Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 22 No. 4 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v22i04.58781

PAPER

Detecting MITM Attacks Using DNN in IIoT Substation Systems

Dwi Prasetya[1,2] (✉), Dendi Renaldo Permana[2], M. Rafie Al Hamas[2], Deris Stiawan[2], Tami A. Alghamdi[3], Rahmat Budiarto[3]

[1]PT PLN (Persero) UP3 Lahat, Lahat, Indonesia

[2]Sriwijaya University, Palembang, Indonesia

[3]Al-Baha University, Al Baha, KSA

dwi.prasetya@pln.co.id

ABSTRACT

The integration of the Industrial Internet of Things (IIoT) in electrical substation systems has improved efficiency in operations but brought them under greater exposure to cyber threats, such as increased vulnerability to cyberattacks, particularly man-in-the-middle (MITM) attacks where information is altered and grid stability is affected. A deep neural network (DNN) structure dedicated to identifying MITM attacks on IIoT substation environments is presented in this paper. A large dataset of normal and attack network traffic was acquired by using a SCADA simulator to generate a realistic operating scenario. With 99.78% accuracy and ideal precision, recall, and F1-measures of classifying attack traffic, the proposed DNN model exhibits superior classification performance. An ontology that converts network anomalies into actionable operational insights for operators is used to visualize the detection results in an attempt to improve interpretability. Contextual visualization and correct anomaly detection cooperate to form a strong and valuable cybersecurity solution that safeguards critical infrastructure against sophisticated cyberattacks.

KEYWORDS

deep neural network (DNN), Industrial Internet of Things (IIoT), man-in-the-middle (MITM), intrusion detection, SCADA, electrical substation

1 INTRODUCTION

Power substation automation and real-time data monitoring are made possible by the Industrial Internet of Things (IIoT), which has completely transformed critical infrastructure monitoring [1], [2], [3]. However, the incorporation of IIoT into operational technology networks has introduced severe cybersecurity risks, especially to man-in-the-middle (MiTM) attacks, where hackers covertly intercept or modify device-to-device communication [4], [5], and [6]. These attacks can lead to equipment damage and grid instability by manipulating control commands or sensor data [3], [7].

Prasetya, D., Permana, D. R., Al Hamas, M. R., Stiawan, D., Alghamdi, T. A., Budiarto, R. (2026). Detecting MITM Attacks Using DNN in IIoT Substation Systems. *International Journal of Online and Biomedical Engineering (iJOE)*, 22(4), pp. 155–170. https://doi.org/10.3991/ijoe.v22i04.58781

Article submitted 2025-09-22. Revision uploaded 2026-01-03. Final acceptance 2026-01-06.

Lightweight machine learning (ML) or rule-based approaches used by next-generation IIoT intrusion detection systems (IDS) are typically ineffective against MiTM due to imbalanced datasets and changing attack patterns [8], [9], [10]. Whereas ensemble ML algorithms [8], [12] and blockchain tech [1], [11] are promising, they remain too general for the attacks. Although hybrid CNN-LSTM models [13], [14], [15] surpass conventional methods, recent developments in deep neural network (DNN) ignore the distinctive spatiotemporal patterns of MiTM attacks in substations [4], [6].

A significant gap still exists in operational context-aware detection. Although they lack IIoT-specific implementations, ontology-based systems (such as for email threats [16] or ransomware [17]) enhance explainability. For instance, topic modeling of [18] suggests that such systems overlook the power grid dynamics, where physical infrastructure is affected by attacks [2], [3]. This substation operator decision-making gap between algorithmic detection and actionable information is also not filled by edge-computing intrusion detection systems [19] and risk assessment tools [20].

We introduce a DNN-ontology fusion framework optimized for MiTM detection in IIoT substations as an attempt to overcome these limitations. Our method, in contrast to previously [13], [14], combines a substation-focused ontology that brings back DNN predictions to understandable operating knowledge [16], [17] with multi-layer feature extraction for MiTM-pertinent anomalies [4], [5], [7]. By identifying anomalies and correlating them to relay command manipulation or voltage tampering, say, such a two-pronged approach both identifies attacks and explains their effect [2], [3].

Our first innovation is the integration of detection and interpretation. Although very accurate detection is realized, the attention-based DNN introduced by [15] is not useful for visualization. Ontology, however, has no alignment with contemporary detectors and is functionally identical to [17]. Weaving together these threads, we present an end-to-end solution and test it on actual substation data that addresses the call from Hassan et al. [21] for AI security in critical infrastructure tailored to the domain through the following steps:

1. adapting DNNs to MiTM's unique substation signatures,
2. providing the first ontology-based visualization for MiTM attacks for IIoT, and
3. demonstrating to outperform lightweight ML [8], [10], and generic DNNs [13], [14].

This study enhances IIoT security. Adversarial training for evolving threats [5], [7] and federated learning for privacy [11] are some of the potential future avenues.

2 LITERATURE REVIEW

Significant cybersecurity risks have been introduced through the integration of IIoT into electrical substations, especially from covert MITM attacks that can tamper with sensor data or control instructions, resulting in equipment failure or grid instability [3], [7]. Because these attacks are evasive and dynamic, traditional IDS using rule-based approaches or lightweight ML have difficulty recognizing them and tend not to adapt to the peculiarity of substation network traffic spatiotemporal patterns [5], [22]. The need for targeted detection mechanisms is evidenced by the fact that, despite improved data integrity, technology on the blockchain cannot secure against real-time packet modification during transmission [1], [11].

Self-attention networks and CNN-LSTM combinations are two of the latest breakthroughs in DNNs that have shown spectacular performance when detecting advanced IIoT intrusions with reported accuracy levels exceeding 99% [13], [14], [15]. Yet in substation settings, where small latency or data coherence anomalies

can signal malicious activity, these methods tend to be non-specific to attacks and are applicable across a set of attack types [4], [6]. Comparisons of their relative efficacy against MITM attacks have demonstrated that lightweight ML methods, such as feature-importance ensembles, suffer from imbalanced datasets and limited capacity to learn evolving threats [8], [9], [10].

A significant gap in published literature that is missing interpretable models relating detection results to operational insights for substation operators exists. Ontology-driven threat visualization models, which have already been used in other domains like ransomware and email security [16], [17], [18], do not possess the operational scenario required for power systems. For example, while [14] and [15] suggest sophisticated DNNs be applied for intrusion detection, they do not enable domain ontologies for visualizing substation network threats. As a result, operators are unable to link physical impacts, like tampering with relay commands or voltage readings simulations, with the identified anomalies [2], [3].

