Artificial Intelligent-driven Hybrid Attention-based Distilled Transformer Framework for Cyber-Attack Detection in Cognitive IoT Network

Jamila Ahmed Masoud Al Mashani
College of Computing and Information Sciences,
University of Technology and Applied Sciences -Salalah, Oman
PG46J25244@utas.edu.om

Abstract: The prevalent incorporation of the Cognitive Internet of Things (CIoT) in sophisticated communication systems has significantly increased security concerns due to intrinsic vulnerabilities that attackers can exploit. Through this integration, devices are capable of gathering and distributing data as well as analyzing and making decisions based on that data, improving their operational efficiency of operation and flexibility to variations in the environment. IoT devices' inadequate data processing and storage capabilities make traditional safety measures ineffective, which increases these security concerns. To address these issues, an AI-based security model is developed with the support of a deep learning network within a communication system. Initially, IoT data are collected from the relevant database and applied to the Hybrid boosted Attention-based Distilled Transformer (HA-DT) for the detection of cyberattacks. Here, a distilled transformer captures complex features, which helps to identify cyberattacks in the CIoT. After the extraction of relevant features, hybrid attention is applied to boost the model detection capability. Finally, the proposed HA-DT detects cyberattacks in CIoT and then classifies it into normal behavior or any specific type of attack. Performance evaluation is conducted using metrics such as accuracy, precision, F1-score and detection rate that help to measure the effectiveness of the proposed framework. Thus, the proposed framework enhances the cybersecurity of CIoT devices in an advanced communication system.

Keywords: Cyber-Attack Detection; Cognitive IoT Network; Hybrid Attention-based Distilled Transformer; Artificial Intelligence

Received: November 2, 2025 Revised: November 14, 2025 Accepted: December 3, 2025

Nomenclature

Abbreviation	Description
CIoT	Cognitive Internet of Things
IoT	Internet of Things
AI	Artificial Intelligence
HA-DT	Hybrid Attention Distilled Transformer
CNN	Convolutional Neural Network
1DCNN	One-Dimensional Convolutional Neural Network
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory
GRU	Gated Recurrent Unit
SE	Squeeze-and-Excitation
DRL	Deep Reinforcement Learning
FPR	False Positive Rate
URL	Uniform Resource Locator

1. Introduction

1.1 Background

The advanced CIoT integrates cognitive computing abilities into traditional IoT frameworks [6]. Through this interconnection, devices not only gather and send data, but also evaluate, learn and make choices based on that data, improving their operational efficacy and responsiveness to environmental changes [7]. A theory of CIoT depends on the capability of devices to process enormous amounts of data, allowing them to change and improve their functions on their own [8]. The unique attribute of CIoT is its dependence on cognitive computing strategies, which allow devices to sense and assess their environment [9]. This capacity is essential for scenarios like intelligent cities, where devices have to assess real-time data to successfully handle resources and respond to shifting urban conditions [10].

The cognitive abilities also help to enhance spectrum allocation for wireless networks by allowing devices to identify and make use of the available bandwidth more efficiently, which is especially significant for the increasing number of IoT devices struggling for restricted spectrum resources [11]. Furthermore, the effectiveness of numerous industries, such as industrial automation, food production, and healthcare, could be greatly improved by CIoT [12]. It improves crop monitoring and drip irrigation systems in agriculture through ecological data analysis, resulting in improved environmental sustainability and production [13]. By allowing devices to adapt their communication strategies in real time based on connection conditions, overall system performance increases [14]. These applications are then further enhanced within CIoT platforms via intelligent radio technology.

In traditional methods, the CIoT network is highly vulnerable to various cyberattacks, which limits computational storage resources [15]. CIoT contains a large number of features, which boost the computational cost and slow down the detection process. In traditional models, managing large-scale networks and resources optimization become difficult, which affects system scalability and performance. Thus, to address these challenges, an efficient deep learning mechanism is integrated, which boosts the attack detection accuracy by reducing the computational complexity.

1.2 System Model of Cognitive IoT

The basic system model of CIoT integrates IoT and cognitive computing, which enables automatic decision-making. This system contains several layers that are responsible for intelligent outcomes.

