International Journal of Science and Engineering Science Research
Volume 1 | Issue 2 | April-June 2025 | ISSN 3049-2793

Cyberattack Prevention in Cloud-Based Systems Using Butterfly Optimization Algorithm BOA Based CNN for Intrusion Detection

[1]Narsing Rao Dyavani
Uber Technologies Inc, California, USA
nrd3010@gmail.com

[2]Venkat Garikipati
Innosoft, Maryland, USA
venkat44557@gmail.com

[3]Charles Ubagaram
Tata Consultancy Services, Ohio, USA
charlesubagaram17@gmail.com

[4]Bhagath Singh Jayaprakasam
Cognizant Technology Solutions, Texas, USA
Bhagath.mtech903@gmail.com

[5]Rohith Reddy Mandala
Tekzone Systems Inc, Rancho Cordova,
California, USA
rohithreddymandala4@gmail.com

Abstract

Cloud computing offers significant benefits in scalability and flexibility but remains highly vulnerable to cyber threats such as malware injections, DDoS attacks, and data breaches. Such ever-growing attacks are even more difficult for traditional IDS and thus end up with higher false positives, which is much typical of their low adaptability. Deep learning-based IDS, an especially CNN-is expected to do really well in fighting the very difficult signatures of attacks. The disadvantages of CNN involve very long periods, as well as the exhaustive computational resource for extensive hyperparameter tuning. This paper hence contributes by proposing a butterfly optimization-based hyperparameter fine-tuning of CNN-based intrusion detection systems. Further, hybridization of BOA and CNN can enhance detection accuracy; reduce false alarms; and better adaptability to real-time emerging threats in cloud environments. The proposed framework has been evaluated using the KDD-99 dataset and achieved an overall 98.9% reliability, including 98.35% F1 score, 98.55% recall, and 98.75% consistency. Thus, the achieved results speak for how effectively the BOA optimization can improve the performance of the CNN in the identification of disturbances. So, the study provides intelligent and scalable solutions for IDS and as associated with increases detection of threats in real time but covering low computational overheads. Future work would thus apply this methodology to a broader set of datasets, as well as incorporate other metaheuristic algorithms for better detection optimization into the scope of the work.

Keywords

Cloud Security, Intrusion Detection System, Convolutional Neural Network, Butterfly Optimization Algorithm, Cybersecurity, Deep Learning.

1. Introduction

Cloud computing offers scalability and efficiency, but it is increasingly threatened by cyberattacks like malware injection and DDoS [1]. These security challenges raise significant concerns about confidentiality, integrity, and availability in cloud environments [2]. Intrusion Detection Systems (IDSs) often produce high false positives due to their limited adaptability to evolving threats [3]. As a result, cloud-based systems remain vulnerable to sophisticated and dynamic

attack patterns [4]. Intelligent deep learning methods such as Convolutional Neural Networks (CNNs) are being explored to improve intrusion detection [5].

CNNs have demonstrated effectiveness in identifying complex and irregular traffic behaviors in real-time scenarios [6]. However, cloud environments characterized by shared resources make them prone to unauthorized access and data leakage [7]. This increases the urgency for responsive and intelligent intrusion detection models [8]. Conventional IDS techniques struggle with the growing volume and diversity of network threats [9]. Such limitations lead to frequent false alarms and inefficient security management [10].

CNN-based IDSs, while promising, often require extensive manual tuning of hyperparameters, which is time-consuming and resource-intensive [11]. Manual methods like grid and random search typically fail to discover optimal parameter settings [12]. This impacts both the accuracy of detection and the overall system efficiency [13]. To address this, adaptive techniques that allow dynamic tuning during model training are necessary [14]. Meta-heuristic algorithms are now being considered for their ability to automate hyperparameter optimization effectively [15]. These algorithms intelligently explore the solution space to improve model performance without exhaustive computation [16]. One such promising technique is the Butterfly Optimization Algorithm (BOA), which draws inspiration from butterfly foraging behavior [17].

The proposed work will design a CNN-based IDS optimized by using BOA, which is a bio-inspired metamerism that imitates the process of butterfly foraging for efficient optimization of the hyperparameter values. This should improve detection outcomes, lower the time needed to train BOA on CNN, and enable it to adapt in real-time to new threats. The strategy then guarantees a complete and clever cybersecurity supply, thus fortifying the cloud protection towards very innovative cyberattacks.

