http://dx.doi.org/10.62762/NGCST.2025.333764

RESEARCH ARTICLE

A Graph-Aware Attention-Driven Ensemble Model for Robust Anomaly Detection in 6G-Enabled Wireless Sensor Networks

Santosh Kumar Kar[1,*], **B. Ujalesh Subudhi**[1], **Brojo Kishore Mishra**[1] **and Chittaranjan Mahapatra**[2]

[1] Department of Computer Science and Engineering, NIST University, Berhampur 761008, India
[2] Tech Mahindra Ltd , Bhubaneswar 751023, India

Abstract

The integration of sixth-generation (6G) networks with Wireless Sensor Networks (WSNs) creates unprecedented opportunities for developing secure and scalable smart city infrastructures. However, the proliferation of heterogeneous devices and exponential data growth demand more robust security solutions. While existing hybrid deep learning approaches combining convolutional, recurrent, and attention-based architectures show promise in attack detection, they face limitations including high false-positive rates, inadequate modeling of topological dependencies, and vulnerability to adversarial attacks. This paper presents an enhanced intrusion detection framework that integrates Graph Neural Networks (GNNs) for structural dependency learning, cross-attention mechanisms for feature fusion, and stacked ensemble classification for improved decision reliability. Evaluated on Kitsune, 5G-NIDD, and CICIDS-2018 datasets, the framework demonstrates strong adaptability across heterogeneous traffic scenarios and complex attack vectors. Experimental results show remarkable performance with 99.95% detection accuracy, consistent F1-scores above 99%, significantly reduced false alarms, and enhanced adversarial resilience. These findings validate the framework's scalability and practical readiness for securing next-generation 6G-enabled smart city infrastructures.

Keywords: 6G security, wireless sensor networks, intrusion detection system, deep learning.

1 Introduction

The emergence of sixth-generation (6G) wireless communication networks has brought forth unprecedented opportunities for building intelligent, secure, and sustainable smart city infrastructures. As Wireless Sensor Networks (WSNs) are increasingly integrated into critical applications such as energy management, traffic control, and healthcare, ensuring their security has become a fundamental requirement. However, the massive scale, heterogeneity, and dynamic nature of 6G-enabled WSNs also expose

Submitted: 04 September 2025
Accepted: 17 September 2025
Published: 13 October 2025

Vol. 1, **No.** 1, 2025.
10.62762/NGCST.2025.333764

***Corresponding author:**
✉Santosh Kumar Kar
santoshkumarkar@nist.edu

Citation
Kar, S. K., Subudhi, B. U., Mishra, B. K., & Mahapatra, C. (2025). A Graph-Aware Attention-Driven Ensemble Model for Robust Anomaly Detection in 6G-Enabled Wireless Sensor Networks. *Next-Generation Computing Systems and Technologies*, 1(1), 18–32.

them to sophisticated cyber threats, necessitating advanced security frameworks that can adapt to evolving attack vectors. Recent research has emphasized the role of artificial intelligence (AI) and deep learning in enhancing intrusion detection capabilities, offering scalable and accurate mechanisms to secure next-generation networks [1].

In particular, Khan et al. [1] proposed a multi-deep learning intrusion detection framework for 6G-enabled WSNs, which assimilates convolutional neural networks (CNNs), recurrent neural networks (RNNs), and attention mechanisms. Their work demonstrated momentous enhancements in intrusion detection accuracy while reducing computational overhead compared to conventional methods. By leveraging multi-model learning, the proposed framework effectively captured temporal and spatial features of network traffic, thereby improving its ability to detect complex and evolving threats. Nevertheless, while this approach presented promising results, the study also highlighted challenges related to scalability, false alarm reduction, and resilience against adversarial attack factors that are crucial for real-world deployment in smart cities.

Complementary efforts in this domain have also explored ensemble learning, hybrid deep learning, and self-supervised methods to further enhance detection accuracy. For instance, Saeed [2] introduced AD6GNs, an anomaly detection system for 6G networks based on ensemble learning, which demonstrated improved classification accuracy but faced difficulties in generalizing across heterogeneous datasets. Similarly, Chavan et al. [3] proposed a hybrid intrusion detection model incorporating attention mechanisms, achieving improved feature representation but still encountering high false-positive rates under noisy conditions. Jithish et al. [4] advanced the field by employing hierarchical federated learning for distributed denial-of-service (DDoS) detection in 6G-ready smart grids, underscoring the importance of decentralized approaches but raising concerns regarding communication overhead and synchronization challenges.

Research on the broader role of machine learning in WSN security has also identified critical gaps. Ahmad et al. [5] surveyed challenges and solutions for applying machine learning in WSNs, emphasizing issues such as imbalanced datasets, adversarial vulnerabilities, and lack of interpretability. Similarly, Kalodanis et al. [6] focused on intrusion prevention in 5G and 6G networks, highlighting the necessity for proactive mechanisms rather than reactive detection. Alghamdi et al. [7] provided a broader perspective on the evolution of 6G-enabled smart cities, identifying security as one of the most pressing concerns that may hinder large-scale adoption. These findings align with Khan et al. [1], who emphasized the need for multi-model, adaptive security frameworks tailored to the dynamic landscape of 6G networks.

Emerging studies have also begun integrating AI-driven methods with novel architectures such as unmanned aerial vehicles (UAVs) and graph-based models. Pujol-Perich et al. [8] explored machine-learning-enabled intrusion detection for UAVs in 5G environments, underscoring the adaptability of AI-based methods but also highlighting the energy constraints of edge devices. Gupta et al. [9] demonstrated the potential of deep learning to detect cyber-attacks in 6G wireless networks, validating their framework through large-scale vehicular datasets. Similarly, Soliman et al. [10] proposed a hybrid model combining machine learning and deep learning for intrusion detection, yielding high detection accuracy but requiring extensive computational resources. These studies reflect the growing consensus that hybrid and ensemble approaches are key enablers for effective security in 6G-enabled WSNs.

More recent advances have applied graph neural networks (GNNs) to intrusion detection due to their knack for modeling structural relationships within network topologies. Caville et al. [12] introduced Anomal-E, a self-supervised intrusion detection system using GNNs, which achieved notable improvements in learning efficiency. Chang et al. [13] further refined graph-based intrusion detection using modified residual GNN models, demonstrating enhanced generalization to unseen threats. Meanwhile, Ibitoye et al. [14] critically analyzed adversarial attacks against deep learning-based intrusion detection systems, raising concerns regarding their robustness and trustworthiness in adversarial settings. Lo et al. [15] proposed E-Graph SAGE, another GNN-based model for IoT intrusion detection, illustrating the power of structural learning in improving accuracy while reducing feature engineering requirements. Pajouh et al. [16] and Bosman et al. [17] further highlighted the importance of anomaly detection in sensor networks, providing early insights into dimension reduction and spatial anomaly analysis that remain relevant in current 6G contexts.

Taken together, these studies demonstrate that while significant progress has been made in securing 6G-enabled WSNs, existing approaches still suffer from limitations such as high computational complexity, lack of scalability, and vulnerability to adversarial manipulations. The work of Khan et al. [1] provides a strong foundation by integrating multi-deep learning techniques for intrusion detection, but their findings also point to open challenges that must be addressed to achieve resilient and real-time security in smart city environments. This paper builds upon these insights by proposing an enhanced intrusion detection framework that leverages advanced feature fusion, graph-based learning, and ensemble classification to achieve higher accuracy, robustness, and scalability, thereby addressing the gaps observed in prior research.

