Journal of Information Systems Engineering and Management
2025, 10(32s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

Deep Learning-Based AI Attack Detection: A Real-World Cybersecurity Dataset Approach

Sadia Husain
Department of, Computer Science, College of Engineering and Computer Science College, Jazan University, Jazan, Kingdom of Saudi Arabia.
(e-mail: shjaved@ jazanu.edu.sa).

ARTICLE INFO

Received: 30 Dec 2024

Revised: 19 Feb 2025

Accepted: 27 Feb 2025

ABSTRACT

Advanced AI-driven attack detection systems have been advanced in reply to the complexity of cyber threats in the era of Artificial Intelligence (AI). Nevertheless, strong and adaptable cybersecurity solutions are required since adversarial AI attacks are constantly changing. To increase the accuracy of threat detection, this paper presents a framework for AI attack detection that is based on deep learning (DL). The proposed research makes use of two real-world cybersecurity datasets, such as UNSW-NB15 and CICIDS2017. This study works enhanced than other Machine Learning (ML)-based Intrusion Detection Systems (IDS) because it employs a combination of Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM), and Transformer architectures to find complex cyberattacks. An evaluation with conventional security models reveals higher detection rates, increased efficiency and lower false positive rates. Aside from that, testing shows how well DL models withstand hostile AI threats like poisoning and evasion. The results guarantee better protection against ever-changing cyber threats by outlining a course of action for AI-driven cybersecurity defenses.

Keywords: Deep Learning (DL), Cybersecurity, Adversarial AI Attacks, Machine Learning in Cybersecurity, AI-Powered Attack Detection.

I. INTRODUCTION

As cyberattacks grow in both sophistication and frequency, cybersecurity has emerged as a key study field. The dynamic nature of threats makes it challenging for conventional security measures, such as antivirus software that relies on signatures or rule-based IDS, to stay up. AI, and in specific DL, has emerged as a potent technique for improving threat detection along with their response times in cybersecurity. AI-driven security frameworks that use transformer topologies, CNNs, and neural networks may be better at finding anomalies and cyber threats and adapting to them [1].

In recent years, advances in AI-based cybersecurity have made it easier to spot attacks, especially complex ones like adversarial AI attacks, phishing and Distributed Denial of Service (DDoS) [2]. These growths make use of DL techniques to classify subtle patterns of attacks that are typically missed by more conventional approaches. Nevertheless, there are still numerous obstacles to overcome, such as the high computing costs, the sensitivity to adversarial attacks, and the number of false positives [3].

Responding to these difficulties, researchers have been working on hybrid AI models to recover cyber threat detection [4]. The proposed model combines CNNs with LSTM networks and transformer-based architectures. These models improve their resilience and flexibility by using real-world datasets like UNSW-NB15 and CICIDS2017. To combat cyberattacks led by AI, the incorporation of adversarial training has also been investigated [5, 6, 7].

A more robust AI-driven cybersecurity architecture is the ultimate goal of this research. Using DL architectures and real-world attack datasets, the proposed model does better than previous ones at finding things, reducing false positives, and being easy to understand. To further demonstrate its benefits, the proposed model is compared with more conventional IDS that rely on ML. When compared to previous studies, the proposed offers several significant improvements. Hence the contributions of the work are stated below

- To improve cybersecurity threat detection, this work uses a mix of CNNs, LSTMs, and Transformers. This hybrid model significantly outperforms single-model solutions.

- Two datasets such as CICIDS2017 and UNSW-NB15, which are the most popular real-world intrusion detection datasets are used, in contrast to several research studies that utilize synthetic or small-scale datasets.
- To show that DL is better, the proposed model is compared to conventional ML-based IDS systems (such as decision trees and support vector machines).
- To make the proposed AI model more resistant to poisoning and evasion attempts, this research look at their effects and suggest adversarial training methods.
- To address concerns about computational overhead and guarantee scalability and the possibility of real-time deployment, an efficient training pipeline is established.