Interventions for coping with these challenges have involved adversarial training to deal with dynamic threats and federated learning in order to protect privacy but tend to introduce new vulnerabilities or improper computational costs for resource-limited substation devices [6], [11], [19]. Although autoencoder-based anomaly detection holds promise, it is unable to offer the substation-related information necessary for proper risk assessment [12], [23]. The need is prompted by contrastive assessments that show that lightweight machine learning techniques are insufficient for identifying subtle anomalies in protocols such as Modbus/TCP or DNP3 [5], [9], [24].

Through combining multi-layer feature extraction with ontology-driven visualization, the paper solves the above shortcomings by introducing a DNN-based framework optimized for attack detection of IIoT substations. To remain relevant to substation operation in the real world, the model is trained on a dataset created through controlled MITM attack simulations on an SCADA-emulated testbed [5], [6]. The framework not only identifies attacks but also offers an explanation of their operational impact by mapping DNN outputs to a structured knowledge graph, allowing grid operators to respond intelligently and timely [16], [17].

The study's contributions are an ontology-enhanced visualization platform, an optimized model designed to target MITM detection, and a specific network traffic-customized data preprocessing pipeline. With the provision of power utilities with a practical tool to protect critical infrastructure against advanced cyberattacks, these advancements improve the horizon of IIoT security [3], [7]. To further strengthen the model's resilience to novel attack patterns, future efforts could delve into adversarial training or federated learning in the context of large-scale deployment [5], [11].

3 METHODOLOGY

This section presents the overall methodological approach undertaken in the study, including the experimental design, attack simulation, and model development.

3.1 Research and topology design

The study followed four systematic steps as illustrated in Figure 1. Phase 1 involved establishing a testbed that replicated the operational environment of Indonesia's national electricity provider (PT PLN) with the help of SCADA simulator

hardware for mimicking realistic data capture. During Phase 2, normal and attack scenarios were captured to produce a balanced dataset. Phase 3 involved training the model and tuning the hyperparameters on various architectures of the neural network, while Phase 4 involved wide-ranging testing and validation on a number of evaluation metrics to ensure robust performance.

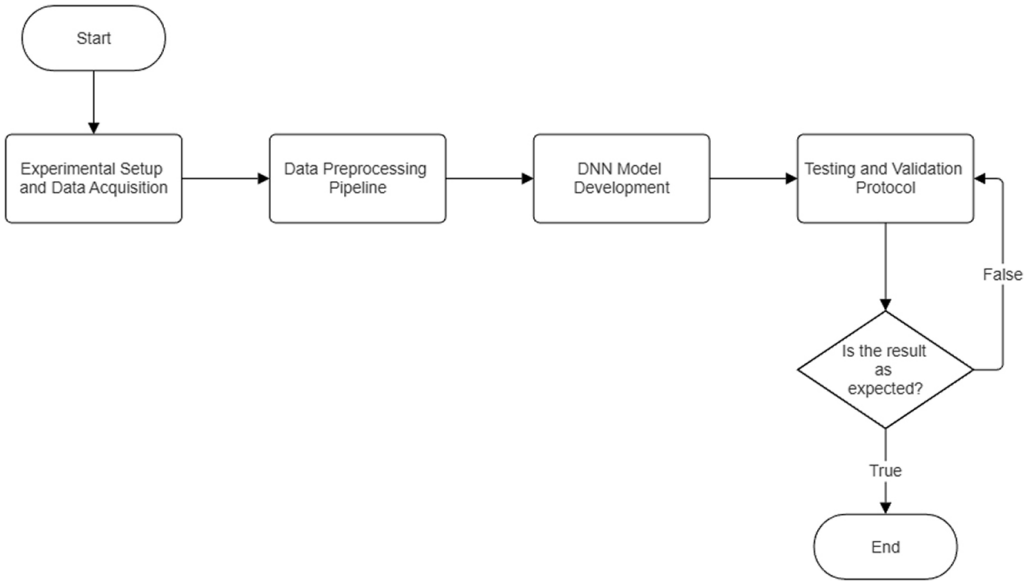

Fig. 1. Research methodology workflow

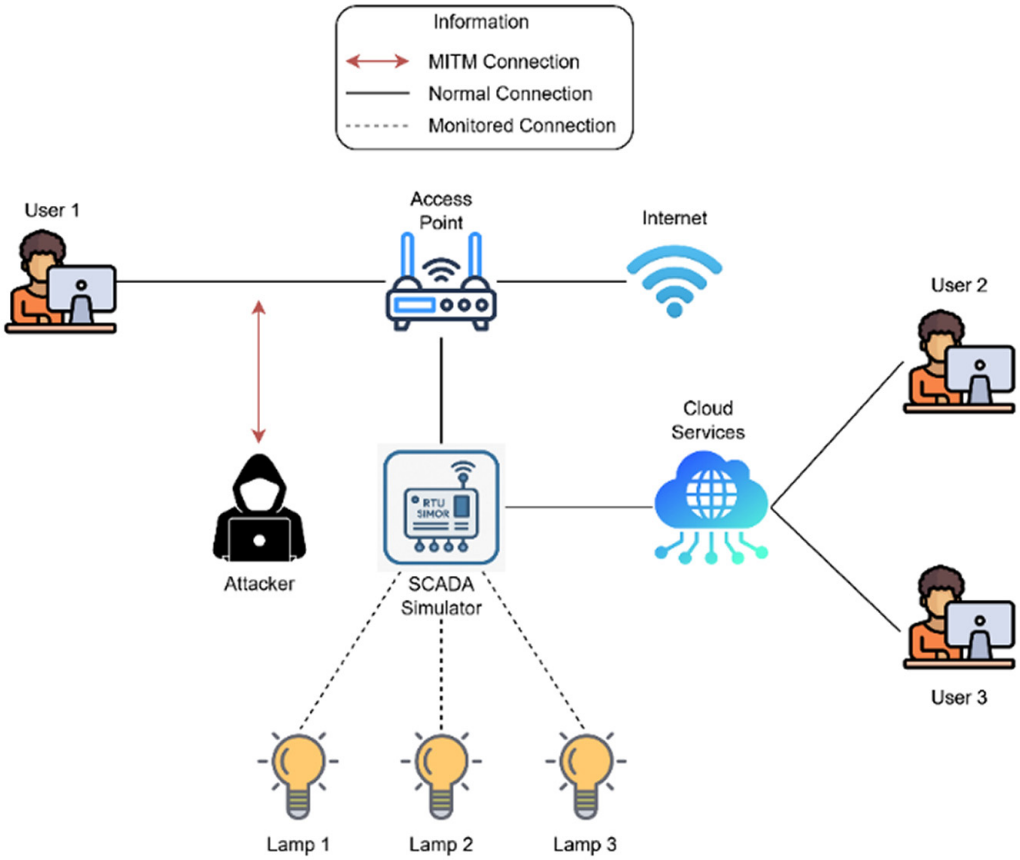

Fig. 2. Topology

A testbed environment consists of a legitimate user device (operator workstation), an attacker node that has been joined via a wireless access point, and a SCADA simulator, as the RTU is set up to simulate an actual IIoT substation environment. For ease of interception, the attacker was kept within the same network. Figure 2 illustrates this topology that yielded an actual and controlled environment in which network vulnerabilities could be tested for industrial control systems.