Sensing layer: It is a data acquisition layer that collects information using different sensors and smart devices. The main purpose of this layer is to collect data and then detect any changes in the environment. The collected data is stored and then further processed for making decisions.

Communication layer: It is suitable for transmitting data from sensors to the processing system for ensuring smooth communication among the components of the system. To ensure reliable communication, different technologies are used for maintaining device connectivity, thereby providing appropriate processing within the system.

Data processing layer: It is the cognitive layer, where data processing is performed. Here, the collected data is analyzed using a deep network that maintains contextual features to predict outcome more effectively. Processing occurs with the support of edge computing, that provides large-scale processing power and storage efficiency.

Cognitive layer: It helps to extract relevant features from the data and then perform predictive analytics to predict the outcome. This layer helps the proposed system to continuously validate new data and boost the decision-making ability over time, which makes IoT applications more intelligent and flexible.

Application layer: It is the final layer that provides services to end-users using the data stored in the CIoT layer. This layer helps users to interact with the system through key information in an effective manner.

1.3 Literature Review on Cyber Attack Detection Methods

In 2024, Deepanramkumar *et al.* [1] have described a Quantum-Secured IoT communication platform for 6G Cognitive Radio Networks (CRNs) to satisfy the rising need for safe and reliable connectivity. Significant features of this architecture involve dual-layer authentication using Secure frequency access laws, appropriate beamforming techniques, and a private key. Additionally, the framework uses a Reinforcement Learning-based Ensemble Regression (RL-ER) model for wavelength sensing. The framework substantially boosts prediction accuracy, decoding and encrypting times, and error rates. This boosts IoT networks with increased signal coverage leads to less latency and strong security.

In 2023, Mishra *et al*. [2] have developed a Congi-Secur, which was employed as a decentralized blockchain and used reinforcement learning to address the issues related to security problems. Hyperledger Fabric was utilized as a blockchain platform that boosts the performance of the proposed scheme. This architecture imitates the cognitive behaviours of miners in a distributed environment and was evaluated for developing the mining agent, which achieved the optimal cumulative reward by minimizing the maximum number of data threats.

In 2025, Tarasenko *et al.* [3] have developed a secure information system model for cognitive data processing in an IoT sensor network with two suitable factors. To boost the data security, two-operand asymmetric operations were performed during transmission. Different algorithms were developed to protect the data transmission from the database to the processing

unit, which address the different security requirements and the overall system structure within the IoT sensor network.

In 2022, Le *et al.* [4] have proposed an AI-aided network to boost trust management and privacy preservation in 5G networks. Based on this analysis, the developed framework ensures secure communication and reliable computation within the 5G system and aims to support secure operation in various sectors.

In 2024, Lilhore *et al*. [5] have developed a hybrid model for 5G networks intrusion detection that utilizes transfer learning and a more effective, lightweight CNN design, MobileNetV3-SVM. The suggested model was integrated with a multi-layered structure, which allows it to sequentially learn from raw network data. Here, different state-of-the-art techniques were compared for maximizing the efficacy of intrusion detection in situations where resources are scarce, such as those related to the IoT and fast 5G networks. The suggested hybrid model uses a lightweight MobileNet to handle network packets in real-time, lowering computational costs and making it appropriate for IoT and 5G edge devices.

1.4 Research Gaps and Contribution

The research gap identified in existing cyber-attack methods and its contribution using the proposed model are presented in the following points.

- ✓ Traditional cyber-attack detection models are highly computationally intensive and unsuitable for resource-constrained devices. To address this issue, the proposed model integrates a distilled transformer architecture that compresses the transformer model while maintaining its learning capability, which boosts its efficient cyber-attack detection process with less computational overhead.
- ✓ Most traditional models do not capture long-range dependencies and contextual features from the data. Hence, the proposed model uses a contextual-learning mechanism to highlight relevant features in the data, which boosts its detection accuracy.
- ✓ Traditional models process all input features simultaneously, which leads to feature redundancy and reduces detection performance. The proposed model integrates an attention mechanism that dynamically assigns weights to features and focuses on the most relevant ones.
- ✓ Most traditional models have limited performance in large-scale IoT networks. The proposed model used a computationally efficient network that scales effectively across large IoT communication scenarios.
- ✓ Due to a separate feature extraction and detection process, traditional models reduce overall efficiency. The proposed model uses a hybrid architecture that performs end-to-end feature extraction and detection simultaneously.

