TransferEdge: Transfer Learning Approach to Detect Evolving DDoS Threats in Edge-IIoT

Mulualem Bitew Anley, Angelo Genovese, Vincenzo Piuri
Department of Computer Science, Università degli Studi di Milano, Milan, Italy
emails: (mulualem.anley,angelo.genovese,vincenzo.piuri)@unimi.it

***Abstract*—The rapid proliferation of edge IoT systems in critical infrastructures, from smart cities to industrial IoT (IIoT) environments, has introduced significant security challenges, particularly Distributed Denial-of-Service (DDoS) attacks. These attacks can degrade service quality and compromise the availability and integrity of services. Although deep learning (DL) models have shown promise in detecting DDoS attacks, their reliance on large, high-quality labeled datasets limits their adaptability in dynamic IoT environments. Transfer learning offers a potential solution; however, existing methods often struggle with domain adaptation and effective knowledge transfer across heterogeneous datasets, leading to suboptimal performance against evolving attack patterns. To address these challenges, we propose TransferEdge, a novel transfer learning-based approach to detect evolving DDoS attacks in industrial IoT edge systems. TransferEdge leverages pre-trained models and describes a novel approach to optimize fine-tuning strategies tailored for DDoS attack detection, so as to align feature spaces and bridge the distributional gap between source and target domains. Experimental evaluations on the UNSW-NB15 and BoT-IoT datasets demonstrate that TransferEdge improves detection accuracy and decreases training time compared to conventional DL methods and current transfer learning approaches.**

***Index Terms*—Transfer Learning, IoT Security, DDoS Detection, Edge IIoT, Fine-Tuning**

I. Introduction

The rapid proliferation of edge IoT systems in critical infrastructures, from smart cities to industrial IoT (IIoT) environments, has revolutionized data processing and decision making at the network edge. However, this interconnectivity introduces a significant security vulnerability, particularly in defending against Distributed Denial-of-Service (DDoS) attacks that can severely degrade service quality and compromise the availability and integrity of essential services [1].

Various approaches exist for DDoS detection, including signature-based and anomaly-based methods [2]. Anomaly-based detection is often preferred for its ability to identify new and evolving attack patterns [3], [4]. Among anomaly-based detection, deep learning (DL) models show promise due to their capability of automatic feature extraction. However, their effectiveness is hindered by the need for large, high-quality labeled datasets, limiting adaptability in dynamic and data-scarce IoT environments [3]. This challenge is further exacerbated in IoT ecosystems due to resource restrictions, privacy concerns, and the high cost of data labeling required to train the DL model [5], [6].

Transfer learning is a promising approach to IoT intrusion detection and DDoS attacks, particularly in resource-constrained edge IIoT devices, which are prevalent in critical infrastructure settings [7]. Several transfer-learning techniques have been proposed for DDoS attack detection, including: [6] introduced a CNN-based approach, [8] leveraged pre-trained transformers to handle distribution shifts, [5] proposed a dual autoencoder, and [9] fine-tuned pre-trained CNN and BiLSTM models for DDoS detection in 5G networks. However, research gaps remain in optimizing the fine-tuning of these pre-trained models for efficient deployment in resource-constrained edge IIoT environments. To address these challenges, we propose TransferEdge, an innovative transfer learning approach for Edge-IIoT systems, leveraging CNN and pre-trained models to introduce a computationally-efficient methodology to optimize the fine-tuning strategies for DDoS attack detection, bridging the distributional gap between source and target domains.

We introduce the following main contributions: *(i)* computationally-efficient transfer learning to optimize fine-tuning strategies for specific datasets, including domain adaptation, differential fine-tuning (using variable learning rates), and selective layer fine-tuning (updating attack-sensitive layers) to reduce computational overhead, and *(ii)* validation using real-world cybersecurity datasets, proving its effectiveness in resource-constrained edge IIoT environments.

The remainder of this paper is structured as follows. Section II reviews related work on transfer learning-based intrusion detection in industrial IoT. Section III describes our methodology, including model architectures, transfer learning framework, and proposed fine-tuning strategies. Section IV outlines the experimental setup, datasets, evaluation metrics, and results. Section V discusses the implications of our findings and suggests future research directions. Finally, Section VI presents the conclusions and future research directions.

