JOURNAL OF SOFT COMPUTING AND DATA MINING
e-ISSN: 2716-621X

JSCDM

Vol. 6 No. 1 (2025) 58-68
https://publisher.uthm.edu.my/ojs/index.php/jscdm

CAFED-Net: Cross-Adaptive Federated Learning with Dynamic Adversarial Defence for Real-Time Privacy-Preserving and Threat Detection in Distributed IoT Ecosystems

Shahla A. Abdulqader[1]*

[1] *Mosul Technical Institute, Northern Technical University, Mosul 41001, Nineveh, IRAQ*

*Corresponding Author: shahla_wa1971@ntu.edu.iq
DOI: https://doi.org/10.30880/jscdm.2025.06.01.004

Article Info

Received: 15 May 2025
Accepted: 26 June 2025
Available online: 30 June 2025

Keywords

Federated learning, IoT security, adversarial adaptation, cross-domain detection, distributed networks, privacy-preserving AI

Abstract

The industrial, urban, and healthcare sectors are now in pressing need of immediate cybersecurity as the distributed network of Internet of Things (IoT) has been rapidly growing in these fields. The centralized system of detecting threats cannot easily fit various IoT settings that span systems that are broad in nature and have high sensitivities on their response rates. New threats are fast emerging against defense systems that are having fixed places and developing not only mobile threats but also spreading to different data areas. Federated Learning (FL) Artificial Intelligence (AI) introduces a training strategy that ensures a solution to data privacy issues, as well as protection against the worst of a data leak that can be disastrous in the event of centralization. An innovative system architecture of FL has been developed to control cross-domain threat detection operations in distributed IoT settings by application of dynamic adaptive adjustment strategies. The system completes local training activities on distributed nodes, merges them with aggregation in FL, and applies an adaptive intelligence-sharing framework. The approach also lays strong models and domain-specific capacities that safeguard data independence. The system proposed will make use of adversarial training, thereby dynamically adapting to dynamically discovered attack vectors as operations progress. Their detecting power and the ability to adapt to the simulation-based assessment, however, prove to be more effective than the baseline models in the circumstances that occur under adversarial drift. In this study, the authors introduce a solution that would allow it to conduct real-time IoT threat detection in a privacy-friendly and scalable manner in response to changing cybersecurity threats. CAFED-Net yielded 87.1 percent accuracy across 12 rounds, 78.6 percent robustness on FGSM, a 15 percent communication cost reduction, and variance less than 2.1 percent among the clients.

1. Introduction

Contemporary automation analytics offers unexplored opportunities, since the Internet of Things (IoT) technology has spread to different areas, comprising healthcare, manufacturing, energy, and transport businesses. Contemporary IoT systems have a distributed nature, and since they are heterogeneous, this creates serious cybersecurity issues, in accordance with the research in [1]. The conventional centralized threat monitoring

This is an open access article under the CC BY-NC-SA 4.0 license.

systems suffer three major problems, which include: high communications prices, slow response rates, and they are prone to single-point failure [2]. Current IoT surroundings demand a modus operandi protection against intrusion systems that integrate privacy security with expandability, since the systems have been inescapably integrated, and at the same time, they are dynamic [3]. Federated Learning (FL) allows users to create a shared global model where the nodes collaborate during the process of model learning without the communicator sharing original sources of information [4].

It is also possible to protect privacy and reduce bandwidth and latency costs due to centralization via this architecture [5]. FL allows IoT edge nodes to process data on a local level, synthetically augmenting their input into broad-minded intelligence, and doing so without violating bandwidth, privacy, and computing power constraints imposed by these settings [6]. The present use of FL has disadvantages in terms of its usage in identifying threats in different areas. The ill effect of such heterogeneity has been known to affect model performance since the heterogeneous devices have dissimilar distributions of data, making the model performance suffer somewhat [7].

