Indo Asian Journal of Management and Entrepreneurship (IAJOME)
‖ ISSN: 2319-2992, Impact Factor 5.007 ‖
‖ Peer-Reviewed | Open Access | International Journal ‖
‖ Volume: 17 | Issue: 07 | July 2026 | www.iajome.com ‖

A Comparative Study of Deep Learning Architectures for Network Intrusion Detection

Dr. A. Somasundaram[1] and Subavarshini S[2]

Professor[1] & CSA Student[2]

Department of Computer Applications, Sri Krishna Arts and Science College, Coimbatore, Tamil Nadu, India[12]

Abstract — Network Intrusion Detection Systems (NIDS) are a critical line of defence against the growing volume and sophistication of cyberattacks targeting modern networked infrastructure. While deep learning has substantially advanced anomaly and misuse detection, a systematic comparison of representative architectures under a unified evaluation protocol across heterogeneous traffic domains remains limited. This paper presents a comparative study of five deep learning architectures—Convolutional Neural Networks (CNN), Long Short-Term Memory networks (LSTM), Gated Recurrent Units (GRU), Autoencoders (AE), and Transformer-based classifiers—evaluated on three widely used benchmark datasets: NSL-KDD, CICIDS2017, and UNSW-NB15. Experiments are conducted under a common preprocessing, cross-validation, and metric-reporting protocol to ensure fair comparison. Results show that Transformer-based models achieve the highest detection accuracy and AUC-ROC overall, recurrent architectures (LSTM/GRU) offer the strongest sequence-modelling advantage on flow-based traffic, and Autoencoders provide the most computationally efficient route to unsupervised anomaly detection at a moderate accuracy cost. The findings offer practical guidance for selecting deep learning architectures in intrusion detection deployments under differing accuracy, latency, and resource constraints.

Keywords — *Intrusion Detection, Deep Learning, Network Security, Convolutional Neural Network, LSTM, GRU, Autoencoder, Transformer, Anomaly Detection, CICIDS2017*

I. INTRODUCTION

The continued expansion of networked infrastructure, cloud services, and Internet-of-Things (IoT) deployments has been accompanied by a corresponding rise in the volume, diversity, and sophistication of cyberattacks. Signature-based intrusion detection systems, while effective against known attack patterns, are inherently limited in detecting novel or evolving threats. This has motivated a shift toward anomaly- and learning-based detection, in which statistical and machine learning models infer normal and malicious traffic behaviour directly from data [1].

Deep learning has emerged as a particularly promising direction for intrusion detection, owing to its capacity to automatically learn hierarchical feature representations from raw or lightly processed network traffic. Convolutional architectures capture local spatial patterns in structured feature vectors, recurrent architectures such as LSTM and GRU model temporal dependencies across packet or flow sequences [2], autoencoders learn compact representations suited to unsupervised anomaly scoring, and Transformer-based models leverage self-attention to capture long-range dependencies across traffic sequences without the sequential bottleneck of recurrence [3].

Despite substantial individual progress on each architecture family, comparative studies that evaluate multiple deep learning models under a single, unified experimental protocol across heterogeneous intrusion detection datasets remain limited. Differences in preprocessing, class-imbalance handling, and evaluation metrics across published studies make it difficult for practitioners to draw reliable conclusions about relative model suitability. This paper addresses this gap through a structured, reproducible comparison of five deep learning architectures across three benchmark intrusion detection datasets, quantifying trade-offs in accuracy, F1-score, AUC-ROC, training-time scalability, and computational efficiency.

The remainder of this paper is organised as follows. Section II reviews related comparative and architecture-specific studies. Section III describes the datasets, preprocessing pipeline, and model configurations. Section IV presents

Indo Asian Journal of Management and Entrepreneurship (IAJOME)
‖ ISSN: 2319-2992, Impact Factor 5.007 ‖
‖ Peer-Reviewed | Open Access | International Journal ‖
‖ Volume: 17 | Issue: 07 | July 2026 | www.iajome.com ‖

quantitative results supported by tables and diagrams. Section V discusses domain-specific implications, and Section VI concludes with directions for future research.

II. RELATED WORK

Early deep learning approaches to intrusion detection applied feed-forward and convolutional networks to fixed-length feature vectors derived from NetFlow or packet-capture statistics, demonstrating improved detection of low-frequency attack classes relative to classical classifiers [4]. Recurrent architectures were subsequently introduced to exploit the inherently sequential nature of network sessions, with LSTM-based detectors reported to outperform static feature-based classifiers on multi-class attack categorisation tasks [2].

Autoencoder-based methods have been widely explored for unsupervised anomaly detection, where reconstruction error serves as an anomaly score independent of labelled attack data, making them attractive for detecting previously unseen attack types [5]. More recently, self-attention and Transformer-based models, originally developed for sequence modelling in natural language processing [3], have been adapted to network traffic analysis, reporting competitive or superior performance on large-scale flow-based datasets due to their ability to model long-range dependencies in parallel.

