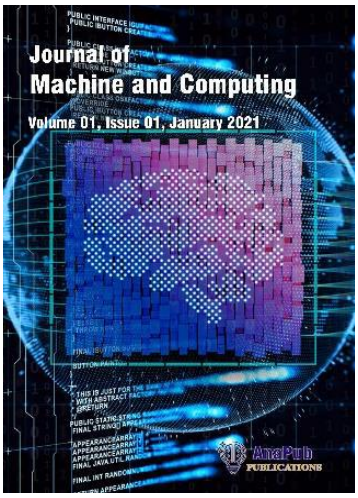
Journal Pre-proof

Advanced Multi-Class Cyber Security Attack Classification in IoT-based Wireless Sensor Networks Using Context-Aware Depthwise Separable Convolutional Neural Network



Bangar Raju Cherukuri

DOI: 10.53759/7669/jmc202505064

Reference: JMC202505064

Journal: Journal of Machine and Computing.

Received 30 May 2024

Revised form 18 December 2024

Accepted 28 January 2025

Please cite this article as: Bangar Raju Cherukuri, “Advanced Multi-Class Cyber Security Attack Classification in IoT-based Wireless Sensor Networks Using Context-Aware Depthwise Separable Convolutional Neural Network”, Journal of Machine and Computing. (2025). Doi: https://doi.org/10.53759/7669/jmc202505064

This PDF file contains an article that has undergone certain improvements after acceptance. These enhancements include the addition of a cover page, metadata, and formatting changes aimed at enhancing readability. However, it is important to note that this version is not considered the final authoritative version of the article.

Prior to its official publication, this version will undergo further stages of refinement, such as copyediting, typesetting, and comprehensive review. These processes are implemented to ensure the article's final form is of the highest quality. The purpose of sharing this version is to offer early visibility of the article's content to readers.

Please be aware that throughout the production process, it is possible that errors or discrepancies may be identified, which could impact the content. Additionally, all legal disclaimers applicable to the journal remain in effect.

AnaPub
PUBLICATIONS

Advanced Multi-Class Cyber Security Attack Classification in IoT-based Wireless Sensor Networks Using Context-Aware Depthwise Separable Convolutional Neural Network

Bangar Raju Cherukuri
Senior Web Developer, Washington, DC, Germantown MD USA,
rajucherukuri5@gmail.com

Abstract

One of the most widely used wireless technologies in recent years has been wireless sensor networks (WSN), which has led to intriguing new Internet of Things (IoT) applications. Internet Protocol IP integration with IoT-based WSN enables any physical item with sensors must have widespread connectivity and transmit data in real time to the server linked to the gate on the internet. WSN security is still a developing area of study that falls under the Internet of Things paradigm. To protect digital infrastructures, strong techniques for precise and effective multi-class classification are required due to the growing frequency and sophistication of cyber-attacks. The proposed method makes use of the CICIDS2017 and UNSW-NB15 datasets alongside IoT-based wireless sensor networks to enhance cyber-security detection. In this work, Boosted Sooty Tern Optimization (BSTO) and Context-Aware Depthwise Separable Convolutional Neural Networks (CA-DSCNN) present an enhanced method for classifying multi-class cyber-security attacks. To guarantee consistent feature scaling, the proposed approach starts by applying Min-Max Scaler Normalization to preprocess the raw attack data. There is a feature selection stage that comes afterwards that uses Banyan Tree Growth Optimization (BTGO) combined with Augmented Snake Optimizer (ASO) to efficiently find and choose the most relevant characteristics to improve classification performance. Because of its strong feature extraction capabilities and computational efficiency, the CA-DSCNN is used; depthwise separable convolutions are used to strike a compromise between processing needs and accuracy. This architecture enhances the ability to extract complicated characteristics from the data and to comprehend those characteristics in context. BSTO is used to optimize the neural network's parameters, improving classification efficiency and accuracy in order to further enhance model performance. By lowering computational expenses and over-fitting, the proposed methodology which integrates IoT-based wireless sensor networks enhances cyber-security attack classification, exhibiting improved accuracy 99.5% and high PDR 99%.

Keywords: multi-class cyber security attack, IoT-based WSN, Min-Max Scaler Normalization, Context-Aware Depthwise Separable Convolutional Neural Networks, Banyan Tree Growth Optimization, Augmented Snake Optimizer, and Boosted Sooty Tern Optimization.

1. Introduction

Cyber-security threats are becoming an increasing issue for everyone in today's society, where the internet plays a major role, including individuals, businesses, and governments. These are attacks that are specifically created to breach, compromise, or penetrate data, networks, or machines. The number of linked gadgets and the Internet of Things (IOT) has increased risk and created a new attack surface due to the exponential growth of internet access [1-5]. The decentralized structure of WSNs (wireless sensor networks) makes security a significant worry. Data and security are frequently compromised by these networks because of the high frequency of security assaults based on node capture and node hacking. The risks to WSNs are also relevant to and dangerous for IoT networks since they are made up of sensor-based networks.

Malware, phishing, denial-of-service (DoS) assaults, and other tactics are some of the ways that cyber security attacks might appear. Significant financial losses, data breaches, and reputational harm can all be brought on by these evil operations [6-8]. Advanced strategies for identifying, categorizing, and mitigating these threats must be developed and put into action since attackers are always improving their techniques.

Cyber-attacks are becoming more harmful due to the advancement of internet technologies. Hackers are increasingly focusing their attacks on Cyber-Physical Systems (CPS) rather than traditional systems. Cyber-attacks targeting intelligent transportation and intelligent homes are growing faster each year. A self-driving car's serious flaws were discovered in 2005 by two security experts [9-13].They were able to stop a self-driving Jeep on a highway by remotely controlling the vehicle's major functions. Cyber-attack methods are evolving into increasingly potent and advanced forms. State-sponsored hackers, as well as individual hackers, are actively planning cyber-attacks. With the use of offensive cyber-security technology, cybercriminals carry out complex attacks. The term "offensive cyber-security" describes a hacking method that targets a system rather than a protection mechanism [14].

Even in the face of unanticipated threats or external attacks, vital facilities like ICS (Internet Industrial Control Systems) and SIPS (Sensitive Industrial Plants and Sites) must continue to function and be dependable. The communication layers, data management, and control are among the systems that are susceptible to cyber-attacks [15-17]. These levels provide malicious individuals with access to sensitive data that they can steal or alter, possibly destroying physical assets and resulting in significant losses. Malicious users have the ability to alter crucial metrics used for managing or observing infrastructure components.

Fighting malicious software is necessary for cyber-security, as it can remain dormant while monitoring compromised assets and infrastructure [18]. The swift advancement of technology such as, IoT and cloud computing boosts confidence in cyber-security. Due to the volume of encrypted traffic and dynamic port allocation, traditional methods of network intrusion detection are no longer effective. Instead, machine learning techniques have replaced port inspection as the method of choice [19-20]. Network anomaly detection in a variety of cloud environments can be addressed with machine learning and deep learning.The study's main contributions are:

- The proposed method uses Min-Max Scaler Normalization to reduce the effect of different feature ranges and normalize feature scales, which improves the model's capacity to learn from the data. The model's capacity to learn efficiently from a variety of IoT-based wireless sensor network data is improved by this standardization.
- In order to provide effective and efficient feature selection that enhances model performance by dimensionality reduction and focusing on the most important features, a hybrid Banyan Tree Growth Optimization with Augmented Snake Optimization is presented. By choosing the most pertinent features from the IoT-based data, this technique reduces dimensionality and boosts model efficiency.

- The proposed method uses a context-aware, depthwise separable convolutional neural network (CA-DSCNN) to minimize computational complexity, resulting in a classification that is more accurate and economical.
- The proposed method uses Boosted Sooty Tern Optimization (BSTO) to adjust network parameters in order to overcome issues like over-fitting and computational complexity and maximize the classification model's accuracy and computational efficiency.
- A methodology for identifying multi-class cyber security assaults is made scalable and effective by the method. The suggested technique boosts speed and precision, two essential elements for quickly identifying and mitigating security threats, by utilizing data from IoT-based wireless sensor networks.

The manuscript is organized as follows: Section 1 outlines the introduction; Section 2 investigates the literature review; Section 3 presents the proposed methods; Section 4 presents the results and discussions; and Section 5 concludes the manuscript.

2. Literature Survey

Jia Y et al. (2023) [21] have suggested the defense of cyber-security for smart cities facilitated by artificial intelligence: A new approach to threat detection established on the MDATA model. This research presents aninnovative architecture for detecting attacks named ACAM, using a suggested mechanism. The outline is built regarding the MDATA model;it describes information that is temporally and spatially dynamic more effectively than the information graph in order to better express the cyber security knowledge.In order to reduce false alerts and enhance multi-step attack recognition capabilities, the framework includes modules for knowledge extraction, subgraph construction, alarm correlation, and attack detection. The suggested method's implementation complexity, which necessitates significant data, is a limitation.

