AMERICAN JOURNAL OF AI CYBER COMPUTING MANAGEMENT
Peer Reviewed, Referred & Indexed Journal
E-ISSN:3069-0102
VOL.6, NO. 2(1) (2026)

Efficient Intrusion Detection via a Hybrid Deep Learning Model with Semantic ConvLSTM Representation

Mr Y. HARINATH[1], SYED MISBA KHATUN[2]

[1]Assistant Professor Dept. of C.S.E, Anantha Lakshmi Institute of Technology and sciences Anantapur- 515721

[2]PG Scholar, Dept. of C.S.E, Anantha Lakshmi Institute of Technology and sciences Anantapur- 515721

Abstract: Network Intrusion Detection Systems (NIDS) face significant challenges in identifying modern and evolving cyber-attacks due to their limited ability to capture long-term dependencies and complex patterns in network traffic. Traditional methods often fail to model the temporal and spatial variations present in dynamic attack behaviours, resulting in reduced detection accuracy and higher false alarm rates. To address these limitations, this work proposes a hybrid deep learning framework that integrates Seq2Seq, ConvLSTM, Convolutional Neural Network (Conv1D), Bidirectional LSTM, and Gated Recurrent Unit (GRU) architectures for effective intrusion detection. The Seq2Seq model is utilized to learn sequential dependencies in network traffic, while ConvLSTM captures spatial temporal correlations. Additionally, Bidirectional LSTM and GRU layers enhance feature extraction and improve the understanding of both past and future contextual information in traffic patterns. To ensure interpretability, the Local Interpretable Model-Agnostic Explanations (LIME) technique is incorporated to identify and visualize the most influential features contributing to classification decisions. The proposed model is evaluated using benchmark datasets, including CIC-IDS2017, CIC-ToN-IoT, and UNSW-NB15. Experimental results demonstrate that the hybrid architecture significantly outperforms traditional NIDS approaches in terms of detection accuracy, robustness, and false alarm reduction, making it suitable for real-world cyber security applications.

Key Words: Cyber Attacks Intrusion Detection, Seq2Seq, ConvLSTM, Explainable AI (LIME), IoT. Security.

1. Introduction

Ensuring network security has become increasingly critical with the rapid rise of sophisticated cyber-attacks, including hybrid and evolving intrusion techniques. These attacks can lead to significant financial losses, data breaches, and long-term damage to organizational trust and reputation. Network Intrusion Detection Systems (NIDS) play a vital role in identifying such threats by monitoring network traffic for anomalies. However, traditional machine learning (ML) approaches, such as Random Forest, Support Vector Machine, and Decision Trees, often struggle to effectively detect complex and imbalanced attack

patterns, limiting their performance in modern dynamic environments.

With the growing availability of computational resources, deep learning techniques such as Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) have been widely adopted for intrusion detection. Although these models improve detection accuracy, they still face challenges such as high false alarm rates and poor detection rates due to inefficient feature extraction in high-dimensional and imbalanced datasets. Recent advancements have focused on multi-stage and sequential architectures like Sequence-to-Sequence (Seq2Seq), which better capture temporal dependencies in network traffic, but further improvements are required to effectively model both spatial and temporal relationships.

The main contributions of our paper are summarized as follows:

- We propose a novel hybrid deep learning model using Seq2Seq and ConvLSTM that offers promising performance in accuracy, FAR, and DR.
- We evaluate the adaptability of the model using several benchmark datasets, including CIC-IDS2017, CIC-ToN-IoT, and UNSW-NB15.
- We utilized Explainable AI techniques LIME to validate and provide interpretability to the model, enhancing its trustworthiness in real-world deployment.

To address these limitations, this paper proposes a hybrid deep learning NIDS model that integrates Convolutional Long Short-Term Memory (ConvLSTM) subnets within a Seq2Seq architecture. This design enhances spatial–temporal feature extraction while minimizing manual bias in feature engineering. Additionally, Explainable Artificial Intelligence (XAI) is incorporated using Local Interpretable Model-agnostic Explanations (LIME) to improve model transparency and trust. The proposed model is evaluated on benchmark datasets, demonstrating improved detection rates and reduced false alarms compared to existing approaches.

