A. Implementation Details of Distribution Invading Attack

In our study, we use the Distribution Invading Attack (DIA) in [54], with a detailed description found in Algorithm 1. Specifically, the test batch $\mathcal{B}^t$ undergoes an update process as outlined in Line 5. Notice that, unlike the general method of using mean, our approach utilizes median calculations as per (7) for these BN statistics. For models with BN layers, executing Line 6 is optional. TTA methods typically perform a single-step update using TTA loss $\mathcal{L}_{\text{TTA}}$ on θ for each $\mathcal{B}^t$, allowing us to estimate $\hat{\theta}$ to be approximately equal to θ. In Line 7, the perturbation δ_{i-1} is updated through projected gradient descent (PGD) [36], where the projection Π_ε is used to clip δ_i within the constraint ε. This process ensures that the images remain valid within the [0, 1] range. $\mathcal{L}_{\text{attack}}$ is replaced by adversary's objectives: targeted attack or indiscriminate attack in Section 4. After N-steps PGD, we get the optimal malicious samples $\hat{\mathcal{B}}^t_{\text{mal}} = \mathcal{B}^t_{\text{mal}} + \delta_N$.

Algorithm 1: Distribution Invading Attack [54]

1: **Input:** Model $f(\cdot;\theta)$ of parameters θ which include BN statistics $(\hat{\mu}_c, \hat{\sigma}^2_c)$, test batch $\mathcal{B}^t = \mathcal{B}^t_{\text{mal}} \cup \mathcal{B}^t_{\text{ben}}$ at time t, a targeted label y^t_{target} on a targeted sample $x^t_{\text{target}} \in \mathcal{B}^t_{\text{ben}}$, learning rate η of TTA update, learning rate α of attack, the number of attack steps N, constraint ε, and perturbation δ_0.
2: **Output:** Perturbed malicious samples $\hat{\mathcal{B}}^t_{\text{mal}} = \mathcal{B}^t_{\text{mal}} + \delta_N$
3: **for** $i = 1, 2, \ldots, N$ **do**:
4: $\quad \mathcal{B}^t \leftarrow (\mathcal{B}^t_{\text{mal}} + \delta_{i-1}) \cup \mathcal{B}^t_{\text{ben}}$
5: $\quad (\hat{\mu}_c, \hat{\sigma}^2_c) \leftarrow (\hat{\mu}_c(\mathcal{B}^t), \hat{\sigma}^2_c(\mathcal{B}^t))$
6: $\quad$ (Optional) $\hat{\theta} \leftarrow \theta - \eta \cdot \partial \mathcal{L}_{\text{TTA}}(\mathcal{B}^t)/\partial\theta$
7: $\quad \delta_i \leftarrow \Pi_\varepsilon(\delta_{i-1} - \alpha \cdot \text{sign}(\nabla_{\delta_{i-1}} \mathcal{L}_{\text{attack}}(f(\cdot\,; \hat{\theta}(\mathcal{B}^t)))))$
8: **end for**
9: **return** $\hat{\mathcal{B}}^t_{\text{mal}} = \mathcal{B}^t_{\text{mal}} + \delta_N$

As we mentioned in Section 4, we discuss the attack objectives $\mathcal{L}_{\text{attack}}$ for two types of attacks: targeted attack and indiscriminate. The targeted attack involves an adversarial input $\hat{\mathcal{B}}^t_{\text{mal}}$ to make the model misclassify a specific sample x^t_{target} to incorrect label y_{target}, formulated as: $\hat{\mathcal{B}}^t_{\text{mal}} = \arg\max_{\mathcal{B}^t_{\text{mal}}} -\mathcal{L}_{\text{CE}}(f(x^t_{\text{target}}; \hat{\theta}(\mathcal{B}^t)), y^t_{\text{target}})$. On the other hand, the objective of indiscriminate attack is to reduce the model's accuracy on all benign data by manipulating the adversarial input $\hat{\mathcal{B}}^t_{\text{mal}}$, given by: $\hat{\mathcal{B}}^t_{\text{mal}} = \arg\max_{\mathcal{B}^t_{\text{mal}}} \sum_{(x,y)\in\mathcal{Z}^t_{\text{ben}}} \mathcal{L}_{\text{CE}}(f(x; \hat{\theta}(\mathcal{B}^t)), y)$.

B. Extended Attack Scenarios

The required knowledge of the white-box attack is excessive but not unattainable. Nevertheless, it is crucial to explore more feasible attacks with constrained knowledge of the adversary. Therefore, we consider two additional attack scenarios: the semi-white box attack scenario, in which the adversary has constrained knowledge, and the adaptive attack, in which the adversary adapts its adversarial objective to obfuscate defense mechanisms.

Semi-white-box attack. We construct a semi-white-box attack that generates malicious samples using only the initial model parameters, while the system continues to adapt its parameters. This approach is more feasible but weaker than the white-box attack. As indicated in Table 7, the malicious samples generated by the semi-white-box attacker are comparably toxic to those from the white-box attacker in an instant attack scenario, and our method demonstrates robustness against such attacks.

Adaptive attack. Since the adversary is aware of defense mechanisms, it can adapt its adversarial objective to obfuscate them. To verify the robustness of our method against such an adaptive attack, we implement it with an additional regularization term, $|\text{med}(\mathcal{B}_{\text{mal}}) - \text{med}(\mathcal{B}_{\text{ben}})|$, which ensures alignment between the median of malicious samples and that of benign samples. However, as shown in Table 7, the adaptive attack is weaker than the white-box attack, and our method (MedBN) is still robust against such adaptive attacks.

Table 7. Attack Success Rate (%) of targeted attack with TENT.

Attack Method	White-box	Semi-white-box	Adaptive white-box
BatchNorm	72.36	53.73	31.87
MedBN (Ours)	**18.36**	**11.20**	**7.47**

C. Experiment Details

Datasets. Three major benchmarks for TTA [25] CIFAR10-C, CIFAR100-C, and ImageNet-C. These benchmarks are designed to measure the robustness of networks in classification tasks. Each dataset includes 15 types of corruption and 5 levels of severity. Our evaluation concentrates on the most severe level 5 of corruption. The CIFAR10-C and CIFAR100-C datasets contain 10,000 test images with 10 and 100 classes, respectively, and the ImageNet-C dataset contains 5,000 test images with 1,000 classes for each type of corruption.

