IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 15, No. 4, August 2026, pp. 3712~3721
ISSN: 2252-8938, DOI: 10.11591/ijai.v15.i4.pp3712-3721

Feature-guided transformer approach for detecting distributed denial of service attacks

Lokeshwaran Kanagaraj[1], Raguraman Purushothaman[2], Sathya Subramanian[3], Durga Devi Saravanan[4], Komal Kumar Napa[5], Cornelius Karunakaran[6], Billa Manindhar[7]
[1]Department of Computer Science and Engineering, School of Computing, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi, India
[2]Department of Computer Science & Engineering (Artificial Intelligence), Madanapalle Institute of Technology & Science, Madanapalle, India
[3]Department of Electronics and Communication Engineering, S.A. Engineering College, Chennai, India
[4]Department of Computer Science and Engineering, Vel Tech High Tech Dr. Rangarajan Dr. Sakunthala Engineering College, Avadi, India
[5]Department of Artificial Intelligence and Data Science, Saveetha Engineering College, Chennai, India.
[6]Department of Artificial Intelligence and Data Science, RMK College of Engineering and Technology, Thiruvallur, India
[7]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, India

Article Info

Article history:

Received Sep 11, 2025
Revised Jul 11, 2026
Accepted Jul 21, 2026

Keywords:

Deep learning for cybersecurity
Distributed denial of service
Intrusion detection systems
Network traffic classification
Transformer model

ABSTRACT

Distributed denial of service (DDoS) attacks continues to pose serious risks to modern networks, with their growing intensity making early detection both critical and challenging. Conventional machine learning (ML) models often struggle with the nonlinear and highly dynamic nature of attack traffic, which motivates the use of advanced architectures. In this study investigate a transformer-based classifier for DDoS detection on the CIC-DDoS2019 dataset. The workflow included preprocessing, feature scaling, and domain-guided feature selection. Logistic regression (LR) was employed as a baseline, achieving 92.1% accuracy and F1-score of 0.90, thereby revealing the limitations of linear models. The transformer, after hyperparameter tuning and 5-fold cross-validation, reached an average accuracy of 99.95% with precision, recall, and F1-scores all above 99.9%. The model demonstrated stable convergence and generalization across folds. These results highlight the strength of attention mechanisms in capturing feature dependencies, while also pointing to future directions such as real-time deployment, explainability, and resilience to zero-day attacks.

Corresponding Author:

Lokeshwaran Kanagaraj
Department of Computer Science and Engineering, School of Computing
Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology
No.42, Avadi-Vel Tech Road, Vel Nagar, Avadi, Chennai, Tamil Nadu 600062, India
Email: k.lokeshwaran@gmail.com

1. INTRODUCTION

Distributed denial of service (DDoS) attacks remain one of the most disruptive threats to modern networked systems, capable of overwhelming target resources and causing severe financial, reputational, and operational damage. The widespread availability of botnets, reflection and amplification techniques, and attack-as-a-service platforms has significantly lowered the barrier to launching large-scale attacks, making timely detection increasingly challenging [1], [2]. Traditional defense mechanisms such as signature-based intrusion detection systems, rule-based firewalls, and threshold-driven anomaly detectors are effective

against known attack patterns but suffer from limited adaptability and poor scalability in dynamic network environments, often leading to high false alarm rates or missed detections [3], [4].

To address these limitations, machine learning (ML) techniques have been widely adopted for DDoS detection. Classical models, including logistic regression (LR), decision trees (DT), random forests (RF), and support vector machines (SVM), can learn discriminative traffic patterns but rely heavily on manual feature engineering and shallow decision boundaries, limiting their ability to capture complex nonlinear relationships in large-scale attack traffic [5], [6]. Deep learning models such as convolutional and recurrent neural networks (RNNs) improve representation learning and temporal modeling but are often computationally expensive, sensitive to hyperparameter choices, and prone to overfitting when applied to high-dimensional tabular flow data [7], [8].

Attention mechanisms, particularly transformer architectures, provide a promising alternative by selectively focusing on informative feature interactions through self-attention, without relying on recurrence or convolution. This capability enables effective modeling of complex dependencies among heterogeneous traffic features and has gained increasing attention in cybersecurity research [9]. The effectiveness of such learning-based approaches, however, depends critically on the availability of realistic and diverse datasets. The CIC-DDoS2019 dataset addresses many limitations of earlier benchmarks by incorporating multiple contemporary attack families, realistic benign traffic, and comprehensive flow-level features, making it a widely adopted benchmark for DDoS detection [10].

Motivated by these developments, this work proposes a feature-guided transformer-based framework for DDoS detection. By integrating domain-informed feature selection with a compact attention-based architecture, the proposed approach aims to achieve high detection accuracy. It also maintains computational efficiency and robustness, supported by reproducible evaluation and fair baseline comparison.

