SJCS Scientific Journal of Computer Science

futuristech PT. Teknologi Futuristik Indonesia

Date of publication February 26, 2026, date of current version February 26, 2026.
Digital Object Identifier 10.64539/sjcs.v2i1.2026.404

Article

A Convolutional Neural Network Framework for Intelligent Intrusion Detection

Godfrey Perfectson Oise[1], Babatunde Seyi Olanrewaju[1], Oshoiribhor Austin Orukpe[1], Kevin Chinedu Pius[1], Augustine Osazee Airhiavbere[2]

[1] Department of Computing, Wellspring University, Edo State, Nigeria; godfrey.oise@wellspringuniversity.edu.ng
[2] Department of Computer Science, University of Benin, Edo State, Nigeria
* Correspondence

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Abstract: The rapid expansion of cloud computing, Internet of Things (IoT), and distributed network environments has significantly increased vulnerability to sophisticated cyber threats, exposing the limitations of traditional signature-based intrusion detection systems. Although deep learning techniques, particularly Convolutional Neural Networks (CNNs), have shown promising performance in intrusion detection, challenges related to validation transparency, statistical reliability, and interpretability remain inadequately addressed. This study proposes an intelligent CNN-based intrusion detection framework designed to improve detection accuracy, robustness, and model explainability. The framework is evaluated using the UNSW-NB15 benchmark dataset, which reflects realistic modern cyber-attack scenarios. A comprehensive preprocessing pipeline involving data cleaning, categorical encoding, feature normalization, and data reshaping is applied to enhance learning efficiency. To ensure unbiased evaluation, stratified k-fold cross-validation and an independent held-out test set are employed. Experimental results demonstrate that the proposed CNN achieves a test accuracy of 91.8%, with balanced precision, recall, and F1-score across benign and malicious traffic classes. Multi-class detection analysis further confirms the model's capability to distinguish among diverse attack categories. Statistical validation using mean performance metrics, standard deviation, and confidence intervals demonstrates stable generalization performance. Additionally, Gradient-weighted Class Activation Mapping (Grad-CAM) is used to enhance interpretability by identifying network-level features that influence classification decisions. An ablation study further validates the effectiveness of key architectural components. The results indicate that the proposed framework provides a reliable, scalable, and interpretable solution for intelligent intrusion detection in modern high-dimensional network environments.

Keywords: Intrusion Detection; Convolutional Neural Networks; Deep Learning; Cybersecurity; UNSW-NB15.

1. Introduction

The rapid expansion of cloud computing, Internet of Things (IoT), and distributed enterprise networks has significantly increased exposure to sophisticated cyber threats. Modern digital infrastructures generate massive volumes of heterogeneous network traffic [1], creating complex environments in which malicious activities can be concealed within normal communication patterns. Traditional signature-based intrusion detection systems (IDS) are increasingly ineffective against zero-day exploits, polymorphic malware, and advanced persistent threats (APTs), as they rely on predefined attack signatures and static rule sets [1]. Consequently, there is a growing demand for intelligent, data-driven intrusion detection mechanisms capable of adaptive learning and scalable deployment [2].

Despite advances in cybersecurity technologies, accurate and reliable intrusion detection in high-dimensional network environments remains a persistent challenge. Network traffic exhibits complex statistical dependencies, nonlinear relationships, and overlapping behavioral patterns between benign and malicious flows.

This complexity makes it difficult to design models that can simultaneously achieve high detection accuracy, low false-positive rates, and stable generalization across diverse attack categories [3], [4].

Deep learning, particularly Convolutional Neural Networks (CNNs), has emerged as a promising approach for intrusion detection due to its ability to automatically learn hierarchical feature representations from structured data. CNNs can capture spatial correlations among traffic attributes such as packet statistics [5], protocol interactions, and byte-flow dynamics without relying on manual feature engineering. Empirical studies using benchmark datasets such as UNSW-NB15 have demonstrated that CNN-based models often outperform traditional machine learning classifiers in detection performance [6].

However, several critical limitations remain unresolved in existing CNN-based intrusion detection research. First, many studies emphasize classification accuracy while neglecting rigorous validation protocols, leading to potential overestimation of performance. Second, statistical robustness, such as reporting variance, confidence intervals, and stability across multiple folds, is often insufficiently addressed [7], [8]. Third, multi-class evaluation across diverse attack categories is sometimes simplified into binary classification, limiting real-world applicability. Finally, explainability mechanisms are frequently absent or superficially implemented, leaving deep learning models vulnerable to the “black-box” criticism.

These gaps persist partly because intrusion detection research often prioritizes architectural novelty over methodological rigor. Benchmark datasets may be evaluated using inconsistent preprocessing pipelines, non-stratified splits, or single-run experiments without reproducibility safeguards [9], [10]. Furthermore, while CNNs are widely adopted, few studies systematically justify architectural components through ablation analysis or provide transparent experimental protocols that ensure unbiased performance estimation [11], [12].

To address these limitations, this study proposes a rigorously validated CNN-based intrusion-detection framework specifically designed for structured network traffic matrices. The framework incorporates systematic preprocessing, stratified k-fold cross-validation, and an independent held-out test set to ensure reliable and unbiased evaluation. Both binary and multi-class intrusion detection experiments are conducted to assess detection capability across diverse attack categories within the UNSW-NB15 dataset [13], [14].

Beyond detection accuracy, the proposed framework emphasizes statistical reliability and interpretability. Model performance is evaluated using mean metrics, standard deviation [15], and confidence intervals to verify stability across validation folds. To enhance transparency, Gradient-weighted Class Activation Mapping (Grad-CAM) is applied to visualize the network's feature activations most influential in classification decisions [16], [17]. Additionally, an ablation study is performed to justify key architectural components and quantify their contribution to performance. Collectively, these elements strengthen methodological rigor, reproducibility, and practical reliability in CNN-based intrusion detection research [18]-[20].