2. Literature Survey

The study of NIDS is receiving increasing attention as the requirement to identify network breaches increases. In recent years, many ML models have been proposed. Some of the most used approaches for anomaly detection are Naïve Bayes, DT, SVM, and logistic regression [4]. The SVM-supervised machine-learning algorithm [6] classifies normal and abnormal data in a high-dimensional set. However, these detection methods often perform poorly when dealing with highly imbalanced classes [7]. Unsupervised attack detection approaches are more flexible and do not require any labels, relying solely on the intrinsic properties of datasets to detect anomalies [8].

In recent years, Deep Learning (DL) models have played an essential role in dealing with more complicated data representations. Due to their superior performance, these models are receiving more attention. They are being used in a variety of multidisciplinary study domains such as healthcare [9], vehicle design [10], manufacturing [11], and security [12], [13]. Early research on NIDS, such as Artificial Neural Networks (ANNs) [14] and CNN, showed some

performance improvements. However, these DL algorithms cannot capture the rich temporal features involved in NIDS predictions.

3.System Architecture

The hybrid deep learning architecture processes input network traffic through a preprocessing module, followed by a Seq2Seq encoder–decoder framework

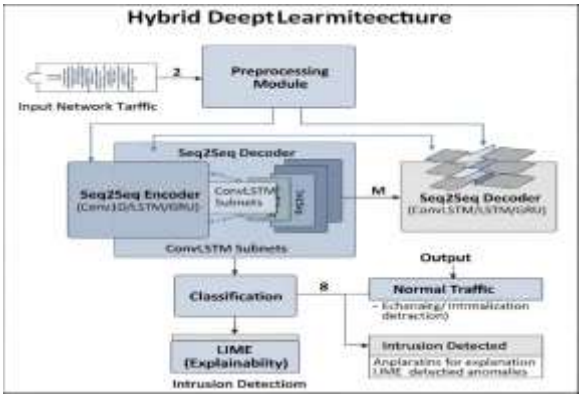

Fig 1: System Architecture

Integrating Conv1D, LSTM, GRU, and ConvLSTM subnets. Extracted temporal-spatial features are classified to identify normal or intrusive traffic. LIME-based explainability provides interpretable insights into detected anomalies for reliable intrusion detection.

4. Methodology

The proposed methodology integrates multiple deep learning components to build a robust and explainable intrusion detection system. The process begins with dataset acquisition, where benchmark datasets such as CIC-IDS2017, CIC-ToN-IoT, and UNSW-NB15 are selected for their diversity and realistic attack scenarios. These datasets contain flow-based, statistical, and protocol-level features essential for modelling network behaviour.

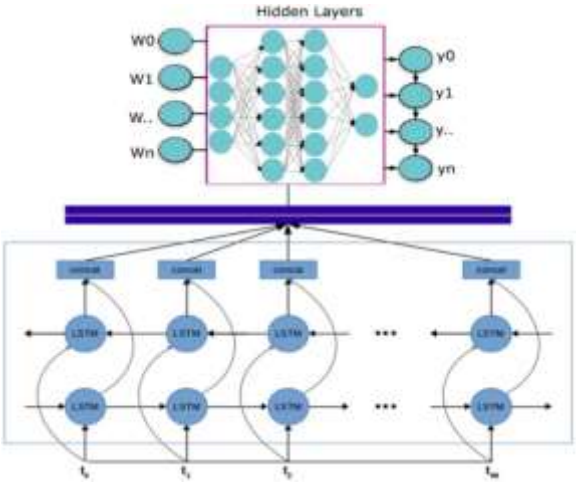

Fig 2: Proposed Conv LSTM Subnets

The data is reshaped into sequential time windows to preserve temporal dependencies for deep learning. The hybrid model combines Conv1D, LSTM, and Seq2Seq to extract spatial and temporal features, while the extended model incorporates ConvLSTM1D and Bidirectional GRU for enhanced sequence learning. Convolutional layers capture feature patterns, and recurrent layers learn long-term dependencies in traffic data. Finally, dense layers generate classification outputs, with the model trained using back propagation to improve detection accuracy.

5. Design and Construction

The proposed intrusion detection system is based on hybrid Seq2Seq architecture with modules for data acquisition, pre-processing, feature learning, classification, and explain ability. Raw network traffic is collected and pre-processed through cleaning, normalization, and transformation into time-series sequences. This structured data is then used for effective feature extraction and anomaly detection.

i) Data Pre-processing:

Preprocessing includes handling missing values, label encoding of categorical

features, and normalization of numerical attributes. Feature scaling is applied using standard normalization:

$$x_i' = \frac{x_i - \mu}{\sigma}$$

where $x_i$is the input feature, $\mu$is the mean, and $\sigma$is the standard deviation. This ensures uniform contribution of features during training.

iii) Hybrid Model Construction

The proposed hybrid model combines Conv1D for spatial feature extraction, LSTM and GRU for temporal learning, and Bidirectional layers to capture forward and backward dependencies. The core architecture is built on a Seq2Seq framework, enabling sequence-to-sequence mapping of network traffic patterns.