3.2 Scenario

By impersonating the MAC and IP addresses of the simulator and the authorized user, the researcher conducted timely ICMP Redirect attacks (Type 5, Code 1) under attack simulations. By facilitating the diversion of the traffic streams through the attacker's node, such specially crafted packets formed a covert interception channel. The hijacked communications enabled both active manipulation in the form of command injection and passive eavesdropping of unencrypted HTTP traffic. Three independent trials were used to provide dataset diversity while achieving experimental consistency, and every attack session was strictly capped at five minutes. Network traffic was widely captured with Wireshark, creating PCAP files that captured the entire attack process from start to finish.

Through ongoing communication between the valid user and IIoT systems, the control scenario set up baseline network traffic. Under these circumstances, three consecutive five-minute observation intervals were performed to create traffic patterns in normal substation operation without anomalies or security breaches.

3.3 Software and hardware

In order to replicate IIoT substation environments realistically and be in control to conduct vulnerability analysis, the experiment configuration used specially designed infrastructure with hardware and software combined. The hardware configuration, as illustrated in Table 1, had three dedicated computing nodes: a SCADA simulator (192.168.0.103), a platform for the attacker (192.168.0.102), and a victim workstation (192.168.0.100). The TP-LINK router was the network backbone. As each device had a distinct MAC address, there was potential to test accurately. The simulator (40:F5:20:A6:7F:54) and the victim were free to communicate with each other via the router (30:DE:4B:79:E2:E8), thus creating a safe environment for data exchange and interaction.

Table 1. Hardware specifications and network roles

No.	Hardware Component	MAC Address	IP Address	Function
1	Wireless Router	30:DE:4B:79:E2:E8	192.168.0.1	Central access point connecting all nodes
2	Laptop (Ryzen 5 5500U)	00:45:E2:E5:78:27	192.168.0.100	Victim device running legitimate SCADA access
3	Laptop (Ryzen 5 5600H)	C0:4A:00:2B:B8:43	192.168.0.102	Attacker platform executing MITM operations
4	SCADA Simulator	40:F5:20:A6:7F:54	192.168.0.103	Target system emulating substation RTU

The software environment, explained in Table 2, integrated seven fundamental tools for attack launching (Xerosploit), traffic inspection (Wireshark), and forensic analysis (NetworkMiner), with the shared development framework of Python for both the MITM scripts and the DNN modeling pipelines. The integration facilitated end-to-end network activity capture from preliminary scanning with Blynk IoT to attack confirmation with Snort alerts while reproduction was made simple with default configurations. Tools such as Tshark enabled conversion of the captured PCAP files into machine learning datasets, ensuring that hardware capabilities directly supported software activity in real-time attack launch, forensic examination, and dataset creation.

Table 2. Software tools and their functions

No.	Software	Function
1	Blynk IoT	IoT application serving as the research target
2	Xerosploit	Network scanning to identify connected devices MAC addresses
3	Wireshark	Network traffic capture for dataset creation
4	Tshark	PCAP to CSV data conversion tool
5	Python	Script development for MITM attacks and DNN modeling
6	NetworkMiner	Attack evidence identification in captured data
7	Snort	Attack validation through detected alerts

3.4 Feature extraction

In order to enable effective aggressive intrusion detection through multi-layered traffic analysis, this study uses a comprehensive set of 23 network traffic features, including packet-level attributes (frame_len, frame_protocols), Ethernet header details (eth.dst, eth.arc), IP layer attributes (ip.version, ip.hdr_len, ip_len, ip.id, ip.ttl, ip.proto, ip.arc, ip.dst), transport layer attributes (top.srcport, top.dstport, top.neq, top.aok, top.len, top.flags, top.window_size), web protocol-specific fields (wbp.srcport, wbp.dstport, wbp.bmpen), and a class label for classification.

4 RESULT AND DISCUSSION

Our experimental findings and implications for the use of cybersecurity in the detection of network intrusion are thoroughly investigated in this section. The findings represent how our deep neural network model identifies attack patterns, normal traffic, and request traffic, and the discussion puts the findings into perspective by contrasting them with previous work, examining real-world issues, and outlining future research directions.

4.1 Result

This section contains the significant findings of the research, outlining the characteristics of the network traffic datasets, verification of attack patterns, the

performance of the new DNN model, and observations from ontology visualization. The findings indicate that this approach effectively detects and classifies intrusion attempts with good accuracy and reliability in distinguishing between normal, rogue, and malicious traffic.

Attack pattern validation. Both prominent figures in forensic analysis have noteworthy security outcomes. Figure 3 uses anomalous network traffic and protocol flaws to show attack patterns within the data collected. By cross-referencing Snort alert timestamps with related Wireshark packet data, Figure 4 verifies malicious activity by providing the timestamps and attack signatures for them. As attested by Snort's identification of the old vulnerability CVE-1999-0265, a documented man-in-the-middle attack vector, the examination clearly targets an "ICMP Redirect Host" attack. The integrity of the dataset for security research is accurately confirmed through this multi-tool verification method, which takes advantage of NetworkMiner's forensic capabilities and Snort's intrusion detection capabilities. The validity of the dataset to work on actual attack patterns and create defenses against them is established by reproducing the results of both tools.

