International Journal of Innovative Computing, Information and Control
Volume **21**, Number **6**, December **2025**

 ISSN 1349-4198
pp. **1807–1824**

RGNIDS: A RESIDUAL-ENHANCED CNN-BIGRU-BASED NETWORK INTRUSION DETECTION SYSTEM WITH SELF-ATTENTION MECHANISM

Kunsan Zhang[1], Song Zhang[1], Chaopeng Li[2,*], Bingjie Xiang[2] and Jiachun Zheng[2]

[1]State Grid Fujian Electric Power Co., Ltd. Zhangzhou Power Supply Company
No. 13, Shengli East Road, Xiangcheng District, Zhangzhou 363000, P. R. China
{ zhangkunsan1991; Zhangsong1983 }@gmail.com

[2]School of Ocean Information Engineering
Jimei University
No. 185, Yinjiang Road, Jimei District, Xiamen 361021, P. R. China
{ jerryxiang; jchzheng }@jmu.edu.cn
*Corresponding author: licp@jmu.edu.cn

Received April 2025; revised July 2025

Abstract. *The rapid advancement of Internet of Things (IoT) networks has intensified cybersecurity challenges in distributed smart environments, where traditional rule-based intrusion detection systems struggle to identify unknown attacks and manage diverse data streams. To address class imbalance and feature redundancy in IoT intrusion data, this study visualizes feature correlations and employs statistical tests to evaluate their relevance. A hybrid sampling technique is then applied to balancing the dataset. We propose a novel IoT-aware network intrusion detection system (RGNIDS) that integrates convolutional neural networks (CNNs), bidirectional gated recurrent units (BiGRUs), and a self-attention mechanism to enhance feature extraction. This architecture is tailored to the unique demands of IoT environments. Evaluations using the CSE-CIC-IDS2018 and Network Security Laboratory-Knowledge Discovery in Databases (NSL-KDD) datasets demonstrate that the RGNIDS outperforms traditional machine learning and existing deep learning methods regarding accuracy, precision, recall, and F1-score. These results demonstrate the model's effectiveness in enhancing intrusion detection for IoT security.*
Keywords: Network intrusion detection, Internet of Things, Bidirectional gated recurrent units, Self-attention mechanism, Hybrid sampling

1. **Introduction.** In modern society, Internet of Things (IoT)-enabled networks have permeated every industry and personal life through interconnected smart devices by triggering extensive data transmission and exchange across IoT ecosystems. As IoT-driven information technology rapidly advances and achieves ubiquitous connectivity, cybersecurity issues are becoming increasingly critical, with significant implications for key IoT-dependent sectors such as smart grids and national security. These attacks, which range from exploitation of vulnerable IoT protocols to complex coordinated intrusions across industrial IoT systems, pose serious threats to the security of distributed IoT architectures, data integrity, and trust in cyber-physical systems. As cybercriminals evolve attack vectors specifically targeting IoT environments, effective security measures have become essential for safeguarding heterogeneous IoT deployments.

DOI: 10.24507/ijicic.21.06.1807

Intrusion detection systems (IDSs) [1] provide essential safeguards for IoT ecosystems by monitoring and analyzing device-to-cloud network traffic to identify suspicious activities in IoT communication protocols. These systems are generally classified into IoT-adaptive network intrusion detection systems (NIDSs) [2] and host-based intrusion detection systems (HIDSs) [3]. HIDSs monitor embedded IoT endpoints, whereas NIDSs analyze IoT device communications across fog and cloud layers. Modern NIDSs must address IoT-specific attack surfaces by examining encrypted message queuing telemetry transport (MQTT) or constrained application protocol (CoAP) traffic and comparing patterns with IoT behavioral baselines. These solutions are crucial for detecting IoT-botnet distributed denial of service (DDoS) attacks [4], malicious firmware updates [5], and edge device spoofing [6], thereby providing real-time security for smart home and industrial IoT networks. Unlike conventional firewalls designed for static information technology (IT) infrastructures, IoT-aware NIDSs dynamically adapt to evolving device topologies while maintaining energy-efficient operation for low-power IoT nodes [7].

Recently, algorithms based on traditional machine learning, including random forest [8] and support vector machine (SVM) [9], have been applied in IoT-enabled IDSs. However, these approaches face limitations in capturing spatiotemporal patterns from IoT device clusters, such as weak generalization across dynamic IoT network topologies and difficulty extracting deep semantic features from encrypted IoT communications. For resource-constrained IoT environments, deep learning-based IDSs have emerged as superior solutions, capable of processing multivariate sensor data streams by leveraging spatiotemporal feature extraction across distributed IoT nodes. This enhances learning capacity and generalization for evolving threat landscapes. Current advanced architectures include lightweight convolutional networks [10] for IoT traffic analysis, recurrent networks [11] modeling device behavior sequences, bidirectional long short-term memory networks (LSTMs) [12] detecting temporal anomalies in smart sensors, generative adversarial networks [13], graph networks [14] monitoring IoT device interdependencies, and sliding-mode observer method [15] applied to sensor fault diagnosis in critical IoT infrastructures (e.g., battery systems).

In IoT security contexts, this study first visualizes network intrusion data using heatmaps [16] to reveal feature correlations in IoT protocol payloads and employs the Pearson correlation coefficient [17] to identify device-type-specific attack indicators. A hybrid sampling strategy is applied to addressing the data imbalance prevalent in IoT event logs, where normal activities dominate. We propose an IoT-optimized intrusion detection system that integrates a residual-enhanced convolutional neural network (CNN) and bidirectional gated recurrent units (BiGRU) architecture with self-attention mechanisms. This architecture combines 1) residual connections for efficient feature propagation in edge computing scenarios; 2) CNNs for local pattern extraction from IoT packet payloads; 3) BiGRUs for bidirectional sequence modeling to capture IoT device state transitions; and 4) attention mechanisms for prioritizing critical features in heterogeneous IoT data streams. Evaluation on public datasets demonstrates that the model achieves superior detection performance than those of traditional methods, thereby showing higher detection accuracy and improved robustness.

The main contributions of this study are as follows:

1) IoT-aware feature engineering was implemented using heatmaps and Pearson's correlation coefficient to identify relationships between device-specific telemetry features and target variables. The ADASYN + Tomek Links hybrid sampling method was employed to address class imbalance in IoT event logs by synthesizing minority attack

samples (e.g., rare IoT botnet activities), removing noise from heterogeneous sensor data, and balancing IoT security datasets dominated by normal device behavior.

2) A residual-enhanced CNN-BiGRU-based intrusion detection system (RGNIDSs) with a self-attention mechanism is proposed. This architecture combines efficient feature propagation and dynamic weighting across different modules to accurately identify and classify potential attacks in the network.
3) Experiments were conducted on the CSE-CIC-IDS2018 and NSL-KDD datasets, and the findings indicate that the proposed approach outperformed most existing methods regarding performance. Ablation studies confirmed the effectiveness of each system module.

2. **Related Works.** Network intrusion detection (NID) techniques have undergone several technological evolution since their inception in the early 1980s. Early methods in the pre-IoT era primarily relied on rule matching and feature comparison, using static analysis of homogeneous network traffic to detect known attack patterns. However, these approaches are inadequate for modern IoT-driven heterogeneous networks, particularly when facing zero-day attacks [18] targeting vulnerable IoT protocols [19] or device-specific variants attacks in smart ecosystems. Their effectiveness is further reduced by the dynamic device behavior of IoT devices and the prevalence of encrypted payloads.

To address IoT-induced data imbalance in multiclass scenarios, oversampling [20] and undersampling [21] methods have been developed to adjust both the distribution of sample quantities across classes and the representation of minority classes. In the 1990s, traditional sampling methods were extensively adopted. Traditional oversampling equalizes class distribution by replicating minority class samples (e.g., rare IoT edge device intrusions), whereas traditional undersampling reduces the number of majority class samples by randomly removing them (e.g., standard sensor telemetry records). In 2002, Chawla et al. proposed the synthetic minority over-sampling technique (SMOTE) [22] and the clustered center-of-mass undersampling method. SMOTE enhances the minority class by creating new synthetic samples rather than merely duplicating the existing ones. The clustered center-of-mass method reduces the majority class size by grouping samples into clusters and using cluster centers as representatives. In 2008, He et al. proposed adaptive synthetic sampling (ADASYN) [23], which focuses on generating additional synthetic samples in regions where the minority class is difficult to classify.

