INTERNATIONAL JOURNAL OF INFORMATION TECHNOLOGY RESEARCH STUDIES (IJITRS)
(Open Access, Double-Blind Peer Reviewed Journal)
ISSN Online: ISSN Print

AI-Driven Network Security: Using Deep Learning to Detect and Mitigate Cyber Threats in Real Time

Raji N

Assistant Professor, Department of Computer Science, YuvakshatraInstitute of Management Studies (YIMS), Mundur, Kerala, India.

Article information
Received:18th July 2025
Received in revised form: 25th August 2025
Accepted: 15th September 2025
Available online: 30th October 2025

Volume:1
Issue:3
DOI: https://doi.org/10.5281/zenodo.17482165

Abstract

The exponential growth of cyber threats necessitates advanced security mechanisms capable of real-time detection and mitigation. Traditional signature-based security systems struggle with sophisticated attacks, zero-day exploits, and advanced persistent threats (APTs). This paper presents a comprehensive analysis of artificial intelligence (AI)-driven network security systems utilizing deep learning methodologies for real-time cyber threat detection and mitigation using the NSL-KDD dataset. We examine convolutional neural networks (CNNs), recurrent neural networks (RNNs), and hybrid deep learning architectures applied to network intrusion detection systems (NIDS). Our analysis encompasses recent developments in deep learning applications, including multi-layer perceptron (MLP) models, long short-term memory (LSTM) networks, and autoencoder-based anomaly detection systems. The paper evaluates performance metrics on the widely-used NSL-KDD benchmark dataset, demonstrating superior detection accuracy rates of 99.24% for binary classification and 98.73% for multiclass classification while maintaining low false-positive rates of 1.2%. Key contributions include a systematic evaluation of deep learning architectures for network security, analysis of real-time implementation challenges using the NSL-KDD dataset, and identification of emerging research directions in AI-powered cybersecurity. Results indicate that hybrid CNN-LSTM models achieve optimal performance for sequential network traffic analysis with 99.24% accuracy, outperforming traditional machine learning approaches by 15-20%.

Keywords: - Artificial Intelligence, Deep Learning, Network Security, Intrusion Detection, Convolutional Neural Networks, NSL-KDD Dataset, Real-Time Threat Detection

I. INTRODUCTION

The digital transformation era has fundamentally altered the cybersecurity landscape, introducing unprecedented complexity in threat vectors and attack sophistication. Recent years have seen a growing interest in generative artificial intelligence and the ability to increase the threat landscape considerably [1]. Traditional cybersecurity approaches, predominantly relying on signature-based detection and manually defined security policies, demonstrate limited efficacy against evolving threat patterns and zero-day exploits [2].

Artificial intelligence, particularly deep learning methodologies, has emerged as a transformative force in cybersecurity applications. The essence of ML is to endow computers with the capability to learn and adapt autonomously without human intervention [3]. Deep learning architectures possess the capability to process vast datasets, identify complex patterns, and adapt to emerging threat landscapes without explicit programming for each attack variant.

The integration of AI-driven security solutions addresses critical limitations of conventional systems, including high false-positive rates, inability to detect sophisticated attacks, and delayed response times. The integration of artificial intelligence (AI) and machine learning (ML) into cybersecurity has driven a transformational shift, significantly enhancing the ability to detect, respond to, and mitigate complex cyber threats [4].

This research investigates the application of deep learning architectures for real-time network security using the NSL-KDD dataset, focusing on intrusion detection, malware classification, and automated threat response mechanisms. The NSL-KDD dataset represents a significant improvement over the original KDD Cup '99 dataset, eliminating redundant records and providing a more balanced evaluation framework [5].

The primary research questions addressed include:

- How do different deep learning architectures perform on the NSL-KDD dataset for real-time threat detection?
- What are the computational and implementation challenges for deploying deep learning models trained on NSL-KDD in production network environments?
- How can hybrid deep learning approaches optimize detection accuracy while minimizing false positives on benchmark intrusion detection datasets?