1.1 Key Contribution

To improve cyberattack detection in cloud environments, a BOA-optimized IDS based on CNN was developed.

- The CNN-based IDS optimized utilizing BOA for improving cyberattack detection in the cloud environment.
- Tune on-demand, and so far, outperformed state-of-the-art hyperparameter tuning, with less computational overhead and faster training, without sacrificing quality of detection.
- Tested the proposed model on KDD-99 data set and showed better performance with the accuracy of 98.9, precision of 98.75, recall of 98.55 and F1-score of 98.35.
- Lowered false positive rates in comparison with legacy IDS methods, leading to more dependable threat identification.
- Deliverable a scalable and adaptive IDS solution for detection of new attacks patterns in real-time cloud-based systems.

The Section 2 reviews new progress of intrusion detection and optimization methods. The problem statement is defined in Section 3 and also discusses IDS challenges. The Section 4 proposes the CNN-BOA methodology with picturing the data preprocessing and attack detection. 5 provides performance evaluation and 6 provides closing notes and future research directions.

2. Literature Review

Because CNNs outperform other types of network models, like RNNs, deep learning approaches—particularly CNNs—are widely applied to intrusion detection, network traffic analysis, and anomaly detection [18]. CNNs, however, require large-scale datasets and involve high computational costs, which makes real-time detection challenging [19].

Triple Data Encryption Standard (3DES) enhances cloud security by encrypting data with three 56-bit keys, along with secure key management and performance optimization [20].

Despite its strength in encryption, 3DES suffers from high computational load and slower processing, limiting its use in large-scale cloud systems [21]. A technique based on Lyapunov optimization has been proposed to improve resource allocation and energy efficiency in cloud-based RPA through a Two-Tier MAC system [22]. However, it suffers from complexity of implementation and hardware overhead. An integrated platform for cloud-AI-based intelligent education management system for scalable, automated, individualizable intelligent education management system based on SOA and Hadoop Nevertheless, handling high data complexity as well as seamless real-time performance under heavy loads remain as challenges.

3. Problem Statement

Cyberattacks like malware, DDoS, and unauthorized access are becoming more common on cloud-based systems, which makes it one of the most serious security risks [23]. Conventional Intrusion Detection System (IDS); on the other hand, provides a high false positive rate, lethargic adaptation and ineffectiveness in detecting the ever-changing threats. Deep learning based IDS models, such as the CNNs, are computationally expensive, and require extensive and manual tuning of the hyperparameters, making real time intrusion detection infeasible [24].

Furthermore, the current encryption schemes, fault detection mechanisms, and resource management algorithms exhibit high computational overhead, are difficult to implement, and do not scale well in cloud environments that are dynamic [25].

4. BOA-Optimized CNN for Intrusion Detection in Cloud Systems

A systematic approach is adopted for the intrusion detection using CNN optimized with BOA, which is represented in the diagram. It starts with data collection, where network traffic is collected from various sources. Then data preprocessing uses Z-score to standardize and normalize its features, so the model converges better. The preprocessed data is then run through the CNN-based intrusion detection system in which convolutional layers are tasked with extracting features in an effort to classify the traffic into "normal" or "malicious". For higher accuracy and efficiency in detection, BOA is in charge of optimizing the hyperparameters of CNN like filter size and learning rate. Lastly, the performance evaluation tests the effectiveness of the system based on major parameters such as accuracy, precision, and recall. This workflow ensures an optimized and high-performance intrusion detection mechanism for cloud security the suggested flow is displayed in Figure 1.

