University of Baghdad
College of Engineering
JOURNAL OF ENGINEERING

Journal of Engineering

journal homepage: www.jcoeng.edu.iq

Volume 32 Number 9 September 2026

SCDF-CNN-BiGRU: A Spatial-Contextual Dual Feature Fusion Framework for Multi dataset Hierarchical Network Intrusion Detection

Aseel M. Mahdi [1, *], **Haider K. Hoomod** [2]

[1] Scientific Research Commission, Ministry of Higher Education and Scientific Research, Baghdad, Iraq
[2] Computer Department, College of Education, Al-Mustansiriyah University, Baghdad, Iraq

ABSTRACT

Network Intrusion Detection Systems (NIDSs) are deployed and constitute an essential component of modern network defense against adversaries employing more and more advanced cyberattack strategies. Nevertheless, maintaining high detection accuracy in heterogeneous network environments still faces great challenges because of the diversity of traffic patterns and attack behaviors. This paper proposes a new Spatial–Contextual Dual-Feature Fusion (SCDF-CNN-BiGRU) framework in the context of hierarchical network intrusion detection employing multiple benchmark cybersecurity datasets. The proposed framework utilizes a lightweight shared convolutional neural network (CNN) feature extraction stage followed by 2 parallel and complimentary learning branches. The first branch employs Bidirectional Gated Recurrent Unit (BiGRU) to extract temporal dependencies and contextual relationships of network traffic while the second one adopts bidirectional convolutional processing to learn discriminative spatial representations. The features extracted from both branches are then fused via feature aggregation module to create an overall traffic representation. It also establishes hierarchical intrusion analysis through binary attack detection and multi-class attack categorization, along with an attack-level mapping strategy to connect different levels of attack types and the corresponding highly representative fine-grained attack descriptions. The framework was evaluated using six heterogeneous benchmark datasets, including UNSW-NB15, NSL-KDD, NF-ToN-IoT-v3, CICIOT2023, BCCC and CIC-UNSW-NB15. Experimental results showed strong detection performance with binary classification accuracies of up to 0.999 and multi-class accuracies of up to 0.988. These findings indicate that the proposed SCDF-CNN-BiGRU framework gives robust representations and consistent performance across multiple heterogeneous datasets via unified learning and ensemble inference.

Keywords: Network intrusion detection system (NIDS), Spatial-contextual feature fusion, CNN-BiGRU, Hierarchical attack classification, Multi-dataset learning.

*Corresponding author
Peer review under the responsibility of University of Baghdad.
https://doi.org/10.31026/j.eng.2026.09.12

Article received: 15/06/2026
Article revised: 28/08/2026
Article accepted: 28/08/2026
Article published: 01/09/2026

1. INTRODUCTION

Network intrusion detection systems (NIDSs) are essential for monitoring network traffic, detecting malicious activities, unauthorized access, malware infections, denial-of-service attacks, etc. **(Liu and Xu, 2025).** As a consequence, a current highly active area of investigation is machine learning and deep learning techniques for designing intelligent intrusion detection systems that have ability to learn complex behaviors of attacks directly from raw network traffic **(Adefemi, 2025)**. CNNs naturally learn discriminative spatial features from the traffic **(Imrana et al., 2024)** and GRU-based architectures implicitly learn the temporal and contextual dependencies **(Sagu et al., 2025)**. Similar observations were reported by **(Hussein et al., 2026)**. Nevertheless, the majority of existing intrusion detection approaches are dataset-specific, and therefore have a limited generalizability across heterogeneous settings **(Mishra et al., 2026)**. Moreover, existing studies mostly focus on the attack detection and classification performance, with many studies not paying sufficient attention to providing structured, systematic and interpretable representations that relate detected categories of attacks to informative fine-grained descriptions of the attacks for security analysis and response.

To overcome these limitations, a novel Spatial–Contextual Dual-Feature Fusion framework, termed SCDF-CNN-BiGRU is proposed. The framework contains a shared CNN feature extraction stage followed by the learning of two complementary branches. The proposed SCDF-CNN-BiGRU framework combines spatial-contextual dual-feature fusion with hierarchical intrusion classification and attack-level mapping to improve representation learning and attack interpretability across heterogeneous datasets.

(Imrana et al., 2024) introduced the CNN-GRU-FF model , which uses a double layer feature fusion technique, achieving high intrusion detection performance on the NSL-KDD and UNSW-NB15 datasets. Likewise, **(Ma et al., 2025)** propose DenseNet1D-BiGRU architectures for IoT intrusion detection that combines dense convolutional feature extraction with bidirectional recurrent learning to improve spatiotemporal representation ability. **(Kavitha and Harini, 2026)** proposed a Dual-path architecture, while **(Xu et al., 2023)** proposed a Hierarchical Deep Learning intrusion detection model that incrementally identify attack categories. However, the framework was evaluated on a single dataset without consideration of unified learning over heterogeneous intrusion environments. Recently, to facilitate generalization, a number of studies have explored cross-datasets and harmonized intrusion detection frameworks. **(Alayash et al., 2026)** evaluated LSTM and GRU models across three different IoT intrusion detection datasets, (NF-ToN-IoT , BoT-IoT, and UNSW-NB15) highlighting the importance of multi-dataset evaluation for robust model selection. **(Mishra et al., 2026)** proposed a cross-dataset harmonized intrusion detection framework that integrates feature harmonization, multi-model benchmarking, and statistical validation across NSL-KDD and CICIDS2017. Their study demonstrated the importance of feature alignment for improving cross-dataset generalization and reproducibility. **(Alalwany et al., 2025)** demonstrated that ensemble deep learning can improve intrusion detection robustness and prediction reliability through classifier aggregation mechanisms. **(Peng et al., 2023)** proposed CDF-IDS, that combines CNN and BiLSTM to capture spatial and temporal characteristics and employ focal loss to alleviate class imbalance issue. But the model has been evaluated on individual datasets and did not address unified learning across them. Several recent works have further advanced the field: attention-based transformer architectures have been applied to network intrusion detection on NSL-KDD and UNSW-NB15 datasets **(Akuthota and Bhargava, 2025)**, while hybrid