II. Related Works

Several studies have explored anomaly-based DDoS attack detection using transfer and deep learning. For instance, the work of [6] addressed the challenge of data scarcity in an intrusion detection system (IDS) by proposing a CNN-based transfer learning approach. Moreover, it leverages knowledge from large labeled datasets to enhance detection accuracy in resource-constrained environments. Building on this concept in the context of critical infrastructure security, the work in [10] introduced a transfer learning-based intrusion detection for communication-based train control (CBTC) systems. Their

framework integrates CNNs of one dimension with LSTM networks to effectively capture both spatial and temporal attack dynamics.

Moreover, to overcome the challenge of limited labeled data, various pre-trained models have recently emerged as powerful tools in cybersecurity. In [8], pre-trained transformers were employed to handle distribution shifts and data scarcity in communication networks. Similarly, [5] introduced a dual autoencoder-based deep transfer learning framework, enabling effective knowledge transfer from labeled to unlabeled IoT datasets. Their approach demonstrates enhanced detection capabilities for novel attacks, highlighting the efficacy of transfer learning in detecting threats.

The relevance of transfer learning has also been demonstrated in emerging networks, including 5G, particularly for DDoS attacks. The works described in [7] and [9] present fine-tuned pre-trained CNN and BiLSTM models using limited real-world 5G datasets, showing considerable results in identifying sophisticated DDoS attacks. The studies underscore transfer learning's ability to generalize across heterogeneous network environments, making it particularly valuable for zero-day intrusion detection scenarios in complex networks.

The integration of ensemble methods with transfer learning has further enhanced detection robustness. In their work [11], the authors discussed an IDS framework combining multiple CNN architectures (VGG16, Inception, Xception), optimized through hyperparameter tuning and ensemble learning, outperforming single-model IDS approaches. Furthermore, cross-network transfer learning, as explored by [12], demonstrated effective intrusion detection by aligning the knowledge of the source networks with the unlabeled target networks, highlighting its suitability for heterogeneous industrial environments. Moreover, [13] proposed a robust DDoS attack detection framework that uses adaptive transfer learning to overcome the heterogeneity and scarcity of data for network traffic, in both tabular and image format datasets, into representations that are amenable to DL and employ hyperparameter optimization and fine-tuning strategies.

Existing studies provide a robust foundation for addressing critical challenges in DDoS detection, particularly in identifying evolving threats within data-constrained industrial IoT environments. Building on these insights, this paper proposes a novel transfer learning approach to detect evolving DDoS attacks in Edge-based IIoT systems.

III. Methodology

This subsection presents our transfer learning methodology for improving DDoS detection in edge-based IIoT. We outline the framework for domain adaptation, describe baseline architectures, introduce EdgeTransfer fine-tuning strategies, and detail our evaluation methods.

A. Overview of the Transfer Learning Framework

Transfer learning is a strategy that leverages the knowledge gained from one problem domain (the source domain) to improve learning in a related but different domain (the target domain). It involves training a model in a source domain D_S with a large dataset and transferring the learned knowledge to a target domain D_T, which has a smaller dataset. This process can be represented as follows: Given a source dataset:

$$D_S = \{(x_i^S, y_i^S)\}_{i=1}^{N_S} \tag{1}$$

with N_S samples, a model f_S is trained to learn representations by optimizing parameters θ_S:

$$\theta_S^* = \arg\min_{\theta_S} \frac{1}{N_S} \sum_{i=1}^{N_S} \ell(f_S(x_i^S; \theta_S), y_i^S) \tag{2}$$

where $\ell(\cdot)$ is the loss function. We then take the learned weights and use them to train the CNN for the target domain (edge IIoT). Because data in this domain is limited, we fine-tuned the model by minimizing the average error on the target samples. The learned parameters θ_S^* are transferred to the target model f_T in the target domain:

$$D_T = \{(x_j^T, y_j^T)\}_{j=1}^{N_T} \tag{3}$$

where $N_T \ll N_S$. The model is then fine-tuned or adapted:

$$\theta_T^* = \arg\min_{\theta_T} \frac{1}{N_T} \sum_{j=1}^{N_T} \ell(f_T(x_j^T; \theta_T), y_j^T) \tag{4}$$

This allows the model to retain useful features from D_S while adapting to the specific patterns in D_T. This adaptation helps the models capture the unique characteristics of the edge IIoT environment while retaining the valuable features learned from the DDoS data.

