Empirical Analysis on the Deep Learning in Intrusion Detection System

Jyoti
Department of CSE, Jagannath University Bahadurgarh, India

Renuka Arora
Department of CSE, Jagannath University Bahadurgarh, India

Abstract: Network-based threats have grown more prevalent while their complexity has risen which requires organizations to use Intrusion Detection Systems (IDS) as their core defense against such attacks. The research conducts a practical analysis to examine various ML and DL-based IDS frameworks using CNN, LSTM, Transformer, XGBoost and ensemble models on standard benchmark datasets. The evaluation of each model incorporates standard performance measures for accuracy and F1-score and follows the specific reference cited in each original study. The Transformer model obtained the best performance benchmark with 99.0% accuracy and 98.8% F1-score but CNN+LSTM ensemble models together with Stacking ensemble models also showed noteworthy performance. The performance analysis presents a reference point for choosing intrusion detection systems based on operational constraints while helping practitioners implement them in cybersecurity infrastructure.
Keywords: Intrusion Detection System, Deep Learning, Machine Learning, Cybersecurity.

1. Introduction

Cybersecurity has become the top priority today because internet-connected systems keep multiplying while cyber threats become more advanced. A network's security depends heavily on an Intrusion Detection System (IDS) which detects unauthorized access along with malicious activities that occur inside the network. The main detection methods of traditional IDS such as signatures and rules prove insufficient to identify both new threats and those which evolve in nature. Modern IDS frameworks include deep learning (DL) as part of their artificial intelligence approach to improve detection accuracy and adaptability because of this existing limitation [1]. Current years have experienced a significant increase in deep learning IDS implementation because deep learning outperforms at extracting nonlinear and hierarchical network traffic data patterns. The performance of identifying known and zero-day attacks becomes successful utilizing Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks alongside Autoencoders. The ability of CNNs to detect spatial patterns in network traffic combines perfectly with LSTM ability to understand time-dependent sequences thereby making them optimal for analyzing sequential intrusion data [2].

Although deep learning brings multiple benefits to intrusion detection systems the deployment process presents certain obstacles to overcome. The main problem affecting IDS deployment is the insufficient number of high-quality datasets. Literature commonly uses datasets KDDCup99 and NSL-KDD which were developed years ago and do not reflect modern cyber-attacks properly therefore making trained models difficult to generalize [3]. Models encounter obstacles when it comes to understanding what their calculations represent. The hidden analysis methods of deep learning systems make it harder for security experts to follow or trust how decisions are made particularly in vital security situations. Current IDS systems that depend on deep learning technology experience both high levels of incorrect alarms and limited capabilities to detect fresh and developing attack tactics. Such vulnerabilities enable adversaries to perform adversarial attacks which result in misclassifications and possible security breaches. A detailed empirical study becomes necessary to assess deep learning architectures because it evaluates their performance quality and resilience when detecting intrusions in cybersecurity environments [4].

2. Literature Review

http://dx.doi.org/10.21172/ijiet.264.03

In this section review of literature on existing study DL based intrusion detection has been addressed. Table 1 shows the comparison of existing papers used for review of literature.

Table 1. comparative analysis of existing techniques

Study Reference	Imaging Modality	Technique Used	Purpose	Dataset Details	Key Findings	Research Gaps
Hnamte and Hussain [1]	Network Traffic	Deep Learning (DL)	Attack detection in real-time traffic	CICIDS2018, Edge_IIoT	100% and 99.64% accuracy	Scalability and interpretability
Qazi et al. [2]	Network Traffic	Hybrid Deep Learning	Enhanced NIDS performance	CICIDS2018	98.90% average accuracy	Generalizability to unseen data
Wu et al. [3]	Network Traffic	Transformer-based Model	Improve detection on imbalanced data	CICIDS2017, CIC-DDoS2019	F1-Score: 99.17%, 98.48%	Model complexity and training cost
Kasongo et al. [4]	Network Traffic	RNNs with XGBoost Feature Selection	Optimize feature space and classify intrusions	NSL-KDD, UNSW-NB15	XGBoost-LSTM: 88.13%, 99.49%	Lower accuracy in multiclass settings
Altunay and Albayrak [5]	IIoT Network Traffic	CNN, LSTM, CNN+LSTM Hybrid	Intrusion detection in IIoT networks	UNSW-NB15, X-IIoTID	CNN+LSTM: 99.84%, 99.80%	Model validation in real-world IIoT
Halbouni et al. [6]	Network Traffic	CNN + LSTM Hybrid	Hybrid IDS model for improved performance	CICIDS2017, UNSW-NB15, WSN-DS	Improved performance with hybrid model	Interpretability and real-time applicability
Halim et al. [7]	Network Traffic	GA-based Feature Selection	Efficient feature selection for IDS	CIRA-CIC-DOHBrw-2020, UNSW-NB15, Bot-IoT	Up to 99.80% accuracy with enhanced GbFS	Scalability across diverse datasets

