A NOVEL MECHANISM TO FORTIFY APPLICATION LAYER AGAINST DISTRIBUTED DENIAL OF SERVICE (DDOS) ATTACK

A Synopsis for the PhD Thesis

submitted to

Gujarat Technological University

for the Award of

Doctor of Philosophy

in

Computer/IT Engineering

by

Dhangar Dharmeshkumar Natthuram

Enrollment No. 219999913039

under supervision of

Dr. Vikram M. Agrawal

Assistant Professor, Computer Engineering Department,

Bhailalbhai & Bhikhabhai Institute of Technology (BBIT),

VallabhVidya Nagar, Anand, Gujarat, India,388120.

GUJARAT TECHNOLOGICAL UNIVERSITY

AHMEDABAD

July-2026

Index

1 Abstract

Distributed Denial-of-Service (DDoS) attacks remain a major cybersecurity challenge for enterprise networks, cloud platforms, and IoT environments. Traditional intrusion detection approaches often rely on static feature selection techniques, limiting their ability to identify evolving attack patterns and complex traffic behavior. To address this issue, this research proposes a Transformer-Based Attention CNN-BiLSTM framework for intelligent DDoS attack detection.

The proposed model combines a Transformer attention mechanism for dynamic feature weighting, CNN for spatial feature extraction, and BiLSTM for temporal dependency learning. This integrated architecture enables effective representation of network traffic characteristics and improves the distinction between benign and malicious flows.

The framework was evaluated on five benchmark datasets, namely CICDDoS2019, BCCC-cPacket-Cloud-DDoS-2024, CICIOT2023, ITDDoS2020, and APA-DDoS, covering enterprise, cloud, IoT, and advanced network environments. Experimental results demonstrate superior performance compared with conventional feature selection methods such as Mutual Information, ANOVA, and RFE. The proposed model achieved 98.74% accuracy, 98.70% F1-score, 0.9997 ROC-AUC, and 0.9842 MCC on CICDDoS2019, while cross-dataset validation produced accuracy ranging from 99.88% to 100%.

The results confirm that dynamic attention-based feature learning significantly enhances DDoS detection performance. The proposed framework demonstrates strong robustness, scalability, and generalization capability, making it a practical solution for real-time protection against modern DDoS attacks.

Keywords: DDoS Detection, Transformer Attention, CNN, BiLSTM, Deep Learning, Intrusion Detection System, Cybersecurity, Cloud Security, IoT Security.

2 State of the art of the research topic

The widespread adoption of cloud computing, IoT devices, and large-scale networked systems has sharply increased the exposure of modern infrastructure to DDoS attacks. As attack strategies grow in complexity, particularly at the application layer where traffic imitates legitimate use, the development of intelligent, adaptive detection mechanisms has become a critical research area.

Early DDoS detection relied on classical machine-learning algorithms — Support Vector Machines, Decision Trees, Random Forests, and k-Nearest Neighbours. Although effective under controlled conditions, their dependence on manually engineered features limited adaptability to evolving traffic and new attack variants. To overcome this, research turned to deep learning capable of learning traffic features automatically. Convolutional models such as LUCID demonstrated that CNN architectures learn spatial traffic characteristics with low computational overhead, suiting practical deployment [7]. Subsequent work applied LSTM and BiLSTM networks to capture temporal relationships and traffic-evolution patterns, though their effectiveness was often constrained by redundant features and an inability to emphasise the most informative attributes dynamically [5], [9].

More recently, attention-based mechanisms have gained interest for their ability to focus on relevant information during training. Hybrid CNN-attention-BiLSTM designs improved feature discrimination on DDoS benchmarks [16], and self-attention-enabled ensemble CNN frameworks showed that integrating attention with deep learning raises classification accuracy over conventional architectures [17]. Transformer-based intrusion detectors extended this further [13], [14]. Despite these gains, most frameworks continue to rely on static feature-selection methods — Mutual Information, ANOVA, and RFE — which assign fixed importance regardless of context, and most are evaluated on a limited set of benchmarks, raising concerns over generalisation [1], [12].

This research proposes a Transformer-based Attention CNN–BiLSTM framework in which the attention mechanism assigns feature importance dynamically per traffic instance, CNN layers extract spatial patterns, and BiLSTM layers model temporal dependencies. The integration of adaptive attention-based feature learning with spatial–temporal modelling provides a robust and scalable basis for application-layer DDoS detection, addressing key limitations of the current state of the art.

2.1 Literature Survey

A total of 38 studies from IEEE, Elsevier, Springer, ACM, Taylor & Francis, and other digital libraries were surveyed during this research, spanning 2019–2026 and covering statistical, classical machine-learning, and deep-learning approaches to DDoS detection. From these, 24 representative and directly relevant studies — selected for their use of hybrid, attention-, Transformer-, or feature-selection-based detection and for their verifiability at the publisher source — are compared in Table 2.1. The remaining surveyed works informed the analysis but are not tabulated individually, because they address adjacent problems. The studies are ordered by relevance, with hybrid, attention-, and Transformer-based detectors placed first, so that the research gap emerges directly from the comparison. A summary of representative research is presented in Table 2.1.

