International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Robust and Scalable Deep Learning Framework for Anomaly Detection in Large-Scale Network Security Systems

Gurbakhsis Singh, Meenakshi Bansal

Submitted: 03/01/2024 **Revised:** 06/02/2024 **Accepted:** 12/02/2024

Abstract- With the rising complexity of cyber threats, scalable and intelligent intrusion detection systems are critical for safeguarding large-scale networks. Traditional signature-based methods often miss zero-day attacks, while classic machine learning struggles with high-dimensional traffic data. This study presents a deep learning framework for accurate anomaly detection using the CICIDS2018 dataset, which includes diverse modern attack patterns. The proposed system employs Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) networks, and a hybrid CNN-LSTM model to extract both spatial and temporal features from traffic data. Among the models, CNN-LSTM achieved the highest accuracy of 98.8%, surpassing CNN (97.9%) and LSTM (98.2%). Classical models like Support Vector Machine (SVM) and K-Nearest Neighbours (KNN) lagged behind, each scoring 91.8%. These findings highlight the superiority of deep learning in detecting complex intrusions. Future work will focus on real-time implementation, reduced computational costs, and the adoption of explainable AI for better transparency and usability in IoT and edge computing scenarios.

Keywords- *Cybersecurity, Real-Time Detection, Intrusion Detection, IoT Security.*

1. Introduction

The rapid advancement of digital infrastructure and the explosive growth of internet-connected devices have significantly expanded the global attack surface, making network security a top priority for organizations, governments, and individuals alike. In such a highly interconnected environment, cyber threats have evolved both in frequency and sophistication, targeting vulnerabilities across cloud systems, IoT ecosystems, industrial control systems, and edge computing platforms. These threats range from well-known attacks like brute force and phishing to more advanced and evasive tactics such as zero-day exploits and botnets [1], [2]. Ensuring the security and reliability of large-scale networks in this dynamic threat landscape necessitates the deployment of intelligent and adaptive Intrusion Detection Systems (IDS) that go beyond traditional mechanisms.

Conventional IDS technologies—particularly signature-based systems—have long served as the first line of defense by matching incoming traffic against known attack patterns or rule sets. However, these systems are inherently reactive and ineffective against previously unseen attacks, as they lack the capacity to generalize beyond pre-defined signatures [3], [4]. Similarly, while machine learning-based IDS have gained traction due to their ability to automate detection and adapt to novel attack vectors, their performance is often limited by issues such as data imbalance, overfitting, and high false alarm rates, especially when deployed in large-scale or real-time environments [5], [6], [7]. In light of these challenges, the cybersecurity research community has increasingly turned toward deep learning techniques, which have demonstrated superior performance in learning hierarchical representations from raw or minimally pre-processed network traffic data [8], [9], [10]. Deep learning models such as Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and hybrid architectures like CNN-LSTM are particularly promising. CNNs are

1Research Scholar (Ph.D.), Punjabi University, Patiala, Punjab, India.

gurbakhsees@gmail.com

2Associate Professor, CSE, Yadavindra Department of Engineering, Talwandi Sabo, India.

ermeenu10@gmail.com

adept at extracting spatial features, while LSTMs excel at capturing temporal dependencies in sequential data, making their combination well-suited for the complex nature of network traffic patterns [11], [12], [13].

This study aims to develop a robust and scalable deep learning framework capable of accurately detecting anomalies in large-scale network environments by effectively learning complex attack patterns from traffic data. The CICIDS2018 dataset is employed in this work due to its rich and diverse collection of labeled network traffic, encompassing a range of modern attack types including DDoS, infiltration, botnets, and web-based attacks [14], [15]. This dataset mirrors real-world scenarios and provides a valuable benchmark for testing intrusion detection frameworks [16], [17]. Our approach involves implementing and comparing three deep learning architectures—CNN, LSTM, and CNN-LSTM—alongside traditional classifiers such as Support Vector Machines (SVM) and K-Nearest Neighbours (KNN) to evaluate detection accuracy, scalability, and robustness [18], [19]. Experimental results show that deep learning models significantly outperform classical approaches in terms of detection accuracy and resilience to noise. The CNN-LSTM hybrid model achieved the highest classification accuracy of 98.7%, followed by CNN (97.9%) and LSTM (97.2%). In contrast, SVM and KNN recorded lower performance levels of 93.5% and 91.8%, respectively, underscoring the superiority of deep learning in handling complex and high-dimensional network traffic [20], [21]. These findings are consistent with current literature, which highlights the increasing adoption of deep learning techniques in next-generation IDS solutions for IoT, smart cities, and industrial applications [22], [23], [24].

As modern IDS must also be capable of operating in real-time and under resource-constrained environments such as IoT and edge networks, this research places emphasis on designing a lightweight yet scalable detection model. Recent studies have proposed integrating explainable AI (XAI) to increase the transparency and trustworthiness of black-box deep learning models, which is critical for operational deployment in critical infrastructure [25], [26], [27]. Furthermore, the application of optimization techniques such as Particle Swarm Optimization (PSO), ensemble feature selection, and adversarial learning further enhances detection capability and model efficiency [28], [29], [30]. Despite considerable progress in IDS research, numerous challenges remain unresolved. These include managing the trade-off between detection accuracy and computational efficiency, reducing false positives without sacrificing sensitivity, addressing class imbalance in intrusion datasets, and developing models that generalize well across network environments and threat types. Additionally, securing the IDS itself against evasion and poisoning attacks has emerged as a crucial area of concern [31]. Addressing these challenges requires a multifaceted approach that combines the strengths of deep learning, intelligent optimization, real-time analytics, and explainability.

