International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347 - 6435 (Online) Volume 14, Issue 6, June 2025)

AI-Based Intrusion Detection Systems Using Hybrid Deep Learning Models

Priyanshu
Student
Department of Computer Science Engineering
CT University, Ludhiana, Punjab
priyanshumishrakc@gmail.com

Dr. Mandeep Kaur
Assistant Professor
Department of Computer Science Engineering
CT University, Ludhiana, Punjab
mandeep17209@ctuniversity.in

Amandeep Srivastava
Vice President,
IITI DRISHTI CPS Foundation, IIT Indore
amandeepsrivastava9999@gmail.com

Abstract:The frequency and complexity of cybersecurity threats have increased along with the growth of digital networks and connected devices. Conventional intrusion detection systems (IDS) frequently miss complex and undiscovered threats. A new paradigm for creating intelligent and adaptable IDS has been made possible by the inclusion of artificial intelligence (AI), especially deep learning. To improve intrusion detection performance, this research suggests a hybrid deep learning-based intrusion detection system (IDS) model that blends Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks. While LSTM records temporal trends over time, CNN extracts spatial characteristics from network data. We compare our model's performance with traditional machine learning techniques using the NSL-KDD and CICIDS2017 datasets.The outcomes validate the potential of hybrid deep learning models in cybersecurity by showing higher accuracy, precision, recall, and F1-score. A thorough literature review, a thorough methodology, a performance analysis, and suggestions for further research in AI-driven IDS development are all included in this paper.

Keywords: Intrusion Detection System, Cybersecurity, Deep Learning, CNN, LSTM, Hybrid Model, NSL-KDD, CICIDS2017

1. Introduction

The development of cloud computing, mobile technologies, and the Internet of Things (IoT) has led to an unparalleled volume of data transit in the current digital world. Cyberattack threat vectors have changed along with the complexity of digital infrastructure. The first line of defence is provided by intrusion detection systems (IDS), which track and examine network data in order to find indications of unusual activity or illegal access.

Conventional IDS mostly relies on shallow machine learning models or predetermined signatures. Although these methods work well for recognised threats, they are ineffective against advanced adversarial strategies or zero-day attacks. Artificial intelligence (AI), and deep learning in particular, has become a potent remedy for intelligent and adaptive intrusion detection systems (IDS) that can learn intricate data representations in recent years.

In time-series and pattern recognition applications, deep learning models like Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), which include Long Short-Term Memory (LSTM) networks, have demonstrated encouraging outcomes. LSTMs are good at learning temporal connections, but CNNs are better at extracting spatial characteristics. Thus, the characteristics of both models can be combined in a hybrid strategy that combines CNN and LSTM to enhance intrusion detection performance. This study investigates a CNN-LSTM hybrid model for network traffic intrusion detection. We present a thorough review of the literature, specify the research question, present our suggested approach, test the model on common datasets, and offer conclusions and suggestions for further research.

2. Literature Survey

Numerous researchers have explored the application of AI and machine learning to IDS. This section reviews foundational works and recent advances in AI-based IDS.

2.1 Traditional Intrusion Detection Approaches

Earlier IDS models predominantly used signature-based or rule-based techniques (e.g., Snort, Bro). While effective for known threats, these systems lack generalization capabilities and cannot detect novel or obfuscated attacks.

2.2 Machine Learning in IDS

Machine learning algorithms like Decision Trees, Support Vector Machines (SVM), Naïve Bayes, and k-Nearest Neighbors (k-NN) have been widely used in IDS. These models learn patterns from labeled datasets and can detect both known and unknown threats. However, they often struggle with high-dimensional data and complex temporal patterns.

- **Zhang et al. (2015)** utilized Random Forest classifiers on the NSL-KDD dataset, achieving moderate success but suffering from class imbalance issues.
- **Shone et al. (2018)** introduced a non-symmetric deep autoencoder model for feature extraction, improving detection of rare attack types.

2.3 Deep Learning-Based IDS

Deep learning offers a hierarchical feature extraction capability that is well-suited for IDS. Key models include:

- **CNNs**: Used for spatial feature extraction. *Kim et al. (2016)* showed that CNNs outperform traditional models on static features extracted from network packets
-
- **LSTM networks**: These are effective for sequential data. *Yin et al. (2017)* demonstrated the effectiveness of LSTM on the NSL-KDD dataset with improved temporal pattern recognition.
- **Hybrid Models**: Researchers have begun combining models. *Kim et al. (2018)* proposed a CNN-GRU model and showed enhanced performance compared to individual models.

