CHAPTER 9

Intelligent Cyber Defense Systems Using Deep Learning for Network Threat Detection

Dr. Syeda Farhath Begum

Associate Professor, Department of Computer Science and Engineering, Nawab Shah Alam Khan College of Engineering and Technology, Hyderabad, Telangana, India.
Email: sdfarhath@gmail.com

https://doi.org/10.58599/GSE.2026.310309

Abstract: The proliferation of network-based attacks has created a critical need for advanced, intelligent, and automated cyber defense systems. Traditional security solutions, such as firewalls and signature-based intrusion detection systems (IDS), are increasingly insufficient to counter the dynamic and sophisticated nature of modern cyber threats. This chapter explores the application of deep learning models for network threat detection, providing a comprehensive overview of the foundations, recent advances, and practical applications of these techniques. We delve into the use of various deep learning architectures, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and hybrid models, for analyzing network traffic data and identifying malicious activities. A novel hybrid deep learning model is proposed for enhanced threat detection, and its performance is evaluated using the benchmark NSL-KDD dataset. The results demonstrate the superior accuracy and efficiency of deep learning-based approaches in comparison to traditional methods, highlighting their potential to revolutionize the field of cybersecurity. The chapter concludes with a discussion of the challenges and future research directions in this rapidly evolving domain.

Keywords: Cyber Defense, Deep Learning, Intrusion Detection, Network Security, Threat Intelligence.

1. Introduction

The digital transformation of modern society has led to an unprecedented reliance on computer networks for communication, commerce, and critical infrastructure. This hyper-connectivity, while offering numerous benefits, has also created a vast and complex attack

ISBN: 978-81-994969-8-9 (Print); 978-81-994969-2-7 (Online)

surface for malicious actors. Cyber threats have evolved from simple, isolated incidents to highly organized and persistent campaigns, capable of causing significant financial, reputational, and societal damage [1]. The increasing volume and sophistication of these threats have overwhelmed traditional security measures, which often rely on predefined rules and signatures to detect known attacks. These methods are largely ineffective against zero-day exploits, polymorphic malware, and advanced persistent threats (APTs), which are designed to evade signature-based detection.

To address these challenges, the cybersecurity community has turned to artificial intelligence (AI) and machine learning (ML) techniques to develop more adaptive and intelligent defense systems. Deep learning, a subfield of machine learning, has emerged as a particularly promising approach for network threat detection. Deep learning models, with their ability to automatically learn hierarchical representations from raw data, are well-suited for analyzing the complex and high-dimensional nature of network traffic. These models can identify subtle patterns and anomalies that may be indicative of malicious activity, without the need for manual feature engineering [2].

This chapter provides a comprehensive exploration of deep learning for network threat detection. We begin with a review of the relevant literature, followed by a detailed description of a proposed hybrid deep learning methodology. We then present the results of our experimental evaluation, using the NSL-KDD dataset, and discuss their implications. Finally, we conclude with a summary of our findings and a discussion of future research directions.

2. Literature Review

The application of machine learning to intrusion detection is not a new concept. Early research in this area focused on traditional machine learning algorithms, such as Support Vector Machines (SVMs), Decision Trees, and Naive Bayes [3]. While these methods showed some success, they often required extensive feature engineering and were not always able to capture the complex, non-linear relationships present in network traffic data. The advent of deep learning has opened up new possibilities for building more accurate and robust intrusion detection systems.

Several deep learning architectures have been proposed for network threat detection. For instance, Convolutional Neural Networks (CNNs), which are widely used in image recognition, have been adapted to analyze network traffic by treating it as a one-dimensional or two-dimensional image [4]. Recurrent Neural Networks (RNNs), and their variants such as Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRUs), are well-suited for modeling the sequential nature of network traffic and detecting temporal patterns associated with attacks [5]. Hybrid models that combine the strengths of different architectures have also been explored. For example, a combination of CNN and

ISBN: 978-81-994969-8-9 (Print); 978-81-994969-2-7 (Online)

LSTM can be used to extract both spatial and temporal features from network traffic, leading to improved detection performance [6].

Recent research has also focused on the use of deep learning for anomaly detection, where the goal is to identify deviations from normal network behavior [7]. Autoencoders, a type of neural network that is trained to reconstruct its input, have been used to learn a model of normal network traffic. Any significant deviation from this model can then be flagged as a potential anomaly [8]. Generative Adversarial Networks (GANs) have also been used for anomaly detection, where a generator network tries to create realistic network traffic that can fool a discriminator network, which is trained to distinguish between real and fake traffic. This adversarial training process can help to improve the robustness of the detection model [9].