The primary challenge in IoT-based intrusion detection is automated feature extraction from heterogeneous device communications. Conventional machine-learning techniques rely on manual feature engineering, which is ineffective for encrypted IoT payloads. Vinayakumar et al. [24] introduced a CNN-based model that significantly improved detection accuracy by extracting relevant features from traffic through convolutional layers. However, CNNs focus on local features and struggle to capture long-range dependencies in IoT device-to-cloud communication sequences – a critical aspect for industrial IoT anomaly detection. To address this, LSTM-based approaches model temporal dependencies across smart device behavioral logs. Laghrissi et al. [25] proposed an LSTM-based model capable of effectively identifying attacks in traffic, thereby achieving strong performance on multiple datasets. Furthermore, bidirectional LSTM (BiLSTM) networks enhance the sequence modeling by processing information in both forward and backward directions. Imrana et al. [26] proposed a BiLSTM-based IDS that achieved better performance in classifying attack types. Jouhari and Guizani [27] introduced a lightweight model that integrated CNN and BiLSTM, thereby significantly reducing computational overhead without compromising detection accuracy. In 2021, Seo et al. [28] presented an intrusion detection model that combines a multi-head self-attention mechanism with

LSTM, which can effectively enhance the detection of intricate attack patterns. The attention mechanism dynamically assigns weights to input features to enhance the model focus on the most important ones. In 2022, Xiao et al. [29] proposed a multisensor data fusion method based on a graph convolutional network (GCN) combined with LSTM. This approach captured spatial correlations among sensors through the GCN and modeled temporal dependencies using LSTM which significantly enhanced the modeling of spatiotemporal features. In 2024, Ullah et al. [30] proposed a model combining a transformer, CNN, and LSTM, which utilizes transformer-based transfer learning to extract semantic features from network traffic. These features are then used by the LSTM-CNN model to identify various types of attacks by deep learning representations.

Building on the concepts mentioned above, this study proposes a residual-enhanced CNN-BiGRU-based intrusion detection system (RGNIDS) that integrates a CNN-BiGRU with a self-attention mechanism. The system leverages residual connections to facilitate efficient information flow between modules, while combining the spatiotemporal feature extraction capabilities of CNN and BiGRU. The self-attention mechanism further enhances classification by assigning dynamic weights to salient features in the input data. This integration improves detection accuracy while retaining computational efficiency. A comparison between previous methods and the proposed model is presented in Table 1.

Table 1. Comparison of existing network intrusion detection model methods

Year	**Methods**	**Datasets**	**Balancing methods**	**Limitations**
2017	CNN-based	KDDcup99, NSL-KDD	–	Limited sequential modeling, Poor handling of temporal dynamics
2021	MHSE	IDS-2012, IDS-2017	–	Insufficient feature selection, Strong data dependence
2022	BiLSTM	NSL-KDD	–	Data imbalance, Limited to NSL-KDD
2024	IDS-INT	UNSW-NB15 CIC-IDS2017 NSL-KDD	SMOTE	High computational cost, Large data requirements, Overfitting risk
2023	CBF-IDS [31]	UNSW-NB15 CIC-IDS2017 NSL-KDD	Focal loss	Long training time, Poor interpretability
2024	CNN-BiLSTM	UNSW-NB15	Weighted loss	Sensitive to long sequences
2025	Ours	NSL-KDD CIC-IDS2018	ADASYN+ Tomek Links	–

3. **Proposed Model.**

3.1. **Feature extraction module based on heatmap and Pearson correlation coefficient (PHM block).**

The PHM module integrates Pearson correlation analysis with heatmap visualization to extract more informative features by modeling both statistical relationships and spatial patterns among input variables. First, it uses Pearson's correlation coefficient to compute the linear relationships between feature pairs and identify those with strong or weak associations. These correlations are then visualized using heatmaps, which enabled features with stronger correlations to be highlighted in the visualization.

This process provides rich and interpretable insights into the intrinsic statistical structure of the feature space and highlights potential synergistic or redundant feature interactions. In contrast to conventional feature extraction methods, PHM explicitly emphasizes foundational statistical dependencies. By filtering and prioritizing features based on their mutual correlations, PHM produces a refined feature set that effectively captures underlying data relationships, thereby facilitating subsequent modeling stages and contributing to a more stable and efficient model performance in complex tasks.

3.2. **Res-CNN module combining self-attention fusion.** The residual-enhanced CNN module integrates edge-optimized depthwise separable convolutions [32] with residual connections and self-attention mechanisms tailored for IoT protocol analysis. The architecture first processes encrypted IoT payloads using channel-wise spatial convolutions to target protocol-specific headers (e.g., MQTT control packets), followed by pointwise convolutions to fuse cross-channel dependencies in heterogeneous IoT communication patterns. Skip connections are added between each convolutional layer to directly combine input and output feature maps to alleviate the gradient vanishing problem [33]. To prioritize IoT attack patterns, a multi-head self-attention [34] layer analyzes contextual relationships in device-to-cloud sessions by computing attention weights across temporal and spatial IoT traffic dimensions. This mechanism intensifies the focus on protocol anomaly transitions (e.g., abrupt CoAP-to-HTTP shifts) and encrypted payload irregularities, while suppressing redundant sensor telemetry noise.

This dynamic context-aware prioritization fundamentally distinguishes the module from static correlation analysis. In contrast to conventional CNNs, it employs lightweight batch normalization and spatial pooling to mitigate overfitting on sparse IoT attack signatures, thereby ensuring effective convergence across diverse IoT ecosystem datasets. The synergistic architecture demonstrates particular efficacy in the real-time analysis of 6LoWPAN fragmentation attacks and detection of masked industrial IoT command injections, which establish foundational capabilities for adaptive edge security in smart city deployments.

3.3. **Bi-directional gated recurrent unit (BiGRU) [35] module for self-attention mechanism synergy.** The BiGRU model is a form of recurrent neural network that comprises two separate GRU units [36]: one processes data in the forward direction of the time sequence, and the other processes it in reverse. This bidirectional architecture enables the BiGRU model to capture both past and future contexts within sequence data, thereby improving its ability to analyze and predict temporal patterns. The GRU regulates information flow using update and reset gate, along with a new candidate activation to manage the flow of information.

To better capture the temporal dependencies in sequential data, we employed a combined modeling approach that integrates a BiGRU with a self-attention mechanism. This integration enables the module to learn temporal relationships while dynamically assigning attention across different time steps during time-series processing. In a standard BiGRU module, the forward and reverse networks process the sequence independently in opposite directions. However, the self-attention mechanism allows the model to focus automatically on the most relevant time points for the current task by adjusting the weights for each moment in the sequence. For IoT attack sequence analysis, the module processes time-aligned features from Res-CNN outputs, transforming spatial patterns into context-aware temporal representations. A self-attention layer subsequently weights critical phases in IoT communication cycles, intensifying focus on anomalous temporal segments such as sudden protocol handshake failures or industrial control command bursts.

This synergy enables the dynamic prioritization of attack-indicative intervals (e.g., botnet activation phases in smart meters) while suppressing benign periodic sensor updates.

It enhances the model's ability to capture meaningful time-series patterns and improves the extraction of long-term dependencies. By incorporating a self-attention layer after the BiGRU output, the model can assign different weights to each time step according to its relevance, thereby allowing it to focus more precisely on critical temporal features. This adaptive weighting mechanism contributes to improved accuracy in detecting complex scanning behaviors.