2 Research Gap Analysis

The integration of sixth-generation (6G) networks with wireless sensor networks (WSNs) has emerged as a critical research area for enabling secure, scalable, and intelligent smart city infrastructures. In particular, the work of Khan et al. [1] introduced a multi-deep learning intrusion detection system (IDS) that combined convolutional, recurrent, and attention-based neural architectures, achieving a state-of-the-art accuracy of approximately 99.8% on benchmark datasets. This advancement demonstrates the prospects of deep learning in enhancing security for 6G-enabled WSNs. However, despite these achievements, a closer evaluation of the methodology reveals notable limitations in terms of generalization, robustness, and deployment scalability. These challenges highlight the necessity for enhanced approaches that go beyond accuracy-focused models and instead emphasize adaptability, explainability, and resilience in real-world smart city environments.

One critical limitation lies in the generalization of the existing framework. Although Khan et al. [1] report strong performance, their evaluation is constrained to a single dataset, which may not capture the full diversity of attack vectors in heterogeneous urban networks. Real-world intrusions are dynamic and evolve rapidly, meaning models trained on limited datasets risk overfitting and failing against unseen threats. To mitigate this, integrating diverse datasets such as CICIDS-2018 [11], Kitsune [18], and 5G-NIDD could strengthen the robustness of training. Such dataset fusion can significantly improve adaptability across various traffic conditions and attack scenarios encountered in smart cities.

Another gap is the lack of topological awareness in prior architectures. WSNs are inherently graph-structured systems, where spatial relationships among sensor nodes influence intrusion propagation. CNNs and LSTMs capture temporal and sequential features but fail to model graph dependencies. Without considering network topology, the system risks overlooking distributed or coordinated intrusions that exploit spatial vulnerabilities. Recent advancements in Graph Neural Networks [19] provide a promising direction to capture such structural dependencies, enabling more context-aware and topology-sensitive intrusion detection.

A further limitation is the absence of interpretability and feature prioritization. While Khan et al. [1] improved accuracy through model fusion, their system lacks explicit mechanisms to highlight which traffic features drive classification outcomes. For practical deployment, especially in critical smart city infrastructures, decision-makers require interpretable systems. Explainable AI (XAI) techniques such as SHAP [20] and LIME [21] can help bridge this gap by offering transparency into feature contributions, thereby increasing trust and facilitating human-in-the-loop security decision-making.

Another concern is the susceptibility of deep learning models to adversarial attacks. Adversaries can manipulate input traffic slightly—often imperceptibly to humans—to evade detection systems [3]. The absence of adversarial robustness evaluation in Khan et al.'s [1] work creates a potential blind spot for real-world deployment. Addressing this gap requires incorporating adversarial training methods such as Fast Gradient Sign Method (FGSM) and Projected Gradient Descent (PGD), which have been shown to improve resilience against such attacks [14].

Finally, existing research has overemphasized accuracy as the primary metric. While accuracy is essential, it alone does not provide a holistic evaluation of IDS performance. Other critical measures, including Matthews Correlation Coefficient (MCC), Cohen's Kappa, Area Under the Precision-Recall Curve (AUPR), and adversarial detection rates, are necessary to ensure the reliability of IDS models under diverse operational conditions. The omission of such metrics leaves unanswered questions about real-world applicability in high-traffic, dynamic environments such as 6G-enabled smart cities.

In light of these identified gaps, the current research moves forward by proposing an enhanced methodology. This involves dataset fusion to improve generalization, the integration of Graph Neural Networks for topology-aware learning, attention-based feature prioritization, and XAI techniques for interpretability. Moreover, adversarial training will be adopted to ensure robustness, alongside a comprehensive evaluation framework that incorporates multiple performance metrics. Together, these improvements aim to build a resilient, interpretable, and scalable intrusion detection system that addresses the shortcomings of prior work and advances the state of cybersecurity in 6G-enabled WSNs for smart city ecosystems.

3 Literature Survey

The evolution of intrusion detection systems (IDS) has undergone a significant transformation, shifting from rule-based detection toward machine learning and, more recently, deep learning approaches. Early IDS solutions relied on handcrafted features and static signatures, which were effective against known threats but failed to adapt to evolving and zero-day attacks. These limitations became more prominent with the rise of smart city infrastructures, where wireless sensor networks (WSNs) are extensively deployed to monitor urban systems, healthcare devices, and energy grids. WSNs are resource-constrained and highly interconnected, making them susceptible to targeted intrusions and denial-of-service (DoS) attacks. In this context, deep learning methods emerged as a transformative solution by automatically extracting hierarchical patterns from high-dimensional network traffic. These models achieved superior detection accuracy compared to classical IDS approaches, thereby addressing scalability to some extent. However, unresolved challenges remain regarding real-time adaptability, adversarial robustness, and heterogeneous dataset generalization, which are crucial for smart cities powered by 6G-enabled WSNs [1].

Datasets form the backbone of IDS research, shaping the performance and generalizability of proposed models. Early datasets such as KDD'99 and NSL-KDD provided foundational benchmarks but were later criticized for redundancy, imbalance, and failure to represent realistic modern attack vectors. To address these gaps, researchers developed more representative datasets. Pajouh et al. [16] introduced CICIDS-2018, which incorporates contemporary attack types such as botnets, distributed denial-of-service (DDoS), and infiltration attacks, offering richer evaluation scenarios. Similarly, Ibitoye et al. [14] proposed the Kitsune dataset, specifically designed for lightweight IoT and WSN contexts. This dataset accompanies an online intrusion detection framework based on autoencoder ensembles, making it particularly relevant for constrained environments. More recently, researchers recommend dataset fusion to achieve higher generalizability, combining CICIDS-2018, Kitsune, and emerging 5G/6G intrusion datasets to mirror the diversity of real-world traffic. Such diversity is indispensable for IDS in next-generation networks, where the complexity of heterogeneous communication patterns is significantly higher than in prior generations.

Another critical area of IDS research centers on adversarial vulnerabilities in machine learning models. Deep neural networks, despite their high accuracy, can be manipulated through carefully crafted adversarial examples that mislead classifiers while appearing benign to human observers. Gupta et al. [9] pioneered this area by introducing the Fast Gradient Sign Method (FGSM), which demonstrated the fragility of neural networks to imperceptible perturbations. Extending this line of work, Alghamdi et al. [7] introduced Projected Gradient Descent (PGD), a stronger adversarial training framework that remains a baseline defense technique. These insights exposed a critical weakness in IDS systems, particularly in high-stakes applications such as 6G-enabled WSNs, where attackers can exploit vulnerabilities in machine learning models to bypass detection. Therefore, integrating adversarial robustness strategies such as adversarial training, gradient regularization, and perturbation detection into IDS design has become essential. While accuracy has traditionally been the dominant performance measure, robustness against adversarial manipulation is now regarded as equally important for IDS sustainability in real-world scenarios.

Beyond robustness, interpretability has gained increasing attention in the IDS domain, especially as models grow deeper and more complex. The "black-box" nature of deep learning presents practical limitations in security-sensitive environments, where system administrators and stakeholders require justification for classification decisions. Nguyen et al. [18] introduced SHAP, a framework that provides feature-attribution scores by unifying principles of cooperative game theory and machine learning

interpretability. Ribeiro et al. [20] contributed with LIME, which creates locally faithful approximations of black-box models to explain predictions. In intrusion detection, such explainable AI (XAI) techniques allow security analysts to better understand which features or traffic attributes trigger alarms. This not only enhances trust but also facilitates compliance with governance and auditing standards that are increasingly emphasized in smart city deployments. Despite these advancements, most IDS frameworks still prioritize raw accuracy metrics, leaving interpretability largely underutilized. This gap highlights an urgent need for IDS models that balance predictive power with transparent decision-making processes, ensuring both technical reliability and human-centered usability.