In 2022 Semwal P and Handa A [22] have suggested the cyber-physical system cyber-attack detection via supervised machine learning.Four distinct supervised machine learning approaches are suggested in this study to develop representations to identify cyber-attack activity on a CPS water treatment facility. The comparison study is carried out by comparing the output of the four classification models, Decision Tree (DT),Random Forest (RF),K-Nearest Neighbors (KNN), and Support Vector Machine (SVM), using evaluation matrices.The suggested method's disadvantage is that it canover-fit complex datasets.

Prabakar D et al. (2023) [23] have demonstrated a cyber-attack detection in a sustainability smart city using energy management and IoT with AI.The study describes a traffic analysis that reduces network traffic and improves data transmission through the use of a kernel polynomial vector classifier. Because there is less traffic, energy efficiency is improved. Next, adversarial Bayesian belief networks are used to detect malicious attacks. Throughput, packet delivery ratio, data traffic analysis,end-end delay, energy efficiency, and quality of service have all been examined experimentally. The potential complexity in model implementation is the method's disadvantage.

Balta EC et al. (2023) [24] have suggested digital twin-based cyber-attack detection system for cyber-physical systems of manufacture. This study tackles two issues related to CPMS cyber-attack identification: the differentiation of cyber-attacks throughout transient response and cyber-attacks from predicted abnormalities. In order to identify cyber-attacks in CPMS through regulated transitory actions and anticipated anomalies, it suggests using a Digital

Twin (DT) paradigm. An experimental case study is offered to illustrate the usefulness of the framework.The complexity of integration is the suggested method's drawback.

In 2022 Li Q et al. [25] have suggested the scalable categorized cyber-attack localization and detection insystems of active dissemination.The study suggested an altered spectrum network partitioning using clustering technique for the "coarse" localization of categorized cyber-attacks. A standardized impact score determined by waveform statistical metrics is then suggested as a way to further refine the cyber-attack site, obtaining a "fine" cyber assault location by describing various waveform attributes. In summary, a thorough quantitative assessment involving two case studies reveals encouraging estimation outcomes for the suggested framework when contrasted with traditional and cutting-edge techniques.

Salam A et al. (2023) [26] have presented deep learning techniques: a novel approach for internet-based assault prevention in sector 5.0.This method focuses on the classification of attacks and the recognition of abnormal behavior using DL (Deep Learning) methods like CNNs, RNNs, and transformer models. Deep learning has proven to be useful in identifying intrusions in Industry 5.0 environments via a transformer-based system that surpasses conventional methods in terms of precision, recall, andaccuracy. This ensures data protection.The suggested method's high computing cost is a disadvantage.

Jullian O et al. (2023) [27]have suggested a scalable attack identification framework for cyber-attacks in IoTnetworks using DL.The distributed framework based on DL that is utilized in this study prevents several sources of vulnerability simultaneously under a single security mechanism. Thefeed-forward neural network and long-short-term memory are two distinct DL models that are assessed. The networks are tested on two distinct datasets (i.e., BoT-IoT and NSL-KDD) for both performance and attack type identification. A drawback of the suggested approach is its high resource consumption and complexity of integration.

Raghunath KK (2022) [28]have introduced the Regression Classifier XGBoost (XRC) model for Inception V4-based cyber-attack identification and categorization.The suggested hybridized classifier, which is utilized in Inception V4 to further develop and evaluate the model, integrates the ideas of both XGBoost and Logistic classifiers. The proposed XRC classifies and predicts a number of prevalent network cyber-attacks, such as phishing, distributed denial of service (DDoS), Internet of Things (IoT), and cross-site scripting (CS). To reduce the erroneous rate and boost efficacy, the hybridized classifier uses the sigmoidal function as a supportive activator.One of the methods' shortcomings is its computationally intensive and complex implementation.

Saghezchi FB (2022) [29]have recommended using machine learning to identifyDDoSassaults in Industry 4.0 CPPSs.The suggested approach makes use of network traffic data that was obtained from an actual semiconductor manufacturing facility. For the purpose of instruction and evaluation of machine learning models, the suggested approach creates several labeled datasets and extracts 45 bidirectional network flow features. The suggested approach examines eleven distinct unsupervised and semi-supervised algorithms and evaluates their efficacy using inclusive simulations. The resultsestablish that supervised algorithms perform better in terms of finding performance than both unsupervised and semi-supervised ones.The suggested method's limitation is restricted to a particular manufacturing setting.

In 2023 Alaca Y and Celik Y [30] have suggested employing lightweight DL algorithms to identify cyber-attacks using QR code descriptions.Initially, substantial data with several classes was produced as QR code images in this investigation. Next, ShuffleNet CNN and

MobileNetV2algorithms were employed for instruction images of QR codes. Following the extraction of features from the training images using Deep CNN models, the Harris Hawk Optimization (HHO) was used to ascertain which characteristics would be most useful for classification. The recommended method's increased computing complexity is a limitation. Table 1 displays the comparison of existing methods.

Table 1: Comparison of existing approaches

References	Method	Advantages	Disadvantages
[21]	ACAM framework with MDATA model	Reduces false alarms, improves multi-stem detection	Implementation complexity, requires extensive data
[22]	KNN, SVM, DT, and RF	Easy to interpret and visualize	Prone to over-fitting with complex datasets
[23]	Kernel quadratic vector discriminant + adversarial Bayesian belief networks	High throughput, improved energy efficient	Potential complexity in model implementation
[24]	Digital twin framework	Real-time detection during system transients	Complexity in integration.
[25]	Deep learning and spectral clustering	Effective at detecting and localizing minor attacks	Complexity in implementation and computation
[26]	CNNs, RNNs, Transformer models.	Enhanced accuracy	High computational cost
[27]	Distributed deep learning framework	High accuracy, comprehensive vulnerability protection	Complexity in integration and high resource consumption
[28]	XGBoost Regression Classifier (XRC) with Inception V4	High accuracy, effective threat detection	Complexity in implementation, computationally intensive
[29]	Machine Learning	High accuracy, real-world data usage	Limited to specific factory environment
[30]	Hybrid HHO, MobileNetV2, and ShuffleNet CNN	High accuracy, efficient feature selection	Increased computational complexity

2.1 Problem Statement

Cyber-security attacks represent significant risks to digital infrastructure; thus, identifying and reducing such hazards requires reliable and precise categorization techniques. The current techniques for classifying cyber-security attacks into many classes have a number of shortcomings, such as difficult implementation, substantial data requirements, over-fitting vulnerability, and expensive computing expenses. These difficulties make it difficult to use them practically, particularly in intricate settings. This research proposes a novel method utilizing a context-aware, depth-wise separable convolutional neural network framework and advanced Boosted Sooty Tern optimization techniques to address these problems. The results include better classification accuracy, lower computational overhead, and increased adaptability in a variety of environments. Furthermore, it uses sophisticated regularization algorithms to prevent over-fitting and reduce the requirement for large amounts of data. The proposed method effortlessly fits into a variety of operational scenarios by optimizing computational efficiency.

3. Proposed Methodology

The proposed method for multi-class cyber-security attack classification initiates with a preprocessing step that uses Min-Max Scaler Normalization to standardize feature scales and improve model performance on raw data. Following normalization, the data is analyzed using feature selection and Banyan Tree Growth Optimization (BTGO) combined with Augmented Snake Optimizer (ASO). By effectively finding and choosing the most pertinent features, this combination lowers dimensionality and raises classification accuracy. After that, the enhanced features are fed into a Context-Aware Depthwise Separable Convolution Neural Network (CA-DSCNN), which takes advantage of depthwise separable convolutions to minimize computational complexity and maximize feature extraction efficiency. In order to improve classification performance, network parameters are adjusted using Boosted Sooty Tern Optimization, which further optimizes the model. This method provides a scalable and effective way to identify and classify various cyber-threats. Figure 1 shows the block schematic illustrates the proposed methodology.

Figure 1: Block diagram of the proposed methodology

3.1 Dataset

The two datasets used in the proposed method, UNSW-NB15 and CICIDS2017, are well known for their ability to classify cyber-security attacks into multiple classes. Additionally the method uses data from the IoT- based wireless sensor networks. These datasets offer a broad variety of attack scenarios, allowing a comprehensive evaluation of the method's efficacy in identifying and categorizing various cyber-threats. Preprocessing based on Min-Max Scaler Normalization is applied to the datasets to provide uniform scaling across features. By minimizing the bias caused by different feature scales, this step improves the performance of the classification that comes next.