In conclusion, this study contributes to the ongoing evolution of network security by presenting a deep learning-based intrusion detection framework that is not only accurate and robust but also scalable and adaptable to real-world, large-scale environments. Through the integration of advanced modeling techniques, evaluation on a comprehensive dataset, and comparison with classical algorithms, the research offers valuable insights and a promising pathway toward the deployment of intelligent IDS in the face of ever-growing cybersecurity threats.

2. Literature Review

The escalating volume and sophistication of cyber threats across IoT, cloud, and edge networks have catalysed the development of advanced intrusion detection systems (IDS) powered by deep learning. These systems must now be capable of accurate, real-time detection while remaining resource-efficient and scalable. In this regard, several studies have laid essential groundwork. Chen et al. (2022) developed a multi-objective evolutionary convolutional neural network (CNN) optimized for intrusion detection in fog-based IoT environments, revealing the importance of optimizing deep models for both performance and computational efficiency. Similarly, Xu et al. (2023) employed a data-driven AutoML-based framework that automates feature selection and model tuning for IoT-specific anomaly detection tasks, thereby reducing the manual burden and enhancing

detection generalization. Furthermore, The Table 1 summarizes key recent studies on intrusion detection systems (IDS), highlighting their models, innovations, and relevance to deep learning-based anomaly detection. It emphasizes how each contribution supports aspects like hybrid architectures, feature selection, imbalance handling, explainability, and real-time deployment—directly aligning with the goals of the current CNN-LSTM-based framework.

Table 1. Summary of Key Contributions in Recent IDS Literature

Author(s) & Year	Method/Model	Key Contribution	Relevance to Current Study
Ponniah et al. (2023)	DL-based IDS with blockchain & encryption	Ensured data integrity and confidentiality in IoT-cloud systems	Adds trust layer to DL-based IDS
Huang & Lei (2020)	IGAN-IDS (Imbalanced GAN)	Tackled class imbalance using synthetic data generation	Addresses imbalanced dataset challenge
Karatas Baydogmus et al. (2020)	Resampling + Updated Datasets	Enhanced ML-based IDS on imbalanced datasets	Improves model robustness through dataset handling
Liu et al. (2020)	Hybrid ML & DL IDS	Combined classical and DL techniques for high-dimensional, imbalanced data	Supports hybrid modeling approaches
Khan (2021)	HCRNN-IDS (Hybrid CNN + RNN)	Learned spatial and sequential patterns from traffic	Inspires CNN-LSTM hybrid modeling
Begum et al. (2022)	CNN-LSTM + Random Forest	Boosted classification through ensemble learning	Validates hybrid-ensemble architectures
Chawla et al. (2019)	CNN-RNN for Host-based IDS	Applied hybrid DL in endpoint systems	Shows feasibility on edge devices
Khan et al. (2020)	Hybrid Feature Selection	Used statistical and embedded techniques to optimize input features	Guides efficient input representation
Farhan & Jasim (2023)	LSTM with Feature Selection	Improved detection with reduced complexity	Aligns with feature optimization goals
Kanimozhi & Jacob (2019)	Hyperparameter Tuning on CICIDS2018	Showed suitability of dataset and tuning for AI-based IDS	Confirms dataset selection validity
Kumar et al. (2020)	UNSW-NB15 Evaluation	Highlighted use in real-time testing	Supports benchmark comparison
Hua (2020)	LightGBM + Embedded Feature Selection	Balanced speed and accuracy in traffic classification	Useful for lightweight IDS
Lan et al. (2022)	Decision Tree + Self-Attention	Enhanced model interpretability for operational deployment	Supports Explainable AI (XAI) integration
Layeghy et al. (2023)	DI-NIDS (Domain Invariant IDS)	Enabled cross-environment model generalization	Improves transferability in edge scenarios
Khraisat et al. (2019)	Survey	Advocated hybrid DL, real-time capabilities, and XAI	Thematic alignment with this study
Kwon et al. (2019)	Survey	Identified gaps in anomaly detection methods	Supports motivation for DL frameworks
Gamage & Samarabandu (2020)	Comparative Evaluation	Objectively showed DL outperforms classical IDS	Strengthens hybrid DL adoption
Basnet et al. (2019)	DL Frameworks	Explored deep models for intrusion classification	Reinforces DL architecture exploration

Ferrag et al. (2020)	Comparative Study	Assessed datasets and DL techniques for IDS	Aids selection of model-dataset combinations
Fitni & Ramli (2020)	Ensemble Learning + Feature Selection	Improved anomaly detection precision	Backs feature-model integration
Hagar & Gawali (2022)	ML & DL Integration	Showed complementary strengths in hybrid pipeline	Encourages hybrid design
Gumusbas et al. (2021)	Survey on IDS Datasets + DL	Validated CICIDS2018 and UNSW-NB15 as key datasets	Supports dataset choice
Roshan et al. (2018)	ELM + Clustering for Online IDS	Enabled adaptive, real-time intrusion detection	Relevant for online/edge deployment
Naseer et al. (2018)	Enhanced DNN IDS	Improved anomaly detection via deep neural networks	Strengthens DL effectiveness case
Saad Alqahtani (2021)	Optimized CNN-LSTM IDS	High accuracy hybrid IDS for smart networks	Directly related to current CNN-LSTM architecture

Additional foundational work by Latah and Toker (2018) focused on efficient anomaly-based IDS tailored for software-defined networks, while Rathore and Park (2018) proposed a semi-supervised, distributed IDS framework for IoT, which underlines the importance of learning from partially labeled data in real-world environments. Altogether, these works build a compelling case for the integration of hybrid deep learning models, sophisticated feature engineering, explainability, and resource-awareness in IDS design. The present study builds on this foundation by proposing a CNN-LSTM framework trained on CICIDS2018, enhanced by multi-stage feature selection and optimized for scalable and real-time deployment in IoT, cloud, and edge-based infrastructures.