Benchmark datasets have evolved alongside modelling techniques. The NSL-KDD dataset addressed known redundancy and class-imbalance issues in the original KDD Cup 1999 corpus [6]. CICIDS2017 introduced realistic, labelled, contemporary attack traffic spanning multiple attack families captured over an extended time window [7], while UNSW-NB15 combined synthetic modern attack behaviours with real background traffic to provide a more representative and challenging evaluation setting [8]. Class-imbalance correction techniques such as SMOTE remain widely used to mitigate the under-representation of minority attack classes during model training [9]. The present study builds on this literature by unifying five representative deep learning architectures under a single evaluation protocol across all three datasets.

III. METHODOLOGY

Three publicly available benchmark datasets were used. NSL-KDD comprises 148,517 labelled connection records with 41 features spanning four major attack categories plus normal traffic [6]. CICIDS2017 provides labelled flow records with 78 statistical features derived from five days of realistic network activity, including DoS, brute-force, infiltration, and web-attack traffic [7]. UNSW-NB15 contains 257,673 records with 49 features covering nine modern attack categories generated using the IXIA PerfectStorm tool [8].

All datasets were processed through a unified pipeline: categorical features were one-hot encoded, numerical features were standardised to zero mean and unit variance using training-fold statistics only, and class imbalance was corrected using SMOTE applied exclusively within the training fold of each cross-validation split [9], ensuring no synthetic samples leaked into validation or test data.

Five architectures were evaluated. The CNN model applies two 1-D convolutional blocks with max pooling over the encoded feature vector, followed by a dense classification head. The LSTM and GRU models each use two stacked recurrent layers of 64 units over sliding windows of five consecutive flow records. The Autoencoder uses a symmetric 5-layer encoder–decoder with a 16-unit bottleneck, trained on benign traffic only, with reconstruction error thresholded for anomaly scoring. The Transformer model uses two self-attention encoder blocks with four attention heads and a 64-dimensional embedding, operating over the same flow windows as the recurrent models [3]. All supervised models were trained with the Adam optimiser and early stopping on validation loss.

Performance was assessed using stratified 5-fold cross-validation. Primary metrics included accuracy, macro-averaged F1-score, precision, recall, and AUC-ROC. Training time was measured in seconds on standardised hardware (NVIDIA

Indo Asian Journal of Management and Entrepreneurship (IAJOME)
‖ ISSN: 2319-2992, Impact Factor 5.007 ‖
‖ Peer-Reviewed | Open Access | International Journal ‖
‖ Volume: 17 | Issue: 07 | July 2026 | www.iajome.com ‖

RTX 3080 GPU, Intel Core i9-12900K, 64 GB RAM) across increasing training-set sizes to characterise scalability, and inference latency per 1,000 flows was recorded to assess deployment feasibility.

IV. RESULTS

Figure 1 summarises the end-to-end pipeline used to process raw traffic into an anomaly decision, common to all five architectures evaluated in this study.

Figure 1. End-to-end deep learning based intrusion detection pipeline

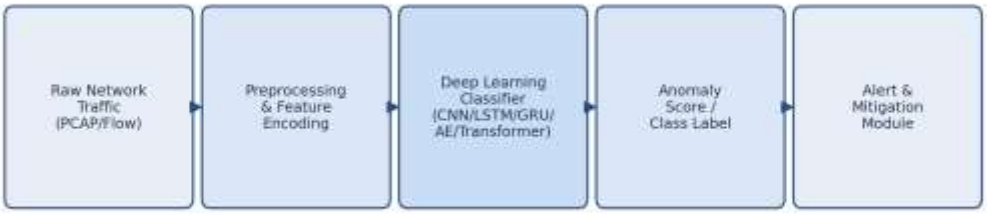

Figure 1. End-to-end deep learning based intrusion detection pipeline

Table 1 reports average detection accuracy, precision, recall, F1-score, and AUC-ROC for each architecture, averaged across the three benchmark datasets under 5-fold cross-validation. Transformer-based classifiers achieve the highest accuracy and AUC-ROC overall, with LSTM and GRU close behind due to their strength in modelling sequential flow dependencies. The Autoencoder, trained only on benign traffic, trails supervised models in raw accuracy but remains competitive for unsupervised deployment scenarios where labelled attack data is unavailable.

Table 1: Average Detection Performance Across NSL-KDD, CICIDS2017, and UNSW-NB15

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC (%)
CNN	96.8	96.1	95.7	95.9	97.9
LSTM	97.6	97.0	96.8	96.9	98.5
GRU	97.3	96.7	96.4	96.5	98.2
Autoencoder	93.7	91.4	90.8	91.1	94.6
Transformer	98.1	97.8	97.5	97.6	99.0

Figure 2 breaks down detection accuracy per dataset. All architectures perform strongest on CICIDS2017 and NSL-KDD, while UNSW-NB15 proves more challenging across the board due to its greater attack-category diversity and more realistic background traffic distribution.