3.2 MinMax scaler normalization based Preprocessing

The datasets are fed into Min-Max Scaler Normalization-based preprocessing to efficiently scale and normalize the feature values, ensuring consistency and relevance for accurate analysis. The normalization procedure ensures that each item of data in the database has a comparable range. When the data has no structure and has a wide range of values, this becomes crucial. Normalization with MinMax scaler is beneficial for high-dimensional data. The feature values in cyber-security might differ greatly because of the variety of attack methods and data sources. Model training may become challenging as a result of this variation. A normalization method called MinMax scaler raises every feature's value to a range of 0 to 1, which enhances the stability and performance of the model. Equations (1) and (2) describe the MinMax scaler normalizing algorithm [31].

$$I_{Std} = \frac{(I - I.Min)}{(I.Max - I.Min)} \tag{1}$$

$$I_{Scaled} = I_{Std} * (I.Max - I.Min) + I.Min \tag{2}$$

The lowest and highest feature values for the dataset under consideration are represented by the min and max values in Equations (1) and (2). These attributes are normalized in the dataset through preprocessing, guaranteeing consistency between various data points. Equations (1) and (2) offer the normalized values corresponding to every feature. Before being used for model training and testing, these normalized values are fit and transformed for the full dataset.The relevant features are then chosen by feeding the preprocessed data into the feature selection process.

3.3 Hybrid Banyan tree growth optimization and Augmented Snake optimizer based feature selection

The important aspects are chosen from the preprocessed data using feature selection. To optimize feature subsets, the hybrid Banyan Tree Growth Optimization (BGTO) and Augmented Snake Optimizer (ASO) based feature selection techniques combine the advantages of both algorithms. Whereas ASO improves the search by concentrating heavily on favorable regions, BGTO expands and grows branches in the solution space to examine a variety of feature combinations. By combining exploration and exploitation, this hybrid strategy produces feature selection that is more precise and effective. In order to promote both high accuracy and low feature count, the fitness function utilized balances predictive performance with feature subset size. As a consequence, a strong feature selection procedure is produced that makes use of the advantages of both optimization techniques.

3.3.1 Banyan tree growth optimization (BTGO) [32]

The ancient species of tropical and subtropical plants known as banyan trees, with their many aerial roots and expansive canopies, served as inspiration. They are sensitive to environmental elements such as water, nutrients, and light and have a strong capability for growth and adaptation. Growth hormones in the tree direct its trunks toward locations with more resources, enabling it to develop in that direction. The concept of optimization is present in the unique growth style of the banyan tree and offers suggestions for remedies. There are several cycles in the growth process, as new leaves and branches emerge and withering branches break down.

3.3.2 Augmented Snake optimizer (ASO) [33]

The behavior of snakes mating in low-temperature environments and in the presence of food serves as the model for the Snake Optimization concept. To improve the global's efficiency, this procedure includes transitional phases. When it's hot outside, snakes concentrate on consuming the food that is accessible. Mating takes place in pairs in cold weather, and females may lay eggs that develop into baby snakes while they are in the search area.

- **Initialization**

The hybrid initialization averages random variables within predefined constraints by combining the BTGO and ASO approaches. For better optimization exploration, this method guarantees a variety of well-balanced starting locations throughout the solution space.

$$X_{a,b} = \frac{1}{2}\left[X_{b,\min} + Rand_{BTGO} \times \left(X_{b,\max} - X_{b,\min}\right) + x_{\min} + Rand_{ASO} \times \left(x_{\max} - x_{\min}\right)\right] \tag{3}$$

Where $X_{b,\min}$ represents the minimum value for $b-th$ dimension, $X_{b,\min}$ denotes the maximum value for $b-th$ dimension, $Rand_{BTGO}$ is the random value for BTGO, $x_{\min}$ is the minimum value for solution space, $x_{\max}$ maximum value for solution space, and $Rand_{ASO}$ is the random value for ASO.

- **Fitness function**

The fitness function of the hybrid BGTO-ASO optimization approach was recently proposed is shown in Equation (4).

$$Fitness(X) = \frac{1}{1 + Error(X)} - \lambda \times \frac{X}{n_{\max}} \tag{4}$$

Where $Error(X)$ denotes the measures model error with selected features, $|X|$ is counts the number of selected features, $n_{\max}$ denotes the maximum allowable feature count, and λ represents the balance accuracy and feature count.

- **Exploration**

The exploration phase in BTGO has been established in order for the algorithm to retain diversity more efficiently. Equations display the phase of exploration (5)–(6).

$$B_i = B_i + \in \times n(0,1) \tag{5}$$

Where $n(0,1)$ indicates the Gaussian distribution's random numbers and $\in$ denotes the exploration factor. This is computed using Equation (6).

$$\in = Step \times Rand \times e^{1 - \frac{\max iter}{\max iter - f + 1}} \tag{6}$$

Where $\max iter$ represents the greatest quantity of repetitions, f is the current generation, and the variable that corresponds to the search space's breadth is the parameter *Step*.

- **Exploitation**

Exploitation in the Snake Optimizer is similar to locating and taking advantage of food sources in that it involves a thorough search around recognized high-quality solutions. By focusing on areas that show promise, this phase improves the search's refinement and increases convergence efficiency and accuracy of the solutions.

$$S_{worst,M} = S_{\min} + Rand \times \left(S_{\max} - S_{\min}\right) \tag{7}$$

$$S_{worst,F} = S_{\min} + Rand \times \left(S_{\max} - S_{\min}\right) \tag{8}$$

Where $S_{worst,M}$ is the worst member in the male group, $S_{worst,F}$ is the worst member in the female group, Metaheuristic algorithms that optimize agent direction can make random position adjustments thanks to the flag direction operator, also called the diversity factor.

- **Termination**

After every step, the termination condition of the hybrid optimization is established by increasing the number of iterations $t = t + 1$.The hybrid BGTO and ASO feature selection method combines the advantages of both techniques to explore and refine feature subsets in an effective manner. In an attempt to streamline the model and improve model performance, this method selects the most important features from the dataset. Following feature selection, a context-aware depthwise separable convolution neural network is employed in the classification stage to categorize the multi-class cyber-security attack based on its optimum properties.

3.4 Context-Aware Depth Wise Separable Convolution Neural Network (CA-DSCNN)

The next step for the feature selection is classification. The proposed method uses Context-Aware Depthwise Separable Convolutional Neural Network (CA-DSCNN): This neural network effectively captures contextual and spatial information with low computational overhead, improving multi-class cyber-security threat categorization. Figure 2 shows the architecture of proposed CA-DSCNN.

Figure 2: Architecture of CA-DSCNN

3.4.1 Depth wise separable convolutional neural network

Decomposition of depth-wise separable convolution yields two different forms: depth-wise convolution and 1x1 convolution, which is also referred to as point-by-point convolution. If point-by-point convolution combines feature maps from several channels in a normal 1x1 convolution process, depth-wise convolution retrieves spatial characteristics on each dimension [34].

Convolutional kernel size H is $h \times h$ for the input feature maps I, which have a size of $C_f \times C_f$. N_{inp} indicates the quantity of input channels and N_{out} indicates the quantity of output channels. The output feature map O has a size of $C_g \times C_g$. The definition of a standard convolutional operation is as follows:

$$O_y = \sum_{x=1}^{N_{inp}} I_x \cdot H_x^y + a_y, \quad y = 1,2,..., N_{out}. \tag{9}$$

Where I_x is the $x-th$ map in I, O_x is the $x-th$ map in O, and H_x^y is the $x-th$ portion in the $y-th$ kernel. The bias of the output map O_x is a_y. Moreover, the notation $\cdot$ represents the convolution operator. Assume that, in a typical convolution process, Fp_1 represents the amount of floating-point computations and Tp_1 represents the total number of trainable parameters (ignoring bias parameters). Equations (10) and (11) can be used to compute them:

$$Tp_1 = h \times h \times N_{inp} \times N_{out}, \tag{10}$$

$$Fp_1 = h \times h \times N_{inp} \times N_{out} \times C_g \times C_g \tag{11}$$

The parameter Tp_2 and the floating-point computation Fp_2 for a depth-wise separable convolution process are the total of the depth-wise and 1x1 point-wise convolutions. Tp_2 and Fp_2 can therefore be computed using the methods provided in Equations (12) and (13) respectively:

$$Tp_2 = h \times h \times N_{inp} + N_{inp} \times N_{out}, \tag{12}$$

$$Fp_2 = h \times h \times C_g \times C_g \times N_{inp} + C_g \times C_g \times N_{inp} \times N_{out}. \tag{13}$$

Equations (14) and (15) display the ratios of Equations (10) and (12) and Equations (11) and (13):

$$\frac{Tp_2}{Tp_1} = \frac{1}{N_{out}} + \frac{1}{h^2}, \tag{14}$$

$$\frac{Fp_2}{Fp_1} = \frac{1}{N_{out}} + \frac{1}{h^2}, \tag{15}$$

It is apparent that the depth-wise separable convolution's parameters and computations are just $\frac{1}{N_{out}} + \frac{1}{h^2}$ times larger than those of the conventional convolution. This significantly lowers the model's parameter and computing expense.

