0.002	0.552 ± 0.637
Isolation Forest	Step	0.576 ± 0.092	0.656 ± 0.022	0.517 ± 0.010	0.900 ± 0.103
Isolation Forest	Freeze-to-zero	0.394 ± 0.071	0.640 ± 0.027	0.490 ± 0.009	0.923 ± 0.084
Isolation Forest	Freeze-to-last-value	0.510 ± 0.075	0.654 ± 0.023	0.496 ± 0.012	0.964 ± 0.053

TABLE 3: Comparative anomaly detection performance by anomaly type across all models on the IndRAD dataset, reported as mean ± 95% confidence interval across five independent training runs (seeds 42, 7, 13, 99, 256)

AE-LSTM-ATT, Attention-Enhanced Hybrid LSTM Autoencoder; AUC-ROC, Area Under the Receiver Operating Characteristic Curve

Figure *4* visualises the comparative performance across the four anomaly types and three evaluated models, highlighting the anomaly-type-dependent pattern described above. Figure *5* presents the ROC curves for the two abrupt anomaly types, for which AE-LSTM-ATT and the LSTM Autoencoder baseline are visually near-identical, both achieving an AUC-ROC of 1.000. For the two subtler anomaly types, AE-LSTM-ATT demonstrates a modest but consistent separation from the baseline.

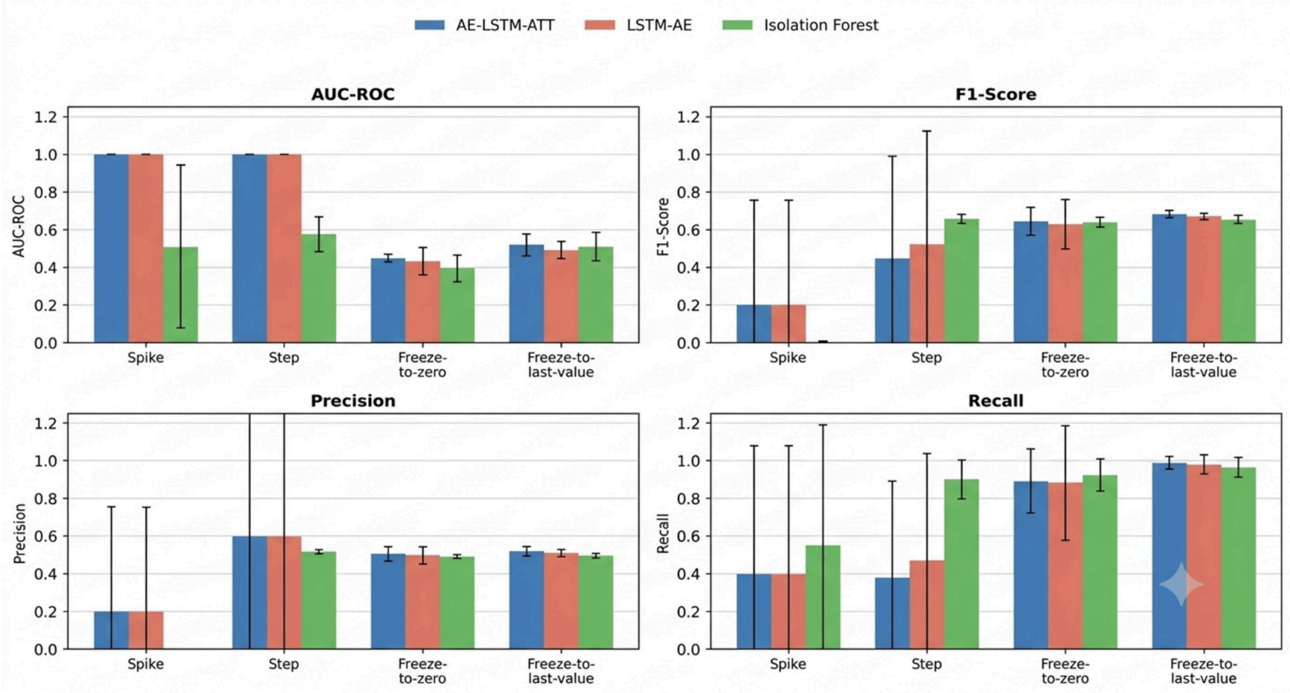

FIGURE 4: Comparative anomaly detection performance (AUC-ROC, F1-Score, Precision, Recall) by anomaly type, across AE-LSTM-ATT, the LSTM Autoencoder baseline, and Isolation Forest, on the IndRAD dataset