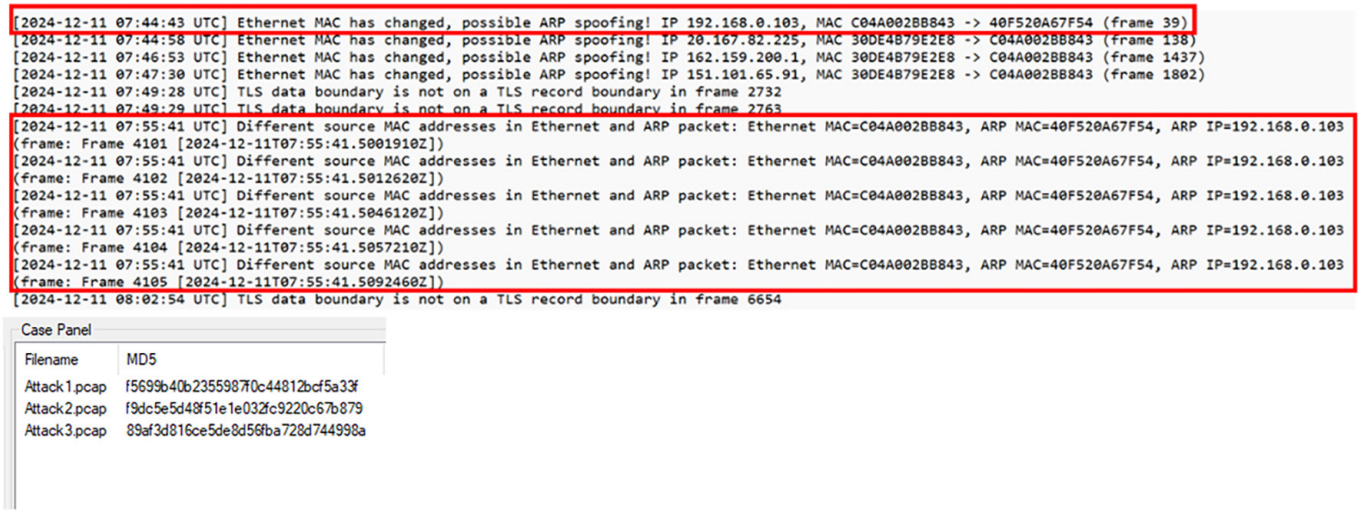

Fig. 3. Network miner result

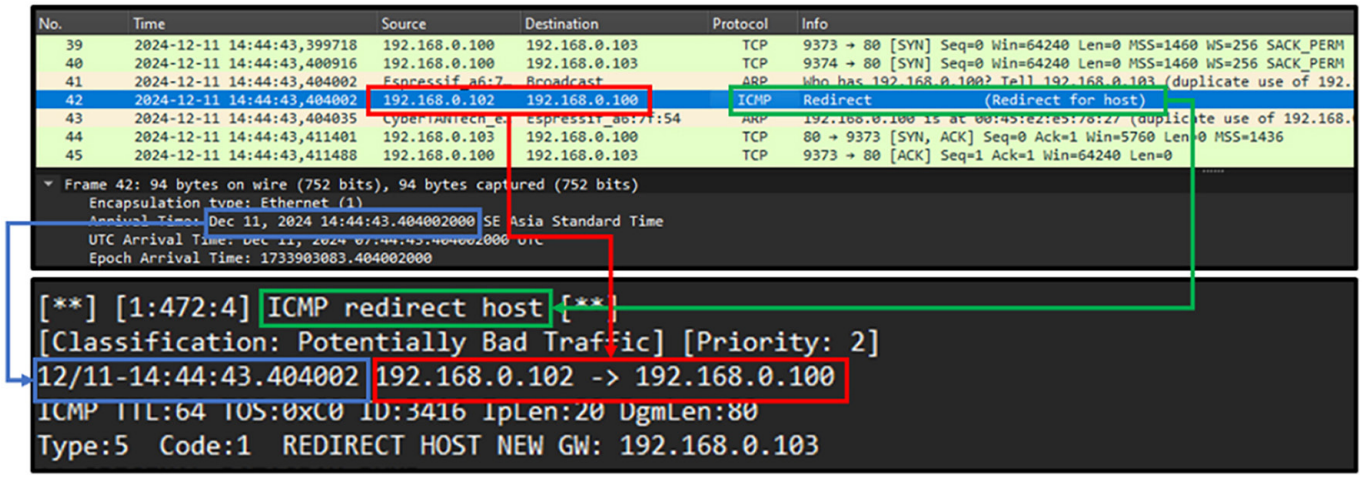

Fig. 4. Network traffic analysis

A preprocessing pipeline was used for DNN analysis based on attack patterns illustrated in Figures 3 and 4 network traffic data. 21,371 instances with 33 features were created after combining six sources of data. There was a class imbalance uncovered following the removal of two empty features: request (64.5%), normal (35.2%), and attack (0.3%). Type conversion (23 features after refinement), label encoding on categorical features (MAC, protocols, etc.), MinMax normalization on numerical variables, and oversampling to balance the classes were all examples of preprocessing. 30,278 training examples and 10,093 test examples made up the 75:25 split of the dataset.

Through the standardization of protocol-specific properties and temporal behavior encoding to identify sequential attack patterns, the transformation process addressed the complexities of network security data sets. Labels were transformed into numerical and one-hot encoded representations for facilitating multiple DNN architectures. Keeping attack signatures' integrity intact through all the transformations was a priority. 40,371 samples with one-hot encoded labels and a 30% reduction in feature space by selective retention of the most informative features were employed in the final dataset. With no loss of richness needed to uncover known and unknown intrusion patterns, this optimized format provides a balanced representation to facilitate efficient computation.

Model performance evaluation. The overall evaluation scores presented in Table 3 reveal that the DNN model performed remarkably well in network traffic classification. The model was optimized with the Adam optimizer (learning rate = 0.00001) and categorical cross-entropy loss and had hidden layers architected with [16, 32, 64] neurons. It was optimized through hyperparameter tuning with GridSearchCV. The model performed well on all three classes—attack, normal, and request traffic—and had almost flawless classification accuracy, registering a global precision, recall, and F1-score of 0.9978. The model generalized strongly without overfitting, as revealed by its flawless classification of attack instances (precision = 1.000, recall = 1.000, F1-score = 1.000) while maintaining high precision (1.000) for normal traffic and flawless recall (1.000) for request traffic. Its stability is also confirmed by the minute difference in training (99.77%) and validation (99.78%) accuracy.

Table 3. Comparative model performance: attack vs. normal vs. request traffic

Metric	Attack Class	Normal Class	Request Class	Overall
Precision	1.000	1.000	0.9934	0.9978
Recall	1.000	0.9935	1.000	0.9978
F1-Score	1.000	0.9967	0.9967	0.9978
Accuracy	–	–	–	0.9978

Its equal performance across traffic categories necessary for accurate intrusion detection is further demonstrated through the comparative evaluation in Table 4. With barely the slightest misclassifications in the normal (recall = 0.9935) and request classes (precision = 0.9934), the DNN achieved 99.78% accuracy on 10,093 samples tested. With only 0.3% attack traffic, it demonstrates how well the selected architecture and preprocessing manage class imbalance. The vast majority of instances are well-classified, with the model suitable for practical cybersecurity where both false

positives and false negatives have significant impacts, as it can accommodate high attack recall without sacrificing precision in normal traffic.

Figure 5 indicates the optimal learning behavior of our deep neural network model when trained. Both validation loss and training loss dropped drastically in the initial 20 epochs, from roughly 1.0 to well below 0.2, as evident from the loss curve of Figure 5a. It was followed by a slow stabilization process that by epoch 100 dropped to values close to zero (less than 0.01). Effective model convergence without overfitting is achieved through this steady decrease and the close alignment of training and validation loss during training. The switchover from fast feature learning to precise parameter fitting is indicated by the clean elbow point that is seen at around epoch 20.