3.4.2 Context-aware attention network

A module for attention transfer and a module for context learning make up the proposed context-aware attention network. Each module has three peeks that use completely convolution and sigmoid layers to forecast an attention map and are tuned for convergence using softmax classification loss [35].

$$p(X) = e\left(f(X) \odot g\left(f(X)\right)\right), \tag{16}$$

Where X denotes the input, $\odot$ represents the way the element-wise product works. Having a layer of softmax to further transform the feature vector into probabilities is also included,

$e(\cdot)$ represents fully linked layers that are used to convert convolutional features into feature vector that might be matched the submissions in each category.

Context learning module

Cyber security attack classification relies heavily on context, and studies in computer networks indicate that accurately modeling context might improve attack comprehension and classification algorithms. The creation of a module for context transfer that transmits contextual details in the right, left, down, and up directions is necessary for effective contextual information learning. The process of context transfer can be written as follows:

$$D_{a,b}^{up} = \max\left(V_{a-1,b}^{up}\, D_{a-1,b}^{up} + D_{a,b}^{up}, 0\right) \quad (17)$$

The transmission processing is depicted in the above equation in an upward direction; comparable operations are carried out in the other directions. In equation (17) $D_{a,b}^{up}$ is one of the input map of features cells, and updating it is the aim. $V_{a-1,b}^{up}$ is a transference parameter that has a range of 0 to 1. Rather of being manually set, the parameter $V_{a-1,b}^{up}$ is learning. For cyber-attack classification, context feature maps $f(X) = concat\left(D^{Left}, D^{Right}, D^{Up}, D^{Down}\right)$ comprise both transmitted and original convolution features.

Attention transfer module

The method creates an attention transfer model generating attention maps through several looks, each containing a unique attention region demonstrating reasoning relations between these regions.Maps with context feature $f(X)$ are produced by the indicated module for context learning and input into the module for attention transfer to produce the predicted attention map.

$$\begin{aligned} EN_t(X) &= EN_{t-1}(X) * \left(1 - AN_{t-1}(X)\right) \\ AN_t(X) &= l(EN_t(X)) \end{aligned} \quad (18)$$

Where the $t-th$ glimpses created attention map is $AN_t(X)$, and the input feature maps are shown by $EN_t(X)$. The attention weight of every input pixel appears on an attention map that the network creates a pixel-by-pixel mask. An inhibition approach is applied for every peek, producing three attention maps from three snapshots, each of which represents a distinct attention zone. Following classification, the neural network is input into an optimization phase wherein its parameters are changed to improve accuracy and performance. By ensuring that the model converges to the most accurate response, optimization raises the model's overall effectiveness and detection capacity.

3.5 Boosted Sooty Tern Optimization (BSTO)

Sooty terns, also known as *Onychoprion fuscatus*, are sea birds with diverse species. They are omnivorous birds that eat various animals, including insects, reptiles, amphibians, fish, and earthworms. They are colonial creatures that locate and hunt prey with intelligence. Sooty terns migrate seasonally to find abundant food sources, grouping together to avoid collisions [36]. They use a flapping mode in flight for air attacks, updating initial positions based on the fittest found sooty tern.Effective error rate minimization is achieved by the use of BSTO. Figure 3 shows the flowchart of BSTO.

$$Fitness\ function = Min(MSE)$$

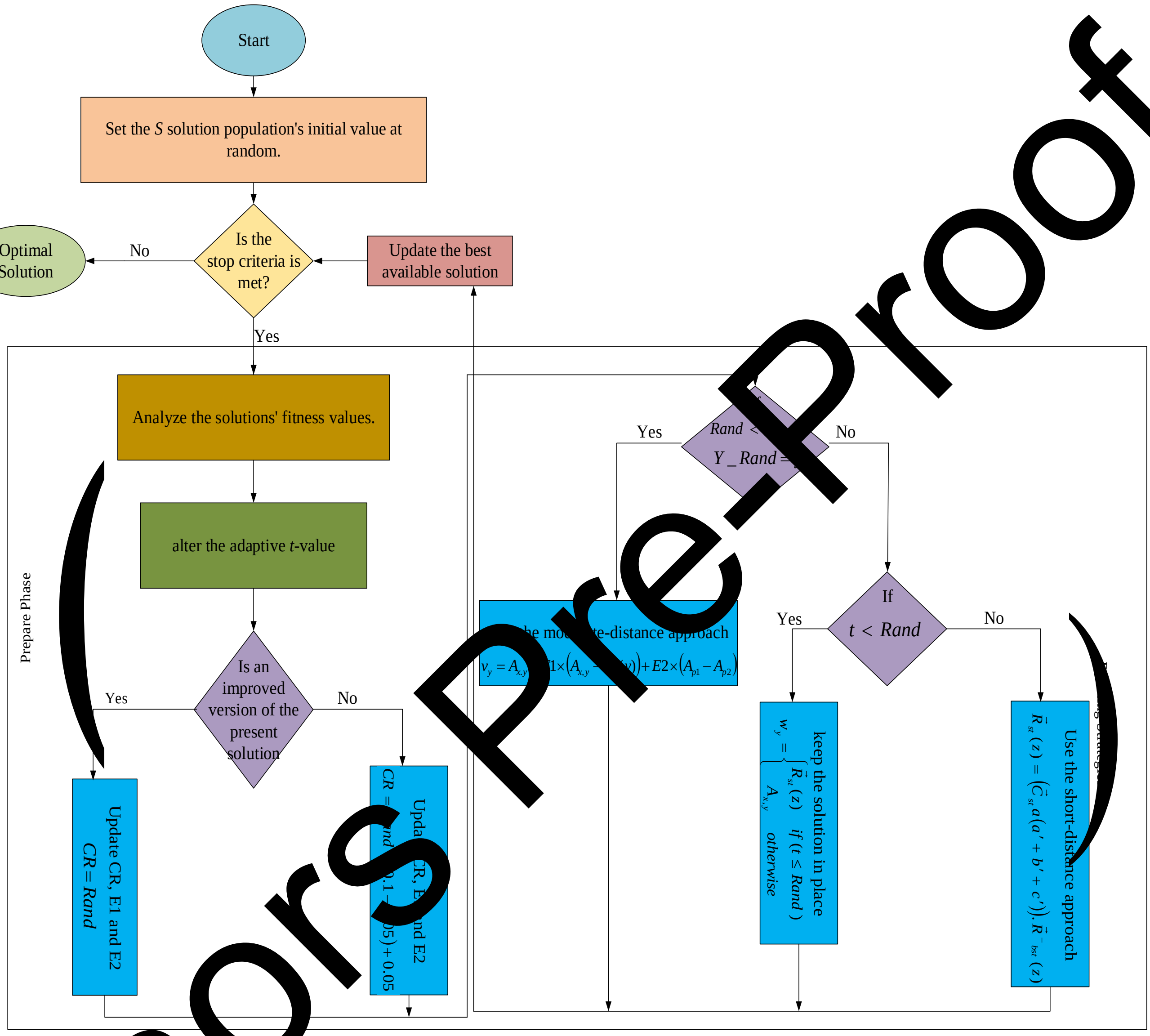


Figure 3: Flowchart of the BSTO

The process for classifying cyber security attacks stated in this proposed method starts with pre-processing the data using min-max scaler normalization. Important features are then selected from the complete set utilizing advanced hybrid optimization techniques. The cyber security attack is then classified by running these chosen features through a CA-DSCNN. To raise efficiency and accuracy in the classification of cyber-security attacks, the BSTO is utilized.

4. Results and Discussion

This section compares the proposed method with the existing approaches using the UNSW-NB15 and CICIDS-2017 datasets. Additionally the method uses data from the IoT- based wireless sensor networks. Regarding validated performance, the proposed method attains superior accuracy, ultra precision, flawless recall, and MAPE, RMSE and MSE efficiency, network lifetime, end-to-end delay, packer delivery ratio (PDR), and throughput and fault tolerance. When it comes to classifying cyber-security attacks into many classes and managing unbalanced datasets, the proposed method regularly performs better than standard models. Incorporating pre-processing, feature selection, and technique optimization into the model-building process is another way to improve stability and reliability. Furthermore, the approach doesn't suffer from a lack of generalization for the categorization of multi-class cyber security attacks and demonstrates its effectiveness. In general, it computes more quickly and has better detection accuracy than the earlier models.Python is used to execute the proposed method.

4.1 Dataset Description

- UNSW-NB15 dataset [37]

The UNSW-NB15 dataset encompasses 10 classes (Normal, Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, and Worms) and contains 42 characteristics (labels excluded).More accurate data for assessing cyber-attack detection systems is intended to be provided by the dataset. With 82,332 examples in the testing set (which includes both attack and normal data), the training set has 175,341 occurrences. The UNSW-NB15 dataset has certain limitations, even though it provides better coverage than its predecessors. These include a small number of network assaults and some obsolete packet information.A comparison of the different forms of data in Table 2 shows the distribution of data from the UNSW-NB15 dataset.

