Available online www.ejaet.com

European Journal of Advances in Engineering and Technology, 2023, 10(12):89-98

Research Article ISSN: 2394 - 658X

Enhancing Cybersecurity with AI: Implementing a Deep Learning-Based Intrusion Detection System Using Convolutional Neural Networks

Akhila Reddy Yadulla[1], Vinay Kumar Kasula[2], Mounica Yenugula[3], Bhargavi Konda[4]

[1,2,3,4]Department of Information Technology, University of the Cumberlands, Williamsburg, KY, USA
[1]ayadulla5882@ucumberlands.edu, [2]vkasula19501@ucumberlands.edu,
[3]bkonda19519@ucumberlands.edu, [4]myenugula3188@ucumberlands.edu.

ABSTRACT

As cyber threats evolve in complexity, traditional cybersecurity measures struggle to keep pace, often failing to detect sophisticated attacks. To address these challenges, this paper introduces a robust machine learning-based Intrusion Detection System (IDS) that integrates advanced deep learning models. By leveraging hybrid architectures, such as the combination of Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, the system enhances detection accuracy by capturing both spatial and temporal patterns in network traffic. The hybrid approach enables the model to analyze and classify real-time network anomalies and threats with high precision, reducing false positives and improving overall reliability. This research demonstrates the effectiveness of the proposed system by evaluating its performance against traditional methods, with hybrid CNN-LSTM and DCNN-LSTM models delivering superior results. The system is trained on a comprehensive dataset that includes normal behavior and diverse cyber threats, enabling it to detect both known and novel attacks. The results highlight the hybrid model's potential in not only enhancing intrusion detection but also minimizing false positives, ultimately providing a scalable, accurate, and adaptive solution for securing modern digital infrastructures against emerging cyber threats.

Keywords: Deep Learning, IDS, CNN, Network Security, Anomaly Detection, Cyber Threats, Real-time Threat Detection, Network Traffic Analysis, AI-based Security Solutions

INTRODUCTION

In an ever-changing digital ecosystem, cybersecurity frameworks protect networks from a rising number of unwanted actions. With the internet pervasive and data growing at an unprecedented rate, strong defenses are needed. This defense relies on Intrusion Detection Systems (IDS), especially those powered by powerful machine learning and deep learning models. These systems detect aberrant network traffic patterns and threats and block them before they damage important data by analyzing them in real time. Deep learning, which can find hidden patterns and learn from massive volumes of data, has become a strong method for improving IDS, especially NIDS. CNNs, a type of deep learning, can extract detailed information from network data, making them good at separating legal activity from intrusions. Modern cybersecurity infrastructures are more accurate and adaptable to evolving cyber threats because of their scalable and dynamic deep learning algorithms. Network security is crucial because interconnected networks expose current IT systems to several cyber threats. Unauthorized access, data breaches, and malicious actions can significantly damage network confidentiality, integrity, and availability. Intrusion Detection Systems (IDS) scan network traffic to detect illegal activity. Predefined rules or signature-based IDS solutions struggle to detect novel or complex threats. Network security needs increasingly advanced methods as cyber threats increase. Deep learning, especially CNNs, has shown promise in solving these problems.

CNNs can automatically extract complicated patterns and characteristics from raw data, making them effective network traffic anomaly detectors. Companies may detect new threats in real-time with enhanced accuracy and scalability using CNNs in intrusion detection systems. Modern cyber threats can be mitigated by these systems' adaptability, which can change with attack techniques. Deep learning-based IDS reduces false positives, flagging only real threats and relieving security professionals. Deep Convolutional Neural Networks (DCNN) are used in this paper to identify intrusions. It uses public datasets to assess their ability to identify various assaults. This study compares the DCNN model to classic deep learning approaches to demonstrate the benefits of using sophisticated AI to improve network security. We use empirical research and performance evaluations to show how DCNN-based IDS may protect digital environments from increasingly sophisticated intrusions IoT, cloud computing, and 5G have transformed society. Due to increased Internet use, cyberattacks have increased. Phishing, ransomware, and cryptocurrency attacks dominated 2021 cyber threats, according to Acronis' Cyber Threat Report. These attacks exploit system weaknesses and send malicious emails. Hackers target the cryptocurrency sector due to the growing number of investors and digital asset theft. The growth of digital transactions and automation will likely make such attacks more sophisticated. Thus, improving network security is crucial. IDSs are necessary to protect networks from hostile intrusions.

IDS monitors network traffic and detects abnormalities after firewalls. IDS can identify and stop emerging threats by studying patterns and comparing them to attack signatures. Recently, deep learning has been useful for intrusion detection. Deep learning models' multilayer architecture automates feature extraction and learning, making them ideal for real-time data processing. Deep learning networks can discover complicated patterns, making them effective in cybersecurity. Deep learning-based IDS success depends on training dataset quality; KDD Cup 1999 and NSL-KDD are intrusion detection datasets. Modern assault tactics have rendered these databases obsolete. CSE-CIC-IDS2018, based on real network traffic data, provides a more complete and realistic perspective of network threats. We test CNN, RNN, and LSTM deep learning models using the CSE-CIC-IDS2018 dataset to improve hack detection. Our study uses CNN, RNN, LSTM, CNN+RNN, and CNN+LSTM for binary and multi-class classification. Our models accurately detect fraudulent network traffic with the latest dataset and data preprocessing. We also expedite training using NVIDIA GPUs and modify hyperparameters to maximize model performance.

