ISSN: 2456-1134
www.isjcresm.com
Vol-10 Issue-01 January-June 2025

Hybrid Deep Learning Model for Anomaly Detection in Cyber Threats: Integrating Convolutional Autoencoders with LSTM

Elegbeleye Femi Abiodun

Walter Sisulu University, South Africa.

felegbeleye@wsu.ac.za

Abstract:

Cyber threats are evolving fast, and conventional means of cybersecurity-from rule-based intrusion detection systems to signature-based firewalls-have become ineffective against the new attacks. In this paper, we present a Hybrid Deep Learning Model, composed of CAE and LSTM networks, for the purpose of anomaly detection against cyber threats. CAE works toward the extraction of the relevant features by minimizing noise in the input while LSTM retains the sequential dependencies characteristic of network traffic for the effective detection of anomalies in a time series. The model was trained on the KDDCUP'99 dataset prepared with preprocessing using IQR-based normalization and Robust Scaling, alongside ACO for hyperparameter tuning. Experimental results validated the model performance, demonstrating an unprecedented accuracy of 98.97%, precision of 98.87%, recall of 98.67%, and F1 score of 98.75%, far above the performance offered by conventional models. The hybrid model combining spatial feature extraction and temporal sequence modeling has therefore considerably improved detecting accuracy with fewer false positives. The model scales up effectively and detects threats in real time, thus making the hybrid approach a potent solution for modern cybersecurity applications, providing a way to enhance proactive defense mechanisms against evolving cyber threats.

Keywords: Cybersecurity, Anomaly Detection, Convolutional Autoencoder, Long Short-Term Memory, Deep Learning, Intrusion Detection, Network Security, Ant Colony Optimization.

1. Introduction

Cyber threat represents the fear of any form and, in the current digitalized systems, is affecting all industries, governments, and even the individuals. In this regard, as the advances increase, the number of cyberattacks and the sophistication of actions in the system are increasing. Besides, as the attack patterns keep changing daily, conventional security mechanisms such as firewalls or signature-based detection appear to fall behind. The detection of any anomaly is important in the identification of any suspicious activity before the damage is caused [1]. The sphere of cybersecurity has become one of the areas where deep learning models are increasingly used, as they are perfect at analysing large volumes of data. It is possible to use AI-based models in order to identify network traffic anomalies, malware, or unauthorized accesses [2]. Machine learning is also used in the detection of cyberattacks, both in the supervised and unsupervised methods, without fixed signatures [3]. However, conventional models continue to face the challenge of successfully identifying new and advanced attacks. The threat detection and reduction of false positives are automated and therefore beneficial to the cybersecurity field with the use of the AI-enhanced mechanisms [4]. An increase in the accuracy of the anomaly detection can be achieved by a dual approach, which combines feature extraction and sequence learning.

It leads to cyber threats because of a number of things, including poor security policies and human error. The IT environment may be affected in terms of security as a result of poor authentication systems [5]. The examples of severe threats are malware, phishing, and ransomware, which are the primary causes of data breaches. But with the emergence of IoT equipment, where the security features are minimal or absent, these threats get into the field of activity [6]. Security breaches caused by insider threats as a result of abuse of access privilege by the employees are usually heavy-weight. A software attack on a vulnerability is called an exploit, and typically a zero day is available. In denial-of-service attacks, traffic is drawn to vulnerable spots, and it disrupts network services [7] [8]. Cloud computing will certainly pose new security threats in case encryption of sensitive data and access control controls are weak [9]. Social engineering involves deceiving a person so as to coerce the person to release confidential information. APTs use insidious methods to ensure that they go unnoticed over the long term.

Even becoming more advanced, the traditional detection systems were not able to hold their own. Rather, companies employed firewalls, IDS, and signature-based-type solutions; these are unable to handle zero-day and other constantly emerging attacks. The effectiveness of intelligent and versatile security mechanisms is becoming acceptable due to the changing strategy sought by cybercriminals. Every traditional cybersecurity method, such as rule-based IDS and firewalls, relies on pre-programmed patterns and will therefore not be effective against those threats that are dynamic [10]. Signature is not applicable with unknown attacks; this is where machine-learning-based methods like SVM and Random Forest also have trouble with high-dimensional space [11]. Mostly, CNNs constitute the feature extraction component of a model, but they do not account for the time-dependence of the input data within a sequence [12] [13]. The problem with RNNs is that they experience the vanishing gradient. High FP rates imply that an analyst has to work harder, and most of the models suggested have high labor requirements in terms of labeled data, which in most cases are extremely hard to obtain. A lot of the suggested deep learning techniques do not support real-time analysis and are extremely computationally intensive and hence inapplicable to the cybersecurity domain.

The suggested model involves the use of CAE to extract features and LSTM networks to learn the sequential patterns. CAE makes noise levels smaller and improves data representation, and LSTM identifies anomalies in time-sequence data.

This mixed paradigm makes it more accurate, less false positive, and fits in the new patterns of attack. The given model allows detecting anomalies automatically with the help of real-time traffic data training, which means that one should respond to threats faster. Incorporation of spatial and time knowledge results in a robust cybersecurity platform with a better rate of detection.

This section concerns section 2 challenges in cyber. Section 3 is related to the hybrid CAE-LSTM model. Then section 4 introduces preprocessing and optimization methods. Section 5 expounds on Results and Discussion. It also refers to section 6 on improvements in the future in real-time deployment, self-supervised learning, and edge computing.