CNN-BiLSTM models with attention demonstrate strong performance on multiple benchmarks **(Dai et al., 2024; Xi et al., 2024)**. Attention-guided dual-path feature fusion combined with residual LSTM **(Alabbadi and Bajaber, 2025)** and explainable AI integration using SHAP and LIME **(Gaspar et al., 2024)** have further improved both performance and interpretability. Deep learning models enhanced with deformable convolution **(Li et al., 2025)**, hybrid CNN-LSTM architectures **(Bamber et al., 2025)**, and BiGRU with Inception-CNN **(Yang et al., 2024)** have shown high accuracy on IIoT environments. Optimized CNN-LSTM systems have been specifically validated on UNSW-NB15 **(Thaljaoui, 2025)**, while Kolmogorov-Arnold network-based IDS presents a novel architectural alternative **(Wang et al., 2025)**. Lightweight IoT-specific approaches with hybrid deep learning **(Yaras and Dener, 2024; Zahid and Bharati, 2025)**, comprehensive AI-for-IoT surveys **(Sharma and Bairwa, 2025)**, industrial IoT federated learning frameworks **(Liu et al., 2024)**, and explainable IDS for industrial cyber-physical systems **(Nandanwar and Katarya, 2024; 2025)** have also been proposed. Feature fusion mechanisms with deep learning architectures **(Ayantayo et al., 2023)** and ensemble learning surveys for multi-class IDS **(Kaur et al., 2025)** further highlight the breadth of current research directions. However, despite such advancements there are still several limitations. Current CNN-GRU and CNN-BiGRU frameworks mainly have sequential feature learning structures, while these methods are performed on one or two datasets only. Cross-dataset generalization studies in hierarchical intrusion detection are rare, and existing multi-dataset frameworks do not leverage deep representation learning but aim at feature harmonization instead. In addition, existing ensemble methods are mostly not combined with hierarchical attack representation and unified multi-dataset architectures. To overcome these limitations, this work proposes a Spatial–Contextual Dual-Feature Fusion CNN-BiGRU (SCDF-CNN-BiGRU) framework combined parallel spatial-contextual feature learning, hierarchical attack representation, unified multi-dataset learning across six heterogeneous intrusion detection datasets and an ensemble decision mechanism based on hard voting and probability aggregation. To achieve an enhanced feature representation, attack classification and generalization for practical intrusion detection environments.
The contributions of this work can be summarized as:

- A new Spatial–Contextual Dual-Feature Fusion (SCDF-CNN-BiGRU) framework is proposed, integrating a shared CNN feature extractor with parallel BiGRU and bidirectional convolution branches to jointly learn contextual and spatial traffic representations.
- propose a hierarchical intrusion detection and attack-level mapping mechanism framework that achieves not only binary attack detection (attack_level1), but also fine-grained multi-class attack categorization (attack_level2). This will reduce the number of target classes through consolidation them under major attack families and mitigate classes unbalancing problem.
- The proposed framework is validated through unified learning and ensemble inference across six heterogeneous intrusion detection datasets, demonstrating robust detection performance and cross-dataset generalization.

2. DATASETS DESCRIPTION

Six publicly available intrusion detection datasets have been utilized in the experimental test, including (UNSW-NB15, BCCC — cPacket-Cloud-DDoS-2024, NSL-KDD, CIC-UNSW-NB15, NF-ToN-IoT-v3, CICIOT2023). These datasets represent diverse network

environments, attack scenarios, traffic distributions, and feature structures, enabling a comprehensive evaluation of the proposed framework across heterogeneous intrusion detection settings. **Table 1** shows the properties of the six datasets used in the proposed system's training.

Table 1. Dataset Properties

Dataset name/reference	No. of Records	No. of Features	Properties
UNSW-NB15 **(Moustafa and Slay, 2015)**	2059413	49	Latest network traffic with various modern attack
BCCC **(Shafi et al., 2024)**	700776	324	A TCP-layer oriented dataset
NSL-KDD**(Dhanabal and Shantharajah, 2015)**	184517	42	Classical well-known benchmark dataset
CIC-UNSW-NB15**(Canadian Institute for Cybersecurity, 2025)**	2350508	55	A large dataset of flow-based traffic.
NF-ToN-IoT-v3 **(Sarhan et al., 2021)**	25704123	55	Introducing attacks targeting interleaving smart devices in IoT environments
CICIOT2023 **(Neto et al., 2023)**	877337	46	Draws a variety of distributed denial-of-service and botnet-related attack tasks

3. HIERARCHICAL MULTI-DATASET UNIFICATION STRATEGY

The six benchmark intrusion detection datasets used in this study differ greatly in feature structures, attack naming conventions and labeling mechanism. A hierarchical data unification strategy has been devised in order to facilitate unified learning across heterogeneous datasets. First, all datasets were harmonized into a common representation using a unified feature alignment process. Attack labels that were derived from various datasets were then mapped into a three-level hierarchy in terms of attack_level1, attack_level2, and attacked_level3. The first level (attack_level1) defines the binary classification labels, identifying network traffic as Normal or Attack. The second level (attack_level2) corresponds to the generalized attack categories common across all datasets. The third level (attack_level3) retains the existing fine-grained attack names that can be found in each dataset. A global categorical mapping was used for the hierarchical mapping attack labels to ensure a consistent representation of attack-level1 and attack level2 categories across the six datasets. This common label representation is required because the six multi class models produce probability vectors over the same attack-level2 class space during ensemble inference. The hierarchical attack representation adopted in this study builds upon an earlier attack taxonomy framework that has been introduced in our previous two-stage intrusion detection framework **(Mohammed and Hoomod, 2026)**, all dataset were harmonized through a unified feature alignment process. Feature alignment was performed through a unified feature dictionary that mapped semantically equivalent attributes across datasets into a common naming schema before encoding and model training. The SCDF-CNN_BiGRU architecture was trained independently for each dataset using its corresponding feature vector dimension and each dataset has two independent trained models, binary and multi-classification model. The resulting framework consists of twelve independently trained models, six binary models and six multi-class models. The input dimensionality of each model is determined by the number of features available in its corresponding dataset. For example, the NSL-KDD model receives a 42- dimensional input vector, whereas the BCCC model receives a 324-dimensional input vector. During runtime inference, the feature columns required by each model are reconstructed according to its stored dataset-specific feature list, missing features are assigned zero values, extra features

are removed, and the columns are reordered to exactly match the feature order used during training. The predictions of the independently trained models are combined through binary hard voting and multi-class probability averaging.

