Deep Learning based Feature Fusion Network for Long Range Attack Detection on Blockchain Consensus Layer

Ritika Shrimali[1], Dr. Pughazendi Narayanan[2]

[1]*Assistant Professor, Department of Computer Science and Engineering, School of Science Studies, CMR University, India*
[2]*Professor and Head, School of Science and Computer Studies, CMR University, India*
Email: shrimali.ritika0@gmail.com

Abstract: A blockchain is recognized as a revolutionary and advanced technology, primarily due to its features of privacy, security, immutability, and data integrity. The consensus layer serves as the foundation and it is the most critical component of blockchain architecture. Identifying Long-Range Attacks (LRA) within a block chain presents significant challenges. Existing studies face various difficulties in detecting these long-range attacks and monitoring the behaviour of validator nodes within the blockchain network. Consequently, this paper introduces a novel deep learning approach designed to accurately detect the nodes as either normal or attack, thereby effectively reducing the risk of long-range attacks. Initially, data are collected, and pre-processing is done to improve the quality of input using Upgraded Min-Max Normalization. Next, high level features are extracted using Improved Non-negative Matrix Factorization (INMF) and Sparse Variational Auto encoder (SVAE) methods. The INMF based features are given as input to the Depth wise Separable Convolutional Resnet (DSC-ResNet) to learn the latent features. The Stacked Bidirectional Gated Recurrent Dropout Network (SBi-GRDN) is trained using the SVAE features to identify complex relationships and interactions among features for capturing attack patterns efficiently. Then, the attention layer is used to fuse features from the DSC-ResNet and SBi-GRDN models. Finally, a fully connected layer with a sigmoid is employed for classifying the attack. The experimental results illustrate that the proposed approach attains an accuracy of 99.1%, Precision of 99.5%, Recall of 99.4%, and F1-score of 99.5%, which provides effective results in detecting long range attacks.

Keywords: Consensus layer, Long range attack, Normalization, Factorization, Auto encoder, Dropout network.

1. Introduction

Blockchain is intended to create a secure distributed database, eliminating the need for a central governing authority [1-3]. It has emerged as a decade-defining technology, with applications in healthcare, voting, supply chain management, banking, the Internet of Vehicles (IoV), art, and more. The information stored in the blockchain is accessible to all participants and is known as nodes. The process of validating and authenticating data on a blockchain is managed by a mechanism referred to as the consensus mechanism [4]. In conventional blockchains, participants engage in competition to solve complex cryptographic and mathematical challenges that are straightforward to verify [5-6]. This activity is referred to as "mining," and the individual who successfully solves the problem is rewarded with new coins for their efforts. Consequently, these blockchains operate on the principle of Proof of Work (PoW) [7-8]. In this context, users are deemed reliable due to the significant computational resources they have invested in transaction verification. In contrast, Proof of Stake (PoS) protocols select users for transaction validation based on the amount of wealth they hold or their stake [9-10].

The actual power consumption could be significantly higher as numerous malware programs engage in mining activities on compromised devices without the users' awareness [11-12]. Although blockchains are often perceived as inherently secure and privacy-oriented, this belief is misleading since various attacks can target them in multiple ways. Some of these attacks are quite technical and may pertain only to certain blockchains [13-14]. It is important to recognize that while blockchains may be a decade-old technology, the majority of users have only recently begun to engage with them, often viewing them as a "black box." Consequently, it is essential to illuminate the different security aspects of blockchains and to analyze the threats they face [15-16]. This detailed investigation intends to shed light on these models' predictive powers as well as their sensitivity to adversarial data manipulation inside supply chain networks that have been strengthened by blockchain technology. Block chain-enabled supply chain networks are complicated systems with many players, like vendors, producers, traders, sellers, and customers [17-18]. These participants engage in a variety of tasks, including manufacturing, distribution, inventory control, and predicting demand, all of which contribute to the timely and effective delivery of goods and services to customers [19]. Nonetheless, these supply chain networks face a variety of obstacles and dangers, like disruptions, uncertainty, cyber attacks, and environmental concerns. A significant problem is to defend the network's security against malicious actors that may seek to negotiation, destroy, or steal important information or resources from the system. Detecting and blocking such attacks is more difficult in block chain-enabled supply chain networks since data is spread and decentralized across several nodes.

Various strategies can be explored to mitigate these attacks. The safety and effectiveness of block chain-enabled supply chain networks result in more robust and secure network detection systems that are personalized. A successful long-range attack allows a hostile node to change the whole history of transactions recorded on the block chain, compromising the core tenet of immutability that blockchains provide [20]. A long-range attack is considered effective when an illegal validator is able to create a secret or alternate chain that starts from the genesis block and mirrors the genuine main chain. Several techniques are proposed for detecting and countering long-range assaults, including software updates as well as the publication of checkpoints for all nodes. However, these alternatives are mostly theoretical and provide implementation issues without jeopardizing the blockchain system's centralization. Validators who sell their outdated keys to hostile actors open the door to long-range attacks. Thus, this works proposes a novel method for detecting long range attacks by analysing the block and supply chain creation.

1.1 Motivation

Blockchain is a popular distributed database that has attracted global interest and acceptance in a variety of industries. The blockchain allows data to be securely recorded and shared among computers on a network. These transactions are unchangeable and impervious to tampering since they are cryptographically linked and secured by consensus techniques. In fact, one well-known consensus technique utilized in Supply chain Blockchain networks is Proof-of-Stake (PoS). Detecting Long-Range Attacks (LRA) in the supply chain Blockchain is critical for ensuring network security as well as integrity. Long-range attacks refer to attacks where an adversary attempts to rewrite the block chain’s history from a point far back in time. Thus, it is necessary for detecting the long range attacks on the supply chain block chain. Recently, deep learning techniques have become more popular for categorizing the normal and attack activities. This is because the deep learning architectures involve several efficient layers that can learn the significant feature information and classify the malicious and non-malicious nodes in an efficient manner. Motivated from this, the proposed study prefers a novel deep learning model to identify malicious activities for mitigating large range attacks on the supply chain blockchain. The main contributions of the work are,

- To enhance the quality of inputs by transforming all data into a similar form in the pre-processing step using an efficient method of upgraded min-max normalization.
- To extract the high level features from the pre-processed output using Improved Non-negative Matrix Factorization (INMF) and Sparse Variational Autoencoder (SVAE) methods.
- To effectively classify the given blockchain node data as attack or normal, a fully connected layer with sigmoid is employed.
- To assess the performance of the suggested study by utilizing diverse metrics and comparing the results with other existing methods.

The document is summarized as: Section 2 offers related works, Section 3 details the suggested approach, Section 4 illustrates the results as well as discussions, and Section 5 ends with the conclusion and future prospects.

2. Related works

Some of the works related to long-range attack detection with various models are described in the following section.

Sanda et al. [21] introduced a technique for classifying nodes in a PoS blockchain for long range attacks. This study used a publicly available dataset. Initially, a semi-synthetic dataset was generated using node validators as well as block transactions. Then, feature selection was utilized to discover critical properties for creating and validating the PoS blocks. Furthermore, the deep learning model accurately categorized nodes as malicious or non-malicious, allowing it to effectively repel long-range assaults. Finally, the suggested model attained an accuracy of 85.20%.

Chauhdary et al. [22] used deep learning to construct a method for detecting cyberattacks in cloud-based supply chain management (SCM). In this study, the input data is analyzed, the complex data is managed, and the most essential features are identified using Evolution Social Spider Optimization (ESSO). The Deep Belief Network (DBN) was then trained with an Extreme Learning Machine (ELM) to detect and identify the attacks. Additionally, the model improved performance by employing Poor and Rich Optimization (PRO) techniques. However, it was noticed that the model faced issues in training.

T soukas et al. [23] developed an end-to-end solution to improve food supply chain security using TinyML. The transparency of food supply chain monitoring systems ensures the safety of all components in the approach. Then, a universal information monitoring method based on block chain technology secured the security of collected data, while a self-governing method for all supply chain actors reduced the number of points of failure. At last, a security system based on the suggested approach TinyML emerging technology was implanted in monitoring devices to reduce a major amount of illegal conduct by supply chain actors.

Ilyas et al. [24] developed an efficient solution for identifying and combating Distributed Denial of Service (DDoS) attacks in a block chain network based on optimization-based deep learning. Initially, verified responses were provided to authenticated users while suspicious traffic was monitored for DDoS detection. The Poaching Raptor Optimization-based Deep Neural Network (DNN) was then used, with the classifier modified to reduce training loss using a suggested optimization approach. This algorithm hybridized the hunting and poaching behaviors of raptors and Lobos to improve detection accuracy. The suggested technique attained an accuracy of 95.12%.

Bassiouni et al. [25] presented an enhanced DL algorithm for predicting supply chain risks under COVID-19 constraints. The dataset was obtained from the publically available Kaggle. The pre-processing step involved removing columns that reflected the identical features as well as distinctive identifiers. The features were then retrieved using four primary models: RNN, CNN, LSTM, and BiLSTM. Furthermore, the classifiers included SoftMax, random trees (RT), random forest (RF), k-nearest neighbor (KNN), artificial neural network (ANN), and support vector machine (SVM). However, it was noticed that the model faced overfitting problems.