Thus, the proposed framework addresses key research gaps by integrating contextual learning, feature extraction and detection, which boosts robustness and scalability in CIoT networks.

2. Proposed Transformer-Driven Cyber-Attack Detection Framework in Cognitive IoT

2.1 Graphical Illustration of Developed Cyber-Attack Detection Framework

The developed cyber-attack detection model starts with IoT data collection, which contains both benign data and malicious attack activities like botnet attacks, spoofing, and abnormal communication. These collected data serve as a primary input for further processing.

The collected data are applied to HA-DT for the detection of cyber-attacks in CIoT. The most relevant features of the network traffic data are selected to reduce data dimensionality and boost computational efficiency. Due to this reason, the proposed CIoT network is more suitable for maintaining high detection accuracy. In this proposed scheme, a hybrid boosting mechanism is integrated to learn complex patterns that form a strong predictive model, which successfully differentiates normal and malicious activities. The hybrid boosting strategy also boosts the framework's robustness by reducing the detection error. Then, the attention mechanism is integrated to identify the most significant features related to cyber-attack detection. It assigns different weights to each feature based on its importance and helps the proposed framework to detect complex attack patterns.

Finally, a distilled transformer performs contextual learning on the network data, which helps to identify the cyber-attack patterns with improved accuracy. Thus, the proposed framework determines whether the network data are benign or malicious in CIoT. This proposed framework supports the security system by allowing a quick response to potential attacks and

sustaining its reliability for a better communication system. The overall architecture diagram of the proposed framework is included in Fig. 1.

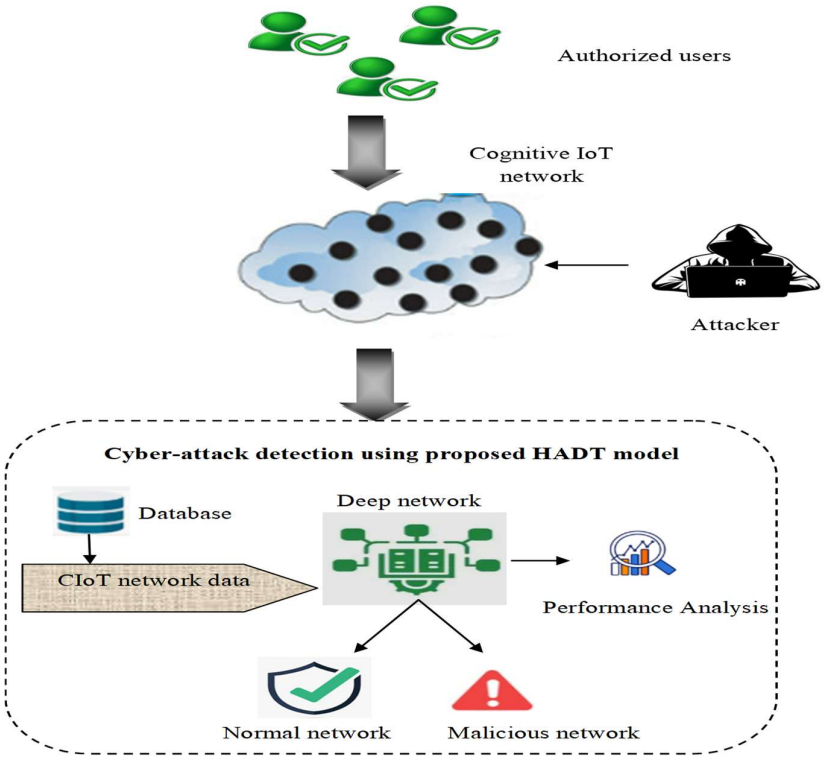

Fig. 1. Overall Architecture Diagram of the Proposed Model

2.2 Dataset Description

Two different datasets are used in this research work to validate the effectiveness and capability of the proposed HA-DT module. The description of these datasets is included below.

Dataset 1(malware/benign permissions Android): In this dataset, the data are generated by creating a binary value of permission. Here, the value 1 is for malware (specific types of attack) and 0 for non-malware. Thus, by analyzing the permission patterns, the proposed framework can detect malicious attacks accurately, and this dataset is commonly used for malware detection systems and IoT-security frameworks.