4. PROPOSED SCDF-CNN-BiGRU FRAMEWORK

The proposed intrusion detection framework operates through a two-stage pipeline: an offline training phase and an online python-based runtime detection phase. During the offline phase, binary and multi class deep learning models are trained. In the online phase, a lightweight detection script monitors incoming network traffic, preprocesses unseen data, and streams them into the loaded pretrained models. When detecting a threat, the program generates an immediate alert containing the attack type and fingerprint.

Building upon the unified feature representation described in the previous section, the proposed SCDF-CNN-BiGRU framework performs hierarchical intrusion detection $X \in R^{T\times 1}$ through a spatial-temporal extraction, feature fusion, and hierarchical intrusion classification pipeline as defined in Eqs. (1-6):

Shared CNN feature extraction

$$(H_{\text{stem}}) = \text{Dropout}(\text{MaxPool1D}(\text{ReLU}(\text{BN}(\text{Conv1D}(\text{X};\text{W}_\text{O}))))) \quad (1)$$

Contextual feature extraction using the BiGRU branch

$$H_{b1} = BN(BiGRU(H_{\text{stem}})) \quad (2)$$

$$F_1 = [\text{GAP}(H_{b1}) \parallel \text{GMP}(H_{b1})] \quad (3)$$

Spatial feature extraction using the bidirectional convolution branch

$$H_{b2} = BN([Conv1D(H_{\text{stem}}) \parallel Conv1D(Reverse(H_{\text{stem}}))]) \quad (4)$$

Spatial-contextual feature fusion

$$F_2 = [GAP(H_{b2}) \parallel GMP(H_{b2})] \quad (5)$$

Fully connected representation learning

$$Z_2 = Dropout\ (Dense(Dropout\ \left(BN\left(Dense(F_1||F_2)\right)\right))) \quad (6)$$

Final hierarchical classification output

$$\left(\hat{Y}\right) = \begin{cases} \sigma(W_{out}Z_2 + b_{out}), & Binary \\ Softmax(W_{out}Z_2 + b_{out}), & Multi-class \end{cases} \quad (7)$$

Where X denotes the input feature sequence, W_O denotes the learnable convolutional kernel of the initial convolution layer, H_{stem} is the shared CNN stem representation after convolution, batch normalization, max pooling, and dropout. H_{b1} denotes the normalized output of the BiGRU branch, while H_{b2} denotes the normalized representation obtained by concatenating the forward and reversed input convolutional outputs. F_1 and F_2 feature representation obtained by concatenating Global Average Pooling (GAP) and Global Max Pooling (GMP) outputs respectively. Z_2 denotes the final fused representation, while W_{out} and b_{out} denotes the weights and bias of the output layer. Conv1D(·), BN(·), MaxPool1D(·), BiGRU(·), GAP(·), and GMP(·) denote one-dimensional convolution, Batch Normalization, one-dimensional max pooling, Bidirectional Gated Recurrent Unit, Global Average Pooling, and Global Max Pooling, respectively. $\hat{Y}$ denotes the predicted output, while $\sigma(.)$ and

$Softmax(.)$ denotes the sigmoid and Softmax activation functions for binary and multi class classification, respectively.

For each dataset, two independent SCDF-CNN-BiGRU models are trained: a binary model based on attack_level1 labels and a multi-class model based on attack_level2 labels. As a result, twelve deep learning models are generated across the six datasets: six binary models and six multi-class models. **Fig. 1** illustrates the overall architecture of the proposed SCDF-CNN-BiGRU framework.