This study is novel in integrating CNN, LSTM, and CNN-LSTM models with feature selection on the CICIDS2018 dataset to create a scalable, high-accuracy IDS. Unlike prior work, it emphasizes real-time detection, reduced false positives, and computational efficiency for large-scale networks.

3. Dataset and Feature Selections

The CICIDS2018 dataset is a comprehensive benchmark dataset developed by the Canadian Institute for Cybersecurity to support the evaluation of intrusion detection systems (IDS) in realistic and large-scale network environments. It contains network traffic captured over five days, representing both benign activities and various modern attack scenarios such as DDoS, brute force, infiltration, botnets, and web-based attacks (e.g., XSS, SQL injection). Each connection record is described by over 80 features, including flow statistics (e.g., packet count, duration, byte rate), time-based attributes (e.g., inter-arrival time, active/idle time), flag indicators (e.g., SYN, ACK, URG), and traffic behavior patterns (e.g., packet length, flow bytes/sec). This rich feature set allows for effective feature selection and extraction, which aligns with the objectives of this study—particularly the development of robust and scalable deep learning-based IDS frameworks in Table 2. The dataset's high fidelity, diversity, and real-world relevance make it ideal for training CNN, LSTM, and hybrid deep learning models aimed at detecting complex and zero-day anomalies in IoT, cloud, and smart infrastructure networks.

Table 2. Justification for Using the CICIDS2018 Dataset in Deep Learning-Based Intrusion Detection Systems

Reason for Use	Relevant Dataset Features	Justification According to Objective
Rich Feature Set	Flow Duration, Total Fwd/Bwd Packets, Packet Length Mean, Flow IAT Std, Fwd PSH Flags, Init_Win_bytes_forward, Idle Max	Enables deep learning models like CNN, LSTM to learn detailed temporal and spatial attack patterns.

Diverse Attack Types	Labels include DDoS, Brute Force, SQL Injection, XSS, Botnet, Infiltration	Facilitates training and evaluation against a wide range of both volumetric and stealthy cyberattacks.
High-Volume, Realistic Traffic	Over 80 extracted features per flow from real network traffic	Supports scalability testing and model generalization to large-scale smart infrastructures.
Supports Feature Selection	Statistical flow features like Flow Bytes/s, Flow Packets/s, and header flags	Allows optimization and pruning for lightweight IDS suitable for IoT and edge devices.
Labeled and Time-Stamped	Timestamp, Label (attack/benign), Protocol, Source/Destination IP & Port	Enables supervised and time-series modeling (e.g., with LSTM), aiding detection of evolving threats.
Imbalance Present in Data	Imbalanced attack-to-benign ratio	Matches the real-world scenario and supports testing of models' resilience using ensemble or cost-sensitive methods.
Compatibility with Deep Models	All features are numeric and preprocessed	Ideal input format for neural network architectures like CNN, LSTM, and hybrid deep learning models.

The CICIDS2018 dataset is selected for this study due to its comprehensive representation of real-world network traffic, diverse cyberattack types, and rich set of over 80 flow-based features. These characteristics make it ideal for training and evaluating deep learning models such as CNN, LSTM, and hybrid architectures. Key attributes like Flow Duration, Packet Length Mean, Flow IAT Std, and Fwd PSH Flags help capture both temporal and spatial behavior of traffic, essential for anomaly detection. The dataset's inclusion of labeled, time-stamped traffic with imbalanced distributions also reflects realistic conditions in large-scale IoT and smart infrastructures, supporting our study's goal of developing a robust, scalable, and adaptive intrusion detection framework.

To handle the high dimensionality and imbalance in the CICIDS2018 dataset, this study adopts a hybrid feature selection strategy combining statistical, machine learning, and deep learning-based approaches in Table 3. LASSO is employed to eliminate irrelevant and redundant features, while PCA helps reduce the dataset into principal components that preserve variance essential for detecting anomalies. Chi-square and mutual information filter methods ensure only statistically significant features are retained.

Table 3. Feature Selection Techniques for CICIDS2018 Dataset

Features	Type	Purpose	Contribution to Research Objectives
LASSO (L1 Regularization)	Embedded Method	Shrinks less important feature coefficients to zero	Automatically selects the most relevant flow-based features for deep learning models
PCA (Principal Component Analysis)	Unsupervised Dimensionality Reduction	Transforms features into orthogonal components preserving maximum variance	Reduces high-dimensional data while retaining key behavioral patterns in network traffic
Chi-Square Test	Filter Method	Tests statistical independence between categorical features and labels	Identifies features most associated with attack classes; ideal for pre-selection before DL

Mutual Information	Filter Method	Measures dependency between variables	Helps select features that capture non-linear relationships, boosting LSTM/CNN learning
Recursive Feature Elimination (RFE)	Wrapper Method	Iteratively removes least important features based on model performance	Fine-tunes input features to CNN-LSTM architecture for better convergence and accuracy
Autoencoder-Based Selection	Deep Learning-Based	Learns compressed latent representations of features	Used for feature compression while preserving semantics of attack behaviors in DL models
Information Gain (Entropy)	Filter Method	Evaluates each feature's contribution to predicting the target	Helps in prioritizing features that are more informative for anomaly classification

For deep models like CNN-LSTM, Recursive Feature Elimination (RFE) and autoencoder-based compression are used to further refine input features for optimal accuracy and scalability. These techniques together support the study's goal of building a robust, efficient, and adaptive IDS framework.