3.4. **Synergy and joint optimization between modules.** First, the input data are processed using a Res-CNN module with self-attentive fusion. This module introduces residual concatenation with depthwise separable convolutions. Through channel-wise convolution, each feature channel is processed independently, followed by pointwise convolution to fuse information across channels. The inclusion of a self-attention mechanism allows feature maps obtained by convolution to be weighted according to their importance, thereby effectively highlighting critical local features and suppressing redundant or irrelevant information. Batch normalization [37] and max-pooling [38] operations ensure stable feature scaling for fluctuating IoT data rates, which is critical in smart-city sensor networks with heterogeneous device sampling frequencies.

Next, the feature maps generated by the Res-CNN module are fed into the BiGRU module, which captures the temporal dependencies by efficiently processing present and past information through its bidirectional structure. The BiGRU output is then passed to a multi-head attention mechanism that operates in parallel across different heads of attention to learn the dependencies between the positions in the sequence, thus capturing the key temporal patterns in the input sequence in a more fine-grained manner. In the residual-enhanced CNN and BiGRU modules, the batch normalization and dropout [39] were jointly utilized to further enhance the model's training process.

Finally, the model fuses and classifies the information using a fully connected layer followed by the output layer. Intermediate features are first regularized through dropout, then mapped to the final classification space via a dense layer. The classification layer features a lightweight, fully connected architecture optimized for resource-constrained gateway hardware, which enables simultaneous detection of protocol-layer vulnerabilities and application-layer data leaks through adaptive fusion of IoT spatiotemporal features. A flow diagram of the overall structure is shown in Figure 1.

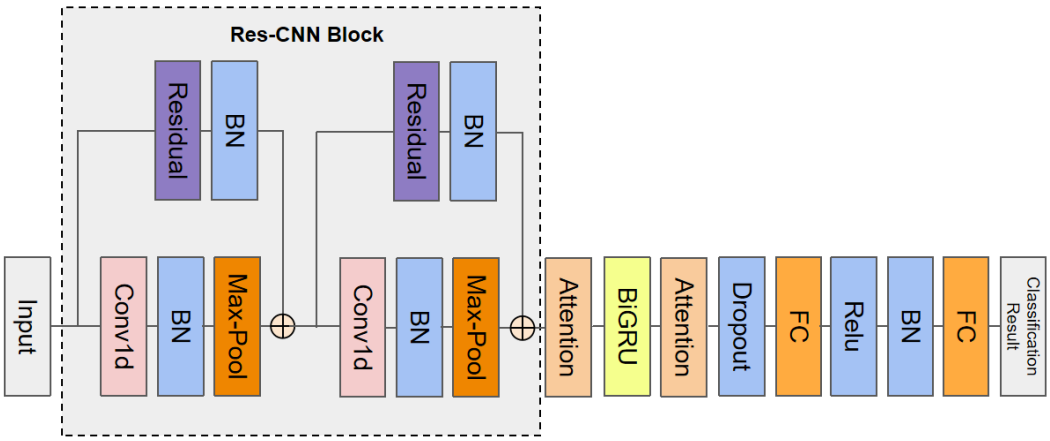


FIGURE 1. The structure of RGNIDS

4. Experiments.

4.1. Datasets.

The NSL-KDD [40] dataset is an updated version of the extensively recognized KDD99 [41] dataset and contains four distinct attack categories: probing, U2R, R2L, and DoS. Specifically, DoS attacks (e.g., Syn Flood) disrupts service availability, whereas probing attacks (e.g., port scanning) focus on network reconnaissance. U2R exploits privilege escalation vulnerabilities through techniques such as buffer overflow, and R2L attacks (e.g., password guessing) attempt unauthorized remote access. The dataset comprises four subsets, with the complete training set containing 125,973 entries and the test set containing 22,444 entries. Each record includes 43 features: 41 traffic characteristics (e.g., duration, protocol type, and flag patterns). The label distribution shows significant class imbalance – DoS dominates 36.46% of training samples, whereas U2R and R2L collectively represent $< 1\%$. The final two attributes denote attack labels and severity levels. The number and proportion of dataset categories are shown in Table 2.

Table 2. NSL-KDD dataset composition

Dataset	Quantity and proportion	DoS	Normal	Probe	R2L	U2R
KDDTrain+	Number	45927	67343	11656	995	52
	scale	36.46%	53.46%	9.25%	0.79%	0.04%
KDDTest+	Number	7458	9711	2421	2654	200
	scale	33%	43%	11%	12.1%	0.9%

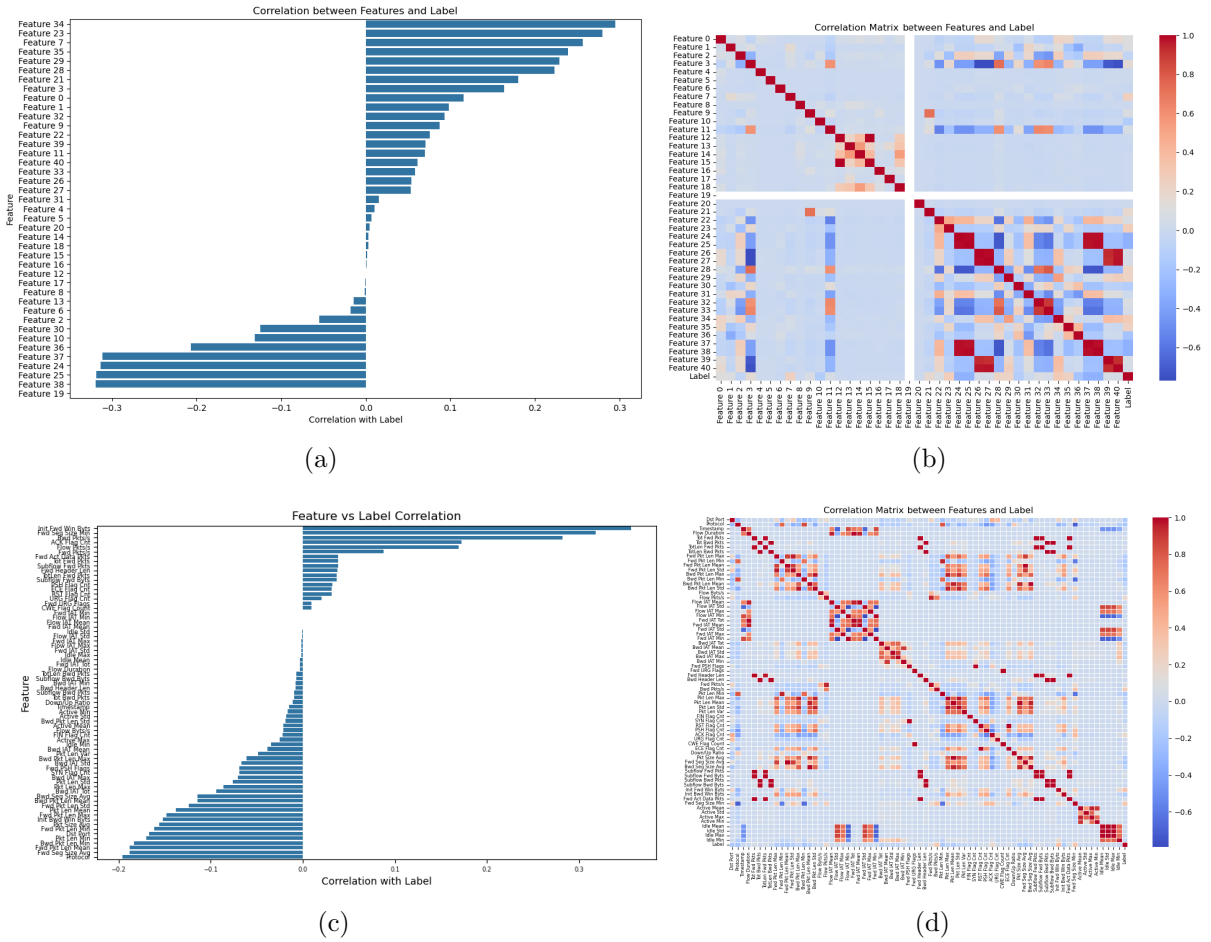
The CSE-CIC-IDS2018 [42] dataset was constructed under the supervision of the Canadian Institute for Cybersecurity. It captures five consecutive days of network traffic data by simulating multiple cyberattacks alongside normal traffic in a controlled environment. The dataset contains 16,232,955 records, including 13,484,708 normal traffic entries, which account for 83.07% of the total. In this study, 10% of the dataset was allocated to the training set, whereas 2% was used for validation and another 2% for testing. The number and proportion of data categories are listed in Table 3.