3. Proposed Methodology

In this section, we propose a hybrid deep learning model for network threat detection that combines the strengths of CNNs and LSTMs. The proposed model is designed to effectively capture both the spatial and temporal characteristics of network traffic, leading to improved detection accuracy and a lower false positive rate. The architecture of the proposed model is shown in Figure 1.

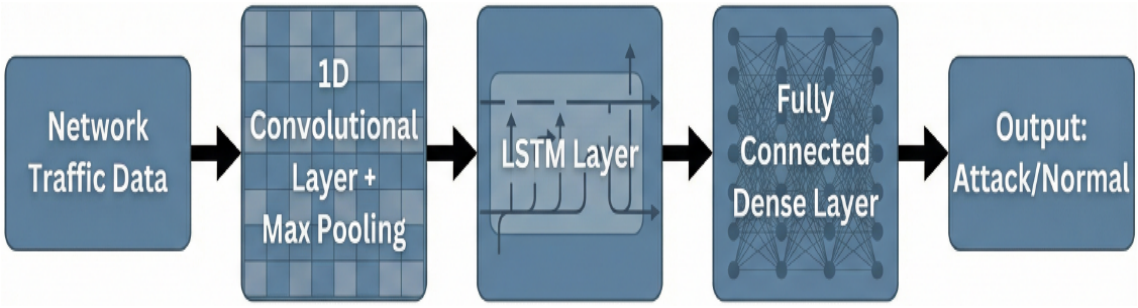

Figure 1: Proposed hybrid CNN-LSTM model for network threat detection.

The proposed model consists of the following layers:

1. **Input Layer:** The input to the model is a sequence of network traffic records, where each record is represented as a vector of numerical features.

2. **Convolutional Layer:** A one-dimensional CNN layer is used to extract local features from the input sequence. The CNN layer applies a set of filters to the input sequence, where each filter learns to detect a specific pattern.

3. **Max Pooling Layer:** A max pooling layer is used to down-sample the output of the convolutional layer, reducing its dimensionality and making the model more robust to small variations in the input.

ISBN: 978-81-994969-8-9 (Print); 978-81-994969-2-7 (Online)

4. **LSTM Layer:** An LSTM layer is used to model the temporal dependencies in the sequence of features extracted by the CNN layer. The LSTM layer is able to learn long-term dependencies, which is important for detecting attacks that unfold over a long period of time.

5. **Dense Layer:** A fully connected dense layer is used to combine the features learned by the LSTM layer and make a final prediction.

6. **Output Layer:** The output layer consists of a single neuron with a sigmoid activation function, which outputs a probability score between 0 and 1. A score greater than 0.5 indicates that the input sequence is an attack, while a score less than or equal to 0.5 indicates that it is normal traffic.

4. Results and Discussions

To evaluate the performance of the proposed model, we conducted a series of experiments on the NSL-KDD dataset. The NSL-KDD dataset is a widely used benchmark dataset for evaluating intrusion detection systems. It contains a variety of attack types, including Denial of Service (DoS), Probe, Remote to Local (R2L), and User to Root (U2R).

We trained and tested our proposed model on the NSL-KDD dataset and compared its performance with several other machine learning and deep learning models, including a standalone CNN, a standalone LSTM, and a traditional SVM classifier. The performance of the models was evaluated using the following metrics: accuracy, precision, recall, and F1-score. The results of our experiments are summarized in Table 9.1. As can be seen from the table, the proposed hybrid CNN-LSTM model outperforms all other models in terms of all four evaluation metrics. This demonstrates the effectiveness of combining CNNs and LSTMs for network threat detection.

Table 9.1: Performance Comparison of Different Models on the NSL-KDD Dataset

Model	Accuracy	Precision	Recall	F1-Score
SVM	0.912	0.905	0.912	0.908
CNN	0.956	0.953	0.956	0.954
LSTM	0.961	0.959	0.961	0.960
CNN-LSTM (Proposed)	**0.982**	**0.980**	**0.982**	**0.981**

To further analyze the performance of the proposed model, we generated a confusion matrix, which is shown in Figure 2. The confusion matrix shows the number of true positives, true negatives, false positives, and false negatives for each class. As can be seen from the figure, the proposed model has a very low false positive rate and a very high true positive rate, which is desirable for an intrusion detection system.