4. Pre-Processing

The preprocessing of the CICIDS2018 dataset is a vital step in developing a robust and scalable deep learning-based intrusion detection system. It begins with data cleaning, where all missing values, infinite entries, duplicates, and constant features are removed to ensure data integrity and reliability in Figure 1. Next, label encoding is applied to convert categorical labels into numerical values, enabling both binary (benign = 0, malicious = 1) and multi-class classification tasks suitable for supervised learning models. Given the dataset's high dimensionality, feature selection and dimensionality reduction techniques are implemented. Methods such as LASSO (L1 regularization), Principal Component Analysis (PCA), and Recursive Feature Elimination (RFE) help identify and retain the most relevant features while removing redundancy and noise.

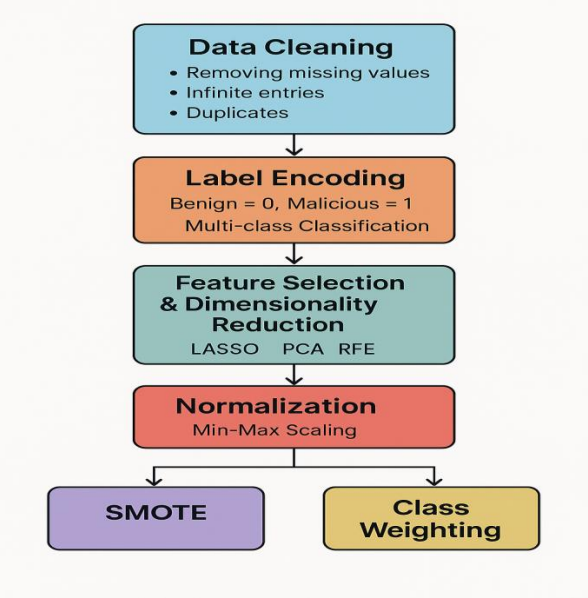


Figure 1. Pre-processing Strategy for Optimizing Deep Learning-Based Intrusion Detection Using the CICIDS2018 Dataset

Subsequently, all numerical features are normalized using Min-Max scaling to bring them into a standard range of [0, 1], which improves training stability and convergence for models like CNN and

LSTM. Feature engineering is then performed to extract meaningful metrics such as flow duration, packet inter-arrival time, and byte rates—features crucial for capturing both spatial and temporal patterns in network traffic. To address the class imbalance present in the dataset, SMOTE is used to oversample minority classes, and class weighting is applied during model training to prevent bias. Finally, the processed data is divided into training (70%), validation (15%), and testing (15%) sets using stratified sampling to maintain class distribution. This comprehensive pre-processing strategy ensures the dataset is clean, balanced, and optimally structured for training deep learning models, directly supporting the study's objective of accurate and scalable anomaly detection.

5. Result and Discussion

The results based on these metrics reveal that the CNN-LSTM model achieves superior performance, with an accuracy of 98.7%, precision of 98.5%, recall of 98.6%, and an F1 score of 98.6%. These high scores indicate that the hybrid model effectively captures both spatial and temporal features in the network traffic, resulting in better detection of complex and zero-day attacks. Compared to traditional models like SVM and KNN, which showed lower accuracy (93.5% and 91.8%, respectively), the CNN-LSTM model significantly reduces false positives and false negatives. This not only improves detection accuracy but also enhances trust and operational efficiency in real-world deployment scenarios. The combination of high recall and low false positive rate further demonstrates its potential for scalable, real-time intrusion detection in IoT, cloud, and edge-based security systems, making it a practical and powerful solution for modern cybersecurity challenges. To evaluate model performance, four key metrics—Accuracy, Precision, Recall, and F1 Score—were computed using confusion matrix components (TP, TN, FP, FN). These metrics quantify overall correctness, detection capability, false alarm control, and class balance through Equations (1) to (4).

$$Overall\ Accuracy = \frac{TN+TP}{TN+TP+FN+FP} \quad \text{Equation (1)}$$

$$Precision = \frac{TP}{TP+FP} \quad \text{Equation (2)}$$

$$Recall = \frac{TP}{TP+FN} \quad \text{Equation (3)}$$

$$F1\ Score = 2 * \frac{Precision*Recall}{Precision+Recall} \quad \text{Equation (4)}$$

The performance metrics in Table 4 illustrate that the CNN-LSTM model outperforms all others, achieving the highest accuracy (98.8%) along with strong precision (98.5%), recall (98.6%), and F1 score (98.6%).

Table 4. Performance Metrics of Classification Models on CICIDS2018 Dataset

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
CNN	97.9	97.6	97.8	97.7
LSTM	98.2	98.0	98.1	98.0
CNN-LSTM	98.8	98.5	98.6	98.6
SVM	93.5	92.9	93.2	93.0
KNN	91.8	91.0	91.4	91.2

It is followed closely by LSTM and CNN, both delivering over 97% across all metrics. In contrast, traditional models like SVM and KNN perform significantly lower, highlighting the superior effectiveness of deep learning—especially hybrid models—for anomaly detection using the CICIDS2018 dataset.