Table 2.1: Comparative review of application-layer DDoS detection studies

Ref	Author(s), Year & Publisher	Technique / Approach	Datasets	Reported Performance	Key Limitations
[16]	Zhao et al., 2023 IEEE	CNN-BiLSTM with attention (CNN-AttBiLSTM)	CICDDoS2019	High accuracy; attention improved discrimination	Compute cost of hybrid; single-dataset
[17]	Kanthimathi et al., 2024 IEEE	Self-attention weighted-ensemble CNN	CICDDoS2019	Acc. 98.32%, P 98.15%, R 98.47%, F1 98.31%	Needs large data; heavy compute
[18]	Kirubavathi et al., 2025 Springer	Self/intersample-attention Transformer (SAINT)	BCCC-Cloud-2024	~97% acc.; scalable, interpretable	TCP-focused; cloud testbed only

Ref	Author(s), Year & Publisher	Technique / Approach	Datasets	Reported Performance	Key Limitations
[24]	Al-Eryani et al., 2025 Springer	GRU-BiLSTM hybrid	CICDDoS2019	Acc. 99.99%, FPR 0.01	Sequential-heavy; cross-domain untested
[25]	Zaidoun & Lachiri, 2025 Springer	Hybrid DNN-LSTM with attention	CICDDoS2019	Val ≈99%, Test 98.84%	Single-dataset validation
[13]	Wu et al., 2022 IEEE	Transformer IDS (RTIDS)	NSL-KDD / CICIDS	Competitive transformer IDS	General IDS, not L7-specific
[14]	Manocchio et al., 2024 Elsevier	FlowTransformer framework	Multiple flow datasets	Generalises transformer IDS	Flow-level abstraction
[15]	Djaidja et al., 2024 IEEE	Attention + RNNs (early detection)	IDS benchmark	Sharpened early detection	Not L7-specific
[22]	Pradeep & Shukla, 2025 Elsevier	Stacked network + optimal feature selection	DDoS benchmark	Acc. 98.5%, P 97.8%, R 98.1%, F1 98.0%	Hyperparameter-sensitive; no cost analysis
[26]	Al-Shukaili et al., 2025 Springer	MI + RFE with hybrid deep learning	CIC-LDDoS	Acc. 97.8%; ~18% FN reduction	Limited multi-category generalisation
[19]	Sharif & Beitollahi, 2023 Elsevier	ML + genetic-algorithm feature selection	App-layer DDoS set	GA selection improved accuracy	Static selection; GA cost
[23]	Yu et al., 2024 Springer	BiGAN + contrastive learning	Anomaly dataset	Acc. 97.2%, P 96.8%, R 97.0%, F1 96.9%	High compute; narrow validation
[8]	Cil et al., 2021 Elsevier	Feed-forward DNN	CICDDoS2019	Acc. 0.9997 (binary); drops multiclass	Degrades on multi-class
[9]	Assis et al., 2021 Elsevier	GRU deep-learning system	CICDDoS2019 / CICIDS2018	Avg. metrics 99.94% (CICDDoS2019)	Detection rate & time not reported
[11]	Amaizu et al., 2021 Elsevier	Dual DNN + feature extraction (B5G)	B5G traffic	Acc. 99.66%, P 99.52%, R 99.30%	Weak real-time performance
[5]	Muraleedharan & Janet, 2021 Elsevier	FC feed-forward network on flow data	CICIDS2017	Acc. 99.61% (slow-DoS)	Limited to HTTP slow-DoS
[21]	Kemp et al., 2023 Springer	Statistical features + ML	App-layer DoS set	Acc. 96.5%, P 95.8%, R 96.0%, F1 95.9%	Limited dataset; high-traffic not validated
[20]	Akgun et al., 2022 Elsevier	Deep-learning IDS model	IDS benchmark	Strong multi-class detection	Single-dataset; thin low-rate coverage
[6]	Fu et al., 2022 Springer	Time–frequency analysis (low-rate)	Low-rate DoS traces	Detects periodic low-rate bursts	Narrow to low-rate class
[4]	David & Thomas, 2021	Information-distance	Flash-crowd vs DDoS	Separates attack from flash	Threshold sensitive to drift

Ref	Author(s), Year & Publisher	Technique / Approach	Datasets	Reported Performance	Key Limitations
	Elsevier	thresholding		crowd	
[3]	Praseed & Thilagam, 2021 IEEE	Behavioural-dynamics modelling	HTTP traces	Effective for asymmetric L7 floods	Needs per-application baseline
[7]	Doriguzzi-Corin et al., 2020 † IEEE	Lightweight CNN (LUCID)	CICIDS2017	Acc. 0.9967, FPR 0.0059, F1 0.9966	Processing time not reported
[10]	Haider et al., 2020 IEEE	Deep CNN ensemble	CICIDS2017	Acc. 99.45%, F1 99.61%	High training time
[12]	Mittal et al., 2023 Springer	Systematic review of DL for DDoS	— (review)	Maps DL landscape & gaps	Review; no new system

2.2 Critical Analysis of Existing Research

The comparison shows a clear trajectory: detection has progressed from single-architecture CNN and DNN models (rows 22–24), through recurrent and generative hybrids (rows 12, 14, 15), to attention- and Transformer-based designs (rows 1–8), while feature selection has remained largely static — MI, ANOVA, RFE, PCA, or genetic-algorithm wrappers (rows 9–11, 18). Deep-learning models have markedly improved detection, yet most still rely on static feature selection and limited benchmarks, and few evaluate class imbalance, cross-dataset generalisation, computational efficiency, and explainability together [1], [12]. Although attention has shown promise, its integration into hybrid architectures as an explicit, instance-wise feature-weighting mechanism — rather than an internal classifier layer — remains limited [15], [16], [17]. These limitations justify a Transformer-based CNN–BiLSTM framework for robust, adaptive, and generalised application-layer DDoS detection.

