Blockchain Frontier Technology (B-Front)
Vol. 6, No. 1, July 2026, pp. 25∼37
E-ISSN: 2808-0009 | P-ISSN: 2808-0831, DOI: 10.34306

Self Supervised Transformers for High Dimensional Time Series Anomaly Detection

Aswadi Jaya[1], Derlina[2], Qurotul Aini[3], Agung Rizky[4*], Richard Evans[5]
[1]Department of English Education, PGRI University Palembang, Indonesia
[2]Faculty of Mathematics and Natural Science, State University of Medan, Indonesia
[3]Faculty of Information Technology, Satya Wacana Christian University, Indonesia
[4]Faculty of Science and Technology, University of Raharja, Indonesia
[5]Department of Computer Science, Adi Journal Incorporation, USA
[1]aswadijaya@univpgri-palembang.ac.id, [2]derlina@unimed.ac.id, [3]aini@raharja.info, [4]agungrizky@raharja.info,
[5]vans.richard@adi-journal.org
***Corresponding Author**

Article Info

Article history:

Submission, 05-02-2026
Revised, 25-02-2026
Accepted, 12-03-2026
Published, 26-03-2026

Keywords:

Self Supervised Learning
Transformer Models
Anomaly Detection
Deep Learning
Software Analytics

ABSTRACT

This study addresses anomaly detection in high dimensional time series data within the context of Artificial Intelligence (AI) driven software development, where modern systems generate large temporal data streams and reliable monitoring remains difficult due to noise, complexity, and limited labeled anomalies. **The objective of this research** is to develop an effective and scalable anomaly detection framework based on self supervised transformer models that can learn meaningful temporal representations without heavy reliance on manual annotation. **The proposed method applies** self supervised pretraining through masked sequence reconstruction and contrastive temporal learning on large scale unlabeled multivariate time series datasets, followed by transformer based attention mechanisms to capture long range dependencies and compute anomaly scores. Experiments are conducted using benchmark datasets and real world system log data implemented with Python based deep learning tools and transformer architectures to evaluate detection performance. **The results indicate** that the proposed approach improves detection accuracy and reduces false positive rates compared to traditional statistical techniques and supervised deep learning models, particularly in high dimensional and low label settings. **In conclusion**, integrating self supervised learning with transformer architectures provides a robust and generalizable solution for time series anomaly detection, contributing to software analytics and monitoring systems by lowering labeling costs and improving adaptability across application domains.

This is an open access article under the CC BY 4.0 license.

DOI: https://doi.org/10.34306/bfront.v6i1.1078
This is an open-access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/)
©Authors retain all copyrights

1. INTRODUCTION

The rapid growth of digital business models has increased dependence on complex software systems. These systems support real time decision making, service delivery, and operational efficiency across industries. Modern digital platforms such as cloud services, financial technology, manufacturing systems and smart infrastructure continuously generate high dimensional time series data originating from sensors logs user interactions and transactional processes [1, 2]. These data streams are critical assets for ensuring system reliability performance optimization and risk mitigation. At the same time automation and artificial intelligence have become central to software engineering practices, enabling predictive monitoring, adaptive control, and autonomous system management. Artificial Intelligence (AI) driven solutions are now embedded throughout the software lifecycle, ranging from development and testing to deployment and maintenance.

In real time monitoring environments, deploying anomaly detection models requires careful integration with streaming data pipelines and operational infrastructures, where the proposed self supervised transformer can function as an inference module processing time series data in sliding windows or micro batches for continuous anomaly scoring without disrupting system operations [3, 4]. However, practical implementation introduces challenges such as high computational demands, latency constraints requiring optimization strategies, and concept drift that necessitates periodic adaptation or retraining. Deployment feasibility also depends on infrastructure compatibility, data throughput, and integration with alert management systems, requiring a balance between detection accuracy, computational efficiency, and responsiveness. More broadly, advances in AI driven development tools demonstrate how machine intelligence is transforming software engineering beyond code generation toward strategic operational monitoring, where anomaly detection in high dimensional time series data becomes essential for system reliability and performance [5, 6]. This capability supports sustainable digital infrastructure and aligns with the objectives of the United Nations Sustainable Development Goals (SDGs), particularly those related to innovation, infrastructure resilience, and efficient resource utilization.

Anomaly detection is increasingly critical in blockchain and distributed ledger systems, where decentralized nodes generate continuous temporal data from transactions, consensus, network synchronization, and smart contracts. Monitoring requires adaptive mechanisms to detect abnormal transactions, consensus issues, node failures, and security threats without extensive labeled data [7, 8]. Self-supervised transformer frameworks suit this task by learning normal behavior from large scale unlabeled activity while capturing long range temporal dependencies and cross component interactions across heterogeneous, asynchronous, and evolving network states, enabling reliable oversight of validation, block propagation, and resource use, and improving reliability, security, and transparency to mitigate costly downtime [9, 10]. In software engineering, traditional rule based or statistical methods struggle with high dimensional noisy time series, while supervised approaches require costly labeled anomalies. AI assisted tools improve productivity but not robustness, and transformer architectures successful in NLP and code understanding remain underexplored for self-supervised anomaly detection, offering potential to enhance detection accuracy, system reliability, and support SDGs 8 and SDGs 11 [11, 12].