Link: https://www.kaggle.com/datasets/xwolf12/datasetandroidpermissions accessed on 2026-03-10.

Dataset 2 (Malicious URLs dataset): This dataset is commonly used for detecting serious threats in cybersecurity. This dataset contains URLs that are labelled based on their types and enable the proposed model to classify whether the URL is safe or malicious. It contains 650,000 URLs that are collected from different cybersecurity sources. Each URL is identified based on its threat types.

Link: https://www.kaggle.com/datasets/sid321axn/malicious-urls-dataset accessed on 2026-03-10.

2.3 Hybrid Boosted Attention Module

To boost the performance of the neural network, hybrid boosted attention networks are used by selecting only the most relevant features in the input data. Here, spatial and channel attention are included for performing efficient cyber-attack detection. The description of both attention mechanisms are include in the points below.

Spatial attention [20]: It focuses on the most relevant features by assigning higher weights to each pixel in the area related to the task. The input is processed to extract features at each pixel, and for each pixel, an attention score is calculated based on its importance. A higher attention score is obtained for relevant features by suppressing the irrelevant features. The average value of each pixel and the maximum pooling features are calculated to focus on the discriminative features.

The two outputs are concatenated and passed through a convolutional operation followed by a sigmoid activation function. The output representation of each pixel in the feature map is included in Eq. (1).

$$G_c^{'} = Conv\left(G_c^{avg}, G_c^{\max}\right) \tag{1}$$

$$G_c = Sig\left(G_c^{'}\right) \tag{2}$$

Here, the term G_c^{avg} represents the average pooling feature map and $G_c^{\max}$ represents the max pooling feature map. The sigmoid function converts the output into normalized weights. Finally, the obtained outcome is multiplied by the input feature map to produce the final outcome.

$$Y = G_c^{'} * y^m \tag{3}$$

This attention network allows the model to identify a specific region where a suspicious pattern occurs, which helps the system to detect attacks more accurately.

Channel Attention: This attention mechanism focuses on the channel of feature maps and assigns weights to the features across different channels.

The most commonly used channel attention approach is SE block, which contains two elements: squeezing and excitation. The spatial dimensions of the feature map are integrated to form a global feature of each channel during the squeeze stage. Next, 1×1 convolutional layer is used to learn the attention weights for each channel. Finally, normal activation function is used to amplify the channels that are more informative and reduce the less relevant features.

2.4 Distilled Transformer Network

In this proposed framework, a distilled transformer [21] is used to process the input data, which is mathematically expressed in Eq. (4).

$$F_m \in \Re^{V\times B} \tag{4}$$

Here, the term V represents the length of the time series and B represents the number of features extracted from each device. These features describe different network characteristics and security-related information. The input features are projected into a higher-dimensional space $V_m^{'}$, which boosts the representation capability of the proposed framework using Eq. (5).

$$V_m^{'} = V_m W_h + V_b \tag{5}$$

Here, the term W_h represents the weighted projection matrix and V_b represents the bias vector, which transforms the original input feature into a richer representation, which is more suitable for a transformer model. Thus, the distilled transformer is effective for handling the time-series data, which provides faster training and quick interferences than traditional models. Finally, the output of hybrid attention with a distilled transformer is used to generate multi-class intrusion detection, which produces a probability distribution over attack classes in resources-constrained settings like CIoT.

2.5 Developed HA-DT for Cyber-Attack Detection

Proposed HA-DT is an advanced network, which is designed to detect cyber-attacks in CIoT network. The proposed HA-DT network is integrated with transformer-based attention, boosting technique, and knowledge distillation process. This network successfully processes sequential data and mainly focuses on suspicious patterns for real-time applications.