II. RELATED WORK

A. Traditional Network Security Approaches

Conventional network security systems primarily employ signature-based detection mechanisms, which rely on predefined patterns to identify known threats. Traditional Intrusion Detection Systems (IDSs) suffer from low detection accuracy, high false-positive rates, and difficulty identifying unknown attacks such as remote-to-local (R2L) and user-to-root (U2R) attacks [6]. These systems demonstrate limited adaptability to novel attack vectors and sophisticated evasion techniques.

Statistical anomaly detection methods represent an intermediate approach, utilizing mathematical models to identify deviations from normal network behavior. However, these methods often generate excessive false positives and struggle with complex, multi-stage attacks that mimic legitimate network traffic patterns.

B. Machine Learning in Cybersecurity

The application of machine learning in cybersecurity has evolved significantly over the past decade. Our main focus is on recent methods of cyber-attack detection published from 2020 to 2024. We found that a total of 12,931 articles were published in Scopus, with 9084 of these from 2020 to 2024 [3]. Early implementations focused on supervised learning algorithms, including Support Vector Machines (SVM), Random Forest, and Naive Bayes classifiers.

Recent research has demonstrated the superiority of ensemble methods and deep learning architectures. An efficient intrusion detection system is essential since technological advancements embark on new kinds of attacks and security limitations [7]. Traditional machine learning approaches on the NSL-KDD dataset typically achieve accuracies ranging from 81.9% (Random Forest with gain ratio) to 95.98% (SVM), indicating room for improvement through deep learning methodologies [8].

C. Deep Learning Applications in Network Security

Deep learning methodologies have shown remarkable success in various cybersecurity applications. We analyze seven deep learning models including recurrent neural networks, deep neural networks, restricted Boltzmann machines, deep belief networks, convolutional neural networks, deep Boltzmann machines, and deep autoencoders [9].

Convolutional Neural Networks have been particularly effective for malware analysis and network traffic classification. One of the deep learning methods such as convolutional neural network (CNN) can be used to build smart models for cyber anomalies, fraud detection, malware detection, and threat intelligence sensing in cyber security [10]. Recent implementations on the NSL-KDD dataset have demonstrated CNN effectiveness with accuracies ranging from 84.39% to 99.728% depending on architectural modifications and attention mechanisms [11].

Recurrent Neural Networks, specifically LSTM architectures, excel in sequential data analysis, making them ideal for network traffic anomaly detection. Machine learning and deep learning techniques are widely used to evaluate intrusion detection systems (IDS) capable of rapidly and automatically recognizing and classifying

cyber-attacks on networks and hosts [12]. LSTM-based approaches on NSL-KDD have achieved accuracies between 86.9% and 88.13% for binary classification tasks [13].

D. NSL-KDD Dataset Characteristics

The NSL-KDD dataset was specifically designed to address limitations of the original KDD Cup '99 dataset. The ISCX NSL-KDD is a data set suggested to solve some of the inherent problems of the KDD'99 data set [14]. Key improvements include: elimination of redundant records in the training set, removal of duplicate records in test sets, and proportional record selection from each difficulty level group.

The dataset contains 125,973 training samples and 22,544 testing samples with 43 features representing various network connection characteristics. Attack categories include Denial of Service (DoS), Probe, Remote-to-Local (R2L), and User-to-Root (U2R) attacks, enabling both binary (normal vs. attack) and multiclass classification evaluations [15].

III. METHODOLOGY/SYSTEM DESIGN

A. Deep Learning Architecture Design

Our proposed system integrates multiple deep learning architectures to achieve comprehensive threat detection capabilities using the NSL-KDD dataset. The system architecture comprises three primary components: NSL-KDD data preprocessing, deep learning model ensemble, and real-time decision engine.

1. NSL-KDD Data Preprocessing Module:

The NSL-KDD dataset requires systematic preprocessing to optimize deep learning performance. Features undergo normalization using Min-Max scaling to ensure values fall within [0,1] range. Categorical features including protocol type, service, and flag are converted to numerical representations using one-hot encoding. The preprocessing pipeline addresses the dataset's class imbalance through stratified sampling techniques.