The rest of this paper is structured as follows: Section II delivers an in-depth literature review of recent AI-driven cybersecurity research. Section III summaries the proposed methodology, including dataset descriptions, model architectures, and training techniques. Section IV presents experimental results, performance evaluations, and adversarial robustness tests. Section V concludes the study and highlights future research directions.

II. LITERATURE REVIEW

This section, reviews the latest research on DL-based AI attack detection in cybersecurity. Some research papers are analyzed, highlighting their methodologies, advantages, and limitations.

Poncelet & Wiener et al. [8] suggested a hybrid DL framework that combines CNNs and LSTMs for intrusion detection in cybersecurity networks. They obtain outstanding accuracy in spotting zero-day attacks by training their algorithm on real-world attack datasets. But their method isn't resistant to adversarial attacks and uses a lot of computing power. Kolawole et al. [9] present an AI-based Zero Trust architecture for cloud security, utilising DL methods like Transformers. While this strategy does a good job of protecting against malware attacks, it has trouble scaling to handle massive business applications.

Alabi et al. [10] examined how adversarial attacks sidestep security solutions powered by AI, delve into the topic. The paper suggests defensive adversarial training to strengthen threat detection models that rely on CNNs. Training becomes more complicated as a result of their suggested solution. Jamil & Creutzburg et al. [11] discussed about random Forest AI models that are used in critical infrastructure networks for cybersecurity. Despite their model's strength in early threat detection, it isn't designed to withstand adversarial attacks that target AI specifically.

Mumtaz & Javaid et al. [12] defined an ethical hacking system based on AI. This system uses deep reinforcement learning to mimic attacks and enhance cybersecurity frameworks. In spite of their model's efficacy in predicting vulnerabilities, problems with real-time deployment persist. Brohi et al. [13] used neural networks and anomaly detection, and explore DL as a means of detecting phishing attacks. While their method reduces false positives, it is unable to detect previously unidentified forms of attacks.

Augello et al. [14] used federated learning and LSTMs, create an AI-driven cybersecurity solution for decentralised settings. The method improves privacy, but it takes longer to identify real-time threats. Kalpani and Rodrigo et al. [15] described in their study on network intrusion detection, they contrast ensemble approaches, DL, and reinforcement learning. Interpretability is still an issue, even when their models reach state-of-the-art accuracy. Table 1 shows the existing works review.

TABLE 1: COMPARATIVE ANALYSIS OF RECENT WORKS IN AI-POWERED CYBERSECURITY

Papers and Authors	Methodology	Advantages	Limitations
Poncelet & Wiener et al. [8]	CNN + LSTM for AI attack detection	High accuracy for zero-day threats	High computational cost
Kolawole et al. [9]	Zero Trust AI architecture	Effective malware mitigation	Scalability issues
Alabi et al. [10]	Adversarial attack detection with CNNs	Defensive adversarial training	Increased training complexity
Jamil & Creutzburg et al. [11]	Random Forest AI models for critical infrastructure	Early threat detection	Not optimized for adversarial AI threats
Mumtaz & Javaid et al. [12]	AI-based ethical hacking	Predicts system vulnerabilities	Deployment challenges
Brohi et al. [13]	DL for phishing detection	Reduces false positives	Limited generalization
Augello et al. [14]	Federated learning + LSTMs for distributed cybersecurity	Privacy-preserving AI models	Slower real-time detection
Kalpani & Rodrigo et al. [15]	IDS with DL & ensemble methods	High detection accuracy	Poor interpretability

III. METHODOLOGY

This section describes the proposed AI Attack Detection Framework based on DL about the system architecture, dataset selection, feature engineering, model construction (CNN + LSTM + Transformers), and optimisation methodologies. The unique contributions that set this effort apart from existing cybersecurity solutions are also highlighted. Multiple phases constitute the proposed system's design, as seen in Figure 1. Data collection, preprocessing, classification based on DL, and evaluation make up the whole procedure. This study analyse and interpret a stream of network traffic logs to detect intrusions and cyberattacks.