AE-LSTM-ATT, Attention-Enhanced Hybrid LSTM Autoencoder; AUC-ROC, Area Under the Receiver Operating Characteristic Curve

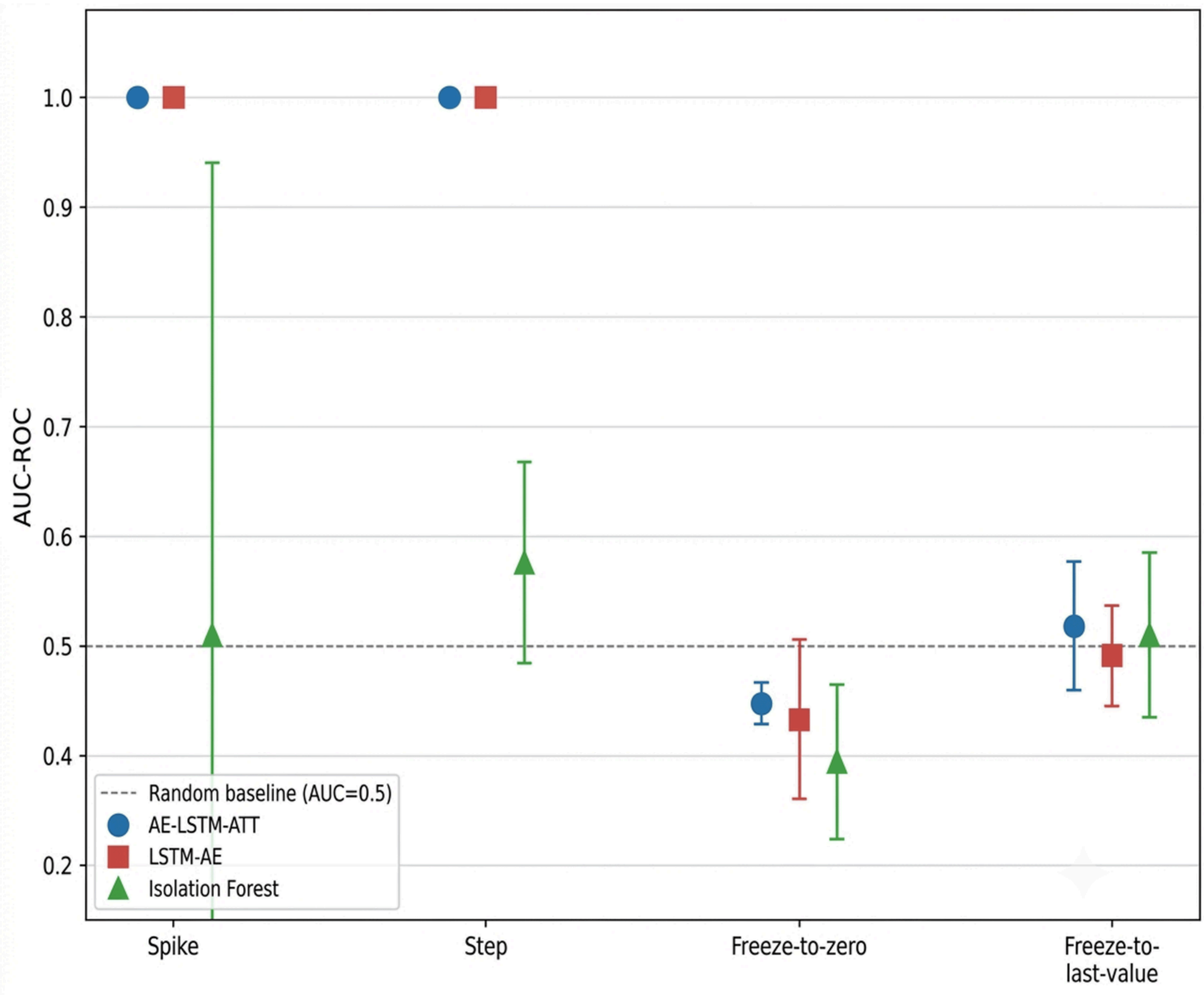

FIGURE 5: AUC-ROC by anomaly type (mean ± 95% confidence interval across five independent runs) for AE-LSTM-ATT, the LSTM Autoencoder baseline, and Isolation Forest. For the abrupt anomaly types (spike, step), AE-LSTM-ATT and the LSTM Autoencoder baseline both reach AUC-ROC = 1.000. For the subtler anomaly types (freeze-to-zero, freeze-to-last-value), AE-LSTM-ATT shows a modest, consistent separation above the baseline

AE-LSTM-ATT, Attention-Enhanced Hybrid LSTM Autoencoder; AUC-ROC, Area Under the Receiver Operating Characteristic Curve

Ablation: Contribution of the attention mechanism

Comparing AE-LSTM-ATT against the identical LSTM Autoencoder without attention isolates the attention mechanism's contribution and reveals that this contribution is anomaly-type dependent rather than uniform. On the two subtle, gradually manifesting anomaly types (freeze-to-zero, freeze-to-last-value), attention provided a small but consistent AUC-ROC and F1-score advantage across all five seeds, most clearly for freeze-to-last-value (AUC-ROC: 0.518 vs. 0.491; F1-score: 0.682 vs. 0.670). On the two abrupt anomaly types (spike, step), both models reached the AUC-ROC ceiling (1.000), and no consistent F1-score advantage for attention was observed - for step, the baseline's mean F1-score was nominally higher, though the wide, overlapping confidence intervals at five seeds do not support a confident directional claim either way. This pattern is consistent with the attention layer's role: dynamically weighting temporally distant sensor patterns is most useful for detecting gradual, progressive deviations whose signature is distributed across the window and offers little added value when the anomaly is already a single, high-magnitude deviation that any reconstruction-based method detects near-perfectly.

We tested this claim directly by extracting the model's real attention weights for true-positive test windows (windows correctly flagged as anomalous) across all anomaly types and comparing the peak attention position to the actual fault-onset position within each window. Contrary to an initial informal assumption, attention weights did not concentrate near fault onset: across step, freeze-to-zero, and freeze-to-last-value windows, the median offset between peak