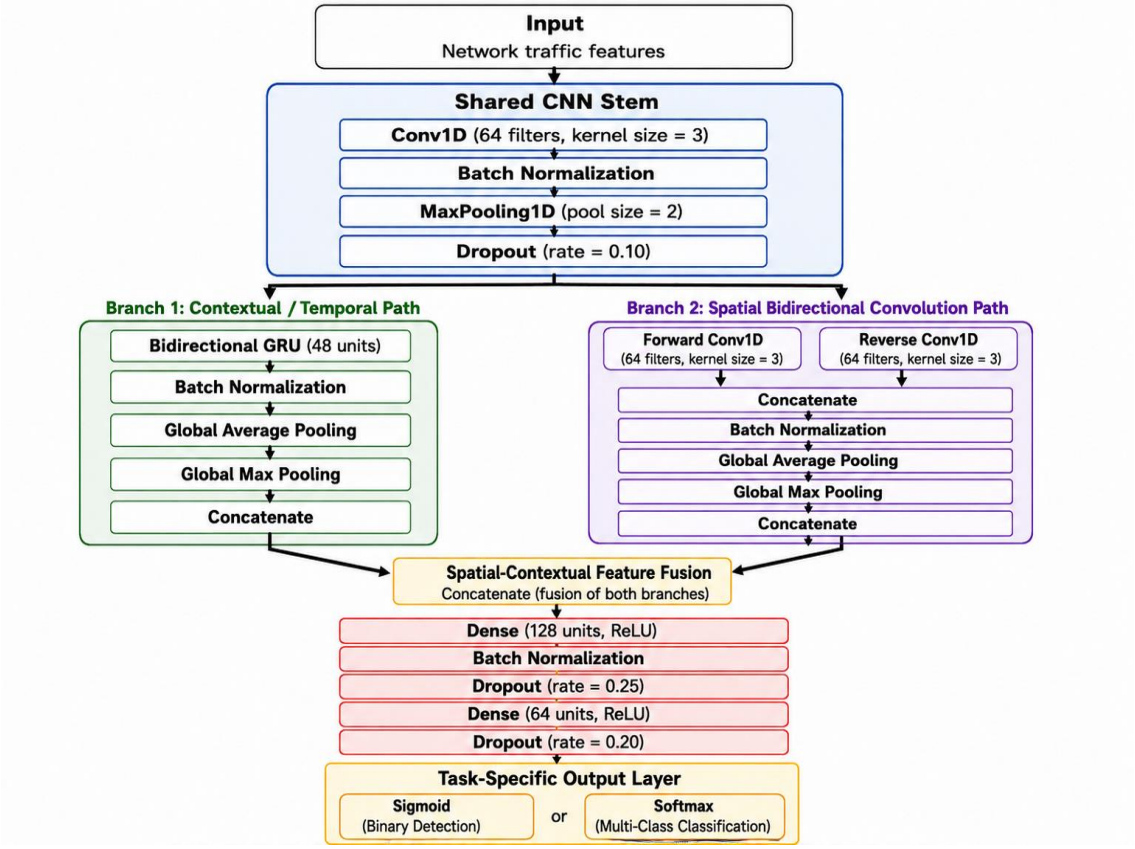

Figure 1. Proposed SCDF-CNN-BiGRU Architecture

4.1 Training Procedure

To ensure consistent feature representation across heterogeneous datasets and to improve the stability and generalization capability of the proposed SCDF-CNN-BiGRU framework, the training procedure involving feature encoding, normalization, dataset partitioning, and model optimization. In feature encoding process, categorical traffic attributes were encoded using corresponding dataset specific categorical mapping on the training partition and subsequently applied to both the training and test partitions to prevent information from the test set from influencing the encoding process. The common categorical mapping was used to maintain a consistent representation of the hierarchical attack labels as defined in Section 4. Min-Max normalization has been applied to normalize numerical features into the range [0,1].

For datasets providing predefined training and testing partitions, the original training and testing partitions were preserved. The validation subset was derived only from the training portion, while the original test set was kept untouched for final evaluation. To prevent the

loss of minority attack categories, attack classes with fewer than six samples are retained within the training subset. The SCDF-CNN-BiGRU models are trained using identical hyperparameter setting across all experiments to ensure fair performance comparison as shown in **Table 2.**

All model training and experimental evaluations were conducted on a workstation equipped with an Intel Core i9-13900K CPU, 128 GB RAM, and an NVIDIA GeForce RTX 4090 GPU with 24 GB VRAM. TensorFlow/Keras with GPU acceleration was used for model training and evaluation.

Table 2. SCDF-CNN-BiGRU Hyperparameter Configuration

Hyperparameter/Component	Value /Configuration
Optimizer	Adam (learning rate = 0.0005)
Batch size	64
epochs	15 with Early Stopping
Loss function (binary/multi)	Binary cross-Entropy/categorical cross-Entropy
Shared CNN Stem	Conv1D (Filters:64, Kernelsize:3, Padding: 'same', Activation: ReLU
BiGRU units (Branch1)	48 units per direction (Bidirectional GRU)
Bi-Conv units (Branch2)	Conv1D (Filters:64, Kernelsize:3) forward & backward
Dropout rates	0.10(stem), 0.25(fusion dense 128), 0.20(fusion dense 64)
Output activations	Sigmoid (Binary), SoftMax (Multi-class)
Early Stopping patience	4 epochs (monitoring val_loss , restore_best_weights=True)
Reduce LR On Plateau Patience	2 epochs (factor:0.5, Min LR:10^{-6})

4.2 Detection Procedure

In the detection phase, captured traffic features undergo the same preprocessing and feature preparation procedure used during model training. The sample is then evaluated by the twelve pretrained SCDF-CNN-BiGRU models. While each model was trained using a dataset-specific feature schema. The input sample is prepared separately for each model according to its stored feature list, scaled using the corresponding data-specific scalar, and reshape to match the model input dimension. For binary intrusion detection, the six binary models produce a binary prediction independently. The sigmoid output of each binary model is converted into a binary decision using a threshold of 0.5, where 0 represents Normal and 1 represents Attack. The six binary decisions are then combined using hard majority voting.

$$D(x) = \begin{cases} 1, & \sum_{j=1}^{6} b_j(x) > 3 \\ 0, & \sum_{j=1}^{6} b_j(x) \leq 3 \end{cases} \quad (8)$$

D(x): the final binary decision,

$b_j(x) \in \{0,1\}$: the binary decision generated by the j-th binary model,

j=1, …, 6 denotes the six-dataset specific binary models.

If the decision is normal, print the result and end detection. If the decision is attack, the multi class attack identification stage is start, each of the six pretrained multi class models produces a probability vector over the unified attack-level2 class space. Because all multi class models use the same attack-level2 label encoding, their output probability vectors correspond to the same class ordering. The probability vectors are then aggregated by element wise averaging.

$$\hat{P}(x) = \frac{1}{6} \sum_{j=1}^{6} P_j(x) \quad (9)$$

The final attack-level2 class is selected according to the maximum value in the aggregated probability vector that can be expressed formally as:

$$\hat{y}_{level2}(x) = \arg\ max\ \widehat{P_k}(x) \quad (10)$$

Where: $P_j(x)$ is the probability vector generated by the j-th multi class model, $\hat{P}(x)$ is the averaged probability vector, N=6 is the number of base multi class models. K indexes the attack-level2 classes. The predicted attack-level2 category is subsequently mapped to its representative attack-level3 description using the dataset-specific attack level mapping.