From an industry perspective, the novelty of the proposed framework lies in its ability to deliver scalable anomaly detection without dependence on labeled anomaly data, making it particularly suitable for software intensive environments within software analytics contexts characterized by continuous data generation and rapidly evolving system behavior [13, 14]. In cloud computing infrastructures, where distributed services generate high volume telemetry streams, self supervised transformer models enable adaptive monitoring that can detect performance degradation, abnormal resource consumption, or emerging system faults before they propagate across services. Based on these challenges, this research defines clear objectives to advance theory and practice. The first objective is to examine how self-supervised transformer models improve anomaly detection in high-dimensional time series data by capturing long-range temporal dependencies and complex feature interactions often missed by traditional models [15, 16]. The second objective is to evaluate how AI-based automation through self-supervised learning enhances monitoring efficiency and scalability by reducing reliance on labeled data and manual configuration. The third objective is to analyze potential risks and limitations, including model interpretability, computational cost, and sensitivity to data drift. These objectives provide insights into AI's impact on software reliability and operational efficiency while supporting the development of resilient and trustworthy AI-driven software systems [17, 18].

2. LITERATURE REVIEW

2.1. Development of Anomaly Detection in Time Series Data

Recent years have seen rapid growth in anomaly detection research for time series data, driven by increasing software complexity and expanding digital infrastructures. Early methods relied on statistical and rule based techniques, such as control charts and threshold monitoring, which worked in stable environments but struggle with large scale, high dimensional, and evolving data streams. As software architectures become more distributed and data intensive, adaptive and intelligent monitoring has become essential. Research now spans domains like blockchain and distributed computing, where continuous multivariate temporal data from transactions, consensus processes, and node interactions create new challenges [19, 20]. These environments demand detection methods that learn dynamic patterns without heavy supervision or extensive labeled datasets. Advances in AI driven blockchain security and distributed machine learning have enabled sophisticated anomaly detection techniques that identify abnormal transactions, consensus irregularities, and smart contract anomalies, while supporting secure, privacy preserving analytics in decentralized systems.

2.2. Transformer Models for Time Series Anomaly Detection

Transformer models, originally for natural language processing, are now effective for sequential data thanks to attention mechanisms and parallel processing [21, 22]. They capture long term dependencies and complex interactions in tasks like time series forecasting and system log analysis. Many transformer based anomaly detection methods, however, rely on supervised or weakly supervised learning, limiting their use with scarce labeled data [23, 24]. Comparing them with traditional statistical approaches, reconstruction based models, generative sequence models, and self-supervised representation learning highlights their strengths and limitations. Table 1 summarizes these methods for high dimensional time series.

Table 1. Critical Comparison of Anomaly Detection Methods

Method Category	**Example Models**	**Strengths**	**Limitations**	**Suitability for High Dimensional Time Series**
Statistical Methods	Control Chart, ARIMA, Threshold Monitoring	Simple implementation, interpretable results	Poor scalability, sensitive to noise, limited ability to capture complex temporal patterns	Low
Reconstruction Based Models	Autoencoder, Variational Autoencoder (VAE)	Effective for unsupervised representation learning, detects reconstruction errors	Difficulty capturing long term dependencies, sensitive to training data distribution	Medium
Generative Sequence Models	GAN based models, LSTM based generative models	Capable of modeling complex data distributions	Training instability, high computational cost	Medium
Contrastive or Self Supervised Models	CPC, SimCLR based temporal models	Learn robust representations without labels	Often require large datasets and careful design of pretext tasks	Medium–High
Transformer Based Models	Transformer, Informer, LogTransformer	Capture long range dependencies and multivariate interactions	Large computational cost and often rely on labeled data	High
Self Supervised Transformer Models	Masked Transformer, Pretraining based architectures	Strong representation learning, scalable to large datasets, reduced reliance on labeled anomalies	Still limited research on deployment, data drift handling, and efficiency	Very High

Table 1 summarizes anomaly detection methods for high-dimensional time series. Statistical methods are simple but scale poorly, reconstruction and generative models capture structure but may miss long term dependencies, contrastive and self-supervised models need large datasets, transformers handle long range dependencies, and self-supervised transformers offer scalable, effective learning with minimal labeled data, though deployment challenges remain.

2.3. Self Supervised Representation Learning for High Dimensional Time Series

Self-supervised learning extracts meaningful features from unlabeled data via masked prediction and contrastive learning, improving performance and robustness in vision and language tasks [25, 26]. Its integration with transformers for high dimensional time series anomaly detection remains limited, as most studies ignore deployment, data drift, and computational trade offs [27, 28]. Self-supervised transformers improve representation quality, generalization, and data efficiency, but cross domain benefits remain underexplored, and few comparisons exist with other unsupervised methods. The growing need for intelligent monitoring to reduce downtime, optimize resources, and ensure service continuity SDGs 9 and 12 highlights a research gap in practical self-supervised transformer frameworks.

3. RESEARCH METHOD

This study uses a structured and systematic methodology to evaluate self-supervised transformer models for anomaly detection in high dimensional time series data generated by modern software systems within software analytics contexts [29, 30]. The proposed methodological framework is intended to support a comprehensive investigation of model performance while maintaining clarity and consistency throughout the research process. In particular, the methodology emphasizes the ability of self-supervised learning approaches to capture complex temporal patterns without relying heavily on manually labeled anomalies, which are often scarce in real-world monitoring environments. The methodology is designed to ensure empirical robustness, reproducibility, and practical relevance by integrating quantitative performance evaluation with qualitative validation [31, 32]. Through this combined evaluation strategy, the study aims to provide a balanced assessment of the effectiveness of transformer-based models in identifying anomalous patterns in complex temporal data streams. The overall workflow of the research process is illustrated in Figure 1, which outlines the sequential stages from data acquisition and preprocessing to model training, evaluation, and final result verification.