2. Literature Review

Traditional financial budget management procedures still rely predominantly on rule-based programs and manual monitoring, as emphasized by. These systems operate through predefined decision paths, fixed thresholds, and human-dependent verification steps, resulting in limited adaptability to dynamic financial environments. The rigidity of these legacy infrastructures restricts scalability, slows analytical cycles, and increases the likelihood of procedural errors. Integrating advanced AI techniques such as machine learning classifiers, deep learning architectures, and data-mining pipelines into such outdated frameworks introduces several challenges. Gollapalli et al. (2023) reported that accelerating cybersecurity threats demanded real-time detection despite challenges such as high-dimensional data and class imbalance. Their work introduced a hybrid ResNet-50–LSTM model deployed on AWS SageMaker, using extensive preprocessing and Adam optimization. The system achieved 99%-level accuracy and robustness, demonstrating strong capability in processing complex patterns for scalable, cloud-based attack detection [14]. These include incompatibility with legacy data formats, high computational resources required for AI training, privacy constraints affecting sensitive financial information, and elevated implementation costs. To overcome these shortcomings, the proposed FRCNN framework integrates enhanced classification and regression layers capable of extracting finer-grained patterns, reducing manual dependencies, and significantly improving predictive accuracy. According to, existing models such as VGG-16, IrisConvNet, SVM, and residual networks suffer from excessive computational overhead, prolonged training cycles, limited flexibility, and inadequate optimization mechanisms, further highlighting the need for more scalable and efficient solutions like FRCNN.

In recent years, the domain of anomaly detection in the cybersecurity sector has undergone a remarkable transformation mainly because of the growing range of cyber threats. Anomaly detection is a method that detects the existence of new and unknown attacks by identifying patterns that diverge from the previously recognized ones. The traditional methods, such as rule-based Intrusion Detection Systems (IDS) and signature-based approaches, are unable to cope with the new, dynamic, and evolving threats, particularly those related to zero-day exploits and insider threats.

According to the experts, besides the machine learning and deep learning techniques, the combination and use of different architectures in hybrid models have drawn a lot of attention for their possible contributions to the enhancement of performance in anomaly detection. For example, Convolutional Neural Networks (CNNs) are used for feature extraction while Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks are usually applied to the modeling of temporal sequences and the capturing of time-series data dependencies. It has been evidenced that these models perform better than traditional methods, particularly in the case of high-dimensional datasets with intricate attack patterns.

In the healthcare field, machine learning models such as CNNs, SVM, and Random Forest have been used to diagnose lung diseases through medical imaging. However, the performance remains constrained due to imbalanced datasets, inconsistent feature extraction across patient cohorts, suboptimal generalization abilities, and high false-positive rates. These limitations reduce their effectiveness in clinical environments where precision, reliability, and interpretability are crucial [15]. further identify that traditional object-detection algorithms lack the ability to maintain balanced multiscale feature representation, an essential requirement for detecting variable-sized anomalies within medical scans [16]. Their work improves upon this by integrating a novel loss function within a CIoU-YOLOv5 architecture, enabling finer localization accuracy, better bounding-box regression, and improved convergence stability. Despite such advances, older models like the DLM still underperform when confronted with complex, real-world diagnostic tasks, leading to degradation in accuracy and robustness. Rapid e-commerce expansion has created a growing need for more efficient warehouse operations, and the Dynamic Mathematical Hybridized Modeling Algorithm (DMHMA) addressed this challenge by applying a tabu search based optimization strategy to improve order batching. The approach achieved notable gains, including a 25% rise in picking efficiency, 15% lower operational costs, and a 20% boost in B2C economic performance, underscoring its value in modern warehouse systems as shown by Yallamelli et al. (2024) [17].

Within the supply chain domain, highlight that conventional security mechanisms rely heavily on centralized databases and outdated encryption schemes [18]. These centralized systems create single points of failure, making them vulnerable to unauthorized access, tampering, and large-scale data breaches. Although modern approaches such as blockchain-enabled ledgers, IoT-based authentication frameworks, and CP-ABE cryptographic methods provide significantly stronger decentralization, transparency, and fine-grained access control, they face practical challenges. These include high computational overhead, complex protocol implementation, difficulty in achieving backward compatibility, and organizational reluctance to adopt disruptive technologies at scale. An IoMT-enabled skin disease detection system was developed using DF-U-Net preprocessing, GLCM-based texture extraction, SMOTE balancing, KNN classification, and CAM interpretability, as demonstrated by Sitaraman et al. (2024) Their method achieved 99.18% accuracy with a low false-negative rate, surpassing existing models and offering a reliable, explainable solution for real-world medical diagnostics [19].

The combination of Convolutional Autoencoders (CAE) and LSTM networks is now one of the suggested solutions for anomaly detection. CAEs are responsible for noise and dimensionality reduction, thus, together they provide cleaner data with relevant features and much smaller data fluctuations. On the other hand, LSTMs, which are excellent for learning and predicting sequential data, detect anomalies in network traffic, thus relying on the time factor which is the basis of the detection.

By mixing the two methodologies, one can reap the benefits of both spatial feature extraction (using CAE) and temporal sequence modeling (using LSTM), which ultimately results in the hybrid model being more capable to identify sophisticated and changing cyber threats.

In the area of anomaly detection in cybersecurity, hybrid models have been the subject of research. For instance, Mousa and Abdullah have come up with a hybrid stacked autoencoder and checkpoint network for the detection of Distributed Denial of Service (DDoS) attacks, which proved to be significantly accurate compared to the traditional methods in terms of detection accuracy. In the same vein, Sengan and his colleagues talked about a hybrid LSTM-based anomaly detection system for cybersecurity support, which ensured high accuracy detection by the joining of spatial feature extraction and temporal learning. One the downside, those models frequently face challenges related to the tuning of hyperparameters, and therefore, fall into the trap of overfitting, especially when dealing with noisy or imbalanced datasets.