Table 3. CSE-CIC-IDS2018 dataset composition

Category	Total size	Total rate	Train size	Test size
Benign	13484708	83.07%	1338811	267762
Bot	286191	1.76%	28524	5705
DoS attacks-Hulk	461912	2.85%	46028	9205
DoS attacks-SlowHTTPTest	139890	0.86%	13976	2795
Brute Force-Web	611	0.004%	64	13
Brute Force-XSS	230	0.001%	30	6
SQL Injection	87	0.001%	10	2
DDoS attacks-LOIC-HTTP	576192	3.55%	57891	11578
Infilteration	161934	1%	15950	3190
DoS attacks-GoldenEye	41508	0.26%	4136	827
DoS attacks-Slowloris	10990	0.07%	1113	223
SSH-Bruteforce	187589	1.16%	18773	3755
FTP-Bruteforce	193360	1.19%	19443	3889
DDoS attacks-HOIC	686023	4.23%	68767	13753
DDoS attacks-LOIC-UDP	1730	0.01%	189	38

4.2. **Data processing.** The NSL-KDD and CSE-CIC-IDS2018 datasets underwent standard preprocessing steps, including data cleaning to address missing values and remove outliers, one-hot encoding [43] of categorical features (e.g., protocol type), and min-max normalization to scale continuous numerical features to the $[0, 1]$ range [44]. For dataset-specific handling, the NSL-KDD dataset required no additional steps beyond standard preprocessing, whereas the CSE-CIC-IDS2018 dataset involved converting timestamps to numerical values and removing irrelevant traffic identifier columns to prevent overfitting.

4.3. **Feature weighting.** Using the PHM module, we quantified the linear correlation between each feature and the target label using Pearson's coefficient. This analysis determined the features most relevant to intrusion detection, as illustrated through heatmap visualization and Pearson correlation coefficient results shown in Figure 2. Based on these results, higher weights were assigned to features with a strong correlation during the model training process. This approach prioritizes key discriminative features to improve the accuracy and robustness of intrusion classification.



(a) (b)

(c) (d)

FIGURE 2. (a) Pearson correlation coefficient for the NSL-KDD dataset; (b) feature correlation heatmap for the NSL-KDD dataset; (c) Pearson correlation coefficient for the CSE-CIC-IDS2018 dataset; (d) feature correlation heatmap for the CSE-CIC-IDS2018 dataset

4.4. **Mixed sampling method selection.** To address the significant class imbalance, we implemented a hybrid sampling strategy that combines oversampling and undersampling techniques. First, ADASYN was used to oversample minority classes by generating

synthetic samples based on their density distribution. ADASYN emphasizes regions near the decision boundary, where classification is more challenging. Subsequently, Tomek Links [45] was applied for boundary cleaning. This technique identifies and removes majority class samples that form ambiguous pseudo-nearest neighbors with minority samples, thereby reducing noise and clarifying class boundaries in overlapping regions. This hybrid approach – adaptive oversampling followed by boundary-aware undersampling – balances the class distribution while improving data separability and enhancing model generalization on complex patterns. The pre-label classification sampling pairs for the NSL-KDD dataset are shown in Figure 3. The pre-label classification sampling pairs for the NSL-KDD dataset are shown in Figure 3, while those for the CSE-CIC-IDS2018 dataset are presented in Figure 4.

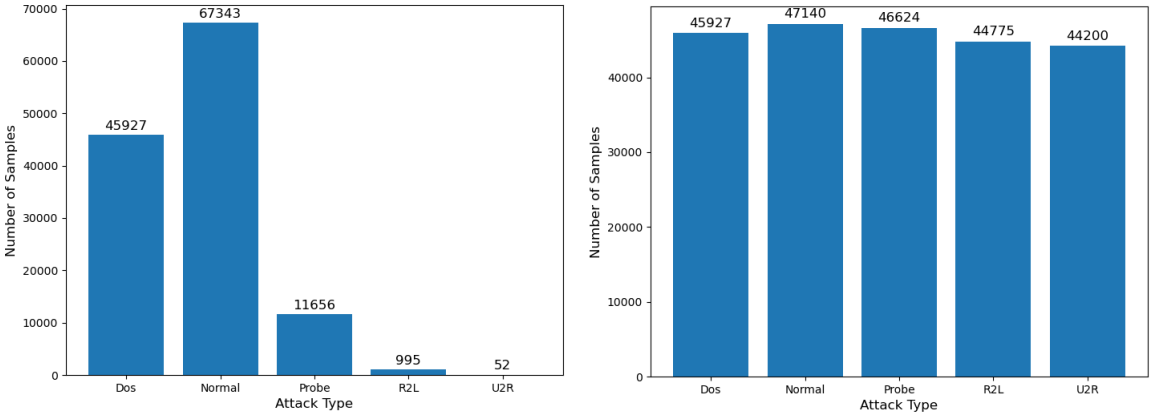


FIGURE 3. Comparison of the number of categories in the NSL-KDD dataset before and after ADASYN+Tomek Links sampling

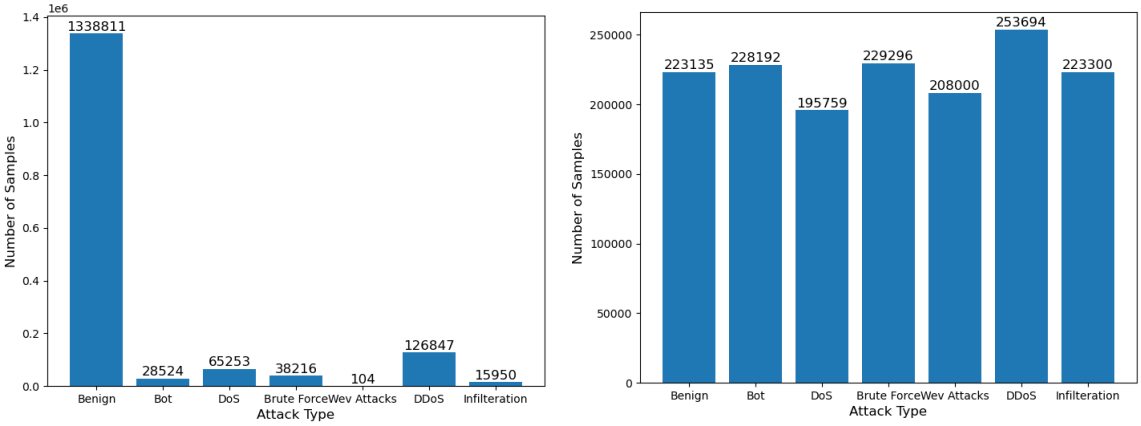


FIGURE 4. Comparison of the number of categories in the CSE-CIC-IDS2018 dataset before and after ADASYN+Tomek Links sampling

4.5. **Selection of experimental hyperparameters and assessment indicators.** In this study, the model training parameters were set as follows: the cross-entropy loss function [46] was used for training process, and the Adam optimizer was selected. The number of training epochs was set to 50, with a batch size of 256. A learning rate of 0.001 was adopted using a cosine annealing strategy, with a 50% decay every 10 cycles, thus enabling the optimizer to converge more accurately.

The assessment indicators in this study were calculated using the confusion matrix [47] presented in Table 4.