2. RELATED WORK

This section reviews existing research on DDoS detection, covering survey studies, classical and deep learning approaches, feature selection techniques, and recent transformer-based intrusion detection systems. Several surveys provide comprehensive overviews of DDoS attack taxonomies, traffic characteristics, and mitigation strategies. These surveys highlight the evolution from volumetric floods to sophisticated multi-vector and application-layer attacks and underscoring the need for adaptive detection mechanisms [11]–[13].

Early learning-based DDoS detection methods primarily employed classical ML models such as LR, SVM, and RF trained on handcrafted flow-level features. While these approaches improved detection accuracy over rule-based systems, their dependence on manual feature engineering and limited ability to model nonlinear traffic behavior restricted generalization under dynamic network conditions [14]–[16]. Hybrid and ensemble-based models were subsequently introduced to enhance robustness by combining multiple classifiers or statistical components; however, these methods often incur higher computational costs and reduced interpretability, limiting real-time applicability [17], [18].

With the adoption of deep learning, convolutional and RNNs were applied to intrusion detection to capture spatial and temporal dependencies in network traffic. Although CNN- and RNN-based models demonstrated improved performance, they are frequently sensitive to data representation and prone to overfitting when applied to high-dimensional tabular features without sufficient regularization or data volume [19]–[21]. Consequently, feature selection has emerged as a critical component, with recent studies exploring filter-based, wrapper-based, and hybrid techniques to reduce redundancy, improve generalization, and enhance computational efficiency for both classical and deep learning models [22]–[24]. More recently, attention-based models and transformer architectures have gained prominence in network security research. By leveraging self-attention mechanisms, transformers can effectively model complex feature interactions and have achieved near state-of-the-art performance on benchmark intrusion detection datasets [25]–[28]. However, many existing approaches rely on high-dimensional feature sets and deep architectures, increasing overfitting risk and limiting deployability. Moreover, evaluations are often restricted to single datasets without explicitly addressing feature redundancy or model compactness. These limitations motivate the development of feature-guided, lightweight transformer-based frameworks that balance accuracy, efficiency, and generalization, forming the basis of the approach proposed in this work [29], [30].

3. METHOD

The methodology of this research has been organized in a step-by-step manner so that each stage can be understood in continuity with the others. The structure broadly follows seven phases: dataset acquisition, data preprocessing, feature selection, baseline model construction, transformer model design,

training with hyperparameter tuning, and evaluation with appropriate performance metrics. Each phase is described in detail as follows.

3.1. Dataset description and preprocessing

The CIC-DDoS2019 dataset was used for evaluation due to its wide adoption and comprehensive coverage of DDoS scenarios. Introduced by Sharafaldin *et al.* [10], includes benign traffic and attacks from 11 DDoS families such as synchronize (SYN) flood, domain name system (DNS), network time protocol (NTP), simple service discovery protocol (SSDP), and WebDDoS, generated in an isolated testbed with realistic background traffic. Eighty flow-level statistical features were extracted using CICFlowMeter. Preprocessing involved removing duplicates, imputing missing or zero values with the median, and applying min–max normalization. Stratified sampling was used to split the data into 80% training and 20% testing sets, with 5-fold cross-validation employed for robust evaluation.

3.2. Feature selection and exploration

Feature selection is critical for efficient intrusion detection, as using all available attributes often introduces redundancy and noise. In this study, domain-guided recommendations from Sharafaldin *et al.* [10] were combined with exploratory analysis of the CIC-DDoS2019 dataset to identify discriminative features. As shown in Figure 1, although DDoS samples dominate, both classes are sufficiently represented. Statistical analysis highlighted packet length–based attributes as strong indicators of attack behavior, which was further confirmed through feature importance analysis, as shown in Table 1. Based on these findings, a refined subset of 11 features was selected and used consistently for both baseline and transformer models.

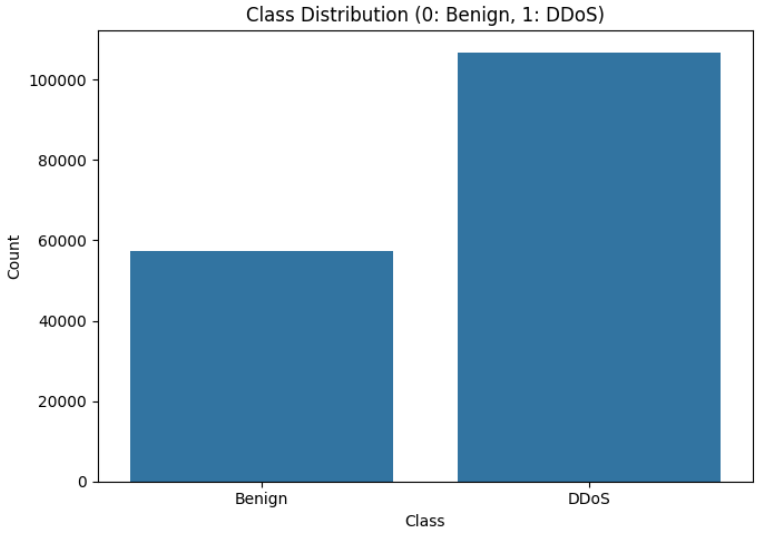

Figure 1. Class distribution

Table 1. Feature set (only selected)