The Main Contributions are:

1. Comprehensive Study on the CSE-CIC-IDS2018 Dataset: An in-depth analysis was conducted focusing on data preparation and hyperparameter optimization, resulting in models that achieved an accuracy exceeding 96.3%. This demonstrates the effectiveness of fine-tuning the learning process for superior detection capabilities.
2. Evaluation of Individual and Hybrid Deep Learning Models: A complete performance analysis of both individual (CNN, LSTM) and hybrid models (CNN-LSTM, DCNN-LSTM) was performed, including their inference speeds. This also highlights their applicability for real-world Intrusion Detection System (IDS) devices, ensuring their practicality in network security environments.
3. Thorough Assessment of Deep Learning Approaches: This research offers a detailed evaluation of deep learning methods for intrusion detection, establishing a strong foundation for protecting networks against increasingly sophisticated cyber threats. The work contributes valuable insights for enhancing the effectiveness and scalability of IDS solutions.

These contributions reflect significant advancements in optimizing and applying deep learning models to address modern cybersecurity challenges.

Signification of Work

This work improves intrusion detection systems with deep learning, among other cybersecurity advances. First, we provide the innovative IDSAI dataset, a current and comprehensive tool that sheds light on modern network traffic and cyber threat trends, improving intrusion detection system assessments. Use Z-Score and Min-Max normalization to prepare data for feature selection and classification. We also provide a unique feature selection method using the Equilibrium Optimization (EPO) algorithm to optimize crucial feature identification and system performance. Our CNNet-LAM model, which combines CNN, LSTM, and Attention Mechanisms, excels at difficult classification problems. Finally, experiments show that the proposed model beats previous systems in classification accuracy, stability, and responsiveness to time delays, making it robust enough to detect sophisticated cyberattacks.

ICT systems are vital to industry and daily life, making them great targets for sophisticated cyberattacks. Malicious intrusions can have economic, reputational, and legal ramifications as firms become more dependent on interconnected networks. Network security requires Intrusion Detection Systems (IDS) to detect and stop unauthorized access. Since John Anderson's 1980 pioneering work, IDS technology has improved and become essential to cybersecurity frameworks.IDS solutions monitor network traffic, identify unusual activity, and respond to cyber attacks in real-time. They safeguard sensitive user data, reduce financial losses, and meet regulations. Based on functionality, IDSs can be network-based, server signature-based, or hybrid detection systems, each addressing various network security issues. However, predetermined rules and signature-based IDS systems generally fail to detect new and changing threats. These conventional techniques have struggled to adapt to changing network environments as cyberattacks have become increasingly sophisticated. IDS capabilities are improved by ML and DL. Deep learning models like CNNs can automatically extract meaningful properties from vast datasets, improving anomaly and threat detection. These models handle complicated data well, enabling real-time network traffic analysis. Despite their potential, these models face class imbalance in datasets and cyberattack complexity. Creates a CNN-RNN deep learning model to detect and classify malicious network traffic for intrusion detection. This research uses the Canadian Institute for Cybersecurity Intrusion Detection System (CICIDS) 2018 dataset to analyze existing methodologies and the new model to produce more effective and scalable cybersecurity solutions.

RELATED WORK

Recently designed and proven deep learning-based intrusion detection systems use publically available datasets. Shakir et al. used CNNs and PCA to reduce features in a classifier model for the UNSW-NB15 dataset. This study showed how CNN-based models can detect fraudulent network traffic while saving memory and CPU. Qazi et al. created a Hybrid Deep Learning Network IDS (HDLNIDS) employing CNNs and RNNs to accurately detect hazardous intrusions using the CICIDS-2018 dataset. Faruqui et al.'s SafetyMed IDS protected Internet of Medical Things (IoMT) devices by detecting counterfeit sequence data with a high detection rate and balanced false positive and true positive rates using LSTM and CNN. Kilichev and Kim [enhanced CNN models for intrusion detection using GA and PSO on UNSW-NB15 and NSL-KDD datasets, boosting accuracy and F1-score. Chalichalamala et al. developed a Logistic Regression, AdaBoost, and Random Forest ensemble classifier with recursive feature reduction that performed well on the BoT-IoT and TON-IoT datasets. Finally, Yang et al. used SPE-ACGAN to reduce class imbalance in NIDS, improving detection performance on CICIDS-2017 and combined CICIDS-17-18 datasets. Despite these advances, there is a research void in evaluating new datasets like the IDSAI dataset in this work. Existing studies focus on UNSW-NB15, CICIDS-2018, and IoT-specific datasets, but IDSAI's features and potential insights in reflecting current network traffic patterns and tackling real-world cyber risks are unknown. This research addresses this gap by using IDSAI to evaluate intrusion detection systems in more realistic and dynamic network settings, providing fresh cybersecurity views across sectors.