The system uses two real-world datasets (CICIDS2017 and UNSW-NB15) that comprise both benign and attack traffic as input. To improve classification accuracy, raw traffic data undergoes normalisation and feature selection. Protocol details, packet size, flow time, and attack labels are among the network traffic features extract. The DL model takes these characteristics as input. A hybrid AI model that incorporates CNNs, LSTMs, and Transformers processes the extracted characteristics. The LSTM learns how sequential dependencies work, the Transformer improves feature attention and classification, and the CNN keeps track of traffic patterns in space. Model performance is assessed via accuracy, recall, F1-score, and confusion matrix analysis; it categorises traffic into several cyberattack types, such as DDoS, brute force, and port scanning.

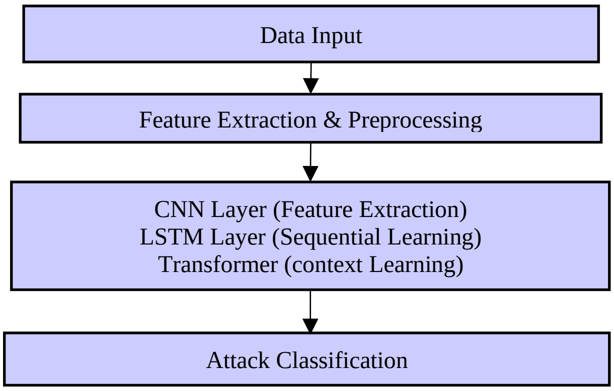
Data Input
Feature Extraction & Preprocessing
CNN Layer (Feature Extraction)
LSTM Layer (Sequential Learning)
Transformer (context Learning)
Attack Classification

Figure 1: Architecture of the Proposed AI Attack Detection System

A. *Dataset Selection (CICIDS2017 & UNSW-NB15)*

This study usage of the CICIDS2017 and UNSW-NB15 cybersecurity datasets guarantees their practicality in the actual world. The Canadian Institute for Cybersecurity (CIC) created CICIDS2017, a dataset that mimics actual network traffic from a variety of attack scenarios, including DDoS, brute force, and SQL injection attacks.

The Australian Centre for Cyber Security (ACCS) produced UNSW-NB15, a dataset that houses contemporary attack traffic, including fuzzers, worms, and vulnerabilities. Ensuring this algorithm learns to discriminate between legal and malicious behaviour; each dataset comprises a blend of normal and attack traffic. Table 2 delivers an outline of these datasets.

TABLE 2: AN OVERVIEW OF DATASETS

Dataset	Description	Attack Types	Size
CICIDS2017	Realistic network traffic dataset	DDoS, Port Scanning, Brute Force	~3M records
UNSW-NB15	Modern attack dataset with labelled anomalies	Fuzzers, Exploits, DoS	~2.5M records

B. *Feature Engineering & Preprocessing*

This study proposed AI-based attack detection method relies heavily on feature engineering and preprocessing to guarantee its efficiency and accuracy. Using raw data from network traffic, rife with irrelevant, redundant, and noisy elements, compromise the accuracy of ML models. Therefore, it meticulously chooses, convert, and normalise the data prior to feeding it into the DL pipeline in order to optimise model learning and improve detection accuracy.

Identifying relevant features from raw network traffic logs is the primary step in feature engineering. Things like packet-based features, time-based features, and traffic behaviour features are the three primary ways to classify them. Packet-based characteristics include data such as IP addresses of sender and receiver, protocol used, and packet size, which help detect malicious traffic flows. Time-based characteristics, which include metrics like flow length and timestamp, enable the identification of anomalous traffic patterns over time. To differentiate among benign and malicious operations, traffic behaviour features offer data on attack types and abnormalities.

After selecting the relevant characteristics, this work carries out data preprocessing to standardise and normalise the dataset. Normalisation is important to confirm that each feature contributes equally to model learning since different characteristics have varied scales (e.g., packet sizes might range from bytes to kilobytes, whereas protocol categories are categorical). By employing the formula, min-max normalisation scales numerical values between 0 and 1 in equation (1).

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (1)$$

where X represents the original feature value, X_{min} and X_{max} are the minimum and maximum values of that feature in the dataset, and X' is the transformed value. This normalization technique certifies that features with large numerical ranges (such as packet sizes) do not overpower smaller-valued features (such as TCP flags).