2.3 Research Gap Identified

Based on the reviewed literature, the following research gaps are identified:

From the survey and critical analysis, the following gaps are identified:

1) Static feature-selection methods cannot adjust feature importance dynamically during inference;
2) Transformer-based attention for adaptive feature weighting in application-layer DDoS detection is under-explored;
3) Cross-dataset validation on recent benchmarks is insufficient, reducing confidence in generalisation;
4) many models emphasise overall accuracy while overlooking imbalance-aware measures such as Balanced Accuracy, MCC, Cohen's Kappa, and G-Mean;
5) CNN, BiLSTM, and Transformer attention are seldom unified within a single end-to-end architecture; and
6) Computational efficiency and explainability are rarely investigated together, limiting real-world deployment.

3 Definition of the Problem

Application-layer Distributed Denial of Service (DDoS) attacks have evolved into highly adaptive and stealthy threats that closely resemble legitimate user behaviour, making their detection significantly more challenging than conventional volumetric attacks. Existing intrusion detection approaches predominantly employ static feature selection techniques and standalone deep learning models, which are often unable to dynamically identify the most discriminative traffic characteristics under evolving attack patterns. Furthermore, many reported solutions are evaluated using a single benchmark dataset and emphasize overall detection accuracy, while providing limited evidence of robustness, generalization, computational efficiency, and reliable detection under diverse application-layer attack scenarios.

Therefore, there is a need for a novel detection mechanism capable of dynamically learning feature importance while simultaneously modelling both spatial and temporal characteristics of network traffic. This research addresses the problem by proposing a Transformer-based Attention CNN–BiLSTM

framework, where the Transformer attention mechanism performs adaptive instance-wise feature weighting, CNN extracts discriminative spatial representations, and BiLSTM captures temporal dependencies. The proposed mechanism is validated across multiple contemporary benchmark datasets using comprehensive performance and robustness metrics to fortify application-layer services against evolving DDoS attacks through improved detection accuracy, generalization, and real-time applicability.

4 Objectives and Scope of work

The primary objective of this research is to develop an adaptive deep learning framework capable of improving the detection of application-layer Distributed Denial of Service (DDoS) attacks through dynamic feature learning and hybrid deep neural network modeling.

The specific objectives are:

1. To design and develop a Transformer-based Attention CNN–BiLSTM framework for accurate detection of application-layer DDoS attacks by jointly learning spatial, temporal, and contextual characteristics of network traffic.
2. To implement an adaptive feature-weighting mechanism using Transformer attention that dynamically identifies feature importance during training, eliminating dependence on static selection.
3. To evaluate the framework using comprehensive imbalance-aware metrics: Accuracy, Precision, Recall, F1-score, Balanced Accuracy, MCC, Cohen's Kappa, ROC-AUC, G-Mean, FPR, FNR, and inference time.
4. To validate the robustness and generalization of the framework through experimentation on contemporary benchmark datasets.

Scope of Work:

The scope of the proposed research is defined as follows:

1. **The research focuses on the detection and classification of application-layer Distributed Denial of Service (DDoS) attacks** using an adaptive deep

learning framework based on Transformer attention, CNN, and BiLSTM architectures.

2. **The proposed framework employs dynamic attention-based feature weighting** to learn feature importance during training, thereby avoiding dependence on conventional static feature selection techniques.
3. **Experimental evaluation is conducted using contemporary benchmark datasets**, namely **CICDDoS2019, BCCC-cPacket-Cloud-DDoS-2024, CICIOT2023, ITDDoS2020, and APA-DDoS**, to assess model robustness across diverse network environments.
4. **The performance of the proposed model is analyzed using multiple evaluation metrics**, including Accuracy, Precision, Recall, F1-score, Balanced Accuracy, ROC-AUC, MCC, Cohen's Kappa, G-Mean, False Positive Rate (FPR), False Negative Rate (FNR), and inference latency.
5. **The research includes comparative performance analysis** against conventional feature selection approaches, namely Mutual Information (MI), Analysis of Variance (ANOVA), and Recursive Feature Elimination (RFE), to demonstrate the effectiveness of the proposed attention-based feature learning mechanism.
6. **The scope of this work is limited to the detection and classification of DDoS attacks using benchmark datasets** and does not include attack mitigation, traffic filtering, or deployment in operational production network environments.

Expected Outcome

1. Development of a novel hybrid deep learning framework integrating Transformer-based attention, CNN, and BiLSTM for effective detection of application-layer Distributed Denial of Service (DDoS) attacks.
2. An adaptive feature learning mechanism capable of dynamically assigning feature importance, thereby improving feature representation compared with conventional static feature selection approaches.
3. Enhanced detection performance with improved robustness, balanced classification, and reduced false alarm rates across diverse application-layer DDoS attack scenarios.