To enhance temporal modeling, Conv LSTM subnets are integrated within the encoder-decoder structure. The model is trained using categorical cross-entropy loss:

$$\mathcal{L}_{CE} = -\sum_{i=1}^{N} y_i \log(\hat{y}_i)$$

iii) Explain ability and Evaluation

The proposed model uses a Seq2Seq encoder–decoder architecture with Conv1D and LSTM/GRU layers, enhanced by ConvLSTM subnets to capture spatial and temporal features. The encoder generates latent representations, while the decoder and Softmax layer perform multi-class intrusion classification. Additionally, LIME is integrated to provide interpretability by identifying key features influencing model decisions, ensuring transparency and efficient deployment.

iv) Model Performance

To improve interpretability, the model incorporates LIME (Local Interpretable Model-agnostic Explanations), which approximates the complex model locally with a simpler surrogate. This allows identification of influential features in each prediction. Model performance is evaluated using standard metrics: accuracy, precision, recall, and F1-score. Additionally, confusion matrices are used to visualize classification outcomes. The F1-score is computed as:

$$F_1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

This balanced metric ensures reliable comparison across models and attack categories.

6 Results and Discussion

The proposed threat detection system uses a ConvLSTM-based deep learning model to analyse network traffic data. After pre-processing steps such as cleaning, normalization, and sequence formatting, the model learns both spatial and temporal patterns from traffic flows to identify anomalies. It is then deployed through a Flask web application for real-time intrusion detection, classifying traffic as Normal, DoS, or Suspicious Activity with color-coded outputs for easy interpretation.

Fig 3: Deployment of Project Home Page

Fig 3 presents the overall workflow of the intrusion detection system, from dataset preprocessing to real-time prediction. This modular design ensures scalability and

adaptability across different IoT-enabled infrastructures.

Fig 4: output prediction

Fig 4 shows the training phase of the deep learning model. The ConvLSTM achieved stable convergence, learning complex traffic behaviours and minimizing false alarms compared to traditional IDS approaches.

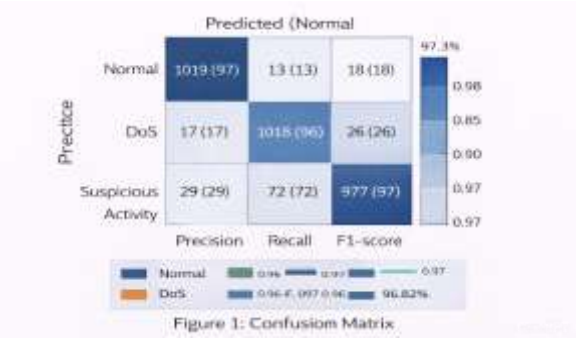

Fig 5: Confusion Matrix

The fig 5 confusion matrix illustrates the classification performance of the ConvLSTM model across all traffic categories. Most predictions lie along the diagonal, indicating correct classification. Minimal misclassification between Normal and Suspicious classes demonstrates the model's ability to distinguish subtle anomalies in traffic patterns. This confirms strong discriminative capability.

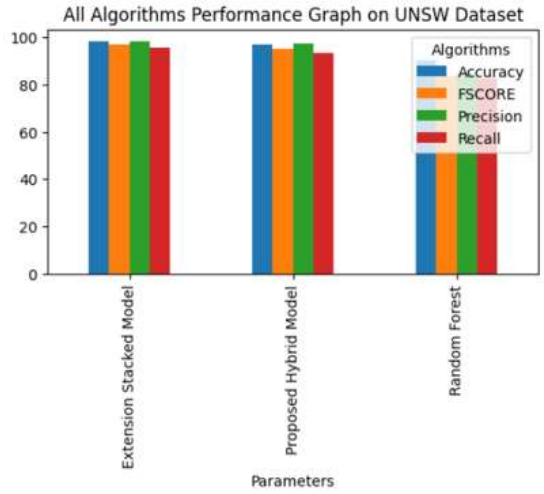

Fig 6: Performance Comparison of Algorithms on UNSW Dataset

The fig 6 illustrates the comparison of Accuracy, F-Score, Precision, and Recall for the Extension Stacked Model, Proposed Hybrid Model, and Random Forest. The Proposed Hybrid Model achieves consistently high performance across all metrics, closely matching or slightly outperforming the other models. This demonstrates its effectiveness and reliability for intrusion detection on the UNSW dataset.