B. Baseline Model Architecture

In this paper, we applied a CNN model trained on the UNSW-NB15 dataset as the source model and transferred it to the BoT-IoT dataset using a set of fine-tuning strategies. Additionally, we implemented and fine-tuned several state-of-the-art DL architectures to evaluate their performance on the target dataset. The details of those model architectures are explained in the following.

1) CNN model: We employ CNNs as our foundational architecture to automatically extract hierarchical features from network traffic. We select CNNs for their simplicity and low computational cost, which we need to meet the resource constraints of Edge-IIoT devices while still providing a reliable baseline for DDoS attack detection in the literature [14].

2) VGG8/VGG16: We use VGG8 as a lightweight variant of the VGG family, designed with fewer layers to balance performance and resource efficiency. It is a straightforward, sequential architecture for rapid inference and interpretability, especially in environments with limited computational resources [15]. We also use VGG16 for its deep architecture that extracts detailed representations of features [16].

3) ResNet18: ResNet18's residual connections, introduced by Microsoft [17], enable deeper architectures while mitigating the vanishing gradient problem, making it efficient at extracting complex features from heterogeneous edge data.

4) EfficientNet: The compound scaling strategy of EfficientNet is particularly advantageous for Edge-IIoT environments [18]. It achieves high accuracy with fewer parameters and lower computational cost, which is essential when processing data at the edge.

5) Xception: We use Xception pre-trained on ImageNet [19] due to its feature extraction capabilities that generalize well to diverse tasks, even with limited labeled data typical in resource-constrained IoT/IIoT environments.

C. Fine-Tuning Strategies

The edgeTransfer approach employs four distinct fine-tuning strategies to optimize knowledge transfer from source to target IIoT systems, as illustrated in Figure 1 and detailed below.

1) TL0 – Baseline Transfer Learning (No Freezing): In this scenario, we employ all layers of the pre-trained model that are fine-tuned on the target domain without freezing any weights. This approach allows the model to fully adapt to the evolving attack patterns in the target domain, enabling it to learn domain-specific features.

2) TL1 – Last Layer Retraining: In the last layer retraining scenario, we assume that the majority of the pre-trained network already possesses robust and transferable features that apply to the target domain. Therefore, instead of updating the entire model, only the final classification layer is re-trained. The generic feature extraction layers remain fixed, and only the decision-making layer that maps these features to specific attack classes is updated.

3) TL2 – Differential Fine-Tuning: Differential fine-tuning applies varying learning rates to different layers based on their role in feature extraction. We update early layers, which capture generic features, with a smaller learning rate $(1e{-}5)$ and higher learning rates $(1e{-}3)$ for deeper layers to effectively preserve stability while learning new features. We fine-tune middle layers with a moderate learning rate while updating deep layers, responsible for domain-specific features, with a higher learning rate.

4) TL3 – Selective Layer Fine-Tuning: Selective layer fine-tuning takes a more targeted approach by identifying and updating only the most critical layers for the target domain. This fine-tuning scenario employs sensitivity analysis using gradient-based methods to determine which layers are most critical to adapting the model to the target domain. In this approach, the loss gradients for the parameters of each layer are calculated to quantify the influence of each layer on the model's performance [20]. Layers with high gradient magnitudes indicate a strong sensitivity to changes in the input data and are thus identified as essential for capturing attack patterns. Only these sensitive layers are subsequently fine-tuned, while the remaining layers are kept frozen, as elaborated on in the following equations.

In particular, we consider the edge IIoT traffic data X, a pre-trained model $f(X;\theta)$ which computes predictions $\hat{y}$ and the cross-entropy loss $\mathcal{L}$:

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N} y_i \log \hat{y}_i, \tag{5}$$

where $\theta = \{\theta_1, \ldots, \theta_L\}$ are model parameters across L layers. During backpropagation, we quantify the sensitivity of layer l as the expectation of gradient magnitudes over training batches:

$$S_l = \mathbb{E}_{\text{batches}}\left[\left\|\frac{\partial \mathcal{L}}{\partial \theta_l}\right\|_2\right]. \tag{6}$$

Layers with higher S_l are prioritized for adaptation, as they most influence threat detection accuracy. We classify layers by S_l and select the top-k for adaptation. Parameters in these layers are updated via:

$$\theta_l^{t+1} = \theta_l^t - \eta \cdot \mathbb{I}(l \in \text{Top-}k) \cdot \frac{\partial \mathcal{L}_{\text{new}}}{\partial \theta_l}, \tag{7}$$

where $\mathbb{I}(\cdot)$ is an indicator function (1 if l is selected, 0 otherwise), and $\mathcal{L}_{\text{new}}$ uses target datasets. Non-selected layers remain frozen, reducing computational costs by $1 - \frac{k}{L}$ for the edge deployment.