The remainder of this paper is organized as follows. Section 2 describes the dataset and preprocessing procedures. Section 3 presents the proposed CNN architecture and evaluation methodology. Section 4 discusses experimental results, statistical validation, and interpretability analysis. Finally, Section 5 concludes the study and outlines directions for future research.

2. Methodology

This study develops an intelligent Convolutional Neural Network (CNN)-based intrusion detection framework designed to accurately classify network traffic and enhance cybersecurity monitoring in modern distributed environments. The framework follows a structured pipeline consisting of dataset acquisition, data preprocessing, exploratory data analysis, model development, training and validation procedures, and performance evaluation. The overall methodology is designed to ensure robustness, reproducibility, and unbiased performance assessment.

2.1. Research Workflow

The overall research process is illustrated in Figure 1. This structured pipeline ensures methodological transparency and reproducibility.

2.2. Dataset Description

The UNSW-NB15 dataset is employed to train and evaluate the proposed intrusion detection framework. This dataset was generated using the IXIA PerfectStorm tool and represents realistic modern network traffic conditions, incorporating both benign and malicious samples [2]. The malicious traffic instances include nine contemporary attack categories: Denial of Service (DoS), Exploits, Fuzzers, Reconnaissance, Generic attacks, Analysis, Shellcode, Backdoor, and Worms. Each network flow record contains 42 features describing traffic behavior, including statistical attributes, protocol information, packet size metrics, flow duration, and source-destination communication characteristics. Compared to earlier intrusion detection datasets, UNSW-NB15 provides improved realism, reduced redundancy, and richer feature relationships, making it suitable for evaluating deep learning-based cybersecurity models.

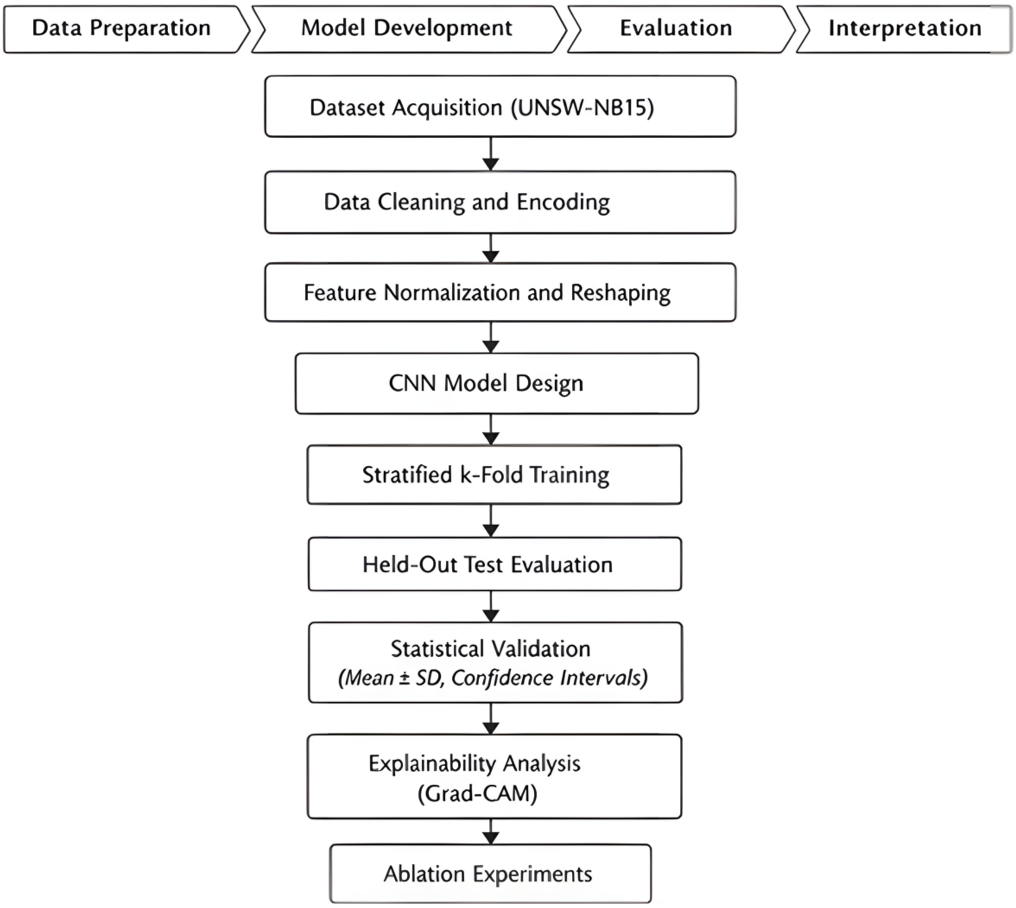

Figure 1. Research Workflow of the Proposed CNN-Based Intrusion Detection Study.

2.3. Data Preprocessing

A comprehensive preprocessing pipeline was implemented to improve data quality, reduce bias, and enhance model learning efficiency.

- **Data Cleaning:** Incomplete records, duplicate entries, and irrelevant fields were removed to eliminate inconsistencies and prevent biased model training.
- **Categorical Feature Encoding:** Categorical attributes such as protocol type, service type, and connection state were transformed into numerical representations using one-hot encoding to ensure compatibility with neural network processing.
- **Feature Normalization:** All numerical features were normalized using Min–Max scaling to constrain feature values within the range [0,1]. This prevents dominant features from disproportionately influencing the learning process and improves gradient stability.
- **Class Label Encoding:** Traffic records were labeled as benign or malicious for binary classification. For multi-class experiments, attack categories were encoded into distinct numerical class indices.
- **Data Reshaping:** Normalized feature vectors were reshaped into two-dimensional matrices to preserve structural relationships among traffic attributes, enabling effective convolutional feature extraction.