Implementation details. In all experiments, we adapt the Adam optimizer with a learning rate of 0.001 and no weight decay. For SAR and SoTTA, we use the SAM optimizer with the Adam optimizer. We follow the baseline papers or official codes to set the hyper-parameters for each TTA method. For data poisoning attacks, we follow the experimental setting of unconstrained attack in [54], which is the most threatening attack. Specifically, we use attack steps of 100 with an attacking optimization rate α of 1/255, the initial perturbation δ_0 of 0.5, and the perturbation constraint ε of 1.0.

Details on semantic segmentation task. Our experimental setup aligns with prior works [9, 34] on semantic segmentation. We utilize DeepLabv3+ [7] with ResNet-101 backbone pre-trained on the Cityscapes training set [12] and evaluate its performance on the validation set of SYNTHIA [42]. In evaluating targeted attack within the segmentation task, we adopt the metric of Attack Success Rate (ASR), akin to image classification. The performance of indiscriminate attacks is evaluated through the mean Intersection over Union (mIoU) on benign samples after the attack.

D. Extended Related Works

Test Time Adaptation (TTA). TTA has been studied to address the issue of distribution shift between source and target domains during the online testing phase, without altering training phase. TTA methods can be broadly categorized into three groups on the specific parameters they update within a model: (i) fully-updated TTA; update all parameters of the model, (ii) BN-updated TTA; update only BN parameters of the model, and (iii) meta-updated TTA; update meta networks attached with frozen pre-trained model. Several studies [6, 10, 14, 35, 53] have improved performance by updating entire model parameters, which may be impractical when the available memory sizes are limited. The majority of fully-updated TTA methods adopt the mean-teacher framework, which largely relies on pseudo-labeling of a more reliable teacher model. The stability of mean-teacher frameworks in changing environments is attributed to their use of an exponential moving average with various loss functions, such as symmetric cross-entropy.

Since fully-updated TTA methods encompass BN-updated TTA approaches, TTA typically involves adapting pre-trained models that include BN layers [28], which often struggle with domain shifts at test time due to their reliance on training statistics optimized for the training distribution. Prior methods in TTA [37, 44] have indicated that adapting BN statistics can effectively mitigate distributional shifts. Moreover, recent TTA approaches [20, 34, 40] have primarily focused on utilizing normalization statistics directly from the current test input, often in conjunction with self-training techniques, such as entropy minimization [39, 52, 61]. Meanwhile, in addition to works [4, 27, 57] focusing on memory efficiency, [47] proposes an architecture that is efficient in terms of memory. This design combines frozen original networks with newly proposed meta networks, requiring an initial warm-up using source data. To address the adversarial risks in TTA methods if BN layers are being adapted, we propose MedBN method that can be integrated into any existing TTA methods if BN layers are being adapted and demonstrate a theoretical analysis of our method. When MedBN is integrated into these methods, they consistently demonstrate robustness against malicious samples.

Data poisoning attacks and defenses. Data poisoning attacks involve injecting poisoned samples into a dataset, causing the model trained with the poisoned dataset to produce inaccurate results at test time. These attacks pose threats to various machine learning algorithms [2, 5, 38, 45]. Furthermore, recently, [11, 54] suggest the risks of data poisoning attacks in the test-time adaptation process, wherein TTA methods adapt the model at test time.

For defense against data poisoning attacks, [48] removes outliers by approximating the upper bounds of loss. This method requires the assumption that the dataset is large enough to approximate the loss. However, for test-time adaptation, the number of test data is insufficient to concentrate statistics of loss, and there are no labels for the test data, which means that this approach is not suitable for adaptation during the test phase. [18] demonstrates that adversarial training is an effective defense method for data poisoning attacks, enhancing the robustness of models in the training phase. However, in test-time adaptation, access to the training process is restricted, primarily due to privacy concerns related to the training data and the substantial computational resources required for training. Additionally, adversarial training leads to performance degradation in test data. Due to the aforementioned limitations, adversarial training is infeasible in this context. To address the above limitations of existing defenses, we propose a robust batch normalization method that is not only simple and effective but also universally

applicable across any existing TTA methods if BN layers are being adapted.

Median aggregation for robust distributed learning. The abundance of collected data has led to the emergence of distributed learning frameworks. In such systems, several data owners or workers collaborate to construct a global model, typically employing the widely used distributed stochastic gradient descent (SGD) algorithm with a central server. This server iteratively updates the model parameter estimated by aggregating the stochastic gradients calculated by the workers. However, this algorithm is susceptible to misbehaving workers, referred to as Byzantine in [32], that may send arbitrarily deceptive gradients to the server, potentially disrupting the learning process [1,49,56]. To address these issues, extensive researches [8,15,22,23,55,58] have been dedicated to robustly aggregating gradients regardless of Byzantine behavior. Among a wide range of aggregation methods, the median is widely used for robust aggregation and its effectiveness has been verified: [8] employs the geometric median for robust aggregation, [55] uses the mean around the median, and [58] utilizes coordinate-wise median. In terms of robust aggregation, the median can also be applied to robustly aggregate batch statistics against malicious samples. To the best of our knowledge, we are the first to use the median for robustly aggregating batch statistics to defend against malicious samples.

E. Effectiveness of MedBN across Different Model Architectures

In the main text, we have focused on ResNet-26. Beyond ResNet-26, our study includes two additional architectures, which are commonly used in TTA: WideResNet-28 (WRN-28) for CIFAR10-C, as referenced in the RobustBench benchmark [13], and ResNext-29 for CIFAR100-C from [26]. Table 8 demonstrates the efficacy of MedBN across various architectures over both attack instant scenarios, indicating that MedBN is independent of specific architectural designs, i.e., architecture-agnostic.

Table 8. Effectiveness of MedBN across various model architectures. We use the batch size of 200 with 40 malicious samples.