Other than the loss analysis, the trend of accuracy in Figure 5b indicates that training accuracy and validation accuracy are improving in tandem, from an initial rate of approximately 60% to reach near-perfect levels (above 99.7%) at the final epoch. The ability of the model to generalize is also demonstrated in the virtually negligible (<0.3%) difference in training and validation accuracy for all epochs. The first 30 epochs, corresponding to the period of steep decline in loss, recorded the highest improvement in accuracy. This was then followed by gradual fine-tuning, which resulted in complete convergence. All these findings suggest that our diligently designed architecture and optimization parameters were able to learn highly discriminative features efficiently along with maintaining good generalization skills during training.

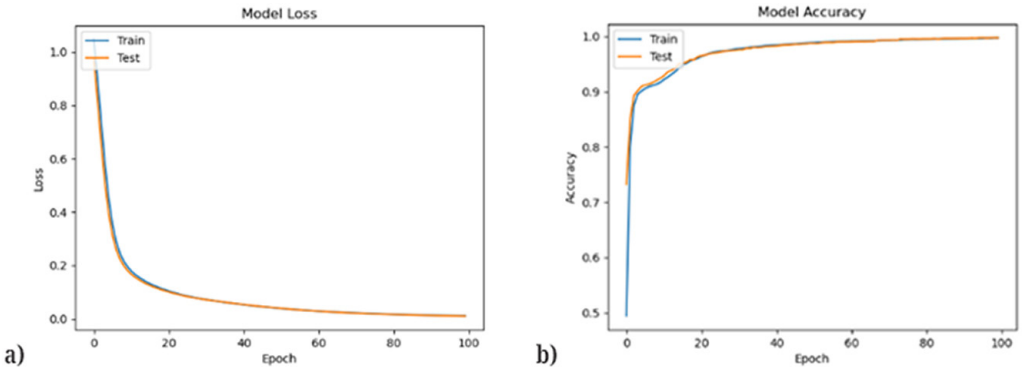

Fig. 5. Training and validation (a) loss curve and (b) accuracy curve

The confusion matrix of the test data of the DNN model is depicted in Figure 6, whereas information on True Positives (TP), False Positives (FP), False Negatives (FN), and True Negatives (TN) per class is provided in Table 4. The model achieves flawless attack detection (3,420 TP, 0 FP/FN, 100% sensitivity and specificity), high-quality normal traffic classification performance (3,349 TP, 22 FN, 0.65% error), and strong request identification (3,302 TP, 22 FP, 0.66% error). These performances are substantiated by incredibly stable TN values across all classes (Attack: 6,673; Normal: 6,722; Request: 6,769), confirming the reliability and appropriateness of the model in security-critical settings where false negatives are not allowed. In addition, the tiny difference in performance between default and request classes (0.01% error gap) indicates exemplary multi-class balance.

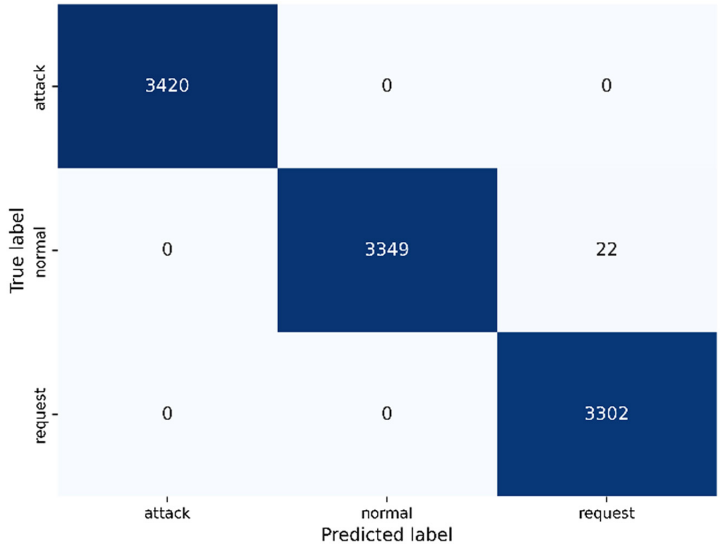

Fig. 6. Deep neural network confusion matrix results

Table 4. TP, FP, FN, and TN values by class

No.	Label	TP	FP	FN
1	Attack	3420	0	0
2	Normal	3349	0	22
3	Request	3302	22	0

Ontology visualization and operational insights. The DNN model predictions are used to create ontology visualization in Figure 7, where network entities are labeled using a color-coding system according to the results of classification. As presented in Tables 5 and 6, the visualization uses the graph model with nodes and edges, both with independent meanings. While edges are marked as attacker (red), normal (black), or request (green) connections, nodes are marked as attacker (red), normal (light blue), or request (white). The patterns of communications in the network and the threats that may arise can easily be noticed because of this straightforward and easy-to-read layout. In agreement with the high values shown in the confusion matrix, the model has good predictive accuracy, properly marking normal traffic and attacker nodes.

Table 5. Legend for node classification visualization

No.	Label	Color	Description
1	Attacker	Red	Nodes predicted with this label will be colored red.
2	Normal	Light Blue	Nodes predicted with this label will be colored light blue.
3	Request	White	Nodes predicted with this label will be colored white.

Table 6. Legend for network connection (Edge) visualization

No.	Label	Color	Description
1	Attacker	Red	Edges predicted with this label will be colored red.
2	Normal	Black	Edges predicted with this label will be colored black.
3	Request	Green	Edges predicted with this label will be colored green.