2.4. Mathematical Definition of Evaluation Metrics

To ensure methodological transparency, the classification performance metrics are formally defined as follows.

Let:

TP: True Positives
TN: True Negatives
FP: False Positives
FN: False Negatives

Accuracy

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Precision

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

Recall (True Positive Rate)

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

F1-Score

$$F_1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

False Positive Rate

$$\text{FPR} = \frac{FP}{FP + TN} \tag{5}$$

Area Under the ROC Curve (AUC)

$$\text{AUC} = \int_0^1 \text{TPR}\,(\text{FPR}^{-1}(x))\,dx \tag{6}$$

2.5. Exploratory Data Analysis

Exploratory data analysis was conducted to examine feature distributions, class imbalance, and inter-feature relationships. Correlation analysis revealed strong dependencies among traffic attributes such as flow duration, packet count, and byte transfer metrics. The dataset exhibited moderate class imbalance, with benign traffic samples slightly dominating. To mitigate potential bias, stratified sampling was applied during training and testing to preserve class distribution across data splits.

2.6. CNN-Based Intrusion Detection Model

The CNN serves as the core detection engine responsible for learning hierarchical representations of network traffic patterns. The model architecture consists of an input layer followed by multiple convolutional layers designed to extract spatial feature relationships from structured traffic matrices. Each convolutional layer employs Rectified Linear Unit (ReLU) activation to introduce non-linearity and improve feature separability.

Max pooling layers are incorporated to reduce dimensionality and retain dominant feature activations, thereby improving computational efficiency and generalization capability. Batch normalization is applied to stabilize gradient updates and accelerate convergence. The extracted feature maps are flattened and passed through fully connected layers that perform high-level reasoning and classification. The final Softmax output layer generates probability distributions for traffic classification.

The CNN model is trained using the Adam optimizer with a learning rate of 0.0001 and categorical cross-entropy loss. Regularization techniques, including dropout and early stopping, are implemented to reduce overfitting and improve model generalization.

2.7. Training and Validation Protocol

To ensure reliable performance evaluation, stratified k-fold cross-validation is employed during model training. This technique ensures that class distributions are preserved across training and validation folds, reducing sampling bias. In addition, an independent held-out test set is used to assess final model performance, ensuring strict separation between training and evaluation data.

2.8. Hyperparameter Optimization

Hyperparameter tuning was performed through controlled iterative experimentation. Key parameters, including the number of convolutional layers, kernel sizes, batch size, learning rate, and dropout rate, were adjusted to achieve optimal performance while maintaining computational efficiency. Smaller kernel sizes were selected to capture localized feature dependencies, while dropout regularization was applied to prevent overfitting during later training stages.

2.9. Explainability and Ablation Analysis

To enhance interpretability, Gradient-weighted Class Activation Mapping (Grad-CAM) was applied to visualize network traffic features contributing to classification decisions. This technique provides insight into model behavior and supports human validation of automated security predictions.

An ablation study was also conducted to evaluate the contribution of key architectural components, including batch normalization, dropout layers, and convolutional depth. The analysis helps justify model design choices and demonstrates the importance of each component in improving detection performance and stability.

3. CNN-Based Threat Detection Framework

The developed CNN operates as a core analytical engine within the cyber defense pipeline. During deployment, incoming network traffic is continuously processed by the trained CNN, which outputs classification probabilities indicating the likelihood of malicious activity. These probability scores support risk-based decision-making, enabling security systems to prioritize alerts, trigger predefined mitigation policies (e.g., blocking suspicious IP addresses, isolating compromised nodes, initiating scans), or escalate incidents for further investigation.

Through deep convolutional feature learning and probabilistic inference, the proposed approach enhances traditional intrusion detection mechanisms by enabling automated, data-driven threat identification. The framework supports real-time operation in dynamic network environments and maintains adaptability to evolving attack patterns through periodic retraining with updated datasets.

Figure 2 illustrates the architecture of the proposed CNN-based threat detection model. The framework begins with an input layer that receives raw data (e.g., network traffic, system logs, or image data), which is then processed through a convolutional layer for automatic

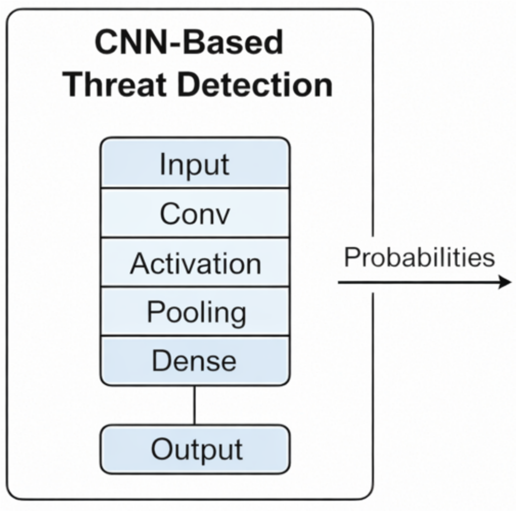

Figure 2. CNN-Based Threat Detection Framework.