Dai et al. [26] suggested an automated DDoS assault traffic recognition solution based on a cross-multilayer CNN model optimized for the blockchain network layer. This solution addressed the difficulties of low generalization, high erroneous rates, and inefficient recognition that are common with current detection strategies. Initially, the model applied a convolution operation to pre-processed traffic within the blockchain network layer, using a cross-layer method that included L2 regularization. This technique allowed the algorithm to capture intricate aspects of attack traffic at several levels while enhancing the representational capabilities of essential features; specifically, parameters with large volatility were penalized in order to stabilize the model's weight changes. The model retrieved very resilient abstract aspects of attack traffic, increasing its generalizability and reducing misreporting rates. The abstract features were then parametrically encoded using a stacked sparse auto encoder based on Kullback-Leibler divergence, with model sparsity changes applied to eliminate unnecessary data and reduce interdependence across abstract features. However, it was noticed that there was difficulty in classifying the attack.

The application of AI-driven methodologies, especially Deep Learning (DL) techniques, can improve the functionality of Uncrewed Aerial Vehicles (UAVs). However, this advancement raises issues related

to their safety and vulnerability to adversarial threats. Hickling et al. [27] used Deep Reinforcement Learning (DRL) to improve model efficiency. This research described a strategy for using the comprehension of DL approaches to develop an effective detection system to secure certain DL frameworks.Two adversarial attack detection techniques were proposed in this paper. The first was a Convolutional Neural Network Adversarial Detector (CNN-AD), which had 80% detection accuracy. The second system used a Long Short Term Memory (LSTM) network, attaining 91% accuracy with faster processing speeds than the CNN-AD, allowing for real-time detection of adversarial attacks. However, it was noticed that there was a difficulty in training. A summary of the existing methods is provided in Table 1.

Table 1: Summary of the existing methods

Authors & references	Techniques	Description	Limitations
Sanda et al. [21]	PoS- proof of stake	Long range attack detection	Data dependency and Model Complexity.
Chauhdary et al. [22]	SCM, ESSO, DBN, ELM and PRO	Attack prediction in supply chain	Dynamic threat environments and over fitting.
T soukas et al. [23]	TinyML	Food supply chain security	Complex integration and Data privacy concerns.
Ilyas et al. [24]	Poaching raptor based optimization	Prevention of DDoS attacks	Ethical and privacy issues and scalability.
Bassiouni et al. [25]	RNN, CNN, LSTM, BiLSTM, KNN, RF and ANN	Predict supply chain risks under COVID-19	Ethical and social considerations, and overfitting.
Dai et al. [26]	Kullback leibler divergence	Improves the overall reliability	Difficulty in distinguishing the classes
Hickling et al. [27]	Convolutional neural network-Adversarial detector	Enhances the security	Difficulty in training

Problem statement

The increasing occurrence of long-range attacks targeting blockchain consensus layers presents a serious risk to the integrity and security of blockchain systems. These attacks take advantage of weaknesses within the consensus mechanism, undermining the trust and reliability of the entire network. Current detection methods often fall short in addressing the complexity and dynamic characteristics of these attacks, as they struggle to effectively capture intricate, non-linear interactions and relationships among various features. This shortcoming results in diminished accuracy and efficiency in detecting attacks, especially when dealing with sophisticated and covert attack strategies. To tackle these issues, a robust, scalable, and intelligent solution is necessary to extract significant high-level features and to accurately identify complex attack patterns. The proposed model confronts these challenges by incorporating advanced deep learning methodologies, such as Improved Non-negative Matrix Factorization (INMF) and Sparse variational Auto encoder (SVAE), for thorough feature extraction. Utilizing a Depth wise Separable Convolutional ResNet (DSC-ResNet) alongside a Stacked Bidirectional Gated Recurrent Dropout Network (SBi-GRDN), the model captures latent and sequential feature interactions. The integration of these features through an attention mechanism further boosts detection capabilities, resulting in accurate classification via a fully connected layer with sigmoid activation. This innovative approach significantly enhances the detection of long-range attacks, providing improved accuracy, adaptability, and scalability, thus reinforcing the resilience of blockchain consensus mechanisms.

3. Proposed methodology

A novel method is developed to detect and classify the attack on blockchain consensus layer. In order to mitigate the long range attack on the consensus layer, various stages are performed, like pre-processing, feature extraction, feature fusion as well as classification. Figure 1 describes the workflow of the proposed approach.

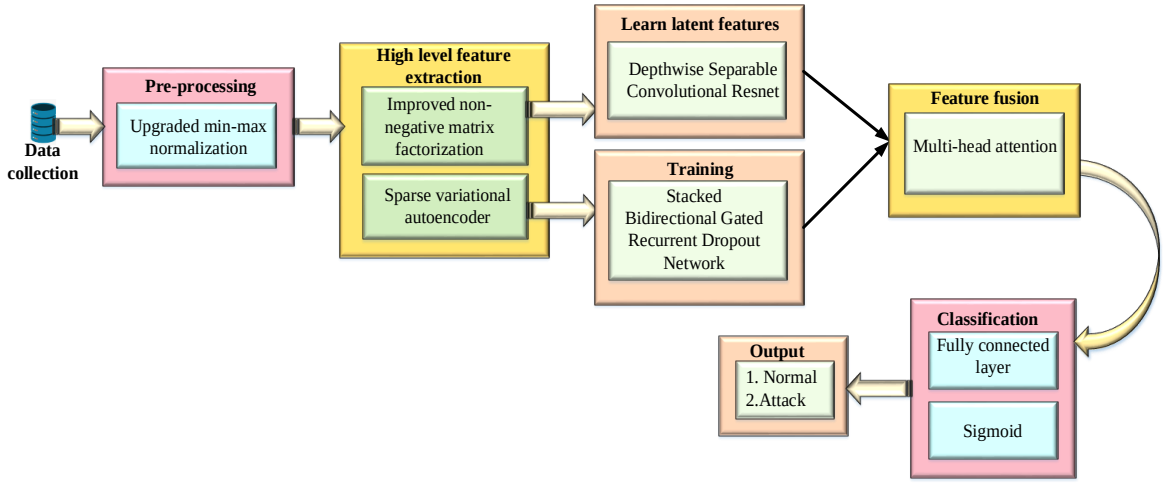

Figure 1: Block diagram of the proposed approach

Initially, data are collected from the proof of stake blockchain dataset. Then, pre-processing phase aims at improving input quality through Upgraded Min-Max Normalization. After that, high-level features are recovered from pre-processed data using Improved Non-negative Matrix Factorization (INMF) as well as Sparse Variational Autoencoder (SVAE) methods. Next, in the construction of the proposed model, features derived from INMF are fed into the Depthwise Separable Convolutional ResNet (DSC-ResNet) to facilitate the learning of latent features. Subsequently, the Stacked Bidirectional Gated Recurrent Dropout Network (SBi-GRDN) is trained with the SVAE features to uncover complex relationships and interactions among the features, which enhances the model's ability to efficiently capture attack patterns. An Attention layer is then implemented to fuse features from both the DSC-ResNet and SBi-GRDN models. Finally, a fully linked layer with a sigmoid activation function is used to classify attacks.

3.1 Pre-processing

Pre-processing is used to reduce the noise and increase the quality of data. Pre-processing is used to improve the quality of input data by enhancing min-max normalization.

- Upgraded min max normalization

Min-max normalization presents significant challenges due to its lack of robustness, which is particularly sensitive to high-value data used in the estimation process. To address this limitation, an upgraded min-max normalization is introduced to preserve the original distribution of matching scores, with the exception of applying a scaling factor to normalize all high scores. For instance, if the range of the first score is [20, 85] and the range of the second score is [35, 100], it is advisable to utilize the raw scores rather than the normalized ones, as the ranges of both scores are nearly identical. In another scenario, where the first score ranges from [20, 85] and the second score ranges from [0.25, 0.75], normalization of the first score is warranted since it is approximately 100% greater than the second score set. The normalization process can be executed using the appropriate formula.

$$Y' = \frac{y - \min(y)}{mean(y^*) + std(y^*) - \min(y)} \qquad (1)$$

Here, y represents the original value in the dataset, $\min(y)$ denotes the minimum value in the dataset, and $std(y^*)$ describes the standard deviation of the upgraded dataset y^*. After pre-processing, feature extraction is performed.

3.2 Proposed framework

In the proposed framework, from the pre-processed output, the high level features are removed using INMF and Sparse Variational Autoencoder (SVAE) methods.