Table 2: Distributed data for the UNSW-NB15 dataset

Data Types	**Description**	**Number of records**
Normal	Typical network information	2,218,761
Fuzzers	Utilizing data feeding that is created at random to suspend applications	24,246
Analysis	Includes assaults such port scans, spam, and HTML page penetration.	2677
Backdoors	Method for getting around system security	2329
DoS	Denial of service attack	16,353
Exploits	Making use of the acknowledged security flaws	44,525
Generic	Method that attacks every block cipher	215,481
Reconnaissance	Attack-simulating strikes to obtain information	13,987
Shellcode	Snippet of code used to take advantage of software vulnerabilities	1511
Worms	In order to infect other computers, worms duplicate themselves.	174

- CICIDS2017 dataset [38]

The Canadian Institute for Cyber-security created the dataset. The dataset includes some modern multi-stage attacks, including DoS assaults and Heartbleed. A range of contemporary protocols are also included. CICIDS2017 simulates seven different attack families, including brute force, heart bleed, botnet, denial-of-service, web, and infiltration attacks. It is designed for use in intrusion detection and network security applications.A comparison of the different forms of data in Table 3 shows the distribution of data from the CICIDS-2017 dataset.

Table 3: Distributed data for the CICIDS-2017 dataset

Data Types	Description	Number of records
Normal	Typical network information	2,358,036
Brute Force Attack	Attempt to guess FTP passwords using a brute force attack.	7938
Heart Bleed Attack	Employing openSSL exploits to inject malicious data into openSSL memory	11
Botnet	Use of the victim system in the Botnet network and trojan-based attacks	1966
Denial-of-Service (DoS)	Excessive use of HTTP get requests in order to limit HTTP use	5499
Web Attack	Using a brute force method to extract personal ID numbers from webpages	1707
Infiltration Attack	Unauthorized access to the system through the use of instruments and penetration techniques	36

- **IoT-based wireless sensor networks data**

IoT-based wireless sensor networks provide the raw data that is utilized to evaluate the proposed method. This dataset captures the intricacies of network traffic and device interactions, encompassing a broad spectrum of attack scenarios pertinent to IoT systems. A thorough evaluation of the approach's effectiveness in identifying and categorizing different cyber security risks unique to IoT-based contexts is made possible by the utilization of IoT-based sensor data.

4.2 Performance comparison with existing approaches

Performance comparison on the UNSW-NB15 Dataset

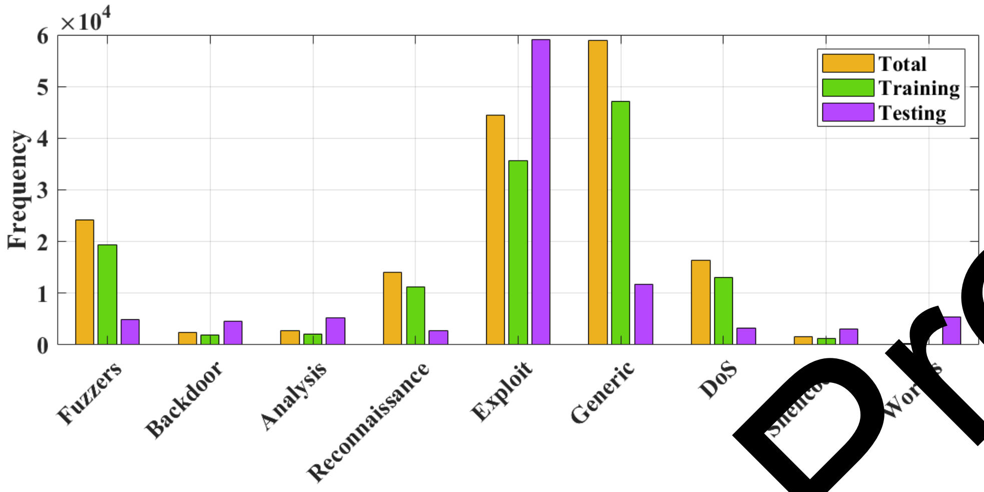


Figure 4: Distribution of Attack Frequencies in the UNSW-NB15 Dataset

Figure 4 shows the prevalence of various attack types, such as backdoors, fuzzers, exploits, and reconnaissance, across all training and testing sets. Initially, the categories "Exploit" and "Generic" are shown with comparatively higher frequency, particularly in the training set. It helps to perceive the distribution of attacks in different phases of the model's development.

Table 4: Data Distribution for Training, Testing, and Validation Sets on UNSW-NB15 dataset

Label	Training Dataset	Testing Dataset	Validation Dataset
Normal data	1,014,221	289,777	144,899
Attack data	157,748	45,071	22,535

Table 4 provides statistical information on normal and attack data instances in the UNSW-NB15 dataset's training, test, and validation sets. The training set consists of 1,014,221 normal training data records and 157,748 attack data records. There are 289,777 records of routine testing and 45,071 records of attacks in the testing set. The validation set consists of 22,535 assault data records and 144,889 normal validation data records.

Table 5: UNSW-NB15 dataset performance evaluation outcomes

Methods	Accuracy(%)	Precision(%)	Recall(%)	F1-Score (%)	Detection rate (%)
DNN	98.8	97.94	97.86	98.76	97.92
CNN	99.47	99.43	99.46	99.44	98.65
SVM	75.21	99.16	75.21	76.60	80.12
RF	99.30	99.09	99.30	99.12	98.51
NB	98.86	99.01	98.86	98.85	97
ANN	99.28	99.37	99.28	99.17	98.02
Proposed CA-DSCNN	99.51	99.49	99.51	99.46	99.33

Table 5 presents a comparison of the performance evaluation results for various methods on the UNSW-NB15 dataset. In comparison with existing models, the proposed CA-DSCNN performs better. CA-DSCNN is the best at classifying multi-class cyber-security attacks, with the highest accuracy (99.51%), precision (99.49%), recall (99.51%), F1-score (99.46%), and detection rate (99.33%).

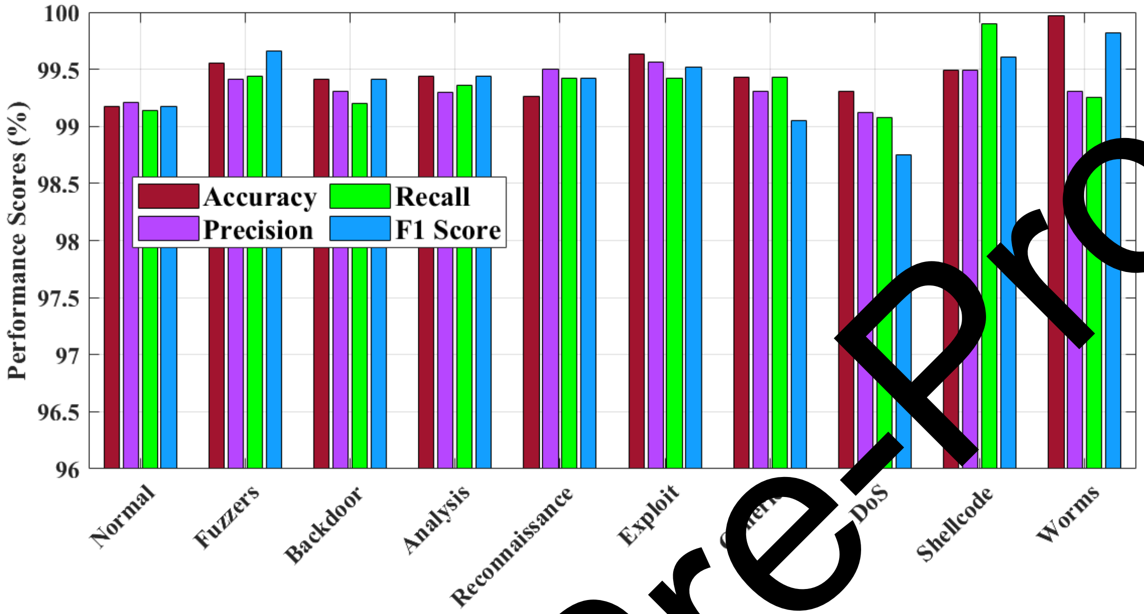


Figure 5: Performance Measures for Classifying Cyber Security Attacks into Multiple Classes Using UNSW-NB15 Dataset

Figure 5 presents the performance metrics of a multi-class cyber security attack classification system in terms of many classes, such as DoS Attack Category, Shellcode, etc., as well as regular traffic classes include accuracy, precision, recall, and F1 score. Every indicator displays high values, with the majority exceeding 98%, suggesting that the model is accurate in characterizing and classifying various cyber-attacks.

Performance comparison on the CICIDS 2017 Dataset.