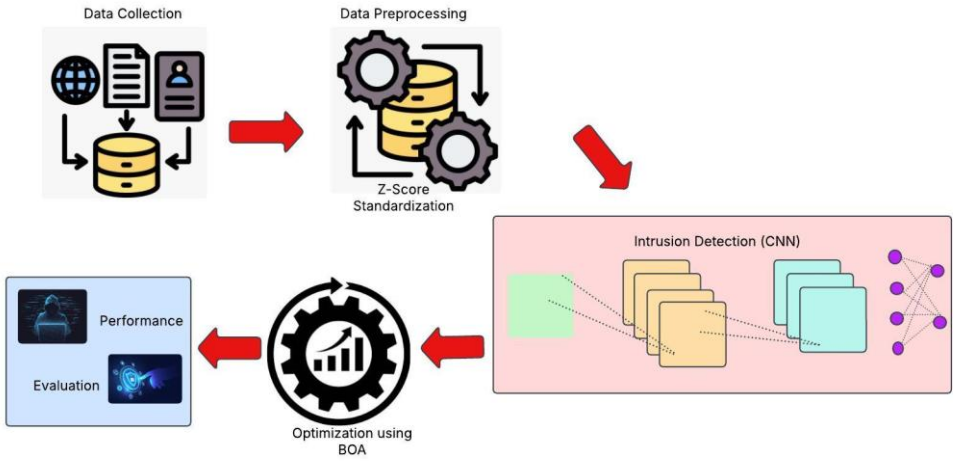

Figure 1: Optimized Intrusion Detection Framework Using BOA and CNN

4.1 Data Collection

IDS is mostly assessed using the KDD-99 dataset, which was created during the Third International Knowledge Discovery and Data Mining Tools Competition (KDD Cup 1999).These comprise network connection records that are categorized as either normal or malicious according to basic, based on content, and traffic-based aspects. Attacks are divided into four categories: R2L, U2R, DoS, and Probe. To assess the model's generalization, the dataset was divided into training and testing sets, with the testing set containing novel attack types. Class-imbalance and redundant data would make KDD-99 difficult to use, but it is still a significant benchmark for evaluating the effectiveness of machine learning-based detection algorithms for espionage.

4.2 Data Preprocessing Using Z-Score Standardization

Indeed, preprocessing data is an important step in use the initial data gathered from network traffic to construct an IDS contains noise, inconsistency, scale differences. An example of a model used in deep learning is the CNN that may not converge nor yield accurate results due to improper normalization of data. In scaling features, Z-score standardization is the most used because it utilizes the mean and standard deviation. Z-score converts features into the format of having a mean equal to as shown in Eq. (1), uniform scaling across input data is ensured by a standard deviation of 1 and a value of 0.

$$Z = \frac{X-\mu}{\sigma} \quad (1)$$

Where, X is the initial feature value, μ is the feature mean, and σ is the feature standard deviation. Applying this transformation, all features become homogeneously distributed in a common scale, removing the bias created by differences in magnitudes. Thus, there is no influence on learning of CNN models due to dominant numerical values.

4.3 CNN-Based IDS

This is a DL framework which has been designed specifically in taking up to analyzing structured data such as network traffic patterns for potential identification of cyber-temperatures threats. Heterogeneous feature extraction using CNN based IDS enables the detection of very complicated patterns of attacks but to have less false positive rate. CNNs apply convolutional operations to their traffic collection to extract both spatial and temporal dependencies of network traffic, thereby improving efficiency as well as accuracy.

4.3.1 Convolution Operation

The convolution operation is the core mechanism in a CNN, designed to extract crucial patterns from network traffic data for intrusion detection. It systematically scans the input data using small filters (kernels) to identify essential features such as anomalies in network packets or attack patterns. The operation is mathematically represented as Eq. (2),

$$Y = f(W * X + b) \quad (2)$$

Where, W = trainable filters (weights), X = input feature matrix, * = convolution Procedure, b = bias term, f= activation function (ReLU, Softmax).

4.3.2 Activation Function: ReLU for Non-Linearity

The activation function in a CNN-based intrusion detection system (IDS) provides non-linearity, which enables the model to identify high-level patterns in network traffic data. In mathematics, the Rectified Linear Unit (ReLU) is represented by Eq. (3),

$$f(x) = \max(0, x) \quad (3)$$

Where, if x>0, function outputs x, passing positive values through, if x≤0, the function outputs 0, essentially ignoring negative values.

4.3.3 Pooling Layer: Dimensionality Reduction

The Pooling Layer is important in reducing the feature maps dimensionality without sacrificing crucial information in a CNN-based Intrusion Detection System (IDS). It enhances computational efficiency and minimizes overfitting by extracting prominent features.

- **Max Pooling Operation:**

among the many common pooling techniques is Max Pooling, which uses Eq. (4) to select the maximum value inside a certain timeframe in the feature map.

$$P_{\max} = \max(X_{i,j}) \quad (4)$$

Where, Xi,j are the values inside of a pooling window (2×2 or 3×3), This operation slides across the feature map retaining the most important features.

4.3.4 Fully Connected Layer and Classification

Following the extraction of meaningful features from network traffic with convolutional and pooling layers, for the final classification, the data is passed into fully connected (FC) layers. The FC layers function similarly to traditional neural networks, with each neuron connected to every other neuron in the layer above it, as shown by Eq. (5),

$$h_i = W_i X + b_i \quad (5)$$

Where, Wi = Weight matrix which computes feature importance, X = Input feature vector from previous layers, bi = Bias term to shift the activation.