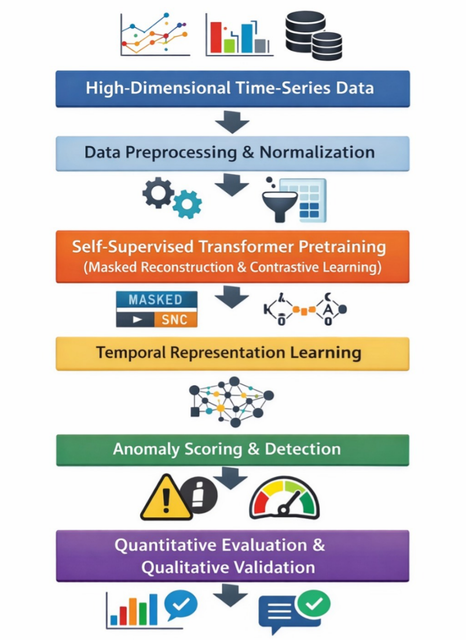

Figure 1. Self-Supervised Transformer Framework for Time-Series Anomaly Detection

Figure 1 clarifies how each data source contributes to the methodological objectives of the study within self supervised learning contexts. By combining operational system data with human centered feedback and benchmark datasets, the research ensures a balanced evaluation that reflects both technical performance and practical relevance across different software contexts [33, 34]. This integrative approach strengthens the validity of the analysis by capturing multiple perspectives, enabling more comprehensive assessment of model effectiveness in both controlled experimental settings and real world operational environments.

3.1. Research Approach

The study uses a mixed methods approach to evaluate the proposed self supervised anomaly detection framework. Quantitative analysis measures performance via precision, recall, F1 score, false positive rate, and detection latency, while qualitative analysis assesses reliability, interpretability, and operational usability. This combination provides insights into both detection accuracy and practical deployment in real world software environments [35, 36]. For reproducibility, model training uses predefined hyperparameters epochs, learning rate, batch size, and transformer settings which influence learning dynamics and detection capability. Training across multiple epochs allows stable temporal representations, with the learning rate controlling optimization steps. The hyperparameter configuration is summarized in Table 2.

Table 2. Experimental Hyperparameter Configuration

Parameter	Value	Description
Training Epochs	50	Number of full training iterations over the dataset
Learning Rate	0.001	Step size used by the optimizer during training
Batch Size	64	Number of samples processed in one training step
Optimizer	Adam Optimizer	Optimization algorithm used for parameter updates
Transformer Layers	4	Number of encoder layers in the transformer model
Attention Heads	8	Number of parallel attention mechanisms
Embedding Dimension	128	Size of latent feature representation
Dropout Rate	0.1	Regularization to prevent overfitting

The Table 2 selected hyperparameter configuration is designed to balance model performance, training stability, and computational efficiency during the anomaly detection experiments. These parameters are carefully chosen to ensure that the model can learn meaningful temporal representations from high dimensional time series data while maintaining stable convergence throughout the training process. In addition, the configuration aims to prevent overfitting and excessive computational cost, thereby enabling the framework to operate effectively in practical monitoring environments.

3.1.1. Data Collection

Data collection is conducted through multiple complementary sources to ensure diversity, representativeness, and robustness within software analytics and self supervised learning contexts. High dimensional multivariate time series data are obtained from software system logs, performance monitoring metrics, and telemetry signals generated by AI assisted and cloud based software systems [37, 38]. These data streams capture complex temporal dynamics and inter variable dependencies essential for anomaly detection tasks. In parallel, qualitative data are collected through structured surveys administered to software developers and system engineers who interact with AI based monitoring tools. To support reproducibility and generalization, publicly available benchmark datasets and open source repositories are also incorporated. The characteristics of the collected data sources are summarized in Table 3.

Table 3. Data Collection Sources and Characteristics

Data Source Type	Data Characteristics	Purpose in Analysis
System Logs	High frequency event sequences	Capture operational anomalies
Performance Metrics	Multivariate numerical time series	Monitor system health
Telemetry Signals	Continuous sensor based data	Detect temporal deviations
Developer Feedback	Qualitative survey responses	Validate practical relevance
Open Benchmark Data	Public time series datasets	Ensure reproducibility

Table 3 clarifies how each data source contributes to the methodological objectives of the study. By combining operational system data with human centered feedback and benchmark datasets, the research ensures a balanced evaluation that reflects both technical performance and practical relevance across different software contexts [33, 34]. This integrative approach strengthens the validity of the analysis by capturing multiple perspectives, enabling more comprehensive assessment of model effectiveness in both controlled experimental settings and real world operational environments. Additionally, the use of diverse data sources enhances the generalizability of the findings by ensuring that the evaluation reflects varied system conditions and usage scenarios.

3.1.2. Analytical Techniques

A range of analytical techniques is applied to address the research objectives in a systematic manner. Quantitative analysis includes evaluation of anomaly detection performance using standardized metrics and regression analysis to examine the relationship between data dimensionality, model complexity, and detection accuracy. Additional indicators, such as reductions in manual monitoring effort and improvements in anomaly response time, are considered where applicable [39, 40]. Qualitative analysis is conducted using thematic analysis of survey responses to identify recurring patterns related to trust, interpretability, and adoption challenges of AI assisted monitoring systems. From a machine learning perspective, self supervised transformer models are trained using masked sequence reconstruction and contrastive temporal learning objectives to capture long range temporal dependencies and complex feature interactions. The conceptual architecture of the proposed model is presented in Figure 2.