D. Evaluation Metrics

We use accuracy, precision, recall, F1 score, and robustness as evaluation metrics to evaluate the results of the proposed TransferEdge approaches.

IV. Experiments and Evaluation

A. Dataset

We use the UNSW-NB15 dataset to build a robust source model and then adapt it for the smaller BoT-IoT dataset to simulate edge-based IIoT scenarios with limited data. To assess the model's adaptability to evolving threats, we introduce unseen attack types in the target dataset. The following sections provide details on the preprocessing steps and each dataset.

1) UNSW-NB15 dataset: A popular benchmark for network intrusion detection, the UNSW-NB15 dataset[1] covers a variety of attack methods and typical traffic patterns. Duplicate data were eliminated, missing values were handled, irrelevant features were filtered out, numerical features were normalized using Min-Max scaling, and categorical features were encoded using one-hot encoding as part of our preprocessing steps. The types of attacks and their distribution in our experiments are shown in Table I.

2) BoT-IoT datasets: The BoT-IoT dataset[2] is a commonly used benchmark for network ID, covering everyday traffic patterns and a variety of attack techniques. We use its smaller-scale records to simulate real-world Edge-IIoT scenarios with limited data availability. As the target data set, BoT-IoT tests the adaptability of our pretrained model in resource-constrained settings. To simulate evolving IIoT threats, we considered previously unseen attack types. We preprocessed

[1]https://research.unsw.edu.au/projects/unsw-nb15-dataset
[2]https://research.unsw.edu.au/projects/bot-iot-dataset