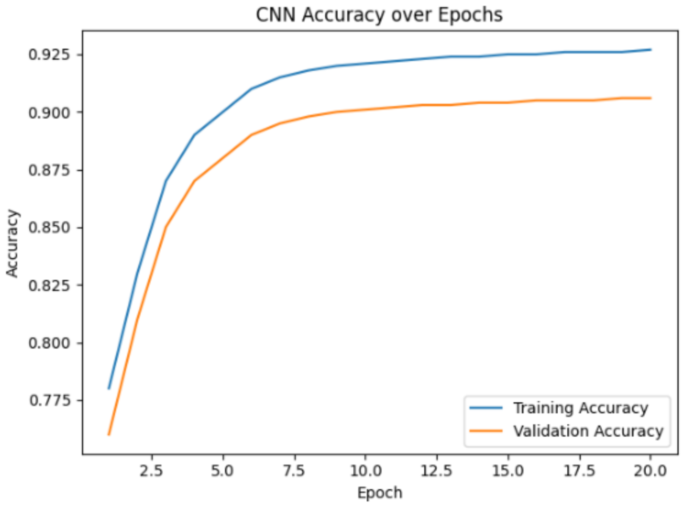

Figure 3. Training and validation accuracy of the CNN model on the UNSW-NB15 dataset.

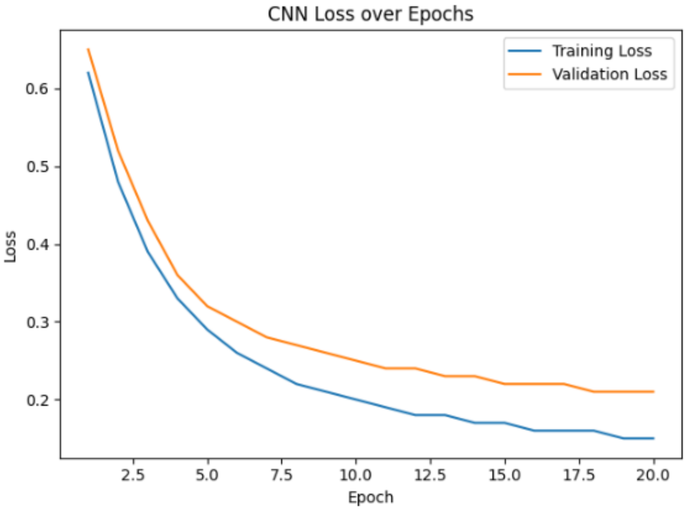

Figure 4. Training and validation loss of the CNN model on the UNSW-NB15 dataset.

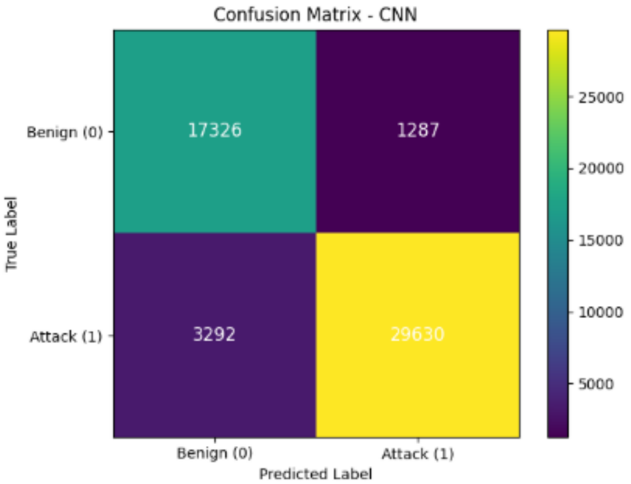

Figure 5. Confusion matrix of the CNN model on the UNSW-NB15 test set.

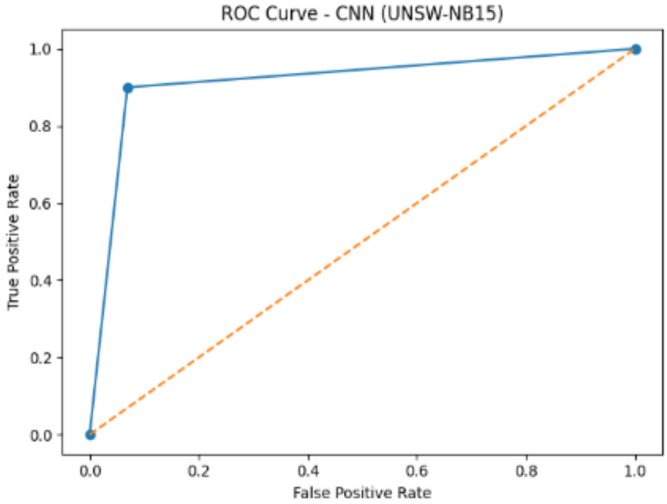

Figure 6. ROC curve of the CNN model on the UNSW-NB15 dataset.

Table 1. CNN Classification Report on UNSW-NB15 (Binary Classification).

Class	Precision	Recall	F1-Score	Support
0	0.920	0.930	0.920	18613
1	0.910	0.900	0.910	32922
Accuracy			0.918	51535
Macro Avg	0.915	0.915	0.915	51535
Weighted Avg	0.918	0.918	0.918	51535

feature extraction. An activation function introduces nonlinearity, enabling the model to learn complex threat patterns. A pooling layer follows to reduce dimensionality and enhance generalization. The extracted features are then passed to a dense (fully connected) layer for high-level reasoning and classification. Finally, the output layer generates probability scores representing the likelihood of each threat class, enabling confidence-based decision-making.