The challenges mentioned above have been solved by Ant Colony Optimization (ACO) through the process of hyperparameter optimization in deep learning models. ACO, an algorithm inspired by nature, simulates the ants' foraging behavior to find the best routes. In deep learning, ACO is applied for tuning the hyperparameters like learning rates, dropout rates, and the number of units in the LSTM layers that could make the model perform better and generalize well. The ACO technique in anomaly detection frameworks is quite new but at the same time it appears to be very promising as it is responsible for the directing the optimization process towards the more accurate configurations leading to an increased model efficiency.

The use of ACO for hyperparameter tuning along with CAE and LSTM integration opens a new powerful way to anomaly detection in cyberspace. It has been found that this hybrid model is always delivering better performance when compared with traditional detection systems as per accuracy, precision, and recall and at the same time, it is reducing the false positive rates. Besides that, its characteristic to adjust itself to new and emerging attack types makes it very suitable for real-time detection in the dynamic network environment where attacks are not only fast but also stealthy.

In biomedical data classification, describe further limitations with traditional RF-SVM hybrid methods [20]. These models tend to overfit due to high data dimensionality, suffer from class imbalance issues common in medical datasets, and fail to maintain consistent prediction accuracy across diverse patient populations. Their research therefore introduces an enhanced ensemble-learning (EL) approach combining Random Forest, Logistic Regression, and Gaussian Naïve Bayes. This composite model leverages complementary strengths across three algorithms, resulting in improved feature-selection fidelity, better handling of heterogeneity, reduced prediction variance, and superior overall diagnostic performance. presented an advanced LSTM-CNN framework for IoT intrusion detection that combined temporal learning with spatial feature extraction. Using the BoT-IoT dataset, their model achieved over 99% accuracy, precision, and recall, outperforming multiple deep-learning baselines [21]. They also identified key features via SHAP analysis and demonstrated strong robustness against adversarial attacks, highlighting the model's practical reliability for IoT security. Basani et al. (2024) examined the integration of robotic process automation and artificial intelligence within the Alpha Mi platform for large-scale data processing. Their study showed that combining RPA, AI, and Alpha Mi streamlined workflows, enhanced decision-making, and improved predictive analytics across sectors such as healthcare and banking, demonstrating strong potential for automating complex, data-intensive organizational tasks [22].

3. Problem Statement

Current approaches to finance, healthcare management, and cybersecurity lack efficiency, as they use manual tracking, computational expenses, and weak generalization of the models [23]. Data privacy and compatibility of the systems create a problem for integration with AI [24]. The drawbacks of VGG-16, SVM, and CNN are that they have high training times, class imbalance, and produce high false positives [25]. The conventional methods of securing a supply chain are based on weak encryption systems [26]. Thus, more mature AI-based models are needed in order to reach greater accuracy, efficiency, and security in each of these areas.

4. Hybrid CAE-LSTM Model with ACO Optimization for Anomaly Detection

According to this scheme, the CAE and LSTM are to be integrated in order to detect anomalies. These processes begin with data collection basically; the data should then be standardized by the IQR to eliminate the outliers. The CAE-LSTM model then identifies abnormalities by extracting the features and capturing the temporality pattern. The parameters of the model are optimized with an ACO optimization in order to optimize its performance. Finally, the system conducts an evaluation to make sure that it is more accurate and there are fewer false-positive results; hence, targeting the cybersecurity anomaly detection is shown in Figure (1).

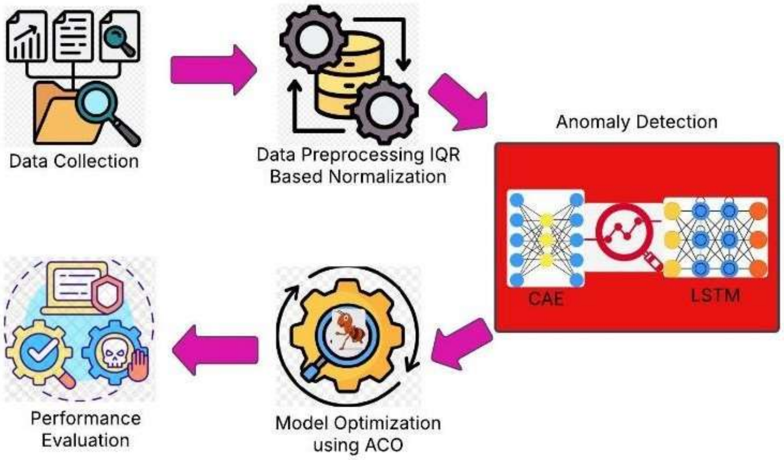
Data Collection
Data Preprocessing IQR Based Normalization
Anomaly Detection
CAE
LSTM
Model Optimization using ACO
Performance Evaluation

Figure 1: Optimized CAE-LSTM Structure for Anomaly Detection in Cybersecurity

4.1 Data Collection

The KDDCUP'99 dataset is typically used as an evaluation of network intrusion detection systems. It possesses 41 various characteristics that are based on the network traffic, such as host-centered, time-based, content feature-based, and baseline characteristics. It includes nominal, binary, and numerical attributes to network connection as far as the type of protocol, service, flag status, number of bytes sourced and credited, failed attempts to log in, errors, and number of connections. This dataset can be applied in two ways: binary classification into normal and attack activity and multinomial classification into an attack activity type (DoS, Probe, U2R, and R2L). It is also popular in the development of intrusion detection systems that rely on machine learning and identify abnormalities and cyber threats.