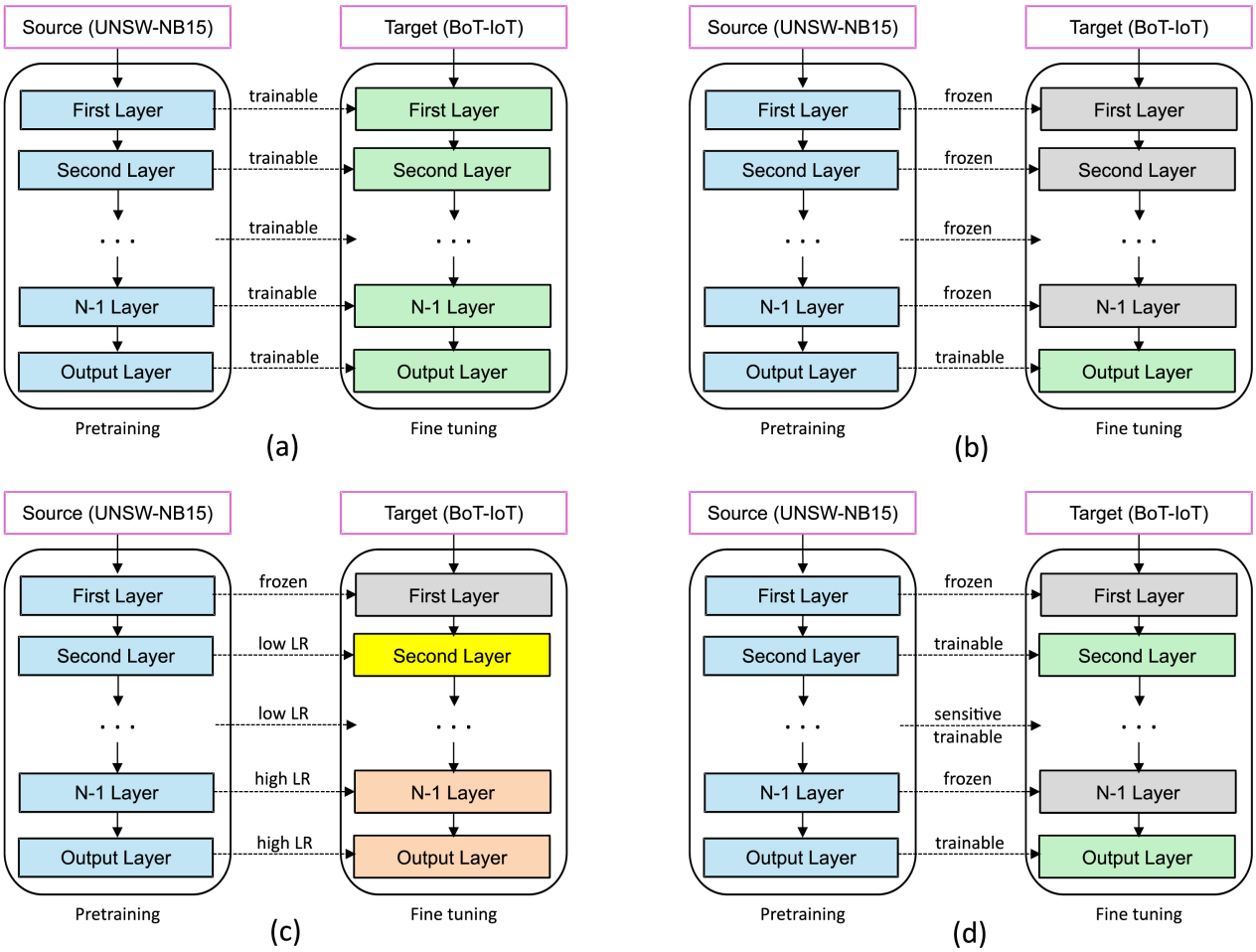

Figure 1: The overall process for all scenarios. (a): TL0 – Baseline Transfer Learning (No Freezing); (b): TL1 – Last Layer Retraining; (c) TL2 – Differential Fine-Tuning; (d): TL3 – Selective Layer Fine-Tuning.

Table I: Class Distribution in UNSW-NB15 and BoT-IoT Datasets after preprocessing

UNSW-NB15	# Records	BoT-IoT	# Records
Benign	2,218,761	Benign	92,543
Generic	215,481	DoS	32,480
Exploits	44,525	DDoS	5,194
DoS	16,353	Theft	1,587
Recon.	13,987	Recon.	21,639

the dataset by handling missing values and duplicates, applying Min-Max scaling to numerical features, and one-hot encoding categorical variables. The attack types and their distribution in our experiments are shown in Table I.

B. Hyper parameters

To find the best values for the hyperparameters, we defined a uniform search space: learning rates from 1×10^{-5} to 1×10^{-4}; batch sizes of 64, 128, and 256; weight decay from 0 to 1×10^{-4}; and fine-tuning epochs between 20 and 100.

Four transfer learning scenarios were evaluated on CNN, VGG8/16, ResNet18, EfficientNet, and Xception using random search (50 trials per model scenario) on Google Colab with a T4 GPU (16 vCPUs, 64 GB RAM).

V. Results and Discussion

A. Comparative Analysis of Fine-tuning Strategies

We evaluated the proposed EdgeTransfer fine-tuning strategies across CNN and pre-trained models, as shown in Table II. In the TL0 setting (no freezing) on the BoT-IoT dataset, VGG16 achieved near-perfect performance (99.99% precision, recall, and F1-score), outperforming CNN (99.90%), ResNet18 (99.83%), and the other models. However, VGG16 required 68.7 minutes for training-nearly three times longer than CNN (22.4 minutes), highlighting the need of a trade-off between detection accuracy and efficiency. In this sense, EfficientNet offered a balance, achieving 99.87% accuracy with a reduced training time of 47.1 minutes.

Table III presents the results for TL1 – Last Layer Retraining, where EfficientNet achieved the highest performance (99.97% across all metrics), followed closely by CNN (99.89%) and ResNet18 (99.84%). VGG16 demonstrated

Table II: Performance Comparison of Models on BoT-IoT Using TL0 – Baseline Transfer Learning (No Freezing)

Model	Accuracy	Precision	Recall	F1-Score	Train Time [m]
CNN	99.90	99.90	99.90	99.87	22.4
VGG8	99.87	99.87	99.87	99.87	41.8
VGG16	99.99	99.99	99.99	99.99	68.2
ResNet18	99.83	99.83	99.83	99.83	72.4
EfficientNet	99.87	99.87	99.87	99.87	47.1
Xception	99.91	99.91	99.91	99.91	53.51

Table III: Performance Comparison of Models on BoT-IoT Using TL1 – Last Layer Retraining

Model	Accuracy	Precision	Recall	F1-Score	Train Time [m]
CNN	99.89	99.89	99.89	99.86	38.7
VGG8	96.28	96.28	99.42	99.35	65.2
VGG16	99.64	99.64	99.64	99.64	108.4
ResNet18	99.84	99.84	99.84	99.84	59.7
EfficientNet	99.97	99.97	99.97	99.96	95.3
Xception	99.64	99.64	99.64	99.64	82.69