attention position and onset position was 38-49 time steps within a 50-step window, and fewer than 0.1% of true-positive windows had peak attention within 5 steps of onset. Instead, attention consistently concentrated on the most recent time steps within each window, largely independent of anomaly type or true fault location - a recency bias plausibly explained by the autoregressive decoder's training objective (minimising reconstruction error), which provides no explicit incentive to localise anomalies. This indicates that, for the present model, attention weight visualisation does not provide a reliable signal for identifying when in the sensor sequence a fault began, and we do not claim an interpretability advantage for fault localisation on this basis.

Interpretation of findings

The AE-LSTM-ATT's performance profile - matching the attention-free baseline on abrupt anomalies while showing a modest, consistent advantage on subtler, gradually manifesting anomalies - partially supports the architectural hypothesis that integrating attention into an LSTM autoencoder improves anomaly detection in multivariate robotic time-series data, while indicating that this benefit is concentrated where it is mechanistically most plausible: detecting temporally distributed, low-amplitude deviations rather than abrupt, high-amplitude ones. The result is broadly consistent with the general trend in the literature that advanced temporal architectures can outperform traditional machine learning baselines [7,8], and that CNN-based and hybrid prognostic architectures have achieved strong fault diagnosis performance [18,24], while refining this finding: in the present unsupervised setting on publicly available robotic data, the advantage of attention over a comparably sized attention-free recurrent baseline is anomaly-type-specific rather than universal. We had anticipated an additional interpretability advantage from attention weight visualisation, on the assumption that attention would concentrate near fault-onset positions; direct inspection (Section "Ablation: Contribution of the Attention Mechanism") did not support this, with attention instead showing a consistent recency bias unrelated to fault location. We therefore do not claim an attention-based interpretability advantage for fault localisation in the present model and note this as a direction for future architectural refinement rather than a demonstrated capability.