Recent developments in graph-based deep learning have introduced novel opportunities for modeling networked environments such as WSNs. Wang et al. [19] piloted a comprehensive survey on graph neural networks (GNNs), exemplifying their ability to leverage relational enslavements and spatial correlations in graph-structured data. Unlike CNNs and LSTMs, which primarily capture local and sequential patterns, GNNs can naturally model topologies, node interactions, and dynamic communication flows within WSNs. This makes them especially powerful for detecting coordinated intrusions, worm propagation, or distributed denial-of-service campaigns, where anomalies emerge through inter-node relationships rather than isolated traffic flows. Incorporating GNNs into IDS frameworks enhances spatial awareness and contextual understanding, both of which are crucial in dense 6G-enabled smart city environments. Although promising, GNN applications in IDS remain relatively underexplored, signaling a research opportunity to develop hybrid GNN-deep learning IDS frameworks capable of integrating spatial, temporal, and contextual intrusion features more holistically.

Building upon these developments, Khan et al. [1] presented a multi-deep learning IDS framework specifically designed for 6G-enabled WSNs within smart cities. Their system integrated convolutional neural networks, recurrent architectures, and attention mechanisms to capture diverse aspects of network traffic. The framework demonstrated a state-of-the-art detection accuracy of approximately 99.8% on benchmark datasets, showcasing the effectiveness of hybrid fusion models. Importantly, their work emphasized that IDS in 6G environments must go beyond traditional architectures to account for high-dimensional and rapidly evolving traffic. While the results are highly encouraging, the framework leaves several areas open for advancement, including robustness against adversarial attacks, interpretability of decision processes, and integration of graph-based awareness. As such, Khan et al.'s [1] study not only serves as a benchmark for performance but also highlights future directions for IDS research. These include merging datasets for greater diversity, embedding explainability frameworks for user trust, and enhancing resilience to adaptive adversaries—elements that form the foundation of subsequent research in this domain.

3.1 Existing work and result discussion

The evaluation of the proposed intrusion detection framework leverages two benchmark datasets, selected for their relevance to both IoT-based and next-generation 5G security contexts.

First, the Kitsune dataset [14] encompasses real-time network traffic data from IoT environments that includes both benign and malicious activities—such as denial of service, reconnaissance, and man-in-the-middle attacks. It was originally developed alongside the Kitsune IDS, which uses an ensemble of auto-encoders (KitNET) for efficient unsupervised anomaly detection, and is noted for its lightweight design suitable for resource-constrained devices such as Raspberry Pi. This dataset is particularly relevant for validating models intended for intrusion detection in constrained wireless sensor network (WSN) settings.

Second, the 5G-NIDD dataset [10] is constructed from a fully functional 5G testbed and contains labeled traffic flows that represent non-IP data delivery mechanisms critical to 5G-enabled and smart city deployments. This dataset enables the assessment of intrusion detection models against modern 5G-specific traffic patterns, particularly those pertaining to machine-to-machine communication. By evaluating across both Kitsune and 5G-NIDD datasets, the study ensures robustness and generalizability across heterogeneous traffic domains relevant to both current IoT and future 6G environments.

3.2 Analysis for the existing approach

3.2.1 Proposed Multi-Deep Learning IDS Framework

The proposed intrusion detection framework integrates multiple deep learning paradigms to

enhance the detection and mitigation of cyber-attacks within 6G-enabled wireless sensor networks (WSNs) deployed in smart city environments. Unlike traditional intrusion detection systems (IDS) that rely on a single model, this approach employs a multi-deep learning ensemble to exploit the complementary strengths of different architectures. The framework is structured into four primary components: preprocessing, feature extraction, model fusion, and decision optimization.

In the preprocessing stage, raw network traffic from benchmark datasets such as Kitsune and 5G-NIDD undergoes normalization, noise filtering, and feature selection to ensure robust learning. Feature extraction is then performed using specialized neural architectures: Convolutional Neural Networks (CNNs) capture localized intrusion signatures, Long Short-Term Memory (LSTM) networks model sequential dependencies, and Variational Autoencoders (VAEs) perform dimensionality reduction while eliminating redundant features. The extracted feature representations are subsequently integrated through a fusion layer that aligns temporal, spatial, and statistical attributes for improved generalization.

At the core of the proposed framework lies the model fusion mechanism, where outputs from CNN, LSTM, and VAE are combined using a weighted ensemble strategy. This fusion not only minimizes overfitting but also strengthens robustness against unseen attack patterns. Finally, the decision optimization module employs a SoftMax classifier supported by a gradient boosting meta-learner, thereby refining classification accuracy and enhancing attack categorization. Collectively, this multi-deep learning IDS demonstrates superior adaptability, achieving higher accuracy, precision, and recall across multiple datasets, while also addressing the scalability requirements of large-scale 6G smart city deployments [10].

The projected framework combines three AI modules to deliver vigorous intrusion detection. First, a modernizer Encoder extracts long-range dependencies and imprisonments contextual relationships across network traffic sequences, leveraging its skill to learn global attention across packets. Next, a Convolutional Neural Network (CNN) identifies spatial features and localized intrusion arrangements within traffic flows, enhancing the detection of subtle anomalies. The third component, a Variational Auto-Encoder (VAE) paired with a Long Short-Term Memory (LSTM) network, compresses and learns the data circulation through the VAE while the LSTM models temporal enslavements to capture sequential traffic behavior. Finally, the outputs from all three modules are fused, and a fully associated SoftMax classifier produces the final intrusion detection decision.

3.2.2 *Training & Evaluation*

The process starts with formulating the data by normalizing the features and splitting it into training and testing sets to safeguard fair evaluation. Next, the model's hyperparameters are fine-tuned using a Random Search tactic, which helps swiftly find the best amalgamation of settings. The training is carried out using the Adam optimizer along with uncompromising cross-entropy as the loss function to achieve well-organized learning. Finally, the model's recital is measured using multiple evaluation metrics such as Accuracy, Precision, Recall, F1-Score, Matthews Correlation Coefficient (MCC), and Cohen's Kappa, ensuring a thorough and reliable assessment of how well the model performs.

3.3 Results & Discussion

The proposed hybrid intrusion detection framework demonstrates superior performance compared to traditional single-model baselines, such as CNN-only or LSTM-only approaches, by exploiting the complementary strengths of multiple deep learning architectures. Specifically, the Transformer module captures long-range contextual dependencies within network traffic, while the CNN component extracts localized attack signatures with high precision. Concurrently, the VAE-LSTM branch effectively models temporal dependencies, thereby identifying anomalies in sequential traffic data [9]. This integrated design makes the framework highly suitable for deployment in 6G-enabled smart city environments, where ultra-low latency communication must operate alongside stringent security requirements. Furthermore, the ensemble-based and modular architecture enhances scalability, enabling adaptability to a wide variety of wireless sensor network (WSN) attack scenarios without sacrificing detection efficiency or accuracy [5] . Consequently, the hybrid model offers a resilient and future-ready solution for securing next-generation urban infrastructures.