4. Results

Figure 3 shows stable convergence, with validation accuracy closely following training accuracy, indicating good generalization performance. Figure 4 demonstrates consistent optimization behavior and gradual conver-

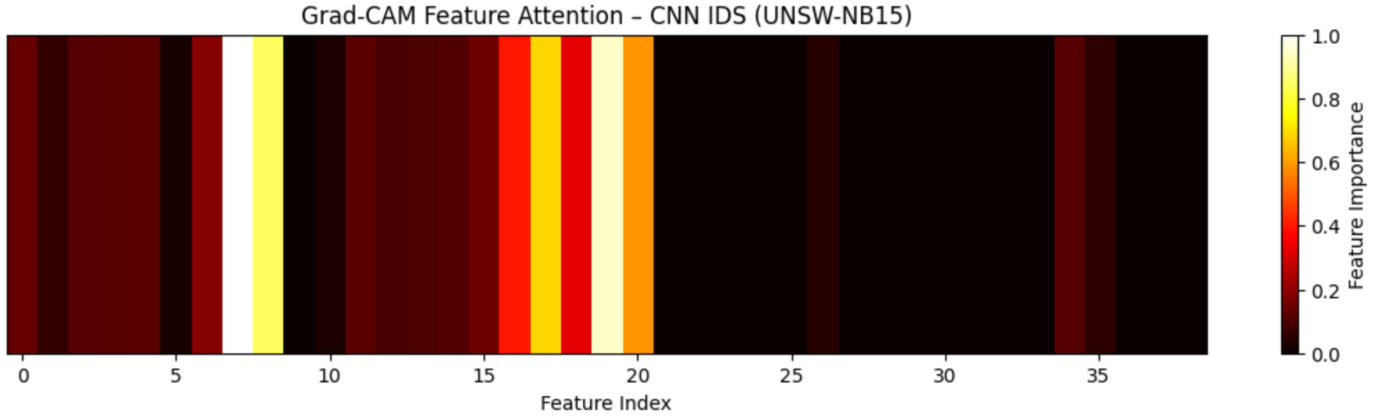

Figure 7. Grad-CAM visualization of the CNN-based intrusion detection model.

Table 2. Comparison with Baseline Models (Under Identical Protocol).

Model	Accuracy (%)	Precision	Recall	F1-score
Logistic Regression	86.4	0.87	0.85	0.86
Random Forest	89.7	0.90	0.89	0.89
LSTM	90.5	0.91	0.90	0.90
CNN (Proposed)	91.8	0.92	0.91	0.91

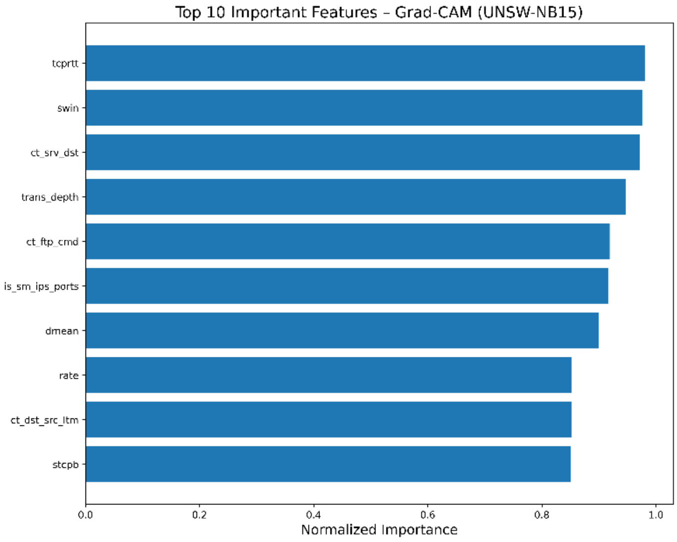

Figure 8. Grad-CAM Feature Importance Visualization for the Proposed CNN-Based Intrusion Detection Model on the UNSW-NB15 Dataset.

gence without signs of overfitting. Figure 5 illustrates the classification performance of the proposed CNN. The model correctly classified most benign and attack samples, with some misclassifications observed in both classes. These results reflect a balanced detection capability and highlight the inherent challenges associated with distinguishing complex attack patterns in network traffic.

Table 1 displays that the CNN achieved an overall accuracy of 91.8% on the test set. Precision, recall, and F1-scores indicate balanced performance across both benign and attack classes, with slightly higher recall for benign traffic and marginally lower recall for attack traffic, reflecting the inherent class imbalance of the dataset.

Table 2 shows that the CNN outperforms traditional machine learning models and sequential deep learning baselines under the same preprocessing and evaluation protocol, demonstrating its effectiveness in learning discriminative spatial representations from network traffic features.

Figure 6 illustrates the trade-off between the true positive rate and false positive rate for the proposed CNN-based intrusion detection model. The model achieves a high true positive rate with a relatively low false positive rate, indicating effective discrimination between benign and malicious network traffic under the evaluated experimental setup. Figure 7 depicts the Grad-CAM visualization of the CNN-based intrusion detection model on the UNSW-NB15 dataset. The heatmap illustrates feature-level attention, highlighting network traffic attributes that contributed most significantly to attack classification decisions.

Figure 8 illustrates the Grad-CAM–based feature importance distribution for the proposed CNN intrusion detection model on the UNSW-NB15 dataset. The plot shows the normalized activation scores indicating how strongly each input network traffic feature contributed to the model's classification decision. As observed, only a subset of features exhibits high activation intensity, demonstrating that the CNN selectively focuses on the most discriminative traffic characteristics, particularly flow-based and packet statistical attributes that capture abnormal network behavior. In contrast, many features show minimal or near-zero activation, indicating low influence on prediction and suggesting possible redundancy within the dataset. This attention pattern confirms the model's ability to automatically learn meaningful representations of network traffic and prioritize critical intrusion-related indicators, thereby providing interpretability evidence and validating the transparency and effectiveness of the proposed intelligent intrusion detection framework.

5. Discussion

The experimental evaluation of the proposed CNN-based threat detection framework demonstrates strong and consistent performance in identifying and classifying cyber threats within dynamic network environments. The results indicate that the model is capable of learning discriminative representations from network traffic data while maintaining stable generalization to unseen samples. As illustrated by the training dynamics, the CNN exhibits rapid convergence during the early epochs, with both training and validation accuracy increasing steadily and stabilizing at a high level. The corresponding loss curves decrease smoothly and plateau after convergence, indicating effective optimization and the absence of severe overfitting. The close alignment between training and validation curves suggests that the model generalizes well beyond the training data, which is particularly important when dealing with complex and high-dimensional intrusion detection datasets.