3.2.1 Improved Non-negative Matrix Factorization

The aim of feature extraction through INMF is to refine the process of identifying significant, high-level features while overcoming the shortcomings of traditional NMF [28]. The adaptive weight matrix and entropy regularization are used to improve the NMF. These advancements allow for a more effective recognition of hidden patterns or structures that are essential for comprehending the underlying data. Furthermore, the enhanced algorithm typically offers improved scalability and efficiency, making it suitable for handling real time data. An entropy regularizer is utilized to impose a penalty on the cost

function of NMF, allowing the weights to fall within the range of [0, 1] rather than being restricted to binary values of 0 or 1. This approach leverages information entropy to assess the uncertainty associated with the updated weights T and is represented as,

$$T_{ji} = \frac{e^{-\frac{[(y_{ji}-WH)_{ji}]^2}{\nu}}}{\sum_{l=1}^{M} e^{-\frac{[(y_{ji}-WH)_{ji}]^2}{\nu}}} \quad (2)$$

Then, the base matrix W is updated as,

$$W \leftarrow W\Theta(T\Theta Y)H^T./\{[T\Theta(WH)]H^T\} \quad (3)$$

In a similar manner, the update rule for H is derived as,

$$H \leftarrow H\Theta W^T(T\Theta Y)./\{W^T[T\Theta(WH)]\} \quad (4)$$

Here, the improved framework not only boosts the accuracy and dependability of subsequent tasks but also guarantees scalability and relevance. These innovations lay a solid groundwork for addressing complex challenges in fields that require high-dimensional data analysis, such as cyber security, image processing, and bioinformatics.

3.2.2 Sparse Variational auto encoder

A sparse variational auto encoder (SVAE) [29] plays a vital role in identifying long-range attacks by proficiently extracting high-level features that reveal underlying patterns and anomalies within the data. The SVAE utilizes the framework of a variational auto encoder (VAE) to develop a probabilistic representation of the input data in a latent space. The structure of SVAE is shown in Figure 2.

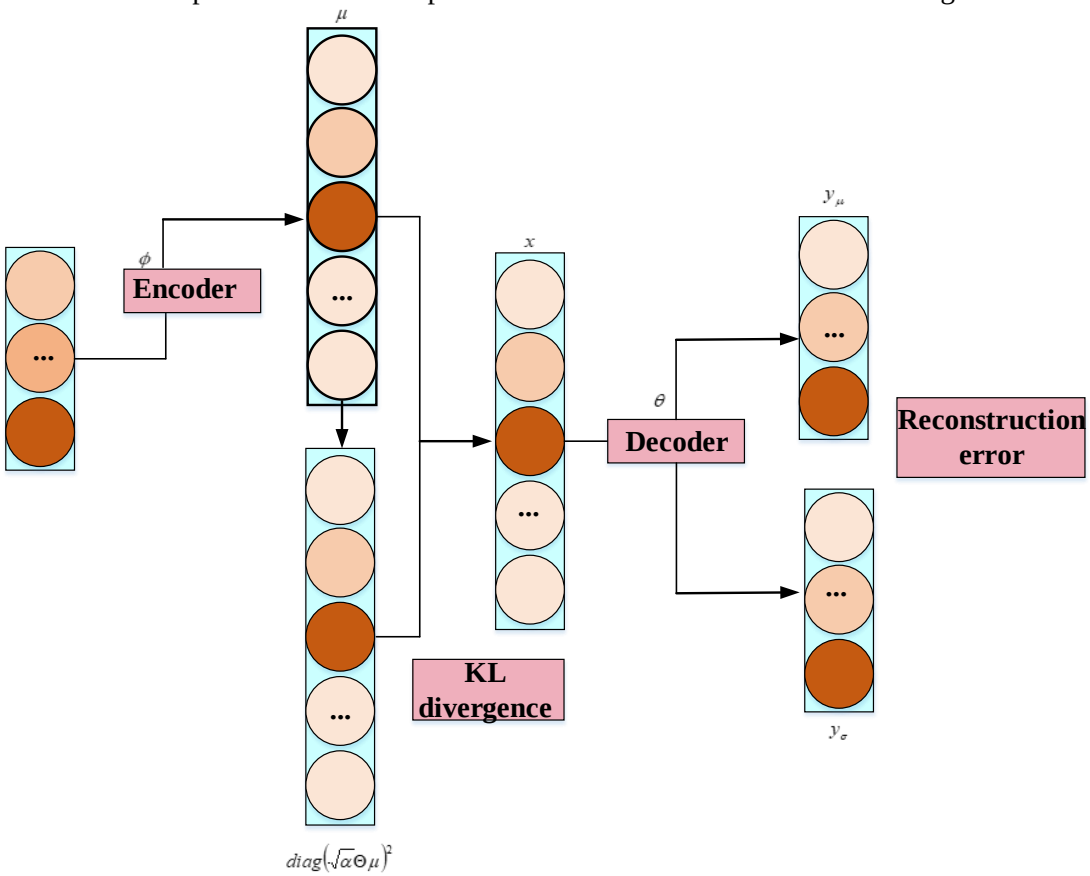

Figure 2: Structure of SVAE

This approach minimizes noise and redundancy, highlighting the essential traits that distinguish normal behavior from attack patterns. By mapping high-dimensional data into a lower-dimensional latent space, the SVAE uncovers subtle, distributed, or correlated anomalies that signal long-range attacks, which may be hidden in the raw data. Furthermore, the probabilistic characteristics of the VAE allow the model to adapt effectively across various scenarios, improving its resilience to different attack methodologies. This advanced feature extraction is critical for subsequent detection and classification processes, facilitating the precise and efficient identification of long-range attacks in complex systems.Variational Autoencoders (VAEs) function as generative models that utilize neural networks as probabilistic

encoders and decoders. The dropout posteriors are incorporated into the VAE framework to regularize the neural networks, resulting in a sparse representation of the variational parameters. For any input y , the Encoder layer performs a linear transformation to derive the necessary outputs.

$$\mu = \phi y \quad (5)$$

Let x represents the hidden layer of the sparse VAE, and the approximate posterior of x follows this distribution as

$$p_\phi(x \mid y) = \left(\mu : diag\left(\sqrt{b}\Theta\mu\right)^2 \right) \quad (6)$$

The approximate posterior described in Equation (6) is specifically known as dropout posteriors when utilized in VAE models. Following this, x is passed to a linear transformation layer known as the Decoder, which produces,

$$y_\mu = \theta x \quad (7)$$

Here,

$$q_\theta(y \mid x) = N(y_\mu; y_\sigma) \quad (8)$$

The calculation of the Kullback-Leibler (KL) divergence is made easier by using the following formula.

$$K\left[p_\phi(x \mid y)\right] \approx -k_1\sigma(k_2 + k_3 \; in \; \alpha) + 0.5 \ln(1 + \alpha^{-1}) + k_1 \quad (9)$$

Here, $k_1 = 0.63576$, $k_2 = 1.87320$, $k_3 = 1.48695$, $\sigma(.)$ is defined as sigmoid function. Based on the equation (9), variational lower bound can be acquired as,

$$L_{VAE}(y) = -K\left[p_\phi(x \mid y) \| \, q(x)\right] + E_{P\phi(x|y)}\left[\log q_\theta(y \mid x)\right] \quad (10)$$

The Sparse VAE is applied to extract the high level features. Within each Sparse VAE, the hidden representations of the encoders are defined as follows.

$$\mu_t = \phi_t y_t$$

$$x_t \sim N(\mu_t; diag(\sqrt{b_t}\Theta\mu_t)^2) \quad (11)$$

Then, the data reconstructed in the decoders are,

$$\mu_t = \theta_t y_t$$

$$\hat{x}_t \sim N(y\mu_t; y_{\sigma t})^2) \quad (12)$$

Here, in this stage, SVAE maximizes the lower bound and extracts the high level features.

3.2.3 Depth wise separable convolutional Resent

The INMF features are given as input to Depthwise Separable Convolutional Resnet (DSC-ResNet) to learn latent features. Here, a deptwise separable convolution is combined with the residual network to learn the latent features. The integration of Depthwise Separable Convolution and Residual Networks (ResNets) in the detection of long-range attacks aims to effectively learn latent features that encapsulate both spatial and temporal dependencies within intricate data. Depthwise separable convolution is a highly efficient convolutional technique that minimizes computational demands while preserving the capability to extract detailed spatial features from the input data. Residual Networks use shortcut connections to bypass one or more layers, which helps to reduce the vanishing gradient problem and allows the network to develop more detailed representations. When combined, these techniques create a robust framework for learning latent features, as depth wise separable convolutions improve computational efficiency while ResNets facilitate effective feature extraction in deeper architectures. This synergy is particularly advantageous for detecting long-range attacks, where recognizing intricate patterns over extended timeframes is essential for establishing strong defence strategies.

- Depth wise separable convolution

Depth wise separable convolution [30] is an efficient approach for extracting latent features from data that divides the usual convolution process into two stages: depth wise convolution as well as point wise convolution. During the depth wise convolution phase, each input channel is treated independently, allowing the network to focus on capturing specific spatial characteristics inside each one. Subsequently, the point wise convolution phase merges these channel-specific features. This two-phase approach significantly lowers both computational complexity and the number of parameters, making depth wise

separable convolution especially beneficial in contexts involving high-dimensional data or limited resources. The deptwise separable convolution is defined in Figure 3.