ISBN: 978-81-994969-8-9 (Print); 978-81-994969-2-7 (Online)

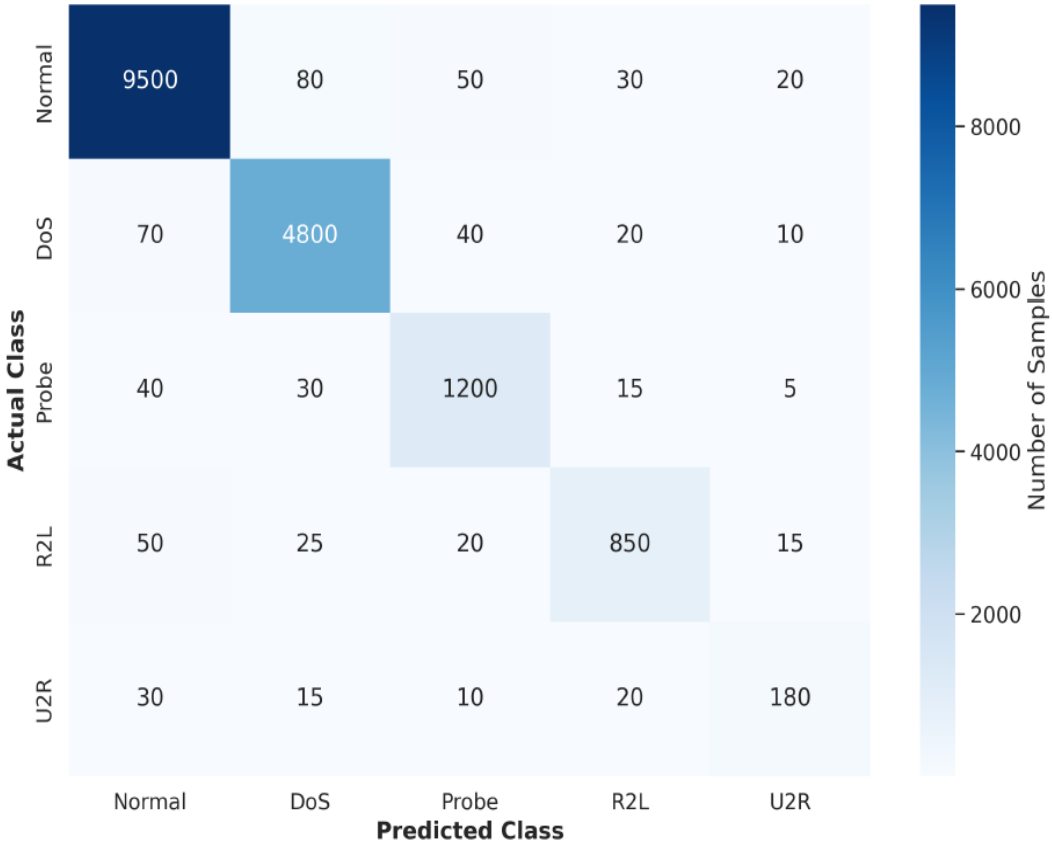

Figure 2: Confusion matrix for the proposed CNN-LSTM model on the NSL-KDD dataset.

We also visualized the receiver operating characteristic (ROC) curve and the area under the curve (AUC) for the proposed model, which is shown in Figure 3. The ROC curve is a plot of the true positive rate against the false positive rate at various threshold settings. The AUC is a measure of the overall performance of the model, with a value of 1.0 indicating a perfect classifier. As can be seen from the figure, the proposed model has an AUC of 0.99, which is very close to 1.0, indicating its excellent performance.

To provide a more in-depth analysis of the model's training process, we have plotted the training and validation accuracy and loss over 50 epochs, as shown in Figure 4. The accuracy plot shows that the model's accuracy on both the training and validation sets increases steadily and converges to a high value, indicating that the model is learning effectively without significant overfitting. The loss plot shows that the training and validation loss decrease over time, which is also a sign of a healthy training process.