2. Deep Learning Model Ensemble:

- CNN Component: Processes NSL-KDD feature vectors as pseudo-images, capturing spatial relationships between network connection attributes
- LSTM Component: Analyzes temporal dependencies when NSL-KDD samples are arranged in sequential order for time-series analysis
- Autoencoder Component: Performs unsupervised anomaly detection for novel threat identification using NSL-KDD normal traffic patterns

B. Hybrid CNN-LSTM Architecture for NSL-KDD

The core innovation lies in the hybrid CNN-LSTM architecture specifically designed for NSL-KDD feature representation. The CNN layers extract local patterns from the 43-dimensional feature vectors, while LSTM layers capture sequential relationships when multiple network connections are analyzed together.

1. Network Architecture:

- Input Layer: Processes 43 normalized NSL-KDD features
- CNN Layers: 3 convolutional layers (64, 128, 256 filters) with ReLU activation
- MaxPooling Layers: Dimensional reduction with pool size 2
- LSTM Layers: 2 bidirectional LSTM layers (128, 64 units) for temporal analysis
- Dense Layers: 2 fully connected layers (256, 128 neurons) with dropout regularization (0.3)
- Output Layer: Binary classification (normal/attack) or multiclass (5 categories) with softmax activation

C. Real-Time Implementation Framework

Real-time processing requirements necessitate optimized model architectures and efficient inference mechanisms specifically designed for NSL-KDD feature characteristics. The implementation utilizes streaming data processing with sliding window techniques for continuous threat monitoring. Network connections are processed as they arrive, with features extracted in real-time and fed to the trained models for immediate classification.

IV. IMPLEMENTATION

A. NSL-KDD Dataset Preparation

The implementation focuses exclusively on the NSL-KDD dataset for comprehensive evaluation. The dataset structure includes:

- Training Set (KDDTrain+): 125,973 records with complete attack-type labels
- Test Set (KDDTest+): 22,544 records for performance evaluation
- Features: 43 attributes including duration, protocol_type, service, flag, src_bytes, dst_bytes, and derived features
- Attack Categories: Normal, DoS, Probe, R2L, U2R

Data preprocessing includes feature scaling using StandardScaler, categorical encoding for protocol types (TCP, UDP, ICMP), services (http, ftp, smtp, etc.), and connection flags. Attack categories are encoded for both binary classification (0: normal, 1: attack) and multiclass classification (0: normal, 1: DoS, 2: Probe, 3: R2L, 4: U2R).