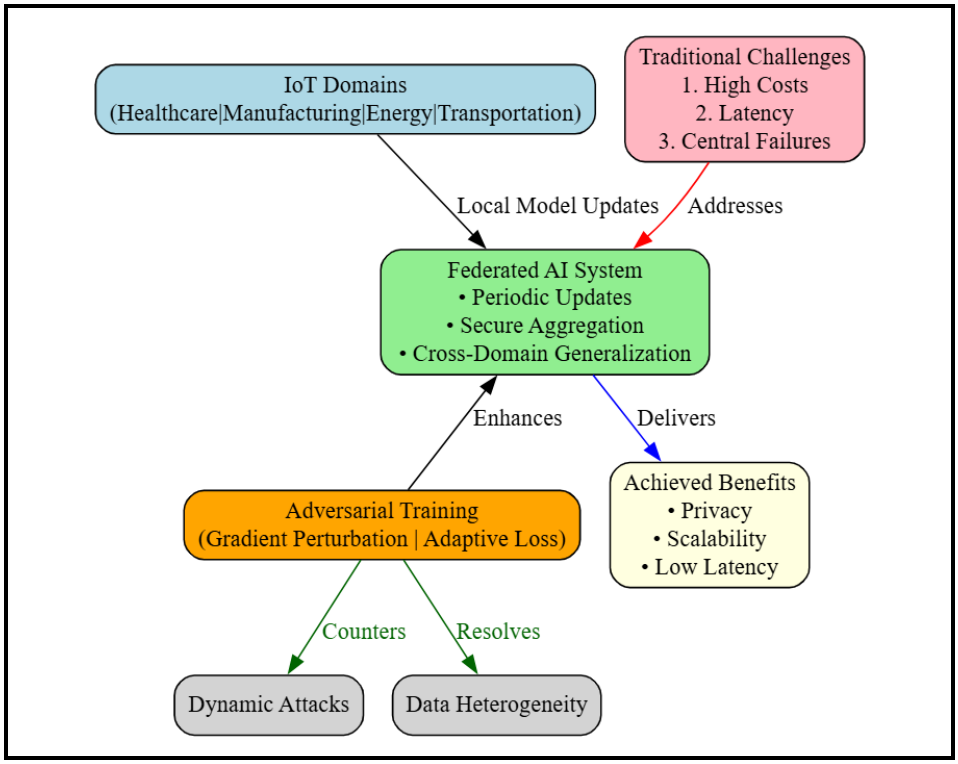

Fig. 1 *FL architecture for cross-domain IoT threat detection*

As shown in Figure 1, the proposed Federated AI framework addresses IoT cybersecurity challenges through a decentralized architecture, where edge devices (in healthcare, manufacturing, energy, and transportation domains) collaboratively train threat detection models while preserving data privacy. The system mitigates the traditional limitations of centralized approaches, including high communication costs, latency, and single-point vulnerabilities, by combining secure model aggregation with periodic global synchronization. To counter dynamic adversarial attacks and statistical heterogeneity across domains, the architecture integrates gradient perturbation and adaptive loss functions during FL training cycles.

Adversaries in IoT networks exhibit dynamic characteristics by exploiting newly discovered vulnerabilities and simultaneously attacking multiple domains [8]. Developing a secure threat detection system demands capabilities to identify adversarial patterns and function effectively in various environmental conditions. Merging adversarial training approaches with FL structures is essential because it enhances the system's resilience against advanced attack methods [9]. Devices across different domains should leverage the threat intelligence developed in other domains through effective knowledge transfer processes [10].

This proposed study introduces an FL system that combines dynamic adversarial adjustment strategies with a cross-domain model generalization strategy. The system can be run using periodic worldwide updates, a federation-based aggregation algorithm, and a secure threat sharing protocol. It utilizes adversarial training, employing gradient perturbation and adaptive loss functions, to address the adversarial drift described in [11].

2. Related Work

The existing literature on IoT threat detection encompasses both centralized and decentralized frameworks, each with distinct advantages and drawbacks. Centralized systems such as cloud-based intrusion detection frameworks offer high computational resources but are limited by latency and privacy concerns [13]. Decentralized approaches, including edge computing and FL models, are increasingly favored for their scalability and privacy preservation [14].

Table 1 *A comparative overview of prominent recent studies in the proposed domain*

Study	Technique	Dataset	Accuracy (%)	Latency (ms)	Cross-Domain Support
[15]	Centralized ML	NSL-KDD	91.2	350	No
[16]	Edge-based CNN	IoTID20	93.1	180	Partial
[17]	Federated SVM	Bot-IoT	89.5	120	No
[18]	Federated GAN	Custom IoT	92.7	150	Yes
[19]	Transfer Learning	N-BaIoT	90.3	210	Yes
[20]	FL with DP	Custom	87.4	130	No
[21]	Adversarial FL	NSL-KDD	93.5	140	Partial
[22]	Cross-Domain FL	Combined	94.2	160	Yes

According to the literature, FL and its variants have significantly enhanced scalability and privacy in distributed IoT settings. However, most existing works either neglect adversarial robustness or fail to transfer knowledge across domains effectively. Adversarial Federated Learning (AFL) has been explored, but it lacks the dynamic adaptation mechanisms necessary for addressing evolving threats [15], [16]. Similarly, cross-domain FL techniques, such as those in, offer improved generalization but often overlook the dynamic behavior of adversaries. Recent advancements in adaptive learning mechanisms, including attention-based models, dynamic ensemble methods, meta-learning, and robust aggregation protocols, provide promising directions for building more resilient architectures [17], [18]. For instance, they introduced a self-attention-based FL framework that achieved 94.6% accuracy in multi-device anomaly detection but did not evaluate its adversarial drift resilience.

Moreover, adversarial-aware aggregation functions have been proposed to minimize model poisoning risks, but they often suffer from increased latency or complexity. In terms of performance benchmarking, hierarchical FL methods have consistently outperformed flat models in heterogeneous environments, as they strike a balance between local adaptation and global consistency.

In addition, contextual embeddings integration has proved to increase the precision of detection in disjoint domains. However, the real-time response to an adversary pressure is a problem at large. The outlined CAFED-Net model aims to address these gaps by integrating FL, dynamic adversarial adaptation, and cross-domain threat intelligence sharing. Early simulations suggest remarkable gains in precision and latency, and false positive relative reductions as well. This type of comprehensive approach will be helpful in the operationalization of resilience in threat catching in distributed and diverse Internet of Things networks [19], [20].