The proposed framework begins with raw data from the CIoT, and each data point is transformed into feature vectors that are suitable for embedding in the transformer. In traditional models, the attention weights are computed equally for all input feature vectors, which is computationally efficient, but makes it difficult to select the most relevant data. Proposed HA-DT addresses this issue by integrating a boosted attention mechanism that helps to select the features related to malicious attacks. This boosting attention mechanism integrates both spatial and channel attention, which ensures that the system captures the global context of the input sequence by merging the relevant features from the spatial and channel attention networks. The main component of HA-DT is a stack of transformer encoders, which contains multi-head self-attention, feed-forward networks, and residual connections. The interaction among different parts of input sequences is captured through multi-head self-attention, and then the feed-forward network applies a position-wise dense network to transform the attention output. The stable training of the proposed network is ensured by using the residual connection that helps to reduce the gradient vanishing issues. Usually, transformers are computationally expensive, which is difficult in CIoT with only limited resources. Hence, a distilled transformer is used for obtaining

high detection accuracy with a few parameters and noise. Thus, the integration of a distilled transformer ensures that the proposed HA-DT model is computationally efficient for IoT devices with limited resources. Finally, the output layer generates the classified outcome as a normal network and a cyber-attack network. The diagrammatic representation of the proposed HA-DT for the cyber-attack detection process is included in Fig. 2.

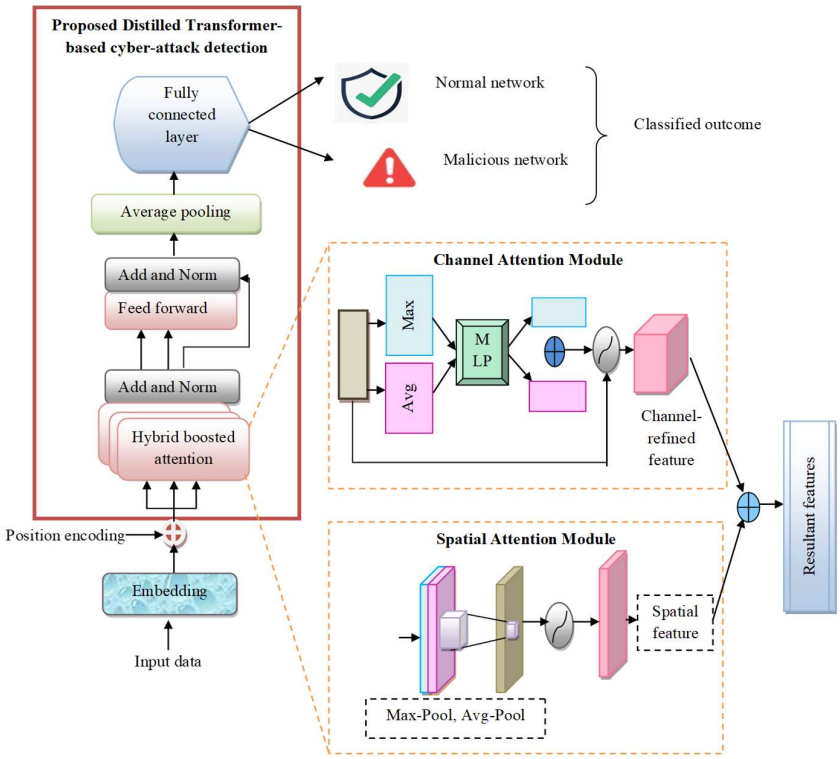

Fig. 2. Diagrammatic representation of Proposed HA-DT for Cyber-Attack Detection

3. Results and Discussion

3.1 Experimental Setup

An effective cyber-attack detection model was developed and implemented using Python software. The experiments were conducted with the support of a highly equipped computing environment. Here, the proposed HA-DT module utilizes a distilled transformer boosted with hybrid attention to capture the significant network patterns rated to cyber-attacks. During training, the proposed HA-DT module was optimized with different batch sizes and hidden neuron counts. The performance of the proposed model was evaluated using different classification metrics like accuracy, precision, FPR and specificity. The resultant outcomes were compared with baseline models like 1DCNN [16], RNN [17], LSTM [18] and GRU [19] to determine the robustness of the proposed HA-DT model in cyber-attack detection.

3.2 Key Performance Analysis of Proposed Cyber-attack Detection Model

Fig. 3 provides the performance analysis of the proposed HA-DT model compared with traditional models under different neuron counts. Here, the proposed HA-DT model achieves a higher accuracy of approximately 98% at a 250 neuron count. This higher accuracy highlights the detection ability of the distilled transformer to focus on essential patterns in network data. If the neuron count is increased, the FPR value decreases for all traditional models, but the proposed HA-DT module achieved the most significant reduction, which indicates its ability to minimize incorrect attack detection.