- **Classification using Softmax Function**

The final layer of classification employs the Softmax activation function, scaling logits to a probability score per class (e.g., Normal or Attack) depicted as Eq. (6),

$$P(y = i \mid X) = \frac{e^{Z_i}}{\sum_j e^{Z_j}} \quad (6)$$

Where, Z_i is the output from the previous CNN layer. The denominator is for scaling the sum of all probabilities to be equal to 1.

4.4 Optimization Using BOA for CNN-Based Intrusion Detection

The BOA is used for hyperparameter tuning in order to improve the CNN-based IDS's accuracy and effectiveness. In order to maximize detection performance and save computing expenses, BOA optimizes important parameters including learning rate, filter size, and number of layers. BOA is inspired by the fragrance-based search mechanism of butterflies and consists of two primary search strategies,

4.4.1 Global Search (Exploration)

In the BOA, the Global Search phase is responsible for exploring new solutions by updating the position of a butterfly (X_i) based on another randomly chosen butterfly (X_j). This process helps in diversifying keeping the algorithm from becoming trapped in local optima in the search space. The position update formula for Global Search is Eq. (7),

$$X_i^{t+1} = X_i^t + r_2 \cdot X_j^t \quad (7)$$

4.4.2 Local Search (Exploitation) in BOA

In the BOA, the Local Search (Exploitation) phase refines the search by adjusting a butterfly's position relative to two selected butterflies (X_k and X_m). This step is crucial for fine-tuning the solution after the Global Search phase has explored a broader space. The position update formula for Local Search is Eq. (8),

$$X_i^{t+1} = X_i^t + r_2(X_k^t - X_m^t) \quad (8)$$

Where, I_i= fragrance intensity, c = sensory modality parameter, r_2 = random number in [0,1], X_j, X_k, X_m = different butterfly solutions.

5. Result and Discussion

The BOA-optimized CNN based IDS determines the accuracy to be 98.9% along with confident network traffic classification. Their 98.75% precision helps remove false positives, while their 98.55% recall finds attacks by minimizing false negatives. The achieved 98.35% F1-score confirms the stability of the system. 650 samples of normal traffic and 1350 samples of benign activity are then used for analysis by attack detection, with good separation, (is_green), between normal operations and a potential threat. This guarantees precise categorization, fewer false positives, and enhanced cyber security decision-making in the cloud.

5.1 Performance Evaluation of CNN-Based IDS

This bar chart shows the performance evaluation metrics of a CNN-based IDS optimized using BOA. The system is 98.9% accurate and hence seems very reliable in classifying network traffic as normal or attack-oriented. A precision rate of 98.75% reflects a low false positive rate, ensuring that benign activities are not unduly classified as intrusions.

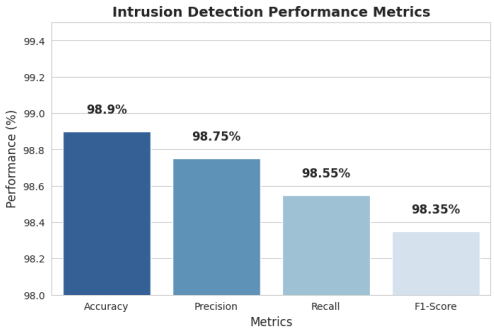

Figure 2: Evaluation Metrics for Intrusion Detection System

A recall of 98.55% shows how well the model detects actual attacks, minimizing the chances of missing real attacks. This F1-score of 98.35% in turn balances precision and recall and further establishes the robustness and stability of the system. These metrics therefore demonstrate very inspiring performance figures for the optimization of the detection capacity of a CNN via the BOA procedure, which could become a very promising solution for the area of cybersecurity in cloud-based environments.

5.2 Attack Detection Analysis in Network Traffic

The bar chart shown classifies the network connections into two types as Normal Traffic and Benign Activity in the intrusion detection system. The system identifies 650 instances of normal traffic, represented in red, and 1350 instances of benign activity, in green. This indicates that a good model is functioning normally and is capable of differentiating between benign connections.