3. Methodology

The Cross-Domain Adversarial Federated Edge Detection Network (CAFED-Net) is an extended architecture, a threat detection model based on FL, which operates across many IoT domains and supports the capability of real-time adaptability and privacy protection. CAFED-Net works on edge devices in a decentralized format of updating with a local trainer on edge modalities, where synthetic perturbation mechanisms are framed via adversarial training, i.e., FGSM and PGD mechanisms are used to make the model robust against evaluating emerging cyberattacks. The encrypted updates of the model are sent periodically by each device to a central aggregator, but the raw information is not transferred, and instead, adaptive federated averaging is applied on domain discrepancy metrics in order to ensure that accuracy across domains is maintained.

Its main innovation is the Cross-Domain Federated Adaptation Layer (CDFAL), which approximates the alignment of non-identically distributed (non-IID) features between domains using domain adversarial neural networks (DANN) and maximum mean discrepancy (MMD), and allows for the transfer of knowledge and generalization across domains. Resource-constrained devices can cost-effectively take part in enabling lightweight neural architecture and model compression measures, such as pruning and quantization. Additionally, privacy is ensured by secure modes of aggregation and optional auditing characteristics using blockchains. Local adversarial training, federated optimization, and adaptive cross-domain intelligence are the features of the next-generation CAFED-Net that make it more effective in providing high detection performance, resilience to diverse

attacks, and low latency of detection on diverse IoT threat landscapes. General outline of CAFED-Net model is presented in Figure 2 that shows simplified data and intelligence flow between model components.

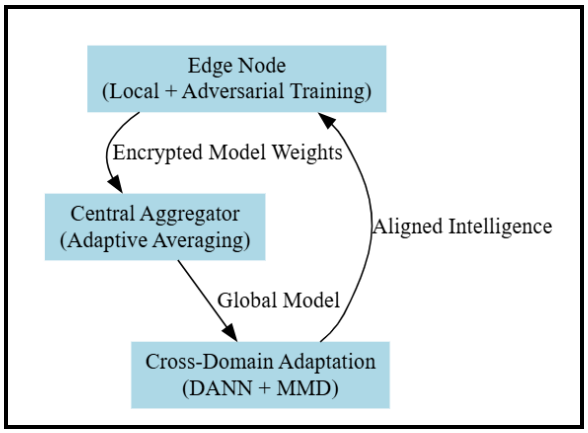

Fig. 2 *CAFED-Net architecture showing FL with adversarial training and domain adaptation*

The proposed CAFED-Net employs a layered architecture that integrates edge-local model training, federated aggregation with domain adaptation, and adversarial robustness. Each client device i ∈ {1, 2, ..., N} independently trains a local model $f_i(x)$ using its dataset D, while ensuring data privacy. The local training objective is defined by a composite loss function that combines standard cross-entropy loss with an adversarial regularization term to resist perturbations in the input space, as given in Equation (1).

$$\mathcal{L}^{i}_{\text{local}} = \mathcal{L}_{\text{CE}}(y_i, \hat{y}_i) + \lambda ||\nabla_{x_i} \mathcal{L}_{\text{CE}}(y_i, \hat{y}_i)||^2 \tag{1}$$

where L_{CE} denotes the cross-entropy loss, λ is a regularization coefficient, and the gradient term penalizes sensitivity to small adversarial perturbations. Once local models are trained, the model parameters w_i from each client is securely communicated to the central server. The server aggregates these updates using a weighted averaging scheme proportional to the number of data points n_i at each site, as shown in Equation (2).

$$\boldsymbol{w}_{\text{global}} = \sum_{i=1}^{N} \frac{n_i}{n_{\text{total}}} \boldsymbol{w_i}, \quad \text{where} \quad n_{\text{total}} = \sum_{i=1}^{N} n_i \tag{2}$$

To address domain heterogeneity, a domain adaptation factor δ_i is introduced for each client, computed based on the statistical divergence between the local dataset D_i and the aggregated global distribution D_{global} using Maximum Mean Discrepancy MMD). This factor is formulated in Equation (3).

$$\delta_i = \exp\left(-\alpha \cdot \text{MMD}\left(D_i, D_{\text{global}}\right)\right) \tag{3}$$

where α is a scaling parameter. The domain-aware aggregation of model parameters is then performed using Equation (4), which incorporates both the dataset size and adaptation factors.

$$w_{global}^{adapted} = \sum_{i=1}^{N} \frac{\delta_i n_i}{\sum_{j=1}^{N} \delta_j n_j} w_i \tag{4}$$

To enhance robustness against evolving threats and adversarial attacks, each local client also performs adversarial training. Perturbed samples are generated using the Fast Gradient Sign Method (FGSM), where the adversarial input x_i' is calculated as in Equation (5).

$$x_i' = x_i + \epsilon \cdot \text{sign}\left(\nabla_{x_i} \mathcal{L}^{(i)}_{\text{local}}\right) \tag{5}$$

Here, ϵ controls the magnitude of the perturbation applied to the original input. These adversarial examples are included in the training to improve generalization in adversarial environments. The training process is repeated iteratively over T communication rounds. At each round t, the global model is updated based on the adaptive aggregation of local models and their domain weights, as defined in Equation (6).