Rank	Feature name	Category	Importance (relative)	Decision
1	Bwd packet length max	Packet length statistics	High	Selected
2	Fwd packet length min	Packet length statistics	High	Selected
3	Bwd packets length total	Packet length statistics	High	Selected
4	Fwd packets length total	Packet length statistics	High	Selected
5	Flow packets/s	Traffic intensity	High	Selected
6	Total Fwd packets	Traffic intensity	High	Selected
7	ACK flag count	Header/flag field	High	Selected
8	Idle mean	Flow timing	Medium	Selected
9	Packet length mean	Packet length statistics	Medium	Selected
10	Bwd packet length min	Packet length statistics	Medium	Selected
11	Flow duration	Flow timing	Medium	Selected

Here, feature-guided detection refers to domain-informed feature selection before training, rather than naively supplying all extracted features. While self-attention can internally reweight inputs, providing

all features increases learning complexity and overfitting risk. By retaining only informative attributes, the proposed approach reduces noise and enables the transformer to focus on meaningful feature interactions, leading to improved accuracy and generalization, as demonstrated by the ablation study.

3.3. Baseline models

A LR classifier was implemented as the baseline. LR is widely used in intrusion detection research due to its simplicity and interpretability. It assumes a linear decision boundary, making it a natural benchmark against which to compare more advanced architectures. In this study, LR was trained with the liblinear solver and balanced class weights. The model provided insight into how well simple linear separation could perform on CIC-DDoS2019.

3.4. Transformer model design

The proposed transformer was designed for DDoS detection on tabular network flow data. Each flow was represented as a fixed-length sequence of 10-time steps with 16 features per step and projected into a 256-dimensional embedding space. A single transformer block with two-head self-attention was used to capture feature dependencies, followed by layer normalization, residual connections, a compact feed-forward layer (dimension 2), and dropout (0.3) for regularization.

Global average pooling was applied across the sequence dimension to obtain a fixed-length representation, which was then passed to a lightweight multilayer perceptron (MLP) with 128 and 64 units and a sigmoid output for binary classification. The overall architecture is illustrated in Figure 2. This compact design balances expressiveness and efficiency, enabling accurate DDoS detection with low computational overhead.

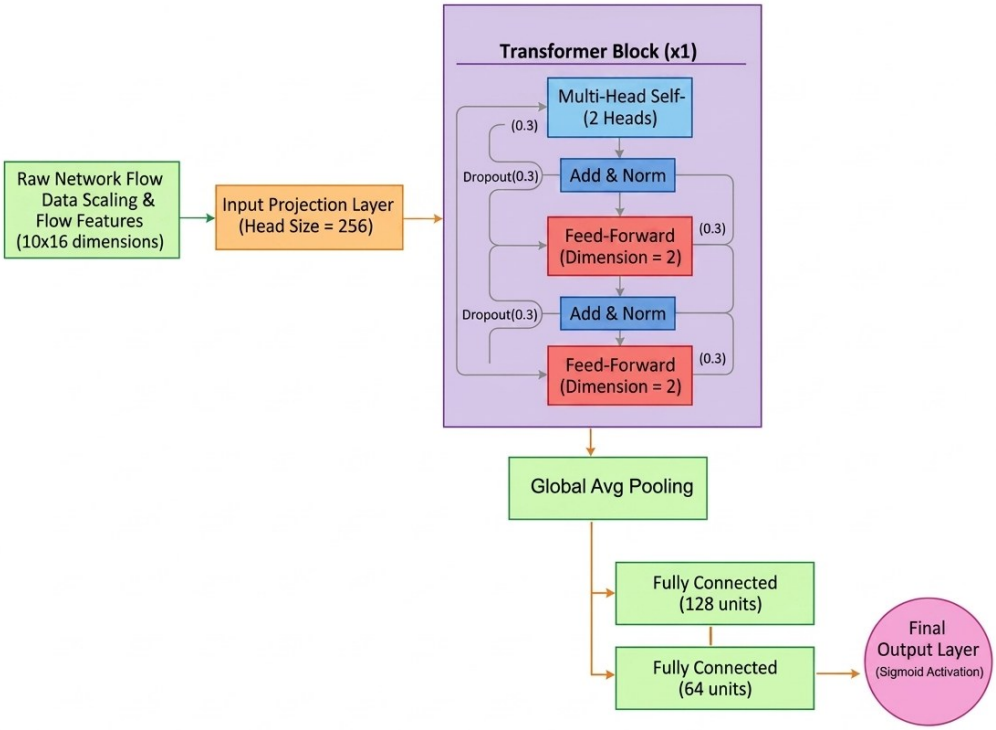

Figure 2. Proposed transformer model architecture

3.5. Hyperparameter tuning

Hyperparameter tuning was performed using Keras tuner with random search, optimizing validation accuracy and validated via 5-fold cross-validation. The search covered attention head size, number of heads, feed-forward dimension, transformer blocks, dropout, and MLP units. The chosen configuration (Table 2) head size 256, two heads, feed-forward dimension 2, one transformer block, MLP [128, 64], and dropout 0.3—reflects a compact design tailored for tabular flow features. Deeper or wider models showed no significant gains and increased overfitting risk, making the shallow, regularized architecture an efficient and generalizable choice.