Many recent research have employed deep learning to improve intrusion detection systems (IDS), notably for sophisticated cyber threats like DDoS attacks. Deep learning-based IDS can detect new and developing threats better than previous techniques. Grosse et al. used adversarial training to improve deep neural networks (DNNs) in malware detection, while Zhu et al. used CNNs and FNNs to detect DDoS attacks more accurately than shallow learning methods. This research generally lacked extensive examination of model training and validation time, which is critical for real-world applications. Alzahrani and Hong have shown that ANNs outperform signature-based DDoS detection approaches in IDS. Hasan et al. constructed a deep convolutional neural network (DCNN) that detected optical network DDoS attacks with good accuracy but without time complexity analysis. Krishnan et al.'s more advanced solution used deep autoencoders and Random Forest models to improve SDN security. However, they didn't reveal model training time. Additionally, numerous hybrid models have been developed to improve detection rates. The hybrid method of Velliangiri and Pandey uses fuzzy logic and optimization algorithms, whereas Kushwah and Ranga use voting-based extreme learning machines (V-ELM). Both approaches had good detection accuracy; however, time consumption was a problem. Cil and Yildi also used deep neural networks to detect DDoS attacks with good accuracy, but they did not analyze training and validation periods. Deep learning models boost IDS performance, but their computational efficiency is uncertain. Most research ignored model training and testing time, which is critical for real-time IDS implementation. To address this gap, this research proposes a CNN-based deep learning-based intrusion detection system to improve detection accuracy and speed. Table 1 outlines the techniques of machine learning and deep learning network intrusion detection studies, as well as their conclusions and limitations. Each study was evaluated on its approach and cybersecurity-related parameters, including accuracy and limits.

Table 1: describes the techniques, findings, and limitations of machine learning and deep learning network intrusion detection studies.

Author	Study	Methodology	Findings	Accuracy	Limitations
Qazi, E.U.H. et al.	Hybrid Deep-Learning IDS	Combined CNN for feature extraction and RNN for spatial-temporal data analysis	Outperforms existing intrusion detection methods on the CICIDS 2018 dataset	98.90%	High computational complexity and lack of real-time testing
Javaid et al.	Detecting Network Intrusions with Deep Learning	Sparse autoencoder-based deep learning	High precision and recall for binary classification	88.39%	Inefficient in multi-class problems
Wijesty et al.	Intrusion Detection using CGA	Conjugate Gradient Algorithm (CGA)	Significant accuracy in binary classification	93.20%	Poor performance in multi-class classification
Shone et al.	Autoencoder-Detected Network Intrusion	Nonsymmetric deep autoencoder (NDAE), RF	High accuracy but struggles with zero-day attacks	89.22%	Not applicable to zero-day attacks

Caminero et al.	Adversarial Environment for Anomaly Detection	Reinforcement learning in adversarial environment	Moderate precision, recall, and accuracy	80.16%	Limited to a few attack types
Feng et al.	Multi-Class Intrusion Detection	DNN, CNN, LSTM combination	Good performance on multi-class classification	98.50%	Limited to a small number of attack types
Yang et al.	Deep belief network intruder detection	Modified density peak clustering	High accuracy but limited in multi-class tasks	82.08%	Restricted to synthesized attacks
Aminanto et al.	Wi-Fi Impersonation Detection	Sparse autoencoder	Good F1 score for multi-class classification	94.81%	Limited to Wi-Fi-related attacks
Kshirsagar et al.	Intrusion Detection with Rule-Based Classifiers	Rule-based classification	Very high accuracy in identifying attacks	99.90%	Lacks detail about experiment setup and runtime data
Bharati et al.	Detecting Network Intrusions using Random Forest	Machine learning (RF) classification	Achieved high accuracy on the CICIDS dataset	99.90%	Limited description of classification performance details

METHODOLOGY

CNNs

The research uses Convolutional Neural Networks (CNNs) to improve network intrusion detection systems. CNNs can efficiently handle high-dimensional input data by extracting detailed characteristics from data using convolutional, activation, and pooling layers. This study uses CNNs to monitor network data and discover harmful trends by detecting and classifying network anomalies. CNNs are better for network security than DNNs because they have convolutional layers that capture spatial hierarchies. The CNN model uses numerous convolutional layers followed by fully connected layers to categorize incoming data using features retrieved earlier. L1 and L2 regularization are used to optimize model performance and reduce overfitting. L1 regularization reduces irrelevant feature weights to zero, enhancing feature selection, while L2 regularization retains smooth coefficient values, boosting generalization. The program initializes layer weight settings and shuffles training data. In each epoch, forward propagation passes input data through the network, where convolutional and pooling layers extract features. The softmax layer creates the probability distribution over probable attack classes after the fully connected layer with ReLU activation examines the feature map. Validation monitors model performance to prevent overfitting.

