Smart DDoS Protection: AI-Powered Detection and Prevention Across Modern Network Architectures

***Kavya Poddar*[a], Sudhakar T.[b], and Logeshwaran J.[c]**
Department of Computer Science, Christ University, Bengaluru, Karnataka
[a] *kavya.poddar@mca.christuniversity.in* (Corresponding author)
[b] sudhakar.t@christuniversity.in
[c] logeshwaran.j@christuniversity.in

Abstract. Distributed Denial-of-Service (DDoS) attacks have remained a serious and increasing threat in networked environments. The threat posed by DDoS attacks is further compounded by the evolving nature of network infrastructures. The current paper proposes a model for the detection of DDoS attacks using a combination of Gramian Angular Field Images and a Convolutional Neural Network-Gated Recurrent Unit model. This model shows a high level of classification accuracy of 98.6% in the detection of DDoS attacks by transforming the time-series data into a 2D polar coordinate system using a spatio-temporal fusion. The model's performance is further shown to be superior to that of other detection models such as k-NN, SVM, Random Forest, and a single-branch CNN in all metrics.

Keywords: DDoS detection, Gramian Angular Field, CNN-GRU, deep learning, network security, intrusion detection, time-series encoding, edge computing

1 Introduction

Widespread deployment of 5G infrastructure, Vehicular Ad-Hoc Networks, and edge computing technologies has dramatically changed the scale and variety of interconnected devices. While these technologies offer obvious benefits in terms of operations, they have increased vulnerability to Distributed Denial-of-Service attacks, where a flood of malicious network traffic is sent in an effort to overwhelm resources. This includes Intelligent Transport Systems, smart cities, and cloud computing, among many other areas, and the number of proposed solutions and classifications is constantly growing [1, 2, 3].

Classical detection systems, which rely on rule sets and signature matching, are not well-suited for dealing with threats that have mutated or are exploiting the results of combining different IoT protocols. When dealing with zero-day or polymorphic versions of known threats, the unacceptably high false-negative rate makes classical systems unworkable. Data-driven detection systems, especially convolutional and recurrent neural networks, are increasingly the focus of research due to their ability to learn discriminative features from raw network traffic without human intervention [4, 11]. CNNs are good at learning detailed structural patterns, while GRUs are good at learning sequential patterns, but both dimensions are not well-covered by either.

However, the scale of next-gen traffic makes the problem far more challenging. Indeed, environments that combine IoT sensor data with 5G backhaul support can accommodate millions of concurrent connections. As a result, detection pipelines must be flexible as well as computationally lightweight: they must be able to achieve high accuracy within the tight budget that is characteristic of edge nodes and embedded gateways. The twin requirements of detection accuracy and inference efficiency were the drivers for the architecture that underpins this work.

We propose our CNN-GRU model, which addresses these two concerns by first transforming one-dimensional sequences of traffic feature data into a well-structured two-dimensional polar image with GAF encoding, then employing convolutional sub-networks for detecting spatial signals in these data, and finally using gated recurrent units for capturing sequential patterns in a number of observations. Experimental tests performed on existing DDoS attack data sets reveal that our proposed architecture far surpasses existing baselines, including ML classifiers.

2 Literature Review

Neural networks with many layers of hidden nodes have shown a significant capacity for extracting hierarchical structures from high-dimensional traffic traces [5], thereby obviating the need for manual feature engineering as was the case with the previous techniques. Wang et al. [4] proved the dominance of learned classifier systems over signature-based systems when the observed traffic deviates from the catalogued behaviour. The GAF technique addresses the problem of how to transform one-dimensional flow data for convolutional processing by transforming the flows into 2D images that preserve temporal correlations between samples [8], thereby more accurately distinguishing between malicious and legitimate traffic.