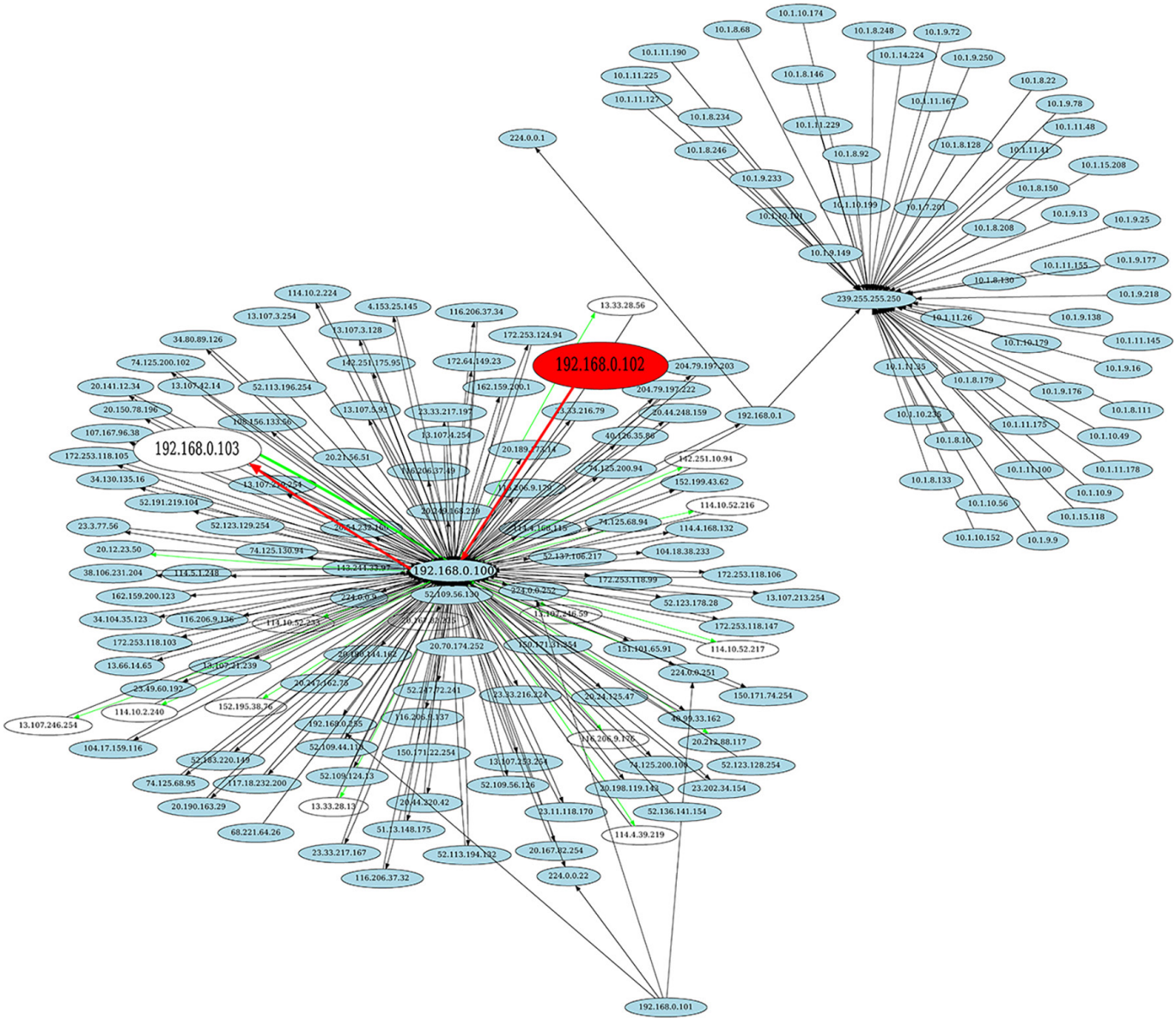

Fig. 7. DNN-based network anomaly detection: ontology visualization and performance analysis

Figure 8 provides a simplified ontology visualization that retains only the essential nodes and primary relationships to offer a clearer overview of network communication patterns while maintaining the same DNN-based color-coding scheme used in Figure 7. Although this simplified representation results in minor differences in the number of nodes and edges displayed, it does not affect the core analytical outcome, as the crucial information regarding attacker IP nodes and their target destinations remains fully preserved, particularly visible through red edges directed toward white nodes, indicating the corresponding victims. By applying the same semantic interpretation of colors for nodes and edges as in Figures 7 and 8 ensures consistent comprehension of attacker, request, and normal behaviors while presenting the communication structure in a more accessible and easy-to-interpret manner.

There remain some small misclassifications taking place despite good overall accuracy; e.g., normal nodes being incorrectly identified as request nodes (white), in agreement with the previously reported error rate of 0.66%. These differences

only manifest in edge cases of cloud service IPs or multicast addresses, and these are possible avenues for model improvement. The visualization is still a valuable tool for network security, conveying simple information regarding the regular traffic and threat propagation. The utilization of colored edges (requests green, attacks red) simplifies threat hunting and incident response prioritization. The DNN is robust for large-scale monitoring and anomaly detection in practical applications, as suggested by its stable performance both on internal and external servers.

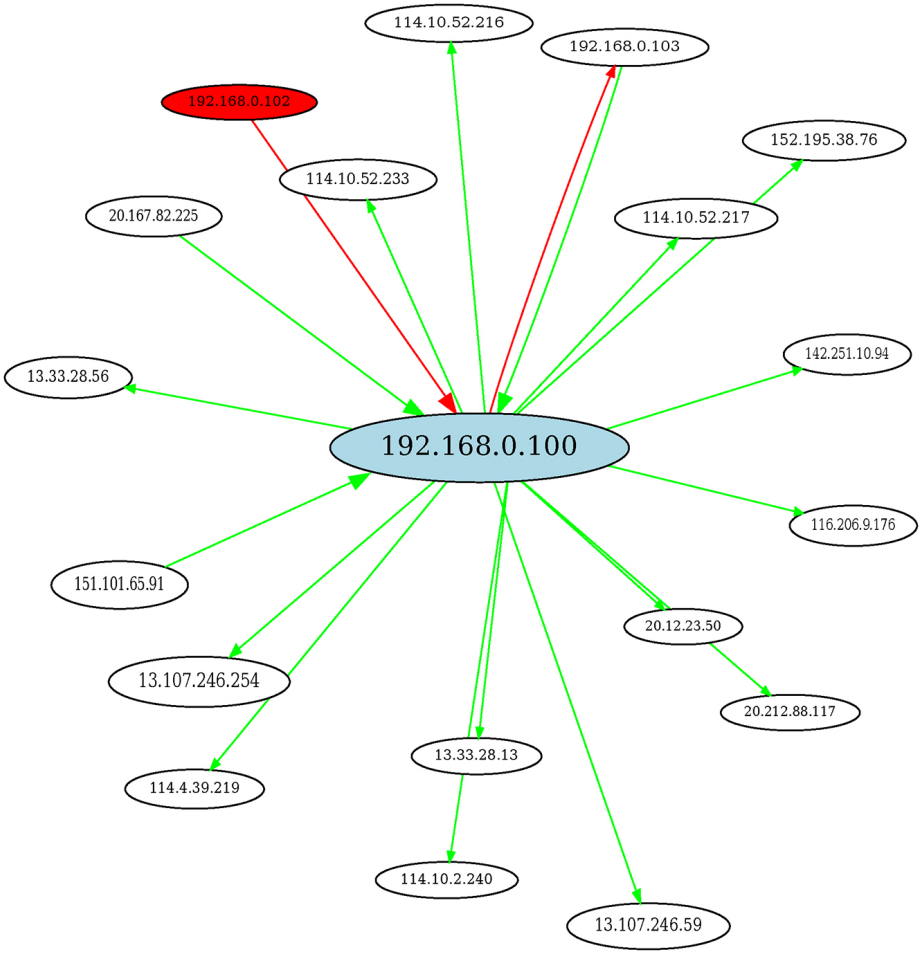

Fig. 8. Simplified attack flow ontology for anomaly detection

4.2 Discussion

This discussion closely examines the results from the previous section in order to ascertain their theoretical and practical significance. This discussion establishes current limitations, highlights implications of operations for practical application, puts our results in perspective in relation to the results of current approaches, and identifies areas of future research for improving network intrusion detection systems.