Table IV: Performance Comparison of Models on BoT-IoT Using TL2 – Differential Fine-Tuning

Model	Accuracy	Precision	Recall	F1-Score	Train Time [m]
CNN	99.91	99.91	99.91	99.88	35.2
VGG8	99.84	99.84	99.84	99.85	62.4
VGG16	99.88	99.89	99.88	99.88	95.6
ResNet18	99.83	99.83	99.83	99.83	89.9
EfficientNet	99.87	99.87	99.87	99.87	71.9
Xception	99.75	99.75	99.75	99.75	77.83

Table V: Performance Comparison of Models on BoT-IoT Using TL3 – Selective Layer Fine-Tuning

Model	Accuracy	Precision	Recall	F1-Score	Train Time [m]
CNN	99.89	99.89	99.89	99.86	8.4
VGG8	99.85	99.68	99.84	99.76	12.1
VGG16	99.92	99.92	99.92	99.92	24.7
ResNet18	99.88	99.88	99.88	99.88	27.5
EfficientNet	99.84	99.67	99.84	99.76	14.9
Xception	99.91	99.91	99.91	99.91	16.59

strong but slightly lower results (99.64%), while VGG8 lagged with 96.28%, despite maintaining high recall.

In scenario TL2 – Differential Fine-Tuning, as shown in Table IV, CNN achieved the highest performance (99.91%), followed closely by VGG16 (99.88%), VGG8 (99.84%), and ResNet18 (99.83%). EfficientNet also performed well (99.87%) maintaining a strong balance across all metrics.

Table V illustrates the metrics of various models in the BoT-IoT dataset under TL3 – Selective Layer Fine-Tuning. VGG16 achieves the highest accuracy (99.92%) among the models, closely followed by Xception (99.91%). These models also demonstrate strong precision, recall and F1 score metrics, indicating their robustness in detecting diverse types of DDoS attacks with minimal false positives.

These results underscore the importance of transfer learning in DDoS detection capabilities within resource-constrained Edge IIoT systems. In particular, the comprehensive fine-tuning strategy TL0 applied across all network layers enabled the improvement of model performance, with VGG16 achieving the highest detection accuracy despite its increased computational overhead. In contrast, TL3 significantly reduced training time compared to other scenarios (TL0, TL1, TL2), demonstrating its practical efficiency for deployment in Edge-IIoT environments where computational resources are limited.

B. Analysis of Training Dynamics

Figure 2 illustrates the training dynamics of CNN-based transfer learning from the source to the target domain. TL0 – Baseline Transfer Learning (No Freezing) overfits source domain patterns, exhibiting validation loss instability (fluctuations of ± 0.02 at epoch 15) despite rapid early convergence ($\sim$99% accuracy by epoch 5). In contrast, TL2 – Differential Fine-Tuning achieves optimal balance, maintaining $\sim$99% validation accuracy with minimal divergence (training-validation gap $< 2\%$, $\Delta_{\text{loss}} < 0.1$ after epoch 20), indicating stable cross-domain adaptation. While TL0 attains near-perfect training accuracy (99.9%), its validation performance aligns incrementally, reflecting gradual refinement of hierarchical characteristics. These results highlight the critical trade-off between computational efficiency and accuracy: TL2 enables resource-efficient adaptation, whereas TL0 maximizes accuracy at a higher computational cost, emphasizing the role of layer-specific flexibility in preserving domain-invariant features for cross-domain attack detection.

C. Adaptability Analysis

The transfer learning capabilities of the proposed approach demonstrate the possibility of CNNs to adapt to changing conditions, by showing how a network trained on one dataset can be effectively tuned on a different one, in some cases with limited training times. In particular, we evaluated the adaptability across six backbone architectures (CNN, VGG8, VGG16, ResNet18, EfficientNet, and Xception) in both transfer directions (UNSW-NB15->BoT-IoT and BoT-IoT->UNSW-NB15). To further evaluate how the proposed methodology can adapt to changing network conditions, we perform a double-stage transfer learning experiment, by training the CNNs on the first dataset, tuning on the second, then again tuning on the first dataset (e.g., UNSW-NB15->BoT-IoT->UNSW-NB15). We experiment in both forward and reverse directions, with no significant differences in accuracy with respect to training a CNN on a single dataset. These results demonstrate that EdgeTransfer's staged double transfer strategy consistently mitigates domain shift, making it well-suited for real-time DDoS detection at the IIoT edge under dynamic network conditions.