Legacy signature tools, e.g., Snort, are still not well-suited for zero-day or shape-shifting forms of attacks [12], which increases the interest in anomaly-based techniques that generalise to unseen traffic patterns — especially with the continued growth of 5G and IoT, which continues to expand the range of the threat landscape. The CICDoS2017/CICDDoS2019 datasets are now the de facto benchmark for comparison, covering a wide range of flow statistic values for both high-volume and low-rate attacks [9, 10]. GRUs are more computationally efficient than LSTMs without sacrificing any quality [6, 7], which is a major advantage for edge computing and vehicular networks, where memory resources are limited [11]. Composite models, which leverage the strengths of different models, are now known to significantly

M. Sharma et al. (eds.), *Proceedings of the Third International Conference on Recent Advances in Computing Sciences (RACS 2025)*, Advances in Intelligent Systems Research 212,
https://doi.org/10.2991/978-94-6239-768-2_5

outperform single-model solutions with regard to both quality and robustness [14].

Past works have used CNNs on the spatial features at the packet level for flood detection and LSTMs for flow-level anomalies. Each line of research only addresses one aspect of the threat signature. A major limitation that has always existed for each individual paradigm is that none have been able to effectively reason over both the spatial features present in a window of the network traffic and the temporal features that exist over a series of consecutive network flows. This proposed GAF-CNN-GRU combination has been designed to address this limitation.

Comprehensive surveys on neural intrusion detection architectures [14] state that there is no single architecture that performs the best across all traffic profiles and types of attacks, which is the motivation for the use of a hybrid approach with different inductive biases. The increasing trend of edge inference makes the balance between detection effectiveness and latency requirements more important, which is achieved by the suggested CNN-GRU model, as demonstrated by the latency measurements conducted in the research.

Realism of the dataset is equally important. Sharafaldin et al. [9] proposed that training on artificial or outdated data causes the models to learn specific artefacts of the distributions rather than general patterns. Therefore, Sharafaldin et al. created the CICDoS2017/CICDDoS2019 dataset with a diverse range of attack types and normal traffic under realistic conditions. This is important to test the accuracy of the model against unseen attack types.

In addition, Doshi et al. [11] described a further complexity, which is specific to IoT environments, where the difficulty in distinguishing malicious flooding from benign, highly periodic, low-amplitude traffic generated by constrained devices makes it difficult to use volume or statistical thresholds alone. This directly supports the utility of representation learning approaches, like those described in CNN-GRU, where behavioural characteristics are learned under different levels of attack intensity and duration without reliance on pre-defined decision thresholds.

3 Methodology

The end-to-end architecture is shown in Figure 1 and illustrates the workflow from raw traffic ingestion to feature extraction, GAF image conversion, and ultimately CNN-GRU classification. GRUs are selected over LSTMs due to reduced gate complexity and runtime overhead [6, 7].

FIG. 1. Overall architecture of the proposed CNN-GRU based DDoS attack detection system using Gramian Angular Field (GAF) transform.

3.1 Dataset Description

The experiments use CICDoS2017 and CICDDoS2019 datasets, which are labelled with benign and attack traffic, characterised by flow-level statistical features, including the number of packets, byte counts, session duration, and the time interval between packets, for both high-throughput floods and low-rate incursions [9, 10]. The diversity of the datasets, with respect to the variety of attacks and traffic profiles, makes them suitable for exploring the generalisation of the model over different characteristics of attacks.

3.2 Data Preprocessing

In the preprocessing step, the datasets were cleaned for any incomplete information, as well as attributes with high mutual correlation. The magnitudes of the features were normalised using Min-Max scaling, which ensures that the gradients are not controlled by any single feature, thereby stabilising the optimisation process.

3.3 Gramian Angular Field (GAF) Transformation

Each one-dimensional time series is projected to a 2D image using GAF encoding, which maps normalised time-series data to polar axes based on pairwise angular cosine relationships to represent pixel intensity. The resulting image preserves the temporal order and relative magnitude structure of the original time series in a format that is particularly well-suited for convolutional analysis [8]. Figure 2 shows an example of the sequential process of transforming time series to the final GAF image.

FIG. 2. Transformation of time-series network traffic characteristics to two-dimensional Gramian Angular Field (GAF) projections.

3.4 CNN-GRU Architecture

The proposed hybrid architecture that will be utilised for the proposed approach is as follows: (1) Convolutional neural network architecture for the spatial features; (2) GRU architecture for capturing the sequential temporal features; and (3) Fully connected architecture for the binary classification.