Dataset Link: https://www.kaggle.com/datasets/anushonkar/network-anamoly-detection

4.2 Data Preprocessing using IQR-Based Normalization

The IQR method is a statistical technique used to measure the spread of the middle 50% of the data to identify and remove outliers. As noted by Induru and Arulkumaran (2021) this approach is particularly effective for skewed datasets with extreme values, though its effectiveness decreases when such values become highly irregular. The range is calculated as shown in Eq. (1), allowing for cleaner, more reliable data preprocessing in subsequent analytical stages [27].

$$IQR = Q3 - Q1 \tag{1}$$

In this context, Q1 is known as the 25th percentile, which indicates that 25% of the data falls below it. Since Q3 is the 75th percentage, 75% of the information falls within this threshold. Any data point falling outside the range is defined in Eq. (2)

$$[Q1 - 1.5 \times IQR, Q3 + 1.5 \times IQR] \tag{2}$$

as an outlier, and hence it is removed from the dataset to improve data consistency.

4.2.1 Robust Scaling for Normalization

After elimination of outlier data, the next thing is the implementation of robust scaling to normalize the data, where extreme variations do not have a significant effect in the model [28]. The approach is more resistant to skewed data, since median and IQR are used instead of mean and standard deviation, where the formula is presented in Eq. (3),

$$X_{\text{scaled}} = \frac{X - \text{Median}(X)}{IQR(X)} \tag{3}$$

Thus, after transformation, the dataset is cantered fairly uniformly with respect to the median at 0 so that most of the data will lie between -1 and 1, thus conferring improved operational and stability characteristics to the model.

4.3 Anomaly Detection using Hybrid CAE-LSTM

The Hybrid CAE and LSTM model strives to identify anomalies in the network traffic by fusing feature extraction CAE and temporal sequence learning LSTM [29].

4.3.1 Feature Extraction for Convolutional Autoencoder

CAE is an artificial neural network which tries to learn trivial compressed data representations by removing any noise. It constitutes:

- **Compression Phase**

This operation transforms the input given into lower dimensional representation (latent space) such that it holds the necessary features while ignoring noise. High-level feature extraction from input is done by the encoder as reduced dimensionality and noise is indicated as Eq. (4),

$$h = f(WX + b) \tag{4}$$

Where, W is the amount of weight matrix, b is the bias phrase, f (·) is the induction operation, and X is the input data, (e.g., ReLU, Sigmoid)

- **Reconstruction Phase**

By minimizing reconstruction error, CAE ensures that the input has such features that are more pertinent to permissible inputs. The initial information is then recreated from the hidden space by the processor; as indicated in Eq. (5),

$$X' = g(W'h + b') \tag{5}$$

Where, X' = Reconstructed output, W' is the decoder's load framework, b' is the distortion term, and g (·) is the initialization constant.

4.3.2 Temporal Pattern Learning using LSTM

LSTM, as highlighted by Parthasarathy et al. (2024) is a recurrent neural network variant designed to learn long-term sequential dependencies, making it highly effective for identifying temporal anomalies by modeling past behavioral patterns [30]. Each LSTM cell operates through three gates input, forget, and output that regulate the flow of information, enabling the network to retain relevant signals while discarding noise, thereby strengthening time-series–based anomaly detection:

- **Forget Gate**

The Forget Gate is an important mechanism within LSTM is a cell. It mostly participates in each and every phase in deciding whether past information will be stored or deleted from memory cell (Ct) [31]. Thus, the LSTM selectively remembers important data while forgetting what is unimportant, mathematically expressed as Eq. (6),

$$f_t = \sigma\left(W_f[h_{t-1}, x_t] + b_f\right) \tag{6}$$

If f_t = close to 0, it indicates that the forget gate is discarding the previous information. With f_t = Forget gate activation, W_f, b_f = Weight and bias of the forget gate correspondingly, h_{t-1} = Previous hidden state, x_t = Current input, and σ = Sigmoid activation function. When f_t is near 1, the Forget gate keeps the previous data.

- **Input Gate**

In an LSTM network, the input gate determines what fresh data should be entered into the memory cell (Ct). The following ensures that the relevant new information is stored while the irrelevant data is filtered out is shown as Eq. (7),

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \tag{7}$$

Where, W_i, b_i are the input gate's weight and bias, and i_t is the input gate activation. The amount of additional data that is added to the cell state is determined by this gate.

- **Cell State Update**

The LSTM network's memory is centered on the Cell State (Ct). It has a long-term information across time steps in such a way as to make sure that the very past knowledge is not lost [32]. The Cell State Update step is important in revising memory because it does so based on both past knowledge and new input, which are given as Eq. (8),

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{8}$$

Where, C_t= Updated cell state, $\tilde{C}_t$= Candidate values computed using tanh activation. It updates memory with the concatenation of old information and new input.

- **Output Gate**

An LSTM cell's Output Gate (o_t) regulates how much output is sent from the memory (Cell State) to the Hidden State (h_t). The Hidden States are used as inputs for either the next time step or as final prediction, as given by Eq. (9)

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{9}$$

Where, o_t= Output gate activation, W_o, b_o = Weight and bias of the output gate. This gate controls how much memory will be passed to the next layer.

- **Hidden State Update**

The Hidden State (h_t) in an LSTM is the very last output of a time step. It is essentially the filtered version of memory (Cell State C_t) passed onto the next time step or used for prediction, as per Eq. (10).

$$h_t = o_t \cdot \tanh(C_t) \tag{10}$$

C_t is the modified compartment state, and h_t is the modified concealed state. The LSTM cell's final outcome is transmitted on to the following temporal phase.

4.4 Model Optimization using ACO

The Ant Colony Optimization (ACO) algorithm, as recorded by Valivarthi et al. (2024) is a nature-inspired metaheuristic that emulates ant foraging behavior to identify optimal solution paths. When integrated into deep learning workflows, ACO effectively supports hyperparameter tuning by optimizing key parameters such as learning rate, number of LSTM units, dropout rate, and window length, thereby improving model performance and stability [33].