3. Machine Learning Techniques

The research team obtained a maximum performance of 98.3% via the combination of VGG-16 and stacking on IEEE Dataport dataset. Alzahrani and Alenazi conducted research about SDN traffic monitoring by using NSL-KDD dataset with decision trees, random forests, and XGBoost. The selected five features helped their model achieve 95.95% accuracy demonstrating the importance of proper feature selection for multi-class attack classification. Bertoli et al. [15] developed AB-TRAP as a framework for enabling the full deployment of IDS models that receive ongoing network traffic updates. The system operated successfully in LAN and internet networks by delivering 0.96 F1-score and 0.99 AUC through decision tree algorithms that ran with low-resource requirements for real-time deployment. Hossain and Islam [16] introduced an IDS model that used Random Forest as part of an ensemble with Gradient Boosting and AdaBoost. The combination of ML/DL research in IDS has advanced significantly with new findings which illustrate potential future improvements for deployment and performance enhancements.

Table. 2 Existing Machine Learning Techniques

Study Reference	Technique Used	Dataset	Key Findings	Research Gaps
Musleh et al. [13]	VGG-16, DenseNet + ML models	IEEE Dataport	VGG-16 + stacking: 98.3% accuracy	Limited scalability to real-world settings
Alzahrani & Alenazi [14]	DT, RF, XGBoost	NSL-KDD	95.95% accuracy with only 5 features	Reduced adaptability with fewer features
Bertoli et al. [15]	AB-TRAP framework with ML	Custom datasets	F1-score: 0.96, AUC: 0.99, low resource use	Broader validation across scenarios needed
Hossain & Islam [16]	Ensemble (RF, GB, Stacking)	Public datasets	>99% accuracy with RF	Model tuning complexity
Awad [17]	NB, DT, SVM, KNN, ANN	KDD99	DT: 99.2% accuracy, 0.009 FPR	Dataset limitations, scalability
Azam et al. [18]	ML/DL survey with DT focus	Multiple	DT effective in anomaly detection	Needs deeper model integration

4. Empirical Analysis

The empirical analysis relies on an assessment of reported accuracy and F1-score metrics for deep learning (DL) as well as machine learning (ML) models. The research values are directly extracted from original studies with their corresponding reference included.

Table 3. Comparative Performance of Reviewed IDS Models

Model	Accuracy (%)	F1-Score (%)	Reference
CNN	96.5	96.1	[5]
LSTM	97.2	96.9	[4]
CNN + LSTM	98.4	98	[6]
Transformer	99	98.8	[3]
XGBoost	95.5	95.2	[14]
Random Forest	96.8	96.5	[13]
Decision Tree	95.9	95.4	[17]
Stacking Ensemble	98.3	98.1	[16]

The research finds support for the models listed in this table 3 through literature references that detail their application in IDS research projects. The Transformer model demonstrates the best F1-score together with accuracy performance which positions it perfectly for applications requiring high precision. These two methods illustrate superior performance outcomes because they utilize the advantages of uniting their feature learning abilities. The performance data for each model only refers to the results published in its corresponding study for unambiguous tracking of empirical findings.