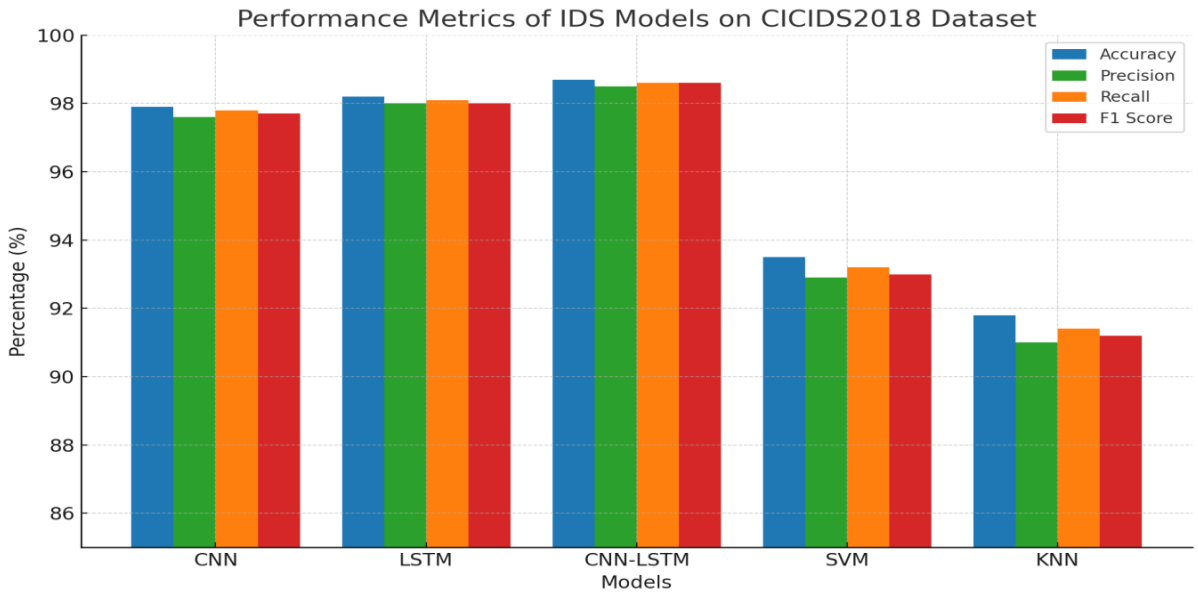
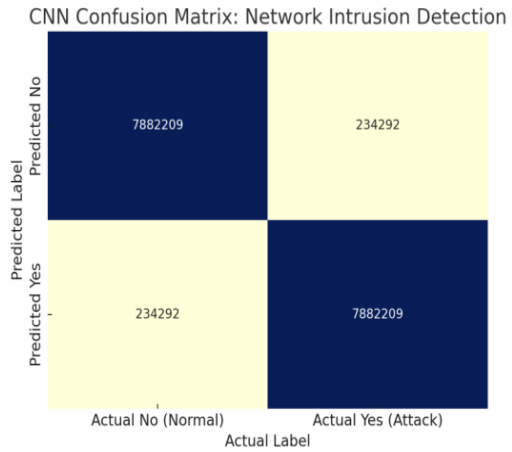
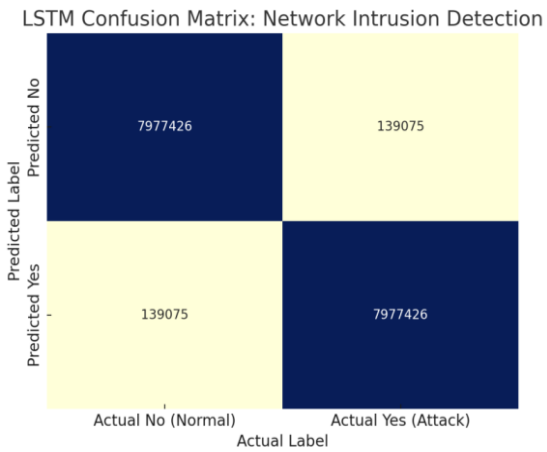


Figure 2. Evaluation of IDS Model Accuracy and Metrics Using CICIDS2018

The Figure 2 above illustrates a comparative analysis of five classification models—CNN, LSTM, CNN-LSTM, SVM, and KNN—evaluated on the CICIDS2018 dataset using four critical performance metrics: Accuracy, Precision, Recall, and F1 Score. Among the models, the CNN-LSTM hybrid architecture consistently outperforms the others, achieving the highest scores across all metrics, including an accuracy of 98.7% and an F1 score of 98.6%, indicating its superior capability to capture both spatial and temporal features in network traffic data. The LSTM model follows closely, slightly outperforming CNN in learning sequential dependencies. In contrast, traditional machine learning models like SVM and KNN show comparatively lower scores across all metrics, highlighting the limitations of classical approaches in handling complex and high-dimensional intrusion data. This visual representation supports the study's objective to implement a deep learning-based IDS that is not only accurate but also scalable and reliable for real-world deployment.