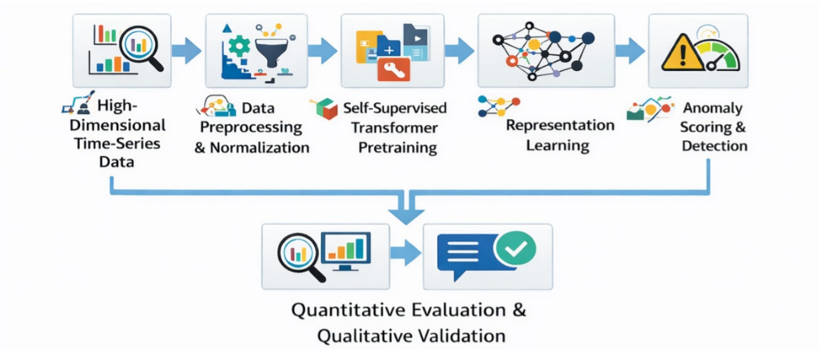

Figure 2. Self Supervised Transformer Architecture for Anomaly Detection

Figure 2 illustrates how the transformer architecture processes multivariate time series inputs through attention mechanisms, while self supervised learning enables the model to identify normal and anomalous patterns without labeled data, improving scalability and robustness in high dimensional environments. Table 4 summarizes the analytical techniques and evaluation goals, clarifying the relationship between model design, procedures, and criteria. Integrating architectural explanation with methodological summary enhances transparency, allowing readers to trace how theoretical concepts are implemented and empirically evaluated [41, 42].

Table 4. Analytical Techniques and Evaluation Metrics

Analysis Type	Technique	Evaluation Focus
Quantitative	Precision, Recall, F1-score	Detection accuracy
	False Positive Rate	Reliability of detection
	Detection Latency	Real-time applicability
Qualitative	Thematic Analysis	Interpretability and trust
Machine Learning	Self-Supervised Learning	Representation robustness

Table 4 highlights the alignment between analytical techniques and research objectives, demonstrating how quantitative metrics, qualitative insights, and machine learning methods jointly support a comprehensive evaluation of the proposed framework. This alignment ensures that each analytical component directly contributes to addressing the study's key research questions in a structured and systematic manner [43, 44]. Furthermore, it strengthens the overall coherence of the evaluation process by integrating multiple forms of evidence to provide a more holistic understanding of the framework's performance and practical relevance. In addition, this structured mapping enhances methodological transparency by clearly showing how evaluation criteria correspond to specific analytical procedures, thereby improving the rigor and interpretability of the research design.

3.2. Validity and Reliability

To ensure validity and reliability, multiple control measures are applied throughout the research process. Data preprocessing is carefully reviewed to minimize noise and bias, while Transformer Models are trained across multiple runs to ensure stability and consistency. Selected anomaly detection results are manually inspected to validate model behavior against domain expectations, and qualitative findings are supported by inter rater reliability checks [45, 46]. Cross validation and result verification confirm reproducibility, as illustrated in Figure 3. These measures establish a structured framework that ensures methodological rigor, strengthens confidence in model stability across datasets, and maintains transparency in the research methodology. Together, these procedures provide a comprehensive quality assurance mechanism that supports both analytical accuracy and interpretive credibility.

Figure 3. Validation and Reliability Control Process

Figure 3 emphasizes the layered validation strategy adopted in this study, ensuring that both quantitative results and qualitative interpretations are reliable and reproducible. This systematic approach strengthens the credibility of the findings and supports the applicability of self-supervised transformer models for anomaly detection in high dimensional time series data. Furthermore, the structured validation process enhances confidence in the consistency of the model's performance across different datasets and evaluation scenarios, reinforcing the robustness of the overall research framework. In addition, this comprehensive validation design helps minimize potential bias and uncertainty, ensuring that the reported results accurately reflect the model's true analytical and operational capabilities [47, 48].

4. RESULTS AND DISCUSSION

This section presents the experimental results and analytical discussion of the proposed self-supervised transformer framework for anomaly detection in high dimensional time series data. The evaluation emphasizes comparative detection performance and scalability under increasing data dimensionality, supported by quantitative metrics and interpretative discussion. The results demonstrate the effectiveness of the proposed approach in addressing key challenges associated with unlabeled and high dimensional monitoring data in software intensive systems. In addition, the findings provide empirical evidence of the model's robustness across varying data conditions, offering deeper insight into its practical suitability for complex and large scale monitoring environments [49, 50]. Moreover, the analysis highlights how the framework maintains stable performance as system complexity increases, reinforcing its potential for deployment in dynamic and continuously evolving operational settings. Furthermore, the integration of self supervised learning plays a central role in enabling the model to learn meaningful temporal representations from unlabeled data, thereby enhancing adaptability and reducing dependence on manual annotation. This characteristic further supports the framework's applicability in real world environments where data complexity, scale, and variability continue to grow over time.

To strengthen the reliability of the experimental findings, statistical validation is also incorporated through significance testing and confidence estimation. Specifically, statistical tests such as the *t-test* are applied to examine whether the observed performance differences between the proposed framework and baseline methods are statistically significant. In addition, confidence intervals are calculated for the main evaluation metrics to quantify the uncertainty and stability of the results across experimental runs. These statistical analyses provide additional empirical support for the reported performance improvements and ensure that the conclusions drawn from the experiments are not solely based on point estimates but are supported by rigorous statistical evidence.

4.1. Comparative Anomaly Detection Performance

The first set of experiments evaluates the anomaly detection performance of the proposed model against representative baseline methods, including statistical thresholding, supervised LSTM models, and unsupervised autoencoder based approaches. Performance is assessed using the F1 score, which balances precision and recall and is widely adopted in anomaly detection research. The comparative results are visualized in Figure 4. This comparison provides a clear quantitative basis for assessing the relative effectiveness of each method and highlights the performance differences across varying modeling approaches. Furthermore, the evaluation framework ensures a consistent experimental setting, allowing for a fair and systematic comparison of detection capabilities across all tested models.