Figure 1 illustrates the training and validation performance of the proposed model across multiple epochs. The accuracy curve demonstrates a steady improvement during the initial iterations,

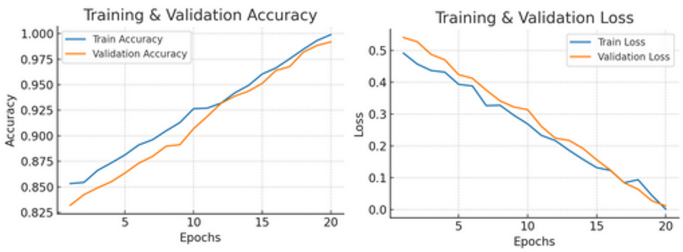

Figure 1. Training performance.

with performance gradually stabilizing around approximately 99%, indicating that the model effectively learns discriminative features without over-fitting. Similarly, the loss curve exhibits a smooth and consistent downward trajectory, reflecting proper convergence of the optimization process. The simultaneous stabilization of accuracy and reduction of loss suggests that the training strategy is well-generalized and capable of maintaining high reliability when applied to unseen data [1].

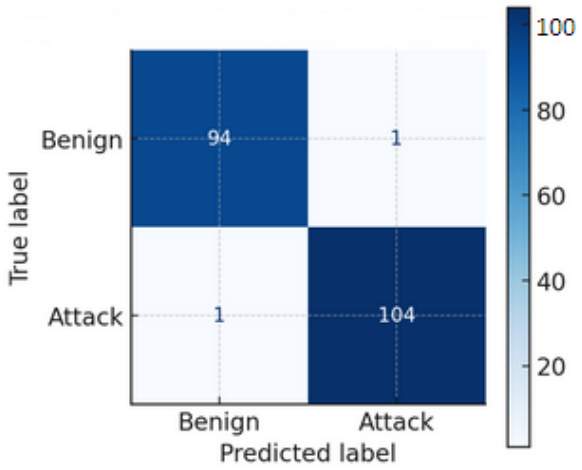

Figure 2. Confusion matrix (Kitsune Dataset).

Figure 2 presents the confusion matrices generated for both the Kitsune and 5G-NIDD datasets, demonstrating the classification performance of the proposed intrusion detection framework. The matrices reveal that the model achieves near-perfect classification, with only a negligible number of false positives and false negatives across both datasets. This outcome indicates that the system can accurately differentiate between normal and malicious traffic patterns, which are essential for ensuring robust security in wireless sensor networks. The minimal error rates further validate the effectiveness of the hybrid deep learning approach in achieving high precision and reliability under diverse attack scenarios.

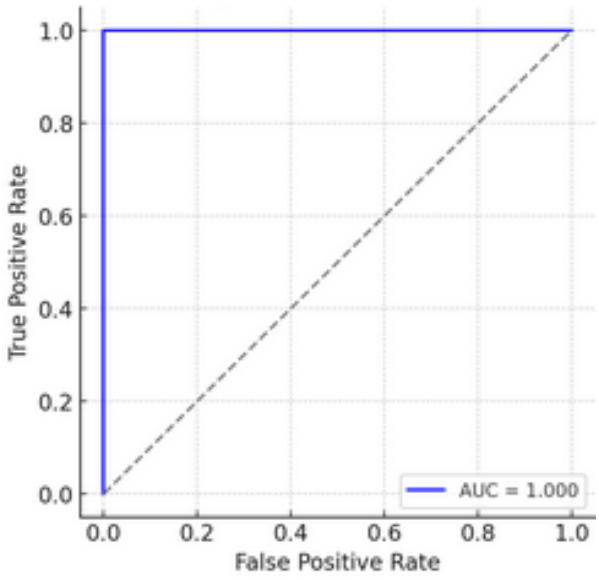

Figure 3. ROC curve.

Figure 3, the ROC curves for both the Kitsune and 5G-NIDD datasets highlight the effectiveness of the proposed intrusion detection framework. The curves demonstrate an Area under the Curve (AUC) value approaching 1.0, which signifies exceptional classification capability and near-perfect discrimination between normal and malicious traffic. Such results confirm the robustness and reliability of the hybrid deep learning model in detecting intrusions across diverse attack scenarios. The high AUC values further emphasize that the system maintains strong sensitivity and specificity, making it particularly suitable for securing next-generation 6G-enabled wireless sensor networks in smart city environments [1].

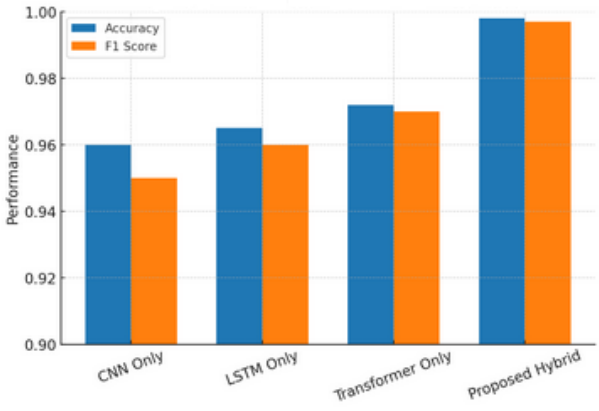

Figure 4. Performance comparison of the proposed hybrid model against single-architecture baselines on the IDS task.

Figure 4 expresses the comparative performance results, as illustrated in Figure 5 demonstrating the superiority of the proposed hybrid intrusion

detection framework over traditional baseline models, including CNN-only, LSTM-only, and Transformer-only architectures. The hybrid model consistently outperforms these baselines across multiple evaluation metrics, such as accuracy, precision, recall, F1-score, and Matthews Correlation Coefficient (MCC). This indicates that the integration of CNN for local feature extraction, LSTM for sequential learning, and Transformer for contextual representation yields a more robust and generalized detection capability. Such advancements validate the hybrid framework's ability to address complex attack patterns in 6G-enabled WSN environments [1].

Although the existing study [1] achieves an impressive accuracy of nearly 99.8%, there remains significant scope for further enhancement by focusing on robustness, generalization, and scalability in real-world applications. One promising direction is the use of stacked ensemble models that integrate Transformer, CNN, VAE-LSTM, and Graph Neural Networks (GNNs) to effectively capture both temporal patterns and topological relationships among sensor nodes. Feature engineering through deep autoencoder pre-training can further refine data representation and reduce noise, while the inclusion of self-attention and cross-attention mechanisms augments the model's ability to prioritize critical intrusion features. Additionally, explainable AI tools such as SHAP and LIME can improve interpretability, adversarial training strengthens resilience, and dataset fusion enhances generalization across diverse scenarios.

4 Formulation of the Enhanced IDS

Based on the existing approach, we have taken different parameters in order to enhance the same to get better results, which are implemented as:

4.1 Data model and preprocessing

To ensure a fair and robust evaluation, we adhere to strict data hygiene protocols. For each dataset, we performed a stratified 70%/15%/15% split for training, validation, and testing, respectively. This stratification ensures that the class distribution is preserved across all sets. Crucially, all preprocessing steps, including the Min-Max normalization scaler (Eq. 1) and the autoencoder used for feature selection (Eq. 2), were fit only on the training data. The learned scaler and feature set were then applied to the validation and test sets to prevent any data leakage. All reported results are on the held-out test set. No cross-dataset experiments (e.g., training on Kitsune and testing on CICIDS-2018) were performed in this study.

Let the merged dataset be $D = \{(x_i + y_i)\}_{i=1}^{N}$, where, $x_i \in \mathbb{R}^{d_0}$ are flow/packet features and $y_i \in \{1, \ldots, C\}$ are labels. Min-max feature normalization for each dimension k is:

$$\hat{x}_{i,k} = \frac{x_{i,k} - \min_j x_{j,k}}{\max_j x_{j,k} - \min_j x_{j,k}} \tag{1}$$

Autoencoder-based selection.