Moreover, the precision performance indicates that the proposed HA-DT module achieved a high precision of 96.7% at a 150 neuron count, which indicates that the model produces more accurate detection of cyber-attacks in CIoT. Similarly, the specificity metric shows that the proposed HA-DT module outperformed traditional models by obtaining high specificity across

all neuron counts. Thus, the overall analysis proves that the proposed HA-DT module significantly boosts cyber-attack detection performance by obtaining better accuracy, precision, and specificity with minimum error rate.

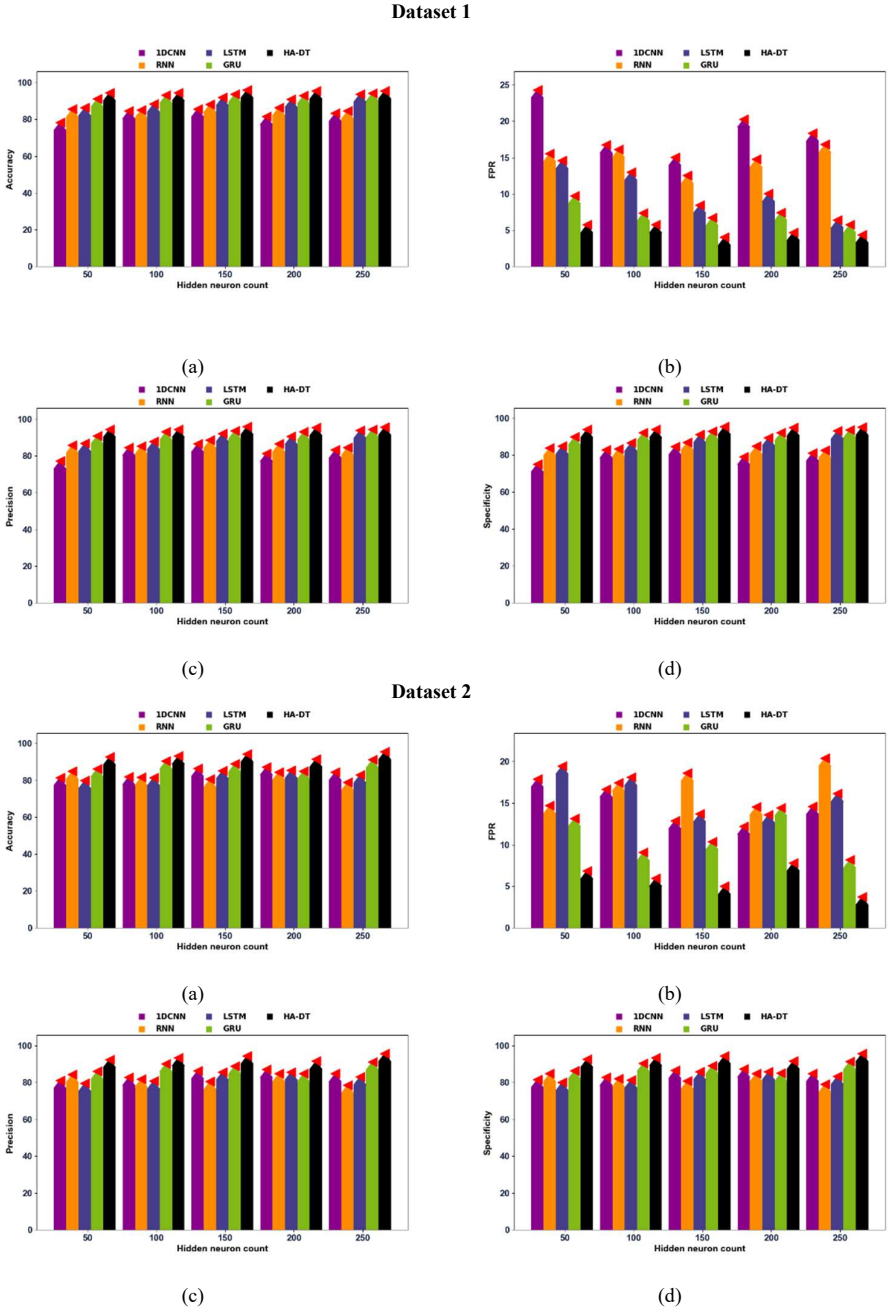

Fig. 3. Diagrammatic representation of Proposed HA-DT for Cyber-Attack Detection

3.3 Accuracy Analysis of Proposed Cyber-attack Detection Model

Table I provides the numerical outcomes of the accuracy analysis using the proposed cyber-attack detection model, and its performance is compared with traditional models. Using Dataset 1, the proposed HA-DT module consistently achieved higher accuracy across all batch sizes. At the batch size of 4, the proposed HA-DT module obtained 94.8% accuracy, which is better than traditional models like 1DCNN, RNN, LSTM, and GRU. If the batch size increased to 16, the proposed HA-DT module achieved 96.4% accuracy, indicating that the designed module effectively captures discriminative features from the raw data. Even at batch size 32 and 48, the proposed HA-DT module obtained accuracies of 95.8% and 96% respectively, which highlight

the stability and robustness of the proposed HA-DT module, thereby boosting the flexibility of cyber-attack detection in CIoT.

Table .1 Accuracy Analysis of the Proposed Cyber-Attack Detection Model

Batch sizes	1DCNN [16]	RNN [17]	LSTM [18]	GRU [19]	HA-DT
Dataset 1					
4	88.8	88.48	88.16	92.96	94.88
8	79.68	87.2	85.28	88.96	94.88
16	81.92	82.88	92.64	91.52	96.48
32	90.08	82.88	88.32	89.92	95.84
48	86.4	81.6	93.92	89.92	96
Dataset 2					
4	85.12	77.92	81.12	92.32	93.12
8	88.32	88.96	90.56	93.44	93.6
16	87.36	84	87.84	87.84	94.72
32	83.04	80.32	89.6	92	92
48	78.88	83.68	89.44	93.28	96

4. Conclusion