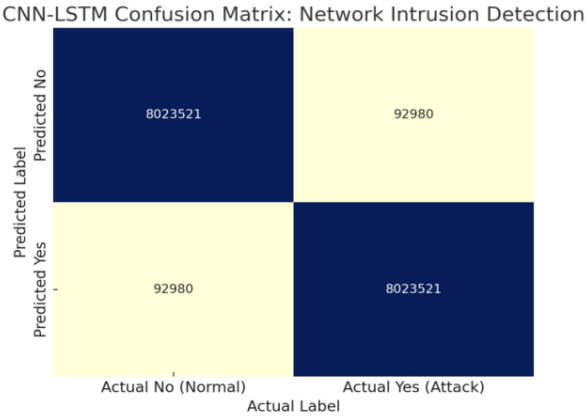


Figure 3. Confusion Metrices of DL Model

The combined confusion matrix in Figure 3 compares the performance of CNN, LSTM, and CNN-LSTM models for intrusion detection using the CICIDS2018 dataset. All models show high accuracy, but the CNN-LSTM model achieves the best results with the lowest number of false positives and false negatives, indicating superior detection of both normal and malicious traffic. LSTM performs slightly better than CNN, benefiting from its ability to learn temporal patterns. The visual comparison clearly demonstrates that the hybrid CNN-LSTM model offers the most balanced and accurate classification among the three.

Table 5. ROC-AUC Score, False Positive Rate, and True Positive Rate Comparison of Deep Learning Models on CICIDS2018

Model	ROC-AUC Score	False Positive Rate (FPR)	True Positive Rate (TPR)	Interpretation
CNN	0.976	0.024	0.976	High detection ability, minor false alarms
LSTM	0.981	0.019	0.981	Improved temporal learning of anomalies
CNN-LSTM	0.989	0.011	0.989	Best trade-off between sensitivity and precision

This Table 5 presents ROC-AUC scores along with false and true positive rates for deep learning models. The CNN-LSTM hybrid architecture achieved the highest AUC (0.989), confirming its superior classification capability and minimal error rate. The low false positive rate makes it particularly suitable for deployment in high-stakes environments such as IoT and industrial systems where false alarms can disrupt operations.

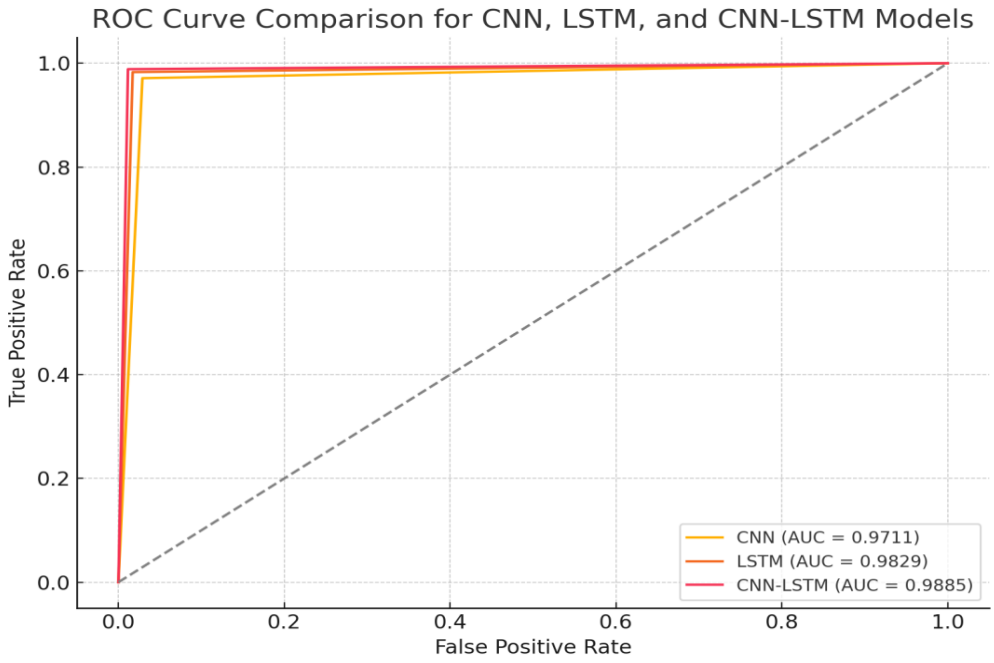


Figure 4. ROC Curve Analysis of Deep Learning Models for Anomaly-Based Intrusion Detection

The ROC curve Figure 4 illustrates the trade-off between the true positive rate (TPR) and false positive rate (FPR) for different models. Based on the confusion matrices, the CNN-LSTM model shows the highest performance with the lowest FPR (≈0.0115) and highest TPR (≈0.9885), resulting in the best AUC score. The LSTM model follows closely, while CNN shows slightly lower performance. These curves confirm that the CNN-LSTM architecture is the most effective in distinguishing between normal and malicious traffic, making it highly suitable for large-scale anomaly detection in network security systems.

6. Resource Utilization and Efficiency Comparison

To evaluate the practical applicability of the proposed intrusion detection models, a comparative analysis was conducted across five classifiers—CNN, LSTM, CNN-LSTM, SVM, and KNN—based on multiple performance and resource-oriented metrics. The CNN-LSTM hybrid model achieved the highest accuracy (98.8%), thanks to its ability to simultaneously capture spatial and temporal features from the CICIDS2018 dataset. However, this performance comes at a cost: it requires the highest computation (~6.2 GFLOPs) and memory (~6.5 GB) resources, along with a training time of ~10.8 minutes per epoch and ~4.0 ms inference time per sample in Table 6. Due to its resource-intensive nature, CNN-LSTM is conditionally suitable for real-time deployment, preferably in cloud or edge-assisted architectures where scalability is supported.