4. A generalized detection model validated on multiple contemporary benchmark datasets to demonstrate improved adaptability across heterogeneous network environments.
5. Comprehensive performance evaluation using accuracy, precision, recall, F1-score, Balanced Accuracy, ROC-AUC, MCC, Cohen's Kappa, G-Mean, FPR, FNR, and inference latency to establish the effectiveness of the proposed framework.
6. A scalable and interpretable intrusion detection framework that contributes to strengthening the security of application-layer services and provides a foundation for future intelligent network security solutions.

5 Original Contribution of the Thesis

The original contributions of the proposed research are summarized as follows:

1. **A novel Transformer-based Attention CNN–BiLSTM framework** is developed for application-layer DDoS attack detection by integrating adaptive attention-based feature learning with spatial and temporal deep learning models within a unified architecture.
2. **An adaptive instance-wise feature weighting mechanism** is introduced using Transformer attention, enabling the model to learn the relative importance of traffic features dynamically during training instead of relying on conventional static feature selection methods such as Mutual Information (MI), ANOVA, Principal Component Analysis (PCA), and Recursive Feature Elimination (RFE).
3. **A comprehensive comparative evaluation framework** is established to assess the effectiveness of dynamic attention-based feature learning against widely used feature selection approaches using multiple classification, robustness, and computational performance metrics.
4. **Cross-dataset validation** of the proposed framework is performed using multiple contemporary benchmark datasets, namely **CICDDoS2019, BCCC-cPacket-Cloud-DDoS-2024, CICIOT2023, ITDDoS2020, and APA-**

DDoS, to demonstrate its generalization capability across heterogeneous network environments.

5. **An imbalance-aware performance evaluation methodology** incorporating Balanced Accuracy, MCC, Cohen's Kappa, G-Mean, False Positive Rate (FPR), False Negative Rate (FNR), and inference latency is presented to provide a comprehensive assessment of model reliability beyond conventional accuracy-based evaluation.
6. **The proposed framework provides an adaptive and scalable mechanism** for strengthening application-layer DDoS detection, thereby contributing toward the development of more robust intelligent intrusion detection systems for modern network environments.

6 Methodology of Research

The Proposed Transformer-CNN_BiLSTM framework follows a five-phase processing pipeline: (i) Dataset Preprocessing and dataset normalization, (ii) Transformer-based Attention for adaptive feature selection, (iii) Spatial feature learning using CNN, (iv)Temporal feature learning using BiLstm, and (v) Final Classification using fully connected layers. The complete algorithmic steps are presented in Algorithm 1.

Algorithm 1 Transformer-CNN-BiLSTM Framework for DDoS Detection

Input: Preprocessed Dataset $\mathbf{X} \in \mathbb{R}^{B \times F}$, Ground Truth $\mathbf{Y}$

Output: Classification Labels $\hat{y}$

// Phase 1: Transformer-based Attention for Feature Selection

Initialize Weights $W^Q, W^K, W^V \in \mathbb{R}^{F \times d_k}$

$Q \leftarrow \mathbf{X}W^Q, K \leftarrow \mathbf{X}W^K, V \leftarrow \mathbf{X}W^V$

$Score \leftarrow \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)$ ▷ Inter-feature dependency

$\mathbf{M} \leftarrow \text{Sigmoid}(\text{Mean}(Score))$ ▷ Feature Saliency Mask

$\mathbf{X}_{attn} \leftarrow \mathbf{X} \odot \mathbf{M}$ ▷ Element-wise gating

// Phase 2: Spatial Feature Learning (CNN)

$\mathbf{X}_{res} \leftarrow \text{Reshape}(\mathbf{X}_{attn}, (B, T, F_{red}))$

$C \leftarrow \text{ReLU}(\text{Conv1D}(\mathbf{X}_{res}, \text{filters} = 64, k = 3))$

$P \leftarrow \text{MaxPool}(C, \text{pool\_size} = 2)$

// Phase 3: Temporal Feature Learning (BiLSTM)

$[\overrightarrow{h}_t, \overleftarrow{h}_t] \leftarrow \text{BiLSTM}(P)$

$\mathbf{H}_{final} \leftarrow \text{Concatenate}(\overrightarrow{h}_t, \overleftarrow{h}_t)$

// Phase 4: Classification

$v \leftarrow \text{GlobalAveragePooling}(\mathbf{H}_{final})$

$z \leftarrow \text{Dense}(v, \text{activation} = \text{'gelu'})$

$\hat{y} \leftarrow \text{Softmax}(z)$

return $\hat{y}$

The proposed research adopts a systematic methodology for developing and validating an intelligent framework for detecting application-layer Distributed Denial of Service (DDoS) attacks. The research workflow consists of the following stages:

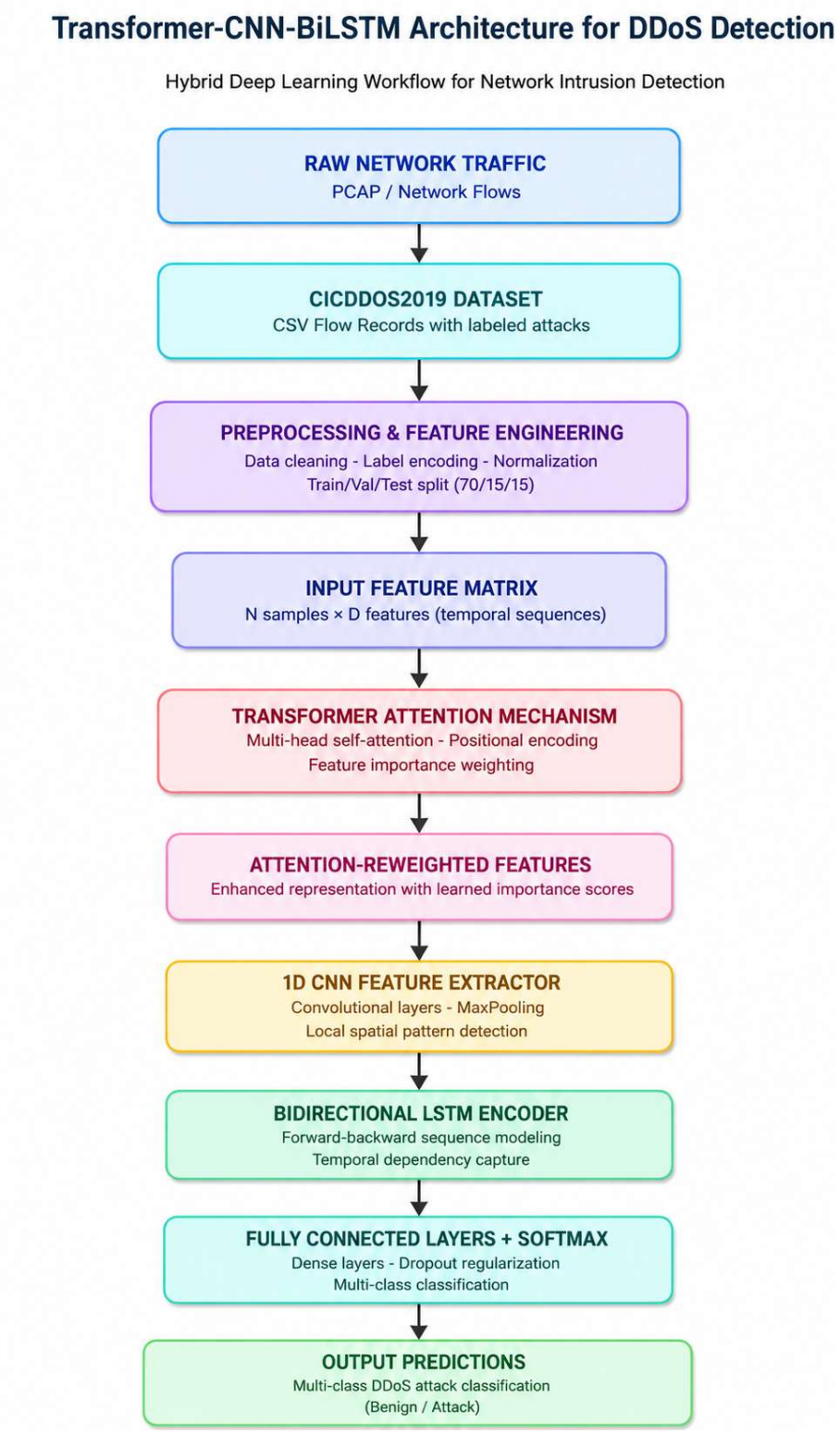

FIGURE 1. — End-to-End architecture of the Proposed Model

7 Results and Comparisons

7.1 Experimental Results

The proposed Transformer-based Attention CNN–BiLSTM framework was experimentally evaluated using the **CICDDoS2019** dataset and subsequently validated on four additional benchmark datasets, namely **BCCC-cPacket-Cloud-DDoS-2024, CICIOT2023, ITDDoS2020, and APA-DDoS**, to assess its robustness and generalization capability. The experimental evaluation employed multiple classification, robustness, and computational performance metrics, including Accuracy, Precision, Recall, F1-score, Balanced Accuracy, ROC-AUC, Matthews Correlation Coefficient (MCC), Cohen's Kappa, G-Mean, False Positive Rate (FPR), False Negative Rate (FNR), and inference latency. Per-dataset results are reported in Tables 7.1(a)–(e); the consolidated cross-dataset comparison follows in Table 7.2.

The experimental observations demonstrate that the proposed framework effectively combines Transformer-based adaptive feature learning with CNN-based spatial feature extraction and BiLSTM-based temporal sequence modelling. The proposed model consistently achieved high classification performance while maintaining robust behaviour across heterogeneous network environments. Cross-dataset validation further confirms the ability of the proposed framework to generalize beyond the primary training dataset, thereby improving its applicability for practical application-layer DDoS detection.

Table 7.1(a). Performance on CICDDoS2019 (Primary Dataset)

Metrics	Results	Metrics	Results
Accuracy	98.74%	MCC	0.98
Balanced Accuracy	83.61%	Cohen's Kappa	0.98
Precision	98.81%	Specificity	99.90%
Recall	98.74%	ROC-AUC	0.99
F1-score	98.70%	G-Mean	99.33%
Inference latency	0.9964 ms/sample		

Table 7.1(b). Performance on BCCC-cPacket-Cloud-DDoS-2024
(cloud environment; test set 20,000 samples; TN/FP/FN/TP = 19,997 / 3 / 0 / 20,000).