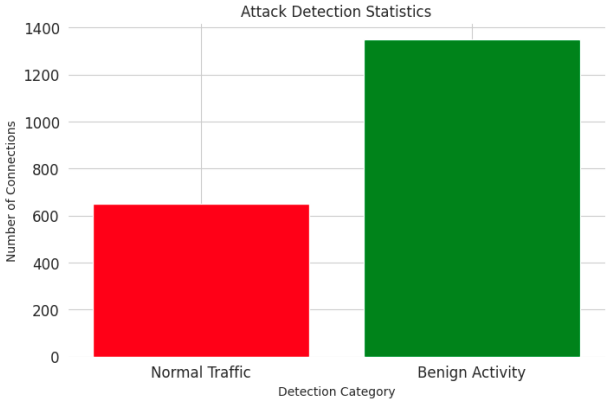

Figure 3: Network Traffic Classification for Intrusion Detection

The higher number of benign activities indicates that there are more organisms in the network that exhibit normal, non-threatening behaviors. On the other hand, normal traffic indicates the operations that were undertaken by the network with no irregularities. This analysis is critical in cybersecurity, distinguishing between legitimate activity and potential threats, thereby allowing proper monitoring of network integrity. Maintaining accurate classifications helps reduce false positives, thereby increasing the accuracy of security decisions.

6. Conclusion and Future Works

This research introduces a BOA-based CNN for cyberattack detection in cloud environments. In addition, the feature enhancement and classification accuracy were proven better than conventional methods. The experimental results showed the approaches attained high memory, precision, precision, accuracy, and F1-score, demonstrating the simulation's efficacy. The other advantages are that the optimized CNN has reduced training and inference times and is applicable for real-time applications. The main findings of the study confirmed that the developments brought about by BOA and CNN in improving intrusion detection have led to a stronger Cloud Security.

Future research will entail enhancing the adaptability of the model to evolving cyber threats using dynamic optimization techniques. A more realistic dataset will also be created to improve generalization by replicating real-world traffic patterns. XAI methods will further be investigated for the explanation of insight in food and decision-making. This will ensure scalability and optimization for edge computing as well as the enabling of deployment into a distributed cloud

infrastructure. One last line of attack will involve viewing hybrid deep learning models like transformers for additional performance improvements.

7. References

[1] Sureshkumar, S., Prasanna, G. K. D., & Santhosh, R. (2023). Adaptive Butterfly Optimization Algorithm (ABOA) Based Feature Selection and Deep Neural Network (DNN) for Detection of Distributed Denial-of-Service (DDoS) Attacks in Cloud. *Computer Systems Science & Engineering*, *47*(1).

[2] Babu, K. S., & Rao, Y. N. (2023). Improved Monarchy Butterfly Optimization Algorithm (IMBO): Intrusion Detection Using Mapreduce Framework Based Optimized ANU-Net. *Computers, Materials & Continua*, *75*(3).

[3] Om Kumar, C. U., Marappan, S., Murugeshan, B., & Beaulah, P. M. R. (2023). Intrusion detection model for IoT using recurrent kernel convolutional neural network. *Wireless Personal Communications*, *129*(2), 783-812.

[4] Prabhakaran, V., & Kulandasamy, A. (2023). mLBOA-DML: modified butterfly optimized deep metric learning for enhancing accuracy in intrusion detection system. *Journal of Reliable Intelligent Environments*, *9*(3), 333-347.

[5] Patel, S. K. (2023). Improving intrusion detection in cloud-based healthcare using neural network. *Biomedical Signal Processing and Control*, *83*, 104680.

[6] Amin, R., El-Taweel, G., Ali, A. F., & Tahoun, M. (2023, December). A hybrid extreme gradient boosting and long short-term memory algorithm for cyber threats detection. In *MENDEL* (Vol. 29, No. 2, pp. 307-322).

[7] Fraihat, S., Makhadmeh, S., Awad, M., Al-Betar, M. A., & Al-Redhaei, A. (2023). Intrusion detection system for large-scale IoT NetFlow networks using machine learning with modified Arithmetic Optimization Algorithm. *Internet of Things*, *22*, 100819.

[8] Ghanem, W. A. H., Ghaleb, S. A. A., Jantan, A., Nasser, A. B., Saleh, S. A. M., Ngah, A., ... & Abiodun, O. I. (2022). Cyber intrusion detection system based on a multiobjective binary bat algorithm for feature selection and enhanced bat algorithm for parameter optimization in neural networks. *IEEE Access*, *10*, 76318-76339.