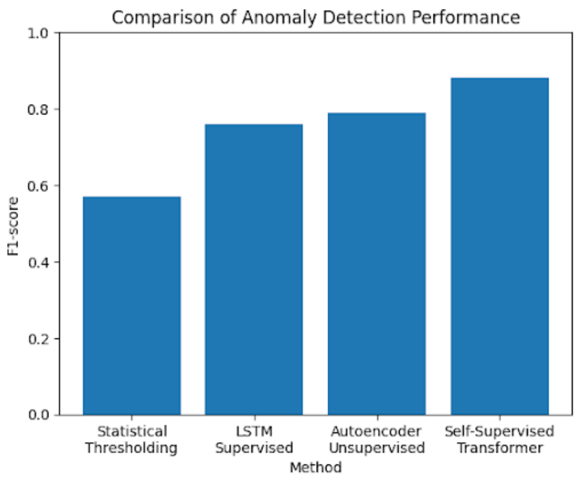

Figure 4. Comparison of Anomaly Detection Performance (F1-score)

Figure 4 shows the comparison of anomaly detection performance measured by F1 score across statistical thresholding, supervised LSTM, unsupervised autoencoder, and the proposed self supervised transformer model, highlighting the relative improvement in detection accuracy achieved by attention based temporal modeling and representation learning without labeled anomaly data.

4.2. Scalability under Increasing Data Dimensionality

To assess robustness and scalability, additional experiments are conducted by increasing the number of monitored variables from low dimensional to highly multivariate settings. This evaluation is critical for modern software systems, where hundreds of metrics are continuously collected. The scalability results are illustrated in Figure 5

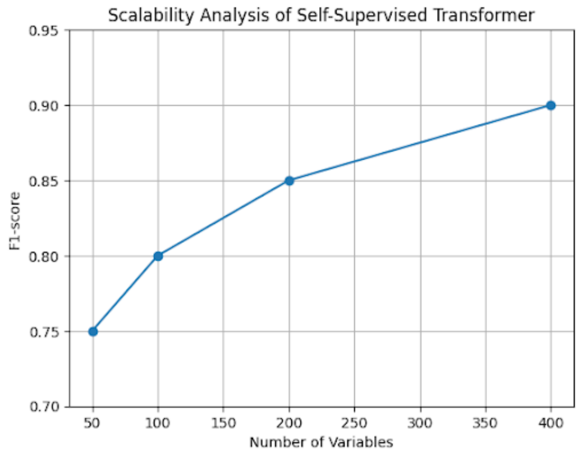

Figure 5. Scalability Analysis of the Self-Supervised Transformer Model

Figure 5 scalability analysis of the self-supervised transformer model under increasing data dimensionality. The figure illustrates how detection performance changes as the number of monitored variables grows, for example from 50, 100, to 200 features, demonstrating the model's ability to utilize additional contextual information while maintaining stable and improved anomaly detection accuracy in high dimensional environments.

4.3. Discussion and Practical Implications

The combined results from Figures 4 and 5 validate the proposed self supervised transformer as a robust and scalable solution for anomaly detection in high dimensional time series data. The model shows superior performance compared to both supervised and unsupervised baselines, while the scalability analysis confirms its adaptability to increasingly complex data environments. These findings are relevant for real world software monitoring where labeled anomalies are limited and system complexity continues to increase. Higher F1 scores indicate lower false detection rates, reducing alert fatigue and increasing trust in AI driven monitoring systems. This capability enables earlier detection of performance degradation, abnormal resource utilization, and system faults without relying on manually defined thresholds or labeled incident data, allowing organizations to shift from reactive incident response toward predictive maintenance.

From a policy perspective, adaptive AI driven monitoring systems also influence governance and infrastructure management standards. Cloud providers and infrastructure regulators may increasingly require intelligent monitoring mechanisms capable of detecting operational risks in real time, particularly in critical service environments. Continuous autonomous monitoring supports risk management frameworks, service level assurance policies, and resilience planning guidelines. Reduced dependence on labeled anomaly data also lowers implementation barriers, enabling broader adoption across organizations with different technical capacities. These advantages support more resilient monitoring practices and evidence based policy frameworks for managing large scale digital infrastructure, while also aligning with SDGs 9 and SDGs 12 through improved infrastructure reliability and more efficient resource utilization.

Compared with real time anomaly detection approaches such as streaming autoencoders, online recurrent models, and lightweight statistical monitoring pipelines, the proposed self supervised transformer achieves competitive detection accuracy while maintaining stable performance under high dimensional input conditions. The attention based architecture enables richer temporal context modeling without significantly reducing detection responsiveness. Experimental observations show that detection latency remains within acceptable monitoring intervals, indicating feasibility for near real time deployment. However, transformer based attention and large scale self supervised pretraining require higher computational resources, including greater memory usage and longer initialization time. Therefore, practical deployment may require optimized inference pipelines, hardware acceleration, or model compression to balance detection accuracy, latency, and infrastructure cost.

5. MANAGERIAL IMPLICATIONS

The findings of this study provide practical implications for organizations managing complex digital infrastructures. The proposed self-supervised transformer framework enables a shift from reactive monitoring toward more proactive and predictive system management. By detecting anomalies without relying on labeled data, the framework helps reduce operational risks, prevent service disruptions, and improve system reliability across large-scale digital environments. It also supports more efficient monitoring by reducing manual analysis and providing real time insights into system behavior.

From an operational perspective, implementing this framework requires consideration of infrastructure readiness, computational resources, and integration with existing monitoring pipelines. Organizations must balance detection performance with implementation costs such as hardware capacity, model training resources, and latency constraints in real time environments. Effective deployment may therefore involve model optimization, scalable system architecture, and periodic model updates to adapt to evolving system patterns.

At the strategic level, the framework contributes to improved digital governance and infrastructure resilience through more reliable monitoring and risk management practices. AI driven anomaly detection can support service reliability, compliance monitoring, and operational stability across centralized and distributed systems. By reducing dependency on labeled anomaly data, the approach also lowers adoption barriers and enables broader implementation in organizations with different levels of technical capability.