Table 2. Hyperparameter configuration

Hyperparameter	Selected value	Rationale
Head size	256	Large embedding dimension for feature interactions
Num heads	2	Captures sufficient attention diversity without high cost
Feed-forward dimension	2	Compact layer, reduces overfitting risk
Num transformer blocks	1	Shallow but effective, enabled faster convergence
MLP units	[128, 64]	Balanced dense layers for classification
Dropout	0.3	Regularization to improve generalization

3.6. Training and evaluation protocol

The transformer and LR models were trained using supervised learning on the CIC-DDoS2019 dataset with an identical feature set to ensure fair comparison. The transformer was optimized using Adam with a learning rate of 0.001 and binary cross-entropy loss, trained with a batch size of 64 for up to 50 epochs, and regularized using early stopping based on validation loss. Model robustness was ensured through 5-fold cross-validation, with final results reported as the average across all folds.

Performance was evaluated using accuracy, precision, recall, and F1-score to account for class imbalance and asymmetric error costs. Confusion matrices, receiver operating characteristic (ROC)-area under the curve (AUC), and precision–recall curves were also analyzed to assess misclassification behavior. These metrics further demonstrate the discriminative capability under imbalanced traffic conditions.

4. RESULTS AND DISCUSSION

4.1. Baseline model performance

LR was adopted as the baseline classifier. It delivered reasonable performance, achieving an accuracy of 92.1%, an F1-score of 0.90, and a ROC-AUC of 0.91. These results indicate that certain attack patterns in CIC-DDoS2019 are linearly separable. However, analysis of the confusion matrix showed that LR struggled to detect complex traffic, particularly MSSQL amplification and NetBIOS floods, where the separation between benign and malicious flows is inherently nonlinear.

4.2. Feature selection ablation

An ablation study was performed to assess the impact of feature selection using three configurations: all features, the proposed feature-guided subset, and a randomly selected subset of equal size. All experiments used the same transformer architecture and 5-fold cross-validation. As shown in Table 3, the feature-guided subset achieves the highest accuracy and F1-score with substantially fewer features. The all-feature configuration shows slightly weaker generalization due to redundancy, while the random subset performs notably worse, confirming the benefit of informed feature selection.

Table 3. Feature selection ablation results

Feature set	Features	Accuracy (%)	F1-score (%)
All features	80	99.72	99.70
Feature-guided (proposed)	11	99.95	99.96
Random subset	11	98.40	98.10

4.3. Transformer model performance

The proposed transformer classifier, incorporating dense projections, multi-head self-attention, a compact feed-forward layer, dropout, and global average pooling, significantly outperformed the LR baseline. Using the tuned hyperparameters, the model achieved consistently high performance under 5-fold cross-validation, with an average accuracy of 99.95%, precision of 99.97%, recall of 99.95%, and F1-score of 99.96%, indicating strong generalization across folds. The confusion matrix in Figure 3 shows that misclassifications were negligible for both benign and attack flows. In addition, the training curve and validation curve as in Figure 4 demonstrate stable convergence with closely aligned accuracy and loss trends, confirming that the compact transformer design achieves near-perfect detection without overfitting.

The ROC curve (Figure 5) further validates this performance, with the AUC approaching 1.0. Likewise, the precision-recall curve (Figure 6) demonstrates that the transformer sustains both high recall and high precision, even under imbalanced conditions. Together, these results confirm that the transformer offers highly reliable detection of DDoS traffic, outperforming classical baselines across all evaluation perspectives.