An AI-based framework was developed for effective cyber-attack detection in the CIoT network. The proposed HA-DT module integrates a distilled transformer with a hybrid attention mechanism to boost the detection capability by maintaining the computational efficiency. The proposed HA-DT module was designed to process raw network data and successfully classify the normal and malicious activities in the CIoT network. Experiments were conducted and compared with traditional models like 1DCNN, RNN, LSTM, and GRU. The resultant outcomes prove that the proposed HA-DT module obtained 96.4% accuracy at a batch size of 16 with a low error rate. This better accuracy indicates that the proposed HA-DT module provides a robust and efficient solution for cyber-attack detection in the CIoT environment in terms of detection accuracy and flexibility.

Research limitation and future scope: The proposed HA-DT module achieved higher accuracy, but the computational complexity during the training process is slightly higher, particularly in large-scale IoT data. The use of a distilled transformer reduces model size, but the integration of hybrid boosted attention leads to increased training time. To address these issues, an edge-optimized version of the model would be suitable for real-time cyber-attack detection. Also, the framework within cloud collaborative architecture will boost the scalability and flexibility of the CIoT network against cyber threats.

References

[1] Deepanramkumar, P. and Sharmila, A.H., "AI-enhanced quantum-secured IoT communication framework for 6G cognitive radio networks," IEEE Access, vol.12, pp.144698-144709, 2024.

[2] Mishra, S., Chakraborty, S., Sahoo, K.S. and Bilal, M., "Cogni-Sec: a secure cognitive enabled distributed reinforcement learning model for medical cyber–physical system," Internet of Things 24, pp.100978, 2023.

[3] Tarasenko, Y., Chervotoka, O., Orlov, S., Lada, N., Shapoval, V. and Piskozub, A., "A model of a secure information system for cognitive data processing in IoT sensor networks for laboratory climatic testing," 2025.

[4] Le, T. and Shetty, S., "Artificial intelligence-aided privacy preserving trustworthy computation and communication in 5G-based IoT networks," Ad Hoc Networks 126, pp.102752, 2022.

[5] Lilhore, U.K., Dalal, S. and Simaiya, S., "A cognitive security framework for detecting intrusions in IoT and 5G utilizing deep learning," Computers & Security, vol.136, p.103560, 2024.