6. CONCLUSION

This study presents a self supervised transformer based framework for anomaly detection in high dimensional time series data, addressing key limitations of traditional statistical and supervised learning approaches. By leveraging attention mechanisms and self supervised learning objectives, the model captures complex temporal dependencies and inter variable relationships without relying on labeled anomaly data. Experimental results show consistent improvements over baseline methods in detection accuracy and scalability, while maintaining robustness as data dimensionality increases. However, the experiments are conducted under controlled experimental settings and specific datasets, which may not fully represent the diversity of real world operational environments. These findings demonstrate that self supervised transformers provide a reliable and effective solution for anomaly detection in modern software intensive systems. Furthermore, the proposed framework contributes to the advancement of software analytics by enabling more intelligent and data driven monitoring of complex system behavior. Its ability to extract meaningful representations from large-scale unlabeled data supports more efficient analysis of operational patterns, system performance, and emerging risks, thereby strengthening analytical capabilities in software driven environments.

From a practical and societal perspective, the proposed framework supports the development of resilient digital infrastructures by enabling proactive system monitoring, reducing operational risk, and minimizing downtime in large scale platforms such as cloud services and industrial systems. It is also highly relevant to decentralized environments, including blockchain based infrastructures, where continuous synchronization, real time transaction validation, and consensus processes generate complex high dimensional time series data. By learning system behavior without labeled anomalies, the model effectively detects malicious activity, abnormal transaction flows, and network instability in distributed settings. Nevertheless, practical deployment in real world systems may require additional consideration of computational resources, system scalability, and latency constraints in large scale monitoring pipelines. Its ability to capture long range dependencies and cross node temporal interactions further enables adaptive monitoring across heterogeneous and dynamically evolving decentralized architectures.

By improving monitoring reliability in distributed and trust minimized environments, this research supports the development of intelligent digital infrastructure and aligns with the United Nations sustainable development goals, particularly SDGs 9 and SDGs 12. The framework also reflects progress in deep learning, showing that advanced representation learning can enable scalable, adaptive, and data efficient monitoring in complex digital ecosystems. Despite these contributions, future research should focus on improving model interpretability, developing adaptive or continual learning to address concept drift, and extending the framework to multimodal data and real time deployment scenarios. Further validation across diverse datasets and real world environments is also needed to strengthen the generalizability of the proposed approach and support the advancement of sustainable and intelligent software systems.

7. DECLARATIONS

7.1. About Authors

Aswadi Jaya (AJ) https://orcid.org/0000-0001-5706-0977

Derlina (DD) https://orcid.org/0000-0002-8650-129X

Qurotul Aini (QA) https://orcid.org/0000-0002-7546-5721

Agung Rizky (AR) https://orcid.org/0009-0006-7046-8639

Richard Evans (RE) https://orcid.org/0009-0007-7280-8323

7.2. Author Contributions

Conceptualization: AJ, DD, QA, AR, and RE; Methodology: AJ, DD, QA, AR, and RE; Software: AJ, DD, QA, AR, and RE; Validation: AJ, DD, QA, AR, and RE; Formal Analysis: AJ, DD, QA, AR, and RE; Investigation: AJ, DD, QA, AR, and RE; Resources: AJ, DD, QA, AR, and RE; Data Curation: AJ, DD, QA, AR, and RE; Writing Original Draft Preparation: AJ, DD, QA, AR, and RE; Writing Review and Editing: AJ, DD, QA, AR, and RE; Visualization: AJ, DD, QA, AR, and RE; All authors, AJ, DD, QA, AR, and RE, have read and agreed to the published version of the manuscript.

7.3. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

7.4. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

7.5. Declaration of Conflicting Interest

The authors declare that they have no conflicts of interest, known competing financial interests, or personal relationships that could have influenced the work reported in this paper.

REFERENCES

[1] S. Arshad, M. Ha, and T.-H. Park, "Stad: Self-supervised transformer for anomaly detection in multivariate time series data," *IEEE Access*, 2025.

[2] Y. Cho, J.-H. Lee, G. Ham, D. Jang, and D.-s. Kim, "Generality-aware self-supervised transformer for multivariate time series anomaly detection: Y. cho et al." *Applied Intelligence*, vol. 55, no. 7, p. 604, 2025.

[3] S. Gong, T. Kim, and J. Jeong, "Spt-ad: Self-supervised pyramidal transformer network-based anomaly detection of time series vibration data," *Applied Sciences*, vol. 15, no. 9, p. 5185, 2025.

[4] D. L. Marino, C. S. Wickramasinghe, C. Rieger, and M. Manic, "Self-supervised and interpretable anomaly detection using network transformers," *IEEE Transactions on Industrial Informatics*, 2025.

[5] J. Siswanto, Hendry, U. Rahardja, I. Sembiring, E. Sediyono, K. D. Hartomo, and B. Istiyanto, "Deep learning-based lstm model for number of road accidents prediction," in *AIP Conference Proceedings*, vol. 3234, no. 1. AIP Publishing LLC, 2025, p. 050004.

[6] A. Sánchez-Ferrera, B. Calvo, and J. A. Lozano, "A review on self-supervised learning in time series anomaly detection: Recent advances and open challenges," *ACM Computing Surveys*, vol. 58, no. 5, pp. 1–35, 2025.

[7] C. Wang, H. Shi, J. Hu, X. Yang, J. Zhang, S. Du, and T. Li, "Pre-training enhanced transformer for multivariate time series anomaly detection," *Information Fusion*, vol. 121, p. 103171, 2025.