Table 4. Confusion matrix of network intrusion detection

Confusion matrix		**Predicted value**	
		Normal	**Attack**
True value	Normal	TN	FP
	Attack	FN	TP

In Table 4, TP is the count of samples where both the predicted and true values are attacks; FN denotes the number of samples whose predicted values are normal but whose true values are attacks; FP is the number of samples where the predicted values are attacks but whose true values are normal; and TN represents the number of samples whose predicted and true values are both normal. Accuracy, Precision, Recall, and F1-score were used to evaluate the classification performance of the model. The calculation formulas are given in Equations (1)-(4):

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} * 100\% \tag{1}$$

$$Pre = \frac{TP}{TP + FP} * 100\% \tag{2}$$

$$Recall = \frac{TP}{TP + FN} * 100\% \tag{3}$$

$$F1 = \frac{2 * Pre * Recall}{Pre + Recall} * 100\% \tag{4}$$

4.6. **Experimental results.** Table 5 and Table 6 provide the main experimental results for the NSL-KDD and CSE-CIC-IDS2018 datasets. The performance of the proposed model in the classification task is compared with that of other models under identical experimental conditions.

Binary and multi-class classification experiments were conducted on the NSL-KDD dataset, which comprises 41 features and 23 subcategories of attack types. These subcategories were consolidated into four main attack categories: DoS, R2L, U2R, and probing. The experimental hyperparameters were outlined in the previous section. The proposed RGNIDS model was compared against several state-of-the-art models from recent literature. As shown in Table 5 and Figure 5, RGNIDS outperforms the comparison models across all four evaluation metrics in both binary and multi-class classification tasks.

To further evaluate the advantages of our model in multiclass intrusion detection tasks, we evaluated the area under the curve receiver operating characteristic (AUC-ROC) for

Table 5. Performance comparison between RGNIDS and other models on NSL-KDD dataset

	Accuracy	**Precision**	**Recall**	**F1**
GMM-WGAN [48]	84.65%	85.13%	84.65%	83.95%
XGB [49]	95.54%	92.61%	95.54%	93.41%
IDS-INT [30]	98.45%	98.00%	99.00%	98.00%
Res-TranBiLSTM [50]	90.99%	91.39%	90.94%	90.89%
Transformer-based [51]	97.84%	97.95%	97.72%	97.83%
Ours	99.63%	99.62%	99.63%	99.62%

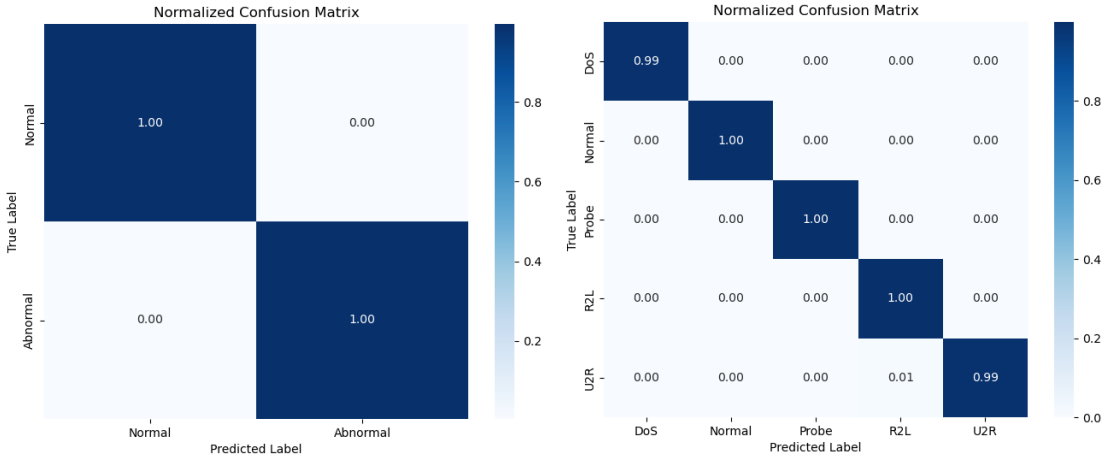


FIGURE 5. RGNIDS binary and multiclassification confusion matrices on the NSL-KDD dataset

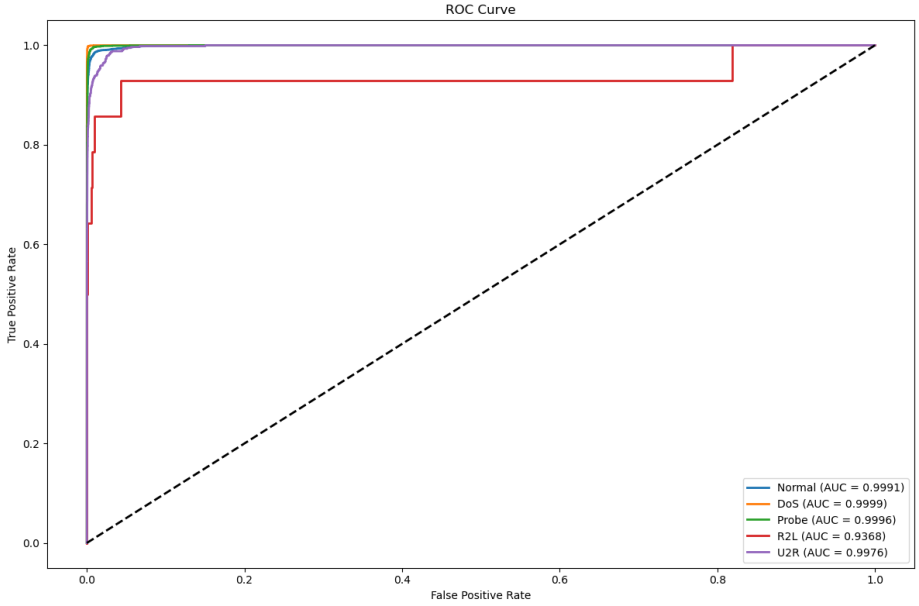


FIGURE 6. (color online) AUC-ROC scores for different classes in NSL-KDD dataset

each attack category on the NSL-KDD dataset. The AUC-ROC represents the trade-off between the true positive rate (TPR) and false positive rate (FPR) across various classification thresholds. The results, shown in Figure 6, demonstrate that our model achieves robust performance across the normal, DoS, probe, and U2R categories. Notably, for more challenging attack types such an R2L, the model exhibits significant performance improvements than those of other existing methods. The results indicate that the proposed model outperforms other models in terms of enhancing intrusion detection performance and plays a crucial role in enhancing the discrimination capability of intrusion detection systems against network attacks.

Samples were extracted and processed from the CSE-CIC-IDS2018 dataset. Given the large size and significant class imbalance within the dataset, proportional sampling was conducted for each file. Owing to the large difference in the number of categories in this dataset, we used ADASYN + Tomek Links for mixed sampling. The model performance

comparison and confusion matrix of RGNIDS in binary and multicategorization tasks are shown in Table 6 and Figure 7. Notably, the RGNIDS model is more balanced and better than the other models in terms of the four-evaluation metrics Acc, Pre, Recall, and F1. As described in the previous section, we plotted the AUC-ROC curves for each category of the CSE-CIC-IDS2018 dataset. As shown in Figure 8, the model maintained an excellent performance under most features.

TABLE 6. Performance of RGNIDS and other models on CSE-CIC-IDS2018 dataset vs. other models

	Accuracy	**Precision**	**Recall**	**F1**
CNN+LSTM	99.00%	98.00%	99.00%	92.00%
U-Net [52]	97.77%	97.94%	97.53%	97.73%
HPO+DNN [53]	95.79%	95.38%	95.79%	95.11%
Novel CNN [54]	97.20%	99.10%	97.20%	95.30%
HDLNIDS [55]	98.90%	98.60%	99.16%	98.83%
Ours	99.16%	99.07%	99.16%	98.85%