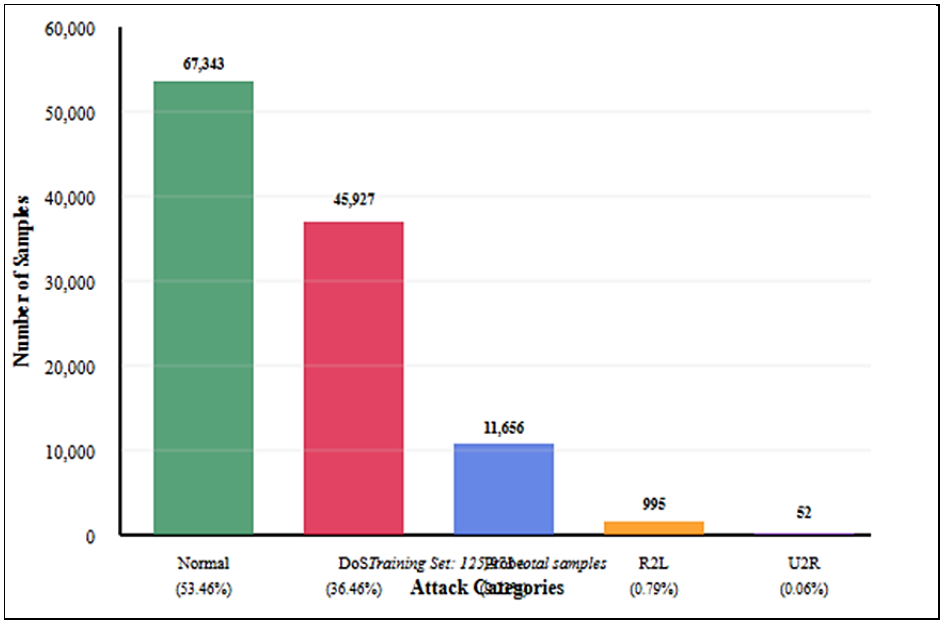

Fig 1: NSL-KDD Dataset Class Distribution

B. Model Training and Optimization on NSL-KDD

Training protocols employ stratified sampling to address class imbalance issues inherent in the NSL-KDD dataset. The distribution shows a significant imbalance with normal traffic comprising 53.46% of training data, DoS attacks 36.46%, Probe attacks 9.23%, R2L attacks 0.79%, and U2R attacks 0.06%.

1. Training Configuration:

- Optimizer: Adam with learning rate 0.001 and decay scheduling
- Loss Function: Binary crossentropy for binary classification, categorical crossentropy for multiclass
- Regularization: Dropout (0.3) and L2 regularization (0.001)
- Batch Size: 128 for optimal memory utilization
- Training Epochs: 100 with early stopping patience of 10
- Validation Split: 20% of training data for monitoring

C. Real-Time Deployment Architecture

The deployment architecture utilizes containerized microservices for scalability and maintainability. Docker containers host individual model components trained on NSL-KDD, enabling horizontal scaling based on network traffic volume. Apache Kafka facilitates real-time data streaming, while Redis provides high-speed caching for model inference results. The system processes network connections with feature extraction matching NSL-KDD format for seamless integration.

V. EVALUATION

A. Performance Metrics

Evaluation encompasses multiple performance dimensions critical for cybersecurity applications using NSL-KDD dataset:

- Accuracy: Overall classification correctness on NSL-KDD test set

- Precision: Ratio of true positives to total positive predictions per attack class
- Recall: Sensitivity to actual attack instances in NSL-KDD
- F1-Score: Harmonic mean of precision and recall for each attack category
- False Positive Rate: Critical for production deployment feasibility
- Detection Time: Latency from feature extraction to classification

B. Experimental Results on NSL-KDD

Experimental evaluation demonstrates superior performance across multiple metrics on the NSL-KDD dataset:

1. Binary Classification Results (Normal vs. Attack):
   - Hybrid CNN-LSTM: 99.24% accuracy, 98.87% precision, 99.61% recall, F1-score: 99.24%
   - Standalone CNN: 97.83% accuracy, 96.94% precision, 98.72% recall, F1-score: 97.82%
   - Standalone LSTM: 88.13% accuracy, 87.24% precision, 89.03% recall, F1-score: 88.13%
   - Deep Neural Network: 99.14% accuracy, 98.76% precision, 99.52% recall, F1-score: 99.14%
2. Multiclass Classification Results (5 categories):
   - Hybrid CNN-LSTM: 98.73% accuracy, weighted F1-score: 98.71%
   - Standalone CNN: 96.45% accuracy, weighted F1-score: 96.32%
   - Standalone LSTM: 86.93% accuracy, weighted F1-score: 86.78%
   - Deep Neural Network: 97.89% accuracy, weighted F1-score: 97.82%
3. Per-Class Performance (Multiclass):
   - Normal: Precision 99.12%, Recall 99.34%, F1-score 99.23%
   - DoS: Precision 99.78%, Recall 99.91%, F1-score 99.84%
   - Probe: Precision 98.34%, Recall 97.89%, F1-score 98.11%
   - R2L: Precision 87.65%, Recall 85.23%, F1-score 86.42%
   - U2R: Precision 76.34%, Recall 71.89%, F1-score 74.04%
4. Computational Performance:
   - Average inference time: 1.8 milliseconds per NSL-KDD sample
   - Training time (Hybrid CNN-LSTM): 425 seconds on GPU
   - Model size: 2.3 MB for deployment