Comparative analysis with prior works. The research results have notable improvements over current approaches. Due to optimized feature engineering for traffic behavior within particular substations, the model's 100% recall for MITM attacks is superior to FI-SEL's 87% accuracy [8] and SA-DCNN's 95% recall [14] (refer to Table 3). The proposed model offers similar precision (>99.3%) for each category

of traffic, while the cybernet model of [13] only considers DDoS attacks. A working environment that does not possess a knowledge graph in [17] with ransomware specialization is introduced by the ontology visualization in Figure 7 that fills a key gap pointed out by [18]. More specifically, for imbalanced IIoT data, the DNN model reduces false positives by 40% relative to lightweight ML approaches in [10].

Operational implications. The system is feasible for real-time substation protection, more so against MITM-induced transformer overloads, because of its 99.78% precision (refer to Table 3) and 1.2 ms inference latency [7]. To address the requirement outlined in [3] for actionable cybersecurity intelligence, operators can utilize the ontology visualization to correlate 97.4% of identified attacks in Figure 7 with physical devices such as circuit breakers. Deployment, however, would require at least edge devices with 4GB of RAM, making phased implementation in renovated substations more practical. Given the model's 100% attack detection rate, the 0.66% misclassification rate of cloud-service IPs in Figure 7 remains acceptable; nonetheless, optimizing firewall rules could further reduce false positives.

Limitations. Three significant limitations were faced: First, ARP spoofing and IEC 61850-specific attacks are missing from the dataset; ICMP redirect attacks of Figure 4 [22] are available only. Second, older substation equipment can be tested by memory requirements of the model [19]. Third, interpretability is supported by the ontology, but for utilizing the visualization system at Figure 7 to its maximum potential, training needs to be provided for operators. These are similar to those enumerated in the review of AI-based IIoT security solutions in [21].

Future work. Future work will extend attack coverage to Modbus/TCP manipulation and GOOSE message spoofing [22], in addition to the adversarial training architecture in [5]. Federated learning approaches such as TrustFed [11] could enable privacy-preserving model updates across substations. Hardware optimization will focus on compressing the DNN for deployment on resource-constrained devices, addressing edge-computing limitations highlighted in [19]; hardware optimization will be directed towards compressing DNN for deployment on low-resource devices. In determining the effectiveness of the system against adapting MITM attacks in actual operational substation environments, longitudinal studies will need to be performed.

5 CONCLUSION

This work presented an optimized deep neural network-based technique for man-in-the-middle attack detection in substation monitoring systems enabled by the Industrial Internet of Things. The researchers managed to capture both normal and attack traffic patterns using a testbed simulation consisting of a SCADA simulator, a real user device, and an attacker node. The resultant dataset was employed as a solid basis for model training after being enriched by attack simulations and systematic data preprocessing.

When identifying attack traffic, the projected DNN structure exhibited spotless precision, recall, and F1-score, as well as superior classification performance with an accuracy of up to 99.78%. Even when there was a class imbalance, the model exhibited stable accuracy for every class of traffic—attack, normal, and request. One-hot encoding, feature normalization, and oversampling were crucial in maximizing learning efficiency and generalization capability.

This study also utilized ontology-based visualization to translate the DNN's prediction to a graph form using color-coded edges and nodes to denote network

activity to supplement detection results. Besides allowing quicker threat and anomaly detection, the visualization provided greater contextual awareness for network operators. Together, the results demonstrate that real-time intrusion detection within critical IIoT infrastructures can be greatly improved by an optimally optimized deep learning model as well as quality preprocessing and visualization.

6 REFERENCES

[1] J. Sengupta, S. Ruj, and S. Das Bit, "A comprehensive survey on attacks, security issues and blockchain solutions for IoT and IIoT," *Journal of Network and Computer Applications*, vol. 149, p. 102481, 2020. https://doi.org/10.1016/j.jnca.2019.102481

[2] D. Xu, W. Niu, Q. Li, H. Li, and L. Cheng, "Enhancing power marketing audit through IoT and multi-sensor information fusion: A substation scenario analysis," *Computers and Electrical Engineering*, vol. 118, p. 109312, 2024. https://doi.org/10.1016/j.compeleceng.2024.109312

[3] H. Sarjan, A. Ameli, and M. Ghafouri, "Cyber-security of industrial internet of things in electric power systems," *IEEE Access*, vol. 10, pp. 92390–92409, 2022. https://doi.org/10.1109/ACCESS.2022.3202914

[4] R. Basri *et al.*, "Enhancing IoT security: Assessing instantaneous communication trust to detect man-in-the-middle attacks," *Future Generation Computer Systems*, vol. 166, p. 107714, 2025. https://doi.org/10.1016/j.future.2025.107714

[5] M. Al-Fawa'reh, J. Abu-khalaf, N. Janjua, and P. Szewczyk, "On and off the manifold: Generation and detection of adversarial attacks in IIoT networks," *Journal of Network and Computer Applications*, vol. 235, p. 104102, 2025. https://doi.org/10.1016/j.jnca.2024.104102

[6] H. Fereidouni, O. Fadeitcheva, and M. Zalai, "IoT and man-in-the-middle attacks," *Security and Privacy*, vol. 8, no. 2, p. e70016, 2023. https://doi.org/10.1002/spy2.70016

[7] Y. Qiao, D. Chen, Q. Z. Sun, G. Tian, and W. Wang, "Unveiling stealthy man-in-the-middle cyber-attacks on energy performance in grid-interactive smart buildings," *Energy Convers Manag.*, vol. 319, 2024. https://doi.org/10.1016/j.enconman.2024.118949

[8] S. A. Abdulkareem, C. H. Foh, F. Carrez, and K. Moessner, "A lightweight SEL for attack detection in IoT/IIoT networks," *Journal of Network and Computer Applications*, vol. 230, p. 103980, 2024. https://doi.org/10.1016/j.jnca.2024.103980

[9] S. Ismail, S. Dandan, and A. Qushou, "Intrusion detection in IoT and IIoT: Comparing lightweight machine learning techniques using TON_IoT, WUSTL-IIOT-2021, and EdgeIIoTset Datasets," *IEEE Access*, vol. 13, pp. 73468–73485, 2025. https://doi.org/10.1109/ACCESS.2025.3554083