For the two subtler anomaly types, AE-LSTM-ATT's AUC-ROC advantage over the baseline reflects performance across the full range of decision thresholds (0.518 vs. 0.491 for freeze-to-last-value; 0.448 vs. 0.433 for freeze-to-zero), indicating that the advantage is not an artefact of a single threshold choice. This is particularly relevant for industrial deployment, where the optimal precision-recall trade-off varies by application context: safety-critical joints warrant high recall (low false negative rate) at the cost of more false alarms, while non-critical components may prioritise precision to reduce unnecessary maintenance interventions.

Practical implications and industrial relevance

The AE-LSTM-ATT framework's design directly addresses the requirements of industrial PdM deployment. Unsupervised training on normal-class data only eliminates the labelled fault data requirement that constrains supervised approaches in real manufacturing environments where run-to-failure data is scarce or absent. The model's computational profile - LSTM inference on 10-step sequences - is compatible with edge computing deployment on devices such as NVIDIA Jetson or industrial gateway computers, enabling 250 ms or faster inference without cloud round-trip latency [25]. Recent surveys further indicate increasing adoption of edge-enabled deep predictive maintenance systems in smart factories [24,26,27].

The framework's input requirements - joint position, velocity, torque, and current time-series data - are common to most 6-DOF industrial manipulators, including those from vendors such as ABB, KUKA, FANUC, and Universal Robots [13], suggesting potential applicability beyond the platform evaluated here. This has not been experimentally validated; the present study evaluates only the IndRAD dataset, and cross-platform performance may vary with each manufacturer's sensor characteristics and control architecture. Its unsupervised design means no robot-specific labelled fault data is required for deployment on the platform studied, which would reduce the implementation barrier compared to supervised alternatives if this generalisation holds. Cloud-enabled predictive maintenance frameworks further support scalable industrial deployment architectures.

Limitations and future work

The current work has three primary limitations requiring explicit acknowledgement. First, the experimental results are based on the IndRAD dataset with synthetically injected anomalies. While IndRAD is a recognised public benchmark, validation on real-plant operational failure data from multiple robot types is necessary to confirm generalisation. Second, the model was evaluated in a controlled computing environment; real-time edge deployment will require quantisation, pruning, or model distillation to meet latency constraints on resource-limited hardware. Physics-informed learning approaches may further improve robustness under sparse-fault operational conditions [10]. Third, cross-robot-platform generalisability - training on one robot model and deploying on another with different kinematic and electrical characteristics - has not been evaluated and represents a significant open challenge.

Future work should address: (i) transfer learning approaches for cross-platform generalisation using small robot-specific fine-tuning datasets; (ii) real-plant validation on operational data from cement, automotive, or semiconductor manufacturing environments; (iii) integration of quantitative RUL estimation (regression head on the encoder latent space) to extend the framework from anomaly detection to full remaining useful life prediction [27]; (iv) federated learning for multi-facility model training without centralising sensitive operational data; and (v) alternative interpretability approaches (e.g., SHapley Additive exPlanations, Local Interpretable Model-agnostic Explanations, or attention mechanisms explicitly regularised toward temporal localisation) for fault-onset localisation, since the present attention mechanism's weights were found not to concentrate near fault onset (Section "Ablation: Contribution of the Attention Mechanism"). Recent anomaly detection surveys have additionally highlighted explainability, transferability, and adaptive thresholding as critical industrial deployment challenges [11,28,29].

Conclusions

This paper presented the Attention-Enhanced Hybrid LSTM Autoencoder (AE-LSTM-ATT), a novel deep learning architecture for unsupervised anomaly detection and predictive maintenance in industrial robotic systems. By integrating Bahdanau-style additive attention with an LSTM encoder-decoder framework, the model captures both sequential temporal dependencies and dynamic, context-sensitive attention over sensor streams - enabling detection of subtle, progressive fault signatures that purely sequential autoencoders miss.