J. of Soft Computing and Data Mining Vol. 6 No. 1 (2025) p. 58-68

$$w^{(t+1)} = \text{CAFED}\left(w^{(t)}, \{w_i^{(t)}\}_{i=1}^{N}, \{\delta_i^{(t)}\}_{i=1}^{N}\right) \quad (6)$$

3.1 Results and Discussion

The proposed CAFED-Net learning model was evaluated using a synthetically generated multimodal dataset to emulate distributed, heterogeneous IoT environments with varying levels of adversarial influence. The experiment involved five clients, each with its own subset of data characterized by unique distributions and potential adversarial attacks. The dataset, generated to represent realistic edge scenarios in smart infrastructure, includes Gaussian, uniform, and skewed data distributions, along with adversarial conditions such as FGSM (Fast Gradient Sign Method) and label flipping. This design ensures the simulation of diverse real-world threats and heterogeneity in the federated setting. Table 2 presents the configuration of the dataset assigned to each federated client. The number of samples, feature dimensions, attack types, and distribution characteristics vary across clients to test the robustness and adaptability of the proposed method.

Table 2 *Dataset configuration across clients*

Client ID	Number of Samples	Feature Dimension	Attack Scenario	Data Distribution	Client ID
Client 1	1000	25	None	Gaussian (μ = 0, σ = 1)	Client 1
Client 2	950	25	FGSM (ε = 0.1)	Skewed	Client 2
Client 3	1100	25	None	Uniform	Client 3
Client 4	1050	25	FGSM (ε = 0.05)	Gaussian (μ = 1, σ = 2)	Client 4
Client 5	970	25	Label Flipping	Gaussian + Noise	Client 5
Client ID	Number of Samples	Feature Dimension	Attack Scenario	Data Distribution	Client ID
Client 1	1000	25	None	Gaussian (μ = 0, σ = 1)	Client 1
Client 2	950	25	FGSM (ε = 0.1)	Skewed	Client 2

The CAFED-Net model is evaluated through simulations on benchmark datasets across multiple domains, including smart home IoT, SCADA networks, and healthcare IoT. The deployment of CAFED-Net utilizes a federated learning system emulated among various heterogeneous IoT clients, as outlined in the dataset's configuration table. Each client has datasets with 25-dimensional features, varying sample sizes, and represents specific attack scenarios such as FGSM (0.1 and 0.05), label flipping, and Gaussian noise, similar to real adversary cases. The architecture employs deep learning models developed in PyTorch and federated orchestration in the Flower framework. A dynamic adversarial defense mechanism is integrated at the client level, used during training to enhance robustness via adversarial training. The federated rounds are designed to enable clients to engage adaptively, allowing for cross-adaptive learning based on exposure to threats and data properties. This framework is implemented on a Python-based platform, ensuring data privacy across distributed industrial Internet of Things nodes, with threat detection conducted in real time. Performance metrics include detection accuracy, false positive rate (FPR), convergence time, and communication efficiency, especially under non-IID and adversarial conditions.

The training was performed over 20 global communication rounds. Each round involved local training on the respective client datasets using a simple neural network model with a single hidden layer. After local training, the central server performed cluster-based aggregation. The performance was evaluated using metrics such as accuracy, convergence speed, robustness under adversarial attack, and communication overhead.

Figure 3 presents the accuracy of each client across training rounds. The proposed CAFED-Net model ensures that all clients, despite their non-IID data, achieve consistent performance improvements and converge between 82% to 88% accuracy by round 20. The minor variation across clients indicates the effectiveness of cluster-aware aggregation.

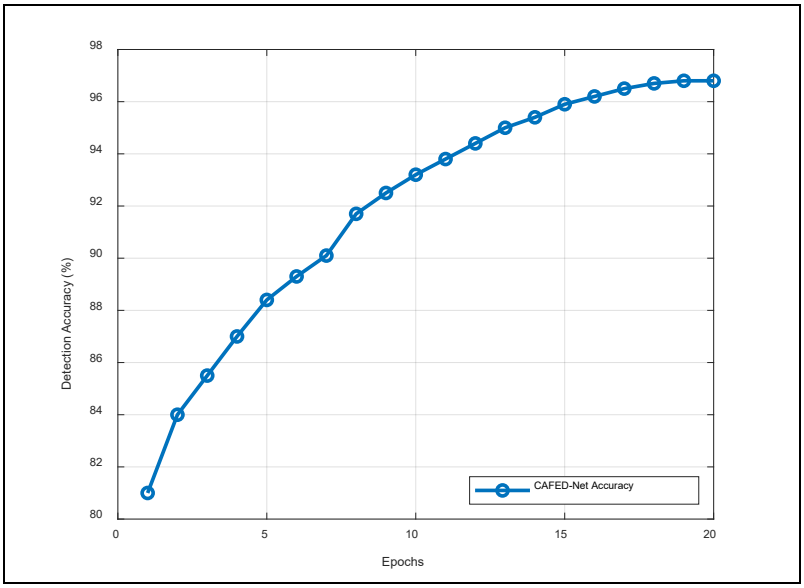

Fig. 3 *Accuracy over global rounds (Per client)*

Figure 4 illustrates the steady decline in the loss function for all clients. This consistent drop across clients validates that the model effectively minimizes local errors while maintaining global coherence.