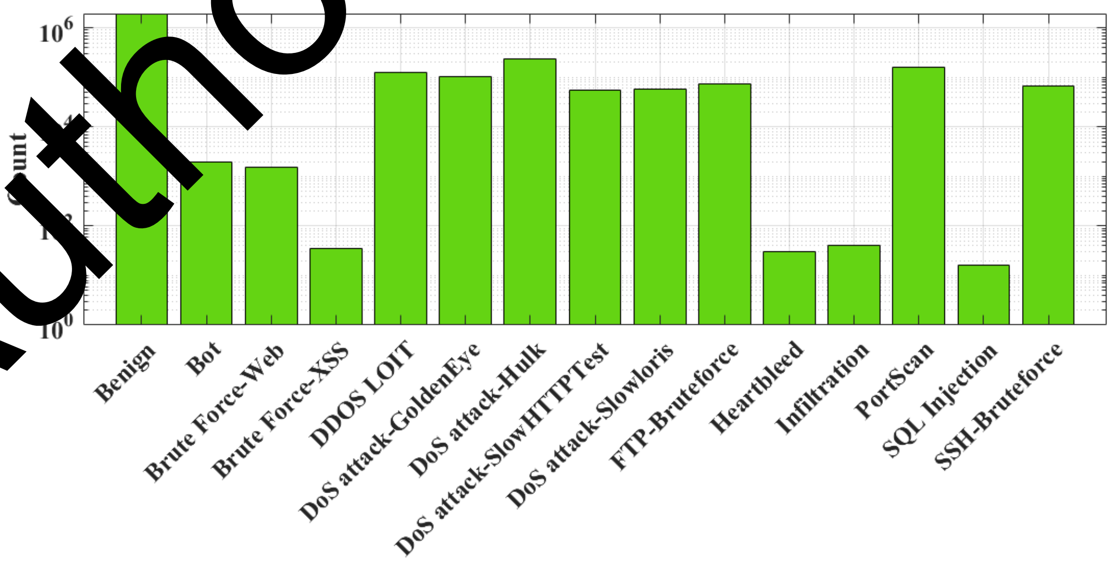


Figure 6: Cyber-security Attack Frequency in the CICIDS 2017 Dataset.

The figure 6 depicts the frequency of cyber security attacks in the CICIDS 2017 dataset on a logarithmic scale. This enables us to clearly see the representation of these types of attacks in relation; for example, "Benign," ”Bot,” and "Dos attack-Hulk“are all included, making it available for cyber-- security analysis. This visualization enables us to see and prioritize a selection of the most common cyber-attacks. Moreover, it emphasizes the necessity of focusing on both ordinary and rare attack vectors in order to guarantee strong security measures.

Table 6: Data Distribution for Training, Testing, and Validation Sets on CICIDS-2017 dataset

Label	Training Dataset	Testing Dataset	Validation Dataset
Normal data	318,014	90,861	45,431
Attack data	7,800	2,229	1,114

Table 6 provides statistics on normal and attack data instances from the CICIDS2017 dataset's training, test, and validation sets. There are 7,800 assault data records and 318,014 regular training data records. There are 2,229 test set-based attack data records and 90,861 normal testing data records. There are 1,114 validation set-based attack data records and 45,431 normal validation data records.

Table 7: CICIDS-2017 dataset performance evaluation outcomes

Methods	Accuracy	Precision	Recall	F1-Score	Detection rate
DNN	97.02	96.[illegible]	96.6	96	92.80
CNN	98.22	98.23	98.21	98.20	94.65
SVM	73.41	96.78	73.99	74.55	75.88
RF	98.15	97.88	98.66	98.54	97.64
NB	96.78	96	96.68	96.58	94
ANN	98.49	98.5[illegible]	98.60	98.11	97.66
Proposed CA-DSCNN	99.48	99.23	99.15	99.66	99.13

Table 7 shows the CICIDS-2017 dataset performance evaluation outcomes. The higher accuracy (99.48%) and F1-Score (99.66%) are attained by the proposed CA-DSCNN, which performs better than the existing approaches. Additionally, it outperforms techniques like CNN and ANN in terms of precision (99.23%) and recall (99.15%). Its effectiveness and reliability are demonstrated by the 99.13% detection rate, which emphasizes the way well it performs in identifying situations when compared to other models.

4.3 Comparative analysis of the proposed method's performance with existing approaches

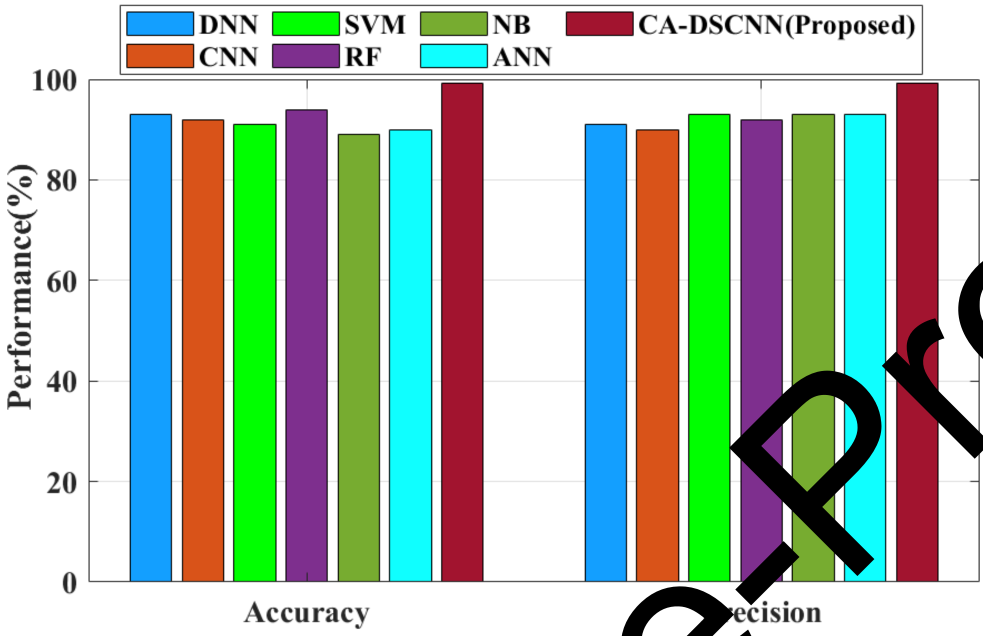


Figure 7: accuracy and precision performance comparison between proposed and existing methods

The performance of the following cyber security attack models is compared in the figure 7: DNN, SVM, CNN, RF, NB, ANN, and CA-DSCNN (proposed). It showsshows98% accuracy and almost 97% precision, with colored bars for each model. In terms of both measures, the proposed CA-DSCNN model performs similarly to alternative models.

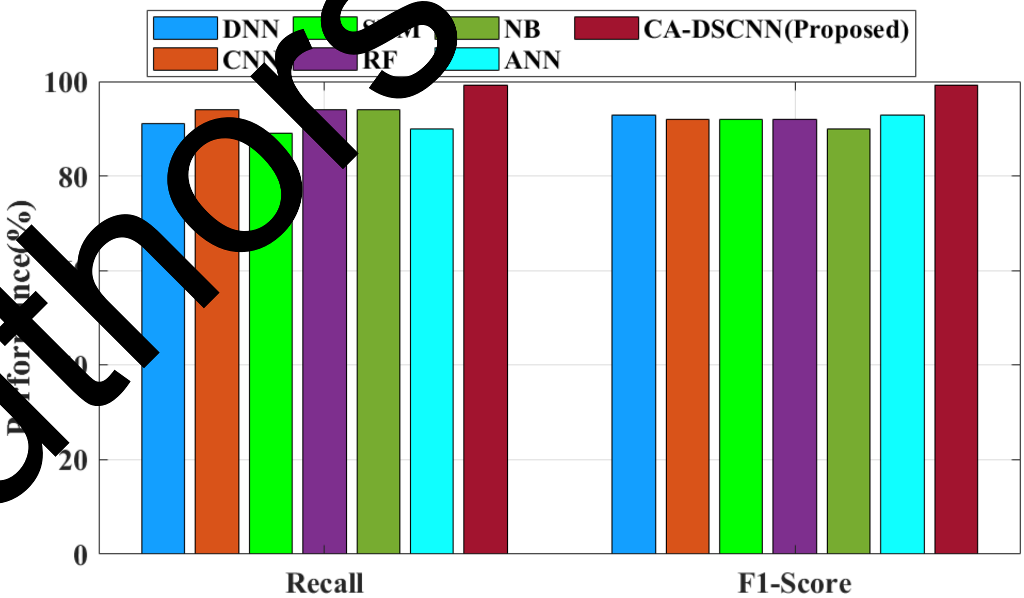


Figure 8: Recall and F1-Score performance comparison between proposed and existing methods

Figure 8 shows the Recall and F1-Score performance comparison between proposed and existing methods. Various algorithms, such as DNN, SVM, NB, CNN, RF, ANN, and the proposed CA-DSCNN, are compared with respect to recall and F1-Score performance. The proposed CA-DSCNN achieves the highest recall (99%) and F1-score (99.5%). However, other models have F1-Score and recall values that range from 80% to 95%, indicating that the CA-DSCNN technique performs better in cyber-security attack classification tasks.