[8] N. K. Alapati, "Enhancing anomaly detection in time-series data using transformer-based models," in *2025 5th International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT)*. IEEE, 2025, pp. 1–5.

[9] C. Shimillas, K. Malialis, K. Fokianos, and M. M. Polycarpou, "Transformer-based multivariate time series anomaly localization," in *2025 ieee symposium on computational intelligence on engineering/cyber physical systems (cies)*. IEEE, 2025, pp. 1–8.

[10] P. H. P. Tan, S. Wijaya, U. Rahardja, B. N. Henry, and A. Anjani, "Modeling the impact of digital literacy on ai based learning adoption through perceived usefulness and easeof use," *Sundara Advanced Research on Artificial Intelligence*, vol. 1, no. 2, pp. 56–64, 2025.

[11] Z. Z. Darban, G. I. Webb, S. Pan, C. C. Aggarwal, and M. Salehi, "Carla: Self-supervised contrastive representation learning for time series anomaly detection," *Pattern Recognition*, vol. 157, p. 110874, 2025.

[12] L. Zhang, J. Zhu, G. Han, B. Jin, P. Wang, and X. Wei, "Self-supervised disentangled representation learning for time series anomaly detection," *IEEE Internet of Things Journal*, 2025.

[13] E. Edozie, A. N. Shuaibu, B. O. Sadiq, and U. K. John, "Artificial intelligence advances in anomaly detection for telecom networks," *Artificial Intelligence Review*, vol. 58, no. 4, p. 100, 2025.

[14] H. Qiao, H. Tong, B. An, I. King, C. Aggarwal, and G. Pang, "Deep graph anomaly detection: A survey and new perspectives," *IEEE Transactions on Knowledge and Data Engineering*, 2025.

[15] M. A. Syari, U. Rahardja, T. Wellem, H. D. Purnomo, and R. Buaton, "Iot enabled smart farming system for optimizing crop management using sensors and machine learning," in *2025 4th International Conference on Creative Communication and Innovative Technology (ICCIT)*. IEEE, 2025, pp. 1–7.

[16] H. Huang, P. Wang, J. Pei, J. Wang, S. Alexanian, and D. Niyato, "Deep learning advancements in anomaly detection: A comprehensive survey," *IEEE Internet of Things Journal*, 2025.

[17] J. Zhang, G. Wang, Y. Jin, and D. Huang, "Towards training-free anomaly detection with vision and language foundation models," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2025, pp. 15 204–15 213.

[18] Z. Li, Y. Yan, X. Wang, Y. Ge, and L. Meng, "A survey of deep learning for industrial visual anomaly detection," *Artificial Intelligence Review*, vol. 58, no. 9, p. 279, 2025.

[19] J. Zheng, C. Yang, T. Zhang, L. Cao, B. Jiang, X. Fan, X.-m. Wu, and X. Zhu, "Dynamic spectral graph anomaly detection," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 39, no. 12, 2025, pp. 13 410–13 418.

[20] N. Azizah, P. A. Sunarya, U. Rahardja, A. B. Mutiara, P. Prihandoko, and C. Pasha, "Improving smear-negative tuberculosis detection using data augmentation and faster r-cnn," *International Journal of Cyber and IT Service Management (IJCITSM)*, vol. 6, no. 1, pp. 65–77, 2026.

[21] C. Li, W. Xie, B. Zheng, Q. Yi, L. Yang, B. Hu, and C. Deng, "An enhanced clkan-rf framework for robust anomaly detection in unmanned aerial vehicle sensor data," *Knowledge-Based Systems*, vol. 319, p. 113690, 2025.

[22] X. Qiu, Z. Li, W. Qiu, S. Hu, L. Zhou, X. Wu, Z. Li, C. Guo, A. Zhou, Z. Sheng *et al.*, "Tab: Unified benchmarking of time series anomaly detection methods," *arXiv preprint arXiv:2506.18046*, 2025.

[23] J. Paparrizos, P. Boniol, Q. Liu, and T. Palpanas, "Advances in time-series anomaly detection: Algorithms, benchmarks, and evaluation measures," in *Proceedings of the 31st ACM SIGKDD conference on knowledge discovery and data mining v. 2*, 2025, pp. 6151–6161.

[24] D. A. Dewi, H. K. R. Singh, J. Periasamy, T. B. Kurniawan, Henderi, M. S. Hasibuan, and Y. Nathan, "Incorporate transformer-based models for anomaly detection," *Journal of Applied Data Sciences*, vol. 6, no. 3, pp. 2046–2055, 2025. [Online]. Available: https://garuda.kemdiktisaintek.go.id/documents/detail/5284270

[25] S. Shaumiwaty, H. R. C. MOCHAMAD, and N. HENI, "Enhancing personalized learning using artificial intelligence and machine learning approaches," *BLOCKCHAIN FRONTIER TECHNOLOGY: Pandawan*, vol. 4, no. 2, pp. 156–170, 2025.

[26] X. Ren, W. Wei, L. Xia, and C. Huang, "A comprehensive survey on self-supervised learning for recommendation," *ACM Computing Surveys*, vol. 58, no. 1, pp. 1–38, 2025.

[27] Y. Chen, Z. Wang, J. Wang, Y. Chu, Q. Zhang, Z. A. Li, and X. Zeng, "Self-supervised learning in drug discovery," *Science China Information Sciences*, vol. 68, no. 7, p. 170103, 2025.

[28] C. Liu, G. Gui, Y. Wang, T. Ohtsuki, D. Niyato, and X. S. Shen, "A comprehensive survey on self-supervised learning for specific emitter identification," *IEEE Communications Surveys & Tutorials*, 2025.