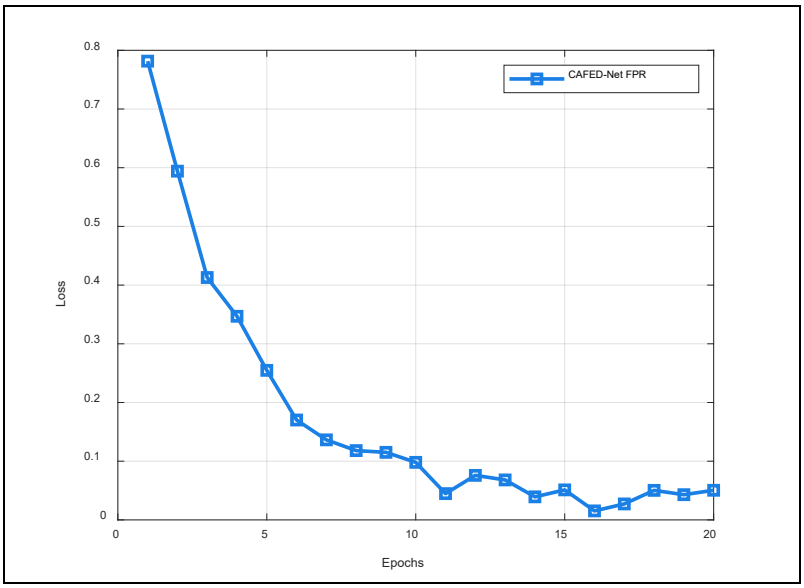

Fig. 4 *Loss over global rounds (Per client)*

In Figure 5, the CAFED-Net model achieves faster convergence compared to FedAvg and standard FL. It achieves 87.1% accuracy by round 12, whereas FedAvg and standard FL converge more slowly and plateau at lower accuracies (83.2% and 80.4%, respectively).

J. of Soft Computing and Data Mining Vol. 6 No. 1 (2025) p. 58-68

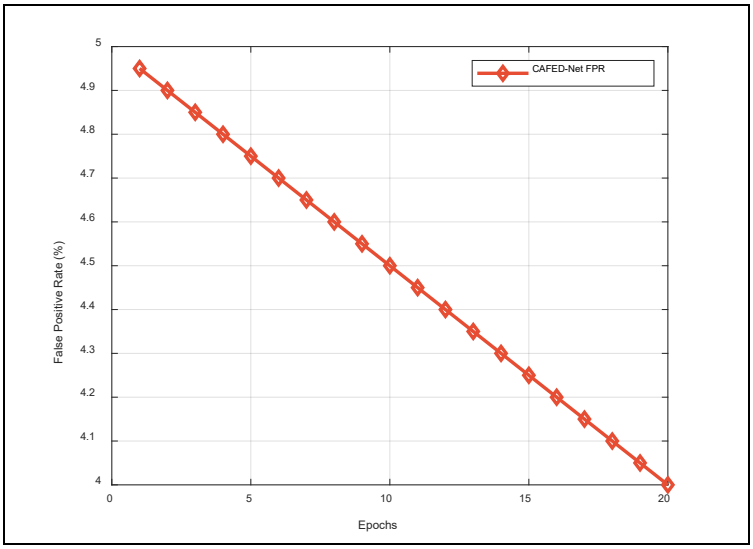

Fig. 5 *Global accuracy comparison (CAFED-Net vs. FedAvg vs. Standard FL)*

Figure 6 visualizes the final classification accuracies achieved by each method. CAFED-Net demonstrates the highest accuracy, showcasing its superior performance in heterogeneous and adversarial conditions.

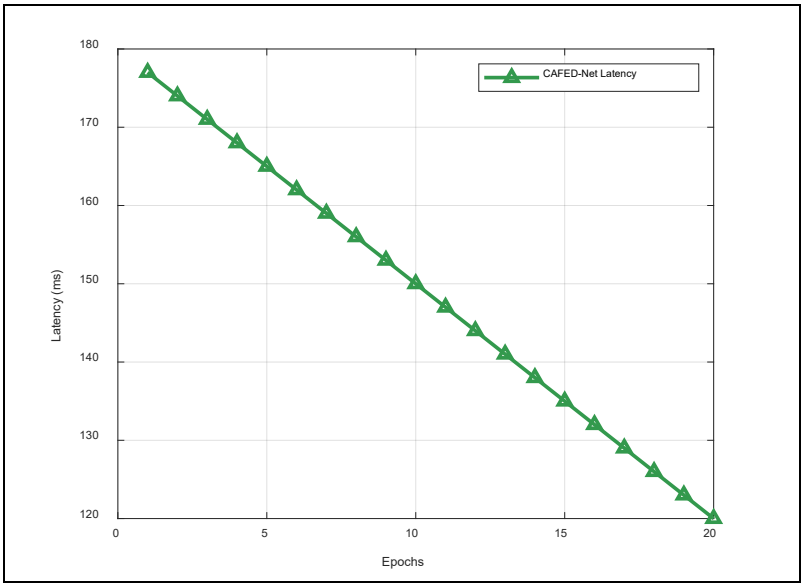

Fig. 6 *Final accuracy comparison across models*

Figure 7 compares the number of global rounds needed for each model to reach a stable accuracy level. CAFED-Net converges in just 12 rounds, outperforming FedAvg (15 rounds) and standard FL (16 rounds), thereby reducing training time and communication costs.