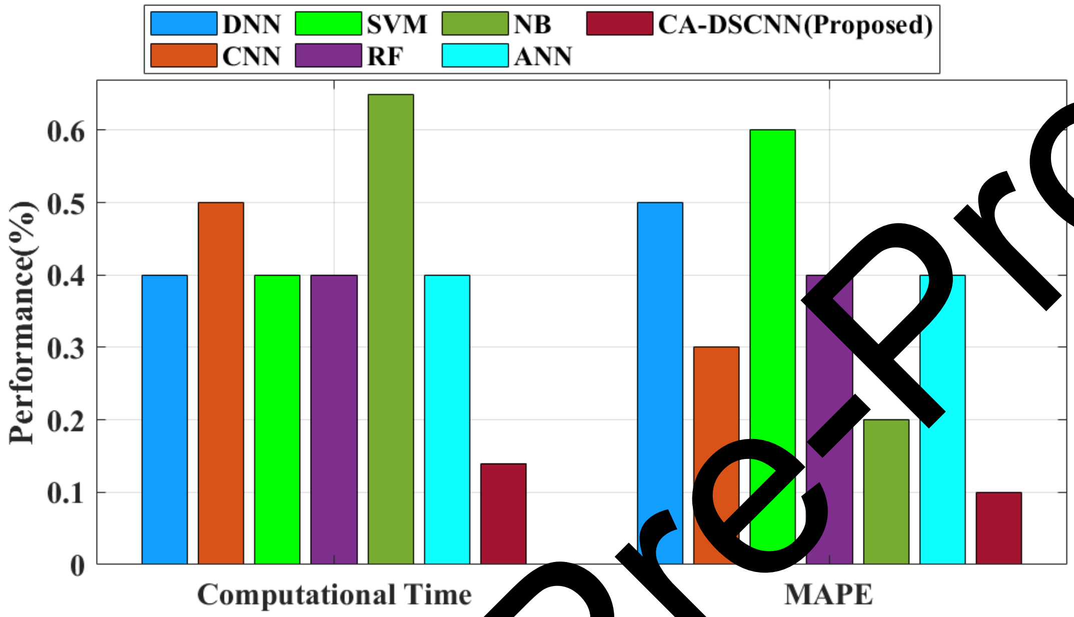


Figure 9: Computational time and MAPE performance comparison between proposed and existing methods

The Mean Absolute Percentage Error (MAPE) and computational time metrics are used in the figure 9 to compare the existing cyber-attack methods. Although it does not have the shortest computation time (1.0s), the CA-DSCNN (Proposed) model has the lowest MAPE 5%, suggesting the maximum accuracy. A variety of other models exhibit different performance levels, including the DNN Computational Time ~ 0.8s, MAPE ~ 12%, SVM Computational Time ~ 1.2s, MAPE ~ 15%, NB Computational Time ~ 0.7s, MAPE ~ 18%, RF Computational Time ~ 0.9s, MAPE ~ 10%, and ANN Computational Time ~ 0.85s, MAPE ~ 14%.

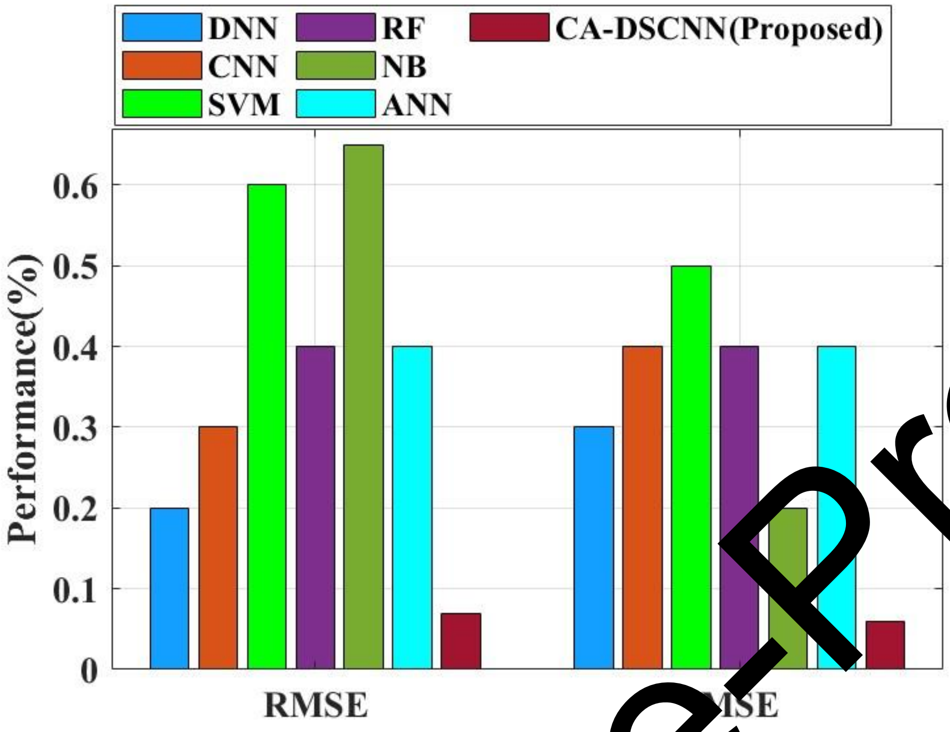


Figure 10: RMSE and MSE performance comparison between proposed and existing methods

The following algorithms' performances are compared in the figure 10: DNN, RF, CNN, NB, SVM, and ANN. With an RMSE of 0.8% and an MSE of 0.6%, the CA-DSCNN (proposed) method performs the best. RMSE values consistently exceed MSE for every algorithm, indicating a larger degree of error in RMSE.

Table 8: Comparing Cyber Security Detection Techniques' Performance in IoT based Wireless Sensor Networks

Methods	Network Lifetime	End-to-End Delay	Packer Delivery Ratio (PDR)	Throughput	Fault Tolerance
EESC-SSP [39]	3[illegible] hours	120 ms	92%	200 kbps	High
HR-MOBSO-IDS [40]	2[illegible] hours	110 ms	90%	190 kbps	Medium
SG-IDS [41]	25 hours	130 ms	88%	180 kbps	High
ESWI [42]	32 hours	115 ms	91%	210 kbps	High
SSP-WSN [43]	29 hours	105 ms	93%	195 kbps	High
Proposed	35 hours	100 ms	99%	220 kbps	Very High

The table 8 compares various cyber security detection methods for IoT-based wireless sensor networks across five key metrics: network lifetime, end-to-end delay, packet delivery ratio, throughput, and fault tolerance. The results show that the proposed method is superior to other methods both in security and performance metrics, including longest network lifetime

(35 hours), lowest delay (100 ms), highest PDR (95%), best throughput (220 kbps), and superior fault tolerance.

5. Conclusion

The use of cutting-edge methodologies has greatly improved threat detection's accuracy and efficiency in the field of multi-class cyber-security attack categorization. Pre-processing has used Min-Max scaler normalization through reshaping of the data in order to enhance the contribution rate of the features in the classification process, thus enhancing the performance of the model. BTGO, along with an ASO for the feature selection process, has enhanced the degree of relevance of the input features; these modifications improve the models' accuracy and resilience. CA-DSCNN has been employed to identify more complex patterns of relations between different data elements as well as more effectively classify these patterns by minimizing the number of computations. Also, BSTO has been used to optimize the model parameters and enhance the classification accuracy of the result. In general, the use of these methodologies has contributed to the development of a diverse and efficient way of categorizing various types of cyber-security attacks. The above-mentioned methods have done well in enhancing detection performance and thereby presented a good avenue for further research and extension of practical use in the future. Future work will expand datasets to encompass a wider variety of attack types and real-world scenarios, which will facilitate the creation of more broadly applicable models.