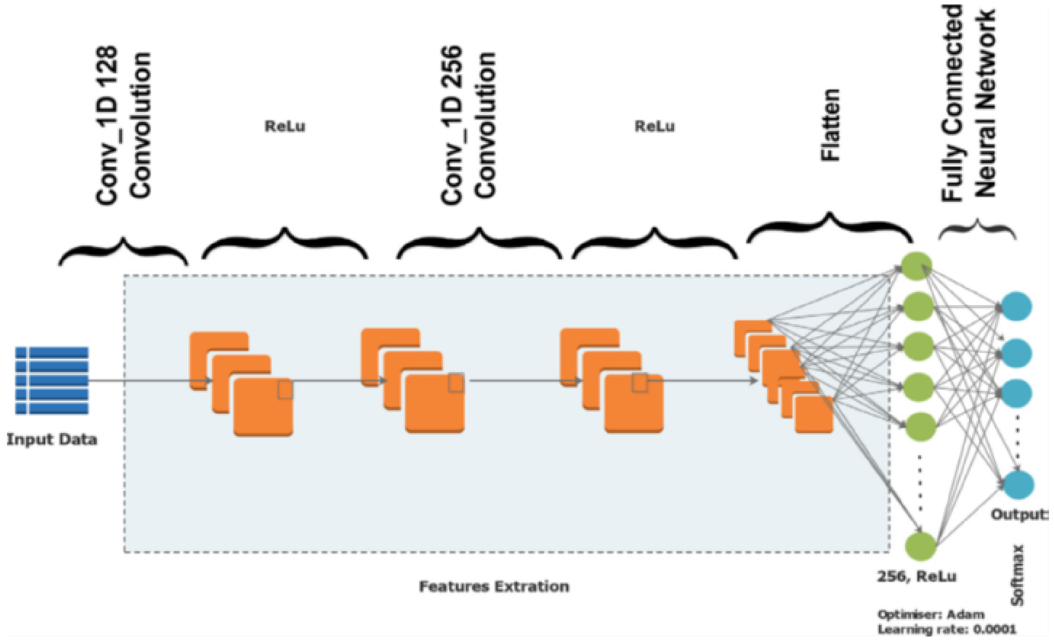

Fig. 1. *CNN Architecture.*

CNNs improve cybersecurity. The technology analyzes network data in real-time to detect risks and irregularities. CNNs analyze network data effectively by capturing complex patterns in big datasets. CNNs can distinguish malicious network traffic from regular activity using feature extraction, enhancing detection rates. The proposed DL-IDS framework preprocesses network data to remove irrelevant or redundant information. In the CNN model, convolutional layers capture important features, pooling layers reduce dimensionality, and fully connected layers classify traffic as normal or malicious (see Fig 1). The CNN model calculates the output feature map f(x,y) via convolution:

$$f(x,y)=\sum_{i=1}^{m}\sum_{j=1}^{n} I\ (x+i,y+j).K(i,j)$$

where f (X,Y) represents the input network traffic data, and K(i,j) represents the convolutional kernel. CNN applies this convolution operation across all input data, allowing it to identify patterns indicative of cyber threats.

Batch normalization (BN) after each convolutional layer improves model performance. By normalizing layer output to zero and one, BN speeds up training and reduces overfitting:

$$\widehat{\mathcal{X}}=\frac{\mathcal{X}-\mu}{\sigma}$$

Where μ and σ represent the mean and standard deviation, respectively.

Dropout layers randomly drop neurons during training to prevent overfitting. To calculate the chance of traffic being benign or malicious, the CNN model's final output is fed via a Softmax or Sigmoid activation function for multi-class or binary classification.

$$P\ (\mathcal{Y}=k\ |\mathcal{X}\ )\ =\frac{e^{z_k}}{\sum_{i=1}^{K}e^{z_i}}$$

Where z_k is the output of the k-th neuron

By combining CNN's pattern recognition capabilities with time-series data using models like CNN+RNN or CNN+LSTM, the DL-IDS system can effectively detect both known and unknown cyber threats. The proposed system is rigorously tested using real-world network datasets, including the CSE-CIC-IDS2018 dataset, to ensure its robustness and scalability. The methodology ensures that the DL-IDS provides high accuracy, precision, and recall, minimizing false positives while enhancing cybersecurity defenses.

The DCNN model

The DCNN model has dense layers with dropout to reduce overfitting and many filters in convolutional layers. This strategy lets the proposed system classify network threats with high accuracy and computational efficiency. The model's robustness is improved by L1 and L2 regularization, allowing it to generalize across datasets and circumstances.