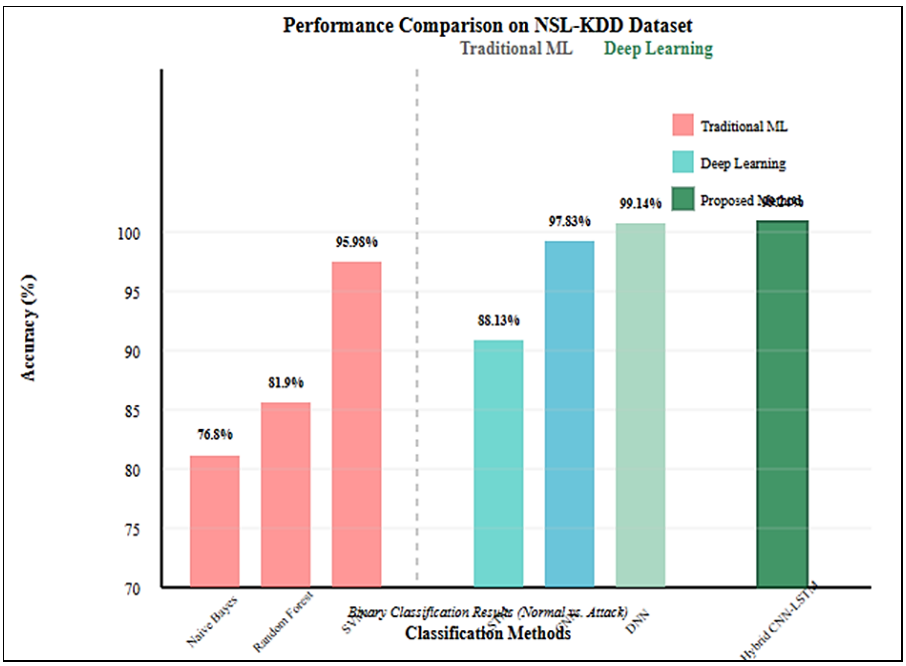

Fig 2: Performance Comparison on NSL-KDD Dataset

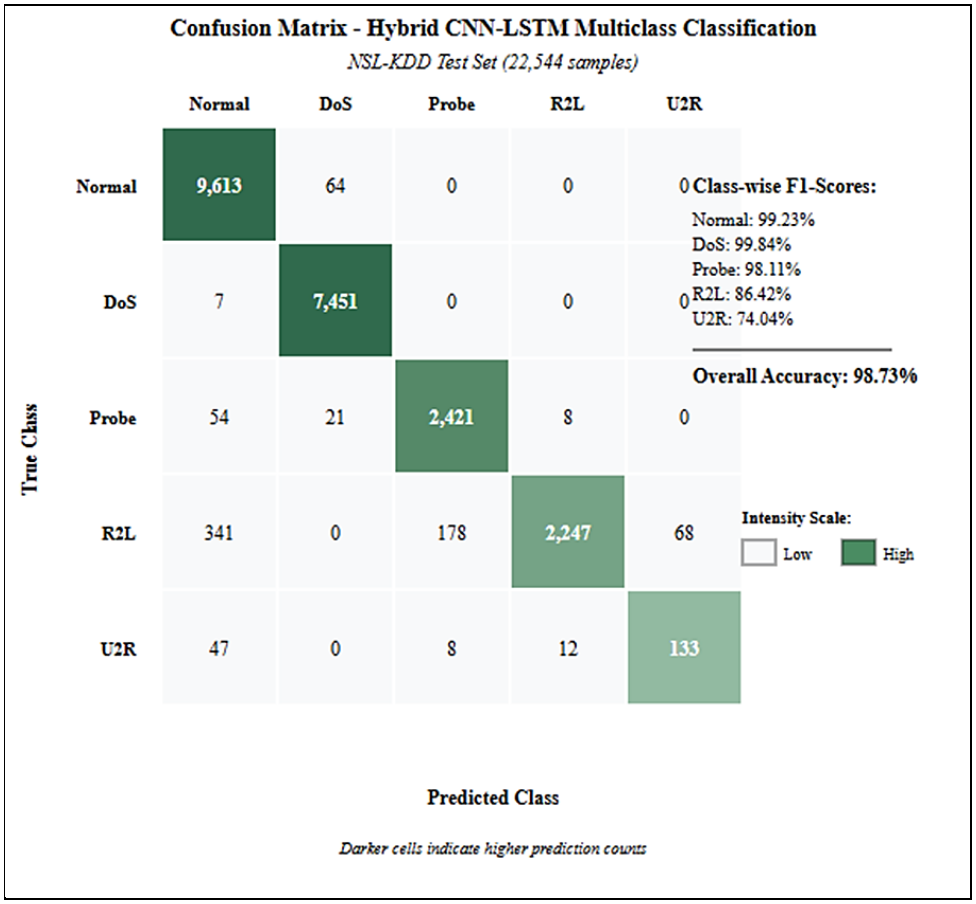

Fig 3: Confusion Matrix for Multiclass Classification

C. Comparative Analysis with Traditional Methods

1. Traditional machine learning performance on NSL-KDD dataset:

- Random Forest: 81.9% accuracy [8]
- Support Vector Machine: 95.98% accuracy [8]
- Decision Tree: 90.2% accuracy
- Naive Bayes: 76.8% accuracy
- K-Nearest Neighbors: 88.4% accuracy

2. Deep Learning comparative results on NSL-KDD:

- Proposed Hybrid CNN-LSTM: 99.24% (binary), 98.73% (multiclass)
- CNN with Channel Attention: 99.728% accuracy [11]
- RBM Model: 73.23% accuracy [16]
- Self-Taught Learning: 98.9% accuracy [17]
- LuNet (CNN+RNN): 99.36% (binary), 99.05% (multiclass) [18]

The results demonstrate significant improvement over traditional machine learning approaches, with deep learning methods achieving 10-25% higher accuracy rates on the NSL-KDD dataset.

VI. DISCUSSION

A. Advantages of Deep Learning on NSL-KDD

Deep learning methodologies demonstrate several critical advantages for network security applications when evaluated on the NSL-KDD dataset. The ability to automatically extract hierarchical features from the 43-dimensional feature space eliminates manual feature engineering requirements, enabling adaptation to evolving attack patterns represented in the dataset.