5. RESULTS AND DISCUSSION

5.1 Overall Detection Performance

The performance of the proposed SCDF-CNN-BiGRU framework was evaluated using six heterogeneous intrusion detection datasets. A variety of evaluation metrics was included such as Accuracy, Precision, Recall, F1-score, Matthews Correlation Coefficient (MCC), Area under the ROC Curve (AUC) and training time measurements. The Receiver Operating Characteristic (ROC) curves illustrate the trade- off between the true positive rate and false positive rate in classification mode. **Fig. 2** shows ROC corresponding to the performance metrics reported in **Table 3.**

Table 3. Evaluation Metrics for Binary Classification (SCDF-CNN-BiGRU)

Dataset	Accu.	Precision weighted	Recall weighted	F1-Score weighted	MCC	AUC
UNSW-NB15	0.991	0.995	0.995	0.995	0.904	0.999
NSL-KDD	0.988	0.990	0.986	0.988	0.975	0.999
NF-ToN-IoT-v3	0.942	0.935	0.969	0.951	0.881	0.982
CICIOT2023	0.985	0.985	0.985	0.985	0.780	0.995
BCCC	0.950	0.930	0.989	0.959	0.898	0.984
CIC-UNSW-NB15	0.999	0.999	0.999	0.999	0.999	0.999

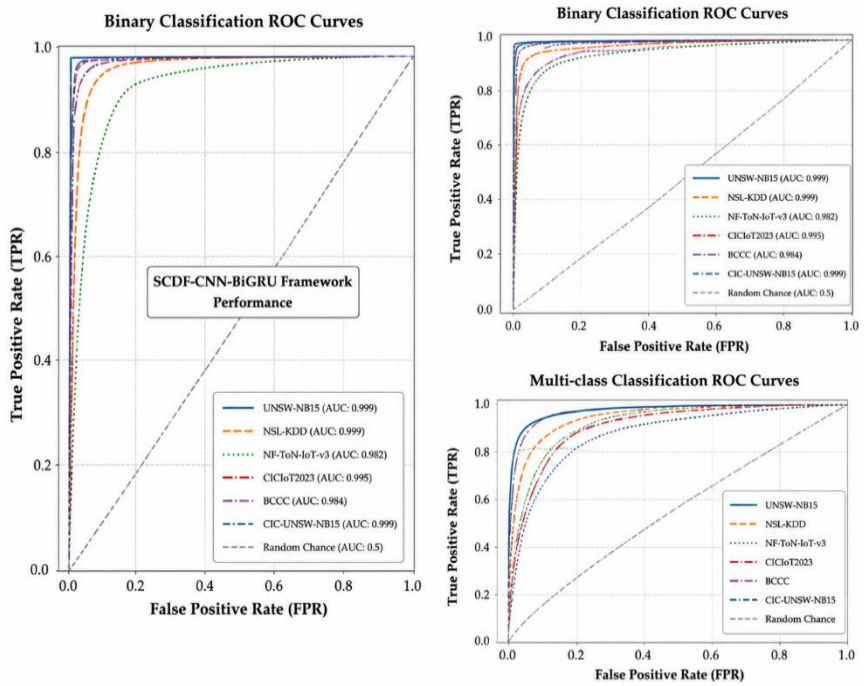

Figure 2. ROC Curves for Classification Across the Evaluated Datasets

For CICIoT2023, the MCC was 0.780 while accuracy, weighted precision, weighted recall and weighted F1 score were approximately 0.985. this difference occurs because MCC considers the overall relationship between the predicted and actual classes and is sensitive to class imbalance and errors affecting minority classes, whereas the other listed metrics are influenced by the contribution of each class according to its support. The multi-class classification performance of the framework is reported in **Table 4.** It is observed that accuracies ranging from 0.910 to 0.988 confirming the effectiveness of the proposed Spatial-Contextual Dual-Feature Fusion (SCDF)architecture.

Table 4. Evaluation Metrics for Multi Classification (SCDF-CNN-BiGRU)

Dataset	Accu.	Precision weighted	Recall weighted	F1-Score Weighted	MCC
UNSW-NB15	0.980	0.979	0.980	0.979	0.784
NSL-KDD	0.988	0.987	0.988	0.987	0.980
NF-ToN-IoT-v3	0.932	0.928	0.931	0.925	0.887
CICIOT2023	0.978	0.978	0.978	0.977	0.837
BCCC	0.910	0.909	0.910	0.898	0.830
CIC-UNSW-NB15	0.984	0.982	0.984	0.982	0.844

The obtained performance is also consistent with recent hybrid deep learning approaches that combine convolutional and recurrent representations for intrusion detection **(Dai et al., 2024; Yang et al., 2024). Table 5** summarizes the total training time results. Although larger datasets required longer training durations, the framework consistently maintained strong detection performance across all evaluated environments.

Table 5. Training Time Metrics for SCDF-CNN-BiGRU

Dataset	Task	Total Training Time (all Epochs) (sec)	Average Time per Epoch(seconds)
UNSW-NB15	Binary	2,196	146.4
	Multi-class	2,235	149.0
NSL-KDD	Binary	534	35.6
	Multi-class	467	31.1
NF-ToN-IoT-v3	Binary	4,146	276.4
	Multi-class	4,813	343.8
CICIOT2023	Binary	4,343	361.9
	Multi-class	3,939	262.6
BCCC	Binary	2,565	171.0
	Multi-class	2,562	170.8
CIC-UNSW-NB15	Binary	4,217	281.1
	Multi-class	5,473	364.9

5.2 Computational Complexity and Limitations

The experimental evaluations show that an ultra-low average inference latency of 24.54 ms per incoming network traffic, verifying that the proposed architecture is computationally lightweight in real time intrusion monitoring environments. coupled with the overall offline training duration listed in **Table 5**, the framework balances high classification performance with manageable computational overhead. Certain practical limitations involve handling exceptionally high-density traffic bursts within a single processing thread may constrain real time network stream unless supported by multi-threading or hardware acceleration. Future

research direction will focus on applying model optimization techniques such as model pruning to accelerate runtime execution, also parallel multi-threaded inference pipeline can be utilized to speedup detection process.

5.3 Ablation Study and Statistical Significance Analysis

A comparative validation of each core component of the proposed architecture is performed:

- Variant A (contextual-only): only the CNN stem and BiGRU branch.
- Variant B (spatial only): CNN stem and the bidirectional CNN branch.
- Variant C (full SCDF): the complete proposed framework.