Objective	Dataset	Architectures	Normalization	Method							$m = 0$
				TeBN	TENT	ETA	SAR	SoTTA	sEMA	mDIA	TeBN (ER %)
Targeted Attack	CIFAR10-C	WRN-28	BatchNorm	86.67	80.53	82.00	82.00	27.47	20.53	25.87	20.43
			Ours (MedBN)	**24.53**	**23.33**	**23.07**	**22.27**	**9.07**	**8.53**	**11.87**	**21.49**
	CIFAR100-C	ResNext-29	BatchNorm	96.67	80.00	79.73	84.00	12.93	9.20	7.87	35.56
			Ours (MedBN)	**3.07**	**2.13**	**2.27**	**2.13**	**1.60**	**2.00**	**1.07**	**37.62**
Indiscriminate Attack	CIFAR10-C	WRN-28	BatchNorm	37.30	34.35	33.70	34.20	27.69	28.45	42.95	20.43
			Ours (MedBN)	**29.63**	**27.24**	**26.69**	**26.96**	**23.98**	**25.42**	**35.21**	**21.49**
	CIFAR100-C	ResNext-29	BatchNorm	62.35	52.02	51.14	52.75	44.44	45.93	46.90	35.56
			Ours (MedBN)	**43.81**	**39.15**	**39.32**	**40.20**	**37.47**	**40.51**	**40.06**	**37.62**

F. Extended Ablation Study Cases

In this section, we present detailed results of the three additional cases in our ablation studies, which were not included in Section 7.5. Each case explores different combinations of datasets and attack scenarios, providing further insights into the robustness of our method.

The number of malicious samples. We investigate MedBN's robustness against different ratios of malicious samples using a batch size of 200. In the instant attack scenario, MedBN demonstrates robustness across all malicious ratios, performing well under both targeted attacks (Table 9) and indiscriminate attacks (Table 10).

Table 9. Attack Success Rate (%) of targeted and instant attack for different numbers of malicious samples m with batch size of 200.

Dataset	Normalization	# of Malicious Samples (m)				
		10	20	40	60	80
CIFAR10-C	BatchNorm	21.60	42.00	84.00	96.67	99.47
	Ours (MedBN)	**7.07**	**10.27**	**19.20**	**26.80**	**38.27**
CIFAR100-C	BatchNorm	16.80	42.13	92.00	99.73	99.87
	Ours (MedBN)	**1.73**	**2.00**	**2.93**	**3.60**	**4.27**

Table 10. Error Rate (%) of indiscriminate and instant attack for different number of malicious samples m with batch size of 200.

Dataset	Normalization	# of Malicious Samples (m)				
		10	20	40	60	80
CIFAR10-C	BatchNorm	19.07	22.98	31.02	40.14	50.52
	Ours (MedBN)	**16.42**	**18.00**	**22.34**	**28.00**	**34.24**
CIFAR100-C	BatchNorm	45.35	50.03	59.84	69.21	78.99
	Ours (MedBN)	**43.31**	**44.38**	**48.58**	**53.86**	**61.44**

Test batch size. We assess the effect of varying batch sizes with a fixed ratio of malicious smaples around 20%. In the case of targeted attacks, MedBN consistently achieves significantly lower ASR compared to BN across all batch sizes (refer to Table 11. Similarty, in the case of indiscriminate attacks, MedBN consistently outperforms BN with lower error rates across all tested batch sizes (refer to Table 12).

Table 11. Attack Success Rate (%) of targeted and instant attack for different batch size B with a consistent 20% of malicious samples.

Dataset	Normalization	Batch-size (B)				
		200	128	64	32	16
CIFAR10-C	BatchNorm	83.91	87.76	84.84	83.87	84.60
	MedBN (Ours)	**19.16**	**20.51**	**17.83**	**20.19**	**29.14**
CIFAR100-C	BatchNorm	91.78	88.44	89.43	90.46	90.47
	MedBN (Ours)	**2.80**	**4.72**	**5.01**	**8.20**	**12.65**

Table 12. Error Rate (%) of indiscriminate and instant attack for different batch size B with a consistent 20% of malicious samples.

Dataset	Normalization	Batch-size (B)				
		200	128	64	32	16
CIFAR10-C	BatchNorm	31.02	33.14	35.01	40.67	49.85
	MedBN (Ours)	**22.34**	**23.83**	**24.78**	**28.58**	**34.81**
CIFAR100-C	BatchNorm	59.80	62.35	67.07	73.73	83.08
	MedBN (Ours)	**48.55**	**49.86**	**52.88**	**58.80**	**67.63**

G. Discussion on Median Absolute Deviation (MAD)

We further explore the feasibility of using Median Absolute Deviation (MAD) as an alternative to the mean of squared deviations $(z_{bchw} - \eta_c)^2$, used in our MedBN. The MAD is calculated as the median of the absolute deviations from the median of data, formulated as:

$$\text{med}\left(|z_{bchw} - \eta_c|\right)_{bhw} \ . \tag{16}$$

As discussed in Section 5.1, ρ_c typically computes the mean of squared deviations $(z_{bchw} - \eta_c)^2$, opting for MAD presents an alternative method. Our findings reveal that while adopting MAD enhances defense capabilities, specifically in the targeted attack, it also results in a notable decrease in performance, particularly over ImageNet-C, as detailed in Table 13.

H. Extension of Theorem 1

In this appendix, we extend Theorem 1 for multi-dimensional vectors. For median of multi-dimensional vectors, we consider coordinate-wise median (cwmed) and geometric median (geomed). The coordinate-wise median is the median along each dimension. The geometric median is a vector that minimizes the sum of the distances to vectors in $\mathcal{B} = \{x_i \in \mathbb{R}^C : i \in [n]\}$ with a set of n numbers, which is defined as follows:

$$\text{geomed}(\mathcal{B}) = \arg\min_{z \in \mathbb{R}^d} \sum_{x_i \in \mathcal{B}} \|z - x_i\|_2 \ . \tag{17}$$

Note that cwmed is a solution of $\arg\min_{z \in \mathbb{R}^d} \sum_{x_i \in \mathcal{B}} \|z - x_i\|_1$.

Table 13. Comparison of BatchNorm, MedBN (our method), and MAD in terms of Attack Success Rate (%) for the targeted and instant attack scenario and Error Rate (%) for the indiscriminate and instant attack scenario using TeBN. This table also includes Error Rate (%) on benign samples without attack as per TTA benchmarks.