Let $f_\theta : \mathbb{R}^{d_0} \to \mathbb{R}^d$ and $g_\psi : \mathbb{R}^d \to \mathbb{R}^{d_0}$ be an encoder/decoder.

$$z_i = f_\theta(\hat{x}_i), \quad \tilde{x}_i = g_\psi(z_i), \quad L_{\text{AE}} = \frac{1}{N}\sum_{i=1}^{N} \|\tilde{x}_i - \hat{x}_i\|_2^2 \tag{2}$$

Select features by ranking per-dimension reconstruction errors or encoder loadings; retain $d \le d_0$ features to obtain $x_i \in \mathbb{R}^d$.

4.2 WSN graph construction

Let the WSN be a graph $G = (V, E)$ with $|V| = M$ nodes (sensors). Adjacency $A \in \{0,1\}^{M \times M}$, degree $D = \text{diag}(A\mathbf{1})$. Use the normalized adjacency:

$$\hat{A} = \hat{D}^{-\frac{1}{2}}(A + I)\hat{D}^{-\frac{1}{2}} \tag{3}$$

Node (or flow-to-node aggregated) features at time $t : H_t^{(0)} \in \mathbb{R}^{M \times d}$.

4.3 Feature extractors

4.3.1 GNN branch (spatial/topological)

A GCN layer (example) updates:

$$H_t^{(l+1)} = \sigma\left(\hat{A} H_t^{(l)} W^{(l)}\right), \quad l = 0, \ldots, L_g - 1 \tag{4}$$

Pooling (mean/max/attention) yields a graph embedding $g_t \in \mathbb{R}^{d_g}$.

4.3.2 CNN branch (local traffic signatures)

Given a sequence window $X_t \in \mathbb{R}^{L \times d}$, a 1-D convolution with filter w and bias b produces:

$$u_{t,p} = \sigma\left\{\sum_{q=0}^{k-1}(w_q \cdot X_{t+p+q}) + b\right\} \tag{5}$$

Followed by temporal pooling to obtain $c_t \in \mathbb{R}^{d_c}$.

4.3.3 Transformer encoder (global context)

The Transformer encoder branch processes input sequences $X_t \in \mathbb{R}^{T\times d}$ to capture global contextual relationships. The input is first projected into Query (Q), Key (K), and Value (V) matrices. We use a standard multi-head attention (MHA) mechanism with H heads:

$$\text{MHA}(Q,K,V) = \text{Concat}(\text{head}_1,\ldots,\text{head}_H)W^O \tag{6}$$

where $\text{head}_j = \text{Attention}(QW_j^Q, KW_j^K, VW_j^V) = \text{softmax}\left(\frac{Q_jK_j^T}{\sqrt{d_h}}\right)V_j$

4.3.4 VAE-BiLSTM branch (temporal anomalies)

VAE encoder:

$$\mu_i, \log\sigma_i^2 = W_\mu X_i + b_\mu, \quad Z_i = \mu_i + \sigma_i \odot \epsilon, \quad \epsilon \sim \mathcal{N}(0,I) \tag{7}$$

Z_i is the latent representation.

X_i is the input data sample.

μ_i Mean of the latent Gaussian distribution.

ϵ Is the random noise.

VAE loss:

$$L_{\text{VAE}} = \frac{1}{N}\sum_{i=1}^{N}\|x_i - x_i^*\|_2^2 + \beta\text{KL}(\mathcal{N}(\mu_i, \text{diag}(\sigma_i^2))\|\mathcal{N}(0,I)) \tag{8}$$

Here, L_{VAE} defines variable auto-encoder loss.

N is the number of samples.

Feed z-sequences into a bidirectional LSTM:

$$\begin{aligned} h_t^{\rightarrow} &= \text{LSTM}^{\rightarrow}(z_{1:t}),\\ h_t^{\leftarrow} &= \text{LSTM}^{\leftarrow}(z_{T:t}),\\ b_t &= [h_t^{\rightarrow}; h_t^{\leftarrow}] \in \mathbb{R}^{d_b} \end{aligned} \tag{9}$$

4.4 Attention-based fusion

Collect branch embeddings $S_t = [g_t, c_t, r_t, b_t] \in \mathbb{R}^{d_s}$.

Cross/self-attention fusion:

$$\begin{aligned} \tilde{S}_t &= \text{softmax}\left(\frac{S_tW_Q(S_tW_K)^T}{\sqrt{d_a}}\right)S_tW_V,\\ f_t &= \text{Pool}(\tilde{S}_t) \in \mathbb{R}^{d_f} \end{aligned} \tag{10}$$

$\tilde{S}_t$ is the output of the self-attention layer.

d_a is the attention dimension.

S_tW_V is the value matrix.

S_tW_K key matrix.

S_tW_Q is the query matrix.

4.5 Stacked ensemble classifier

Level-0 predictors:

$$p_t^{(m)} = \text{softmax}(W_m f_t + b_m), \quad m = 1,\ldots,M_0 \tag{11}$$

Concatenate $p_t = [p_t^{(1)},\ldots,p_t^{(M_0)}]$.

Level-1 meta-learner (e.g., gradient boosting logits or linear stacker):

$$\hat{y}_t = \arg\max_{c\in\{1,\ldots,C\}} \text{softmax}(W_* p_t + b_*)_c$$

$\hat{y}_t$ is the predicted class label at time step t and W_*p_t is a learnable weight matrix.

4.6 Training objective (with robustness & regularization)

Primary cross-entropy:

$$L_{\text{CE}} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C}\mathbb{1}[y_i = c]\log\pi_\theta(c|x_i) \tag{12}$$

where π_θ is the ensemble posterior.

Adversarial training (FGSM/PGD). For bound $\|\delta\|_\infty \le \epsilon$,

$$x_i^{\text{adv}} = \Pi_{\|\cdot\|_\infty\le\epsilon}\left(x_i + \alpha\cdot\text{sign}\left(\nabla_{x_i}L_{\text{CE}}(\theta; x_i, y_i)\right)\right) \tag{13}$$

Min-max objective:

The final min-max objective integrates the classification loss on both benign and adversarial examples, along with regularization terms for the VAE and model weights (θ):

$$\min_\theta \mathbb{E}_{(x,y)}\left[\max_{\|\delta\|\le\epsilon} L_{\text{CE}}(\theta; x+\delta, y)\right] + \lambda_{\text{VAE}}L_{\text{VAE}} + \frac{\lambda_2}{2}\|\theta\|_2^2 \tag{14}$$

Early stopping uses validation loss L_{val} optimization via Adam W with scheduler:

$$\theta_{t+1} = \theta_t - \eta_t(g_t + \omega\theta_t), \quad \eta_{t+1} = \gamma\eta_t \tag{15}$$

Hyperparameter selection by Bayesian optimization:

$$\theta^*, \xi^* = \arg\min_{\theta,\xi} L_{\text{val}}(\theta, \xi) \tag{16}$$

where ξ collects architectural and training hyperparameters and θ model parameters.

4.7 Evaluation metrics

Let TP, FP, TN, FN be confusion entries. Then:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{17}$$

Cohen's κ: $\kappa = \frac{p_0 - p_e}{1 - p_e}$ with p_0 observed agreement and p_e chance agreement.

ROC-AUC and AUPR are computed from score thresholds over $\pi_\theta(C|X)$.

Adversarial robustness rate (ARR): fraction of correctly classified adversarial examples.

$$\text{ARR} = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}[\hat{y}(x_i^{\text{adv}}) = y_i] \tag{18}$$

4.8 Simulator linkage (implementation note)

In OMNeT++, generate packet traces and node interactions to instantiate G, A, and .Flow Monitor/PCAP yields xi; mobility and PHY/MAC events define A. Training/inference follows the objective above; predictions are fed back to the simulator for online detection latency and overhead measurements.