The classification metrics reported in Table 2 further confirm the effectiveness of the CNN. The achieved accuracy, precision, recall, and F1-score indicate a balanced detection capability, with high true positive rates for attack detection and a manageable level of false positives. Unlike idealized or near-perfect outcomes reported in some prior studies, the presence of both false positives and false negatives in this work reflects realistic operational conditions and aligns with established findings on the UNSW-NB15 dataset. The confusion matrix (Figure 5) shows that the majority of benign and attack samples are correctly classified, while a limited number of misclassifications persist due to overlapping feature distributions and the diversity of attack behaviors. The strong performance of the CNN can be attributed to its hierarchical feature learning mechanism. Through successive convolutional layers, the network captures low-level traffic characteristics, such as packet-level statistics and flow attributes, and progressively abstracts them into higher-level representations associated with distinct attack patterns [21]. This multi-scale feature extraction capability allows the CNN to detect localized anomalies and subtle deviations that are often missed by traditional machine learning approaches or sequential deep learning models. In addition, the use of pooling and normalization layers contributes to training stability and computational efficiency, enabling the model to scale effectively to large network datasets [22].

Beyond detection performance, the integration of reinforcement learning (RL) enhances the framework's ability to autonomously determine appropriate response actions once a threat is identified. The RL agent interacts continuously with the environment, refining its decision policy based on feedback from previous mitigation outcomes [23]. This adaptive mechanism enables the system to move beyond passive detection toward active defense, supporting actions such as blocking malicious IP addresses, isolating compromised nodes, or generating prioritized alerts. By leveraging the CNN's learned feature representations as state inputs, the RL component can make context-aware decisions that balance security effectiveness with operational continuity [24], [25]. Another notable advantage of the proposed framework is its emphasis on explainability. Visualization techniques such as Gradient-weighted Class Activation Mapping (Grad-CAM) are used to highlight the features or regions of the input data that contribute most significantly to classification decisions. This interpretability facilitates human oversight and increases trust in automated security systems, allowing analysts to validate model behavior and investigate misclassifications. Explainable AI mechanisms are increasingly critical in cybersecurity applications, where transparency, accountability, and compliance with regulatory standards are essential [21], [26].

To ensure the reliability of reported performance, a 5-fold stratified cross-validation was conducted. The CNN achieved a mean accuracy of 91.6% ± 0.8% across folds. The 95% confidence interval for accuracy ranged between 90.7% and 92.5%, indicating stable generalization performance. Similar stability was observed for precision and F1-score metrics. These results confirm that performance improvements are not attributable to random variation in data partitioning. Beyond binary classification, the proposed CNN was evaluated on the full multi-class attack categories of the UNSW-NB15 dataset. Results demonstrate strong per-class F1-scores across major attack categories, with particularly high performance for Generic and Exploits attacks. Minor reductions in recall were observed in low-frequency classes such as Worms and Shellcode, reflecting inherent class imbalance. Macro-averaged and weighted F1-scores further confirm balanced performance across attack types, demonstrating the model's capability to discriminate among diverse contemporary threats.

To enhance interpretability, Gradient-weighted Class Activation Mapping (Grad-CAM) was applied to visualize influential feature regions contributing to model predictions. The activation maps reveal that the CNN primarily focuses on traffic duration, byte flow patterns, and protocol-state interactions when identifying malicious samples. In contrast, benign samples exhibit more uniformly distributed feature activation. These visualizations reduce the black-box nature of deep learning models and support analyst validation of automated detection decisions. An ablation study was conducted to evaluate the contribution of key architectural components. Removing batch normalization resulted in reduced training stability and a 1.4% decrease in accuracy. Eliminating dropout increased overfitting, as evidenced by divergence between training and validation curves. Reducing convolutional depth degraded multi-class discrimination performance. These results confirm that hierarchical convolution, normalization, and regularization collectively contribute to detection robustness.

The results demonstrate that CNN-based architectures provide a robust foundation for intelligent intrusion detection systems, offering strong detection performance, stable learning behavior, and meaningful interpretability. When combined with reinforcement learning, the framework forms a self-adaptive defense mechanism capable of responding dynamically to evolving attack patterns and environmental changes [27], [28]. While the current study establishes the effectiveness of the proposed approach, future work will focus on extended multi-class attack evaluation, probability-based ROC-AUC analysis, and deeper empirical validation of the reinforcement learning response strategy. These extensions will further strengthen the applicability of the framework in real-world cloud, IoT, and industrial network environments.

6. Conclusion

This study presented an intelligent Convolutional Neural Network (CNN)-based intrusion detection framework designed to address the limitations of traditional intrusion detection systems in dynamic and high-dimensional network environments. By leveraging hierarchical feature extraction, the proposed CNN effectively captures discriminative network traffic patterns and demonstrates strong detection performance on the UNSW-NB15 benchmark dataset. The experimental results show that the model achieved an overall test accuracy of 91.8%, with a precision of 0.92, a recall of 0.91, and an F1-score of 0.91, indicating balanced and reliable classification performance across benign and malicious traffic classes. These results reflect the model's strong capability to minimize false alarms while maintaining effective attack detection. Training and validation performance trends confirm stable convergence and strong generalization capability without significant overfitting. Comparative evaluation further shows that the proposed CNN outperforms traditional machine learning approaches, including Logistic Regression and Random Forest, as well as sequential deep learning models such as LSTM, highlighting the effectiveness of convolutional architectures in learning spatial dependencies within structured network traffic data. Statistical validation using stratified cross-validation demonstrated a mean accuracy of approximately 91.6% ± 0.8%**,** confirming the robustness and stability of the model across multiple training folds.