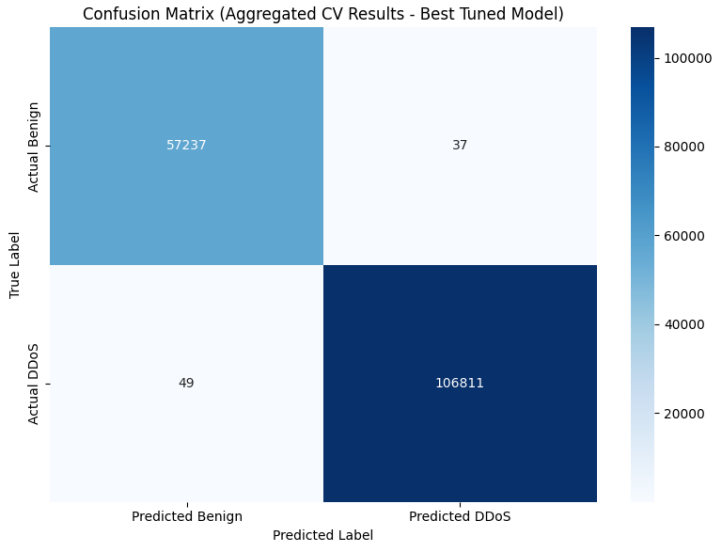

Figure 3. Confusion matrix

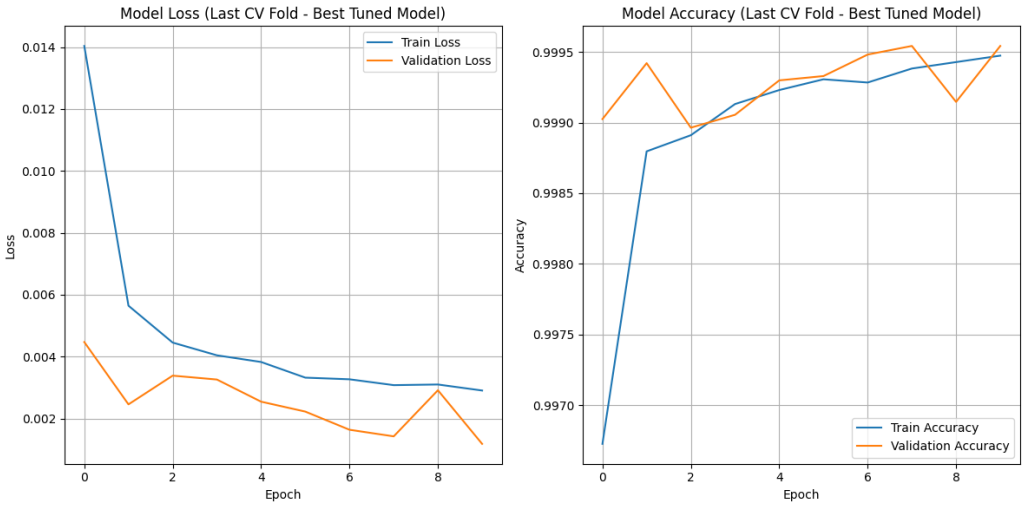

Figure 4. Training and validation curves

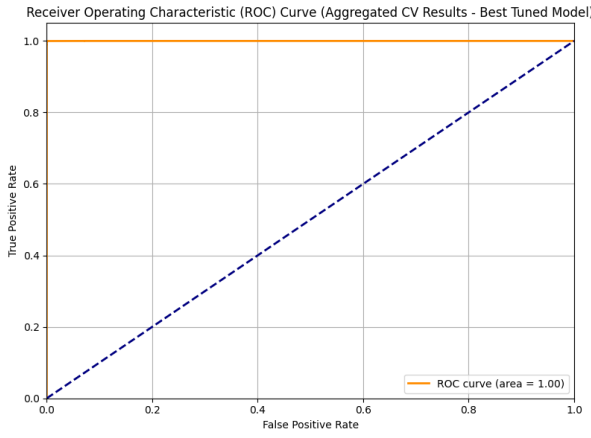

Figure 5. ROC curve

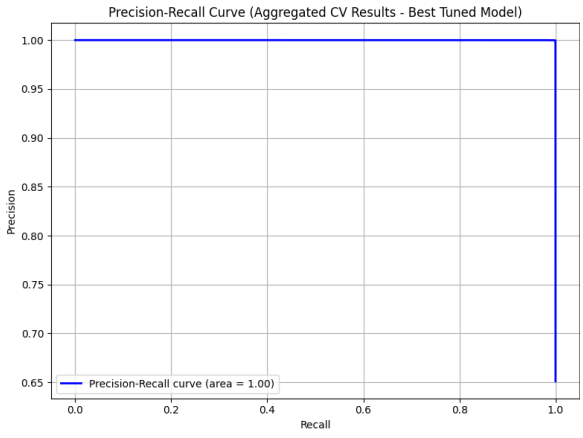

Figure 6. Precision-recall curve

4.4. Comparative evaluation

The comparative analysis highlights the clear advantage of the transformer model over the LR baseline for DDoS detection. LR, trained on the selected 11 features, achieved an accuracy of 92.1% with an F1-score of 0.90, indicating limited ability to model the nonlinear boundaries present in complex attack scenarios. In contrast, the transformer demonstrated consistently high and stable performance across 5-fold cross-validation, as shown by the fold-wise results in Table 4. The average performance summarized in Table 5 (accuracy = 99.95%, precision = 99.97%, recall = 99.95%, and F1-score = 99.96%) demonstrates a substantial improvement over the baseline while maintaining a well-balanced trade-off between false positives and false negatives. These consistently high and balanced metrics confirm the effectiveness of the proposed transformer architecture, particularly its self-attention mechanism, in capturing complex dependencies and discriminative patterns within network traffic for reliable DDoS detection.