Metric	Result	Metric	Result
Accuracy	99.99%	FPR	0.0150%
Precision	99.98%	FNR	0.0%
Recall (TPR)	100%	MCC	0.99
F1-score	99.99%	Cohen's Kappa	0.99
Specificity (TNR)	99.98%	ROC-AUC	1.00
Balanced Accuracy	99.99%	G-Mean	0.99
Inference latency	1.026 ms/sample		

Table 7.1(c). Performance on CICIOT2023
(IoT environment; test set 5.00 lakh samples; TN/FP/FN/TP = 21,226 / 22 / 591 / 478,161).

Metric	Result	Metric	Result
Accuracy	99.88%	MCC	0.9852
Precision	100.00%	Cohen's Kappa	0.9851
Recall	99.88%	G-Mean	0.9989
F1-score	99.94%	ROC-AUC	1.0000
Specificity	99.90%	Inference time	88.29 s (0.177 ms/sample)

Table 7.1(d). Performance on ITDDoS-2020
(IoT/enterprise environment; 83 features; TN/FP/FN/TP = 8,015 / 0 / 0 / 117,142).

Metric	Result	Metric	Result
Test Accuracy	100%	DDoS Detection Rate	100%
Precision	100%	False Positive Rate	0.00%
Recall	100%	False Negative Rate	0.00%
F1-score	100%	ROC-AUC	1.00
Training epochs	~50 (early stopping)	—	—

Table 7.1(e). Performance on APA-DDOS-2021
(advanced-network environment; test set 30,240 samples)

Metric	Result	Metric	Result
Accuracy	100%	MCC	1.00
Precision	100%	Cohen's Kappa	1.00
Recall	100%	G-Mean	1.00
F1-score	100%	ROC-AUC	1.00
Balanced Accuracy	100%	Average latency	1.358 ms/sample
Training time	80.59 min	Inference time	41.06 s

7.2 Comparative Analysis

Cross-dataset comparison. Table 7.2 consolidates the framework's performance across the five environments, demonstrating consistent detection and generalization beyond the primary training dataset.

Table 7.2. Comprehensive cross-dataset comparative analysis

Dataset	CICDDoS2019	BCCC-Cloud-DDoS-2024	CICIOT2023	ITDDoS-2020	APA-DDoS-21
Environment	Enterprise	Cloud	IoT	IoT/Enterprise	Advance Network
Samples	5.60 Cr+	7.0 L+	4.67 Cr+	6.26 L	1.51 L
Features	88	315	46	83	23
Accuracy (%)	98.74	99.99	99.88	100	100
Bal. Acc. (%)	83.61	99.99	99.89	100	100
F1 (%)	98.7	99.94	99.94	100	100
ROC-AUC	0.99	1	1	1	1
MCC	0.98	0.99	0.98	1	1
Kappa	0.98	0.99	0.98	1	1
Latency	0.99 s	0.17 ms/sample	1.02 ms/sample	—	1.35 ms/sample

7.3 Discussions

The proposed framework enhances application-layer DDoS detection by integrating adaptive feature learning with hybrid deep learning: attention learns feature relevance dynamically, CNN captures localised spatial characteristics, and BiLSTM models temporal dependencies, yielding a more discriminative representation than static selection [16], [17], [25]. The accuracy margin over the best static selector (ANOVA) is under one percentage point, but the Balanced-Accuracy margin is large — 0.835 versus 0.718–0.741 — and the method leads on G-Mean, MCC, and Cohen's Kappa, indicating superior performance under class imbalance. The trade-off is a higher inference time (0.996 s vs 0.272–0.286 s), justified by these gains. Against PCA, the method offers higher threshold-independent separability, a smaller footprint, and interpretability. A false-negative rate of 0.1621 observed in the PCA comparison is noted for reduction, and extended validation across additional benchmarks (CICIoT2023, BCCC-cPacket-Cloud-DDoS-2024, APA-DDoS) is in progress.

8. Achievements with respect to objectives

All the defined research objectives of this work were successfully achieved and are summarized in below table as follows.

Table 8.1 Achievements with respect to objectives

Research Objectives	Achievements
Objective 1: Develop a Transformer-based Attention CNN–BiLSTM framework	A unified end-to-end architecture integrating multi-head self-attention, 1-D CNN, and BiLSTM was designed, implemented, and trained.
Objective 2: Implement adaptive feature learning to replace conventional static feature selection techniques.	An instance-wise attention saliency gate was implemented and shown to outperform MI, ANOVA, RFE, and PCA.
Objective 3: Evaluate the proposed framework using comprehensive performance metrics.	The framework was assessed on eleven-plus metrics, with Balanced Accuracy, MCC, Kappa, and G-Mean decisive in interpretation.
Objective 4: Validate the robustness and generalization capability of the proposed framework.	The proposed framework has been validated using CICDDoS2019, BCCC-cPacket-Cloud-DDoS-2024, CICIOT2023, ITDDoS2020, and APA-DDoS benchmark datasets. Experimental observations demonstrate consistent detection performance across enterprise, cloud, and IoT traffic environments, confirming the framework's robustness and cross-dataset generalization capability. Comparative evaluation with MI, ANOVA, and RFE further demonstrates the effectiveness of adaptive attention-based feature learning over conventional feature selection approaches.