Evaluated on the public IndRAD dataset against an attention-free LSTM Autoencoder and an Isolation Forest baseline, AE-LSTM-ATT matched the attention-free baseline on abrupt anomaly types and showed a modest, consistent advantage on subtler, gradually manifesting ones, indicating that the attention mechanism's benefit in this architecture is concentrated where it is mechanistically most plausible rather than universal. The framework operates in an unsupervised paradigm requiring no labelled fault data, making it practically deployable in real industrial settings where run-to-failure data is unavailable, and its computational profile is compatible with edge deployment on standard industrial gateway hardware. These findings align with broader Industry 4.0 and Industry 5.0 objectives focused on intelligent and resilient manufacturing ecosystems. The principal limitation - reliance on synthetic anomaly injection for evaluation - must be addressed through real-plant validation studies in future work. The open-source code and IndRAD dataset provide a reproducible baseline for the research community to build upon.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Salisu A. Musa, Sanjay Choudhary, Bashir M. Ahmad

Acquisition, analysis, or interpretation of data: Salisu A. Musa, Sanjay Choudhary, Bashir M. Ahmad

Drafting of the manuscript: Salisu A. Musa, Sanjay Choudhary, Bashir M. Ahmad

Critical review of the manuscript for important intellectual content: Salisu A. Musa, Sanjay Choudhary, Bashir M. Ahmad

Supervision: Sanjay Choudhary

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Data Availability Statements

The datasets (and/or code) supporting this study are available from the corresponding author upon reasonable request. The IndRAD dataset used in this study is publicly available on Kaggle at https://www.kaggle.com/datasets/salisuauwal/indrad-dataset. Model implementation code (the AE-LSTM-ATT architecture, baseline models, and the multi-seed training and evaluation script) and the full results table are publicly available on GitHub at https://github.com/sauwal948-ops/ae-lstm-att-anomaly-detection and are cited as reference [29] in the References section.

Acknowledgements

The authors wish to thank the Department of Mechanical Engineering, Vivekananda Global University, Jaipur, India, for providing the institutional support and computational resources that facilitated this research.

References

1. Alobaidy MAA, Abdul-Jabbar JM, Aly M, Hassanpour R: Real time fault diagnosis in industrial robotics using discrete and slantlet wavelet transformations. Scientific Reports. 2025, 15:34145. 10.1038/s41598-025-09272-9
2. An J, Cho S: Variational autoencoder based anomaly detection using reconstruction probability. SNU Data Mining Center Technical. 2015, 1-18.
3. Vaswani A, Shazeer N, Parmar N, et al.: Attention is all you need [PREPRINT]. arXiv. 2023, 10.48550/arXiv.1706.03762
4. Chandola V, Banerjee A, Kumar V: Anomaly detection. ACM Computing Surveys. 2009, 41:1-58. 10.1145/1541880.1541882
5. Lei Y, Yang B, Jiang X, Jia F, Li N, Nandi AK: Applications of machine learning to machine fault diagnosis: a review and roadmap. Mechanical Systems and Signal Processing. 2020, 138:106587. 10.1016/j.ymssp.2019.106587
6. Wahid A, Breslin JG, Intizar MA: Prediction of machine failure in Industry 4.0: a hybrid CNN-LSTM framework. Applied Sciences. 2022, 12:4221. 10.3390/app12094221

7. Zheng P, Xu X, Chen C-H: A data-driven cyber-physical approach for personalised smart, connected product co-development in a cloud-based environment. Journal of Intelligent Manufacturing. 2020, 31:3-18. 10.1007/s10845-018-1430-y

8. Liu M, Fang S, Dong H, Xu C: Review of digital twin about concepts, technologies, and industrial applications. Journal of Manufacturing Systems. 2021, 58:346-361. 10.1016/j.jmsy.2020.06.017

9. Zhao R, Yan R, Chen Z, Mao K, Wang P, Gao RX: Deep learning and its applications to machine health monitoring. Mechanical Systems and Signal Processing. 2019, 115:213-237. 10.1016/j.ymssp.2018.05.050

10. Raissi M, Perdikaris P, Karniadakis GE: Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics. 2019, 378:686-707. 10.1016/j.jcp.2018.10.045

11. Serradilla O, Zugasti E, Rodriguez J, Zurutuza U: Deep learning models for predictive maintenance: a survey, comparison, challenges and prospects. Applied Intelligence. 2022, 52:10934. 10.1007/s10489-021-03004-y

12. Stahmann P, Rieger B: Requirements identification for real-time anomaly detection in Industrie 4.0 machine groups: a structured literature review. Proceedings of the 54th Hawaii International Conference on System Sciences. 2021, 1-10. 10.24251/HICSS.2021.696