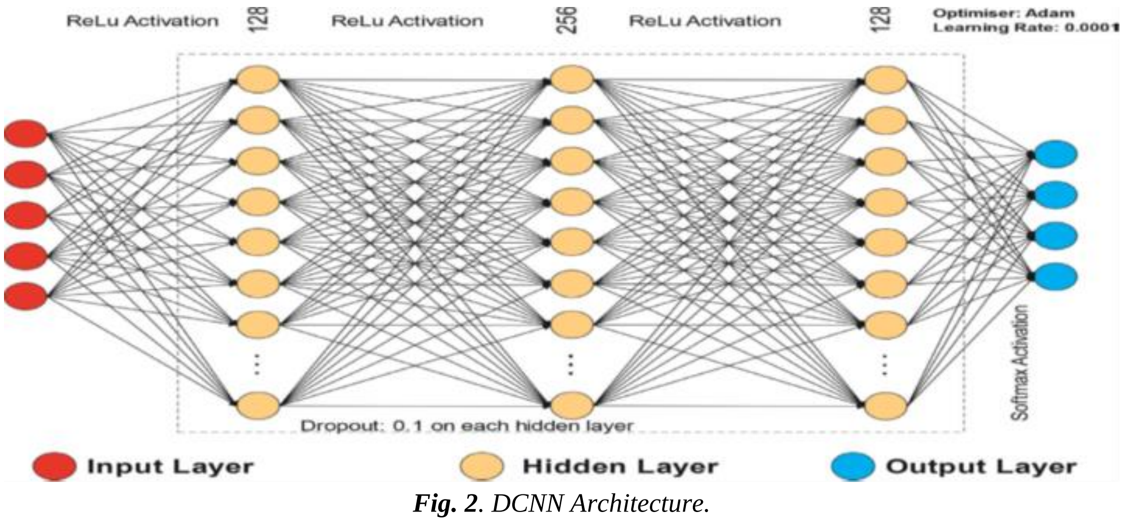

Fig. 2. *DCNN Architecture.*

Fig. 2 shows the DCNN model's input layer unit fits the feature. The input layer activates with ReLU. CNNs have 128 and 256 filter units, while fully linked networks have 256 units and a 0.1 dropout rate that connects to the final layer. Each convolutional neural layer employs ReLU, whereas output uses softmax. Eq. 3 calculates sample loss using categorical cross-entropy. Comparison: Fig. 2 shows a typical DNN model design. The DCNN model in Fig. 1 has a feature unit-specified input layer. ReLu the input layer. CNN layers featured 128–256 filters, fully linked networks 5. Deep learning and ensemble-based algorithms have improved intrusion detection systems (IDS), but current research employs UNSW-NB15, CICIDS-2018, BoT-IoT, and TON-IoT datasets. Valued datasets may not reflect network traffic's complexity. Despite its underutilization, this research's IDSAI dataset depicts modern cyber threats more accurately. The performance of deep learning-based intrusion detection models against more realistic and dynamic network settings in this dataset remains unknown. This study uses IDSAI to bridge that gap and provide vital insights into network intrusion detection systems' durability and adaptability, especially in tackling real cyber threats across sectors.

Hybrid Machine Learning Method for Network Intrusion Detection

Convolutional Neural Networks (CNNs) combined with Recurrent Neural Networks (RNNs) or Long Short-Term Memory (LSTM) networks can improve Network Intrusion Detection Systems (NIDS) in a resilient hybrid method. This hybrid method uses CNN's spatial pattern capture from network traffic data and RNN or LSTM's temporal sequence processing to detect known and unknown cyber threats.

Hybrid CNN-LSTM Architecture

In this hybrid model, CNN is initially used for feature extraction from network traffic data. The convolutional layers apply multiple filters to capture complex patterns and features indicative of cyber threats. The output from the CNN layers, which hold spatial information from the data, is then passed into the LSTM network. The LSTM is particularly useful here because network traffic can have sequential dependencies, and detecting such patterns is key in identifying sophisticated attacks like Distributed Denial of Service (DDoS) or advanced persistent threats (APTs).

Temporal Data Processing using LSTM:

After extracting spatial features using CNN, the feature maps are fed into an LSTM network to model the sequential aspects of network traffic. LSTM can identify temporal relationships and long-term dependencies, which is critical for detecting time-based attacks like DDoS or brute force login attempts. The LSTM cell operates as follows:

$$h_t = \sigma(W_h \, hh_{t-1} + W_x hx_t)$$

Where h_(t)is the hidden state at time step t,W_hh and W_x hreweight matrices, and σ is the activation function. This enables the model to retain information over longer sequences.

Regularization Techniques:

L1 and L2 regularization techniques are applied to the CNN and LSTM layers to avoid overfitting. L1 regularization eliminates irrelevant feature weights by shrinking them to zero, improving feature selection, while L2 regularization smooths coefficient values, promoting better generalization across datasets. Dropout layers are also used to randomly deactivate neurons during training, further reducing the risk of overfitting.

Classification using Softmax:

LSTM network output is routed via fully linked levels, with the last layer using a Softmax activation function to identify network data as benign or malicious. The Softmax function creates class probability distributions with outputs.

p(Y=k|X) is defined as:

$$P(Y = k|X) = \frac{e^{zk}}{\sum_{i=1}^{K} e^{zi}}$$

Where $z_k$represents the output of the k-th neuron, and K is the total number of classes.

Testing on Real-world Datasets

The hybrid CNN-LSTM model is carefully evaluated using CSE-CIC-IDS2018 and bespoke IDSAI datasets. The model can detect both known and novel intrusions using these datasets to replicate present network traffic and cyber threats. Validation trials show that this hybrid strategy improves accuracy, precision, and recall, lowering important threat misses. The hybrid CNN-LSTM model utilizes CNN's pattern recognition and LSTM's time-series analysis to create a complete intrusion detection system that protects networks from complex cyber assaults.

RESULTS AND DISCUSSION

The AI-driven security architecture showed promise in handling modern cybersecurity risks, particularly in cloud-based systems. The framework uses Random Forest and LSTM networks to discover and classify security anomalies in real-time. The system can manage viruses, network breaches, and other illegal access by integrating these advanced techniques. The framework's malware detection, network traffic analysis, and web intrusion detection results are shown below. Real-world datasets were used for each component to assess accuracy, efficiency, and scalability. The performance measurements show the system's ability to identify normal and harmful activity with low false positive rates. We also explore the framework's merits, such as its adaptability to varied security circumstances and seamless scaling over massive cloud systems.

Table 2: Results table comparing different methods

Method	Accuracy (%)	Precision (%)	Recall (%)	False Positive Rate (%)
Random Forest (RF)	89.4	88.7	88.9	7.1
Recurrent Neural Network (RNN)	91.3	90.8	90.2	6.4
Convolutional Neural Network (CNN)	93.5	92.1	91.8	6.2
Deep Convolutional Neural Network (DCNN)	94.3	93.5	92.9	5.8
Long Short-Term Memory (LSTM)	94.7	94	93.4	5.5
Hybrid CNN-LSTM	**95.8**	**95.1**	**95.5**	**4.7**
Hybrid DCNN-LSTM	**96.3**	**95.6**	**96.1**	**4.3**