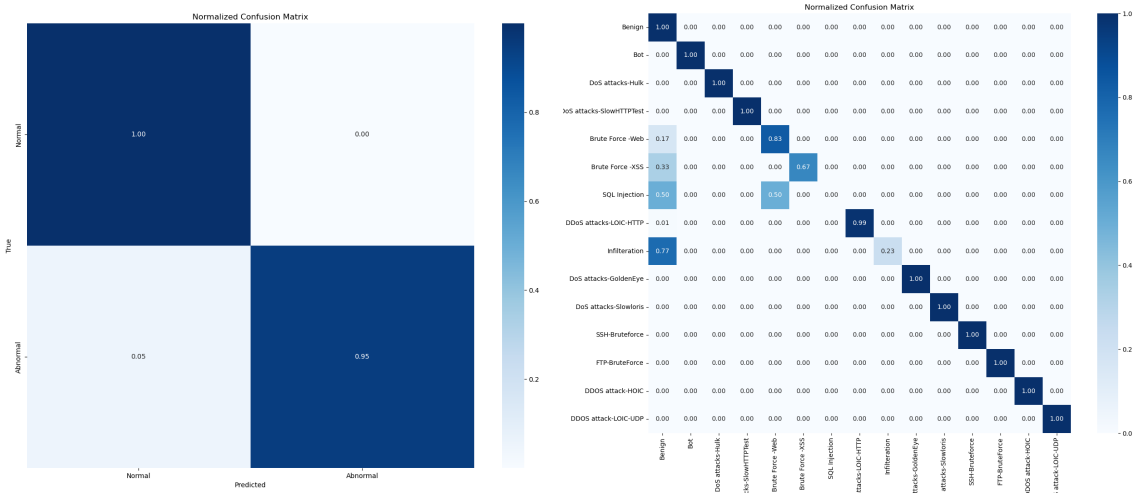


FIGURE 7. RGNIDS binary and multiclassification confusion matrices on the CSE-CIC-IDS2018 dataset

4.7. **Training efficiency analysis.** To evaluate the proposed RGNIDS model, we compared its training time with several existing benchmark models. Under identical hardware and software configurations, we recorded the training time per epoch on the NSL-KDD dataset. As shown in Figure 9, the RGNIDS model required a significantly longer training time than the other models. This increased computational cost is primarily due to the model's complex architecture, which integrates multiple advanced components. This multifaceted approach, although more computationally intensive than simpler models, enables a deeper analysis of spatial patterns (e.g., packet payload anomalies) and temporal dynamics (e.g., sequence-based attack signatures), which significantly enhance detection performance. Notably, the training time of RGNIDS is shorter than that of transformer-based models, despite both incorporating complex sequence modeling and feature interaction mechanisms. This represents a carefully considered tradeoff: by leveraging deep spatiotemporal feature extraction and attention-based weighting, RGNIDS sacrifices training efficiency to achieve substantial improvements in prediction accuracy

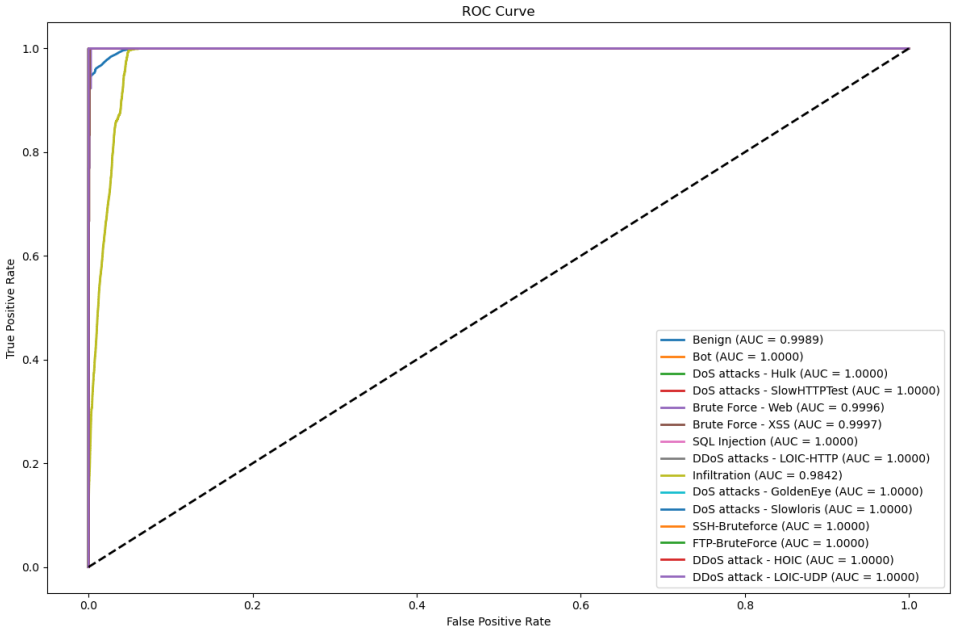


FIGURE 8. (color online) AUC-ROC scores for different classes in CSE-CIC-IDS2018 dataset

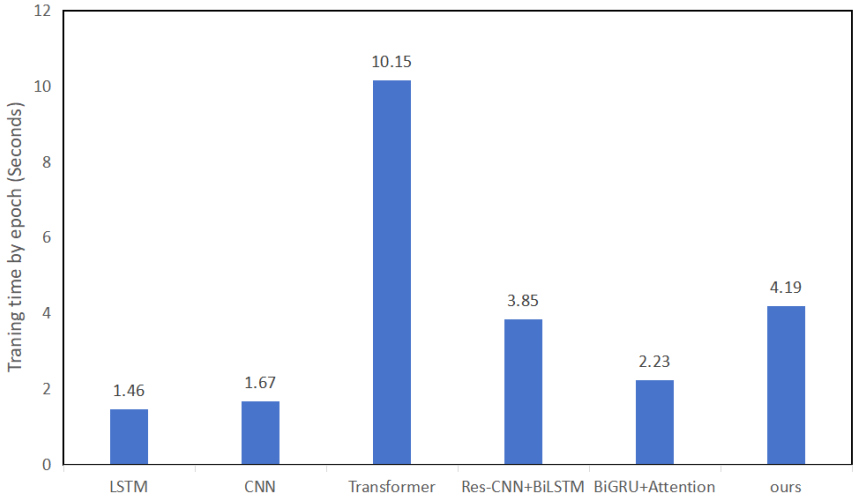


FIGURE 9. Comparison of training time per epoch on NSL-KDD dataset

and robustness, which are critical requirements for effective intrusion detection in complex IoT environments.

4.8. **Ablation experiments.** We propose RGNIDS, a residual-enhanced CNN-BiGRU based NIDS incorporating a self-attention mechanism. To evaluate the contribution of each module, we conducted ablation experiments on the NSL-KDD dataset using a transformer + CNN + LSTM model as the baseline. The full model without ablation is denoted as Res-CNN+BiGRU+Attention. We used the Adam optimizer and trained 50 epochs with a learning rate that varied throughout training. We compared the training efficiency and evaluation metrics of different model variants: 1) CNN + BiGRU, 2) CNN + Attention, and 3) BiGRU + Attention, as shown in Table 7. The results demonstrate that each component of RGNIDS contributes positively to the overall performance. Specifically, CNN and attention primarily assess the importance of feature regions, whereas BiGRU

TABLE 7. Performance of different modules of RGNIDS on NSL-KDD dataset

	Accuracy	**Precision**	**Recall**	**F1**
Transformer+CNN+LSTM	98.45%	98.00%	99.00%	98.00%
Res-CNN+BiGRU	97.96%	97.92%	97.96%	97.86%
Res-CNN+Attention	96.22%	96.02%	96.22%	95.81%
BiGRU+Attention	98.24%	98.24%	98.24%	98.20%
Res-CNN+BiGRU+Attention	99.63%	99.62%	99.63%	99.62%

effectively captures contextual temporal information, which is then weighted by attention to generate the final classification.