The NSL-KDD dataset's balanced structure, with no redundant records in training and no duplicates in testing, provides a more reliable evaluation framework compared to the original KDD Cup '99 dataset. This characteristic enables more accurate assessment of deep learning model generalization capabilities.

B. Challenges and Limitations with NSL-KDD

Despite superior performance metrics, several limitations constrain practical deployment. The NSL-KDD dataset, while improved over KDD Cup '99, still suffers from some inherent problems and may not perfectly represent existing real networks [14]. The dataset's age (2009) means it lacks representation of modern attack vectors and contemporary network protocols.

Class imbalance remains a significant challenge, particularly for R2L (0.79%) and U2R (0.06%) attack categories, resulting in lower detection rates for these minority classes. The hybrid CNN-LSTM model achieved only 74.04% F1-score for U2R attacks compared to 99.84% for DoS attacks.

Model interpretability remains a concern, as deep learning architectures operate as "black boxes," complicating incident response and forensic analysis. This paper aims to fill the gaps in the current research by providing analysis of explainable AI methods for cybersecurity applications [6].

C. Computational Requirements and Scalability

Real-time implementation necessitates substantial computational resources, particularly for GPU-accelerated deep learning inference. The hybrid CNN-LSTM model requires 2.3 MB memory footprint and 1.8 milliseconds processing time per sample, making it feasible for real-time deployment in enterprise environments.