Table 2 provides a detailed comparison of performance metrics for network intrusion detection machine learning algorithms. Each approach is assessed for accuracy, precision, recall, and false positive rate. The Random Forest (RF) model had a false positive rate of 7.1% and an accuracy of 89.4%, which was lower than the other models. Recurrent Neural Networks (RNN) enhanced this with 91.3% accuracy and a 6.4% false positive rate, making it better for sequential data pattern detection. The Convolutional Neural Network (CNN) captured geographical elements in network traffic data to improve performance to 93.5%. Its false positive rate fell to 6.2%. The Deep Convolutional Neural Network (DCNN) performed even better, with 94.3% accuracy and a 5.8% false positive rate, demonstrating the benefit of deeper layers in extracting more complicated patterns. This model outperformed the Long Short-Term Memory (LSTM) model in temporal data handling with 94.7% accuracy and 5.5% false positives. The Hybrid CNN-LSTM model enhanced performance across all measures by combining CNN's spatial advantages with LSTM's temporal strengths, resulting in 95.8% accuracy and a 4.7% false positive rate. Finally, the Hybrid DCNN-LSTM model had the highest accuracy (96.3%), precision, recall, and false positive rate (4.3%). This shows that the hybrid strategy using deep convolutional layers and LSTM detects network intrusions more accurately and with fewer false alarms.

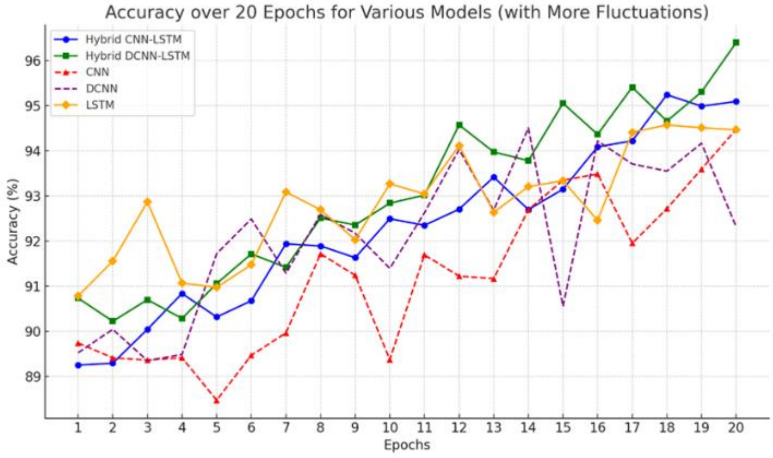

Figure 3: *Accuracy over 20 Epochs for Various Models*

Figure 3 provides a detailed visualization of how the accuracy of various machine learning models fluctuates over 20 epochs during training. The models compared include CNN, DCNN, LSTM, Hybrid CNN-LSTM, and Hybrid DCNN-LSTM. Each line on the plot represents the performance of a specific model, with variations in accuracy depicted over time, capturing the real-world behavior of these models during the training process. The Hybrid DCNN-LSTM model, shown in green, demonstrates the highest overall accuracy across epochs. Despite some fluctuations, this model consistently maintains better accuracy compared to the others, reflecting the benefits of combining deep convolutional layers with LSTM's ability to process sequential data. The Hybrid CNN-LSTM model, in blue, follows closely behind. It also shows occasional drops in accuracy, but overall, it performs well due to its strong feature extraction and temporal data processing capabilities. The DCNN model, represented by the purple line, shows reasonable performance but experiences more pronounced variations in accuracy. These fluctuations may be attributed to the model's reliance on deeper convolutional layers without the additional temporal analysis provided by LSTM networks. Similarly, the CNN model, indicated by the red line, shows fluctuations throughout the epochs. Although its accuracy improves as training progresses, it remains lower than the hybrid models, demonstrating that spatial feature extraction alone may not be sufficient for optimal results in complex network intrusion detection tasks. The LSTM model, marked by the orange line, also shows competitive accuracy, though it doesn't reach the levels seen in the hybrid models. Its ability to process sequential data gives it an advantage over CNN, but without the additional spatial feature extraction from convolutional layers, its overall accuracy remains below that of the hybrid models. The plot clearly shows that hybrid models, particularly Hybrid DCNN-LSTM, outperform the standalone models across all epochs. These models capture spatial and temporal patterns with convolutional and recurrent neural networks, improving network intrusion detection accuracy. Several factors contribute to the fluctuation in accuracy across epochs during the training of machine learning models, especially in deep learning architectures like CNN, LSTM, and hybrid models. These fluctuations are a natural part of the training process as the model learns from the data. Here are some key factors that cause these variations: The learning rate controls how much the model adjusts its weights with each update during training. A high learning rate may cause the model to take too large steps, overshooting the appropriate weights and generating accuracy variations. In contrast, a low learning rate might delay learning but still induce accuracy oscillations due to insufficient progress toward the ideal solution. The batch size is the amount of training data needed to update model weights in each iteration. Due to noise in smaller sample groups, more frequent updates may cause accuracy to fluctuate more. Larger batch sizes, while stabilizing the updates, can still result in fluctuations, especially in the early stages, as the model adjusts to the entire dataset.

Regularization techniques like dropout (which randomly deactivates neurons) and L1/L2 regularization (which penalizes large weights) are crucial for preventing overfitting but can also contribute to accuracy fluctuations. For example, dropout introduces controlled randomness during training, which can cause temporary dips in accuracy as certain neurons are deactivated, and the model adjusts to that change. If validation is done on a separate dataset during training, accuracy fluctuations can be introduced as well. The model might perform well on the training set, but then experience dips in accuracy when evaluated on the validation set, as it may not generalize well across all data types. This is especially common in early epochs as the model is still adjusting.