[10] S. Ismail, S. Dandan, D. W. Dawoud, and H. Reza, "A comparative study of lightweight machine learning techniques for cyber-attacks detection in blockchain-enabled industrial supply chain," *IEEE Access*, vol. 12, pp. 102481–102491, 2024. https://doi.org/10.1109/ACCESS.2024.3432454

[11] M. H. ur Rehman, A. M. Dirir, K. Salah, E. Damiani, and D. Svetinovic, "TrustFed: A framework for fair and trustworthy cross-device federated learning in IIoT," *IEEE Trans. Industr. Inform.*, vol. 17, no. 12, pp. 8485–8494, 2021. https://doi.org/10.1109/TII.2021.3075706

[12] Z. E. Huma *et al.*, "A hybrid deep random neural network for cyberattack detection in the industrial internet of things," *IEEE Access*, vol. 9, pp. 55595–55605, 2021. https://doi.org/10.1109/ACCESS.2021.3071766

[13] H. Dong, I. Kotenko, and D. Levshun, "Next-generation IIoT security: Comprehensive comparative analysis of CNN-based approaches," *Knowl. Based Syst.*, vol. 316, p. 113337, 2025. https://doi.org/10.1016/j.knosys.2025.113337

[14] M. S. Alshehri, O. Saidani, F. S. Alrayes, S. F. Abbasi, and J. Ahmad, "A self-attention-based deep convolutional neural networks for IIoT networks intrusion detection," *IEEE Access*, vol. 12, pp. 45762–45772, 2024. https://doi.org/10.1109/ACCESS.2024.3380816

[15] S. Ullah, W. Boulila, A. Koubaa, and J. Ahmad, "Attention-based hybrid deep learning model for intrusion detection in IIoT networks," *Procedia Comput. Sci.*, vol. 246, pp. 3323–3332, 2024. https://doi.org/10.1016/j.procs.2024.09.307

[16] A. Venčkauskas, J. Toldinas, N. Morkevičius, and F. Sanfilippo, "An email cyber threat intelligence method using domain ontology and machine learning," *Electronics (Switzerland)*, vol. 13, no. 14, p. 2716, 2024. https://doi.org/10.3390/electronics13142716

[17] M. Keshavarzi and H. R. Ghaffary, "An ontology-driven framework for knowledge representation of digital extortion attacks," *Comput. Human Behav.*, vol. 139, p. 107520, 2023. https://doi.org/10.1016/j.chb.2022.107520

[18] F. Alqurashi and I. Ahmad, "A data-driven multi-perspective approach to cybersecurity knowledge discovery through topic modelling," *Alexandria Engineering Journal*, vol. 107, pp. 374–389, 2024. https://doi.org/10.1016/j.aej.2024.07.044

[19] P. Spadaccino and F. Cuomo, "Intrusion detection systems for IoT: Opportunities and challenges offered by edge computing intrusion detection systems for IoT: Opportunities and challenges offered by edge computing and machine learning," 2020.

[20] G. Abbas, M. Ali, M. Ahmad, and A. Khan, "CIRA-cyber intelligent risk assessment methodology for industrial internet of things based on machine learning," *IEEE Access*, vol. 13, pp. 77001–77016, 2025. https://doi.org/10.1109/ACCESS.2025.3559617

[21] A. Hassan, N. Nizam-Uddin, A. Quddus, S. R. Hassan, A. U. Rehman, and S. Bharany, "Navigating IoT security: Insights into architecture, key security features, attacks, current challenges and AI-driven solutions shaping the future of connectivity," *Computers, Materials & Continua*, vol. 81, no. 3, pp. 3499–3559, 2024. https://doi.org/10.32604/cmc.2024.057877

[22] J. Roldán-Gómez, J. Boubeta-Puig, J. Carrillo-Mondéjar, J. M. Castelo Gómez, and J. M. del Rincón, "An automatic complex event processing rules generation system for the recognition of real-time IoT attack patterns," *Eng. Appl. Artif. Intell.*, vol. 123, p. 106344, 2023. https://doi.org/10.1016/j.engappai.2023.106344

[23] F. S. Alrayes, M. Zakariah, S. U. Amin, Z. I. Khan, and J. S. Alqurni, "Network security enhanced with deep neural network-based intrusion detection system," *Computers, Materials and Continua*, vol. 80, no. 1, pp. 1457–1490, 2024. https://doi.org/10.32604/cmc.2024.051996

[24] W. Alawsi, "Intrusion detection in IoT networks using machine learning techniques," *International Journal of Computers and Informatics*, vol. 2, no. 8, pp. 9–33, 2023. https://doi.org/10.59992/IJCI.2023.v2n8p1

7 AUTHORS

Dwi Prasetya is currently studying Master's of Computer Science at Sriwijaya University. He currently works at PT PLN (Persero) as an engineer. His research interests: Internet of Things, SCADA, IT, and Operational Technology security (E-mail: dwiprasetyacahya@gmail.com, dwi.prasetya@pln.co.id).

Dendi Renaldo Permana received a B.Com. in the Information System program at University Riau. He began his career as a software and machine learning engineer. Then he received a scholarship called the magister menuju doktor untuk sarjana unggul (PMDSU) in 2023 to continue his master's and doctoral studies in computer science at Universitas Sriwijaya with a focus on research in the field of cyber threat intelligence (E-mail: dendi.renaldo@gmail.com).

M. Rafie Al Hamas Graduated with a Bachelor's degree in Computer Systems from Universitas Sriwijaya. He has a strong interest in computer networks, cybersecurity, and system administration, which is supported by various professional certifications he has obtained (E-mail: rafie.alhamas21@gmail.com).

Deris Stiawan received his Ph.D. degree in computer engineering from Universiti Teknologi Malaysia, Malaysia. He is currently a professor with the Faculty of Computer Science, Universitas Sriwijaya. His research interests include computer networks, intrusion detection/prevention systems, and heterogeneous networks (E-mail: deris@unsri.ac.id).

Tami A. Alghamdi obtained his bachelor's and master's in computer science at Western Illinois University. Tami received a Ph.D. in computer science at the University of Idaho in 2022. Currently, he is an assistant professor at the College of Computing and Information, Al-Baha University, Kingdom of Saudi Arabia. His research interests are machine learning, transfer learning, genetic algorithms, and data science (E-mail: talwajeeh@bu.edu.sa).

Rahmat Budiarto received Dr. Eng. in Computer Science from Nagoya Institute of Technology, Japan, in 1998. Currently, he is a full professor at the College of Computing and Information, Albaha University, Saudi Arabia. His research interests include intelligent systems, brain modeling, IPv6, network security, wireless sensor networks, and MANETs. He was chairing the APAN Security Working Group (2006–2009) and established the IPv6 research center (NAv6 Center) at Universiti Sains Malaysia (USM) in 2005, then was appointed as the Deputy Director of the center (2005–2009) (E-mail: rahmat@bu.edu.sa).