All variants have been trained and evaluated five times using different random seeds (42,43,44,45,46) on all six datasets as shown in **Table 6**. mean F1 score ±standard deviation for binary classification. A paired two-tailed t-tests (α= 0.05) was performed between the full model and each ablated variant to demonstrated the statistical significance analyses. The results in **Table 6** show that the full SCDF framework outperforms both single branch variants across all datasets. Variant B outperforms variant A, indicating that spatial features are more discriminative for network traffic compared to temporal alone. The statistical improvements ($p<0.05$) achieved by the full mode over variant B indicates that the fusion of BiGRU and BiCNN provides more robust traffic representation.

Table 6. Ablation Study Results (Binary Mean F1 score ±Standard Deviation) Over 5 Seeds

Dataset	Variant A (BiGRU)	Variant B (CNN)	Variant C (full SCDF)	p-value (C vs A)	p-value (C vs B)
UNSW-NB15	0.978 ± 0.003	0.988 ± 0.002	0.995 ±0.001	<0.001	<0.001
NSL-KDD	0.972 ± 0.004	0.983 ± 0.002	0.988 ± 0.001	<0.01	<0.05
NF-ToN-IoT-v3	0.935 ± 0.005	0.942 ± 0.004	0.951 ± 0.002	<0.01	<0.05
CICIOT2023	0.971 ± 0.003	0.979 ± 0.002	0.985 ±0.001	<0.001	<0.01
BCCC	0.941 ± 0.006	0.952±0.003	0.959±0.002	<0.01	<0.05
CIC-UNSW-NB15	0.989 ±0.002	0.995±0.001	0.999± 0.000	<0.001	<0.01

5.4 Comparative Analysis with the State of Art

A comparative analysis is performed of the proposed SCDF-CNNBiGRU with different recent state-of-the-art intrusion detection methods reported in literature, **Table 7**. We contrast results with these existing studies to showcase the unique features of the framework proposed and its ability to overcome various limitations observed in such studies. The direct quantitative comparison is not performed because the existing studies adopt different train-test splits and preprocessing. The qualitative comparison in **Table 7** highlights that: Most of the existing studies are limited to one or only few aspects of intrusion detection for example hybrid deep learning architecture, and hierarchical classification, ensemble learning cross-dataset evaluation. Nevertheless, none unify these approaches into a single framework that jointly leverages dual-feature fusion, hierarchical attack representation, multi-dataset learning together with ensemble decision mechanisms and achieve cross-dataset generalization. However, the proposed SCDF-CNN-BiGRU framework incorporates all of these capabilities into a single unified framework to offer more effective intrusion detection across heterogeneous network environments.

Table 7. Comparison with Related Work Studies

Study	Hybrid DL	Feature Fusion	Hierarchical Classification	Multi-Dataset Evaluation	Ensemble Learning	Cross-Dataset Generalization
CNN-GRU-FF **(Imrana et al., 2024)**	✓	✓	✗	✗	✗	✗
DenseNet1D-BiGRU **(Ma et al., 2025)**	✓	✓	✗	✗	✗	✗
Dual-Path Hybrid Framework **(Kavitha and Harini, 2026)**	✓	✓	✗	Partial	✗	✗
Hierarchical IDS with Attention **(Xu et al., 2023)**	✓	✓	✓	✗	✗	✗
Multi-Dataset LSTM/GRU **(Alayash et al., 2026)**	✓	✗	✗	✓	✗	Partial
Cross-Dataset Harmonized IDS **(Mishra et al., 2026)**	Partial	Partial	✗	✓	✓	✓
Stacking Ensemble IDS **(Alalwany et al., 2025)**	✓	✗	✗	✗	✓	✗
Proposed SCDF-CNN-BiGRU	✓	✓	✓	✓ (6 datasets)	✓	✓

6. CONCLUSIONS

The experimental results demonstrated that the proposed SCDF-CNN-BiGRU framework successfully combines hierarchical dataset unification, spatial-contextual dual-feature fusion, and ensemble decision-making within a unified intrusion detection framework. The proposed architecture achieved binary classification accuracies of up to 0.999 and multi-class accuracies of up to 0.988 across six heterogeneous intrusion detection datasets. These results indicate that the SCDF framework provides robust feature representation and strong cross-dataset generalization for modern network intrusion detection environments. The ablation analysis further confirmed that the fusion of contextual and spatial representations contributed to the overall detection performance. The framework also provides a practical hierarchical representation that connects binary detection with generalized and fine-grained attack descriptions. The framework also achieved an average inference latency of 24.54 ms, indicating its potential for real time intrusion monitoring while maintaining strong detection performance across heterogeneous datasets.

Credit Authorship Contribution Statement

Aseel M. Mahdi: conceptualization, Methodology, Software, Data Curation, Formal Analysis, Investigation, Writing-Original Draft, Visualization. Haider K. Hoomod: Supervision, Validation, Methodology, Writing-Review & Editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

Adefemi, K.O., M.B.M., and Alimi, O.A., 2025. A hybrid CNN–GRU deep learning model for IoT network intrusion detection. *Journal of Sensor and Actuator Networks*, 14(5), 96. https://doi.org/10.3390/jsan14050096

Akuthota, U.C., and Bhargava, L., 2025. Transformer-based intrusion detection for IoT networks. *IEEE Internet of Things Journal*, 12, pp. 6062–6067. https://doi.org/10.1109/JIOT.2025.3525494.

Alabbadi, A., and Bajaber, F., 2025. X-FuseRLSTM: A cross-domain explainable intrusion detection framework in IoT using the attention-guided dual-path feature fusion and residual LSTM. *Sensors*, 25(12), P. 3693. https://doi.org/10.3390/s25123693.

Alalwany, E., Alsharif, B., Alotaibi, Y., Alfahaid, A., Mahgoub, I., and Ilyas, M., 2025. Stacking ensemble deep learning for real-time intrusion detection in IoMT environments. *Sensors*, 25(3), P. 624. https://doi.org/10.3390/s25030624 .

Alayash, W., Rahrouh, M., Ibrahim, A.A., Mohamed, M.H., Ahmed, S.T., Albarri, M.H., and Ahmed, M.H., 2026. Assessing LSTM and GRU for multi-dataset intrusion detection in IoT environments. *Statistics, Optimization & Information Computing*, 15(4), pp. 3155–3173. https://doi.org/10.19139/soic-2310-5070-3226.