[6] Ronanki, S.K. and Killi, B.R., "Machine Learning based Cyber Attack Detection and Classification in O-RAN," Journal of Network and Systems Management, vol. 34, no. 2, pp.41, 2026.

[7] Jullian, O., Otero, B., Rodriguez, E., Gutierrez, N., Antona, H. and Canal, R., "Deep-learning based detection for cyber-attacks in iot networks: A distributed attack detection framework," Journal of Network and Systems Management, vol. 31, no. 2, pp.33, 2023.

[8] Rajkumar, S., Sheeba, S.L., Sivakami, R., Prabu, S. and Selvarani, A., "An IoT-based deep learning approach for online fault detection against cyber-attacks," SN Computer Science 4, no. 4, pp.393, 2023.

[9] Bitirgen, K. and Filik, Ü.B., "A hybrid deep learning model for discrimination of physical disturbance and cyber-attack detection in smart grid," International Journal of Critical Infrastructure Protection 40, pp.100582, 2023.

[10] Khoa, T.V., Hoang, D.T., Trung, N.L., Nguyen, C.T., Quynh, T.T.T., Nguyen, D.N., Ha, N.V. and Dutkiewicz, E., "Deep transfer learning: A novel collaborative learning model for cyberattack detection systems in IoT networks." IEEE Internet of Things Journal 10, no. 10, pp.8578-8589, 2022.

[11] Raghunath, K.K., Kumar, V.V., Venkatesan, M., Singh, K.K., Mahesh, T.R. and Singh, A., "XGBoost regression classifier (XRC) model for cyber attack detection and classification using inception V4," Journal of Web Engineering 21, no. 4, pp.1295-1322, 2022.

[12] Behera, B.B., Mohanty, R.K. and Pattanayak, B.K., "A deep fusion model for automated industrial iot cyber attack detection and mitigation," International Journal of Electrical and Electronics Research 10, no. 3, pp.604-613, 2022.

[13] Huma, Z.E., Latif, S., Ahmad, J., Idrees, Z., Ibrar, A., Zou, Z., Alqahtani, F. and Baothman, F., "A hybrid deep random neural network for cyberattack detection in the industrial internet of things," IEEE access 9, pp.55595-55605, 2021.

[14] Upadhyaya, B., Sun, S. and Sikdar, B., "Multihypothesis sequential testing for illegitimate access and collision-based attack detection in wireless IoT networks," IEEE Internet of Things Journal 8, no. 14, pp.11705-11716, 2021.

[15] Gopalakrishnan, T., Ruby, D., Al-Turjman, F., Gupta, D., Pustokhina, I.V., Pustokhin, D.A. and Shankar, K., "Deep learning enabled data offloading with cyber attack detection model in mobile edge computing systems," IEEE Access 8, pp.185938-185949, 2020.

[16] Shankar, T.S., Shamsudeen, K.V. and Ramadass, R., "1D-CNN-based Real-Time Network Intrusion Detection with Privacy-Preserving for IoT," Journal of Wireless Networks and Communication Systems, 2025.

[17] Kwon, S., Yoo, H. and Shon, T., "IEEE 1815.1-based power system security with bidirectional RNN-based network anomalous attack detection for cyber-physical system," IEEE Access 8, pp. 77572-77586, 2020.

[18] Yeom, S., Choi, C. and Kim, K., "LSTM-based collaborative source-side DDoS attack detection." IEEE Access 10, pp.44033-44045, 2022.

[19] Ghani, A.A. and Alasadi, S.A., "A Deep Learning Algorithm to Cybersecurity: Enhancing Intrusion Detection with a Hybrid GRU and BiLSTM Model," Engineering, Technology & Applied Science Research, vol.15, no. 3, pp.23605-23612, 2025.

[20] Dang, L., Weng, L., Dong, W., Li, S. and Hou, Y., “Spectral-Spatial Attention Transformer with Dense Connection for Hyperspectral Image Classification," Computational Intelligence and Neuroscience 2022, no. 1, pp.7071485, 2022.

[21] Nadiah, A.N., Alamri, A., Aljuhani, A. and Kumar, P., "Transformer-based knowledge distillation for explainable intrusion detection system," Computers & Security 154, pp.104417, 2025.