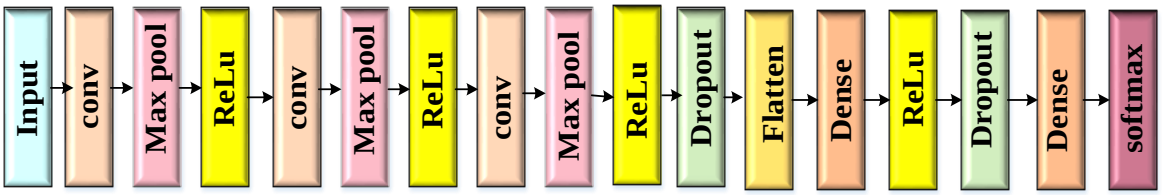

Figure 3: Depth wise separable convolution

The presence of numerous convolution parameters in the kernel necessitates extensive matrix calculations, leading CNN models that handle high-resolution data to require increased memory resources. Consequently, it may be necessary to simplify several CNN models to reduce the number of convolution parameters. One effective approach is Depthwise Separable Convolution (DSC), which minimizes the overall parameter count, decreases the complexity of matrix calculations, and maintains accuracy. Unlike traditional CNNs that utilize a convolution kernel with identical input channels, allowing for channel-by-channel matrix operations, DSC consists of two distinct convolution stages: Depth wise as well as Point wise. In the depth wise convolution stage, each input channel of the neural network is processed independently using a separate filter, which corresponds to spatial convolution. Subsequently, point wise convolution, which employs 1×1 convolutional windows, combines the features across channels to create new representations. Thus, the total number of trainable parameters in depth-wise separable convolution can be stated as follows:

$$D_m^2 \times D_f^2 \times P \quad (13)$$

$$D_f^2 \times P \times Q \quad (14)$$

$$\left(D_f^2 \times P\right)\left(D_m^2 + N\right) \quad (15)$$

The overall count of parameters for depth wise convolution is represented by Equation (13). The total parameters for pointwise convolution are indicated by Equation (14). Consequently, the total parameters for depthwise separable convolution (DSC) are outlined in Equation (15). This leads to a substantial reduction in both the model's parameters and computational costs.

- Residual network

Residual Networks (ResNets) [31] represent a significant advancement in deep learning architecture, specifically engineered to effectively learn latent features while overcoming the difficulties associated with training very deep neural network. The primary innovation of ResNets is their use of residual connections, which allow certain layers to bypass one or more subsequent layers, establishing shortcut paths that enable the direct addition of a layer's input to its output. These latent features are vital for tasks that demand detailed analysis. The capability to develop deep and meaningful representations is often critical for achieving high-precision outcomes. Each of these residual units can be represented as follows.

$$x_j = m(y_j) + R(y_j, v_j) \quad (16)$$

$$y_{j+1} = +e(x_j) \quad (17)$$

Here, R represents a residual function, e denotes a ReLU function, v_j signifies weight matrix, while y_j and x_j correspond to the inputs and outputs of the j^{th} layer. The function m is defined as an identity mapping that is shown below:

$$m(y_j) = y_j \quad (18)$$

Here, the residual function R is described as,

$$R(y_j, v_j) = v_j.\sigma(C(v_j').\sigma(C(y_j))) \quad (19)$$

Here, $C(y_j)$ represents batch normalization, "." indicates convolution and $\sigma(y) = \max(y,0)$. The core concept of residual learning revolves around the branching of paths for gradient flow. Residual networks exhibit certain similarities to highway networks, including the use of residual blocks and

shortcut connections. ResNets can be conceptualized as an ensemble of multiple paths rather than a deep structure. Notably, these paths within ResNets vary in length with only one path traversing all residual units. Additionally, not all signal paths contribute to gradient propagation, which facilitates the quicker optimization and training of ResNets.

3.2.4 Stacked Bidirectional gated recurrent dropout network

The Stacked Bidirectional Gated Recurrent Dropout Network (SBi-GRDN) is essential for detecting long-range attacks, utilizing high-level features derived from a Sparse Variational Autoencoder (SVAE) to uncover intricate relationships and interactions among features. Here, bidirectional gated recurrent unit is in a stacked connection and combines with a dropout network to identify the complex relationship among features. Its bidirectional design allows the network to analyze sequential data in both forward and backward directions, effectively capturing the complex temporal dependencies typical of long-range attacks. The inclusion of gated recurrent units (GRUs) enhances the network's capacity to model long-term dependencies and nonlinear dynamics, while the stacked architecture facilitates a hierarchical approach to learning deeper and more abstract feature representations. By merging the comprehensive feature representations from the SVAE, the system establishes a powerful mechanism for detecting and analyzing long-range attack patterns within complex environments.

- Stacked bi-directional gated recurrent unit

Bidirectional GRU (BiGRU) uses the relationship between preceding and following signal states in the same sequence to better forecast the current state. The stacked BiGRU [32] layers are described in Figure 4.

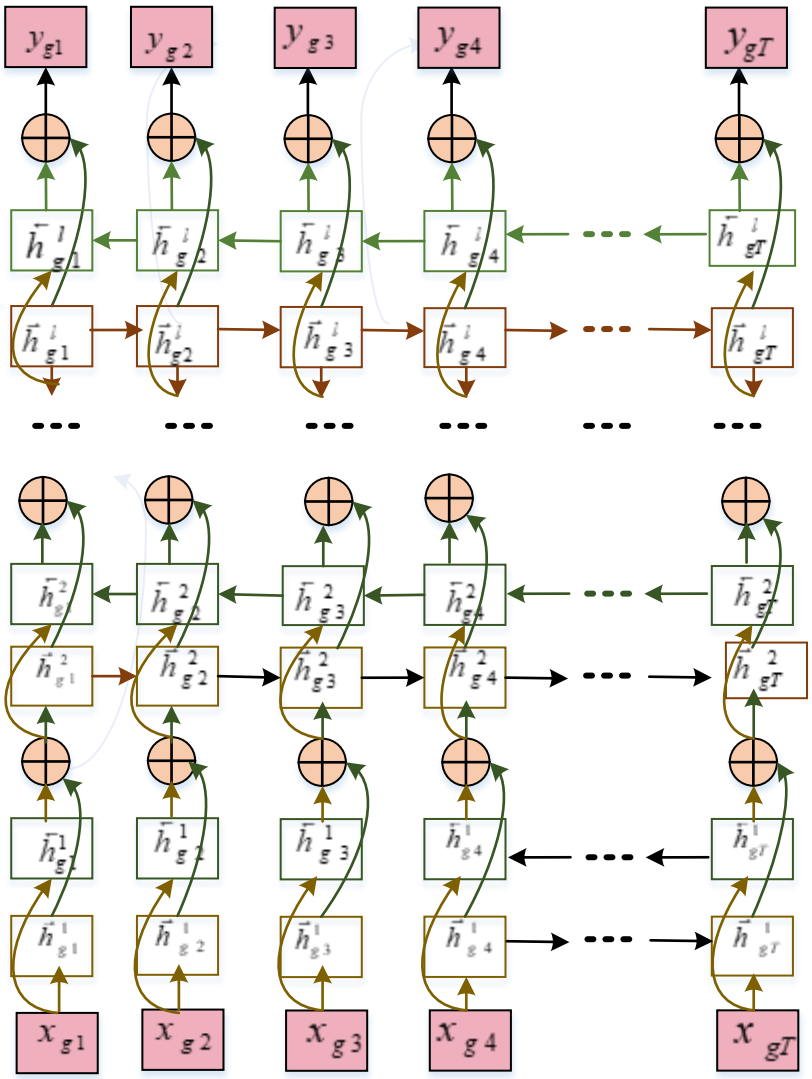

Figure 4: Stacked BiGRU layers

Additionally, stacking multiple BiGRU layers enables the effective extraction of higher-level features from the sequence. The input for the module is defined as $y_G = \{y_{g1},....y_{gT} \mid y_{gT} \in R^d\}$ and is

processed by the stacked BiGRU. With the number of layers in the stacked BiGRU module denoted by M_G, the forward and backward computations are represented by the arrows $\rightarrow$and $\leftarrow$, respectively. In the m^{th} layer, forward and backward hidden units are denoted by $\vec{h}^m$ and $\overleftarrow{h}^m$. Upon completion of the computations in the m^{th} layer, the results from both directions are combined and forwarded to the subsequent layer. For the input at the current timestep t in the m^{th} layer, represented as $y_{gt}^m \in R^{d^t}$, the precise circulation method of a single GRU is illustrated as,

$$\begin{aligned} R_{gt}^m &= \sigma\left(y_{gt}^m V_{yr}^m + H_{gt-1}^m V_{hr}^m + c_r^m\right) \\ X_{gt}^m &= \sigma\left(y_{gt}^m V_{yx}^m + H_{gt-1}^m V_{hy}^m + c_y^m\right) \\ \tilde{H}_{gt}^m &= \tanh\left(y_{gt}^m V_{yh}^m + \left(R_{gt}^m \Theta H_{gt-1}^m\right) V_{hh}^m + c_h^m\right) \\ H_{gt}^m &= X_{gt}^m \Theta H_{gt-1}^m + \left(1 - X_{gt}^m\right) \Theta H_{gt}^m \end{aligned} \tag{20}$$

Here, R_{gt}^m and X_{gt}^m denote the outputs of the reset and update gates, correspondingly. The term $\tilde{H}_{gt}^m$ signifies the generated memory, while H_{gt-1}^m indicates the hidden state from the previous time step. The hidden state for the current time step is represented by H_{gt}^m. The symbol Θ denotes element-wise multiplication. The weight matrices and biases are defined as V_{yr}^m, V_{yr}^m, V_{yh}^m and c_r^m, c_y^m, c_h^m, respectively. Additionally, σ and tanh represent the sigmoid as well as hyperbolic tangent activation functions, correspondingly.