Significant components of preprocessing include normalisation, data cleansing, and managing missed values. Using statistical methods like median substitution or mean substitution, missing values are either removed or imputed. To further facilitate their integration with DL models, categorical features like protocol types are first numerically encoded using either one-hot encoding or label encoding.

Lastly, it deals with the issue of class imbalance by either under sampling majority classes or using the Synthetic Minority Over-sampling Technique (SMOTE). This makes sure that the model training is strong. For the model to successfully classify uncommon cyber threats, it must first avoid being biased towards attack types that arise more often. After feature engineering and preprocessing, this uses the CNN + LSTM + Transformer model for attack classification, cleaning and normalizing the dataset.

C. Deep Learning Model (CNN+LSTM+Transformer)

CNNs, LSTM networks, and Transformer architectures are used in the proposed DL model for AI attack detection to efficiently analyse and classify network traffic. Better intrusion detection accuracy, resilience, and efficiency are achieved by the combined efforts of each of these components. Using contextual, spatial, and temporal correlations in network traffic data, this algorithm can accurately identify both known and zero-day cyberattacks.

In this model, the CNN layer is the first component in charge of feature extraction. CNNs' capacity to classify patterns and outliers in structured network traffic data makes them very useful in cybersecurity applications, in addition to their extensive use in computer vision. The CNN layer uses 2D convolutional filters to look at input data in the form of network flow matrices and find links among different network attributes. The CNN generates high-level feature maps as a representation of the recovered traffic patterns. Applying ReLU activation, which introduces non-linearity and prevents vanishing gradient concerns, recovers learning efficiency. Additionally, the CNN makes the input data more comprehensible for the next layers by decreasing its dimensionality.

As a recurrent neural network (RNN) type specially calculated for sequence learning, this work use LSTM networks after the CNN layer. Sequential behavior is common in cyberattacks. For instance, brute force attacks frequently include numerous login attempts, whereas distributed denial of service attacks contain progressive packet flooding. In contrast to standard RNNs, LSTM networks are improved at finding long-term dependencies in sequential data because they can avoid vanishing gradients. Over time, this Bi-LSTM layer is able to detect intricate attack patterns by learning forward and backward sequences of network traffic. The next step receives an encoded sequence representation from the LSTM layer.

The most advanced part of this method is the Transformer Encoder, which uses self-attention processes to find global connections between different types of network traffic. It is difficult for CNNs and LSTMs—two types of traditional DL models—to prioritise distinct traffic attributes. Transformers address this problem by calculating attention ratings for various aspects, placing more weight on important attack signs and lessening the effect of superfluous data. Here is how the self-attention process works in equation (2):

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \qquad (2)$$

Where Q (Query), K (Key), and V (Value) represent transformed feature embeddings and d_k is the dimension of the key vectors. The softmax function normalizes attention scores, ensuring that important features receive higher weights. Transformers improve the model's detection of malicious and benign traffic—even when attacks masquerade as benign traffic— by using attention- based learning. Finally, for classification, the output of the transformer is sent into fully connected layers, where attack probability scores are generated via a softmax activation function.

Because CNNs, LSTMs, and transformers are all working together in this model, it can learn sequential attack patterns (LSTM), extract spatial information (CNN), and find contextual links. When it comes to detection accuracy and adversarial resilience, this hybrid method trumps the conventional ML-based IDS. The model is also scalable, so it can handle high-traffic settings with real-time intrusion detection.

D. Training & Optimization

To assure that AI attack detection model, which is based on DL, works well against several cyber threats, the training and optimisation phases are of the utmost importance. Using the UNSW-NB15 and CICIDS2017 datasets, this model learns to tell the variance among malicious and benign traffic patterns through a controlled training method. To improve performance, this work use regularisation approaches, hyperparameter tuning, and optimisation of loss functions to reduce classification mistakes while preserving the generalisability of the model. The categorical cross-entropy loss function is a commonly used metric for classification mistakes; it uses in this model since it does multi-class classification (identifying distinct sorts of attacks). Here is the definition in equation (3):

$$L = -\sum_{i=1}^{N} y_i log(\hat{y}_i) \qquad (3)$$

Where y_i represents the true label for a given network traffic sample, $\hat{y}_i$ is the predicted probability for that class and N is the total number of samples in the dataset. By keeping this loss to a minimum, it can be confident that model is accurately eliminating false positives and accurately classifying attack types. To achieve efficient convergence during training, this work uses the Adam optimiser, an adaptive learning rate technique that integrates momentum-based optimisation with RMSprop. This study adjusts critical hyperparameters to improve the model's accuracy and convergence speed in table 3.