9. Conclusion

This research presents a Transformer-based Attention CNN–BiLSTM framework that strengthens application-layer DDoS detection through adaptive, instance-wise feature learning and hybrid deep neural modelling. The framework replaces static feature selection with a learned attention saliency gate, jointly capturing spatial and temporal traffic characteristics. On the CICDDoS2019 benchmark it outperformed Mutual Information, ANOVA, RFE, and PCA, with its advantage concentrated in imbalance-aware reliability- Balanced Accuracy 0.835, MCC 0.985, Cohen's Kappa 0.984, and G-Mean 0.993 — rather than in overall accuracy, where the margin was slim; the attention weights additionally provide interpretable feature importance. The contribution is bounded to detection and classification and, on present evidence, to a single well-characterized benchmark; multi-dataset validation, reduction of the false-negative rate, real-time and edge deployment, HTTP/3-over-QUIC coverage, and federated/online learning are identified as future work. Within these bounds, the work provides evidence that adaptive attention-based feature learning is a principled, interpretable, and more imbalance-robust alternative to static feature selection for fortifying the application layer against DDoS attacks.

List of Publications

1. Dharmesh Dhangar, “A Review of Defense Mechanisms against Distributed Denial-of-Service (DDoS) Attacks on the Application Layer, as well as Machine Learning (ML)-based Mechanisms to Defend Against DDoS Attacks”, Int J Intell Syst Appl Eng, vol. 12, no. 21s, pp. 4583–4590, Mar. 2024, Indexed in **SCOPUS**.

2. D. Dhangar and V. Agrawal, “Adaptive Transformer-Gated Feature Selection With Cnn–Bilstm Hybrid Network For Efficient Ddos Attack Detection”, IJASIS, pp. 110–129, Dec. 2025, doi: 10.29284/s2gavn02, Indexed in **SCOPUS**.

3. D. N. Dhangar and V. Agrawal, “TAFS-Net: A Transformer Attention-Based Feature Selection and BiLSTM Framework for Intelligent Cloud DDoS Attack Detection Using the BCCC-Packet-2024 Dataset,” International Scientific Journal of Engineering and Management (ISJEM), vol. 5, no. 7, pp. 1–10, Jul. 2026, doi: 10.55041/ISJEM08216.

Bibliography

[1] T. F. Deressu, A. M. Beyene, and A. Y. Gemechu, "A survey of application layer distributed denial of service detection: Approaches, models, and datasets," Artif. Intell. Rev., art. no. 139, 2026, doi: 10.1007/s10462-026-11528-3.

[2] N. Tripathi and N. Hubballi, "Application layer denial-of-service attacks and defense mechanisms: A survey," ACM Comput. Surv., vol. 54, no. 4, pp. 1–33, 2021, doi: 10.1145/3448291.

[3] A. Praseed and P. S. Thilagam, "Modelling behavioural dynamics for asymmetric application layer DDoS detection," IEEE Trans. Inf. Forensics Security, vol. 16, pp. 617–626, 2021, doi: 10.1109/TIFS.2020.3017928.

[4] J. David and C. Thomas, "Discriminating flash crowds from DDoS attacks using efficient thresholding algorithm," J. Parallel Distrib. Comput., vol. 152, pp. 79–87, 2021, doi: 10.1016/j.jpdc.2021.02.019.

[5] N. Muraleedharan and B. Janet, "A deep learning based HTTP slow DoS classification approach using flow data," ICT Express, vol. 7, no. 2, pp. 210–214, 2021, doi: 10.1016/j.icte.2020.08.005.

[6] Y. Fu, X. Duan, K. Wang, and B. Li, "Low-rate denial of service attack detection method based on time-frequency characteristics," J. Cloud Comput., vol. 11, no. 1, art. no. 31, 2022, doi: 10.1186/s13677-022-00308-3.

[7] R. Doriguzzi-Corin, S. Millar, S. Scott-Hayward, J. Martínez-del-Rincón, and D. Siracusa, "LUCID: A practical, lightweight deep learning solution for DDoS attack detection," IEEE Trans. Netw. Service Manag., vol. 17, no. 2, pp. 876–889, 2020, doi: 10.1109/TNSM.2020.2971776. [FOUNDATIONAL]

[8] A. E. Cil, K. Yildiz, and A. Buldu, "Detection of DDoS attacks with feed forward based deep neural network model," Expert Syst. Appl., vol. 169, art. no. 114520, 2021, doi: 10.1016/j.eswa.2020.114520.

[9] M. V. O. de Assis, L. F. Carvalho, J. Lloret, and M. L. Proença, "A GRU deep learning system against attacks in software defined networks," J. Netw. Comput. Appl., vol. 177, art. no. 102942, 2021, doi: 10.1016/j.jnca.2020.102942.