- Dropout

Dropout [33] serves as a solution to mitigate the overfitting issues. The fundamental concept of dropout involves randomly deactivating a portion of the units with a probability of $1-\rho$ (or retaining them with a probability of ρ) during each training iteration, with ρ being determined through experimentation. Consequently, the networks that utilize dropout differ from one another and are less complex than traditional neural networks, thereby improving the model's ability to resist overfitting and accelerating the training process.

3.3 Feature fusion using multi head attention

Multi-head attention [34] is an effective mechanism for integrating features in detecting long-range attacks, especially within complex, distributed systems where attack patterns may change over time. This technique, frequently employed in transformer architectures, effectively addresses the need to capture both temporal and spatial dependencies in extensive and varied datasets. By leveraging multiple attention heads, the model can simultaneously concentrate on various elements of the input data, enabling it to identify subtle and prolonged attack patterns across different features. For example, in the context of long-range attack detection, multi-head attention facilitates the integration of information from diverse data sources, including network logs, application metrics, and system performance indicators, while preserving the temporal relationships among these features. Each attention head is trained to recognize distinct interactions and correlations, ensuring the model remains responsive to a wide range of attack strategies. Although it requires significant computational resources, the benefits of multi-head attention in feature integration, scalability, and adaptability render it a vital component in the detection of long-range attacks within contemporary cybersecurity frameworks. Initially, the inputs Q, K, and V undergo a linear transformation. This process computes one attention head at a time, which necessitates the execution of multi-head operation in n times. Each linear transformation for Q, K, and V utilizes distinct parameters P. The corresponding formula is presented below,

$$head_j = attention(QP_j^Q, KP_j^K, VP_j^Q) \tag{21}$$

$$multihead(Q,K,V) = concat(head_j, ..., head_r)P^O \tag{22}$$

Here, the outputs from the features r are concatenated, and the resulting value is transformed linearly to serve as the output of the MHA.

3.4 Fully connected layer with sigmoid function

In the context of classifying long-range attacks within a consensus layer, a fully connected layer [35] utilizing a sigmoid activation function is essential for converting high-dimensional feature

representations into practical classifications. This layer guarantees that each neuron is linked to all activations from the previous layer, enabling the model to identify intricate patterns and relationships throughout the entire feature set. The sigmoid activation function compresses outputs to a range between 0 and 1 for binary classification tasks. Collectively, the fully connected layer and sigmoid function are vital for improving predictions and ensuring accurate classification in the realm of consensus-layer security. The function performed by the fully connected layer is represented as $x = fc(y, v, c)$, where y is the input to the layer, v signifies the weight matrix, c is the bias, and x indicates the output. The input data y consists of n features, with each element y_j corresponding to feature j. It is important to note that the weight matrix v has dimensions of $N \times K$, here K denotes the dimensionality of x. This function processes the vector y, producing the element j' in x through a specific operation as,

$$x_{j'} = \sum_{j} v_{jj'} y_j + c_{j'} \quad (23)$$

Here, Sigmoid function is described as follows,

$$\sigma(y) = \frac{1}{1+e^{-y}} \quad (24)$$

The sigmoid activation function transforms its input values from the range of $[-\infty;+\infty]$to [0; 1]. Finally, a fully connected layer, along with the sigmoid activation function, classifies the attack.

4. Results and discussion

This section describes in depth the suggested model's performance analysis and results evaluations. The implementation utilizes the Python programming tool. Table 2 denotes the system configuration.

Table 2: System configuration

Parameters	Configuration
Processor	Intel(R) Core(TM) i5-6500 CPU @ 3.20GHz 3.19 GHz
Installed RAM	16.0 GB (15.9 GB usable)
Device ID	2FBF9ABF-16D4-452A-9A96-AB694B2B7888
System type	64-bit operating system, x64-based processor

4.1 Dataset description

The proposed study comprises a total data of 10000 and a total features of 17. The feature names are Block height, Unix Timestamp, TxnFee(Binary), Status(Tags), Block Generation Rate, Stake Reward, Coin Stake, Stake distribution rate, TxnSize, Coin Days, Coin Age, Block density (%), Block score, Coin Day Weight, Node Label, as well as Validation status. The two classes are normal and attack.

4.2 Performance metrics

The following part includes in-depth study of several performance indicators. The metrics relevant to the model under investigation are evaluated through Accuracy, Precision, Recall, F1-score, mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), Training accuracy, Training loss, Testing accuracy, Testing loss, and Receiver operating characteristics (ROC) [36].

4.3 Performance analysis with existing models

The performance analysis is compared with the existing models such as depth CNN, attention CNN methods, bidirectional gated recurrent unit (BI-GRU), and DenseNet.

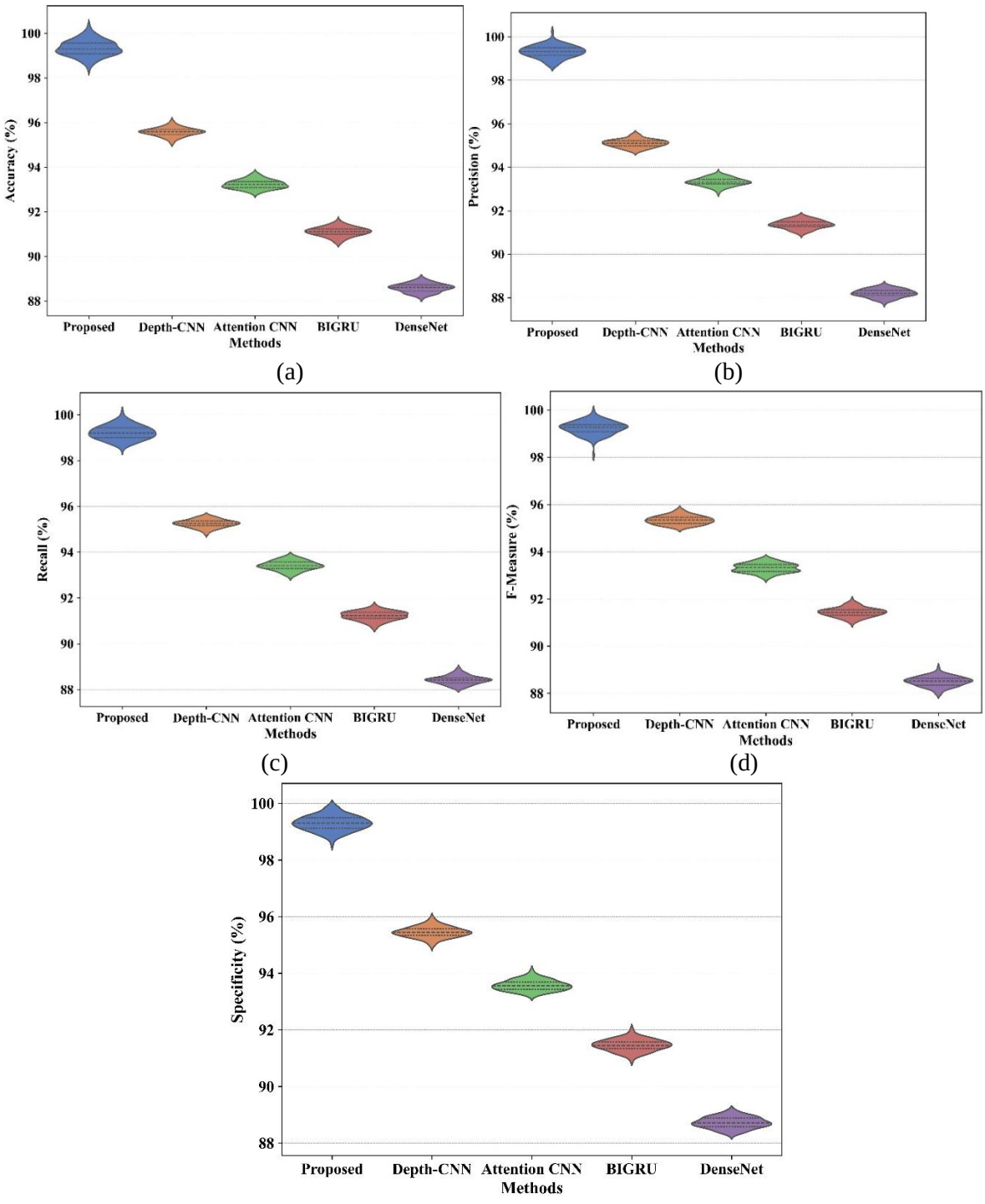

Figure 5: Performance evaluation of the proposed model with the existing methods

Figure 5 illustrates the performance evaluation of the suggested strategy. Figure 5 (a) shows a comparison of accuracy, emphasizing the benefits of the suggested approach. According to the results, the suggested approach outperforms existing methods in terms of accuracy. The x-axis depicts different methodologies, and the y-axis shows the related accuracy levels. This investigation indicates the usefulness of the suggested method in identifying long-range attacks. Figure 5 (b) highlights the precision of the proposed method, which surpasses that of the other techniques. In Figure 5 (c), recall rates are compared, with the proposed method also achieving a higher rate than its competitors. Figure 5 (d) presents a comparison of F1 scores. Finally, Figure 5 (e) shows the specificity. The existing model faces issues with limited feature representation, inefficient feature fusion, and inconsistent performance. This method not only demonstrates enhanced accuracy but also improves processing speed in classifying

the attack, which makes it more effective for detection purposes. A detailed performance analysis in relation to existing methods is provided in Table 3.