Ayantayo, A., Kaur, A., Schmid, A., and Waismann, B., 2023. Network intrusion detection using feature fusion with deep learning. *Journal of Big Data*, 10, P. 167. https://doi.org/10.1186/s40537-023-00834-0.

Bamber, S.S., Katkuri, A.V.R., Sharma, S. and Angurala, M., 2025. A hybrid CNN-LSTM approach for intelligent cyber intrusion detection system. *Computers & Security*, 148, P. 104146. https://doi.org/10.1016/j.cose.2024.104146.

Canadian Institute for Cybersecurity, 2025. UNSW-NB15 Augmented Dataset. University of New Brunswick. Available at: https://www.unb.ca/cic/datasets/cic-unsw-nb15.html [Accessed 27 June 2025].

Dai, W., Li, X., Ji, W., and He, S., 2024. Network intrusion detection method based on CNN, BiLSTM, and attention mechanism. *IEEE Access*, 12, pp. 53099–53111. https://doi.org/10.1109/ACCESS.2024.3385928.

Dhanabal, L., and Shantharajah, D.S.P., 2015. A study on NSL-KDD dataset for intrusion detection system based on classification algorithms. https://doi.org/10.17148/IJARCCE.2015.4696 .

Gaspar, D., Antunes, A., and Bacao, F., 2024. Explainable AI for intrusion detection systems: LIME and SHAP applicability on multi-layer perceptron. *IEEE Access*, 12, pp. 30164–30178. https://doi.org/10.1109/ACCESS.2024.3368377.

Hussein, M.A., Alazawi, S.A.H., and Hoomod, H.K., 2026. Analysis of hybrid OS security using machine learning and deep learning. In: *Advanced Engineering, Technology and Applications on Power Systems. Cham: Springer Nature Switzerland*. https://doi.org/10.1007/978-3-032-13921-4_47 .

Imrana, Y., Xiang, Y., Ali, L., Noor, A., Sarpong, K., and Abdullah, A. A., 2024. CNN-GRU-FF: A double-layer feature fusion-based network intrusion detection system using convolutional neural network and gated recurrent units. *Complex & Intelligent Systems*, 10, pp. 3353–3370. https://doi.org/10.1007/s40747-023-01313-y.

Kaur, J., Randhawa, S.K., and Singh, S., 2025. Ensemble learning approaches for multi-class intrusion detection systems for the Internet of Vehicles: A comprehensive survey. *Future Internet*, 17(7), P. 317. https://doi.org/10.3390/fi17070317.

Kavitha, P., and Harini, S.R., 2026. A dual-path hybrid deep learning framework with BiLSTM, multi-head attention, and intelligent feature engineering in IoT networks. *Research Square Preprint*. https://doi.org/10.21203/rs.3.rs-9302453/v1.

Li, B., Li, J., and Jia, M., 2025. ADFCNN-BiLSTM: A deep neural network based on attention and deformable convolution for network intrusion detection. *Sensors*, 25(5), P. 1382. https://doi.org/10.3390/s25051382.

Liu, L., and Xu, M., 2025. A network intrusion detection method based on contrastive learning and Bayesian Gaussian Mixture Model. *Cybersecurity*, 8(1), P. 59. https://doi.org/10.1186/s42400-025-00364-7 .

Liu, S., Yu, Y., Zong, Y., Yeoh, P.L., Guo, L., Vucetic, B., Duong, T.Q., and Li, Y., 2024. Delay and energy-efficient asynchronous federated learning for intrusion detection in heterogeneous industrial internet of things. *IEEE Internet of Things Journal*, 11(8), pp. 14739–14754. https://doi.org/10.1109/JIOT.2023.3335112.

Ma, X., Liang, J., and Yin, G., 2025. IoT intrusion detection technology based on DenseNet1D and Bi-GRU hybrid model. *Discover Applied Sciences*, 8(2), P. 115. https://doi.org/10.1007/s42452-025-08126-3 .

Mishra, S., Alshammari, N.S., Hussain, H., and Alfahidah, R.A., 2026. A cross-dataset harmonized intrusion detection framework with statistically validated multi-model learning. *PLOS ONE*, 21(4), P. e0346982. https://doi.org/10.1371/journal.pone.0346982 .

Mohammed, A.M., and Hoomod, H.K., 2026. A two-stage hybrid intrusion detection framework based on hierarchical attack mapping and pruned CNN-GRU models. *Engineering, Technology & Applied Science Research*, 16(1), pp. 32342–32347. https://doi.org/10.48084/etasr.16210.

Moustafa, N., and Slay, J., 2015. UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). In: 2015 Military Communications and Information Systems Conference (MilCIS), Australia, pp. 1–6. https://doi.org/10.1109/MilCIS.2015.7348942 .

Nandanwar, H., and Katarya, R., 2024. Deep learning enabled intrusion detection system for industrial IoT environment. *Expert Systems with Applications*, 249, P. 123808. https://doi.org/10.1016/j.eswa.2024.123808.

Nandanwar, H., and Katarya, R., 2025. Securing Industry 5.0: An explainable deep learning model for intrusion detection in cyber-physical systems. *Computers & Electrical Engineering*, 123, P. 110161. https://doi.org/10.1016/j.compeleceng.2025.110161.

Neto, E.C.P., Dadkhah, S., Ferreira, R., Zohourian, A., Lu, R., and Ghorbani, A.A., 2023. CICIoT2023: A real-time dataset and benchmark for large-scale attacks in IoT environment. *Sensors*, 23(13), https://doi.org/10.3390/s23135941 .

Peng, H., Wu, C., and Xiao, Y., 2023. CBF-IDS: Addressing class imbalance using CNN-BiLSTM with focal loss in network intrusion detection system. *Applied Sciences*,13(21), P. 11629. https://doi.org/10.3390/app132111629.