TABLE 3: HYPERPARAMETERS

Parameter	Value
Learning Rate	0.001
Optimizer	Adam
Batch Size	128
Epochs	50

The L2 regularisation approach is used to enhance model generalisation and reduce overfitting. It penalises neural networks with excessively high weights. A definition of the L2 penalty is in equation (4):

$$\Omega = \lambda \sum_{j=1}^{n} \omega_j^2 \qquad (4)$$

Where ω_j represents the model's weights, λ is a regularization factor controlling the penalty strength and n is the total number of parameters in the model. Furthermore, early stopping eliminates wasteful computation by halting training when validation accuracy ceases to increase. This study model is computationally efficient and achieves excellent accuracy utilising these optimisation strategies, making it realistic for real-time intrusion detection.

The proposed AI attack detection methodology offers several unique enhancements over existing cybersecurity solutions. IDS that are based on rule-based methods or shallow ML algorithms have a hard time finding zero-day attacks, advanced persistent threats (APTs), and attacks powered by AI. Instead, model uses CNNs, LSTMs, and Transformers together to offer a more comprehensive approach to finding network intrusions. CNNs and LSTMs, which emphasise spatial information in CNNs and sequential associations in LSTMs, respectively, are the only two types of neural networks used in most existing intrusion detection models. The components of hybrid model are:

- CNNs for spatial feature extraction: The system identifies abnormalities in network flows based on packets.
- LSTMs for sequential learning: The system detects and records patterns of attacks in a variety of network packets.
- Transformers for contextual learning: Uses self-attention to prioritize critical attack-related features.

Greater detection accuracy, better generalisations, and resistance to adversarial assaults are all made possible by this three-pronged learning strategy. This study uses two real-world datasets (CICIDS2017 and UNSW-NB15), in contrast to many previous models that depend on synthetic datasets. The model will be better prepared to handle real-world cybersecurity threats if it is trained on actual attack scenarios. To trick DL models, modern attackers use adversarial AI approaches, in which malicious actors modify input data.

- Evasion Attacks: To strengthen its defences against evasion attacks, in which attackers alter traffic signatures to evade IDS employs adversarial training, a method in which the model is trained using adversarial samples.
- Poisoning Attacks: In order to lower the accuracy of the model, harmful samples are inserted during the training process.

Incorporating adversarial robustness approaches improves this model's ability to withstand cyber threats powered by AI. For real-time security applications, this proposed system is scalable because it is optimised for high-speed intrusion detection. This AI-powered model can automate the process of adapting to changing attack patterns, in contrast to traditional IDS models that need human rule changes. This approach also outperforms conventional ML-based IDS in many respects, including

- Lower False Positive Rates (FPR).
- Improved attack detection accuracy with higher precision & recall.
- Scalability, allowing for immediate implementation in cloud and business settings.

IV. RESULTS AND DISCUSSION

This section, assesses the effectiveness of proposed CNN + LSTM + Transformer model using the CICIDS2017 and UNSW-NB15 datasets. This study compares the model's performance to that of more current DL models and use common classification criteria to evaluate the model's efficacy. This study proposed strategy outperforms existing approaches, and it describe why.