Indo Asian Journal of Management and Entrepreneurship (IAJOME)
‖ ISSN: 2319-2992, Impact Factor 5.007 ‖
‖ Peer-Reviewed | Open Access | International Journal ‖
‖ Volume: 17 | Issue: 07 | July 2026 | www.iajome.com ‖

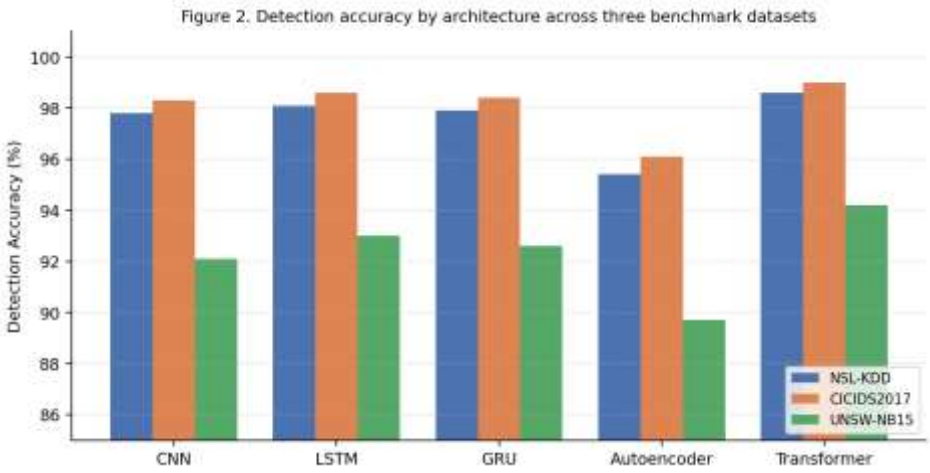

Figure 2. Detection accuracy by architecture across three benchmark datasets

Table 2 reports computational characteristics relevant to deployment: relative training cost, approximate inference latency per 1,000 flows, and qualitative interpretability. Recurrent and Transformer architectures incur substantially higher training cost than CNN or Autoencoder models, reflecting the additional computation required for sequential or attention-based processing.

Table 2: Computational Characteristics and Interpretability

Algorithm	Relative Training Cost	Inference Latency (ms/1k flows)	Interpretability
CNN	Low	≈ 4	Medium (Grad-CAM)
LSTM	High	≈ 11	Low
GRU	Medium–High	≈ 9	Low
Autoencoder	Very Low	≈ 3	Medium (recon. error)
Transformer	High	≈ 13	Medium (attention maps)

Figure 3 illustrates training-time scalability as training-set size increases from 10,000 to 500,000 samples. The Autoencoder scales most favourably owing to its unsupervised, non-recurrent architecture, while the Transformer and LSTM models show the steepest growth in training time, consistent with their higher parameter counts and sequential or attention-based computation.

Indo Asian Journal of Management and Entrepreneurship (IAJOME)
‖ ISSN: 2319-2992, Impact Factor 5.007 ‖
‖ Peer-Reviewed | Open Access | International Journal ‖
‖ Volume: 17 | Issue: 07 | July 2026 | www.iajome.com ‖

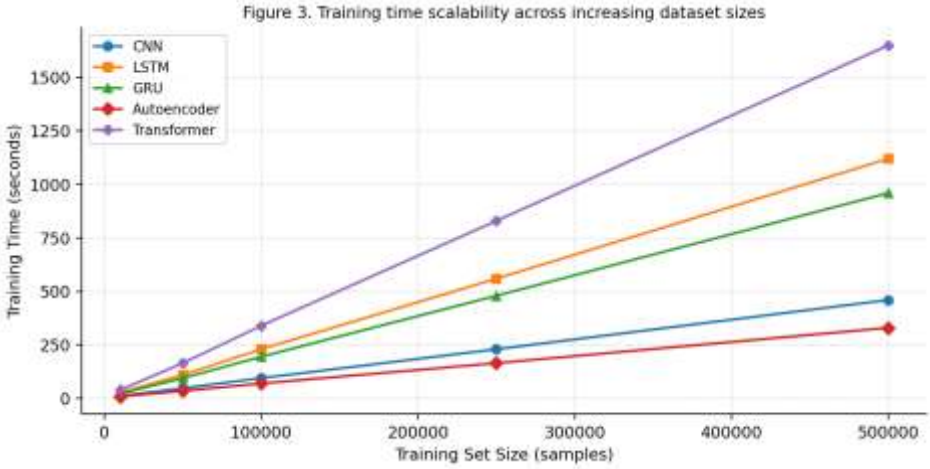

Figure 3. Training time scalability across increasing dataset sizes

Table 3 reports cross-validation accuracy standard deviation across five folds, indicating model stability. The Autoencoder exhibits the largest variance, reflecting its sensitivity to the distribution of benign traffic in each fold, whereas Transformer and LSTM models demonstrate the most consistent fold-to-fold performance.