[10] S. Haider, A. Akhunzada, I. Mustafa, T. B. Patel, A. Fernandez, K.-K. R. Choo, and J. Iqbal, "A deep CNN ensemble framework for efficient DDoS attack detection in software defined networks," IEEE Access, vol. 8, pp. 53972–53983, 2020, doi: 10.1109/ACCESS.2020.2976908. [FOUNDATIONAL]

[11] G. C. Amaizu, C. I. Nwakanma, S. Bhardwaj, J. M. Lee, and D. S. Kim, "Composite and efficient DDoS attack detection framework for B5G networks," Comput. Netw., vol. 188, art. no. 107871, 2021, doi: 10.1016/j.comnet.2021.107871.

[12] M. Mittal, K. Kumar, and S. Behal, "Deep learning approaches for detecting DDoS attacks: A systematic review," Soft Comput., vol. 27, no. 18, pp. 13039–13075, 2023, doi: 10.1007/s00500-021-06608-1.

[13] Z. Wu, H. Zhang, P. Wang, and Z. Sun, "RTIDS: A robust transformer-based approach for intrusion detection system," IEEE Access, vol. 10, pp. 64375–64387, 2022, doi: 10.1109/ACCESS.2022.3182333.

[14] L. D. Manocchio, S. Layeghy, W. W. Lo, G. K. Kulatilleke, M. Sarhan, and M. Portmann, "FlowTransformer: A transformer framework for flow-based network intrusion detection

systems," Expert Syst. Appl., vol. 241, art. no. 122564, 2024, doi: 10.1016/j.eswa.2023.122564.

[15] T. E. T. Djaidja, B. Brik, S. M. Senouci, A. Boualouache, and Y. Ghamri-Doudane, "Early network intrusion detection enabled by attention mechanisms and RNNs," IEEE Trans. Inf. Forensics Security, vol. 19, pp. 7783–7793, 2024, doi: 10.1109/TIFS.2024.3441862.

[16] J. Zhao, Y. Liu, Q. Zhang, and X. Zheng, "CNN-AttBiLSTM mechanism: A DDoS attack detection method based on attention mechanism and CNN-BiLSTM," IEEE Access, vol. 11, pp. 136308–136317, 2023, doi: 10.1109/ACCESS.2023.3334916.

[17] S. Kanthimathi, S. Venkatraman, K. S. Jayasankar, T. P. Jiljith, and R. Jashwanth, "A novel self-attention-enabled weighted ensemble-based convolutional neural network framework for distributed denial of service attack classification," IEEE Access, vol. 12, pp. 158629–158646, 2024, doi: 10.1109/ACCESS.2024.3478764.

[18] G. Kirubavathi et al., "Detection and mitigation of TCP-based DDoS attacks in cloud environments using a self-attention and intersample attention transformer model," J. Supercomput., 2025, doi: 10.1007/s11227-025-06940-5.

[19] D. M. Sharif and H. Beitollahi, "Detection of application-layer DDoS attacks using machine learning and genetic algorithms," Comput. Secur., vol. 135, art. no. 103511, 2023, doi: 10.1016/j.cose.2023.103511.

[20] D. Akgun, S. Hizal, and U. Cavusoglu, "A new DDoS attacks intrusion detection model based on deep learning for cybersecurity," Comput. Secur., vol. 118, art. no. 102748, 2022, doi: 10.1016/j.cose.2022.102748.

[21] C. Kemp, C. Calvert, T. M. Khoshgoftaar, and J. L. Leevy, "An approach to application-layer DoS detection," J. Big Data, vol. 10, no. 1, art. no. 22, 2023, doi: 10.1186/s40537-023-00699-3.

[22] K. J. Pradeep and P. K. Shukla, "Designing a novel network anomaly detection framework using multi-serial stacked network with optimal feature selection procedures over DDoS attacks," Int. J. Intell. Netw., vol. 6, pp. 1–13, 2025, doi: 10.1016/j.ijin.2024.11.001.

[23] H. Yu, W. Yang, B. Cui et al., "Enhanced anomaly traffic detection framework using BiGAN and contrastive learning," Cybersecurity, vol. 7, art. no. 71, 2024, doi: 10.1186/s42400-024-00297-7.

[24] M. Al-Eryani et al., "A deep learning GRU-BiLSTM for DDoS attack detection," SN Comput. Sci., 2025, doi: 10.1007/s42979-025-04110-1.

[25] A. S. Zaidoun and Z. Lachiri, "A hybrid deep learning model for multi-class DDoS detection in SDN networks,"Ann. Telecommun., vol. 80, no.5–6, pp.459–472, 2025, doi: 10.1007/s12243-025-01085-1.

[26] N. A. Al-Shukaili, M. L. M. Kiah, and I. Ahmedy, "Optimizing feature selection and deep learning techniques for precise detection of low-rate distributed denial of service (LDDoS) attack," Discover Internet Things, vol. 5, no. 1, art. no. 80, 2025.

[27] I. Sharafaldin, A. H. Lashkari, S. Hakak, and A. A. Ghorbani, "Developing realistic distributed denial of service (DDoS) attack dataset and taxonomy," in Proc. IEEE Int. Carnahan Conf. Security Technol. (ICCST), 2019, pp. 1–8, doi: 10.1109/CCST.2019.8888419. [FOUNDATIONAL]

[28] D. Pinto, I. Amorim, E. Maia, and I. Praça, "A review on intrusion detection datasets: Tools, processes, and features," Comput. Netw., vol. 262, art. no. 111177, 2025, doi: 10.1016/j.comnet.2025.111177.