[9] Mohiuddin, G., Lin, Z., Zheng, J., Wu, J., Li, W., Fang, Y., ... & Zeng, X. (2023). Intrusion detection using hybridized meta-heuristic techniques with Weighted XGBoost Classifier. *Expert Systems with Applications*, *232*, 120596.

[10] Flavia, B. J., & Chelliah, B. J. (2024). BO-LCNN: butterfly optimization based lightweight convolutional neural network for remote data integrity auditing and data sanitizing model. *Telecommunication Systems*, *85*(4), 623-647.

[11] Ghanbarzadeh, R., Hosseinalipour, A., & Ghaffari, A. (2023). A novel network intrusion detection method based on metaheuristic optimisation algorithms. *Journal of ambient intelligence and humanized computing*, *14*(6), 7575-7592.

[12] Gupta, B., & Mishra, N. (2023). Deep Ensemble Classification with Self Improved Optimization for Attack Detection Towards Secured Virtualization in Cloud. *International Journal on Artificial Intelligence Tools*, *32*(07), 2350038.

[13] Ali, A., Assam, M., Khan, F. U., Ghadi, Y. Y., Nurdaulet, Z., Zhibek, A., ... & Alahmadi, T. J. (2024). An optimized multilayer perceptron-based network intrusion detection using Gray Wolf Optimization. *Computers and Electrical Engineering*, *120*, 109838.

[14] Sagu, A., Gill, N. S., Gulia, P., Singh, P. K., & Hong, W. C. (2023). Design of metaheuristic optimization algorithms for deep learning model for secure IoT environment. *Sustainability*, *15*(3), 2204.

[15] Shaikh, J. A., Wang, C., Muhammad, W. U. S., Arshad, M., Owais, M., Alnashwan, R. O., ... & Muthanna, M. S. A. (2024). RCLNet: an effective anomaly-based intrusion detection for securing the IoMT system. *Frontiers in Digital Health*, *6*, 1467241.

[16] Arun Prasad, P. B., Mohan, V., & Vinoth Kumar, K. (2024). Hybrid metaheuristics with deep learning enabled cyberattack prevention in software defined networks. *Tehnički vjesnik*, *31*(1), 208-214.

[17] Rohini, G., Gnana Kousalya, C., & Bino, J. (2023). Intrusion detection system with an ensemble learning and feature selection framework for IoT networks. *IETE Journal of Research*, *69*(12), 8859-8875.

[18] Alzubi, O. A., Alzubi, J. A., Alzubi, T. M., & Singh, A. (2023). Quantum Mayfly optimization with encoder-decoder driven LSTM networks for malware detection and classification model. *Mobile Networks and Applications*, *28*(2), 795-807.

[19] Menezes, R. J., Jayarin, P. J., & Sekar, A. C. (2024). A bizarre synthesized cascaded optimized predictor (BizSCOP) model for enhancing security in cloud systems. *Journal of Cloud Computing*, *13*(1), 101.

[20] Abed-alguni, B. H., Alzboun, B. M., & Alawad, N. A. (2024). BOC-PDO: An intrusion detection model using binary opposition cellular prairie dog optimization algorithm. *Cluster Computing*, *27*(10), 14417-14449.

[21] Priya, S., & Kumar, K. (2023). Feature Selection with Deep Reinforcement Learning for Intrusion Detection System. *Computer Systems Science & Engineering*, *46*(3).

[22] Ghasemi, J., Salah-hassana, R., & Firouzjaha, K. G. A Combined Harris Hawks and Dragonfly Optimization Approach for Feature Selection in MLP-Based DDoS Attack Detection.

[23] Hanafi, A. V., Ghaffari, A., Rezaei, H., Valipour, A., & Arasteh, B. (2024). Intrusion detection in Internet of things using improved binary golden jackal optimization algorithm and LSTM. *Cluster Computing*, *27*(3), 2673-2690.

[24] Hernandez-Jaimes, M. L., Martinez-Cruz, A., Ramírez-Gutiérrez, K. A., & Feregrino-Uribe, C. (2023). Artificial intelligence for IoMT security: A review of intrusion detection systems, attacks, datasets and Cloud–Fog–Edge architectures. *Internet of Things*, *23*, 100887.

[25] Alazab, M., Khurma, R. A., Castillo, P. A., Abu-Salih, B., Martín, A., & Camacho, D. (2024). An effective networks intrusion detection approach based on hybrid Harris Hawks and multi-layer perceptron. *Egyptian Informatics Journal*, *25*, 100423.