Furthermore, the incorporation of Grad-CAM-based explainability improves model transparency by identifying influential traffic features contributing to classification decisions, thereby enhancing trust and supporting human-in-the-loop cybersecurity analysis. Although the reinforcement learning-based response mechanism provides a conceptual foundation for adaptive cyber defense, future work will focus on implementing and empirically validating this component, extending multi-class intrusion detection performance, and incorporating probability-based evaluation metrics such as ROC-AUC for comprehensive performance assessment. These enhancements will further strengthen the framework's applicability in modern cloud, IoT, and large-scale enterprise network environments.

7. Declarations

7.1. Author Contributions

Godfrey Perfectson Oise: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources; **Babatunde Seyi Olanrewaju:** Resources, Data Curation, Writing – Original Draft; **Oshoiribhor Austin Orukpe:** Resources, Data Curation, Writing – Original Draft; **Kevin Chinedu Pius:** Writing – Review & Editing, Visualization, Supervision, Project administration, Funding acquisition; **Augustine Osazee Airhiavbere:** Investigation, Visualization, Formal analysis.

7.2. Institutional Review Board Statement

Not applicable.

7.3. Informed Consent Statement

Not applicable.

7.4. Data Availability Statement

The data presented in this study are available in the manuscript.

7.5. Acknowledgment

Not applicable.

7.6. Conflicts of Interest

The authors declare no conflicts of interest.

8. References

[1] A. Anandita Iyer and K. S. Umadevi, "Role of AI and Its Impact on the Development of Cyber Security Applications," *Artificial Intelligence and Cyber Security in Industry 4.0*, 2023, pp. 23–46. https://doi.org/10.1007/978-981-99-2115-7_2.

[2] M. Sarhan, S. Layeghy, N. Moustafa, and M. Portmann, "NetFlow Datasets for Machine Learning-Based Network Intrusion Detection Systems," *Big Data Technologies and Applications,* 2021, pp. 117–135. https://doi.org/10.1007/978-3-030-72802-1_9.

[3] G. P. Oise, S. A. Oyedotun, O. C. Nwabuokei, A. E. Babalola, and N. B. Unuigbokhai, "Enhanced Prediction of Coronary Artery Disease Using Logistic Regression," *Fudma Journal of Sciences*, vol. 9, no. 3, pp. 201–208, Mar. 2025, https://doi.org/10.33003/fjs-2025-0903-3263.

[4] N. B. Unuigbokhai, G. P. Oise, B. E. Akilo, O. C. Nwabuokei, J. A. Odimayomi, S. K. Bakare, O. M. Atake, "Advancements in Federated Learning for Secure Data Sharing in Financial Services," *Fudma Journal of Sciences*, vol. 9, no. 5, pp. 80–86, May 2025, https://doi.org/10.33003/fjs-2025-0905-3207.

[5] A. M. K. Adawadkar and N. Kulkarni, "Cyber-security and reinforcement learning — A brief survey," *Eng Appl Artif Intell*, vol. 114, p. 105116, Sep. 2022, https://doi.org/10.1016/j.engappai.2022.105116.

[6] G. Oise and S. Konyeha, "E-Waste Management Through Deep Learning: A Sequential Neural Network Approach," *Fudma Journal of Sciences*, vol. 8, no. 3, pp. 17–24, Jul. 2024, https://doi.org/10.33003/fjs-2024-0804-2579.

[7] S. Dasgupta, A. Piplai, P. Ranade, and A. Joshi, "Cybersecurity Knowledge Graph Improvement with Graph Neural Networks," in *2021 IEEE International Conference on Big Data (Big Data)*, IEEE, Dec. 2021, pp. 3290–3297. https://doi.org/10.1109/BigData52589.2021.9672062.

[8] T. Bilot, N. El Madhoun, K. Al Agha, and A. Zouaoui, "Graph Neural Networks for Intrusion Detection: A Survey," *IEEE Access*, vol. 11, pp. 49114–49139, 2023, https://doi.org/10.1109/ACCESS.2023.3275789.

[9] B. Lakha, S. L. Mount, E. Serra, and A. Cuzzocrea, "Anomaly Detection in Cybersecurity Events Through Graph Neural Network and Transformer-Based Model: A Case Study with BETH Dataset," in *2022 IEEE International Conference on Big Data (Big Data)*, IEEE, Dec. 2022, pp. 5756–5764. https://doi.org/10.1109/BigData55660.2022.10020336.

[10] A. Shruti and Sreekumar, "Fintech and Financial Inclusion—A Review of Risk Management Strategies," *International Program and Project Management — Best Practices in Selected Industries*, 2025, pp. 199–216. https://doi.org/10.1007/978-3-031-80275-1_9.

[11] R. W. Idayani, R. Nadlifatin, A. P. Subriadi, and Ma. J. J. Gumasing, "A Comprehensive Review on How Cyber Risk Will Affect the Use of Fintech," *Procedia Comput Sci*, vol. 234, pp. 1356–1363, 2024, https://doi.org/10.1016/j.procs.2024.03.134.

[12] R. Gadge, A. Masharkar, A. Singh, N. Shelke, and A. Pimpalkar, "Managing Cybersecurity Risks in Emerging Technologies: Challenges and Solutions," in *2024 International Conference on Artificial Intelligence and Quantum Computation-Based Sensor Application (ICAIQSA)*, IEEE, Dec. 2024, pp. 1–9. https://doi.org/10.1109/ICAIQSA64000.2024.10882358.

[13] S. AlBenJasim, H. Takruri, R. Al-Zaidi, and T. Dargahi, "Development of cybersecurity framework for FinTech innovations: Bahrain as a case study," *International Cybersecurity Law Review*, vol. 5, no. 4, pp. 501–532, Dec. 2024, https://doi.org/10.1365/s43439-024-00130-4.