[29] X. Wu, D. DeTone, D. Frost, T. Shen, C. Xie, N. Yang, J. Engel, R. Newcombe, H. Zhao, and J. Straub, "Sonata: Self-supervised learning of reliable point representations," in *Proceedings of the Computer Vision and Pattern Recognition Conference*, 2025, pp. 22 193–22 204.

[30] N. Lutfiani, D. A. Astrieta, V. Wildan, H. Sulistyaningrum, M. R. Anwar, and E. D. Astuti, "Emotional well-being and psychological support in infertility a multi-modal ai approach," *International Journal of*

Cyber and IT Service Management (IJCITSM), vol. 5, no. 1, pp. 81–92, 2025.

[31] Y. Zhang, Y. Zhang, W. Liao, X. Li, and X. Wang, "Multi-view self-supervised learning on heterogeneous graphs for recommendation," *Applied Soft Computing*, vol. 174, p. 113056, 2025.

[32] R. Bushuiev, A. Bushuiev, R. Samusevich, C. Brungs, J. Sivic, and T. Pluskal, "Self-supervised learning of molecular representations from millions of tandem mass spectra using dreams," *Nature Biotechnology*, pp. 1–11, 2025.

[33] Z. Zamanzadeh Darban, G. I. Webb, S. Pan, C. Aggarwal, and M. Salehi, "Deep learning for time series anomaly detection: A survey," *ACM Computing Surveys*, vol. 57, no. 1, pp. 1–42, 2024.

[34] P. Wu, C. Pan, Y. Yan, G. Pang, Q. Yan, P. Wang, and Y. Zhang, "Deep learning for video anomaly detection: A review," *IEEE Transactions on Neural Networks and Learning Systems*, 2026.

[35] Y. Ma, W. Liu, C. Xu, L. Bai, E. Zhang, and J. Wang, "Multivariate time series anomaly detection based on inverted transformer with multivariate memory gate," *Entropy*, vol. 27, no. 9, p. 939, 2025. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC12469927/

[36] Q. Liu, S. Lee, and J. Paparrizos, "Tsb-autoad: Towards automated solutions for time-series anomaly detection," *Proceedings of the VLDB Endowment*, vol. 18, no. 11, pp. 4364–4379, 2025.

[37] T. Wald, C. Ulrich, J. Suprijadi, S. Ziegler, M. Nohel, R. Peretzke, G. Kohler, and K. Maier-Hein, "An openmind for 3d medical vision self-supervised learning," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2025, pp. 23 839–23 879.

[38] Y. Zhang, X. Wu, Y. Lao, C. Wang, Z. Tian, N. Wang, and H. Zhao, "Concerto: Joint 2d-3d self-supervised learning emerges spatial representations," *arXiv preprint arXiv:2510.23607*, 2025.

[39] A. Rozi, J. Junengsih, S. Alam, A. Pawar, W. Sumarjo, and D. Sunarsi, "Impact of hr management on ai implementation and data protection in indonesian manufacturing," *International Journal of Cyber and IT Service Management (IJCITSM)*, vol. 6, no. 1, pp. 51–64, 2026.

[40] D. Khettaf, D. Djenouri, Z. Rezaeifar, and Y. Djenouri, "Deep learning for collective anomaly detection," *ACM Computing Surveys*, 2026.

[41] C. Fan, Y. Ding, F. Geng, Z. Lv, and K. Yang, "Anomaly warning classification based on deep learning method and temperature-induced effect of bridge crack monitoring data," *Engineering Structures*, vol. 352, p. 122158, 2026.

[42] S. A. Moghaddas, S. Ajayi, X. Wang, Y. Liu, Y. Huang, and Y. Bao, "Automatic multi-anomaly detection of pipelines with ensemble deep learning-based computer vision," *Advanced Engineering Informatics*, vol. 70, p. 104201, 2026.

[43] A. S. A. R. Dapit, M. Othman, M. Azuddin, A. A. Puzi, and U. Rahardja, "A computational model for stress intervention using affective brain-computer interfaces," in *2025 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET)*. IEEE, 2025, pp. 114–119.

[44] S.-K. S. Fan and L.-C. Lai, "A novel diffusion model-based deep learning approach for anomaly detection in computer connector image analysis," *Engineering Applications of Artificial Intelligence*, vol. 166, p. 113604, 2026.

[45] J. Yin, X. Hu, T. Nan, and C. Lu, "Spatiotemporal prediction and attribution of groundwater storage anomaly using enhanced hybrid deep learning modeling with uncertainty quantification," *Journal of Environmental Management*, vol. 400, p. 128766, 2026.

[46] C. Zargane, A. Guezzaz, S. Benkirane, and M. Azrour, "An overview of deep learning-based anomaly detection in environmental data," *Smart Technologies for a Sustainable Environment*, pp. 103–112, 2026.

[47] E. T. Rusmiati, L. Febrina, Y. Sari, and E. M. S. Sakti, "Adoption of ai driven ecological preaching systems using sem pls analysis," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 8, no. 1, pp. 284–295, 2026.

[48] Y. Li, B. Gan, R. Zhu, X. Chen, Y. Cui, H. Wang, and L. Wu, "Significance of atlantic sea surface temperature anomalies to arctic sea ice variability revealed by deep learning," *npj Climate and Atmospheric Science*, 2026.

[49] V. Rani, S. T. Nabi, M. Kumar, A. Mittal, and K. Kumar, "Self-supervised learning: A succinct review," *Archives of Computational Methods in Engineering*, vol. 30, no. 4, pp. 2761–2775, 2023.

[50] M. Lou, Y. Fu, and Y. Yu, "Sparx: A sparse cross-layer connection mechanism for hierarchical vision mamba and transformer networks," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 39, no. 18, 2025, pp. 19 104–19 114.