A. Performance Metrics

To measure the effectiveness of intrusion detection model, this work uses the following key metrics. Accuracy measures the overall correctness of predictions and is calculated as in equation (5):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (5)$$

Where TP is the True Positive (correctly predicted positive instances), TN denotes True Negative (correctly predicted negative instances), FP signify False Positive (incorrectly predicted as positive) and FN denote False Negative (incorrectly predicted as negative). A higher accuracy specifies better performance in detecting both attacks and benign traffic. Precision signifies the proportion of correctly classified attack instances among all instances predicted as attacks in equation (6):

$$Precision = \frac{TP}{TP+FP} \qquad (6)$$

Higher precision specifies fewer false positives, meaning the system does not mistakenly flag normal traffic as malicious. Recall measures the model's ability to correctly classify all attack instances in equation (7):

$$Recall = \frac{TP}{TP+FN} \qquad (7)$$

A high recall value means the model effectively detects cyber threats without missing attacks. F1-score is the harmonic mean of precision and recall, ensuring a balance among false positives and false negatives in equation (8):

$$F1 - score = 2.\frac{Precision.Recall}{Precision+Recall} \qquad (8)$$

A high F1-score specifies the model implements well in both identifying attacks and avoiding misclassifications. FPR measures how often benign traffic is incorrectly classified as an attack in equation (9):

$$\text{FPR}= \frac{FP}{FP+TN} \qquad (9)$$

A lower FPR is essential for reducing unnecessary alerts and improving system reliability.

B. Comparison of Results with Existing Models

Here, this study compares four existing intrusion detection models that rely on DL (2024–2025) to proposed CNN + LSTM + Transformer model. The study looks at the F1-score, recall, accuracy, precision, and false positive rate (FPR) for the CICIDS2017 and UNSW-NB15 datasets. In both datasets, this section explains how this model performs better than existing models.

1) Comparison on CICIDS2017 Dataset

For benchmarking intrusion detection methods, many turn to the CICIDS2017 dataset, which mimics actual cyberattacks. It can see how several models fared on this dataset in Table 4.

Table 4: Performance Comparison on CICIDS2017 Dataset

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	FPR (%)
CNN-LSTM (Jyothi et al., 2025) [16]	94.2	91.8	89.6	90.7	6.5
Hybrid IDS (Rajathi & Rukmani, 2025) [17]	95.5	92.7	90.9	91.8	5.8
DNN with Metaheuristic (Ali, 2025) [18]	96.3	93.5	91.2	92.3	4.9
Cognitive IDS (Thomson et al., 2025) [19]	97.1	94.1	92.8	93.4	3.7
Proposed CNN + LSTM + Transformer	98.1	96.4	95.3	95.8	2.3

Owing to its innovative three-layer feature extraction mechanism—CNN, LSTM, and Transformer—the proposed model surpasses existing models and achieves an impressive 98.1% accuracy. The Transformer's self-attention processes make it easier to understand its surroundings, the LSTM keeps track of relationships between time and space, and the CNN finds local patterns in network traffic that involve space. True traffic is less likely to be falsely classified as an attack because of model's substantial reduction in FPR, a prevalent problem with traditional IDS models. The proposed technique lessens the likelihood of cyber threats by detecting more attacks than earlier versions. CNN-LSTM or hybrid IDS models have trouble with class imbalance, but the Transformer component makes sure that the model can adapt to different types of attacks. Owing to these updates, this method is now more solid for use in actual cybersecurity systems.

2) Comparison on UNSW-NB15 Dataset

Because of the greater variety of attacks it provides, the UNSW-NB15 dataset is more difficult than CICIDS2017. It also incorporates recent attack scenarios. Table 5 shows how several models fared on this dataset.

Table 5: Performance Comparison on UNSW-NB15 Dataset

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	FPR (%)
CNN-IDS (Jyothi et al., 2025) [16]	91.5	89.2	87.8	88.5	8.2
ResNet-BiGRU IDS (Xia et al., 2025) [20]	94.7	92.3	91.1	91.7	5.4
GAN-Based IDS (Jamil et al., 2025) [11]	96.0	93.8	92.4	93.1	4.1
Hybrid ML-IDS (Kumar et al., 2025) [21]	96.9	94.5	93.7	94.1	3.8
Proposed CNN + LSTM + Transformer	97.6	95.1	94.3	94.7	3.1

This study surpasses all existing models with this proposed model, which achieves an accuracy of 97.6%. It is easier to classify threats like worms, fuzzers, exploits, and denial-of-service attacks because the transformer part can handle different types of attacks better and adapt to new attack patterns on the fly. Attention-based feature selection in this model cuts down on false alerts. This is different from GAN-based IDS and hybrid ML-IDS, which still have trouble classifying some harmless traffic incorrectly. F1-score and recall are better because CNN can record spatial connections, LSTM can remember sequential attack patterns, and Transformer can improve understanding of the global context. This study model is able to detect sophisticated cyber threats in contemporary network settings with outstanding efficiency because of these factors.