5. **Conclusion and Future Work.** This study presents an IoT-aware intrusion detection framework that addresses the critical security challenges in interconnected smart ecosystems. By integrating heatmap visualization with Pearson correlation analysis tailored to IoT protocol heterogeneity, our method effectively identifies discriminative features across encrypted MQTT/CoAP traffic and device behavioral patterns. The ADASYN + Tomek Links hybrid sampling strategy demonstrates particular efficacy in balancing industrial IoT datasets dominated by routine operational data, successfully mitigating false negatives for stealthy edge device compromises while preserving temporal integrity in smart grid communication sequences. The proposed RGNIDS architecture, which synergizes residual-enhanced CNNs with bidirectional GRUs and self-attention mechanisms, achieves robust performance across both the legacy NSL-KDD dataset (simulating IoT protocol reconnaissance) and the modern CSE-CIC-IDS2018 dataset (resembling cross-layer IoT ecosystem breaches). It demonstrates adaptability to resource-constrained fog computing environments through optimized feature processing pipelines. However, several areas remain for improvement in future research. In IoT network security, existing methods require improvements in the dynamic adaptability of data preprocessing and feature extraction. This includes using dynamic feature weighting to handle time-varying attack patterns by optimizing the processing of high-dimensional sparse data, designed targeted sampling strategies informed by domain knowledge to reduce false samples. Additionally, enhancing multimodal feature fusion by integrating device metadata, physical-layer signals, and network behavior data is crucial for capturing multidimensional attack features. Finally, realizing collaborative optimization of preprocessing and feature learning through an end-to-end framework can improve data quality and model robustness.

In the future, we plan to optimize the model by reducing its complexity while enhancing performance to improve computational efficiency. We will explore the integration of transfer learning (TL) techniques to enable faster adaptation and enhanced detection accuracy for novel attacks. Additionally, we plan to implement a multimodal data-fusion approach that combines diverse data sources to improve identification of complex attack patterns. Finally, we aim to develop an end-to-end framework that integrates preprocessing and feature learning, thereby enabling joint optimization to improve data quality and overall model robustness.

Acknowledgment. We sincerely thank Renguang Zheng, Jiawen Yang and Xinyi Wu for their valuable contributions to this work. Mr. Zheng provided ingenious ideas in this research, and Mr. Yang offered important insights and feedback throughout the research process, significantly enhancing the depth and quality of this study. Ms. Wu's assistance in data analysis and methodology is indispensable. Their support is crucial to ensuring the robustness of our research results. We are extremely grateful for their time, efforts

and professional knowledge. They have made a tremendous contribution to the success of this article. This research was funded by Research on Key Technologies for Intelligent Diagnosis of Power Information Network Security (No. 521350240008).

REFERENCES

[1] A. S. Ashoor and S. Gore, Importance of intrusion detection system (IDS), *International Journal of Scientific and Engineering Research*, vol.2, no.1, pp.1-4, 2011.

[2] S. Rubin, S. Jha and B. P. Miller, Automatic generation and analysis of NIDS attacks, *The 20th Annual Computer Security Applications Conference*, pp.28-38, 2004.

[3] P. Deshpande, S. C. Sharma, S. K. Peddoju and S. Junaid, HIDS: A host based intrusion detection system for cloud computing environment, *International Journal of System Assurance Engineering and Management*, vol.9, pp.567-576, 2018.

[4] T. Peng, C. Leckie and K. Ramamohanarao, Survey of network-based defense mechanisms countering the DoS and DDoS problems, *ACM Computing Surveys (CSUR)*, vol.39, no.1, 2007.

[5] A. Cui, M. Costello and S. J. Stolfo, When firmware modifications attack: A case study of embedded exploitation, *NDSS*, vol.1, 2013.

[6] F. Galtier, R. Cayre, G. Auriol, M. Kaâniche and V. Nicomette, A PSD-based fingerprinting approach to detect IoT device spoofing, *2020 IEEE 25th Pacific Rim International Symposium on Dependable Computing (PRDC)*, pp.40-49, 2020.

[7] K. K. S. Liyakat, Machine learning approach using artificial neural networks to detect malicious nodes in IoT networks, *International Conference on Machine Learning, IoT and Big Data*, pp.123-134, 2023.

[8] L. Breiman, Random forests, *Machine Learning*, vol.45, pp.5-32, 2001.

[9] M. Mohammadi, T. A. Rashid, S. H. T. Karim, A. H. M. Aldalwie, Q. T. Tho, M. Bidaki, A. M. Rahmani and M. Hosseinzadeh, A comprehensive survey and taxonomy of the SVM-based intrusion detection systems, *Journal of Network and Computer Applications*, vol.178, 102983, 2021.

[10] P. Sun, P. Liu, Q. Li, C. Liu, X. Lu, R. Hao and J. Chen, DL-IDS: Extracting features using CNN-LSTM hybrid network for intrusion detection system, *Security and Communication Networks*, 8890306, 2020.

[11] C. Yin, Y. Zhu, J. Fei and X. He, A deep learning approach for intrusion detection using recurrent neural networks, *IEEE Access*, vol.5, pp.21954-21961, 2017.

[12] S. Siami-Namini, N. Tavakoli and A. S. Namin, The performance of LSTM and BiLSTM in forecasting time series, *2019 IEEE International Conference on Big Data (Big Data)*, pp.3285-3292, 2019.

[13] E. Seo, H. M. Song and H. K. Kim, GIDS: GAN based intrusion detection system for in-vehicle network, *2018 16th Annual Conference on Privacy, Security and Trust (PST)*, pp.1-6, 2018.

[14] X. Deng, J. Zhu, X. Pei, L. Zhang, Z. Ling and K. Xue, Flow topology-based graph convolutional network for intrusion detection in label-limited IoT networks, *IEEE Transactions on Network and Service Management*, vol.20, no.1, pp.684-696, 2022.

[15] D. Xu, Y. Ma, W. Yang, T. Pan and Z. Dou, Sliding mode observer-based sensor fault diagnosis for lithium-ion battery packs, *International Journal of Innovative Computing, Information and Control*, vol.19, no.5, pp.1455-1470, 2023.

[16] L. Wilkinson and M. Friendly, The history of the cluster heat map, *The American Statistician*, vol.63, no.2, pp.179-184, 2009.

[17] J. Benesty, J. Chen, Y. Huang and I. Cohen, Pearson correlation coefficient, *Noise Reduction in Speech Processing*, pp.1-4, 2009.

[18] L. Bilge and T. Dumitraş, Before we knew it: An empirical study of zero-day attacks in the real world, *Proc. of the 2012 ACM Conference on Computer and Communications Security*, pp.833-844, 2012.

[19] J. Wang, S. Lu and C. Li, Uneven clustering routing protocols for multi-hop cognitive radio sensor networks: General design principles and an illustrative example, *International Journal of Innovative Computing, Information and Control*, vol.21, no.1, pp.153-172, 2025.

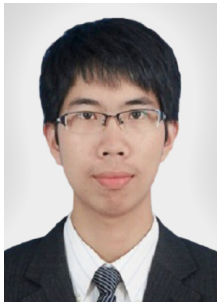
[20] J. C. Candy and G. C. Temes, Oversampling methods for A/D and D/A conversion, *Oversampling Delta-Sigma Data Converters*, pp.1-25, 1992.

[21] X.-Y. Liu, J. Wu and Z.-H. Zhou, Exploratory undersampling for class-imbalance learning, *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol.39, no.2, pp.539-550, 2008.