Table 4. 5-fold cross-validation results

Fold	Accuracy	Precision (DDoS)	Recall (DDoS)	F1-score (DDoS)
1	0.9994	0.9998	0.9992	0.9995
2	0.9995	0.9999	0.9994	0.9996
3	0.9995	0.9994	0.9998	0.9996
4	0.9995	0.9998	0.9994	0.9996
5	0.9995	0.9994	0.9999	0.9996
Avg.	0.9995	0.9997	0.9995	0.9996

Table 5. Baseline and transformer model comparison

Model	Accuracy	Precision	Recall	F1-score
LR	0.9210	0.9100	0.9000	0.9000
Transformer (Avg. CV)	0.9995	0.9997	0.9995	0.9996

4.5. Computational complexity and efficiency

The proposed transformer model was intentionally designed to be lightweight to support practical deployment. Due to its compact architecture (single transformer block and reduced feed-forward dimension), training converged within a small number of epochs under 5-fold cross-validation. Inference involves a single forward pass through a shallow attention module and a compact MLP, resulting in low latency and modest memory usage. Compared to deeper transformer variants, the proposed model significantly reduces computational overhead while maintaining high detection performance, making it suitable for near real-time intrusion detection scenarios.

4.6. Statistical significance and confidence analysis

To assess the stability of the reported performance, we computed 95% confidence intervals for the evaluation metrics across the 5-fold cross-validation. The narrow confidence bounds observed for accuracy and F1-score indicate low variance across folds, confirming that the near-perfect performance of the proposed transformer is consistent and not attributable to random data partitioning. This further supports the model's generalization capability and alleviates concerns regarding overfitting.

4.7. Critical discussion and limitations

The results reveal a substantial performance gap between LR and the proposed transformer. LR achieved 92.1% accuracy and an F1-score of 0.90, reflecting its limitations on complex DDoS traffic. In contrast, the transformer achieved 99.95% average accuracy with balanced precision and recall (≈99.9%) under 5-fold cross-validation. Low variance indicates strong generalization due to regularized design, feature selection, and self-attention. Despite strong results, evaluation is limited to CIC-DDoS2019, motivating future real-world, cross-dataset, and zero-day studies. The evaluation is limited to the CIC-DDoS2019 dataset, chosen for its realistic traffic and diverse DDoS attacks. Variations in features, traffic patterns, and labeling across datasets such as CIC-IDS2017, UNSW-NB15, and Bot-IoT may cause dataset shift and affect generalization. Leave-one-family-out testing, a strong zero-day evaluation, requires extensive retraining and is beyond this revision’s scope.

5. CONCLUSION

This work presented a transformer-based framework for DDoS attack detection using the CIC-DDoS2019 dataset, addressing the limitations of traditional classifiers in modeling complex attack

behaviors. A structured pipeline involving preprocessing, scaling, and feature selection was adopted, with LR serving as a baseline. The proposed model integrated dense projections, multi-head attention, feed-forward layers, dropout, and global average pooling. Experimental results demonstrated superior performance, achieving 99.95% average accuracy with balanced precision and recall, supported by confusion matrix and ROC/precision–recall analyses. Feature selection further improved accuracy while reducing complexity, highlighting the efficiency of compact transformer designs. Although evaluation was limited to offline experiments on a single dataset, the feature-guided transformer shows strong potential for cross-dataset and zero-day detection, which will be investigated in future work.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Lokeshwaran Kanagaraj	✓	✓	✓	✓	✓	✓			✓	✓	✓			
Raguraman Purushothaman		✓		✓	✓	✓	✓			✓				
Sathya Subramanian	✓	✓	✓		✓			✓	✓		✓			
Durga Devi Saravanan		✓			✓	✓			✓	✓		✓	✓	
Komal Kumar Napa	✓			✓	✓		✓	✓		✓	✓	✓	✓	
Cornelius Karunakaran			✓		✓	✓	✓			✓				
Billa Manindhar					✓			✓	✓					

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The author declares that there are no known conflicts of interest associated with this publication. There are no financial or personal relationships that could inappropriately influence or bias the content of this work.

DATA AVAILABILITY

The experiments in this study were carried out using the CIC-DDoS2019 dataset, which is openly available on Kaggle at https://www.kaggle.com/datasets/dhoogla/cicddos2019.