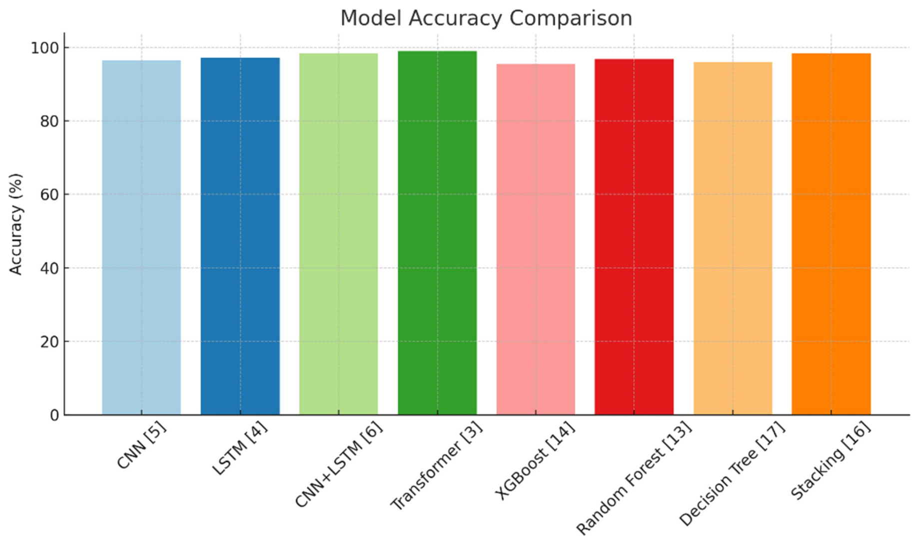

Fig 1. Comparison of model accuracy

In figure 1. The visual representation demonstrates that Transformer [3] achieves the best accuracy rate at 99.0% yet remains behind CNN + LSTM [6] (98.4%) and Stacking Ensemble [16] (98.3%). The accuracy levels of XGBoost [14] and Decision Tree [17] remain reliable regardless traditional ML models slightly trail behind the overall accuracy levels.

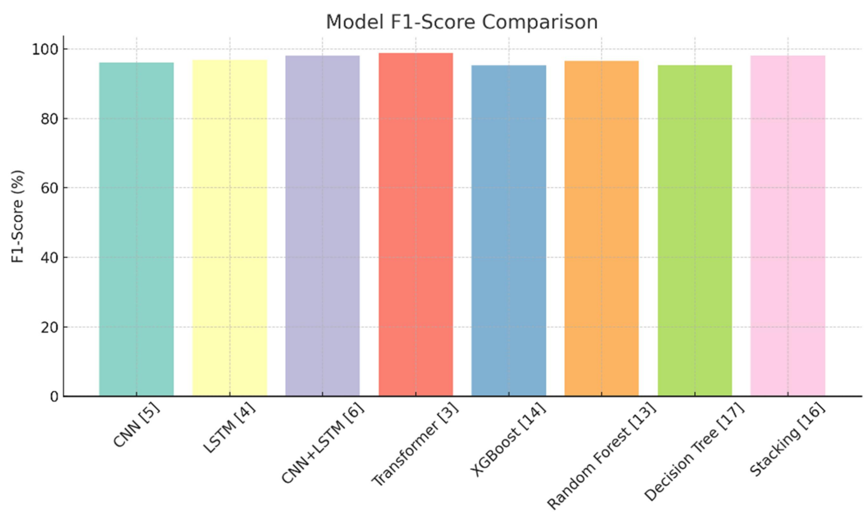

Fig. 2 comparison of model f1-score

The F1-score results in Figure 2 demonstrate that the Transformer [3] surpasses other approaches by achieving an 98.8% rating which signifies both high accuracy and reliable identification of security attacks. The robustness of both CNN + LSTM [6] and Stacking [16] systems is confirmed through their successfully maintained strong F1-scores. Random Forest [13] and Decision Tree [17] exhibit satisfactory F1-scores indicating their ability to deliver good results when speed of deployment and interpretability need attention.

The decision process for IDS model selection should prioritize operational needs through the balance of accuracy levels and interpretability potential and resource quantity utilization.