	Dataset	CIFAR10-C		CIFAR100-C		ImageNet-C	
	m / B	40 / 200		40 / 200		20 / 200	
Objective	Normalization	ER (%) w/o Attack	ASR (%)	ER (%) w/o Attack	ASR (%)	ER (%) w/o Attack	ASR (%)
Targeted Attack	BatchNorm	14.92	83.90	40.08	91.78	66.62	97.78
	Ours (MedBN)	15.19	19.16	40.77	2.80	69.55	0.36
	MAD	18.40	2.93	46.13	0.13	85.08	0.27
Objective	Normalization	ER (%) w/o Attack	ER (%)	ER (%) w/o Attack	ER (%)	ER (%) w/o Attack	ER (%)
Indiscriminate Attack	BatchNorm	14.92	31.02	40.08	59.80	66.62	81.46
	Ours (MedBN)	15.19	22.34	40.77	48.55	69.55	69.74
	MAD	18.40	23.46	46.13	53.41	85.08	84.99

Theorem 2 (Extension of Theorem 1) *Consider a set of n numbers $\mathcal{B} = \{x_i \in \mathbb{R}^C : i \in [n]\}$ and $1 \leq m \leq n$ where the first m numbers are possibly manipulated by adversaries. Let $\mathcal{B}_{\text{mal}} = \{x_i : i \in [m]\}$, and $\mathcal{B}_{\text{ben}} = \mathcal{B} \setminus \mathcal{B}_{\text{mal}}$.*
(i) The mean can be arbitrarily manipulated by a single malicious sample, i.e., for any $1 \leq m \leq n$,

$$\sup_{\mathcal{B}_{\text{mal}}} \|\text{mean}(\mathcal{B}_{\text{mal}} \cup \mathcal{B}_{\text{ben}}) - \text{mean}(\mathcal{B}_{\text{ben}})\|_2 = \infty \,. \tag{18}$$

(ii) The cwmed *or* geomed *are robust against malicious samples unless they are not the majority, i.e., for any $1 \leq m < n/2$. For the simplicity, we denote the* med *instead of* cwmed *or* geomed*,*

$$\sup_{\mathcal{B}_{\text{mal}}} \|\text{med}(\mathcal{B}_{\text{mal}} \cup \mathcal{B}_{\text{ben}}) - \text{med}(\mathcal{B}_{\text{ben}})\|_2 < \infty \,, \text{ and} \tag{19}$$

$$\sup_{\mathcal{B}_{\text{mal}}} \|\text{med}(\mathcal{B}_{\text{mal}} \cup \mathcal{B}_{\text{ben}}) - \text{mean}(\mathcal{B}_{\text{ben}})\|_2 < \infty \,. \tag{20}$$

Proof of Theorem 2. First, we prove the vulnerability of mean (18). The k-th coodrinate of $\|\text{mean}(\mathcal{B}_{\text{mal}} \cup \mathcal{B}_{\text{ben}}) - \text{mean}(\mathcal{B}_{\text{ben}})\|$ is $\text{mean}(\mathcal{B}_{\text{mal}} \cup \mathcal{B}_{\text{ben}})_k - \text{mean}(\mathcal{B}_{\text{ben}})_k$. Then, $\|\text{mean}(\mathcal{B}_{\text{mal}} \cup \mathcal{B}_{\text{ben}}) - \text{mean}(\mathcal{B}_{\text{ben}})\| = \sqrt{\sum_{k=1}^{C} |\text{mean}(\mathcal{B}_{\text{mal}} \cup \mathcal{B}_{\text{ben}})_k - \text{mean}(\mathcal{B}_{\text{ben}})_k|^2}$. Consequently, by (9), the equation (18) holds.

For the second part on the robustness of the median, particularly for the cwmed, we can demonstrate (19) and (20) by using (11) and (10), similar to the proof of (18). Regarding the geomed, we can use Lemma 9 in [16],

$$\|\text{med}(\mathcal{B}_{\text{mal}} \cup \mathcal{B}_{\text{ben}}) - \text{med}(\mathcal{B}_{\text{ben}})\|_2 = \frac{1}{\sqrt{1 - \frac{m^2}{(n-m)^2}}} \max_{x_j \in \mathcal{B}_{\text{ben}}} \|x_j - \text{med}(\mathcal{B}_{\text{ben}})\|_2 < \infty \,. \tag{21}$$

Therefore, by (21), the equation (19) holds. Similarly, we can demonstrate the (20).

Remarks. In contrast to cwmed, calculating geomed is computationally expensive as it necessitates an optimization procedure. Therefore, although geomed can be considered for robust batch normalization, it is challenging to apply it to the neural network, which generally operates with high dimensional features.

I. Comprehensive Analysis of Malicious Samples on Every BN Layers

For analyzing the effect of MedBN, we plot the t-SNE of features before going through BN layers. For evaluation, we use Gaussian corruptions in CIFAR10-C with ResNet26 and TeBN for the adaptation method. The attack is implemented for targeted and instant attack scenario and we use $\varepsilon = 1$ for the attack. Figure 8a shows that for the deeper layer, the malicious samples tend to be clustered and distant from the benign samples to mislead the output of the model.

Additional analysis with constrained ε. We conduct a comparative analysis of BN and MedBN under the same setup, except for using a constrained ε value of $8/255$. Table 14 shows that the reduced ε leads a lower ASR compared to $\varepsilon = 1$, indicating a weaker attack. Moreover, our methods outperforms in both cases, with $\varepsilon = 1$ and $\varepsilon = 8/255$. We plot the L_1 distance as outlined in Section 7.4. Figure 6 and Figure 7 show that MedBN statistics is less influenced by malicious samples than BN statistics. Comparing the early layers (specifically, in bn1) between attack with $\varepsilon = 1$ (the left of the Figure 6) and attack with $\varepsilon = 8/255$ (the right of the Figure 6), we can observe that a smaller ε value leads to reduced perturbations in the early layers. In other words, as the weaker attack, the perturbation for early layers is reduced.

Additionally, we visualize t-SNE of all layers and ε's in Figure 8 and Figure 9. In contrast to BN layers in $\varepsilon = 1$ (Figure 8a), BN layers in $\varepsilon = 8/255$ (Figure 9a) shows that malicious samples tend to be clustered and become more distant from benign samples at deeper layers than under $\varepsilon = 1$, indicating a weakened attack. However, as shown in MedBN layers in $\varepsilon = 1$ (Figure 8b), MedBN layers in $\varepsilon = 8/255$ (Figure 9b) demonstrates that MedBN effectively mitigates the malicious samples to not be outlier against the benign samples, i.e., malicious samples are closed from the benign samples.

Table 14. Attack Success Rate (%) of targeted and instant attacks for different ε by using TeBN.