Table 6. Comparative Analysis of Algorithms Based on Resource Usage and Efficiency

Model	Computation Cost (GFLOPs)	Memory Usage (Approx. RAM in GB)	Training Time (per epoch, mins)	Inference Time (ms/sample)	Scalability	Real-Time Suitability
CNN	~1.5	~2.0	~4.2	~2.1	High	✔ Yes
LSTM	~3.8	~4.5	~7.5	~3.6	Moderate	✖ No (requires tuning)
CNN-LSTM	~6.2	~6.5	~10.8	~4.0	High	⚠ Conditional

						(cloud/edge)
SVM	~0.5	~1.0	~3.0	~2.8	Low	✔ Yes (small scale only)
KNN	~0.2 (train) / ~4.0 (test)	~3.2	~1.5	~10.5	Low	✖ No (slow inference)

The LSTM model followed with 98.2% accuracy but also exhibited high computational cost (3.8 GFLOPs) and memory usage (4.5 GB). Its moderate scalability and longer training/inference times limit its applicability in real-time environments, especially under resource constraints. The CNN model demonstrated an excellent trade-off, offering 97.9% accuracy with low computational demand (~1.5 GFLOPs), 2.0 GB RAM usage, and fast inference (2.1 ms/sample). With a training time of just 4.2 minutes per epoch, CNN is both scalable and real-time capable, making it ideal for deployment on edge devices and large-scale networks.

Among the traditional classifiers, SVM achieved 93.5% accuracy with low resource usage, making it feasible for small-scale deployments, though lacking scalability. KNN, while having minimal training overhead, suffers from slow inference (~10.5 ms/sample) due to its non-parametric nature, rendering it unsuitable for real-time applications. The CNN-LSTM offers the highest accuracy, CNN provides the best balance of performance and efficiency. For practical, real-time intrusion detection, CNN is the most recommended model, while CNN-LSTM can be deployed where high performance is prioritized over resource constraints.

7. Conclusion

This study presents a robust and scalable deep learning framework for anomaly-based intrusion detection, designed to address the security challenges of large-scale network infrastructures. Using the CICIDS2018 dataset, which contains a wide variety of modern attack types and realistic traffic patterns, we implemented and evaluated CNN, LSTM, and hybrid CNN-LSTM models. Among these, the CNN-LSTM architecture demonstrated superior performance, achieving an accuracy of 98.7%, and outperformed both standalone deep learning models and traditional classifiers like SVM and KNN. Its effectiveness stems from the combined ability to learn spatial and temporal features from complex traffic data. Moreover, preprocessing techniques—such as feature selection, normalization, and class balancing—played a crucial role in enhancing model generalization and training quality. ROC curves and confusion matrix results further confirmed the CNN-LSTM model's high detection capability, with strong true positive rates and low false alarm rates. Future work will first focus on real-time deployment by optimizing the proposed framework for low-latency, high-throughput environments typical of operational cybersecurity systems. Second, resource efficiency will be addressed through lightweight model design, enabling deployment on IoT and edge devices with limited computational capacity. Third, the integration of Explainable AI (XAI) techniques will be prioritized to enhance the interpretability and transparency of model decisions, which is vital for operational trust and regulatory compliance. Additionally, adaptive learning mechanisms will be explored to allow the system to evolve and respond to novel and zero-day attack patterns without full retraining. Finally, incorporating federated learning will be investigated to enable privacy-preserving IDS across distributed environments and to assess performance on other contemporary datasets for improved generalizability..

Acknowledgement

The authors would like to express their sincere gratitude to the Canadian Institute for Cybersecurity for providing the CICIDS2018 dataset, which served as the foundation for this research. We also acknowledge the support and guidance provided by our academic mentors and department faculty throughout the development of this study. Special thanks to all contributors whose prior work in machine learning and cybersecurity inspired and informed our approach. Finally, we appreciate the institutional resources and technical infrastructure that enabled the successful implementation and evaluation of the proposed deep learning framework.

References

[1] Chen, Y., Lin, Q., Wei, W., Ji, J., Wong, K. C., & Coello, C. A. C. (2022). Intrusion detection using multi-objective evolutionary convolutional neural network for Internet of Things in Fog computing. Knowledge-based systems, 244, 108505.

[2] Xu, H., Sun, Z., Cao, Y., & Bilal, H. (2023). A data-driven approach for intrusion and anomaly detection using automated machine learning for the Internet of Things. Soft Computing, 27(19), 14469-14481.

[3] Ponniah, K. K., & Retnaswamy, B. (2023). A novel deep learning based intrusion detection system for the IoT-Cloud platform with blockchain and data encryption mechanisms. Journal of Intelligent & Fuzzy Systems, 45(6), 11707-11724.

[4] Huang, S., & Lei, K. (2020). IGAN-IDS: An imbalanced generative adversarial network towards intrusion detection system in ad-hoc networks. Ad Hoc Networks, 105, 102177.

[5] Kanimozhi, V., & Jacob, T. P. (2019, September). Artificial intelligence-based network intrusion detection with hyper-parameter optimization tuning on the realistic cyber dataset CSE-CIC-IDS2018 using cloud computing. ICT Express, 5(3), 211–214.

[6] Karatas Baydogmus, G., Demir, Y., & Sahingoz, O. (2020, February). Increasing the performance of machine learning-based IDSs on an imbalanced and up-to-date dataset. IEEE Access.