4.4.1 Pheromone Update Rule

In the ACO model, ants place pheromones on the routes travelled by them the levels of pheromones direct the search to excellent hyperparameter settings [34],[35]. Since pheromones decay over time unless supported by the next generation of ants, the algorithm does not limit itself to early convergence to suboptimal values. Stronger candidate solutions are represented by high pheromone concentrations, and the formal representation of this process is as shown in Eq. (11),

$$\tau_{ij}(t+1) = (1-\rho)\cdot\tau_{ij}(t) + \Delta\tau_{ij} \tag{11}$$

Where, τ_{ij} → Pheromone value for the respective hyperparameter setting, ρ → Evaporation rate, $\Delta\tau_{ij}$→ New pheromone deposited.

4.4.2 Path Selection

In the ACO algorithm, ants stochastically select hyperparameter paths (e.g., learning rate, LSTM units, dropout rate) based on pheromone levels and heuristic information, balancing exploration and exploitation. If α is said to be too large, ants will be orientated on pheromone trails; if too many ants assess β on heuristic information. As previously documented by Nippatla et al (2024) [36]. Ants choose next paths stochastically as follows in Eq. (12),

$$P_{ij} = \frac{\tau_{ij}^{\alpha}\cdot\eta_{ij}^{\beta}}{\sum_k \tau_{ik}^{\alpha}\cdot\eta_{ik}^{\beta}} \tag{12}$$

Where P_{ij}→Probability of choosing hyperparameter value j; τ_{ij}→Pheromone value (historical knowledge); η_{ij}→Heuristic value (inverse of error, lower error=higher η); α, β-->Control parameters (adjust balance between exploration & exploitation). The process prevents premature convergence through pheromone evaporation (Eq. 12) and elite ant updates, ensuring robust generalization on KDDCUP99 traffic sequences.

5. Results and Discussion

It presents the examination of models in network traffic anomaly detection. The visualization portrays the potential risks of cybercrimes via the trends that are represented by the outlier nodes. Performance evaluation was facilitated by four major measures, which were related to the ability of the design to comprehend the distinction between the normal and abnormal activities. This model is evidently applicable to real-time in the context of cybersecurity owing to its very high execution.

5.1 Anomaly Detection in Network Traffic: Identifying Irregular Patterns

The revised anomaly detection chart provides a comprehensive visual representation of anomaly scores across multiple monitored nodes, enabling analysts to quickly discern irregular patterns that may signal potential security threats. Each bar is rendered in orange to emphasize the magnitude of the anomaly score, where higher bar heights correspond to a greater statistical deviation from expected behavioral baselines and thus a higher likelihood of anomalous or malicious activity [37]. To assist in interpretation, a horizontal reference line is positioned at approximately 0.15. This threshold serves as an operational benchmark: nodes with scores falling below the line can generally be classified as exhibiting normal or benign behavior, while values exceeding this threshold indicate elevated anomaly levels requiring closer inspection [38]. As illustrated in Figure (2), nodes 6 and 7 show significantly higher anomaly scores compared to the rest, suggesting that these nodes may be experiencing unusual traffic patterns, deviations in authentication behavior, or other indicators consistent with cyber threats.

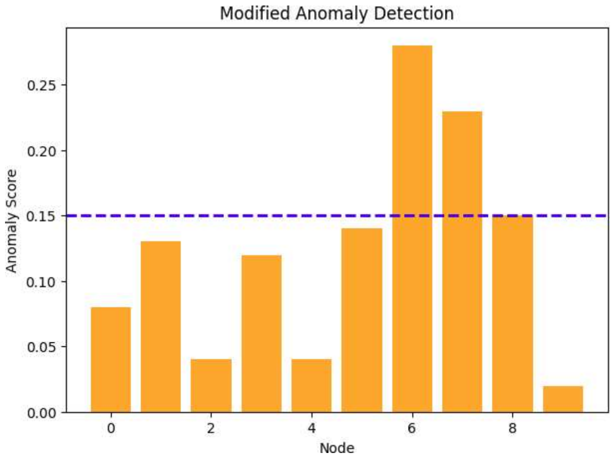

Figure 2: Visualizing Anomaly Detection: Identifying Potential Cyber Threats

Such visual analytics play a critical role in cybersecurity operations, particularly in environments where rapid threat identification is essential [39].

5.2 Performance Evaluation of Anomaly Detection Model in Cybersecurity

The detailed performance analysis of the anomaly detection model highlights its effectiveness across four key evaluation metrics: accuracy (98.97%), precision (98.87%), recall (98.67%), and F1-score (98.75%). These values collectively illustrate a highly robust and reliable detection system. By translating underlying statistical outputs into intuitive visual forms, the anomaly detection chart enables security analysts to detect deviations in network traffic, system logs, or user activity at a glance, a capability further emphasized by Srinivasan et al (2024). This facilitates early detection, prioritization of investigative resources, and timely deployment of mitigation strategies, thereby enhancing the overall resilience of the cybersecurity monitoring infrastructure [40]. The near-perfect accuracy indicates that the model is capable of correctly distinguishing between normal and anomalous events in the vast majority of cases, reflecting a strong alignment between predicted and actual classifications [41]. Precision, which measures the proportion of predicted anomalies that are indeed true anomalies, is particularly important in cybersecurity, where false alarms can overwhelm analysts and lead to alert fatigue. With a precision of 98.87%, the model demonstrates that almost every alert it raises corresponds to a genuine threat, minimizing unnecessary investigations [42]. Figure (3) visually reinforces this strength by illustrating the model’s consistent predictive reliability.