VI. Conclusion

This paper introduced TransferEdge, a novel approach to optimizing transfer learning and fine-tuning algorithms for detecting evolving DDoS attacks in resource-constrained Edge-IIoT environments. By integrating pre-trained models with fine-tuning strategies optimized for the specific datasets considered, TransferEdge addresses domain shift and data scarcity while optimizing detection performance. Experimental validation, using UNSW-NB15 as the source dataset and BoT-IoT as the target domain, demonstrates that TransferEdge-enhanced architectures (CNN, VGG8, VGG16, ResNet18,

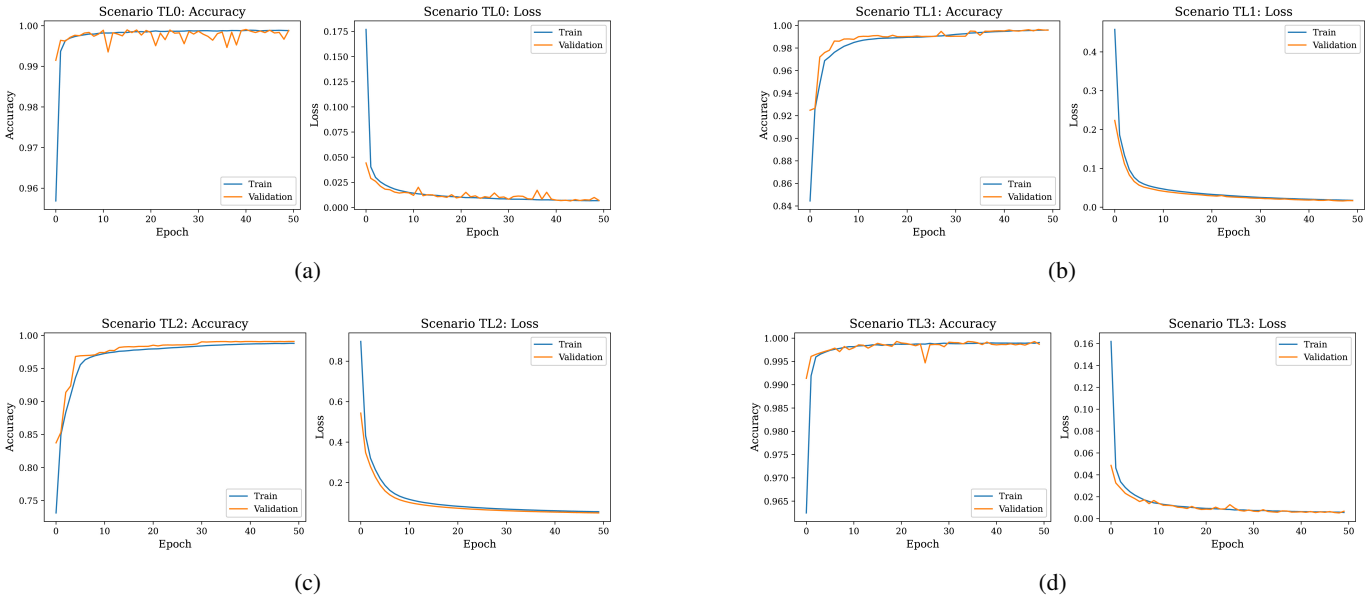

Figure 2: Training and validation performance of TransferEdge strategies (TL0-TL3) for DDoS attack detection: Transfer from UNSW-NB15 (source) to BoT-IoT (target). (a): TL0; (b): TL1; (c): TL2; (d): TL3 metrics of accuracy and loss curves.

EfficientNet, and Xception) improve accuracy in identifying novel attack patterns. These results highlight TransferEdge's effectiveness in strengthening industrial IoT cybersecurity, offering a scalable solution against DDoS threats. Future works will consider online learning to further assess performance in changing and dynamic network conditions, as well as consider a security analysis by modeling adversarial attacks and integrating cryptographic safeguards in large-scale, heterogeneous deployments.