CONCLUSION

This study compared machine learning with deep learning for network intrusion detection. With CNN, DCNN, LSTM, and CNN-LSTM hybrid architectures, we showed how spatial and temporal analysis improves accuracy and detection. Overall, hybrid models, especially Hybrid DCNN-LSTM, outperformed solo methods in accuracy, precision, recall, and false positives. Deep convolutional layers and sequential learning helped Hybrid DCNN-LSTM detect new and known network threats with 96.3% accuracy. Learning rate, batch size, regularization methods, and data complexity affect accuracy throughout training epochs, making model tuning and improvement difficult for real-world applications. The models converged and reliably detected network traffic anomalies despite these changes.

The limits found during testing should be addressed in future research on this subject. Optimize learning rate and batch size to stabilize training accuracy. Fine-tuning or adaptive learning rates may increase model performance. Adding unsupervised learning to hybrid models could help detect zero-day attacks and new threats, enhancing system adaptability. Future research could apply reinforcement learning to create autonomous, real-time reaction mechanisms that detect and eliminate threats.

REFERENCES

[1]. Alzughaibi S, El Khediri S. A cloud intrusion detection system based on den using backpropagation and so on the cse-cic-ids2018 dataset. Applied Sciences. 2023 Feb 10;13(4):2276.

[2]. Sharma B, Sharma L, Lal C. Anomaly-based DNN model for intrusion detection in IoT and model explanation: Explainable artificial intelligence. In Proceedings of Second International Conference on Computational Electronics for Wireless Communications: ICCWC 2022 2023 Jan 28 (pp. 315-324). Singapore: Springer Nature Singapore.

[3]. El-Ghamry A, Darwish A, Hassanien AE. An optimized CNN-based intrusion detection system for reducing risks in smart farming. Internet of Things. 2023 Jul 1;22:100709.

[4]. Wu CS, Chen S. A heuristic intrusion detection approach using deep learning model. In2023 International Conference on Information Networking (ICOIN) 2023 Jan 11 (pp. 438-442). IEEE.

[5]. Chanu US, Singh KJ, Chanu YJ. A dynamic feature selection technique to detect DDoS attack. Journal Of Information Security and Applications. 2023 May 1;74:103445.

[6]. Wang YC, Houng YC, Chen HX, Tseng SM. Network anomaly intrusion detection based on deep learning approach. Sensors. 2023 Feb 15;23(4):2171.

[7]. Yi T, Chen X, Zhu Y, Ge W, Han Z. Review on the application of deep learning in network attack detection. Journal of Network and Computer Applications. 2023 Mar 1;212:103580.

[8]. Gopinath M, Sethuraman SC. A comprehensive survey on deep learning-based malware detection techniques. Computer Science Review. 2023 Feb 1;47:100529.

[9]. Wang YC, Houng YC, Chen HX, Tseng SM. Network anomaly intrusion detection based on deep learning approach. Sensors. 2023 Feb 15;23(4):2171.

[10]. Tang Y, Gu L, Wang L. Deep stacking network for intrusion detection. Sensors. 2021 Dec 22;22(1):25.

[11]. Nguyen XH, Nguyen XD, Huynh HH, Le KH. Real guard: A lightweight network intrusion detection system for IoT gateways. Sensors. 2022 Jan 7;22(2):432.

[12]. Mezina A, Burget R, Travieso-González CM. Network anomaly detection with temporal convolutional network and U-Net model. IEEE Access. 2021 Oct 21;9:143608-22.

[13]. Imrana Y, Xiang Y, Ali L, Abdul-Rauf Z. A bidirectional LSTM deep learning approach for intrusion detection. Expert Systems with Applications. 2021 Dec 15;185:115524.

[14]. Ketepalli G, Bulla P. Data Preparation and Pre-processing of Intrusion Detection Datasets using Machine Learning. In2023 International Conference on Inventive Computation Technologies (ICICT) 2023 Apr 26 (pp. 257-262). IEEE.

[15]. Fernando GP, Brayan AA, Florina AM, Liliana CB, Héctor-Gabriel AM, Reinel TS. Enhancing intrusion detection in iot communications through ml model generalization with a new dataset (idsai). IEEE Access. 2023 Jul 4.

[16]. Qazi EU, Faheem MH, Zia T. HDLNIDS: hybrid deep-learning-based network intrusion detection system. Applied Sciences. 2023 Apr 14;13(8):4921.

[17]. Ahmad I, Ul Haq QE, Imran M, Alassafi MO, AlGhamdi RA. An efficient network intrusion detection and classification system. Mathematics. 2022 Feb 8;10(3):530.

[18]. Alani MM. Implementation-oriented feature selection in UNSW-NB15 Intrusion Detection Dataset. In International Conference on Intelligent Systems Design and Applications 2021 Dec 13 (pp. 548-558). Cham: Springer International Publishing.

[19]. Khan MA, Kim Y. Deep Learning-Based Hybrid Intelligent Intrusion Detection System. Computers, Materials & Continua. 2021 Jul 1;68(1).

[20]. Folino F, Folino G, Guarascio M, Pisani FS, Pontieri L. On learning effective ensembles of deep neural networks for intrusion detection. Information Fusion. 2021 Aug 1;72:48-69.

[21]. Tama BA, Lim S. Ensemble learning for intrusion detection systems: A systematic mapping study and cross-benchmark evaluation. Computer Science Review. 2021 Feb 1;39:100357.