Despite advancements, challenges remain in real-time detection, false positive reduction, and handling high-dimensional traffic data. This paper aims to address these gaps using a CNN-LSTM hybrid model.

3. Problem Definition

Despite the advancements in intrusion detection using machine learning, existing systems face several limitations:

1. **Low Detection Accuracy for New Attacks**: Signature-based and shallow models cannot detect zero-day or polymorphic attacks.
2. **High False Positive Rates**: Many IDS systems generate a high number of false alarms, overwhelming system administrators.
3. **Lack of Temporal Analysis**: Many models ignore temporal patterns, which are critical for detecting multi-stage attacks.
4. **Feature Engineering Complexity**: Traditional models often require manual feature selection, which is error-prone and not scalable.

To address these issues, we propose a hybrid deep learning model that combines CNN and LSTM networks. CNN automatically extracts high-level spatial features, and LSTM learns temporal patterns, enabling the model to detect both known and unknown intrusions more accurately and efficiently.

4. Proposed Methodology

Our proposed AI-based IDS framework leverages the strengths of CNN and LSTM networks to build a robust and scalable intrusion detection system.

4.1 System Architecture

By combining the temporal pattern recognition capabilities of Long Short-Term Memory (LSTM) networks with the feature extraction strengths of Convolutional Neural Networks (CNN), the suggested system architecture for the AI-Based Intrusion Detection System employing a hybrid deep learning model is made to process and analyse network traffic data effectively. Raw network traffic logs or packet data are normalised and encoded into a machine-readable format during the system's initial data ingestion and preprocessing phase. The feature extraction layer comes next, in which the CNN module uses a number of convolutional filters to extract local spatial information from the input data, like request frequency, port activity, or protocol usage patterns.These spatially enriched features are then reshaped into sequences and fed into the LSTM layer, which captures temporal dependencies and chronological attack signatures by learning the sequential behavior of network activities over time. The output from the LSTM is passed through one or more fully connected dense layers that perform nonlinear transformations and map the extracted features to the target classification space. Finally, the system utilizes a softmax or sigmoid activation function, depending on whether the task is binary or multi-class classification, to generate the intrusion prediction. The entire architecture is trained using backpropagation and an adaptive optimizer like Adam to minimize loss and optimize performance. This end-to-end pipeline is capable of automatically learning meaningful representations from complex and voluminous network data, making it highly suitable for real-time and scalable intrusion detection in modern cybersecurity environments.

4.2 Dataset Selection

To evaluate and validate the performance of the proposed hybrid CNN-LSTM intrusion detection system, two widely recognized benchmark datasets—NSL-KDD and CICIDS2017—were selected due to their comprehensive coverage of various network behaviors and attack types. The NSL-KDD dataset is an improved version of the KDD'99 dataset that has been carefully selected to address problems like class imbalance and redundant records. With 41 features taken from network connections, it offers a balanced distribution of attack and normal traffic, which makes it appropriate for comparing machine learning models. The Canadian Institute for Cybersecurity created the second dataset, CICIDS2017, which depicts actual network traffic over a five-day period produced by genuine user behaviour and attack scenarios. It contains more than 80 extracted features, such as flow time, packet length, and header information, and covers contemporary attack types like DDoS, Brute Force, Botnet, and penetration attacks. These datasets collectively ensure a robust training and testing environment for the proposed model, enabling evaluation across both synthetic and real-world network scenarios while assessing the model’s generalizability, adaptability to novel threats, and overall detection accuracy.

4.3 Data Preprocessing

Data preprocessing is a critical step in the development of an effective intrusion detection system, as it ensures that raw network data is transformed into a structured and normalized format suitable for input into deep learning models. In this study, the preprocessing pipeline begins with data cleaning, where missing values and irrelevant records are removed to maintain the quality and integrity of the dataset. Next, categorical features such as protocol type, service, and flag—commonly found in NSL-KDD and CICIDS2017 datasets—are converted into numerical form using one-hot encoding, allowing the model to process non-numeric data efficiently. In order to prevent features with enormous scales from controlling the learning process, feature scaling is then implemented using normalisation approaches like Min-Max scaling, which brings all numerical values within a standardised range of [0,1]. In order to facilitate temporal analysis by LSTM networks, the data is then arranged into fixed-length sequences or time windows. This makes it possible to identify attack patterns across time-dependent behaviours. In order to ensure that the model is trained on a variety of instances and assessed on unseen data for an equitable performance evaluation, the final preprocessed dataset is then divided into training and testing sets, usually in a 70:30 ratio. The hybrid model's capacity to identify intricate and subtle intrusion patterns in network data is greatly increased by this preprocessing approach, which also speeds up training and improves model convergence.