Dataset	Normalization	value of ε	
		8/255	1
CIFAR10-C	BatchNorm	58.00	83.91
	MedBN (Ours)	**16.13**	**19.16**

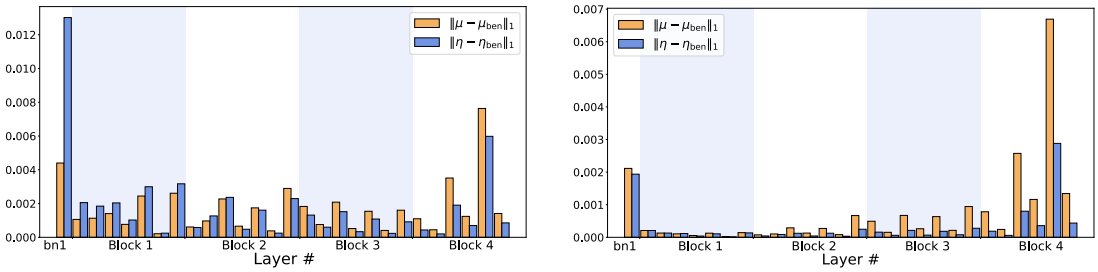


Figure 6. L1 distance for measuring the amount of perturbation $\|\mu - \mu_{\text{ben}}\|_1$ and $\|\eta - \eta_{\text{ben}}\|_1$ by malicious samples across various layers, with $\varepsilon = 1$ on the left and $\varepsilon = 8/255$ on the right.

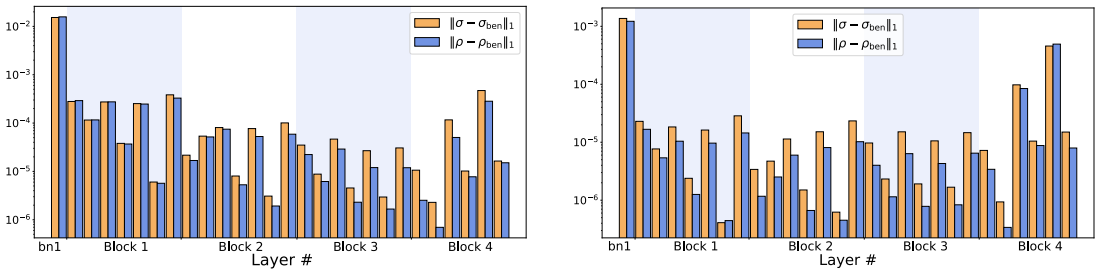
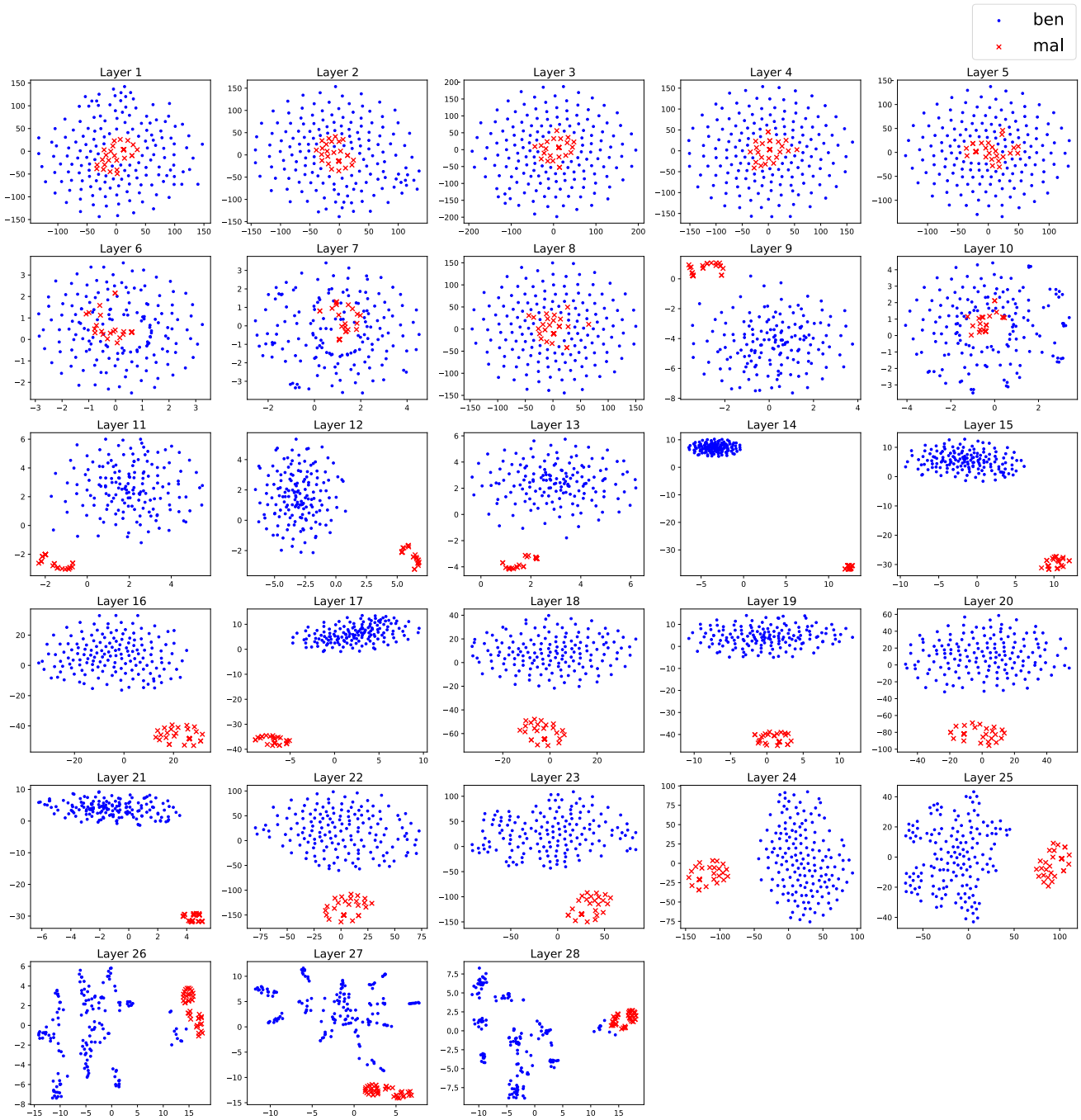
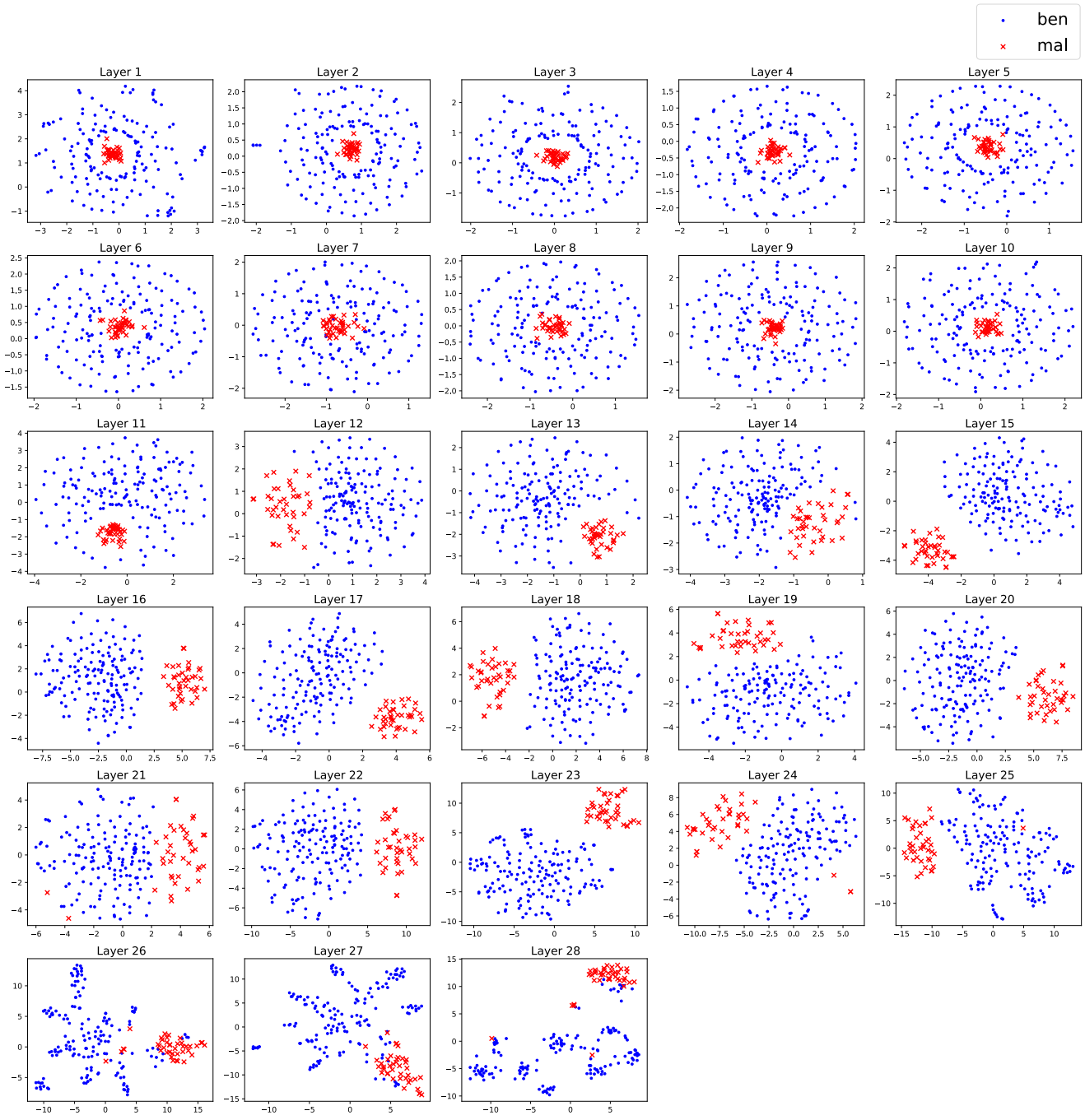