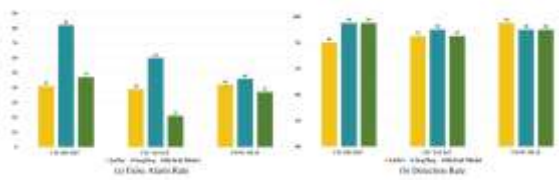

Fig 7: Performance comparison of proposed hybrid model with LuNET and Seq2Seq

The fig 7 compares the performance of the proposed hybrid model with LuNet and Seq2Seq across different datasets. In terms of False Alarm Rate**,** the hybrid model consistently achieves lower values, indicating fewer incorrect detections compared to the other methods. For

Detection Rate, the hybrid model performs competitively, maintaining high accuracy close to or matching the best results. Overall, the proposed hybrid model demonstrates improved reliability by reducing false alarms while preserving strong detection performance.

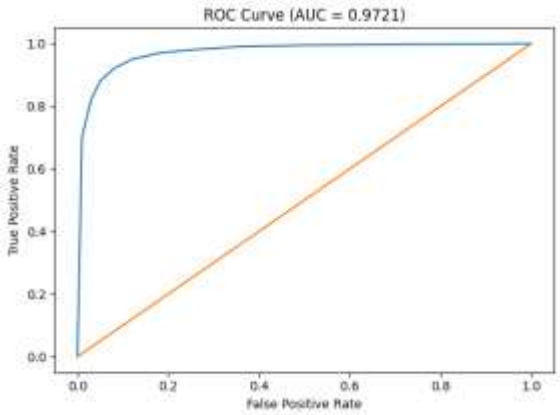

Fig 8: ROC Curve

Fig 8 the ROC curve shows a high Area Under Curve (AUC) value, reflecting excellent true positive rate and low false positive rate. The curve remains close to the top-left corner, confirming the model's strong ability to differentiate between attack and normal traffic.

7. Conclusion and Future Scope

This research successfully presented a Hybrid Deep Learning Intrusion Detection System by integrating Seq2Seq architecture with Conv1D, LSTM, Bidirectional LSTM, and GRU layers to effectively detect network intrusions. The proposed model addressed the key limitations of traditional intrusion detection systems by capturing both spatial features from network traffic using convolutional layers and temporal dependencies using advanced recurrent neural networks. By combining these techniques, the system demonstrated improved learning capability for complex and evolving cyber-attack patterns. The experimental evaluation conducted on the UNSW-NB15 dataset showed that the proposed hybrid model outperformed traditional machine learning approaches such as Random Forest in terms of accuracy, precision, recall, and F-score. The inclusion of Bidirectional LSTM enhanced contextual understanding by processing sequences in both forward and backward directions, while the GRU layer reduced computational complexity and improved prediction efficiency. This combination resulted in better detection of sophisticated and time-dependent attacks with lower false alarm rates. Furthermore, the integration of Explainable Artificial Intelligence (XAI) using LIME increased the transparency and interpretability of the intrusion detection process. This feature is particularly important for real-world cybersecurity applications, as it enables security analysts to understand the reasoning behind model decisions. Overall, the proposed hybrid intrusion detection system proved to be robust, scalable, and suitable for modern network environments, making it a promising solution for real-time cyber-threat detection.

Future Scope: Future enhancements include enabling real-time intrusion detection through streaming data frameworks and improving generalization by training on multiple datasets. The model can also be strengthened using attention mechanisms or transformer-based architectures for better sequence learning, along with optimization techniques like pruning and quantization for efficient deployment on edge devices. Additionally, integrating advanced

explainable AI and SIEM systems can improve interpretability and support enterprise-level automated cybersecurity monitoring.