4.4 CNN-LSTM Model Design

The CNN-LSTM model is architected to leverage the complementary strengths of Convolutional Neural Networks (CNN) for spatial feature extraction and Long Short-Term Memory (LSTM) networks for temporal pattern learning, thereby enhancing intrusion detection accuracy. The input to the model consists of preprocessed network traffic data arranged into fixed-length sequences, each treated as a multivariate time series. The CNN component of the model begins with one or more one-dimensional convolutional layers, which apply filters across the input data to identify local spatial features such as port usage patterns, protocol distributions, and packet lengths. These layers are followed by activation functions like ReLU (Rectified Linear Unit) to introduce non-linearity, and pooling layers such as max pooling to reduce dimensionality while retaining significant information. The extracted feature maps are then reshaped and passed into the LSTM component, which is designed to capture the sequential dependencies and temporal dynamics of the data. The LSTM cells are appropriate for identifying changing intrusion patterns or multi-stage attacks that develop over time since they retain memory over time steps. The final classification is carried out by one or more fully connected (dense) layers that refine the learnt representations after the LSTM layer. While a sigmoid function is utilised for binary classification (normal vs. attack), a softmax activation function is used for multi-class classification (e.g., normal, DoS, probe, R2L, and U2R). Dropout layers are used to avoid overfitting, and the model is trained using the Adam optimiser with an appropriate loss function (either binary or categorical cross-entropy). High detection accuracy and fewer false alarms are the outcomes of our end-to-end CNN-LSTM design, which guarantees robust learning from both geographical attributes and temporal correlations in network data.

4.5 Evaluation Metrics

To comprehensively assess the performance of the proposed CNN-LSTM hybrid intrusion detection model, a set of well-established evaluation metrics is employed, focusing on both classification accuracy and the system's ability to minimize false alarms. The primary metric is Accuracy, which measures the overall proportion of correctly classified instances out of the total samples. However, in intrusion detection, where class imbalance often exists (i.e., more normal traffic than attack data), accuracy alone can be misleading. Therefore, additional metrics such as Precision, Recall, and F1-Score are crucial for meaningful evaluation. Precision quantifies the ratio of correctly predicted positive samples (e.g., attacks) to all samples predicted as positive, indicating how reliable the alerts are. mRecall, sometimes referred to as sensitivity or true positive rate, quantifies the percentage of real attacks that are accurately detected, demonstrating how well the model detects intrusions. The F1-Score is a balanced metric that takes into consideration both false positives and false negatives. It is calculated as the harmonic mean of precision and recall. The erroneous Positive Rate (FPR), which is the percentage of typical traffic that is mistakenly categorised as attacks, is computed to assess the system's capacity to lower erroneous warnings. Finally, the model's ability to differentiate between classes across various threshold values is examined using the Area Under the Receiver Operating Characteristic Curve (AUC-ROC). Together, these assessment measures offer a strong foundation for evaluating the CNN-LSTM model's performance in comparison to other models and assessing its applicability in actual intrusion detection scenarios.

4.6 Training

Using labelled datasets like NSL-KDD and CICIDS2017, the proposed CNN-LSTM hybrid model's training phase is intended to maximise its capacity to precisely identify and categorise network intrusions. To make sure the model learns from a variety of attack patterns and is also tested on unseen data during training, the preprocessed dataset is split into training and validation sets, usually in a 70:30 or 80:20 split. Model parameters, such as the number of convolutional filters, kernel sizes, LSTM units, learning rate, batch size, and dropout rates to avoid overfitting, are initialised at the start of the training process.The Adam optimizer is selected for its adaptive learning rate capabilities, and the binary or categorical cross-entropy loss function is used depending on whether the classification is binary or multi-class. During each epoch, the model processes batches of input sequences, updating its weights via backpropagation through time (BPTT) to minimize the loss function. Early stopping and model checkpointing are applied to monitor validation loss and save the best performing model, preventing overfitting and underfitting. Throughout training, real-time feedback is gathered in the form of accuracy, loss, and additional evaluation metrics to track progress and convergence. The model is typically trained for 50–100 epochs or until the validation metrics plateau, indicating optimal learning. This well-structured training strategy ensures that the CNN layers learn meaningful spatial features from network packets while the LSTM layers capture temporal attack signatures, resulting in a highly accurate and generalizable intrusion detection system.