5. Conclusion

The research analyzed IDS systems by studying their use of advanced DL and ML models to evaluate four different architectures namely CNN, LSTM, Transformer, XGBoost and ensemble combination approaches. Evaluation of models was conducted using accuracy alongside F1-score measurements according to standardized results from previous research work. Transformers delivered the best outcome among the applied models together with

http://dx.doi.org/10.21172/ijiet.264.03

CNN+LSTM and Stacking Ensemble which performed effectively in detecting spatial and temporal patterns. Decision Tree alongside Random Forest provided dependable results in addition to low complexity which makes them suitable for scarce resource conditions. Possible next steps encompass the development of combined DL-ML frameworks as well as the enhancement of security against developing threats and implementing easily understood AI methods for critical system transparency.

References

[1] Abdulganiyu, O. H., Ait Tchakoucht, T., & Saheed, Y. K. (2023). A systematic literature review for network intrusion detection system (IDS). International journal of information security, 22(5), 1125-1162.

[2] Heidari, A., & Jabraeil Jamali, M. A. (2023). Internet of Things intrusion detection systems: a comprehensive review and future directions. Cluster Computing, 26(6), 3753-3780.

[3] Lansky, J., Ali, S., Mohammadi, M., Majeed, M. K., Karim, S. H. T., Rashidi, S., ... & Rahmani, A. M. (2021). Deep learning-based intrusion detection systems: a systematic review. IEEE Access, 9, 101574-101599.

[4] Ozkan-Okay, M., Samet, R., Aslan, Ö., & Gupta, D. (2021). A comprehensive systematic literature review on intrusion detection systems. IEEE Access, 9, 157727-157760.

[5] Ashiku, L., & Dagli, C. (2021). Network intrusion detection system using deep learning. Procedia Computer Science, 185, 239-247.

[6] Hnamte, V., & Hussain, J. (2023). DCNNBiLSTM: An efficient hybrid deep learning-based intrusion detection system. Telematics and Informatics Reports, 10, 100053.

[7] Qazi, E. U. H., Faheem, M. H., & Zia, T. (2023). HDLNIDS: hybrid deep-learning-based network intrusion detection system. Applied Sciences, 13(8), 4921.

[8] Wu, Z., Zhang, H., Wang, P., & Sun, Z. (2022). RTIDS: A robust transformer-based approach for intrusion detection system. IEEE Access, 10, 64375-64387.

[9] Kasongo, S. M. (2023). A deep learning technique for intrusion detection system using a Recurrent Neural Networks based framework. Computer Communications, 199, 113-125.

[10] Altunay, H. C., & Albayrak, Z. (2023). A hybrid CNN+ LSTM-based intrusion detection system for industrial IoT networks. Engineering Science and Technology, an International Journal, 38, 101322.

[11] Halbouni, A., Gunawan, T. S., Habaebi, M. H., Halbouni, M., Kartiwi, M., & Ahmad, R. (2022). CNN-LSTM: hybrid deep neural network for network intrusion detection system. IEEE Access, 10, 99837-99849.

[12] Halim, Z., Yousaf, M. N., Waqas, M., Sulaiman, M., Abbas, G., Hussain, M., ... & Hanif, M. (2021). An effective genetic algorithm-based feature selection method for intrusion detection systems. Computers & Security, 110, 102448.

[13] Musleh, D., Alotaibi, M., Alhaidari, F., Rahman, A., & Mohammad, R. M. (2023). Intrusion detection system using feature extraction with machine learning algorithms in IoT. Journal of Sensor and Actuator Networks, 12(2), 29.

[14] Alzahrani, A. O., & Alenazi, M. J. (2021). Designing a network intrusion detection system based on machine learning for software defined networks. Future Internet, 13(5), 111.

[15] Bertoli, G. D. C., Júnior, L. A. P., Saotome, O., Dos Santos, A. L., Verri, F. A. N., Marcondes, C. A. C., ... & De Oliveira, J. M. P. (2021). An end-to-end framework for machine learning-based network intrusion detection system. IEEE access, 9, 106790-106805.

[16] Hossain, M. A., & Islam, M. S. (2023). Ensuring network security with a robust intrusion detection system using ensemble-based machine learning. Array, 19, 100306.

[17] Awad, N. A. (2021). Enhancing network intrusion detection model using machine learning algorithms. Computers, Materials & Continua, 67(1), 979-990.

[18] Azam, Z., Islam, M. M., & Huda, M. N. (2023). Comparative analysis of intrusion detection systems and machine learning-based model analysis through decision tree. IEEE Access, 11, 80348-80391.