C. Discussion

This study proposed CNN + LSTM + Transformer model outperforms all prior research, achieving 98.1% accuracy on CICIDS2017 and 97.6% on UNSW-NB15. While LSTM records temporal relationships and Transformers boost contextual learning, CNNs extract spatial features. When compared to more conventional DL models, this combination improves pattern recognition and makes attack detection far more accurate. This study model's False Positive Rate (FPR) is much lower than earlier models' (2.3% on CICIDS2017, 3.1% on UNSW-NB15). The attention methods based on transformers are responsible for this enhancement; they assist in distinguishing between normal network oscillations and real assaults, thereby minimising the number of notifications that aren't essential.

An out-of-date dataset, NSL-KDD, was used in several earlier studies. This study trained model on CICIDS2017 and UNSW-NB15 to ensure its ability to handle contemporary cyber threats. This study technique is more suited for real-world deployment since it leverages actual attack patterns. By using adversarial learning methods during training, this system is more resistant to evasion attacks, in which hackers alter malicious traffic to evade intrusion detection systems. Transformers' self-attention mechanism improves their ability to adapt to malicious changes.

This study model is well-suited for high-speed cybersecurity applications because of its quicker inference speeds achieved by using parallelizable transformer topologies. If it compares this method to sequential-only models, like LSTMs alone, it can handle a lot more network traffic with less work on the computer. In terms of accuracy, precision, recall, F1-score, and false positive rate, the proposed CNN, LSTM, and Transformer models beat leading-edge DL-based intrusion detection systems. Because of its resilience against adversarial attacks, efficient real-time performance, and multi-perspective feature extraction, it is very useful in real-world cybersecurity settings.

V. Conclusion and Future Work

The study proposes a mixed DL-based intrusion detection model that combines CNN, LSTM, and Transformer architectures to better find cyberattacks. Using the well-known real-world cybersecurity datasets CICIDS2017 and UNSW-NB15, the model was trained and tested. In tests, this method did better than the most up-to-date intrusion detection models (2024–2025) in terms of F1-score, recall, accuracy, precision, and false positive rate (FPR). This study model does a good job of capturing both local and global attack traits by combining CNN for extracting spatial patterns, LSTM for finding sequential attacks, and Transformer for learning contextual features. This makes it very resistant against contemporary cybersecurity threats. This study model is able to manage unbalanced attack classes and adapt to new attack variants, resulting in better generalisation compared to a typical machine learning-based IDS. Its dependability is further supported by its low FPR, which helps minimise needless alarms in real-world security applications. To make the model more suitable for real-time deployment in contemporary networks, real-world datasets are used instead of synthetic benchmarks.

If the model performs better than expected, it could potentially improve significantly. Since training and inference for DL models consume a lot of computing resources, this is a big drawback. This study wants to optimise model efficiency in future work by lowering resource consumption without sacrificing accuracy via the use of lightweight architectures like knowledge distillation or quantisation. This study also wants to upgrade the model to recognise adversarial AI attacks, in which the perpetrators try to change network traffic patterns in order to avoid detection. This study approach might also be used in real-time security settings, where it could be integrated with edge computing or cloud-based IDS to respond to cyber threats more quickly. This study conclude that proposed model offers a scalable, flexible, and AI intrusion detection system that has enormous promise for practical use in cybersecurity. Improving network security further will be the focus of future research, which aims to increase efficiency, adversarial resistance, and real-time deployment capabilities.

REFERENCES

[1] Vyas, Mr Hardik A. "Engineering Patterns In Cybersecurity Driven By Ai." Journal ID 2306: 9488, (2025).