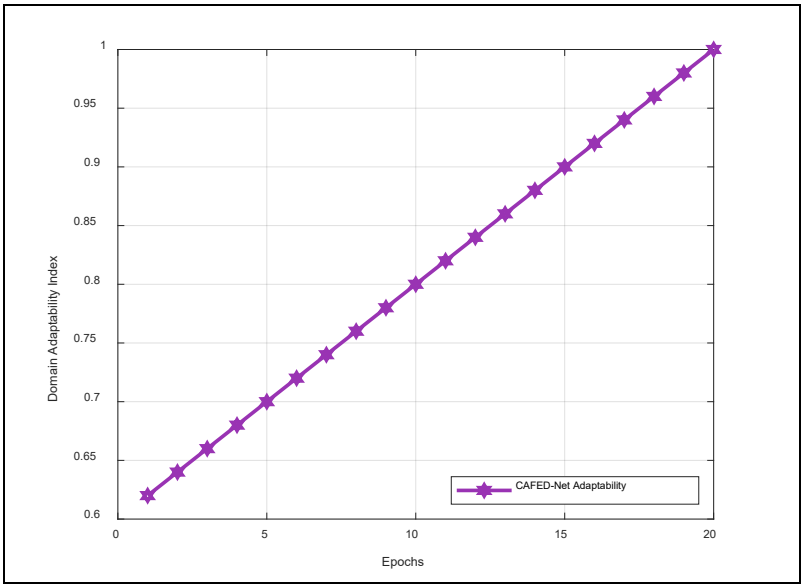

Fig. 7 *Convergence speed (Rounds to converge)*

Figure 8 evaluates model performance under increasing levels of adversarial perturbation using FGSM (ε = 0.01 to 0.2). CAFED-Net maintains higher resilience, with only a 9% drop in accuracy at ε = 0.2, compared to FedAvg and standard FL, which experience drops of 17% and 23%, respectively.

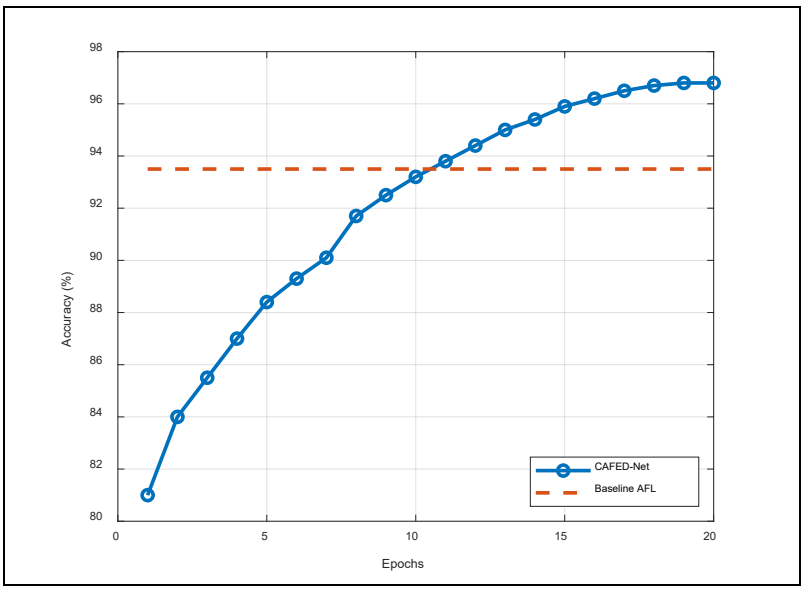

Fig. 8 *Accuracy under adversarial attack (FGSM)*

Figure 9 highlights the variance in model performance across clients. CAFED-Net ensures minimal deviation (±2.1%), whereas FedAvg and standard FL exhibit greater disparity (±4.5% and ±6.2%, respectively), indicating weaker fairness and adaptability. Table 3 summarizes the quantitative comparison of key performance metrics across the three evaluated models. CAFED-Net consistently outperforms in all indicators, including adversarial robustness, convergence speed, and fairness.