REFERENCES

[1] T. Mahjabin, Y. Xiao, G. Sun, and W. Jiang, "A survey of distributed denial-of-service attack, prevention, and mitigation techniques," *International Journal of Distributed Sensor Networks*, vol. 13, no. 12, 2017, doi: 10.1177/1550147717741463.
[2] O. Osanaiye, K.-K. R. Choo, and M. Dlodlo, "Distributed denial of service (DDoS) resilience in cloud: review and conceptual cloud DDoS mitigation framework," *Journal of Network and Computer Applications*, vol. 67, pp. 147–165, May 2016, doi: 10.1016/j.jnca.2016.01.001.
[3] W. Khreich, B. Khosravifar, A. H.-Lhadj, and C. Talhi, "An anomaly detection system based on variable N-gram features and one-class SVM," *Information and Software Technology*, vol. 91, pp. 186–197, 2017, doi: 10.1016/j.infsof.2017.07.009.
[4] P. D. Bojović, I. Bašičević, S. Ocovaj, and M. Popović, "A practical approach to detection of distributed denial-of-service attacks using a hybrid detection method," *Computers and Electrical Engineering*, vol. 73, no. 14, pp. 84–96, 2019, doi: 10.1016/j.compeleceng.2018.11.004.
[5] Q. Niyaz, W. Sun, and A. Y. Javaid, "A deep learning-based DDoS detection system in software-defined networking (SDN)," *ICST Transactions on Security and Safety*, vol. 4, no. 12, 2017, doi: 10.4108/eai.28-12-2017.153515.
[6] M. Najafimehr, S. Zarifzadeh, and S. Mostafavi, "A hybrid machine learning approach for detecting unprecedented DDoS attacks," *Journal of Supercomputing*, vol. 78, no. 6, pp. 8106–8136, 2022, doi: 10.1007/s11227-021-04253-x.
[7] A. Aldweesh, A. Derhab, and A. Z. Emam, "Deep learning approaches for anomaly-based intrusion detection systems: a survey, taxonomy, and open issues," *Knowledge-Based Systems*, vol. 189, 2020, doi: 10.1016/j.knosys.2019.105124.

[8] C. Ma, X. Du, and L. Cao, "Analysis of multi-types of flow features based on hybrid neural network for improving network anomaly detection," *IEEE Access*, vol. 7, pp. 148363–148380, 2019, doi: 10.1109/ACCESS.2019.2946708.
[9] A. Vaswani *et al.*, "Attention is all you need," *Advances in Neural Information Processing Systems*, vol. 2017, pp. 5999–6009, 2017, doi: 10.1201/9781003561460-19.
[10] I. Sharafaldin, A. H. Lashkari, S. Hakak, and A. A. Ghorbani, "Developing realistic distributed denial of service (DDoS) attack dataset and taxonomy," in *2019 International Carnahan Conference on Security Technology (ICCST)*, 2019, pp. 1–6, doi: 10.1109/CCST.2019.8888419.
[11] A. Garg, S. Kamalanathan, P. Suhasaria, and A. Agarwal, "Distributed denial of service attacks: a survey," *Imperial Journal of Interdisciplinary Research*, vol. 3, no. 3, pp. 31–37, 2017.
[12] S. Hajj, R. El Sibai, J. B. Abdo, J. Demerjian, A. Makhoul, and C. Guyeux, "Anomaly-based intrusion detection systems: the requirements, methods, measurements, and datasets," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 4, 2021, doi: 10.1002/ett.4240.
[13] A. Garg and P. Maheshwari, "A hybrid intrusion detection system: a review," in *10th International Conference on Intelligent Systems and Control, ISCO 2016*, 2016, pp. 1–5, doi: 10.1109/ISCO.2016.7726909.
[14] N. Bindra and M. Sood, "Evaluating the impact of feature selection methods on the performance of the machine learning models in detecting DDoS attacks," *Romanian Journal of Information Science and Technology*, vol. 23, no. 3, pp. 250–261, 2020.
[15] K. Bouzoubaa, Y. Taher, and B. Nsiri, "Predicting DOS-DDOS attacks: review and evaluation study of feature selection methods based on wrapper process," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 5, pp. 132–145, 2021, doi: 10.14569/IJACSA.2021.0120517.
[16] O. Osanaiye, H. Cai, K. K. R. Choo, A. Dehghantanha, Z. Xu, and M. Dlodlo, "Ensemble-based multi-filter feature selection method for DDoS detection in cloud computing," *Eurasip Journal on Wireless Communications and Networking*, vol. 2016, no. 1, 2016, doi: 10.1186/s13638-016-0623-3.
[17] S. Almutairi, S. Mahfoudh, S. Almutairi, and J. S. Alowibdi, "Hybrid botnet detection based on host and network analysis," *Journal of Computer Networks and Communications*, vol. 2020, 2020, doi: 10.1155/2020/9024726.
[18] S. Sokkalingam and R. Ramakrishnan, "An intelligent intrusion detection system for distributed denial of service attacks: a support vector machine with hybrid optimization algorithm based approach," *Concurrency and Computation: Practice and Experience*, vol. 34, no. 27, 2022, doi: 10.1002/cpe.7334.
[19] M. A. H. Azmi, C. F. M. Foozy, K. A. M. Sukri, N. A. Abdullah, I. R. A. Hamid, and H. Amnur, "Feature selection approach to detect DDoS attack using machine learning algorithms," *International Journal on Informatics Visualization*, vol. 5, no. 4, pp. 395–401, 2021, doi: 10.30630/JOIV.5.4.734.
[20] J. David and C. Thomas, "DDoS attack detection using fast entropy approach on flow-based network traffic," *Procedia Computer Science*, vol. 50, pp. 30–36, 2015, doi: 10.1016/j.procs.2015.04.007.
[21] R. Bitit, A. Derhab, M. Guerroumi, and F. A. Khan, "DDoS attack forecasting based on online multiple change points detection and time series analysis," *Multimedia Tools and Applications*, vol. 83, no. 18, pp. 53655–53685, 2024, doi: 10.1007/s11042-023-17637-3.
[22] L. Hong, K. Wehbi, and T. H. Alsalah, "Hybrid feature selection for efficient detection of DDoS attacks in IoT," in *ACM International Conference Proceeding Series*, 2022, pp. 120–127, doi: 10.1145/3556677.3556687.
[23] F. Kamalov, S. Moussa, Z. El Khatib, and A. Ben Mnaouer, "Orthogonal variance-based feature selection for intrusion detection systems," in *2021 International Symposium on Networks, Computers and Communications, ISNCC 2021*, 2021, pp. 1–5, doi: 10.1109/ISNCC52172.2021.9615656.
[24] A. Maslan, K. M. B. Mohamad, A. Hamid, H. Pangaribuan, and S. Sitohang, "Feature selection to enhance DDoS detection using hybrid N-gram heuristic techniques," *International Journal on Informatics Visualization*, vol. 7, no. 3, pp. 815–822, 2023, doi: 10.30630/joiv.7.3.1533.
[25] H. Wang and W. Li, "DDosTC: a transformer-based network attack detection hybrid mechanism in SDN," *Sensors*, vol. 21, no. 15, 2021, doi: 10.3390/s21155047.
[26] Y. Liu and L. Wu, "Intrusion detection model based on improved transformer," *Applied Sciences*, vol. 13, no. 10, 2023, doi: 10.3390/app13106251.
[27] Z. Long, H. Yan, G. Shen, X. Zhang, H. He, and L. Cheng, "A transformer-based network intrusion detection approach for cloud security," *Journal of Cloud Computing*, vol. 13, no. 1, 2024, doi: 10.1186/s13677-023-00574-9.
[28] 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, doi: 10.1016/j.dcan.2023.03.008.
[29] M. A. Hossain and M. S. Islam, "Enhancing DDoS attack detection with hybrid feature selection and ensemble-based classifier: a promising solution for robust cybersecurity," *Measurement: Sensors*, vol. 32, 2024, doi: 10.1016/j.measen.2024.101037.
[30] J. Li, X. Tong, J. Liu, and L. Cheng, "An efficient federated learning system for network intrusion detection," *IEEE Systems Journal*, vol. 17, no. 2, pp. 2455–2464, 2023, doi: 10.1109/JSYST.2023.3236995.