Figure 7. L1 distance for measuring the amount of perturbation $\|\sigma - \sigma_{\text{ben}}\|_1$ and $\|\rho - \rho_{\text{ben}}\|_1$ by malicious samples across various layers, with $\varepsilon = 1$ on the left and $\varepsilon = 8/255$ on the right.

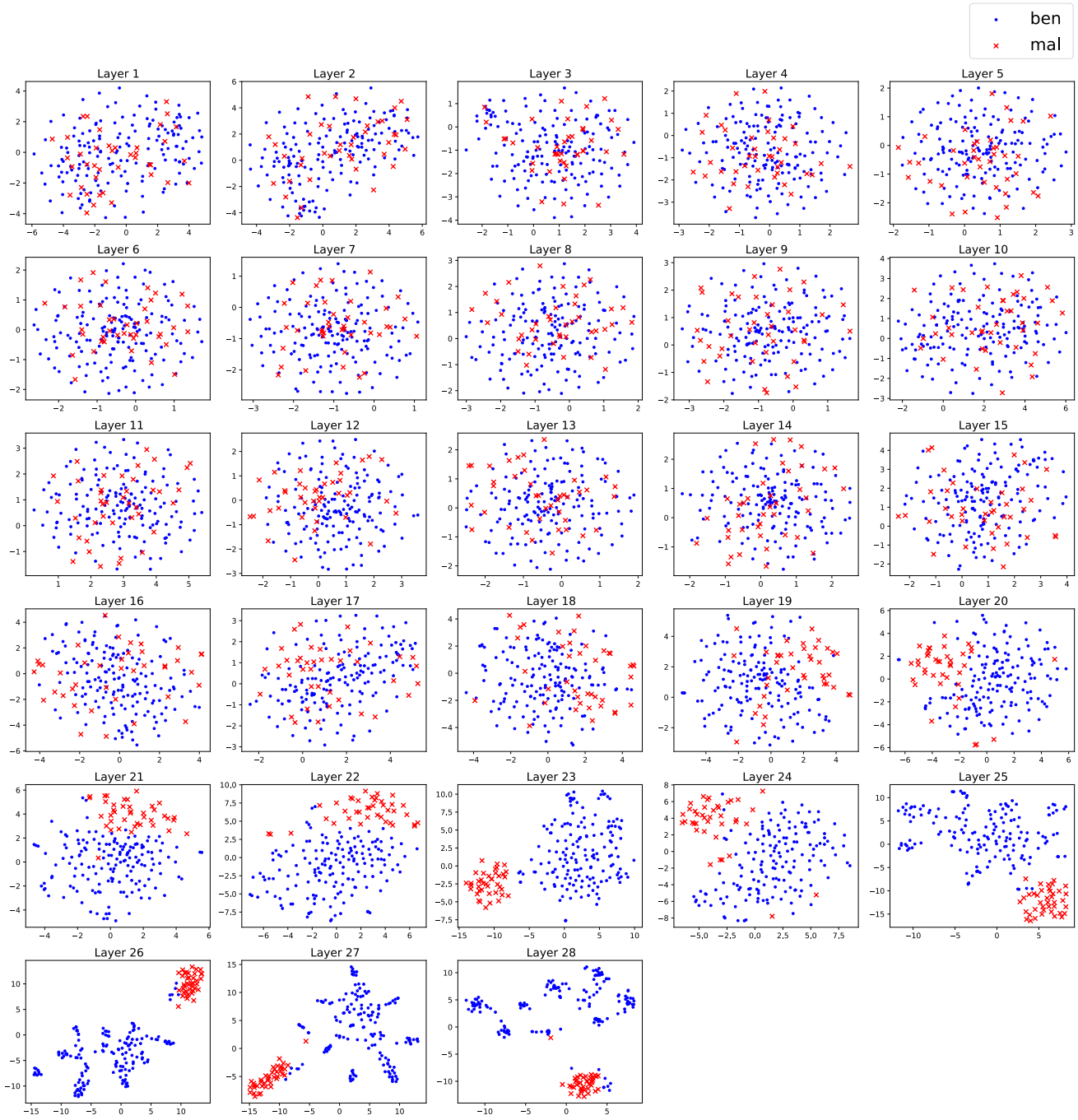


(a) t-SNE visualization of all BN layers ($\varepsilon = 1$).

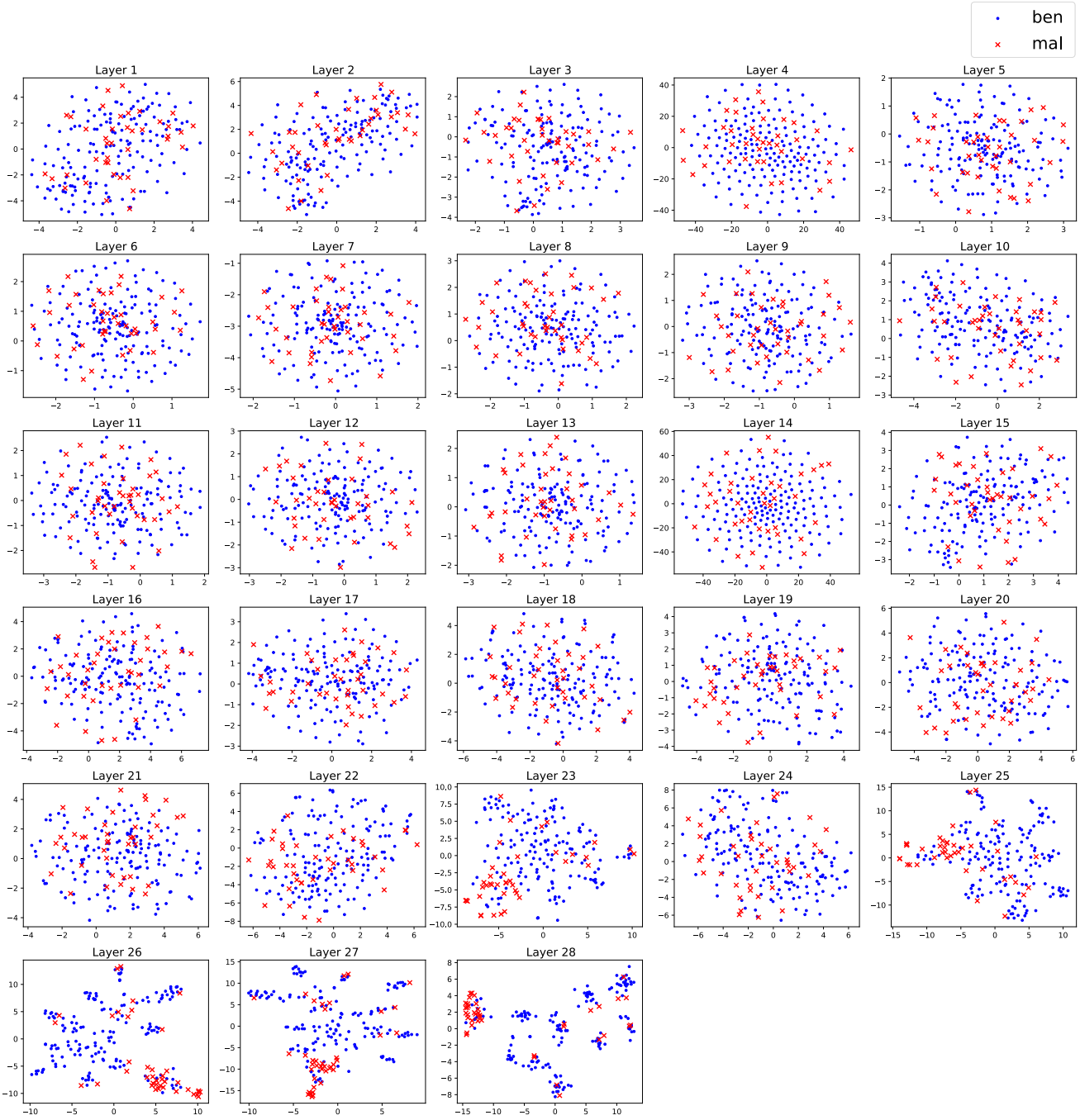


(b) t-SNE visualization of all MedBN layers ($\varepsilon = 1$).

Figure 8. t-SNE visualization of all BN layers (Figure 8a) and all MedBN layers (Figure 8b) with $\varepsilon = 1$. In Figure 8a, for deeper layers, the features of malicious samples tend to be distant from benign samples to mislead the outputs of model. However, when we apply MedBN, Figure 8b demonstrates that malicious samples are closed to benign samples, i.e. the effect of malicious samples is significantly mitigated.



(a) t-SNE visualization of all BN layers ($\varepsilon = 8/255$).



(b) t-SNE visualization of all MedBN layers ($\varepsilon = 8/255$).

Figure 9. t-SNE visualization of all BN layers (Figure 9a) and all MedBN layers (Figure 9b) with $\varepsilon = 8/255$. For the early layers, the features of malicious samples tend to be more close to those of benign samples as the ε is reduced. For deeper layers, similar to Figure 8a, the malicious samples tend to move away from the benign samples to mislead the model. However, when we apply the MedBN, the impact of malicious samples is significantly mitigated as shown in Figure 9b.

J. Comprehensive Results of Instant Attack Scenario

We provide detailed results of instant and targeted attack scenario in Table 15 and instant and indiscriminate attack scenario in Table 16 across all types of corruptions in the TTA benchmark datasets.