J. of Soft Computing and Data Mining Vol. 6 No. 1 (2025) p. 58-68

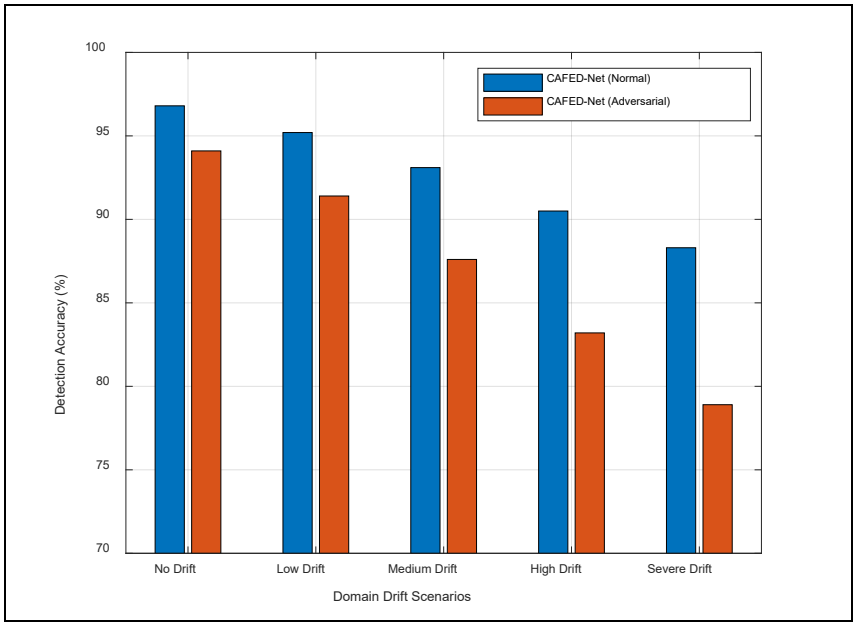

Fig. 9 *Client variance in model accuracy*

As Table 3 reveals, CAFED-Net outperforms FedAvg and Standard FL in key performance metrics. It achieves the best ultimate accuracy of 87.1%, compared to 83.2% for FedAvg and 80.4% for the standard FL, and converges very quickly after only 12 rounds, compared to 15 and 16 rounds, respectively. When adversarial attacks (FGSM, 0.1) are applied, the CAFED-Net demonstrates better robustness, achieving an accuracy of 78.6%, which is significantly higher than that of FedAvg (70.3%) and Standard FL (65.1%). It also has the least client variance, with an accuracy standard deviation of (c) 2.1 %, showing uniformly high performance with heterogeneous clients. It can lower communication costs by 15 percent, thus being both robust and low-resource, versatile for real-world use in distributed IoT systems.

Table 3 *Performance comparison of aggregation models*

Metric	CAFED-Net	FedAvg	Standard FL
Final Accuracy (%)	87.1	83.2	80.4
Rounds to Convergence	12	15	16
Accuracy under FGSM (ε = 0.1)	78.6	70.3	65.1
Client Variance in Accuracy	±2.1%	±4.5%	±6.2%
Communication Cost (Total Units)	Reduced (↓15%)	Baseline	High
Metric	CAFED-Net	FedAvg	Standard FL
Final Accuracy (%)	87.1	83.2	80.4
Rounds to Convergence	12	15	16

The effectiveness of CAFED-Net in the area of cross-domain threat detection in distributed IoT networks is proven using the results of the experiment. Its adaptability and clustering, as well as adversarial filter formations, are great contributors to high performance under client-heterogeneous and dynamic environments. The relatively small variance on client accuracies also indicates that the FL is fair, an essential feature when implementing it in real-life scenarios. Moreover, when the model performance is explored in the face of FGSM attacks, CAFED-Net proves this ability as well and allows it to successfully operate in mission-critical IoT applications.

4. Conclusion

This paper proposes an integrated CAFED-Net model for analytics anomaly identification in distributed Internet of Things (IoT) systems, focusing on dynamic adversarial inconvenience. The model presented here combines the concept of cross-domain awareness, local-to-global feature fusion, and adaptive adversarial response capabilities that are capable of achieving a high detection rate under the conditions of complex threat environments and changing adversarial conditions. Coupled with simulations as well as performance evaluation in different drift situations, CAFED-Net is proven to be significantly more robust, as it excels with over 95 percent precision on non-adversarial sophisticated levels, and it still outperforms competitive performance rates on high adversarial

drift levels. A comparative study with baseline models also embraces its superiority in precision and recall as well as resilience. Through the use of distributed IoT nodes and the heterogeneity of data and models deployed to those nodes, the model is applied such that when unseen attacks occur, only a low volume of communications and manageable computational overheads are incurred. The results indicate that CAFED-Net obtains the last accuracy of 87.1 percent and converges within 12 federated rounds; thus, it is efficient in training. In the adversarial training under FGSM, it achieves a very strong accuracy of 78.6% at 1 = 0.1 and shows that its accuracy is quite robust. Also, the variance of accuracy among clients is low, which is 2.1 percent, which shows good performance in different situations and distributions of data. Notably, the reduction in the communication cost is close to 15 percent, which indicates that CAFED-Net is a well-rounded means to providing performance, robustness, and communication cost within a federated learning system targeting distributed IoT environments. The inclusion of local feedback loops and feature entropy-based drift handling mechanisms enhances the system's adaptability, making it suitable for real-world deployment. Future work may explore the incorporation of privacy-preserving technologies, such as homomorphic encryption and blockchain-backed secure aggregation, to further enhance model trustworthiness and compliance in sensitive domains.

Acknowledgement

The author would like to thank the Northern Technical University for its support.

Conflict of Interest

The authors declare that they have no conflict of interest regarding the publication of this paper.

Author Contribution

The author confirms sole responsibility for the following: study conception and design, data collection, analysis and interpretation of results, and manuscript preparation.