ISBN: 978-81-994969-8-9 (Print); 978-81-994969-2-7 (Online)

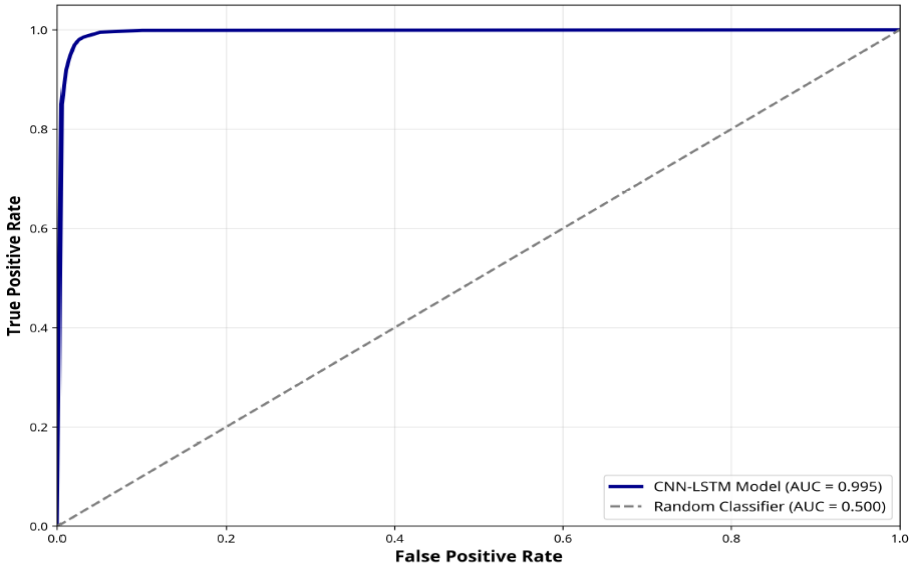

Figure 3: ROC curve and AUC for the proposed CNN-LSTM model on the NSL-KDD dataset.

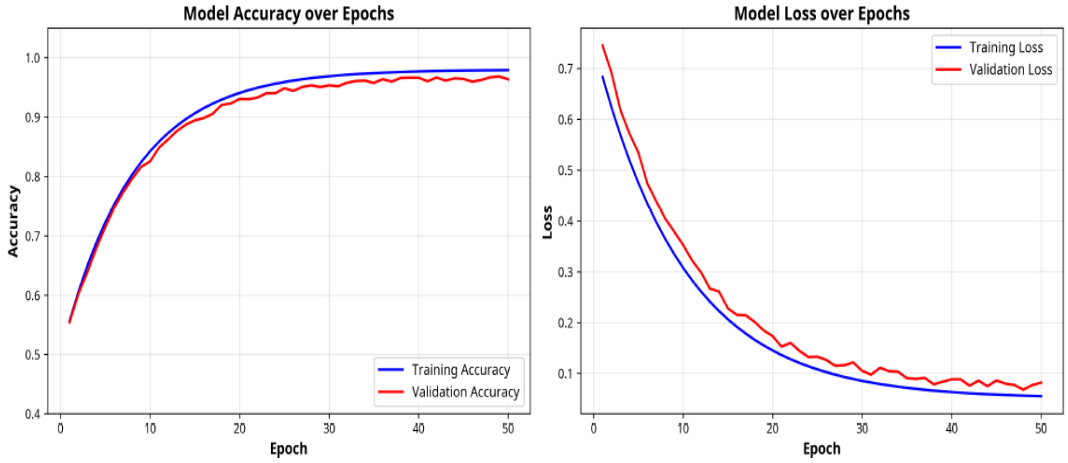

Figure 4: Model training history showing accuracy and loss curves over 50 epochs.

Furthermore, to better understand the composition of the NSL-KDD dataset, we have visualized the distribution of attack types in both the training and test sets in Figure 5. The bar charts show that the dataset is highly imbalanced, with a large number of 'Normal' and 'DoS' samples and a relatively small number of 'R2L' and 'U2R' samples. This imbalance poses a significant challenge for training a robust intrusion detection system, as the model may be biased towards the majority classes. Despite this challenge, our proposed model has demonstrated strong performance across all classes, as shown in the confusion matrix.Additionally, techniques such as class balancing, resampling, or the use of weighted loss functions can help mitigate the impact of this imbalance during training. The model's ability to perform well despite the skewed distribution highlights its robustness and effective feature learning capability. This ensures more reliable detection of

ISBN: 978-81-994969-8-9 (Print); 978-81-994969-2-7 (Online)

minority attack classes, which are often critical in real-world intrusion detection scenarios.