[22] N. V. Chawla, K. W. Bowyer, L. O. Hall and W. P. Kegelmeyerm, SMOTE: Synthetic minority over-sampling technique, *Journal of Artificial Intelligence Research*, vol.16, pp.321-357, 2002.
[23] H. He, Y. Bai, E. A. Garcia and S. Li, ADASYN: Adaptive synthetic sampling approach for imbalanced learning, *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, pp.1322-1328, 2008.
[24] R. Vinayakumar, K. P. Soman and P. Poornachandran, Applying convolutional neural network for network intrusion detection, *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp.1222-1228, 2017.
[25] F. Laghrissi, S. Douzi, K. Douzi and B. Hssina, Intrusion detection systems using long short-term memory (LSTM), *Journal of Big Data*, vol.8, no.1, 65, 2021.
[26] Y. Imrana, Y. Xiang, L. Ali and Z. Abdul-Rauf, A bidirectional LSTM deep learning approach for intrusion detection, *Expert Systems with Applications*, vol.185, 115524, 2021.
[27] M. Jouhari and M. Guizani, Lightweight CNN-BiLSTM based intrusion detection systems for resource-constrained IoT devices, *2024 International Wireless Communications and Mobile Computing (IWCMC)*, pp.1558-1563, 2024.
[28] S. Seo, S. Han, J. Park, S. Shim, H.-E. Ryu, B. Cho and S. Lee, Hunt for unseen intrusion: Multi-head self-attention neural detector, *IEEE Access*, vol.9, pp.129635-129647, 2021.
[29] B. Xiao, X. Xie and C. Yang, Multi-sensor data fusion based on GCN-LSTM, *International Journal of Innovative Computing, Information and Control*, vol.18, no.5, pp.1363-1381, 2022.
[30] F. Ullah, S. Ullah, G. Srivastava and J. C.-W. Lin, IDS-INT: Intrusion detection system using transformer-based transfer learning for imbalanced network traffic, *Digital Communications and Networks*, vol.10, no.1, pp.190-204, 2024.
[31] H. Peng, C. Wu and Y. Xiao, CBF-IDS: Addressing class imbalance using CNN-BiLSTM with focal loss in network intrusion detection system, *Applied Sciences*, vol.13, no.21, 2023.
[32] L. Dang, P. Pang and J. Lee, Depth-wise separable convolution neural network with residual connection for hyperspectral image classification, *Remote Sensing*, vol.12, no.20, 3408, 2020.
[33] S. Hochreiter, The vanishing gradient problem during learning recurrent neural nets and problem solutions, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol.6, no.2, pp.107-116, 1998.
[34] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser and I. Polosukhin, Attention is all you need, *Advances in Neural Information Processing Systems*, vol.30, 2017.
[35] J. Liu, Y. Yang, S. Lv, J. Wang and H. Chen, Attention-based BiGRU-CNN for Chinese question classification, *Journal of Ambient Intelligence and Humanized Computing*, pp.1-12, 2019.
[36] K. Cho, B. V. Merriënboer, D. Bahdanau and Y. Bengio, On the properties of neural machine translation: Encoder-decoder approaches, *arXiv Preprint*, arXiv: 1409.1259, 2014.
[37] S. Ioffe and C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, *International Conference on Machine Learning*, pp.448-456, 2015.
[38] J. Nagi, F. Ducatelle, G. A. D. Caro, D. Cireşan, U. Meier, A. Giusti, F. Nagi, J. Schmidhuber and L. M. Gambardella, Max-pooling convolutional neural networks for vision-based hand gesture recognition, *2011 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, pp.342-347, 2011.
[39] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov, Dropout: A simple way to prevent neural networks from overfitting, *The Journal of Machine Learning Research*, vol.15, no.1, pp.1929-1958, 2014.
[40] M. Tavallaee, E. Bagheri, W. Lu and A. A Ghorbani, A detailed analysis of the KDD CUP 99 data set, *2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications*, pp.1-6, 2009.
[41] D. D. Protić, Review of KDD Cup 99, NSL-KDD and KYOTO 2006+ datasets, *Vojnotehnički glasnik/Military Technical Courier*, vol.66, no.3, pp.580-596, 2018.
[42] J. L. Leevy and T. M. Khoshgoftaar, A survey and analysis of intrusion detection models based on CSE-CIC-IDS2018 big data, *Journal of Big Data*, vol.7, pp.1-19, 2020.
[43] P. Rodríguez, M. A. Bautista, J. Gonzalez and S. Escalera, Beyond one-hot encoding: Lower dimensional target embedding, *Image and Vision Computing*, vol.75, pp.21-31, 2018.
[44] P. J. M. Ali, R. H. Faraj, E. Koya, P. J. M. Ali and R. H. Faraj, Data normalization and standardization: A technical report, *Mach. Learn. Tech. Rep.*, vol.1, no.1, pp.1-6, 2014.
[45] M. Zeng, B. Zou, F. Wei, X. Liu and L. Wang, Effective prediction of three common diseases by combining SMOTE with Tomek links technique for imbalanced medical data, *2016 IEEE International Conference of Online Analysis and Computing Science (ICOACS)*, pp.225-228, 2016.

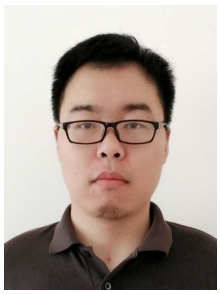
[46] P.-T. de Boer, D. P. Kroese, S. Mannor and R. Y. Rubinstein, A tutorial on the cross-entropy method, *Annals of Operations Research*, vol.134, pp.19-67, 2005.
[47] S. Visa, B. Ramsay, A. L. Ralescu and E. V. D. Knaap, Confusion matrix-based feature selection, *MAICS*, vol.710, no.1, pp.120-127, 2011.
[48] J. Cui, L. Zong, J. Xie and M. Tang, A novel multi-module integrated intrusion detection system for high-dimensional imbalanced data, *Applied Intelligence*, vol.53, no.1, pp.272-288, 2023.
[49] A. Srivastava, D. Sinha and V. Kumar, WCGAN-GP based synthetic attack data generation with GA based feature selection for IDS, *Computers & Security*, vol.134, 103432, 2023.
[50] S. Wang, W. Xu and Y. Liu, Res-TranBiLSTM: An intelligent approach for intrusion detection in the Internet of Things, *Computer Networks*, vol.235, 109982, 2023.
[51] U. C. Akuthota and L. Bhargava, Transformer based intrusion detection for IoT networks, *IEEE Internet of Things Journal*, 2025.
[52] A. Mezina, R. Burget and C. M. Travieso-González, Network anomaly detection with temporal convolutional network and U-Net model, *IEEE Access*, vol.9, pp.143608-143622, 2021.
[53] Y. N. Kunang, S. Nurmaini, D. Stiawan and B. Y. Suprapto, Attack classification of an intrusion detection system using deep learning and hyperparameter optimization, *Journal of Information Security and Applications*, vol.58, 102804, 2021.
[54] R. Selvam and S. Velliangiri, An improving intrusion detection model based on novel CNN technique using recent CIC-IDS datasets, *2024 International Conference on Distributed Computing and Optimization Techniques (ICDCOT)*, pp.1-6, 2024.
[55] E. U. H. Qazi, M. H. Faheem and T. Zia, HDLNIDS: Hybrid deep-learning-based network intrusion detection system, *Applied Sciences*, vol.13, no.8, 4921, 2023.

Author Biography

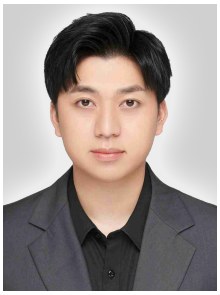


Kunsan Zhang graduated from Fuzhou University, China, in 2014, with a Bachelor's degree in Electrical Engineering and Automation. His main research interests include cyber security attack and defense, cyber security penetration, artificial intelligence and data analysis. Currently, he is working as a senior expert in State Grid Fujian Electric Power Co., Ltd. Zhangzhou Power Supply Company.

Song Zhang graduated from Wuhan University, China, in 2016, with a Master's degree in Electrical Engineering and Automation. His main research interests include network security management and security risk assessment. Currently, he is the deputy general manager of State Grid Fujian Electric Power Co., Ltd. Zhangzhou Power Supply Company.



Chaopeng Li graduated from the Institute of Acoustics of the Chinese Academy of Sciences, China, in 2019, and received a Ph.D. degree in Signal and Information Processing Engineering. His main research interests include deep learning theory, gradient optimization theory, time series data modeling and network security situation awareness. Currently, he serves as an associate professor and master's tutor at the School of Ocean Information Engineering, Jimei University, China.



Bingjie Xiang graduated from Jimei University, China, with a bachelor's degree in 2021. Currently, he is a master's student at Jimei University, China, specializing in Artificial Intelligence. His primary research interests encompass deep learning neural networks, model optimization algorithms, target recognition, and radar signal processing.

Jiachun Zheng graduated from Dalian Maritime University, China, with a bachelor's degree in 1986. He is currently a professor with the School of Ocean Information Engineering, Jimei University, China. He is also a master's supervisor and the director of Xiamen Key Laboratory of Marine Intelligent IoT Terminal Research and Development and Application; mainly engaged in intelligent information processing, marine traffic informatization, satellite communication navigation, intelligent IoT, and sensing research work.