Table 3: Values for the proposed and existing approaches

Models	Proposed	Depthcnn	Attention cnn	BiGRU	Densenet
Accuracy (%)	99.13	95.56	93.23	91.09	88.56
Precision (%)	99.54	95.13	93.31	91.34	88.21
F1-score (%)	99.52	95.23	93.41	91.24	88.41
Recall (%)	99.49	95.33	93.32	91.44	88.51
Specificity (%)	99.49	95.45	93.58	91.46	88.73

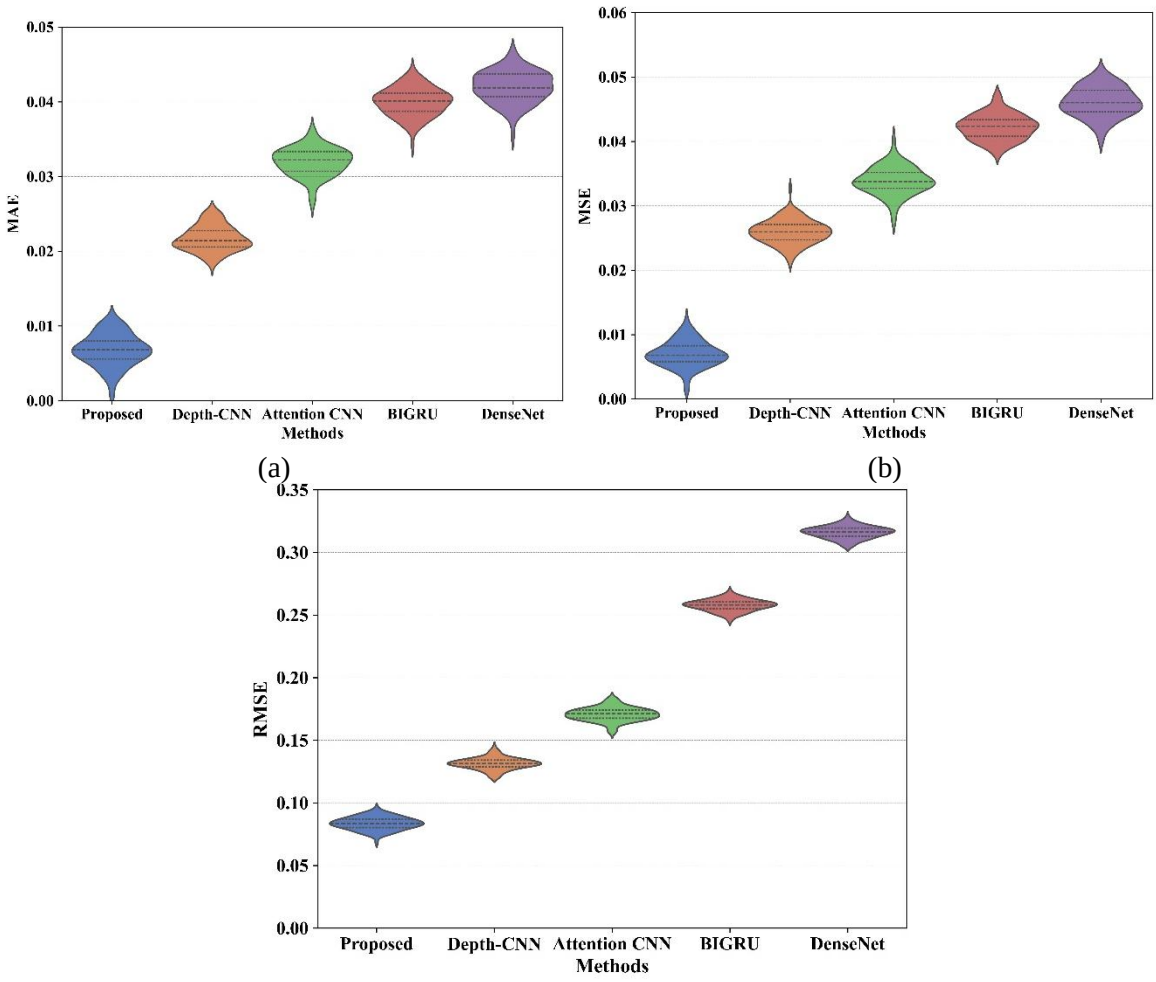

Figure 6: Error metrics examination of the proposed and existing models

Figure 6 provides an MAE, MSE, and RMSE. In Figure 6 (a), the MAE for the proposed method is recorded at 0.004, which is lower than the values observed for the other existing techniques. Figure 6 (b) showcases the MSE for the proposed method, which stands at 0.004, indicating its superior performance relative to the existing methods. Finally, Figure 6 (c) presents the RMSE for the proposed method, measured at 0.06. This evaluation demonstrates that the proposed method consistently achieves lower average errors compared to alternative techniques, underscoring its improved error performance. A detailed performance analysis of these error metrics can be found in Table 4.

Table 4: Error metrics values of the proposed and existing models

Models	Proposed	Depth-CNN	Attention CNN	BiGRU	Densenet
MSE	0.00466	0.022	0.032	0.04	0.042
RMSE	0.06831	0.026	0.034	0.042	0.046
MAE	0.00466	0.132	0.171	0.257	0.316

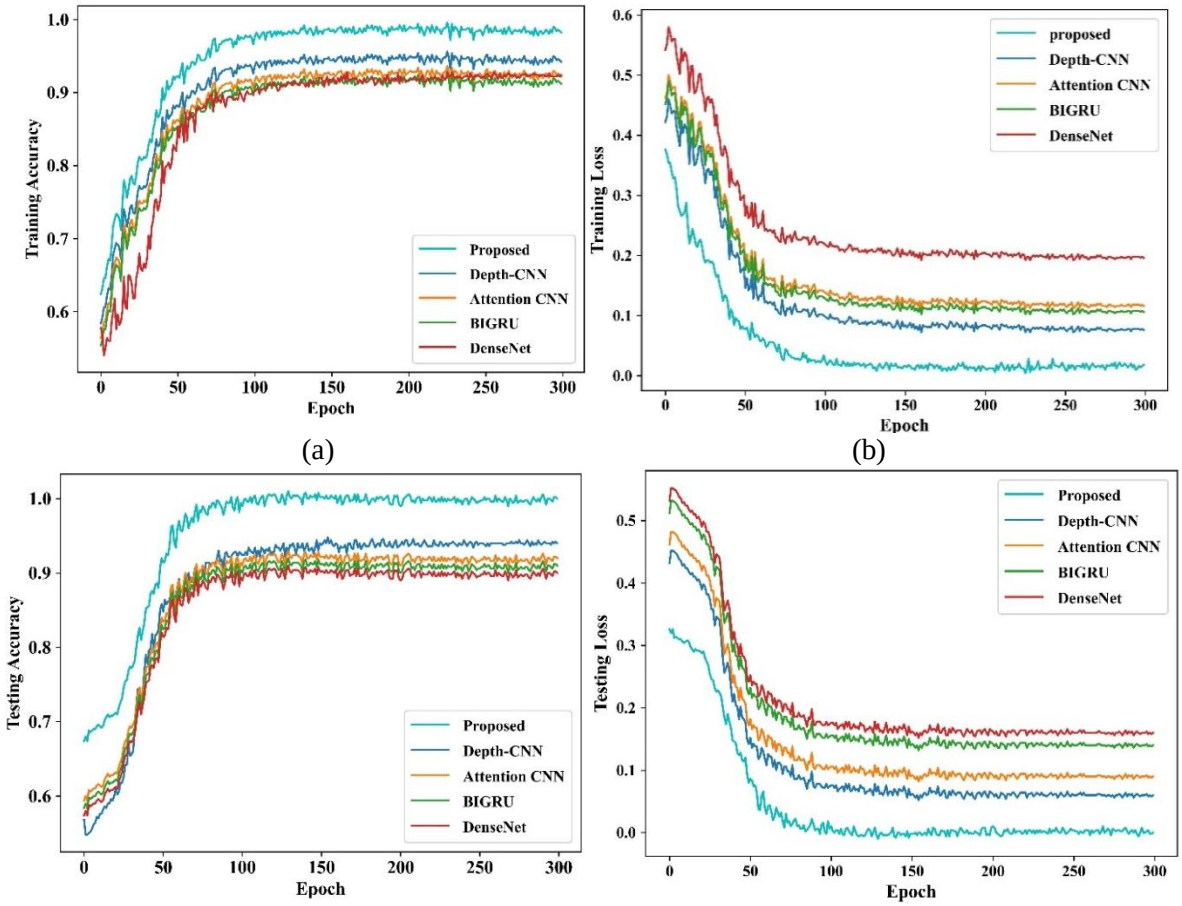

Figure 7: Accuracy and loss in training and testing

Figure 7 denotes the accuracy and loss in training and testing. Figure 7 (a) depicts the training accuracy, while Figure 7 (b) presents the training loss. Figure 7 (c) describes the testing accuracy. Finally, Figure 7 (d) shows the loss in testing. The model demonstrates improved accuracy throughout the training process. The suggested approach yields elevated accuracy levels. Additionally, this method indicates a loss reduction, making it more suitable for application within the proposed framework. The evaluation findings show that the model obtains higher accuracy and lower loss rates in its predictions.