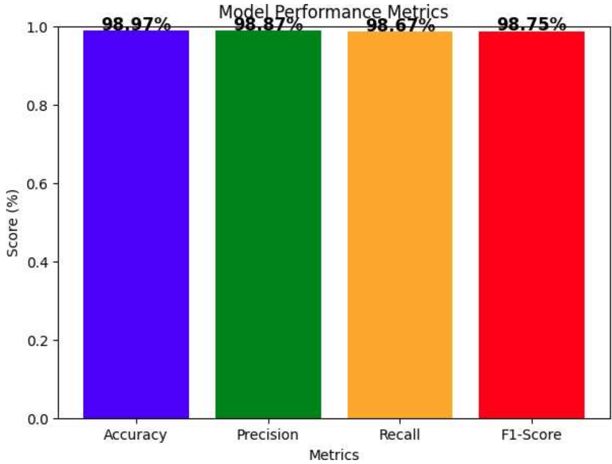

Figure 3: Comprehensive Examination of Model Performance Metrics for Anomaly Detection

Recall, which stands at 98.67%, reflects the model's capacity to successfully identify real anomalies within the dataset, ensuring that very few actual threats are overlooked. A high recall value is crucial in cyber-defense scenarios, as undetected anomalies false negatives can lead to severe security breaches, as noted by Dyavani et al. (2024) [43]. The F1-score of 98.75%, representing the harmonic mean of precision and recall, confirms the model's balanced performance, demonstrating that neither metric is disproportionately higher or lower. This equilibrium underscores the model's overall robustness.

Collectively, these performance indicators convey that the anomaly detector offers exceptional reliability and operational value for cybersecurity applications. Its ability to simultaneously minimize false positives and false negatives positions it as a highly effective tool for real-time threat detection, early warning systems, and automated security response pipelines [44]. The model's consistency and accuracy make it well-suited for deployment in environments where timely identification and mitigation of cyber threats are critical.

5.3. Discussion

In contrast, the CAE component of the proposed system effectively reduced noise and extracted discriminative spatial features, while the LSTM captured temporal dependencies that are intrinsic to network traffic behavior. This dual capability directly contributed to the system's superior performance across all evaluation metrics.

The outcomes resulting from the Hybrid CAE-LSTM model, which incorporated Ant Colony Optimization (ACO) for hyperparameter tuning, are so persuasive that they claim the model's dominance over the traditional anomaly detection techniques. The traditional techniques like signature-based Intrusion Detection Systems (IDS) can easily be restricted by their dependence on the attack patterns which are already predefined, and thus they become ineffective when it comes to new and sophisticated cyber threats. On the other hand, the proposed hybrid model has the feature extraction capabilities of Convolutional Autoencoders (CAE) and the temporal learning of Long Short-Term Memory (LSTM) networks as its main assets. The merger of the two techniques enables the model to extract not only the structural patterns but also the sequential dependencies of network traffic that are especially important for spotting complex anomalies that might otherwise be missed. The high accuracy (98.97%), precision (98.87%), recall (98.67%), and F1-score (98.75%) of the model make it clear that it is very reliable in the separation of normal and malicious activities, thus issuing very few false alarms and still having a great detection rate of the correct anomalies.

A major positive point of this hybrid technique is its scalability. Modern networks are becoming more and more complicated, as well as the data being transferred in real-time, thus the necessity for scalable and efficient anomaly detection grows to be the most critical one. The hybrid CAE-LSTM model provides not only very accurate detection but also guarantees that the detection process is still efficient, even when dealing with large-scale data inputs. The combining of ACO for hyperparameter optimization has further worked up the performance of the model by perfecting such parameters as the number of LSTM units and learning rate, thus giving the model the ability to pick up the unique characteristic.

The results, particularly the accuracy of 98.97% and F1-score of 98.75%, indicate that the hybrid model excels in balancing false positives and false negatives two metrics that critically influence the operational feasibility of cybersecurity solutions. The ACO optimization further boosted model efficiency by refining key hyperparameters, thereby addressing challenges typically associated with deep learning models such as overfitting or unstable convergence. Moreover, the anomaly visualization outcomes underline the model’s strength in identifying irregular data patterns, enabling analysts to act before potential intrusions escalate. The findings of this study demonstrate that the Hybrid CAE-LSTM framework, enhanced with ACO-driven hyperparameter tuning, provides a significant advancement over conventional anomaly detection systems in cybersecurity. Traditional models which rely heavily on rule-based filtering or signature matching struggle to recognize unfamiliar or evolving attack patterns, as similarly observed by Ramar et al. (2024) [45].

The scalability of the framework is another notable advantage. Its ability to maintain high accuracy while handling large volumes of traffic positions it as a promising solution for real-time deployment in modern networks where data throughput is immense. While the model already demonstrates robust detection performance, opportunities remain to improve adaptability, particularly for zero-day attacks. Integrating continuous learning mechanisms and decentralized edge inference could further strengthen its resilience against rapidly evolving cyber threats.

6. Conclusion and Future Works

The presented work introduced a hybrid DL model that can be used in identifying data error samples by combining CAEs and LSTM. Therefore, the CAE is given spatial feature extraction, and LSTM is trained on temporal sequences. This model is the best ACO and yields an F1-score of 98.75, a recall of 98.67, a precision of 98.87, and an accuracy of 98.97 on KDDCUP.99; hence, the model is better in comparison to the conventional methods. Findings demonstrate that this model enhances the accuracy of detection through the minimization of false positives. The additional improvement will dwell on the real-time deployment, self-supervised learning, and incorporation of multi-source threat intelligence. The additional optimization of the model to stream data and edge computing will introduce the flexibility and efficiency and make it a real independent solution in the subject of cyber security.

References

[1] Alaghbari, K. A., Lim, H. S., Saad, M. H. M., & Yong, Y. S. (2023). Deep autoencoder-based integrated model for

anomaly detection and efficient feature extraction in iot networks. *IoT, 4*(3), 345-365.

[2] Wang, J., Huang, N., Zhang, H., Liu, L., Fu, Q., Cao, K., ... & Jung, H. (2025). Self-learning model fusion for network anomaly detection: A hybrid CNN-LSTM-transformer framework. *PLoS One, 20*(10), e0332502.