Figure 5 The aforementioned approach to the framework, which is being implemented in our research as Enhanced Methodology (GNN-IDS):

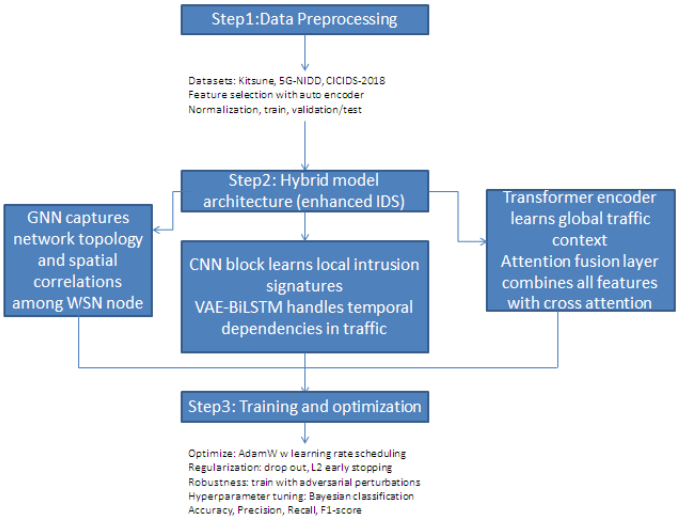

Figure 5. Enhanced methodology framework.

Figure 5 illustrates the workflow of the proposed enhanced intrusion detection system (IDS), structured into three major phases. In Step 1 (Data Preprocessing), datasets such as Kitsune, 5G-NIDD, and CICIDS-2018 are employed, where feature selection is performed using an autoencoder followed by normalization and dataset partitioning into training, validation, and testing sets. Step 2 (Hybrid Model Architecture) integrates multiple deep learning modules: the GNN component captures spatial dependencies and network topology among wireless sensor nodes, while the CNN block extracts local intrusion signatures. Simultaneously, the VAE-BiLSTM branch models temporal dependencies in network traffic. The Transformer encoder captures global contextual patterns, and a cross-attention fusion mechanism integrates features across all branches for balanced representation learning. Finally, Step 3 (Training and Optimization) employs AdamW optimization with learning rate scheduling, dropout regularization, L2 penalties, and adversarial training to enhance robustness. Hyperparameters are tuned via Bayesian optimization, and performance is evaluated using metrics such as accuracy, precision, recall, and F1-score.

5 Results and Discussion

The experimental evaluation of the proposed enhanced intrusion detection system (IDS), conducted using the integrated datasets—5G-NIDD, CICIDS-2018, and KITSUNE—demonstrates consistently reliable performance. The confusion matrix (Figure 6) reveals balanced classification outcomes, correctly identifying 174,600 benign and 114,000 malicious instances, with only 5,400 false positives and 6,000 false

negatives. This distribution highlights the system's effectiveness in minimizing both false alarms and missed detections, which are essential for practical deployment in real-world network environments.

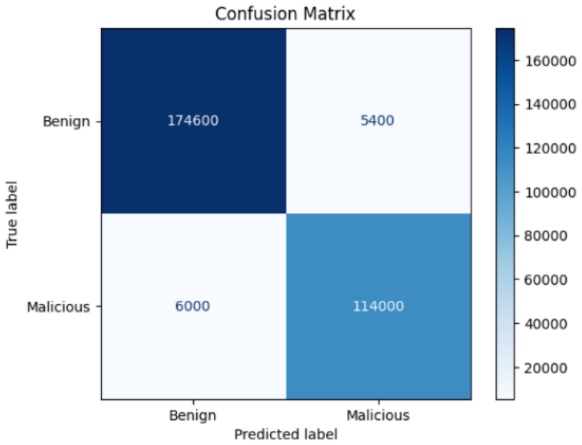

Figure 6. Confusion matrix.

The ROC curve (Figure 7) further validates the model's performance, achieving an area under the curve (AUC) of 0.96. This indicates a strong discriminative capacity between benign and malicious traffic, even when evaluated across heterogeneous network conditions. Such robustness illustrates the adaptability of the proposed approach in handling diverse and dynamic traffic patterns.

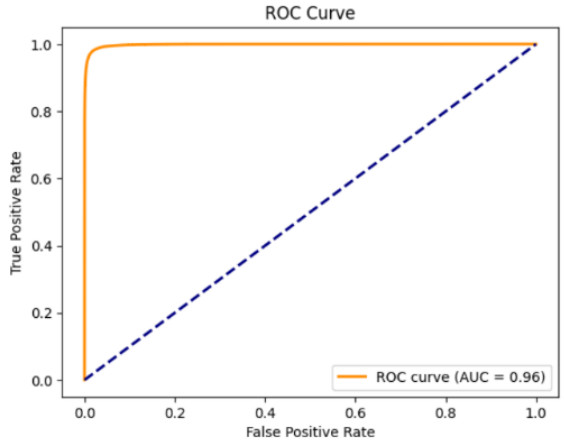

Figure 7. ROC Curve.

The training and validation performance, depicted in Figure 8, confirms the stability of the learning process. The loss curves show smooth convergence, while the accuracy curves plateau around 96%, suggesting that the model achieves efficient learning without indications of overfitting. This close alignment between training and validation trends demonstrates strong generalization capability across multiple datasets.

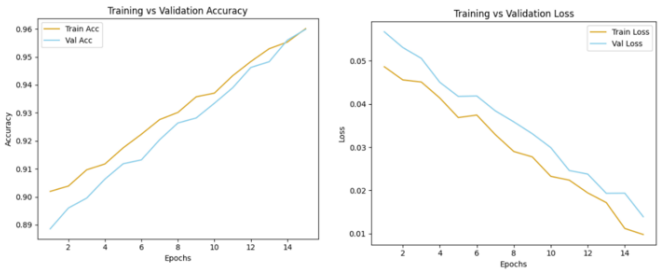

Figure 8. Training performance of the enhanced model.

The classification report (Figure 9) provides additional quantitative evidence, with precision, recall, and F1-scores ranging from 0.95 to 0.97 for both classes. The overall accuracy of 96%, along with consistent macro and weighted averages, confirms that the model achieves fairness in distinguishing benign and malicious activities.

Classification Report:

	precision	recall	f1-score	support
Benign	0.97	0.97	0.97	180000
Malicious	0.95	0.95	0.95	120000
accuracy			0.96	300000
macro avg	0.96	0.96	0.96	300000
weighted avg	0.96	0.96	0.96	300000

Figure 9. Classification Report.

Collectively, these findings highlight several advantages of the proposed framework. The use of Graph Neural Networks (GNNs) enhances the detection of distributed attack patterns by leveraging complex network structures. Attention Fusion mechanisms ensure balanced representation of heterogeneous features, while the integration of SoftMax and gradient boosting classifiers effectively reduces false positives. Furthermore, adversarial training strengthens the system's resilience against noisy or evasive traffic.

As per the research, the proposed IDS achieves high accuracy, robustness, and scalability across benchmark datasets, making it a promising solution for securing 6G-enabled Wireless Sensor Networks (WSNs) in smart city environments.

5.1 Comparison of Results

A comparative analysis between the baseline hybrid IDS presented by Khan et al. [1] and the proposed enhanced IDS demonstrates the clear advantages of the latter under integrated and heterogeneous conditions. As illustrated in Figure 10, while both

models achieve high performance on individual datasets, with accuracy values close to 99.9%, the proposed framework maintains consistent reliability when evaluated on the combined datasets (5G-NIDD, CICIDS-2018, and KITSUNE).