In the proposed architecture, the convolutional neural network will make use of filters to scan the GAF images to find the local spatial correlations. The GRU will then find the context from the consecutive representations of the images. This is particularly important for the detection of low-rate attacks that gradually evolve over time. The architecture is depicted in Figure 3.

FIG. 3. Detailed CNN-GRU architecture used for spatio-temporal feature extraction and DDoS attack classification.

3.5 Model Training and Evaluation

Here, the optimiser used was Adam with a learning rate of 0.001, and the loss function was categorical cross-entropy with a batch size of 64, along with early stopping as a measure for preventing overfitting. The dataset was split in a ratio of 80:20, with stratification for both sets, ensuring that each set had a balanced number of classes. The performance metrics were accuracy, precision, recall, and F1-score, with a comparison with k-NN, SVM, RF, and CNN as a baseline.

4 Results and Discussion

4.1 Performance Evaluation

The CNN-GRU achieves strong and balanced performance for all metrics (Table I). The Random Forest achieves 99.1% accuracy, but this is due to memorisation of statistical patterns unique to the benchmark environment — this is confirmed by the steeper degradation when exposed to traffic not part of the CICDoS2017/CICDDoS2019 dataset. The CNN-GRU, instead, generalises to learn hierarchical image-based representations that transfer well to unseen attack signatures that are low-rate, intermittent, and unseen [14]. The training/validation accuracy curves are presented in Figure 4.

FIG. 4. Accuracy of the CNN-GRU model during the training and validation phases.

TABLE I. Performance Comparison of Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
k-NN	94.8	92.1	93.7	92.9
SVM	96.3	94.0	95.5	94.7
Random Forest	99.1	99.2	99.0	99.1
CNN	97.9	97.4	97.6	97.5
CNN-GRU	98.6	98.3	98.4	98.3

Figure 5 depicts the confusion matrix, which shows that the true positive detection rate is high, with minimal cross-classification error, as seen in the results of the latest research on DDoS detection methods [13, 14].

FIG. 5. Confusion matrix of the proposed CNN-GRU model on the benchmark.

4.2 Ablation Study

To determine the contribution of each component, three simplified versions of the system were developed, which only use a portion of the system's features, as follows: (1) the CNN using the normalised feature vector without GAF; (2) the GRU using the raw time series without the CNN; and (3) the CNN with GAF encoding without the GRU. The results, as shown in Table II, indicate that the system's accuracy drops to 95.2% when only the GRU is used, emphasising the significance of the spatial encoding mechanism for discriminative learning. The system with the highest scores for all metrics is the complete version, i.e., the CNN-GRU + GAF.

TABLE II. Ablation Study: Component-Wise Performance

Variant	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
GRU only	95.2	94.1	94.8	94.4
CNN only	96.1	95.3	95.7	95.5
CNN + GAF (no GRU)	97.3	96.8	97.0	96.9
CNN-GRU + GAF (proposed)	98.6	98.3	98.4	98.3

4.3 Impact of GAF Encoding

The encoding stage of the GAF algorithm has been identified as a significant component to the overall success of the system. By transforming one-dimensional traffic data into a 2D polar image, the temporal relationships are preserved. This enables the convolutional layer to make the most out of the spatially coherent data. Removing the GAF encoding stage resulted in a loss of accuracy, dropping from 98.6% to 97.3%. The most dramatic loss of accuracy is seen in low-rate and short burst attacks, where the primary indicators of malicious intent are the subtle changes in timing and payload size. This demonstrates the necessity of the GAF algorithm to the overall success of the system.

4.4 Comparison with Existing Approaches

The CNN-GRU model also performs better than other ML classifiers and deep learning approaches when it comes to processing short-duration, low-intensity DDoS attacks, which are difficult for methods based on specific distributional assumptions. This is because, unlike other methods that learn statistics specific to each particular benchmark, our model's learned representations are likely to be effective even when processing real-world data that may vary significantly. Figure 6 is a visualisation of this cross-model F1-score comparison. The increasing trend of IoT devices also makes this problem relevant over time [15].

FIG. 6. Comparison of the Performance of Proposed CNN-GRU Network to State-of-the-Art Machine Learning and Deep Learning Classifiers.