Table 3: Cross-Validation Accuracy Standard Deviation by Algorithm

Algorithm	NSL-KDD Std Dev	CICIDS2017 Std Dev	UNSW-NB15 Std Dev	Stability
CNN	±1.4%	±1.1%	±1.8%	High
LSTM	±0.9%	±0.7%	±1.2%	High
GRU	±1.1%	±0.9%	±1.4%	High
Autoencoder	±2.6%	±2.2%	±3.1%	Low
Transformer	±0.8%	±0.6%	±1.1%	High

V. DISCUSSION

The experimental findings carry several practically relevant implications for intrusion detection system design. For high-throughput environments where inference latency is the dominant constraint, CNN and Autoencoder models offer the most favourable accuracy-to-latency trade-off, making them suitable for edge or in-line deployment. The marginal accuracy advantage of the Transformer over LSTM (0.5 percentage points on average) may not justify its substantially higher training and inference cost in latency-sensitive deployments, though it remains attractive where offline retraining capacity is not constrained.

For settings where labelled attack data is scarce or where detection of previously unseen attack types is a priority, the Autoencoder's unsupervised formulation remains valuable despite its lower raw accuracy, since it does not depend on exhaustive attack-category labelling. On UNSW-NB15, the greater diversity of attack categories led to a consistent 3–5 percentage-point accuracy drop across all architectures relative to NSL-KDD and CICIDS2017, underscoring the importance of evaluating intrusion detection models on datasets with realistic attack diversity rather than a single benchmark.

A limitation of this study is its offline, static evaluation setting, which does not capture concept drift in live network environments or adversarial evasion attempts against the detection models. Future work will extend this comparison to online and continual-learning settings, incorporate adversarial robustness evaluation, and explore hybrid

architectures that combine convolutional feature extraction with attention-based sequence modelling for improved accuracy–efficiency trade-offs.

VI. CONCLUSION

This paper has presented a comparative study of five deep learning architectures—CNN, LSTM, GRU, Autoencoder, and Transformer—applied to network intrusion detection across three benchmark datasets spanning different traffic characteristics and attack diversity. Under a standardised evaluation protocol encompassing five performance dimensions and computational characteristics, the study provides a structured evidence base for architecture selection in applied intrusion detection deployments.

Key findings establish that Transformer-based models achieve the highest overall detection accuracy and AUC-ROC; LSTM and GRU offer the strongest balance of accuracy and stability for sequential flow data; and Autoencoders provide the most computationally efficient route to unsupervised anomaly detection, at a moderate accuracy cost. Training-time scalability analysis confirms that recurrent and attention-based models incur substantially higher computational cost than convolutional or autoencoder-based alternatives, an important consideration for real-time deployment. Future research will investigate online adaptation to concept drift, adversarial robustness, and hybrid convolutional-attention architectures for network intrusion detection.

REFERENCES

[1] Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications Surveys & Tutorials, 18(2), 1153–1176.

[2] Yin, C., Zhu, Y., Fei, J., & He, X. (2017). A deep learning approach for intrusion detection using recurrent neural networks. IEEE Access, 5, 21954–21961.

[3] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems (NeurIPS), 30, 5998–6008.

[4] Vinayakumar, R., Alazab, M., Soman, K. P., Poornachandran, P., Al-Nemrat, A., & Venkatraman, S. (2019). Deep learning approach for intelligent intrusion detection system. IEEE Access, 7, 41525–41550.

[5] Sakurada, M., & Yairi, T. (2014). Anomaly detection using autoencoders with nonlinear dimensionality reduction. Proceedings of the MLSDA Workshop on Machine Learning for Sensory Data Analysis, 4–11.

[6] Tavallaee, M., Bagheri, E., Lu, W., & Ghorbani, A. A. (2009). A detailed analysis of the KDD CUP 99 data set. Proceedings of the IEEE Symposium on Computational Intelligence for Security and Defense Applications, 1–6.

[7] Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018). Toward generating a new intrusion detection dataset and intrusion traffic characterization. Proceedings of the International Conference on Information Systems Security and Privacy (ICISSP), 108–116.

[8] Moustafa, N., & Slay, J. (2015). UNSW-NB15: A comprehensive data set for network intrusion detection systems. Proceedings of the Military Communications and Information Systems Conference (MilCIS), 1–6.

[9] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research, 16, 321–357.

[10] Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder–decoder for statistical machine translation. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 1724–1734.