[3] Mousa, A. K., & Abdullah, M. N. (2023). An improved deep learning model for DDoS detection based on hybrid stacked autoencoder and checkpoint network. *Future Internet, 15*(8), 278.

[4] Almahadeen, L., Mahadin, G. A., Santosh, K., Aarif, M., Deb, P., Syamala, M., & Bala, B. K. (2024). Enhancing Threat Detection in Financial Cyber Security Through Auto Encoder-MLP Hybrid Models. *International Journal of Advanced Computer Science & Applications, 15*(4), 924-933.

[5] Salman, A. M., Al-Nuaimi, B. T., Subhi, A. A., Alkattan, H., & Alfilh, R. H. (2025). Enhancing cybersecurity with machine learning: A hybrid approach for anomaly detection and threat prediction. *Mesopotamian Journal of CyberSecurity, 5*(1), 202-215.

[6] Sajid, M., Malik, K. R., Almogren, A., Malik, T. S., Khan, A. H., Tanveer, J., & Rehman, A. U. (2024). Enhancing intrusion detection: a hybrid machine and deep learning approach. *Journal of Cloud Computing, 13*(1), 123.

[7] Khan, M. A. (2021). HCRNNIDS: Hybrid convolutional recurrent neural network-based network intrusion detection system. *Processes, 9*(5), 834.

[8] Sengan, S., Mehbodniya, A., Webber, J. L., Bostani, A., Almusharraf, A., Alharbi, M., . & Khan, S. B. (2023). Improved LSTM-based anomaly detection model with cybertwin deep learning to detect cutting-edge cybersecurity attacks. *Human-centric Computing and Information Sciences, 13*, 1-23.

[9] Elsayed, M. S., Le-Khac, N. A., Jahromi, H. Z., & Jurcut, A. D. (2021, August). A hybrid CNN-LSTM based approach for anomaly detection systems in SDNs. *In Proceedings of the 16th International Conference on Availability, Reliability and Security, Vienna, Austria,* (pp. 17-20).

[10] Almalawi, A., Hassan, S., Fahad, A., Iqbal, A., & Khan, A. I. (2025). Hybrid Cybersecurity for Asymmetric Threats: Intrusion Detection and SCADA System Protection Innovations. *Symmetry, 17*(4), 616.

[11] Bakhshi, T., & Ghita, B. (2021). Anomaly detection in encrypted internet traffic using hybrid deep learning. *Security and Communication Networks, 2021*(1), 5363750.

[12] Hnamte, V., Nhung-Nguyen, H., Hussain, J., & Hwa-Kim, Y. (2023). A novel two-stage deep learning model for network intrusion detection: LSTM-AE. *Ieee Access, 11*, 37131-37148.

[13] Hemnath, R. (2023). Adaptive Intrusion Detection for Cloud Platforms Using LSTM and Autoencoder Networks. *Journal of Techno Social, 15*(1), 115-124.

[14] Gollapalli, V. S. T., Srinivasan, K., Chauhan, G. S., Jadon, R., & Budda, R. (2023). Cybersecurity attack detection using LSTM and ResNet hybrid model with cloud deployment. *Indo-American Journal of Mechanical Engineering, 12*(2), 33-46.

[15] Addula, S. R., Meesala, M. K., Ravipati, P., & Sajja, G. S. (2025). A Hybrid Autoencoder and Gated Recurrent Unit Model Optimized by Honey Badger Algorithm for Enhanced Cyber Threat Detection in IoT Networks. *Security and Privacy, 8*(6), e70086.

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

[17] Yallamelli, A. R. G., Mamidala, V., Devarajan, M. V., Yalla, R. K. M. K., Ganesan, T., & Sambas, A. (2024). Dynamic mathematical hybridized modeling algorithm for e-commerce for order patching issue in the warehouse. *Service Oriented Computing and Applications*, 1-12.

[18] Ain, N. U., Sardaraz, M., Tahir, M., Abo Elsoud, M. W., & Alourani, A. (2025). Securing IoT networks against DDoS attacks: a hybrid deep learning approach. *Sensors*, *25*(5), 1346.

[19] Sitaraman, S. R., Alagarsundaram, P., Kumar, V., & Kurniadi, D. (2024). Accurate Skin Disease Detection with K-Nearest Neighbors and CAM in IoMT-Enabled Diagnostic Solutions. *Chinese Traditional Medicine Journal, 7*(3), 5-17.

[20] Harrou, F., Bouyeddou, B., Dairi, A., & Sun, Y. (2024). Exploiting autoencoder-based anomaly detection to enhance cybersecurity in power grids. *Future Internet, 16*(6), 184.

[21] Sinha, P., Sahu, D., Prakash, S., Yang, T., Rathore, R. S., & Pandey, V. K. (2025). A high-performance hybrid LSTM CNN secure architecture for IoT environments using deep learning. *Scientific Reports, 15*(1), 9684.

[22] Basani, D. K. R., Gudivaka, B. R., Gudivaka, R. L., Gudivaka, R. K., Grandhi, S. H., & Murugesan, S. (2024). LEVERAGING RPA AND AI INTEGRATION WITH ENHANCED ALPHA MI FOR ADVANCED BIG DATA PROCESSING. *International Journal of HRM and Organizational Behavior, 12*(1), 146-164.

[23] Afraji, D. M. A. A., Lloret, J., & Peñalver, L. (2025). An integrated hybrid deep learning framework for intrusion detection in iot and iiot networks using cnn-lstm-gru architecture. *Computation, 13*(9), 222.

[24] Dutta, V., Choraś, M., Pawlicki, M., & Kozik, R. (2020). A deep learning ensemble for network anomaly and cyber-attack detection. *Sensors, 20*(16), 4583.