BIOGRAPHIES OF AUTHORS

Dr. Lokeshwaran Kanagaraj is currently working as associate professor in the Department of Computer Science and Engineering at Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi, India. His research interests include machine learning, data mining, semantic web, linked open data, and cloud computing. He can be contacted at email: k.lokeshwaran@gmail.com.

Mr. Raguraman Purushothaman is currently working as an assistant professor in the Department of Computer Science & Engineering (Artificial Intelligence), Madanapalle Institute of Technology & Science Deemed to be University, Madanapalle, India. His area of interest includes theory of computation, design and analysis of algorithms, image processing, and data science. He can be contacted at email: yuvaragu.pt@gmail.com.

Dr. Sathya Subramanian is currently working as associate professor in the Department of Electronics and Communication Engineering, S.A. Engineering College, affiliated to Anna University, Chennai. Her areas of interest and research domain are image processing, embedded systems, and MEMS. She can be contacted at email: sathyas7979@gmail.com.

Dr. Durga Devi Saravanan is associate professor in the Department of Computer Science and Engineering at Vel Tech High Tech Dr. Rangarajan Dr. Sakunthala Engineering College. Her area of interest is network security, wireless systems, and IoT. She has vast experience in academia and administration. She has published and presented papers in various international and national journals/conferences. She is an active member of professional bodies CSI and ISTE. She can be contacted at email: sdurgadevi@velhightech.com.

Dr. Komal Kumar Napa is currently working as assistant professor (SG) in the Department of Artificial Intelligence and Data Science at Saveetha Engineering College, Chennai, India. His research interests include machine learning, data mining, and cloud computing. He can be contacted at email: komalkumarnapa@gmail.com.

Dr. Cornelius Karunakaran is currently working as associate professor in the Department of Artificial Intelligence and Data Science at RMK College of Engineering and Technology, Tiruvallur, Tamil Nadu, India. His research interests include big data analytics, artificial intelligence, machine learning, and cloud computing. He can be contacted at email: cornelius851@gmail.com.

Mr. Billa Manindhar is currently working as an assistant professor in the Department of Computer Science and Engineering at Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India. His research interests include machine learning, data mining, and cloud computing. He can be contacted at email: bmanindhar@kluniversity.in.