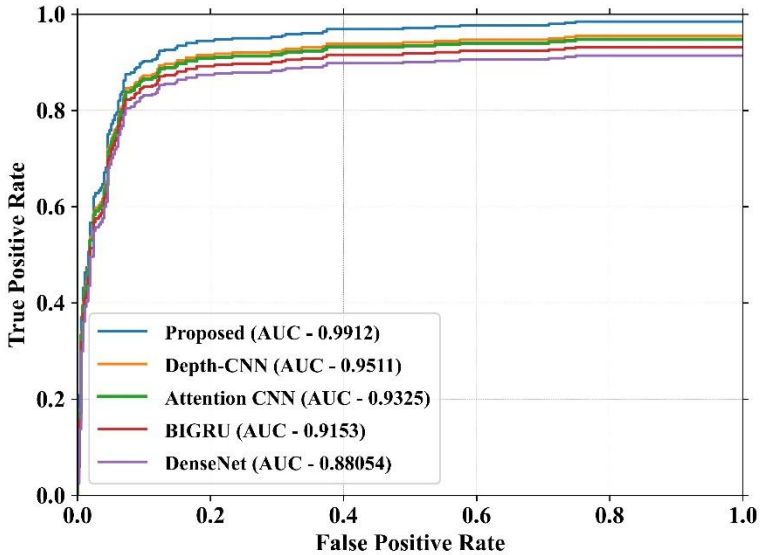

Figure 8: Receiver operating characteristics

Figure 8 illustrates the ROC graph. The method put forward achieves an Area Under the Curve (AUC) of 0.9912, which is a notable improvement over other models. These findings suggest that the proposed method excels in reducing false positives while enhancing true positives. The existing methodologies

exhibit challenges in managing long-range dependencies and intricate feature extraction, resulting in less effective performance in identifying long-range attacks. The proposed method overcomes these limitations by utilizing advanced feature extraction techniques and better modelling of long-range dependencies. Its impressive AUC score underscores its strength and effectiveness in accurately detecting long-range attacks, surpassing both traditional and alternative deep learning frameworks.

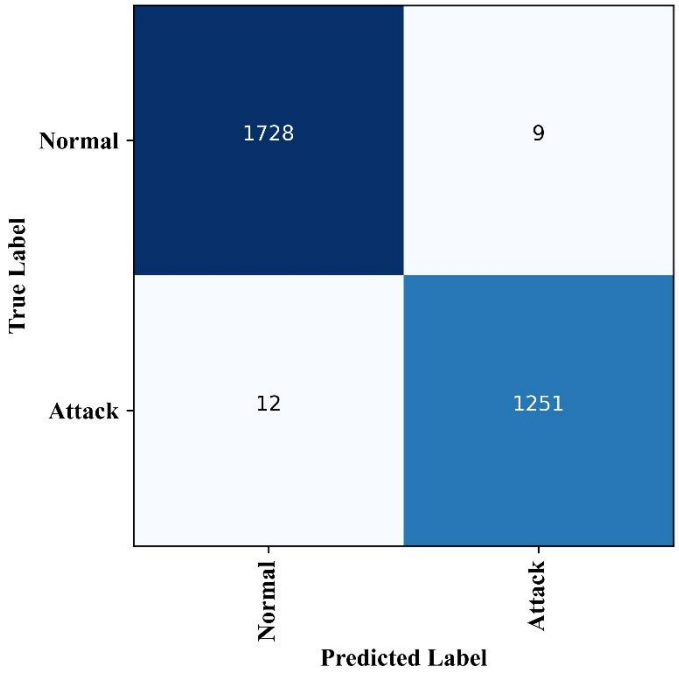

Figure 9: Confusion matrix

Figure 9 depicts the confusion matrix, a critical tool for evaluating the performance of classification approaches. In the realm of long-range attack detection, the confusion matrix underscores the difficulties in differentiating between normal activities and attack behaviours. The proposed approach correctly classifies 1728 labels and 9 are misclassified. Also, 1251 labels are correctly classified as attacks, and only 12 labels are misclassified.

4.4 Discussion

The study introduces a novel deep learning model to solve long range attacks on the blockchain consensus layer. Long-range attack detection is essential for maintaining system security, especially in contexts where malicious activities are intricate and evolve over time. The existing methodologies show varying levels of effectiveness. For instance, Proof of Stake (PoS) achieves an accuracy rate of 85.20%, but its performance limitations reveal challenges in identifying complex attack patterns and long-range dependencies. The Convolutional Neural Network-Adversarial Detector, with a 91% accuracy, demonstrates enhanced capabilities in recognizing adversarial actions. Yet, it encounters difficulties in comprehensive detection due to issues in effectively modeling long-term dependencies. The Kullback-Leibler Divergence method raises accuracy to 93.85% by utilizing distributional metrics to spot anomalies, but it is hindered by computational demands and sensitivity to minor changes. Conversely, the Poaching Raptor Optimization technique reaches the highest accuracy of 95.12%, showcasing its exceptional ability to refine detection processes and adapt to intricate, long-range attack situations. Nevertheless, this method may also experience challenges related to scalability and generalization across various datasets. The proposed solution aims to overcome these limitations by incorporating advanced feature extraction and modelling techniques for long-range dependencies, thereby ensuring robust and scalable detection of sophisticated attacks. Its enhanced performance and adaptability highlight its superiority over the existing approaches, offering a more dependable and efficient solution for long-range attack detection in dynamic settings. Table 5 defines the accuracy of the existing methods.

Table 5: Accuracy

Techniques	Accuracy (%)
PoS- proof of stake [21]	85.20
Poaching raptor optimization [24]	95.12
Kullback leibler divergence [26]	93.85
Convolutional neural network-Adversarial detector [27]	91
Proposed	99.1

5. Conclusion

The proposed model presents a novel deep learning framework for identifying long-range attacks within the blockchain consensus layer by utilizing advanced techniques. The implementation of Upgraded Min-Max Normalization guarantees the integrity and uniformity of the input data. Furthermore, the combined application of Improved Non-negative Matrix Factorization (INMF) and Sparse Variational Autoencoder (SVAE) facilitates the effective extraction of high-level features. The incorporation of Depthwise Separable Convolutional ResNet (DSC-ResNet) alongside the Stacked Bidirectional Gated Recurrent Dropout Network (SBi-GRDN) enables a thorough understanding of latent and intricate feature interactions. The attention process is critical in improving the integration of complementing data from these two models, which results in higher detection accuracy. Finally, a fully linked layer with a sigmoid activation capability ensures reliable detection of prospective threats. This innovative approach underscores the effectiveness of merging deep learning techniques with attention mechanisms to protect the integrity and reliability of blockchain systems against complex attack strategies. The investigational results show that the proposed method shows an accuracy of 99.1%, Precision of 99.5%, Recall of 99.4%, and F1-score of 99.5%. The future scope of the research is to develop a hybrid consensus approach that mitigates the limitations of traditional consensus algorithms while ensuring robust performance and addressing security issues. This paper is intended to assist researchers in conducting more in-depth analyses tailored to specific application areas. A potential path for future research could be to use reinforcement learning to investigate the behaviors of malicious nodes as well as the various types of malicious acts carried out by validator nodes on the PoS blockchain. Enhancing the suggested framework to accommodate multi-modal data and facilitate cross-block chain interoperability could significantly increase its relevance across various blockchain platforms. Furthermore, incorporating explainable AI methodologies can enhance the model's interpretability, thereby building trust and promoting improved decision-making in practical applications. These developments could collectively bolster the resilience and security of blockchain consensus mechanisms, thereby encouraging wider adoption across different sectors.

References

[1] Zaabar, Bessem, Omar Cheikhrouhou, Faisal Jamil, Meryem Ammi, and Mohamed Abid. "HealthBlock: A secure blockchain-based healthcare data management system." Computer Networks 200 (2021): 108500.

[2] Idrees, Sheikh Mohammad, Mariusz Nowostawski, Roshan Jameel, and Ashish Kumar Mourya. "Security aspects of blockchain technology intended for industrial applications." Electronics 10, no. 8 (2021): 951.

[3] Kassen, Maxat. "Blockchain and e-government innovation: Automation of public information processes." Information Systems 103 (2022): 101862.

[4] Han, Rong, Zheng Yan, Xueqin Liang, and Laurence T. Yang. "How can incentive mechanisms and blockchain benefit with each other? a survey." ACM Computing Surveys 55, no. 7 (2022): 1-38.

[5] Murray-Rust, Dave, Chris Elsden, Bettina Nissen, Ella Tallyn, Larissa Pschetz, and Chris Speed. "Blockchain and beyond: Understanding blockchains through prototypes and public engagement." ACM Transactions on Computer-Human Interaction 29, no. 5 (2023): 1-73.