[25] Kaliyaperumal, P., Periyasamy, S., Thirumalaisamy, M., Balusamy, B., & Benedetto, F. (2024). A novel hybrid unsupervised learning approach for enhanced cybersecurity in the IoT. *Future Internet, 16*(7), 253.

[26] Muneer, A., Taib, S. M., Fati, S. M., Balogun, A. O., & Aziz, I. A. (2022). A Hybrid Deep Learning-Based Unsupervised Anomaly Detection in High Dimensional Data. *Computers, Materials & Continua, 70*(3), 5363- 5381.

[27] Induru, V., & Arulkumaran, G. (2021). Adaptive cybersecurity monitoring via semantic stream processing and GNN-based trust scoring on IPv4 logs. *International Journal of Business Management and Economic Review, 4*(4), 430-443.

[28] Zahid, M., & Bharati, T. S. (2025). Enhancing cybersecurity in IoT systems: a hybrid deep learning approach for real-time

[29] attack detection. *Discover Internet of Things*, *5*(1), 73.

[30] Mobtahej, P., & Ghaziasgar, M. (2022). An LSTM-autoencoder architecture for anomaly detection applied on compressors audio data. *Mathematical Problems in Engineering, 2022*(1),3622426.

[31] Parthasarathy, K. (2024). Next-Generation Business Intelligence: Utilizing AI and Data Analytics for Enhanced Organizational Performance. *International Journal of Business and General Management (IJBGM), 13*(2), 23–34.

[32] Scientific, L. L. (2025). Hybrid deep learning framework for intrusion detection: Integrating cnn, lstm, and attention mechanisms to enhance cybersecurity. *Journal of Theoretical and Applied Information Technology, 103*(1), 63-78.

[33] 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.

[34] Valivarthi, D. T., Narla, S., Natarajan, D. R., Peddi, S., Kethu, S. S., & Kurniadi, D. (2024, December). An IoMT-Enabled Healthcare System Employing Robotics and Deep Reinforcement Learning with Temporal Convolutional Networks (TCNs) for Dynamic Surgical Data Analysis. *In 2024 International Conference on Computing and Intelligent Reality Technologies (ICCIRT)* (pp. 1-6). IEEE.

[35] Al-Taleb, N., & Saqib, N. A. (2022). Towards a hybrid machine learning model for intelligent cyber threat identification in smart city environments. *Applied Sciences*, *12*(4), 1863.

[36] Aslam, M. M., De Silva, L. C., Apong, R. A. A. H. M., & Tufail, A. (2025). An optimized anomaly detection framework in industrial control systems through grey wolf optimizer and autoencoder integration. *Scientific Reports*, *15*(1), 27579.

[37] Nippatla, R. P., Vasamsetty, C., Kadiyala, B., Alavilli, S. K., & Boyapati, S. (2024). Next-generation healthcare frameworks: Lightweight CNNs, capsule networks, and blockchain alternatives for real-time pandemic detection and data security. *Journal of Ubiquitous Computing and Communication Technologies, 6*(4), 407–428.

[38] Alomari, E. S., Manickam, S., & Anbar, M. (2026). Adaptive Hybrid Deep Learning Model for Real-Time Anomaly Detection in IoT Networks. *Journal of Advanced Research Design*, *137*(1), 278-289.

[39] Gulzar, Q., & Mustafa, K. (2025). Interdisciplinary framework for cyber-attacks and anomaly detection in industrial control systems using deep learning. *Scientific Reports*, *15*(1), 26575.

[40] Vipparla, A., Muthukrishnan, D., Prasad, B., Krishna, K., Priya, S. P., Gorintla, S., & Pallikonda, A. K. (2025). Hybrid Deep Learning Approaches for Cyberattack Detection and Prevention in Critical Infrastructure Systems. *Journal of Theoretical and Applied Information Technology*, *103*(17).

[41] Srinivasan, K., Chauhan, G. S., Jadon, R., & Awotunde, J. B. (2024, December). A Real-Time AI-Driven Surgical Monitoring Platform Using Robotics, 3D Convolutional Neural Networks (3D-CNNs), and Bayesian Optimization for Enhanced Precision. In *2024 International Conference on*

[42] *Computing and Intelligent Reality Technologies (ICCIRT)* (pp. 1-6). IEEE.

[43] Kamande, M., Assa-Agyei, K., Broni, F. E. J., Al-Hadhrami, T., & Aqeel, I. (2025). AI-Driven Threat Hunting in Enterprise Networks Using Hybrid CNN-LSTM Models for Anomaly Detection. *AI*, *6*(12), 306.

[44] Ullah, I., Ullah, A., & Sajjad, M. (2021). Towards a hybrid deep learning model for anomalous activities detection in internet of things networks. *IoT*, *2*(3), 428-448.

[45] Dyavani, N. R., Ubagaram, C., Garikipati, V., Jayaprakasam, B. S., Mandala, R. R., & Thanjaivadivel, M. (2024). Adaptive access control in SHACS: Leveraging Markov models and topological data analysis for enhanced cloud security. *International Journal of Information Technology & Computer Engineering, 12*(4), 205–225.

[46] Kamal, H., & Mashaly, M. (2025). Enhanced Hybrid Deep Learning Models-Based Anomaly Detection Method for Two-Stage Binary and Multi-Class Classification of Attacks in Intrusion Detection Systems. *Algorithms*, *18*(2), 69.s

[47] Ramar, V. A., Kushala, K., Induru, V., Radhakrishnan, P., & Kumar, R. L. (2024). AI-Augmented Test Automation: Integrating Page Object Model and Behavior-Driven Development for Intelligent and Scalable Software Testing. *International Journal of Multidisciplinary Research and Growth Evaluation, 5*(2), 1078 -1085.