[14] M. Radha, Y. Vasa, A. R. Kumbham, P. Vallurupalli, S. A. Kumar, and D. A, "A Hybrid Graph Neural Network-Based Reinforcement Learning Approach for Adaptive Cybersecurity Risk Management in FinTech," in *2025 International Conference on Computing Technologies &amp; Data Communication (ICCTDC)*, IEEE, Jul. 2025, pp. 1–6. https://doi.org/10.1109/ICCTDC64446.2025.11158732.

[15] J. Alom, M. S. Ullah, M. T. Islam, M. Niloy, R. Islam, and S. Firdaus, “Adaptive Multi-Agent Reinforcement Learning for Intrusion Mitigation Aligned with Smart City,” in *2025 International Conference on Quantum Photonics, Artificial Intelligence, and Networking (QPAIN)*, IEEE, Jul. 2025, pp. 1–6. https://doi.org/10.1109/QPAIN66474.2025.11172093.

[16] G. Oise and S. Konyeha, “Environmental impacts in e-waste management using deep learning,” *Discover Artificial Intelligence*, vol. 5, no. 1, p. 210, Aug. 2025, https://doi.org/10.1007/s44163-025-00376-9.

[17] S. A. Oyedotun, G. P. Oise, and C. E. Ozobialu, “Towards Intelligent Cybersecurity in SCADA and DCS Environments: Anomaly Detection Using Multimodal Deep Learning and Explainable AI,” *Journal of Science Research and Reviews*, vol. 2, no. 3, pp. 20–31, Jul. 2025, https://doi.org/10.70882/josrar.2025.v2i3.76.

[18] S. A. Oyedotun, O. P. Ejenarhome, and G. P. Oise, “Learning Analytics and Predictive Modeling: Enhancing Student Success through Data-Driven Insights,” *Journal of Science Research and Reviews*, vol. 2, no. 3, pp. 42–51, Jul. 2025, https://doi.org/10.70882/josrar.2025.v2i3.77.

[19] D. Azamuke, M. Katarahweire, and E. Bainomugisha, “Financial Fraud Detection Using Rich Mobile Money Transaction Datasets,” *Towards new e-Infrastructure and e-Services for Developing Countries*, 2025, pp. 190–208. https://doi.org/10.1007/978-3-031-81573-7_16.

[20] A. Uprety and D. B. Rawat, “Reinforcement Learning for IoT Security: A Comprehensive Survey,” *IEEE Internet Things J*, vol. 8, no. 11, pp. 8693–8706, Jun. 2021, https://doi.org/10.1109/JIOT.2020.3040957.

[21] T. T. Nguyen and V. J. Reddi, “Deep Reinforcement Learning for Cyber Security,” *IEEE Trans Neural Netw Learn Syst*, vol. 34, no. 8, pp. 3779–3795, Aug. 2023, https://doi.org/10.1109/TNNLS.2021.3121870.

[22] S. Munikoti, D. Agarwal, L. Das, M. Halappanavar, and B. Natarajan, “Challenges and Opportunities in Deep Reinforcement Learning With Graph Neural Networks: A Comprehensive Review of Algorithms and Applications,” *IEEE Trans Neural Netw Learn Syst*, vol. 35, no. 11, pp. 15051–15071, Nov. 2024, https://doi.org/10.1109/TNNLS.2023.3283523.

[23] M. Pham, V. Vaze, and P. Chin, “Strategic Cyber Defense via Reinforcement Learning-Guided Combinatorial Auctions,” in *2025 IEEE High Performance Extreme Computing Conference (HPEC)*, IEEE, Sep. 2025, pp. 1–7. https://doi.org/10.1109/HPEC67600.2025.11196565.

[24] S. A. Oyedotun, G. P. Oise, B. E. Akilo, O. C. Nwabuokei, P. O. Ejenarhome, M. Fole, C. J. Onwuzo, “The Role of Internal Audit in Fraud Detection and Prevention: A Multi-Contextual Review and Research Agenda,” *Journal of Science Research and Reviews*, vol. 2, no. 2, pp. 76–85, May 2025, https://doi.org/10.70882/josrar.2025.v2i2.51.

[25] G. P. Oise, C. J. Onwuzo, M. Fole, S. A. Oyedotun, J. A. Odimayomi,N. B. Unuigbokhai, P. O. Ejenarhome, B. E. Akilo, “Decentralized Deep Learning in Healthcare: Addressing Data Privacy with Federated Learning,” *Fudma Journal of Sciences*, vol. 9, no. 6, pp. 19–26, Jun. 2025, https://doi.org/10.33003/fjs-2025-0906-3714.

[26] B. Blakely, “An Experimental Platform for Autonomous Intelligent Cyber-Defense Agents: Towards a collaborative community approach (WIPP),” in *2022 Resilience Week (RWS)*, IEEE, Sep. 2022, pp. 1–7. https://doi.org/10.1109/RWS55399.2022.9984037.

[27] G. Oise, Cyprian C. Konyeha, O. T. Comfort, S. Konyeha, and C. O. Emmanueld, “The Integration of Internet of Things (IoT) in Smart Classrooms: Opportunities, Challenges, and Future Trajectories,” *Journal of Digital Learning And Distance Education*, vol. 4, no. 3, pp. 1554–1567, Aug. 2025, https://doi.org/10.56778/jdlde.v4i3.537.

[28] G. G. James, G. P. Oise, E. G. Chukwu, N. A. Michael, W. F. Ekpo, and P. E. Okafor, “Optimizing Business Intelligence System Using Big Data and Machine Learning,” *Journal of Information Systems and Informatics*, vol. 6, no. 2, pp. 1215–1236, Jun. 2024, https://doi.org/10.51519/journalisi.v6i2.631.