[6] Kaur, Manpreet, Mohammad Zubair Khan, Shikha Gupta, Abdulfattah Noorwali, Chinmay Chakraborty, and Subhendu Kumar Pani. "MBCP: Performance analysis of large scale mainstream blockchain consensus protocols." Ieee Access 9 (2021): 80931-80944.

[7] Pauletto, Christian. "Blockchain in international e-government processes: Opportunities for recognition of foreign qualifications." Research in globalization 3 (2021): 100034.

[8] Salamé, Léna, Janos J. Bogardi, Zita Sebesvari, Klement Tockner, Burcu Yazici, Fatma Turan, Burcu Calli, Aslıhan Kerç, Olcay Ünver, and Yvonne Walz. "Drivers, pressures and stressors: The societal framework

of water resources management." In Handbook of Water Resources Management: Discourses, Concepts and Examples, pp. 329-364. Cham: Springer International Publishing, 2021.
[9] Shwetha, A. N., and C. P. Prabodh. "A comprehensive review of blockchain based solutions in food supply chain management." In 2021 5th international conference on computing methodologies and communication (ICCMC), pp. 519-525. IEEE, 2021.
[10] Medhi, Pankaj Kumar. "Blockchain-Enabled Supply Chain Transparency, Supply Chain Structural Dynamics, and Sustainability of Complex Global Supply Chains—A Text Mining Analysis." In Information for Efficient Decision Making: Big Data, Blockchain and Relevance, pp. 273-312. 2021.
[11] Hadjiefthymiades, Stathes, Michail Chatzidakis, and Dionisis Reisis. "Ensuring consensus on trust issues in capability-limited node networks with Blockchain technology."
[12] Natoli, Christopher James. "The Road to El Diablo: Towards Secure High Performance Blockchains." PhD diss., 2021.
[13] BABOI, Mihai. "Security of Consensus Mechanisms in Blockchain." Romanian Cyber Security Journal 5, no. 2 (2023): 45-53.
[14] Khoa, Tran Viet, Do Hai Son, Chi-Hieu Nguyen, Dinh Thai Hoang, Diep N. Nguyen, Nguyen Linh Trung, Tran Thi Thuy Quynh, Trong-Minh Hoang, Nguyen Viet Ha, and Eryk Dutkiewicz. "Securing Blockchain Systems: A Novel Collaborative Learning Framework to Detect Attacks in Transactions and Smart Contracts." arXiv preprint arXiv:2308.15804 (2023).
[15] Khoa, Tran Viet, Do Hai Son, Dinh Thai Hoang, Nguyen Linh Trung, Tran Thi Thuy Quynh, Diep N. Nguyen, Nguyen Viet Ha, and Eryk Dutkiewicz. "Collaborative learning for cyberattack detection in blockchain networks." IEEE Transactions on Systems, Man, and Cybernetics: Systems (2024).
[16] Mbaya, Emmanuel Baldwin, Emmanuel Adetiba, Joke A. Badejo, John Simon Wejin, Oluwadamilola Oshin, Olisaemeka Isife, Surendra Colin Thakur, Sibusiso Moyo, and Ezekiel F. Adebiyi. "SecFedIDM-V1: A Secure Federated Intrusion Detection Model With Blockchain and Deep Bidirectional Long Short-Term Memory Network." IEEE Access 11 (2023): 116011-116025.
[17] Yazdinejad, Abbas, Ali Dehghantanha, Reza M. Parizi, Gautam Srivastava, and Hadis Karimipour. "Secure intelligent fuzzy blockchain framework: Effective threat detection in iot networks." Computers in Industry 144 (2023): 103801.
[18] Aljuhani, Ahamed, Prabhat Kumar, Rehab Alanazi, Turki Albalawi, Okba Taouali, AKM Najmul Islam, Neeraj Kumar, and Mamoun Alazab. "A deep learning integrated blockchain framework for securing industrial IoT." IEEE Internet of Things Journal (2023).
[19] Faheem, Muhammad, and Mahmoud Ahmad Al-Khasawneh. "Multilayer cyberattacks identification and classification using machine learning in internet of blockchain (IoBC)-based energy networks." Data in Brief 54 (2024): 110461.
[20] Kumar, Randhir, Prabhat Kumar, Rakesh Tripathi, Govind P. Gupta, AKM Najmul Islam, and Mohammad Shorfuzzaman. "Permissioned blockchain and deep learning for secure and efficient data sharing in industrial healthcare systems." IEEE Transactions on Industrial Informatics 18, no. 11 (2022): 8065-8073.
[21] Sanda, Olanrewaju, Michalis Pavlidis, Saeed Seraj, and Nikolaos Polatidis. "Long-Range attack detection on permissionless blockchains using Deep Learning." Expert Systems with Applications 218 (2023): 119606.
[22] Chauhdary, Sajjad Hussain, Mohammed Saeed Alkatheiri, Mohammed A. Alqarni, and Sajid Saleem. "An efficient evolutionary deep learning-based attack prediction in supply chain management systems." Computers and Electrical Engineering 109 (2023): 108768.
[23] Tsoukas, Vasileios, Anargyros Gkogkidis, Aikaterini Kampa, Georgios Spathoulas, and Athanasios Kakarountas. "Enhancing food supply chain security through the use of blockchain and TinyML." Information 13, no. 5 (2022): 213.
[24] Ilyas, Benkhaddra, Abhishek Kumar, Mohamed Ali Setitra, ZineEl Abidine Bensalem, and Hang Lei. "Prevention of DDoS attacks using an optimized deep learning approach in blockchain technology." Transactions on Emerging Telecommunications Technologies 34, no. 4 (2023): e4729.
[25] Bassiouni, Mahmoud M., Ripon K. Chakrabortty, Omar K. Hussain, and Humyun Fuad Rahman. "Advanced deep learning approaches to predict supply chain risks under COVID-19 restrictions." Expert Systems with Applications 211 (2023): 118604.
[26] Dai, Qian-yi, Bin Zhang, and Shu-qin Dong. "A DDoS-Attack Detection Method Oriented to the Blockchain Network Layer." Security and Communication Networks 2022, no. 1 (2022): 5692820.
[27] R.SenthamilSelvan “Two-Stage Deep Learning - YouTube Video Recommendation Process” by 2023 International Conference on New Frontiers in Communication, Automation, Management and Security (ICCAMS), ISSN:0018-9219,E-ISSN:1558-2256, 27-28 October 2023, 10.1109/ICCAMS60113.2023.10525842 ,15 May 2024.
[28] R.SenthamilSelvan “Development of CNN Model to Avoid the Food Spoiling Level” by 2023 International Conference on New Frontiers in Communication, Automation, Management and Security (ICCAMS), ISSN:0018-9219,E-ISSN:1558-2256, 27-28 October 2023 2023, 10.1109/ICCAMS60113.2023.10525936 ,15 May 2024.

[29] Aonishi, Toru, Ryoichi Maruyama, Tsubasa Ito, Hiroyoshi Miyakawa, Masanori Murayama, and Keisuke Ota. "Imaging data analysis using non-negative matrix factorization." Neuroscience Research 179 (2022): 51-56.
[30] Geadah, Victor, Gabriel Barello, Daniel Greenidge, Adam S. Charles, and Jonathan W. Pillow. "Sparse-coding variational autoencoders." Neural Computation (2024): 1-31.
[31] Liu, Fucong, Hui Xu, Miao Qi, Di Liu, Jianzhong Wang, and Jun Kong. "Depth-wise separable convolution attention module for garbage image classification." Sustainability 14, no. 5 (2022): 3099.
[32] Velayudhan, Nitha C., A. Anitha, and Mukesh Madanan. "An optimisation driven deep residual network for Sybil attack detection with reputation and trust-based misbehaviour detection in VANET." Journal of Experimental & Theoretical Artificial Intelligence 36, no. 5 (2024): 721-744.
[33] Thakur, Ujwala, Ankit Vidyarthi, and Amarjeet Prajapati. "Recognition of Real-Time Video Activities Using Stacked Bi-GRU with Fusion-based Deep Architecture." Journal of Universal Computer Science 30, no. 10 (2024): 1423.
[34] Rustam, Furqan, Ali Raza, Imran Ashraf, and Anca Delia Jurcut. "Deep ensemble-based efficient framework for network attack detection." In 2023 21st Mediterranean Communication and Computer Networking Conference (MedComNet), pp. 1-10. IEEE, 2023.
[35] Ullah, Farhan, Gautam Srivastava, and Shamsher Ullah. "A malware detection system using a hybrid approach of multi-heads attention-based control flow traces and image visualization." Journal of Cloud Computing 11, no. 1 (2022): 75.
[36] Zhang, Zhao, Yong Zhang, Jie Niu, and Da Guo. "Unknown network attack detection based on open-set recognition and active learning in drone network." Transactions on Emerging Telecommunications Technologies 33, no. 10 (2022): e4212.
[37] Dhanya, K. A., Sulakshan Vajipayajula, Kartik Srinivasan, Anjali Tibrewal, T. Senthil Kumar, and T. Gireesh Kumar. "Detection of network attacks using machine learning and deep learning models." Procedia Computer Science 218 (2023): 57-66.