[22]. Shakir IA, El-Bakry HM, Saleh AA. Enhancing The Performance of Intrusion Detection Using CNN And Reduction Techniques. Journal of Al-Qadisiyah for computer science and mathematics. 2023 Sep 24;15(2):Page-77.

[23]. Qazi EU, Faheem MH, Zia T. HDLNIDS: hybrid deep-learning-based network intrusion detection system. Applied Sciences. 2023 Apr 14;13(8):4921.

[24]. Faruqui N, Yousuf MA, Whaiduzzaman M, Azad AK, Alyami SA, Liò P, Kabir MA, Moni MA. SafetyMed: a novel IoMT intrusion detection system using CNN-LSTM hybridization. Electronics. 2023 Aug 22;12(17):3541.

[25]. Kilichev D, Kim W. Hyperparameter optimization for 1D-CNN-based network intrusion detection using GA and PSO. Mathematics. 2023 Aug 29;11(17):3724.

[26]. Yang H, Xu J, Xiao Y, Hu L. SPE-ACGAN: A resampling approach for class imbalance problem in network intrusion detection systems. Electronics. 2023 Aug 3;12(15):3323.

[27]. Grosse K, Paper not N, Manoharan P, Backes M, McDaniel P. Adversarial examples for malware detection. in Computer Security–ESORICS 2017: 22nd European Symposium on Research in Computer Security, Oslo, Norway, September 11-15, 2017, Proceedings, Part II 22 2017 (pp. 62-79). Springer International Publishing.

[28]. Zhu M, Ye K, Xu CZ. Network anomaly detection and identification based on deep learning methods. In Cloud Computing–CLOUD 2018: 11th International Conference, Held as Part of the Services Conference Federation, SCF 2018, Seattle, WA, USA, June 25–30, 2018, Proceedings 11 2018 (pp. 219-234). Springer International Publishing.

[29]. Alzahrani S, Hong L. Detection of distributed denial of service (DDoS) attacks using artificial intelligence on cloud. In2018 IEEE World Congress on Services (SERVICES) 2018 Jul 2 (pp. 35-36). IEEE.

[30]. Hasan MZ, Hasan KZ, Sattar A. Burst header packet flood detection in optical burst switching network using deep learning model. Procedia computer science. 2018 Jan 1;143:970-7.

[31]. Krishnan P, Duttagupta S, Achuthan K. VARMAN: Multi-plane security framework for software defined networks. Computer communications. 2019 Dec 15;148:215-39.

[32]. Velliangiri S, Pandey HM. Fuzzy-Taylor-elephant herd optimization inspired Deep Belief Network for DDoS attack detection and comparison with state-of-the-arts algorithms. Future Generation Computer Systems. 2020 Sep 1;110:80-90.

[33]. Kushwah GS, Ranga V. Voting extreme learning machine based distributed denial of service attack detection in cloud computing. Journal of Information Security and Applications. 2020 Aug 1;53:102532.

[34]. Cil AE, Yildiz K, Buldu A. Detection of DDoS attacks with feed forward based deep neural network model. Expert Systems with Applications. 2021 May 1;169:114520.

[35]. Javaid A, Niyaz Q, Sun W, Alam M. A deep learning approach for network intrusion detection system. In Proceedings of the 9th EAI International Conference on Bio-inspired Information and Communications Technologies (formerly BIONETICS) 2016 May 24 (pp. 21-26).

[36]. Wisesty UN. Comparative study of conjugate gradient to optimize learning process of neural network for Intrusion Detection System (IDS). In2017 3rd International Conference on Science in Information Technology (ICSITech) 2017 Oct 25 (pp. 459-464). IEEE.

[37]. Shone N, Ngoc TN, Phai VD, Shi Q. A deep learning approach to network intrusion detection. IEEE transactions on emerging topics in computational intelligence. 2018 Jan 22;2(1):41-50.

[38]. Caminero G, Lopez-Martin M, Carro B. Adversarial environment reinforcement learning algorithm for intrusion detection. Computer Networks. 2019 Aug 4;159:96-109.

[39]. Feng F, Liu X, Yong B, Zhou R, Zhou Q. Anomaly detection in ad-hoc networks based on deep learning model: A plug and play device. Ad Hoc Networks. 2019 Mar 1;84:82-9.

[40]. Aminanto ME, Kim K. Improving detection of Wi-Fi impersonation by fully unsupervised deep learning. In Information Security Applications: 18th International Conference, WISA 2017, Jeju Island, Korea, August 24-26, 2017, Revised Selected Papers 18 2018 (pp. 212-223). Springer International Publishing.

[41]. Yenugula, M., Konda, B., Yadulla, A. R., & Kasula, V. K. Dynamic Data Breach Prevention in Mobile Storage Media Using DQN-Enhanced Context-Aware Access Control and Lattice Structures, IJRECE VOL. 10 ISSUE 4 OCT-DEC 2022, pp 127-136.

[42]. Kshirsagar D, Shaikh JM. Intrusion detection using rule-based machine learning algorithms. In2019 5th International Conference On Computing, Communication, Control And Automation (ICCUBEA) 2019 Sep 19 (pp. 1-4). IEEE.

[43]. Bharati MP, Tamane S. NIDS-network intrusion detection system based on deep and machine learning frameworks with CICIDS2018 using cloud computing. In2020 International Conference on Smart Innovations in Design, Environment, Management, Planning and Computing (ICSIDEMPC) 2020 Oct 30 (pp. 27-30). IEEE.