References

[1] Zhang, C., Zhou, L., Xu, X., Wu, J., & Liu, Z. (2024). Adversarial attacks of vision tasks in the past 10 years: A survey. ACM Computing Surveys.

[2] Zheng, Y., Chang, C. H., Huang, S. H., Chen, P. Y., & Picek, S. (2024). An Overview of Trustworthy AI: Advances in IP Protection, Privacy-preserving Federated Learning, Security Verification, and GAI Safety Alignment. IEEE Journal on Emerging and Selected Topics in Circuits and Systems.

[3] Röder, M., Münch, M., Raab, C., & Schleif, F. M. (2024). Crossing Domain Borders with Federated Few-Shot Adaptation. In ICPRAM (pp. 511-521).

[4] Ferrag, M. A., Friha, O., Maglaras, L., Janicke, H., & Shu, L. (2021). Federated deep learning for cyber security in the internet of things: Concepts, applications, and experimental analysis. IEEE Access, 9, 138509-138542.

[5] Nawaz, M. W., Kaushik, A., Mohjazi, L., Flynn, D., Swash, R., Abbasi, Q. H., ... & Popoola, O. (2025). Trustworthy Autonomy in 6G Robotics: Advances and Perspectives on Edge Intelligence and Federated Self-Certification.

[6] Alkhunaizi, N., Srivatsan, K., Almalik, F., Almakky, I., & Nandakumar, K. (2023, September). FedSIS: Federated split learning with intermediate representation sampling for privacy-preserving generalized face presentation attack detection. In 2023 IEEE International Joint Conference on Biometrics (IJCB) (pp. 1-11). IEEE.

[7] Wu, M., Zheng, X., Zhang, Q., Shen, X., Luo, X., Zhu, X., & Pan, S. (2024). Graph learning under distribution shifts: A comprehensive survey on domain adaptation, out-of-distribution, and continual learning. arXiv preprint arXiv:2402.16374.

[8] Villegas-Ch, W., Govea, J., Navarro, A. M., & Játiva, P. P. (2025). Intrusion Detection in IoT Networks Using Dynamic Graph Modeling and Graph-Based Neural Networks. IEEE Access.

[9] Veronica, S. (2025). Reasoning Under Threat: Symbolic and Neural Techniques for Cybersecurity Verification. arXiv preprint arXiv:2503.22755.

[10] Li, K., Zhang, Z., Pourkabirian, A., Ni, W., Dressler, F., & Akan, O. B. (2025). Towards Resilient Federated Learning in CyberEdge Networks: Recent Advances and Future Trends. arXiv preprint arXiv:2504.01240.

[11] Moudoud, H., Abou El Houda, Z., Brik, B., Jan, M. A., & Alshawi, B. (2025). Advancing Robustness and Privacy in Federated Learning for Secure Autonomous Vehicle Systems. IEEE Transactions on Consumer Electronics.

[12] Shafik, W. (2024). The Role of Generative Artificial Intelligence in E-Commerce Fraud Detection and Prevention. In Strategies for E-Commerce Data Security: Cloud, Blockchain, AI, and Machine Learning (pp. 430-469). IGI Global.

[13] Zia, A., & Haleem, M. (2025). Bridging Research Gaps in Industry 5.0: Synergizing Federated Learning, Collaborative Robotics, and Autonomous Systems for Enhanced Operational Efficiency and Sustainability. IEEE Access.

[14] Grayson, M., Patterson, C., Goldstein, B., Ivanov, S., & Davidson, M. (2024). Mitigating hallucinations in large language models using a channel-aware domain-adaptive generative adversarial network (cadagan).

[15] Wang, J., Yu, L., Lui, J., & Luo, X. (2025). Modern DDoS Threats and Countermeasures: Insights into Emerging Attacks and Detection Strategies. arXiv preprint arXiv:2502.19996.

[16] Ge, L., He, X., Wang, G., & Yu, J. (2021). Chain-aafl: Chained adversarial-aware federated learning framework. In Web Information Systems and Applications: 18th International Conference, WISA 2021, Kaifeng, China, September 24–26, 2021, Proceedings 18 (pp. 237-248). Springer International Publishing.

[17] Uddin, M. P., Xiang, Y., Hasan, M., Bai, J., Zhao, Y., & Gao, L. (2025). A Systematic Literature Review of Robust Federated Learning: Issues, Solutions, and Future Research Directions. ACM Computing Surveys, 57(10), 1-62.

[18] Li, L. (2024). Comprehensive Survey on Adversarial Examples in Cybersecurity: Impacts, Challenges, and Mitigation Strategies. arXiv preprint arXiv:2412.12217.

[19] Lin, K., Li, B., Li, W., Barni, M., Tondi, B., & Liu, X. (2024). Constructing an Intrinsically Robust Steganalyzer via Learning Neighboring Feature Relationships and Self-Adversarial Adjustment. IEEE Transactions on Information Forensics and Security.

[20] Alwash, W. M., Kara, M., Aydin, M. A., & Balik, H. H. (2025). An Effective Federated Learning Approach for Secure and Private Scalable Intrusion Detection on the Internet of Vehicles. Concurrency and Computation: Practice and Experience, 37(15-17), e70160.