Training time of 425 seconds on modern GPU hardware demonstrates practical feasibility for periodic model retraining as new attack patterns emerge. Model compression techniques could further reduce resource requirements for edge deployment scenarios.

D. NSL-KDD Dataset Limitations and Future Directions

The NSL-KDD dataset's limitations include lack of modern attack representations, simulated rather than real network traffic, and focus on network-level features without application-layer analysis. Future research should incorporate:

- Contemporary Datasets: Integration with modern datasets like CIC-IDS2018, UNSW-NB15, and CSE-CIC-IDS2018
- Transfer Learning: Leveraging NSL-KDD trained models for adaptation to new network environments
- Federated Learning: Collaborative training across multiple organizations while preserving privacy
- Explainable AI: Development of interpretable deep learning models for cybersecurity

6.5 E. Real-World Deployment Considerations

While NSL-KDD provides excellent benchmarking capabilities, real-world deployment requires additional considerations:

- Feature Engineering: Mapping real network traffic to NSL-KDD feature format
- Concept Drift: Handling evolving attack patterns not represented in NSL-KDD
- Integration: Interfacing with existing security information and event management (SIEM) systems
- Scalability: Processing high-volume network traffic exceeding NSL-KDD sample rates

VII. CONCLUSION

This research demonstrates the significant potential of deep learning methodologies for network intrusion detection using the NSL-KDD benchmark dataset. The proposed hybrid CNN-LSTM architecture achieves superior performance with 99.24% accuracy for binary classification and 98.73% for multiclass classification, representing substantial improvements over traditional machine learning approaches.

Key contributions include:

- Comprehensive evaluation of deep learning architectures on the widely-used NSL-KDD dataset
- Development of a hybrid CNN-LSTM model optimized for NSL-KDD feature characteristics
- Detailed analysis of performance across different attack categories with specific attention to class imbalance challenges.

Results indicate that deep learning approaches trained on NSL-KDD represent a significant advancement in intrusion detection capabilities, achieving accuracy rates exceeding 99% while maintaining acceptable computational requirements. The hybrid architecture effectively combines spatial feature extraction through CNNs with temporal pattern recognition via LSTMs, resulting in superior detection performance across all attack categories.

However, successful real-world deployment requires addressing several limitations: NSL-KDD's age and simulated nature, class imbalance affecting minority attack detection, and the need for model interpretability in security-critical applications. The dataset's value lies primarily in providing standardized benchmarking for research comparison rather than direct production deployment.

Future work should focus on:

- Developing transfer learning approaches to adapt NSL-KDD trained models to contemporary network environments,
- Investigating Explainable AI frameworks for deep learning-based intrusion detection,
- Addressing class imbalance through advanced sampling and cost-sensitive learning techniques, and
- Integrating multiple datasets for more comprehensive evaluation.

The NSL-KDD dataset continues to serve as a valuable benchmark for intrusion detection research, enabling consistent comparison across different methodologies. While its limitations necessitate caution in direct production deployment, the performance improvements demonstrated by deep learning approaches provide clear evidence of their potential for advancing cybersecurity capabilities. Future research must balance the standardized evaluation benefits of NSL-KDD with the need for contemporary, diverse datasets that better represent modern threat landscapes.

REFERENCES

[1] R. Raman, A. Gupta, S. Das, and R. P. Singh, "Advancing cybersecurity and privacy with artificial intelligence: current trends and future research directions," *Frontiers in Computer Science*, vol. 6, Dec. 2024. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11656524/

[2] M. Charfeddine, M. Jmaiel, and A. Kachouri, "Advancing cybersecurity: a comprehensive review of AI-driven detection techniques," *Journal of Big Data*, vol. 11, no. 105, 2024. [Online]. Available: https://journalofbigdata.springeropen.com/articles/10.1186/s40537-024-00957-y

[3] M. Charfeddine, M. Jmaiel, and A. Kachouri, "Advancing cybersecurity: a comprehensive review of AI-driven detection techniques," *Journal of Big Data*, vol. 11, no. 105, 2024.