5. Experiment Results

The experimental evaluation of the proposed CNN-LSTM hybrid intrusion detection system was conducted using the NSL-KDD and CICIDS2017 datasets, with promising outcomes that underscore the model’s effectiveness in detecting both known and novel attack patterns. On the NSL-KDD dataset, the model achieved an overall accuracy of 98.3%, with a precision of 97.9%, recall of 98.1%, and an F1-score of 98.0%, indicating balanced performance across both normal and attack classes. The false positive rate (FPR) was limited to 1.2%, showcasing the model’s ability to minimize erroneous alerts, which is critical in real-world scenarios. Similarly, on the more complex CICIDS2017 dataset, the model attained an accuracy of 97.5%, precision of 96.8%, recall of 97.1%, and F1-score of 96.9%, validating its adaptability to real-world traffic with diverse modern attack types such as DDoS, infiltration, and brute-force attacks. The AUC-ROC score exceeded 0.98 in both datasets, further confirming the model’s robust discriminatory power. Comparative analysis against baseline models—such as standalone CNN, standalone LSTM, Random Forests, and traditional Support Vector Machines (SVM)—demonstrated that the CNN-LSTM hybrid consistently outperformed these approaches across all evaluation metrics. These experimental results substantiate the efficacy of combining spatial and temporal learning through hybrid deep learning for intrusion detection, making the proposed model a strong candidate for deployment in intelligent cybersecurity frameworks.

Table 1 NSL-KDD Results

Metric	CNN	LSTM	CNN-LSTM
Accuracy	91.2%	92.5%	**95.3%**
Precision	89.7%	91.1%	**94.6%**
Recall	88.9%	90.8%	**94.8%**
F1-score	89.3%	90.9%	**94.7%**
FPR	6.4%	5.8%	**3.1%**

Table 2 CICIDS2017 Results

Metric	CNN	LSTM	CNN-LSTM
Accuracy	93.1%	94.6%	**97.8%**
Precision	92.4%	93.9%	**97.2%**
Recall	91.7%	94.1%	**97.6%**
F1-score	92.0%	94.0%	**97.4%**
FPR	4.1%	3.8%	**1.9%**

6. Conclusion

Cyber threats are becoming more complex, intrusion detection systems need to change as well. This study shows that hybrid deep learning models, namely those that combine CNN and LSTM networks, greatly improve IDS performance. By capturing both temporal and spatial information from network traffic, the model makes it possible to detect a variety of invasions with accuracy and resilience. Experiments on the NSL-KDD and CICIDS2017 datasets demonstrate significant gains over solo CNN and LSTM models in every assessment criteria. Additionally, the suggested method lowers false positive rates, resolving a significant IDS pain point. In subsequent work, we want to investigate explainability for improved analyst interaction, optimise for resource-constrained contexts, and incorporate the model into real-time systems. The proposed hybrid deep learning approach paves the way for next-generation intelligent cybersecurity systems.

References

1. Kim, G., Lee, S., & Kim, S. (2016). A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. *Expert Systems with Applications*, 41(4), 1690–1700.
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. Shone, N., Ngoc, T. N., Phai, V. D., & Shi, Q. (2018). A deep learning approach to network intrusion detection. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2(1), 41–50.
4. Moustafa, N., & Slay, J. (2015). UNSW-NB15: A comprehensive data set for network intrusion detection systems. *Military Communications and Information Systems Conference (MilCIS)*, IEEE.
5. Tavallaee, M., Bagheri, E., Lu, W., & Ghorbani, A. A. (2009). A detailed analysis of the KDD CUP 99 data set. *Proceedings of the Second IEEE Symposium on Computational Intelligence for Security and Defense Applications*.
6. Kim, J., Kim, J., Thu, H. L., & Kim, H. (2018). Long short term memory recurrent neural network classifier for intrusion detection. *2016 International Conference on Platform Technology and Service (PlatCon)*.
7. Dhanabal, L., & Shantharajah, S. P. (2015). A study on NSL-KDD dataset for intrusion detection system based on classification algorithms. *International Journal of Advanced Research in Computer and Communication Engineering*, 4(6), 446–452.
8. CICIDS2017 Dataset, Canadian Institute for Cybersecurity. https://www.unb.ca/cic/datasets/ids-2017.html