Sagu, A., Gill, N.S., Gulia, P., Alduaiji, N., Shukla, P.K., and Shah, M.A., 2025. Advances to IoT security using a GRU-CNN deep learning model trained on SUCMO algorithm. *Scientific Reports*, 15(1), P. 16485. https://doi.org/10.1038/s41598-025-99574-9.

Sarhan, M., Layeghy, S., Moustafa, N., and Portmann, M., 2021. NetFlow datasets for machine learning-based network intrusion detection systems. In: Deze, Z., Huang, H., Hou, R., Rho, S. and Chilamkurti, N., eds. *Big Data Technologies and Applications*. Cham: Springer. https://doi.org/10.1007/978-3-030-72802-1_9 .

Shafi, M., Lashkari, A.H., Rodriguez, V., and Nevo, R., 2024. Toward generating a new cloud-based distributed denial of service (DDoS) dataset and cloud intrusion traffic characterization. *Information*, 15(4), P. 195. https://doi.org/10.3390/info15040195 .

Sharma, S.B., and Bairwa, A.K., 2025. Leveraging AI for intrusion detection in IoT ecosystems: A comprehensive study. *IEEE Access*, 13, pp. 66290–66317. https://doi.org/10.1109/ACCESS.2025.3556701.

Thaljaoui, A., 2025. Intelligent network intrusion detection system using optimized deep CNN-LSTM with UNSW-NB15. *International Journal of Information Technology*. https://doi.org/10.1007/s41870-025-02416-0.

Wang, Z., Zainal, A., Siraj, M.M., Ghaleb, F.A., Hao, X., and Han, S., 2025. An intrusion detection model based on convolutional Kolmogorov-Arnold networks. *Scientific Reports*, 15(1), P. 1917. https://doi.org/10.1038/s41598-025-85700-4.

Xi, C., Wang, H., and Wang, X., 2024. A novel multi-scale network intrusion detection model with transformer. *Scientific Reports*, 14, P. 23239. https://doi.org/10.1038/s41598-024-74419-5.

Xu, H., Sun, L., Fan, G., Li, W., and Kuang, G., 2023. A hierarchical intrusion detection model combining multiple deep learning models with attention mechanism. *IEEE Access*, 11, pp. 66212–66226. https://doi.org/10.1109/ACCESS.2023.3290613.

Yang, K., Wang, J., and Li, M., 2024. An improved intrusion detection method for IIoT using attention mechanisms, BiGRU, and Inception-CNN. *Scientific Reports*, 14(1), P. 19339. https://doi.org/10.1038/s41598-024-70094-2.

Yaras, S., and Dener, M., 2024. IoT-based intrusion detection system using new hybrid deep learning algorithm. *Electronics*, 13, P. 1053. https://doi.org/10.3390/electronics13061053.

Zahid, M., and Bharati, T.S., 2025. Enhancing cybersecurity in IoT systems: A hybrid deep learning approach for real-time attack detection. *Discover Internet of Things*, 5, P. 73. https://doi.org/10.1007/s43926-025-00083-w.

SCDF-CNN-BiGRU : إطار عمل لدمج الميزات المزدوجة المكانية والسياقية للكشف عن اختراقات الشبكات الهرمية متعددة مجموعات البيانات

اسيل محمد مهدي[1]، *، حيدر كاظم حمود[2]

[1] هيئة البحث العلمي، وزارة التعليم العالي والبحث العلمي، بغداد، العراق

[2] قسم الحاسوب، كلية التربية، الجامعة المستنصرية، وزارة التعليم العالي والبحث العلمي، بغداد، العراق

الخلاصة

تُعدّ أنظمة كشف التسلل الشبكي ركيزة أساسية في حماية الشبكات الحديثة ضد الهجمات الالكترونية المتطورة. ومع ذلك، يواجه الحفاظ على دقة كشف عالية في بيئات الشبكات غير المتجانسة تحديات كبيرة، نظرًا لتنوّع أنماط حركة البيانات وسلوكيات الهجمات. تقدم هذه الدراسة إطار عمل جديدًا لدمج الخصائص المكانية والسياقية يدعى (SCDF-CNN-BiGRU)، للكشف الهرمي عن التسلل الشبكي باستخدام عدة مجموعات بيانات مرجعية في مجال الأمن السيبراني. يعتمد الإطار المقترح على مرحلة خفيفة الوزن لاستخراج الخصائص باستخدام شبكة التفاف عصبية مشتركة (CNN)، يليه فرعان متوازيان ومتكاملان للتعلم، حيث يستخدم الفرع الأول وحدة التكرار البوابية ثنائية الاتجاه (BiGRU) لاستخراج التبعيات الزمنية والعلاقات السياقية، بينما يعتمد الفرع الثاني على معالجة التفافية ثنائية الاتجاه لتعلم التمثيلات المكانية المميزة. ثم تُدمج الخصائص المستخرجة من كلا الفرعين عبر وحدة تجميع متخصصة لتوليد تمثيل شامل لحركة البيانات. كما يتضمن النظام تحليلًا هرميًا للاختراقات، يشمل الكشف الثنائي والتصنيف متعدد الفئات، بالإضافة إلى استراتيجية ربط دقيقة على مستوى الهجمات المختلفة وتفاصيلها الفرعية. تم تقييم الإطار باستخدام ست مجموعات بيانات مرجعية متباينة تشمل: UNSW-NB15 وNSL-KDD وNF-ToN-IoT-v3 وCICIOT2023 وBCCC وCIC-UNSW-NB15. أظهرت النتائج أداءً ممتازاً، حيث بلغت دقة التصنيف الثنائي 0.999 ودقة التصنيف متعدد الفئات 0.988. تشير هذه النتائج إلى أن إطار SCDF-CNN-BiGRU المقترح يوفر تمثيلات قوية وتعميمًا فعالًا عبر مجموعات البيانات المختلفة، مما يجعله حلا واعداً للكشف الهرمي عن الاختراقات في بيئات الشبكات غير المتجانسة.

الكلمات المفتاحية: نظام كشف التسلل الشبكي، دمج الخصائص المكانية والسياقية، CNN-BiGRU، التصنيف الهرمي للهجمات، التعلم على مجموعات بيانات متعددة.