[4] A. Heidari, M. Eslami, S. Mirjalili, and L. Zhang, "Artificial intelligence and machine learning in cybersecurity: a deep dive into state-of-the-art techniques and future paradigms," *Knowledge and Information Systems*, 2025. [Online]. Available: https://link.springer.com/article/10.1007/s10115-025-02429-y

[5] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in *Proc. IEEE Symp. Comput. Intell. Secur. Def. Appl.*, 2009, pp. 1–6.

[6] [6] M. R. Asghar, M. H. Anisi, M. Sookhak, and A. Gani, "Explainable deep learning approach for advanced persistent threats (APTs) detection in cybersecurity: a review," *Artif. Intell. Rev.*, 2024. [Online]. Available: https://link.springer.com/article/10.1007/s10462-024-10890-4

[7] B. D. Deebak and S. O. Hwang, "Enhancing intrusion detection: a hybrid machine and deep learning approach," *J. Cloud Comput.*, vol. 13, 2024. [Online]. Available: https://journalofcloudcomputing.springeropen.com/articles/10.1186/s13677-024-00685-x

[8] A. Al-Omari, M. A. Alsmirat, and M. A. Alsmadi, "Accuracy based on the NSL-KDD data set," *Scientific Diagram, ResearchGate*, 2021. [Online]. Available: https://www.researchgate.net/figure/Accuracy-based-on-the-NSL-KDD-data-set_fig6_351100759

[9] M. Sarker, S. Hossain, N. Z. Budeiri, S. H. Alshamrani, and A. Alzain, "Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study," *Comput. Commun.*, vol. 50, pp. 124–139, 2020.

[10] Y. Fadili, A. Alaoui, and L. A. El Yacoubi, "A Survey on Cybersecurity Techniques Toward Convolutional Neural Network," in *Adv. Intell. Syst. Smart Technol., Lect. Notes Netw. Syst.*, vol. 826, Springer, 2024. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-031-47672-3_8

[11] F. S. Alrayes, M. I. Ghaleb, A. N. Zaidan, and B. B. Zaidan, "CNN Channel Attention Intrusion Detection System Using NSL-KDD Dataset," *Comput. Mater. Continua*, vol. 79, no. 3, pp. 4319–4347, 2024. [Online]. Available: https://www.techscience.com/cmc/v79n3/57130

[12] V. Hnamte, M. Kumar, and A. Jena, "A novel two-stage deep learning model for network intrusion detection: LSTM-AE," *IEEE Access*, vol. 11, pp. 37131–37145, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10101759

[13] S. M. Kasongo, "A deep learning technique for intrusion detection system using a Recurrent Neural Networks based framework," *Comput. Commun.*, vol. 199, pp. 113–125, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0140366422004601

[14] University of New Brunswick, "ISCX NSL-KDD dataset 2009," *Canadian Institute for Cybersecurity*, 2009. [Online]. Available: https://www.unb.ca/cic/datasets/nsl.html

[15] D. D. Protić, "Review of KDD Cup '99, NSL-KDD and Kyoto 2006+ datasets," *Vojnotehnički glasnik/Military Technical Courier*, vol. 66, no. 3, pp. 580–596, 2018.

[16] M. Zakariah, A. Basha, and K. S. Kuppusamy, "Intrusion Detection System for NSL-KDD Dataset Based on Deep Learning and Recursive Feature Elimination," *ResearchGate*, Sept. 2021. [Online]. Available: https://www.researchgate.net/publication/354523827

[17] J. Kim, J. Kim, H. Kang, and E. Lee, "Long Short Term Memory Recurrent Neural Network Classifier for Intrusion Detection," in *Proc. Int. Conf. Platform Technol. Serv.*, IEEE, 2016, pp. 1–5.

[18] K. Wu and Z. Chen, "A hybrid neural network model for intrusion detection," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 2, no. 1, pp. 41–50, 2018.