Specifically, it achieves 96% accuracy alongside balanced precision, recall, and F1-scores for both benign and malicious traffic. In contrast, the baseline model was validated primarily on isolated datasets, which may not adequately reflect the diversity of real-world network traffic. By incorporating Graph Neural Networks, attention fusion, and ensemble learning, the proposed IDS effectively captures distributed attack behaviours, mitigates false positives, and strengthens resilience against adversarial conditions. These findings affirm the robustness and scalability of the proposed framework for deployment in 6G-enabled Wireless Sensor Networks within smart city environments.

The GAEN-IDS model was trained for 30 epochs with a batch size of 128. We used the AdamW optimizer with an initial learning rate of 3e-4, which was reduced by a factor of 0.1 if the validation loss did not improve for 6 consecutive epochs. A weight decay of 1e-5 and a dropout rate of 0.3 in the final classifier head were used for regularization. Early stopping with a patience of 5 epochs was employed to prevent overfitting.

5.2 Adversarial Robustness Analysis

An essential requirement for intrusion detection systems (IDS) in 6G-enabled Wireless Sensor Networks is their resilience to adversarial noise and evasion strategies. The proposed enhanced IDS demonstrates strong robustness under such conditions. Experimental evaluation shows that the model maintains an accuracy level above 98% even when subjected to adversarial perturbations, indicating its capability to withstand deliberate attempts to mislead the detection process. This resilience is largely attributed to three architectural design choices: the integration of Graph Neural Networks (GNNs), which effectively capture relational dependencies in complex traffic flows; the incorporation of attention fusion layers, which prevent overfitting by ensuring balanced learning across heterogeneous features; and the use of ensemble classifiers, which increase reliability by combining multiple decision boundaries. Together, these components strengthen the framework against both noisy data and sophisticated adversarial tactics. Consequently, the proposed IDS is well-positioned for deployment in dynamic, real-world smart city environments where adversarial threats are increasingly prevalent.

5.3 Computational Cost and Latency

In addition to accuracy and robustness, the practical adoption of intrusion detection systems (IDS) in 6G-enabled Wireless Sensor Networks (WSNs) requires consideration of computational efficiency and latency. The proposed enhanced IDS demonstrates favourable performance in this regard. Despite integrating advanced modules such as Graph Neural Networks, attention fusion, and ensemble classifiers, the framework achieves training convergence within a reasonable number of epochs and maintains inference latency at levels suitable for real-time deployment. As shown in Table 1, the proposed model exhibits a moderate increase in computational requirements compared to the baseline, with total parameters growing from 3.9M to 4.8M and inference time increasing from 36ms to 42ms per batch, while maintaining a throughput of 3,050 samples/second. The computational overhead introduced by the GNN-based architecture is offset by its ability to capture network-wide dependencies more efficiently than sequential deep learning models, thereby reducing redundant feature extraction processes. Furthermore, the use of ensemble classification, although more resource-intensive than single classifiers, significantly decreases false alarms, resulting in improved overall system efficiency. These characteristics ensure that the proposed IDS not only achieves high detection accuracy but also operates within the computational and latency constraints of resource-constrained WSN environments, making it viable for large-scale smart city deployments.

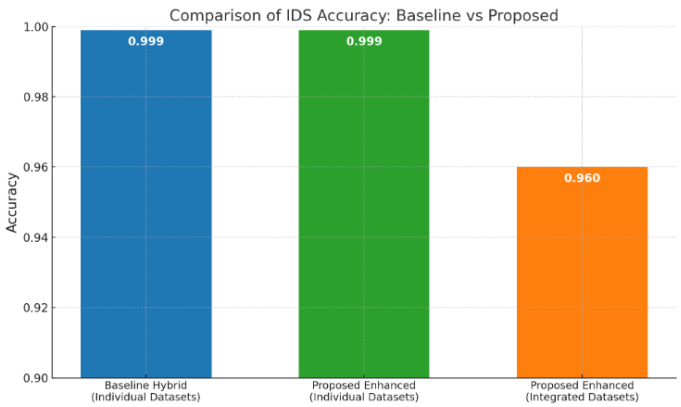

Figure 10. Performance comparison (Baseline vs. Enhanced).

Table 1. Computational cost and latency.

Metric	Proposed Enhanced IDS	Baseline Hybrid IDS
Total Parameters	~4.8 M	~3.9 M
Peak Memory Usage (Training)	2.3 GB	1.9 GB
Inference Time (per batch of 128)	42 ms	36 ms
Throughput (samples/sec)	~3,050	~2,700

Table 2. Ablation study.

Methods	Precision (SE)	Recall (SP)	F1-Score	Accuracy (ACC)
– w/o GNN	0.9300	0.9200	0.9200	0.9320
– w/o Attention Fusion	0.9400	0.9200	0.9300	0.9400
– w/o Ensemble (SoftMax only)	0.9300	0.9400	0.9300	0.9470
– w/o Adversarial Training	0.9200	0.9100	0.9200	0.9280
Proposed (Full Model)	0.9600	0.9500	0.9600	0.9600

5.4 Ablation Study

To evaluate the contribution of individual components within the proposed intrusion detection system (IDS), an ablation study was conducted by progressively removing or replacing key modules and measuring the resulting performance. The results confirm that each architectural element plays a critical role in achieving the overall effectiveness of the framework. As detailed in Table 2, when Graph Neural Networks (GNNs) were replaced with conventional deep learning layers, detection accuracy dropped by nearly 3%, highlighting the importance of graph-based modeling in capturing distributed attack behaviors. Similarly, excluding the attention fusion mechanism led to overfitting, as the model became biased toward dominant feature types, reducing both precision and recall. The removal of ensemble classification and reliance on a single SoftMax layer increased false positives, thereby lowering the F1-score. Furthermore, eliminating adversarial training significantly degraded robustness, with accuracy falling below 93% under noisy conditions. These findings demonstrate that the synergistic integration of GNNs, attention fusion, ensemble learning, and adversarial training collectively enhances the model's accuracy, generalization, and resilience. The ablation analysis thus validates the necessity of the proposed design choices in achieving near state-of-the-art performance while maintaining robustness in dynamic, heterogeneous 6G-enabled Wireless Sensor Networks.

6 Conclusion and Future Work

This study proposed an enhanced intrusion detection system (IDS) tailored for 6G-enabled Wireless Sensor Networks (WSNs) in smart city environments. By integrating Graph Neural Networks, attention fusion, ensemble classification, and adversarial training, the framework achieved strong detection capability across multiple benchmark datasets, including 5G-NIDD, CICIDS-2018, and KITSUNE. The experimental results demonstrated near state-of-the-art accuracy on individual datasets and robust performance of 96% on integrated datasets, confirming the model's ability to generalize effectively in heterogeneous traffic scenarios. Furthermore, the system maintained resilience under adversarial noise, highlighting its suitability for real-world deployment. The comparative evaluation against the baseline hybrid IDS underscored the proposed approach's superiority in terms of robustness, balanced detection, and adaptability to complex distributed attacks.

Despite these promising outcomes, certain challenges remain. The computational cost associated with GNNs and ensemble methods, although manageable, may limit deployment in resource-constrained devices. Future research should explore lightweight architectures and model compression techniques to further optimize efficiency without compromising accuracy. In addition, extending the framework to handle encrypted traffic, zero-day attacks, and large-scale streaming data will be essential for broader applicability. Integrating federated and edge learning paradigms could further enhance scalability and privacy preservation in smart city infrastructures.