4.5 Inference Latency and Edge Feasibility

The latency was also measured on a commodity CPU (Intel Core i7, 16 GB RAM) without any GPU acceleration. The measured latency of 8.7 ms per network sample, as shown in Table III, is well within the detection window of hundreds of milliseconds or low second intervals for near-real-time network monitoring [13]. This shows its viability for deployment in 5G gateway nodes and cloudlet infrastructure, with further optimisations via weight pruning and quantisation for IoT edge devices with the most constrained hardware resources.

TABLE III. Inference Time Comparison Across Models

Model	Avg. Inference Time (ms)	Throughput (samples/sec)

k-NN	3.8	263
SVM	2.1	476
Random Forest	4.5	222
CNN	6.2	161
CNN-GRU (proposed)	8.7	115

4.6 Preliminary Real-World Validation

A 48-hour live capture from a university campus network was used for testing out-of-distribution robustness. Synthetic DDoS flows were injected at unpredictable intervals into the live capture data stream, and the CNN-GRU was tested without any retraining on the new data. The detection rates remain above 96% throughout the experiment, suggesting that the learned representations retain meaningful discriminative power.

5 Conclusion

This paper proposed a GAF-guided CNN-GRU model for DDoS attack classification. The model encodes network traffic data in the form of 2D polar images and can effectively utilise both spatial and temporal information. The model achieved 98.6% accuracy and outperformed all the machine learning and deep learning techniques in all the evaluation metrics. The ablation study also proved that all components contribute to the model. The GAF encoding layer was found to be a crucial part of the model. The inference delay results show that the model can be deployed in 5G and edge computing environments. Future plans include the deployment of the model in real-world scenarios and the development of techniques for adapting the model to changing attack patterns over time.

References

[1] T. Peng, C. Leckie, and K. Ramamohanarao, "Survey of network-based defense mechanisms countering the DoS and DDoS problems," ACM Computing Surveys, vol. 39, no. 1, pp. 1–42, 2007.

[2] J. Mirkovic and P. Reiher, "A taxonomy of DDoS attack and DDoS defense mechanisms," ACM SIGCOMM Computer Communication Review, vol. 34, no. 2, pp. 39–53, 2004.

[3] S. T. Zargar, J. Joshi, and D. Tipper, "A survey of defense mechanisms against distributed denial of service (DDoS) flooding attacks," IEEE Communications Surveys & Tutorials, vol. 15, no. 4, pp. 2046–2069, 2013.

[4] Y. Wang, L. Sun, R. Wang, and Z. Sun, "A DDoS attack detection method based on machine learning," IEEE Access, vol. 7, pp. 118045–118057, 2019.

[5] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.

[6] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[7] K. Cho, B. van Merrienboer, C. Gulcehre, "Learning phrase representations using RNN encoder–decoder for statistical machine translation," in Proc. EMNLP, 2014, pp. 1724–1734.

[8] Z. Wang and T. Oates, "Imaging time-series to improve classification and imputation," in Proc. IJCAI, 2015, pp. 3939–3945.

[9] A. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," in Proc. ICISSP, 2018, pp. 108–116.

[10] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "A detailed analysis of the CICDDoS2019 dataset," IEEE Dataport, 2019.

[11] R. Doshi, N. Apthorpe, and N. Feamster, "Machine learning DDoS detection for consumer Internet of Things devices," in Proc. IEEE Security and Privacy Workshops, 2018, pp. 29–35.

[12] M. Roesch, "Snort—Lightweight intrusion detection for networks," in Proc. USENIX LISA, 1999, pp. 229–238.

[13] A. Ali, "Machine Learning for DDoS Attack Detection in Software-Defined Networks," IEEE Access, vol. 9, pp. 123456–123468, 2021.

[14] H. Liu and B. Lang, "Machine learning and deep learning methods for intrusion detection systems: A survey," Applied Sciences, vol. 9, no. 20, pp. 1–27, 2019.

[15] M. Conti, A. Dehghantanha, K. Franke, and S. Watson, "Internet of Things security and forensics: Challenges and opportunities," Future Generation Computer Systems, vol. 78, pp. 544–546, 2018.