Acknowledgment

This work was supported in part by the EC under projects Chips JU EdgeAI (101097300) and GLACIATION (101070141) and by project SERICS (PE00000014) under the MUR NRRP funded by the EU - NGEU. The project EdgeAI "Edge AI Technologies for Optimised Performance Embedded Processing" is supported by the Chips Joint Undertaking and its members including top-up funding by Austria, Belgium, France, Greece, Italy, Latvia, Netherlands, and Norway under grant agreement No 101097300. Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Chips Joint Undertaking. Neither the European Union nor the granting authority can be held responsible for them. We also thank the NVIDIA Corporation for the GPU donated.

References

[1] J. Mirkovic and P. Reiher, "A taxonomy of DDoS attack and DDoS defense mechanisms," *ACM SIGCOMM Computer Communication Review*, vol. 34, no. 2, pp. 39–53, 2004.

[2] S. Shakya and R. Abbas, "A comparative analysis of machine learning models for DDoS detection in IoT networks," *arXiv preprint arXiv:2411.05890*, 2024.

[3] M. A. Lawal, R. A. Shaikh, and S. R. Hassan, "A DDoS attack mitigation framework for IoT networks using fog computing," *Procedia Computer Science*, vol. 182, pp. 13–20, 2021.

[4] D. Agostinello, A. Genovese, V. Piuri *et al.*, "Anomaly-based intrusion detection system for DDoS attack with deep learning techniques," in *Proc. of ICSC*, 2023, pp. 267–275.

[5] L. Vu, Q. U. Nguyen, D. N. Nguyen, D. T. Hoang, and E. Dutkiewicz, "Deep transfer learning for IoT attack detection," *IEEE Access*, vol. 8, pp. 107 335–107 344, 2020.

[6] P. Wu, H. Guo, and R. Buckland, "A transfer learning approach for network intrusion detection," in *Proc. of ICBDA*, 2019, pp. 281–285.

[7] B. Farzaneh, N. Shahriar, A. H. Al Muktadir, M. S. Towhid, and M. S. Khosravani, "DTL-5G: Deep transfer learning-based ddos attack detection in 5g and beyond networks," *Computer Communications*, vol. 228, p. 107927, 2024.

[8] S. Khanal, S. Tirupathi, M. Dzaferagic, and T. B. Pedersen, "Addressing data scarcity and distribution shifts in communication networks using pre-trained transformers and transfer learning," in *Proc. of AAAI*, 2025.

[9] B. Farzaneh, N. Shahriar, A. H. Al Muktadir, and M. S. Towhid, "DTL-IDS: Deep transfer learning-based intrusion detection system in 5G networks," in *Proc. of CNSM*, 2023, pp. 1–5.

[10] H. Lu, Y. Zhao, Y. Song, Y. Yang, G. He, H. Yu, and Y. Ren, "A transfer learning-based intrusion detection system for zero-day attack in communication-based train control system," *Cluster Computing*, vol. 27, no. 6, pp. 8477–8492, 2024.

[11] F. Yan, G. Zhang, D. Zhang, X. Sun, B. Hou, and N. Yu, "TL-CNN-IDS: transfer learning-based intrusion detection system using convolutional neural network," *The Journal of Supercomputing*, vol. 79, no. 15, pp. 17 562–17 584, 2023.

[12] S. Latif, W. Boulila, A. Koubaa, Z. Zou, and J. Ahmad, "DTL-IDS: An optimized intrusion detection framework using deep transfer learning and genetic algorithm," *Journal of Network and Computer Applications*, vol. 221, p. 103784, 2024.

[13] M. B. Anley, A. Genovese, D. Agostinello, and V. Piuri, "Robust DDoS attack detection with adaptive transfer learning," *Computers & Security*, vol. 144, p. 103962, 2024.

[14] A. A. Najar and S. M. Naik, "Cyber-secure SDN: A CNN-based approach for efficient detection and mitigation of DDoS attacks," *Computers & Security*, vol. 139, p. 103716, 2024.

[15] M. Muhammad, A. S. Alshra'a, and R. German, "An IDS for DDoS attacks in SDN using VGG-based CNN architecture," in *Proc. of SoftCOM*, 2023, pp. 1–7.

[16] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[17] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. of CVPR*, 2016, pp. 770–778.

[18] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. of ICML*, 2019, pp. 6105–6114.

[19] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. of CVPR*, 2017, pp. 1251–1258.

[20] E. Yvinec, A. Dapogny, K. Bailly, and X. Fischer, "Safer: Layer-level sensitivity assessment for efficient and robust neural network inference," *arXiv preprint arXiv:2308.04753*, 2023.