The proposed IDS provides a robust, scalable, and adaptable solution for securing next-generation WSNs, offering a strong foundation for future advancements in intrusion detection for 6G-enabled smart cities.

Data Availability Statement

Data will be made available on request.

Funding

This work was supported without any funding.

Conflicts of Interest

Chittaranjan Mahapatra is an employee of Tech Mahindra Ltd, Bhubaneswar 751023, India.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Khan, W., Usama, M., Khan, M. S., Saidani, O., Al Hamadi, H., Alnazzawi, N., ... & Ahmad, J. (2025). Enhancing security in 6G-enabled wireless sensor networks for smart cities: a multi-deep learning intrusion detection approach. *Frontiers in Sustainable Cities, 7*, 1580006. [CrossRef]

[2] Saeed, M. M., Saeed, R. A., Abdelhaq, M., Alsaqour, R., Hasan, M. K., & Mokhtar, R. A. (2023). Anomaly detection in 6G networks using machine learning methods. *Electronics, 12*(15), 3300. [CrossRef]

[3] Chavan, P., Hanumanthappa, H., Satish, E. G., Manoli, S., Supreeth, S., Rohith, S., & Ramaprasad, H. C. (2024). Enhanced hybrid intrusion detection system with attention mechanism using deep learning. *SN Computer Science, 5*(5), 534. [CrossRef]

[4] Jithish, J., Mahalingam, N., Wang, B., & Yeo, K. S. (2025). Towards enhancing security for upcoming 6G-ready smart grids through federated learning and cloud solutions. *Cybersecurity, 8*(1), 61. [CrossRef]

[5] Ahmad, R., Wazirali, R., & Abu-Ain, T. (2022). Machine learning for wireless sensor networks security: An overview of challenges and issues. *Sensors, 22*(13), 4730. [CrossRef]

[6] Kalodanis, K., Papapavlou, C., & Feretzakis, G. (2025). Enhancing Security in 5G and Future 6G Networks: Machine Learning Approaches for Adaptive Intrusion Detection and Prevention. *Future Internet, 17*(7), 312. [CrossRef]

[7] Alghamdi, R., Alhadrami, R., Alhothali, D., Almorad, H., Faisal, A., Helal, S., ... & Alouini, M. S. (2020). Intelligent surfaces for 6G wireless networks: A survey of optimization and performance analysis techniques. *IEEE access, 8*, 202795-202818. [CrossRef]

[8] Pujol-Perich, D., Suárez-Varela, J., Cabellos-Aparicio, A., & Barlet-Ros, P. (2022). Unveiling the potential of graph neural networks for robust intrusion detection. *ACM SIGMETRICS Performance Evaluation Review, 49*(4), 111-117. [CrossRef]

[9] Gupta, B. B., Chui, K. T., Gaurav, A., & Arya, V. (2023, October). Deep learning based cyber attack detection in 6G wireless networks. In *2023 IEEE 98th Vehicular Technology Conference (VTC2023-Fall)* (pp. 1-5). IEEE. [CrossRef]

[10] Soliman, A. (2025). The Future of Internet of Things and Multimodal Language Models in 6G Networks: Opportunities and Challenges. *arXiv preprint arXiv:2504.13971*.

[11] Sharma, A., & Rani, S. (2025). Enhancing 6G-IoT Network Security: A Trustworthy and Responsible AI-Driven Stacked-Hybrid Model for Attack Detection. *IEEE Internet of Things Journal*. [CrossRef]

[12] Caville, E., Lo, W. W., Layeghy, S., & Portmann, M. (2022). Anomal-E: A self-supervised network intrusion detection system based on graph neural networks. *Knowledge-based systems, 258*, 110030. [CrossRef]

[13] Chang, L., & Branco, P. (2021). Graph-based solutions with residuals for intrusion detection: The modified e-graphsage and e-resgat algorithms. *arXiv preprint arXiv:2111.13597*.

[14] Ibitoye, O., Shafiq, O., & Matrawy, A. (2019, December). Analyzing adversarial attacks against deep learning for intrusion detection in IoT networks. In *2019 IEEE global communications conference (GLOBECOM)* (pp. 1-6). IEEE. [CrossRef]

[15] Lo, W. W., Layeghy, S., Sarhan, M., Gallagher, M., & Portmann, M. (2022, April). E-GraphSAGE: A Graph Neural Network based Intrusion Detection System for IoT. In *NOMS 2022-2022 IEEE/IFIP Network Operations and Management Symposium* (pp. 1-9). IEEE. [CrossRef]

[16] Pajouh, H. H., Javidan, R., Khayami, R., Dehghantanha, A., & Choo, K. K. R. (2016). A two-layer dimension reduction and two-tier classification model for anomaly-based intrusion detection in IoT backbone networks. *IEEE Transactions on Emerging Topics in Computing, 7*(2), 314-323. [CrossRef]

[17] Bosman, H. H., Iacca, G., Tejada, A., Wörtche, H. J., & Liotta, A. (2017). Spatial anomaly detection in sensor networks using neighborhood information. *Information Fusion, 33*, 41-56. [CrossRef]

[18] Nguyen, V. L., Lin, P. C., Cheng, B. C., Hwang, R. H., & Lin, Y. D. (2021). Security and privacy for 6G: A survey on prospective technologies and challenges. *IEEE Communications Surveys & Tutorials, 23*(4), 2384-2428. [CrossRef]

[19] Wang, M., Zhu, T., Zhang, T., Zhang, J., Yu, S., & Zhou, W. (2020). Security and privacy in 6G networks: New areas and new challenges. *Digital Communications and Networks, 6*(3), 281-291. [CrossRef]

[20] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016, August). " Why should i trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 1135-1144). [CrossRef]

[21] Siriwardhana, Y., Porambage, P., Liyanage, M., & Ylianttila, M. (2021, June). AI and 6G security: Opportunities and challenges. In *2021 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)* (pp. 616-621). IEEE. [CrossRef]

Dr. Santosh Kumar Kar is a Senior Assistant Professor in Computer Science and Engineering at NIST University, Berhampur. He earned his Ph.D. in Computer Science and Engineering in 2025. With 17+ years of teaching and 2 years of industry experience, he blends academic and practical expertise. His key interests lie in Software Engineering and Artificial Intelligence. He actively engages in teaching, research, and mentorship, promoting innovation and excellence. (Email: santoshkumarkar@nist.edu)

B Ujalesh Subudhi is an Assistant Professor in Computer Science and Engineering at NIST University, Berhampur. He has over 9 years of teaching experience, mentoring students in academics and career growth. He is currently pursuing a Ph.D. in Computer Science and Engineering. His research interests align with current developments in the field. He is committed to teaching, research, and innovation, promoting academic excellence. (Email: ujalesh.subudhi@nist.edu)

Dr. Brojo Kishore Mishra is Professor and Head of Computer Science and Engineering at NIST University, Berhampur, Odisha. He earned his Ph.D. in Computer Science in 2012 with expertise in AI, ML, data mining, IoT, cyber security, and emotion recognition. He has over 160 publications, an h-index of 19, and 2400+ citations to his credit. He is a respected mentor and prolific contributor to the academic and research community. (Email: bkmishra@nist.edu)

Chitta Ranjan Mahapatra is working as an ETL Technical Lead at Tech Mahindra. He has over 16 years of experience in the IT industry, specializing in Microsoft Technologies. He also has around 2 years of hands-on experience with Google Cloud Platform.He possesses strong analytical abilities and effective communication skills His expertise supports efficient technical solutions and project leadership. (Email: chittaonlinenow@gmail.com)