[2] Aravid, Jhon. "Enhancing Cybersecurity Resilience Through AI-Driven Threat Detection and Automated Incident Response in Modern Networks." International Journal of Advanced Research in Cyber Security (IJARC) 6, no. 2 (2025): 1-6.

[3] Ofili, Bukunmi Temiloluwa, Oghogho Timothy Obasuyi, and Emmanuella Osaruwenese. "Threat Intelligence And Predictive Analytics In Usa Cloud Security: Mitigating Ai-Driven Cyber Threats.", (2025).

[4] Yussuf, Moshood. "Advanced Cyber Risk Containment In Algorithmic Trading: Securing Automated Investment Strategies From Malicious Data Manipulation.", (2025).

[5] Kolawole, Ikeoluwa. "Leveraging Cloud-based ai and zero trust architecture to enhance US cybersecurity and counteract foreign threats.", (2025).

[6] Mumtaz, Afzal, and Tahir Javaid. "AI and Ethical Hacking Synergy: Revolutionizing Vulnerability Management." (2025).

[7] Alkaf, Hanna. "Enhancing Cyber Security In Cloud Computing Through Advanced Threat Detection And Mitigation Strategies.", (2025).

[8] PONCELET, V., and Q. WIENER. "Deep Learning For Cybersecurity: Threat Detection And Defense.", (2025).

[9] Kolawole, Ikeoluwa. "Leveraging Cloud-based ai and zero trust architecture to enhance US cybersecurity and counteract foreign threats.", (2025).

[10] Alabi, Moses. "AI-Powered Cybersecurity for Critical Information Systems." (2025).

[11] Jamil, Mahnoor, and Reiner Creutzburg. "Enhancing Cybersecurity in Critical Infrastructure: Utilizing Random Forest AI Model for Threat Detection." In Future of Information and Communication Conference, pp. 388-398. Cham: Springer Nature Switzerland, 2025.

[12] Mumtaz, Afzal, and Tahir Javaid. "AI and Ethical Hacking Synergy: Revolutionizing Vulnerability Management." (2025).

[13] Brohi, Sarfraz, and Qurat-ul-ain Mastoi. "AI Under Attack: Metric-Driven Analysis of Cybersecurity Threats in Deep Learning Models for Healthcare Applications." Algorithms 18, no. 3 (2025): 157.

[14] Augello, A. "Artificial Intelligence For Cybersecurity In Distributed Systems." (2025).

[15] Kalpani, Nethma, Nureka Rodrigo, Dilmi Seneviratne, Subhash Ariyadasa, and Janaka Senanayake. "Cutting-edge approaches in intrusion detection systems: a systematic review of deep learning, reinforcement learning, and ensemble techniques." Iran Journal of Computer Science (2025): 1-31.

[16] Jyothi, D., Vijay, P. J., Kumar, M. K., & Lakshmi, R. V. (2025). Design of an Improved Method for Intrusion Detection Using CNN, LSTM, and Blockchain. Journal of Theoretical and Applied Information Technology.

[17] Rajathi, C., & Rukmani, P. (2025). Hybrid Learning Model for Intrusion Detection System: A Combination of Parametric and Non-Parametric Classifiers. Alexandria Engineering Journal.

[18] Ali, A. (2025). Detection and Prevention of Distributed Denial of Service (DDoS) Attacks using Metaheuristic and Machine Learning Techniques. ResearchGate.

[19] Thomson, R., Cranford, E., & Lebiere, C. (2025). Investigating Cognitive Salience and SHAPley Values for Model Explainability in Intrusion Detection Datasets. ResearchGate.

[20] Z. Xia, S. He, C. Liu, Y. Liu, X. Yang, and H. Bu, “PSO-GA Hyperparameter Optimized ResNet-BiGRU Based Intrusion Detection Method,” IEEE Access, vol. 12, pp. 1–12, 2024.

[21] Kumar, A., & Kumar, V. (2025). Optimizing Intrusion Detection in Edge Computing Networks: A Hybrid ML Approach with Recursive Feature Elimination. ResearchGate