References:

1. Syed, N.F., Baig, Z., Ibrahim, A. and Valli, C., 2020. Denial of service attack detection through machine learning for the IoT. *Journal of Information and Telecommunication*, *4*(4), pp.482-503.
2. Kim, K., Alfouzan, F.A. and Kim, H., 2021. Cyber-attack scoring model based on the offensive cybersecurity framework. *Applied Sciences*, *11*(16), p.7738.
3. Dutta, V., Choraś, M., Pawlicki, M. and Kozik, R., 2020. A deep learning ensemble for network anomaly and cyber-attack detection. *Sensors*, *20*(16), p.4583.
4. Goyel, H. and Swarup, K.S., 2022. Data integrity attack detection using ensemble-based learning for cyber–physical power systems. *IEEE Transactions on Smart Grid*, *14*(2), pp.1198-1209.
5. Almalaq, A., Albadran, S. and Mohamed, M.A., 2022. Deep machine learning model-based cyber-attacks detection in smart power systems. *Mathematics*, *10*(15), p.2574.
6. Arunkumar, M. and Ashok Kumar, K., 2022. Malicious attack detection approach in cloud computing using machine learning techniques. *Soft Computing*, *26*(23), pp.13097-13107.
7. Aktar, S. and Nur, A.Y., 2023. Towards DDoS attack detection using deep learning approach. *Computers & Security*, *129*, p.103251.
8. Wan, Y. and Dragičević, T., 2022. Data-driven cyber-attack detection of intelligent attacks in islanded dc microgrids. *IEEE Transactions on Industrial Electronics*, *70*(4), pp.4293-4299.
9. Akgun, D., Hizal, S. and Cavusoglu, U., 2022. A new DDoS attacks intrusion detection model based on deep learning for cybersecurity. *Computers & Security*, *118*, p.102748.
10. Alsaade, F.W. and Al-Adhaileh, M.H., 2023. Cyber attack detection for self-driving vehicle networks using deep autoencoder algorithms. *Sensors*, *23*(8), p.4086.
11. Gopalakrishnan, T., Ruby, D., Al-Turjman, F., Gupta, D., Pustokhina, I.V., Pustokhin, D.A. and Shankar, K., 2020. Deep learning enabled data offloading with cyber attack

detection model in mobile edge computing systems. *IEEE Access*, *8*, pp.185938-185949.

12. Vu, L., Nguyen, Q.U., Nguyen, D.N., Hoang, D.T. and Dutkiewicz, E., 2020. Deep transfer learning for IoT attack detection. *IEEE Access*, *8*, pp.107335-107344.
13. Habibi, M.R., Baghaee, H.R., Blaabjerg, F. and Dragičević, T., 2021. Secure MPC/ANN-based false data injection cyber-attack detection and mitigation in DC microgrids. *IEEE Systems Journal*, *16*(1), pp.1487-1498.
14. Aouedi, O., Piamrat, K., Muller, G. and Singh, K., 2022. Federated semisupervised learning for attack detection in industrial Internet of Things. *IEEE Transactions on Industrial Informatics*, *19*(1), pp.286-295.
15. Khosravi, M. and Ladani, B.T., 2020. Alerts correlation and causal analysis for APT based cyber attack detection. *IEEE Access*, *8*, pp.162642-162656.
16. Chen, S., Wu, Z. and Christofides, P.D., 2020. Cyber-attack detection and resilient operation of nonlinear processes under economic model predictive control. *Computers & Chemical Engineering*, *136*, p.106806.
17. Crespo-Martínez, I.S., Campazas-Vega, A., Guerrero-Higueras, Á.M., Riego-DelCastillo, V., Álvarez-Aparicio, C. and Fernández-Llamas, C., 2023. SQL injection attack detection in network flow data. *Computers & Security*, *127*, p.103093.
18. Kravchik, M. and Shabtai, A., 2021. Efficient cyber attack detection in industrial control systems using lightweight neural networks and pca. *IEEE transactions on dependable and secure computing*, *19*(4), pp.2179-2190.
19. Farrukh, Y.A., Ahmad, Z., Khan, I. and Elavarasan, R.M., 2021, November. A sequential supervised machine learning approach for cyber attack detection in a smart grid system. In *2021 North American Power Symposium (NAPS)* (pp. 1-6). IEEE.
20. Tummala, A.S. and Inapakurthi, R.K., 2021. A two-stage Kalman filter for cyber-attack detection in automatic generation control system. *Journal of Modern Power Systems and Clean Energy*, *10*(1), pp.50-59.
21. Jia, Y., Gu, Z., Du, L., Long, Y., Wang, Y., Li, J. and Zhang, Y., 2023. Artificial intelligence enabled cyber security defense for smart cities: A novel attack detection framework based on the MDATA model. *Knowledge-Based Systems*, *276*, p.110781.
22. Semwal, P. and Handa, A., 2022. Cyber-attack detection in cyber-physical systems using supervised machine learning. *Handbook of Big Data Analytics and Forensics*, pp.131-140.
23. Prabakar, D., Sundarrajan, M., Manikandan, R., Jhanjhi, N.Z., Masud, M. and Alqhatani, A., 2023. Energy analysis-based cyber attack detection by IoT with artificial intelligence in a sustainable smart city. *Sustainability*, *15*(7), p.6031.
24. Balta, E.C., Pease, M., Moyne, J., Barton, K. and Tilbury, D.M., 2023. Digital twin-based cyber-attack detection framework for cyber-physical manufacturing systems. *IEEE Transactions on Automation Science and Engineering*, *21*(2), pp.1695-1712.
25. Li, Q., Zhang, J., Zhao, J., Ye, J., Song, W. and Li, F., 2022. Adaptive hierarchical cyber attack detection and localization in active distribution systems. *IEEE transactions on smart grid*, *13*(3), pp.2369-2380.
26. Salam, A., Ullah, F., Amin, F. and Abrar, M., 2023. Deep Learning techniques for web-based attack detection in industry 5.0: a novel approach. *Technologies*, *11*(4), p.107.
27. Jullian, O., Otero, B., Rodriguez, E., Gutierrez, N., Antona, H. and Canal, R., 2023. Deep-learning based detection for cyber-attacks in IoT networks: A distributed attack detection framework. *Journal of Network and Systems Management*, *31*(2), p.33.

28. Raghunath, K.K., Kumar, V.V., Venkatesan, M., Singh, K.K., Mahesh, T.R. and Singh, A., 2022. XGBoost regression classifier (XRC) model for cyber attack detection and classification using inception V4. *Journal of Web Engineering*, *21*(4), pp.1295-1322.
29. Saghezchi, F.B., Mantas, G., Violas, M.A., de Oliveira Duarte, A.M. and Rodriguez, J., 2022. Machine learning for DDoS attack detection in industry 4.0 CPPSs. *Electronics*, *11*(4), p.602.
30. Alaca, Y. and Çelik, Y., 2023. Cyber attack detection with QR code images using lightweight deep learning models. *Computers & Security*, *126*, p.103065.
31. Deepa, B. and Ramesh, K., 2022. Epileptic seizure detection using deep learning through min max scaler normalization. *Int. J. Health Sci*, *6*, pp.10981-10996.
32. Wu, X., Zhou, W., Fei, M., Du, Y. and Zhou, H., 2024. Banyan tree growth optimization and application. *Cluster Computing*, *27*(1), pp.411-441.
33. Khurma, R.A., Albashish, D., Braik, M., Alzaqebah, A., Qasem, A. and Adwan, O., 2023. An augmented Snake Optimizer for diseases and COVID-19 diagnosis. *Biomedical Signal Processing and Control*, *84*, p.104718.
34. Dang, L., Pang, P. and Lee, J., 2020. Depth-wise separable convolution neural network with residual connection for hyperspectral image classification. *Remote Sensing*, *12*(20), p.3408.
35. Zhang, J., Ren, J., Zhang, Q., Liu, J. and Jiang, X., 2023. Spatial context-aware object-attentional network for multi-label image classification. *IEEE Transactions on Image Processing*, *32*, pp.3000-3012.
36. Houssein, E.H., Oliva, D., Celik, E., Emam, M.M. and Ghoniem, R.M., 2023. Boosted sooty tern optimization algorithm for global optimization and feature selection. *Expert Systems with Applications*, *213*, p.119015.
37. Al-Daweri, M.S., ZainolAriffin, K.A., Abdullah, S. and Md. Senan, M.F.E., 2020. An analysis of the KDD99 and UNSW-NB15 datasets for the intrusion detection system. *Symmetry*, *12*(10), p.1666.
38. Dube, R., 2024. Faulty use of the CIC-IDS 2017 dataset in information security research. *Journal of Computer Virology and Hacking Techniques*, *20*(1), pp.203-211.
39. Krishnan, R., Krishnan, R.S., Robinson, Y.H., Julie, E.G., Long, H.V., Sangeetha, A., Subramanian, M. and Kumar, R., 2022. An intrusion detection and prevention protocol for internet of things based wireless sensor networks. *Wireless Personal Communications*, *124*(4), pp.3461-3483.
40. Subramani, S. and Selvi, M., 2023. Multi-objective PSO based feature selection for intrusion detection in IoT based wireless sensor networks. *Optik*, *273*, p.170419.
41. Saleh, H.M., Marouane, H. and Fakhfakh, A., 2024. Stochastic gradient descent intrusions detection for wireless sensor network attack detection system using machine learning. *IEEE Access*.
42. Shahid, H., Ashraf, H., Javed, H., Humayun, M., Jhanjhi, N.Z. and AlZain, M.A., 2021. Energy optimised security against wormhole attack in iot-based wireless sensor networks. *Comput. Mater. Contin*, *68*(2), pp.1967-81.
43. Panahi, U. and Bayılmış, C., 2023. Enabling secure data transmission for wireless sensor networks based IoT applications. *Ain Shams Engineering Journal*, *14*(2), p.101866.