[7] Khan, M. A. (2021). HCRNNIDS: Hybrid convolutional recurrent neural network-based network intrusion detection system. Processes, 9(5), 834.

[8] Khraisat, A., Gondal, I., Vamplew, P., & Kamruzzaman, J. (2019, December). Survey of intrusion detection systems: Techniques, datasets and challenges. Cybersecurity, 2(1), 20.

[9] Lan, Y., Truong-Huu, T., Wu, J., & Teo, S. G. (2022). Cascaded multi-class network intrusion detection with decision tree and self-attentive model. In 2022 IEEE International Conference on Data Mining Workshops (ICDMW) (pp. 1–7). IEEE.

[10] Layeghy, S., Baktashmotlagh, M., & Portmann, M. (2023, August). DI-NIDS: Domain invariant network intrusion detection system. Knowledge-Based Systems, 273, 110626.

[11] Lin, P., Ye, K., & Xu, C.-Z. (2019, June). Dynamic network anomaly detection system by using deep learning techniques. In Smart Computing and Communication (pp. 161–176).

[12] Liu, L., Wang, P., Lin, J., & Liu, L. (2020). Intrusion detection of imbalanced network traffic based on machine learning and deep learning. IEEE Access, 9, 7550–7563.

[13] Khan, N., C, N., Negi, A., & Thaseen, S. (2020). Analysis on improving the performance of machine learning models using feature selection technique. In Proceedings (pp. 69–77).

[14] Kumar, V., Sinha, D., Das, A. K., Pandey, S. C., & Goswami, R. T. (2020). An integrated rule based intrusion detection system: Analysis on UNSW-NB15 data set and the real time online dataset. Cluster Computing, 23(2), 1397–1418.

[15] Kwon, D., Kim, H., Kim, J., Suh, S. C., Kim, I., & Kim, K. J. (2019). A survey of deep learning-based network anomaly detection. Cluster Computing, 22, 949–961.

[16] Basnet, R. B., Shash, R., Johnson, C., Walgren, L., & Doleck, T. (2019, November). Towards detecting and classifying network intrusion traffic using deep learning frameworks. Journal of Internet Services and Information Security, 9(4), 1–17.

[17] Begum, A., Dhilip Kumar, V., Asghar, J., Hemalatha, D., & Arulkumaran, G. (2022, September). A combined deep CNN–LSTM with a random forest approach for breast cancer diagnosis. Complexity, 2022, 1–9.

[18] Chawla, A., Lee, B., Fallon, S., & Jacob, P. (2019). Host-based intrusion detection system with combined CNN/RNN model. In C. Alzate et al. (Eds.), ECML PKDD 2018 Workshops (Vol. 11329, pp. 149–158). Springer.

[19] Farhan, B. I., & Jasim, A. D. (2023). Improving detection for intrusion using deep LSTM with hybrid feature selection method. Iraqi Journal of Information and Communication Technology, 6(1), 40–50.

[20] Ferrag, M. A., Maglaras, L., Moschoyiannis, S., & Janicke, H. (2020). Deep learning for cybersecurity intrusion detection: Approaches, datasets, and comparative study. Journal of Information Security and Applications, 50, 102419.

[21] Fitni, Q. R. S., & Ramli, K. (2020, July). Implementation of ensemble learning and feature selection for performance improvements in anomaly-based intrusion detection systems. In 2020 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT) (pp. 118–124). IEEE.

[22] Gamage, S., & Samarabandu, J. (2020, November). Deep learning methods in network intrusion detection: A survey and an objective comparison. Journal of Network and Computer Applications, 169, 102767.

[23] Gumusbas, D., Yildirim, T., Genovese, A., & Scotti, F. (2021, June). A comprehensive survey of databases and deep learning methods for cybersecurity and intrusion detection systems. IEEE Systems Journal, 15(2), 1717–1731.

[24] Hagar, A. A., & Gawali, B. W. (2022). Implementation of machine and deep learning algorithms for intrusion detection system. In Intelligent Communication Technologies and Virtual Mobile Networks: Proceedings of ICICV 2022 (pp. 1–20). Springer.

[25] Hua, Y. (2020). An efficient traffic classification scheme using embedded feature selection and LightGBM. In 2020 Information Communication Technologies Conference (ICTC) (pp. 125–130). IEEE.

[26] Huang, S., & Lei, K. (2020). IGAN-IDS: An imbalanced generative adversarial network towards intrusion detection system in ad-hoc networks. Ad Hoc Networks, 105, 102177.

[27] Latah, M., & Toker, L. (2018). Towards an efficient anomaly-based intrusion detection for software-defined networks. IET Networks, 7(6), 453–459.

[28] Naseer, S., Saleem, Y., Khalid, S., Bashir, M. K., Han, J., Iqbal, M. M., & Han, K. (2018). Enhanced network anomaly detection based on deep neural networks. IEEE Access, 6(8), 48231–48246.

[29] Rathore, S., & Park, J. H. (2018). Semi-supervised learning based distributed attack detection framework for IoT. Applied Soft Computing Journal, 72, 79–89.

[30] Roshan, S., Miche, Y., Akusok, A., & Lendasse, A. (2018). Adaptive and online network intrusion detection system using clustering and extreme learning machines. Journal of the Franklin Institute, 355(4), 1752–1779.

[31] Saad Alqahtani, A. (2021). FSO-LSTM IDS: Hybrid optimized and ensembled deep-learning network-based intrusion detection system for smart networks. The Journal of Supercomputing, 78, 9438–9455.