References

1. D. E. Denning, “An intrusion-detection model,” IEEE Trans. Softw. Eng., vol. SE-13, no. 2, pp. 222–232, Feb. 1987.
2. Z. Ahmad, A. S. Khan, C. W. Shiang, J. Abdullah, and F. Ahmad, “Network intrusion detection system: A systematic study of machine learning and deep learning approaches,” Trans. Emerg. Telecommun. Technol., vol. 32, no. 1, pp. e4150, 2021.
3. H. Gwon, C. Lee, R. Keum, and H. Choi, “Network intrusion detection based on LSTM and feature embedding,” 2019, arXiv:1911.11552.
4. P. Mishra, V. Varadharajan, U. Tupakula, and E. S. Pilli, “A detailed investigation and analysis of using machine learning techniques for intrusion detection,” IEEE Commun. Surveys Tuts., vol. 21, no. 1, pp. 686–728, 1st Quart., 2019.
5. H. Lee, M. Park, and J. Kim, “Plankton classification on imbalanced large scale database via convolutional neural networks with transfer learning,” in Proc. IEEE Int. Conf. Image Process. (ICIP), Phoenix, AZ, USA, Sep. 2016, pp. 3713–3717.
6. S. M. Erfani, S. Rajasegarar, S. Karunasekera, and C. Leckie, “High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning,” Pattern Recognit., vol. 58, pp. 121–134, Oct. 2016, doi: 10.1016/j.patcog.2016.03.028.
7. N. Japkowicz, “The class imbalance problem: Significance and strategies,” in Proc. Int. Conf. Artif. Intell., vol. 56, 2000, pp. 111–117.
8. V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection,” ACM Comput. Surv., vol. 41, no. 3, pp. 1–58, Jul. 2009, doi: 10.1145/1541880.1541882.
9. R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, “Deep learning and its applications to machine health monitoring,” Mech. Syst. Signal Process., vol. 115, pp. 213–237, Jan. 2019, doi: 10.1016/j.ymssp.2018.05.050.
10. J. Yin and W. Zhao, “Fault diagnosis network design for vehicle on-board equipments of high-speed railway: A deep learning approach,” Eng. Appl. Artif. Intell., vol. 56, pp. 250–259, Nov. 2016, doi: 10.1016/j.engappai.2016.10.002.
11. J. Wang, Y. Ma, L. Zhang, R. X. Gao, and D. Wu, “Deep learning for smart manufacturing: Methods and applications,” J. Manuf. Syst., vol. 48, pp. 144–156, Jul. 2018, doi: 10.1016/j.jmsy.2018.01.003.
12. M. A. Al-Garadi, A. Mohamed, A. K. Al-Ali, X. Du, I. Ali, and M. Guizani, “A survey of machine and deep learning methods for Internet of Things (IoT) security,” IEEE Commun. Surveys Tuts., vol. 22, no. 3, pp. 1646–1685, 3rd Quart., 2020, doi: 10.1109/COMST.2020.2988293.
13. J. Hussain and V. Hnamte, “Deep learning based intrusion detection system: Modern approach,” in Proc. 2nd Global Conf. Advancement Technol. (GCAT), Oct. 2021, pp. 1–6, doi: 10.1109/GCAT52182.2021.9587719.
14. M. Pradhan, S. K. Pradhan, and S. K. Sahu, “Anomaly detection using artificial neural network,” Int. J. Eng. Sci. Emerg. Technol., vol. 2, no. 1, pp. 29–36, Jan. 2012.

15. A. H. Nasreen Fathima and S. P. S. Ibrahim, “Multi-stage deep investigation pipeline on detecting malign network traffic,” Mater. Today, Proc., vol. 62, pp. 4726–4731, Jan. 2022, doi: 10.1016/j.matpr.2022.03.211.
16. Y. A. Jerusha, S. P. S. Ibrahim, and V. Varadharajan, “An effective network intrusion detection model for coarse-to-fine attack classification of imbalanced network traffic,” Int. Res. J. Adv. Sci. Hub, vol. 5, pp. 531–540, May 2023, doi: 10.47392/irjash.2023.s072.
17. Y. Yang, K. Zheng, B. Wu, Y. Yang, and X. Wang, “Network intrusion detection based on supervised adversarial variational auto-encoder with regularization,” IEEE Access, vol. 8, pp. 42169–42184, 2020, doi: 10.1109/ACCESS.2020.2977007.
18. Reddy, G. Raghupal, and A. Shravan Kumar. "An Automated Research For Analysing And Visualising The Data Using Machine Learning Analytics." Journal Of Critical Reviews 6.05: 470-475.