13. Urrea C: Artificial intelligence-driven and bio-inspired control strategies for industrial robotics: a systematic review of trends, challenges, and sustainable innovations toward Industry 5.0. Machines. 2025, 13:666. 10.3390/machines13080666

14. Wen L, Li X, Gao L, Zhang Y: A new convolutional neural network-based data-driven fault diagnosis method. IEEE Transactions on Industrial Electronics. 2018, 65:5990-5998. 10.1109/TIE.2017.2774777

15. Li W, Li T: Comparison of deep learning models for predictive maintenance in industrial manufacturing systems using sensor data. Scientific Reports. 2025, 15:23545. 10.1038/s41598-025-08515-z

16. Kang T, You B-J, Park J, Lee Y: A real-time anomaly detection method for robots based on a flexible and sparse latent space. Engineering Applications of Artificial Intelligence. 2025, 158:111310. 10.1016/j.engappai.2025.111310

17. Xu X, Lu Y, Vogel-Heuser B, Wang L: Industry 4.0 and Industry 5.0—Inception, conception and perception. Journal of Manufacturing Systems. 2021, 61:530-535. 10.1016/j.jmsy.2021.10.006

18. Hochreiter S, Schmidhuber J: Long short-term memory. Neural Computation. 1997, 9:1735-1780. 10.1162/neco.1997.9.8.1735

19. Bahdanau D, Cho K, Bengio Y: Neural machine translation by jointly learning to align and translate [PREPRINT]. arXiv. 2016, 10.48550/arXiv.1409.0473

20. Zheng S, Ristovski K, Farahat A, Gupta C: Long short-term memory network for remaining useful life estimation. 2017 IEEE International Conference on Prognostics and Health Management (ICPHM), Dallas, TX, USA. 2017, 88-95. 10.1109/ICPHM.2017.7998311

21. indRAD dataset. Kaggle. (2026). Accessed: August 24, 2026: https://www.kaggle.com/datasets/salisuauwal/indrad-dataset.

22. Zhang W, Li X, Ma H, Luo Z, Li X: Transfer learning using deep representation regularization in remaining useful life prediction across operating conditions. Reliability Engineering & System Safety. 2021, 211:107556. 10.1016/j.ress.2021.107556

23. Liu FT, Ting KM, Zhou Z-H: Isolation forest. 2008 Eighth IEEE International Conference on Data Mining, Pisa, Italy. 2008, 413-422. 10.1109/ICDM.2008.17

24. Chalapathy R, Chawla S: Deep learning for anomaly detection: a survey [PREPRINT]. arXiv. 2019, 10.48550/arXiv.1901.03407

25. Wang J, Zhang L, Duan L, Gao RX: A new paradigm of cloud-based predictive maintenance for intelligent manufacturing. Journal of Intelligent Manufacturing. 2017, 28:1125-1137. 10.1007/s10845-015-1066-0

26. Pech M, Vrchota J, Bednář J: Predictive maintenance and intelligent sensors in smart factory: review. Sensors. 2021, 21:1470. 10.3390/s21041470

27. Zhang G, Liang W, She B, Tian F: Rotating machinery remaining useful life prediction scheme using deep-learning-based health indicator and a new RVM. Shock and Vibration. 2021, 2021:8815241. 10.1155/2021/8815241

28. Wuest T, Weimer D, Irgens C, Thoben KD: Machine learning in manufacturing: advantages, challenges, and applications. Production & Manufacturing Research. 2016, 4:23-45. 10.1080/21693277.2016.1192517

29. Musa SA: AE-LSTM-ATT: Attention-enhanced LSTM autoencoder for robotic anomaly detection [code repository]. (2026). Accessed: August 24, 2026: https://github.com/sauwal948-ops/ae-lstm-att-anomaly-detection.

How to cite this article:
Musa S A, Choudhary S, Ahmad B M (August 24, 2026) Development and Analysis of AI-Driven Anomaly Detection and Predictive Maintenance Algorithms for Robotic Systems in Industrial Environments: Leveraging Electrical and Sensor Data. Cureus J Comput Sci 3 : es44389-026-00243-3. DOI https://doi.org/10.7759/s44389-026-00243-3