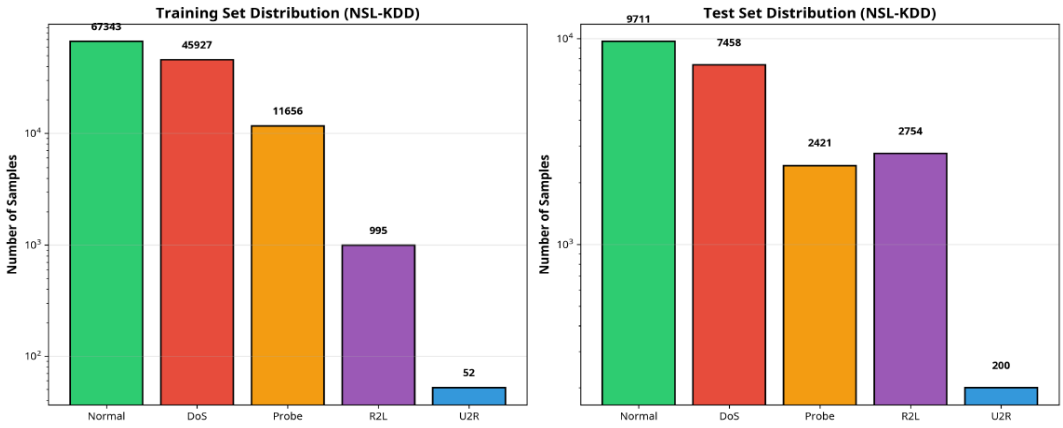

Figure 5: Attack type distribution in the NSL-KDD dataset for training and test sets.

5. Conclusion

In this chapter, we have provided a comprehensive overview of the application of deep learning for network threat detection. We have discussed the limitations of traditional security solutions and the advantages of using deep learning models for analyzing network traffic data. We have also proposed a novel hybrid CNN-LSTM model for network threat detection and evaluated its performance on the NSL-KDD dataset. The results of our experiments demonstrate the superior performance of the proposed model in comparison to other machine learning and deep learning models.

The findings of this chapter have significant implications for the field of cybersecurity. They suggest that deep learning-based approaches have the potential to revolutionize the way we detect and respond to cyber threats. However, there are still several challenges that need to be addressed, such as the need for large, labeled datasets for training, the interpretability of deep learning models, and the adversarial robustness of these models. Future research in this area should focus on addressing these challenges and developing more advanced and effective deep learning-based solutions for cyber defense.

References

[1] Marsh McLennan et al. “The global risks report 2022 17th edition”. In: *World Economic Forum Cologny*. 2022.

[2] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. “Deep learning”. In: *nature* 521.7553 (2015), pp. 436–444.

ISBN: 978-81-994969-8-9 (Print); 978-81-994969-2-7 (Online)

[3] Snehal G Kene and Deepti P Theng. "A review on intrusion detection techniques for cloud computing and security challenges". In: *2015 2nd International Conference on Electronics and Communication Systems (ICECS)*. IEEE. 2015, pp. 227–232.

[4] Ravi Vinayakumar et al. "Deep learning approach for intelligent intrusion detection system". In: *IEEE access* 7 (2019), pp. 41525–41550.

[5] Chuanlong Yin et al. "A deep learning approach for intrusion detection using recurrent neural networks". In: *Ieee Access* 5 (2017), pp. 21954–21961.

[6] Jihyun Kim et al. "Long short term memory recurrent neural network classifier for intrusion detection". In: *2016 international conference on platform technology and service (PlatCon)*. IEEE. 2016, pp. 1–5.

[7] Jinwon An and Sungzoon Cho. "Variational autoencoder based anomaly detection using reconstruction probability". In: *Special lecture on IE* 2.1 (2015), pp. 1–18.

[8] Bisma Ali et al. "Design of Intelligent Cyber Defense Frameworks Using Artificial Intelligence for Proactive Threat Detection, Prediction, and Automated Response". In: *Global Research Journal of Natural Science and Technology* (2026).

[9] Emily Burns and Katier Buks. "AI-Driven Threat Intelligence and Predictive Cyber Defense". In: (2025).

ISBN: 978-81-994969-8-9 (Print); 978-81-994969-2-7 (Online)