Table 15. Extended analysis of Attack Success Rate (%) for targeted and instant attack scenario over all types of corruption (detailed version of Table 1).

		Noise			Blur				Weather				Digital				
	Method	Gauss.	Shot	Impul.	Defoc.	Glass	Motion	Zoom	Snow	Frost	Fog	Brit.	Contr.	Elastic	Pixel.	JPEG	Avg.
CIFAR10-C	TeBN	82.00	90.00	91.33	76.67	94.00	82.67	80.00	82.00	84.00	89.33	73.33	92.00	86.00	77.33	78.00	83.91
	+MedBN	**26.00**	**23.33**	**22.00**	**16.00**	**28.67**	**16.00**	**14.67**	**21.33**	**16.00**	**16.67**	**12.00**	**16.67**	**21.33**	**10.67**	**26.00**	**19.16**
	TENT	73.33	78.00	77.33	57.33	76.00	67.33	68.67	72.67	71.33	72.67	67.33	82.67	74.67	70.00	76.00	72.36
	+MedBN	**23.33**	**18.67**	**19.33**	**16.00**	**30.67**	**16.00**	**14.00**	**16.67**	**12.67**	**14.67**	**14.00**	**19.33**	**21.33**	**16.67**	**22.00**	**18.36**
	ETA	70.67	80.67	82.67	71.33	80.00	74.67	75.33	71.33	76.00	81.33	56.67	88.00	76.67	66.67	74.00	75.007
	+MedBN	**24.67**	**22.00**	**21.33**	**14.67**	**26.00**	**14.67**	**12.00**	**20.00**	**12.00**	**18.00**	**10.67**	**16.00**	**20.00**	**14.67**	**23.33**	**18.00**
	SAR	74.00	78.67	82.67	69.33	86.00	80.67	77.33	75.33	76.00	77.33	66.67	86.67	83.33	72.67	74.67	77.42
	+MedBN	**24.00**	**24.00**	**17.33**	**15.33**	**28.00**	**14.00**	**11.33**	**17.33**	**14.00**	**14.67**	**14.67**	**16.67**	**19.33**	**16.00**	**24.00**	**18.04**
	SoTTA	25.33	20.67	30.00	20.00	24.00	22.00	15.33	21.33	17.33	18.67	22.67	18.00	22.67	21.33	22.67	21.47
	+MedBN	**7.33**	**16.67**	**10.67**	**6.67**	**6.67**	**3.33**	**6.00**	**8.00**	**5.33**	**6.00**	**4.00**	**8.00**	**12.00**	**6.67**	**10.00**	**7.82**
	sEMA	24.67	25.33	23.33	14.00	24.67	13.33	14.00	20.00	12.00	12.67	14.67	12.67	19.33	16.00	26.00	18.18
	+MedBN	**14.00**	**14.00**	**14.67**	**2.00**	**12.00**	**6.00**	**4.00**	**8.00**	**6.00**	**10.00**	**4.00**	**6.00**	**9.33**	**2.00**	**18.00**	**8.67**
	mDIA	44.00	34.00	52.67	24.00	52.00	28.67	26.00	25.33	20.00	34.00	22.00	36.00	42.00	34.00	34.00	33.91
	+MedBN	**12.00**	**14.00**	**16.00**	**2.00**	**10.00**	**6.00**	**4.00**	**8.00**	**4.00**	**10.00**	**4.00**	**8.00**	**12.00**	**3.33**	**18.00**	**8.76**
CIFAR100-C	TeBN	96.00	96.00	98.00	77.33	91.33	88.67	86.67	98.00	98.67	99.33	94.00	98.00	88.00	83.33	83.33	91.78
	+MedBN	**2.67**	**2.00**	**2.00**	**2.00**	**5.33**	**2.00**	**2.00**	**2.67**	**2.00**	**4.00**	**4.00**	**2.67**	**4.00**	**2.67**	**2.00**	**2.80**
	TENT	78.67	84.67	73.33	81.33	73.33	81.33	65.33	74.67	78.67	88.67	80.67	92.00	84.00	84.00	68.67	79.29
	+MedBN	**3.33**	**4.00**	**4.00**	**4.00**	**6.67**	**3.33**	**4.67**	**4.67**	**2.00**	**6.67**	**4.67**	**4.67**	**4.67**	**3.33**	**2.00**	**4.18**
	ETA	78.00	80.67	81.33	84.67	74.00	78.00	71.33	77.33	84.00	92.00	72.67	90.67	84.67	76.67	73.33	79.96
	+MedBN	**2.00**	**3.33**	**2.00**	**4.00**	**5.33**	**3.33**	**3.33**	**2.00**	**2.00**	**6.00**	**3.33**	**2.00**	**2.67**	**2.00**	**2.00**	**3.02**
	SAR	86.00	83.33	86.00	74.67	78.67	76.00	70.67	81.33	85.33	93.33	76.00	96.67	79.33	86.00	71.33	81.64
	+MedBN	**2.00**	**2.00**	**2.00**	**4.00**	**7.33**	**1.33**	**4.00**	**2.00**	**1.33**	**6.00**	**4.00**	**2.00**	**3.33**	**2.67**	**1.33**	**3.02**
	SoTTA	6.67	10.00	7.33	7.33	12.67	8.00	5.33	8.67	6.67	8.67	4.00	7.33	10.00	3.33	8.00	7.60
	+MedBN	**2.00**	**2.00**	**3.33**	**2.00**	**4.67**	**2.00**	**2.00**	**2.00**	**1.33**	**1.33**	**4.00**	**3.33**	**4.00**	**2.67**	**2.00**	**2.58**
	sEMA	10.00	14.67	10.00	8.00	9.33	8.67	4.00	5.33	11.33	10.00	6.67	8.00	9.33	6.00	9.33	8.71
	+MedBN	**2.00**	**2.00**	**2.00**	**2.00**	**4.00**	**2.00**	**0.00**	**0.00**	**0.00**	**0.00**	**4.00**	**2.00**	**2.00**	**2.00**	**0.00**	**1.60**
	mDIA	15.33	18.00	22.00	14.00	14.00	16.00	12.00	16.00	18.00	24.00	10.00	24.00	16.00	12.00	18.00	16.62
	+MedBN	**2.00**	**4.00**	**6.00**	**2.00**	**4.00**	**0.00**	**2.00**	**2.00**	**0.00**	**0.00**	**4.00**	**2.00**	**2.00**	**0.00**	**0.00**	**2.00**
ImageNet-C	TeBN	100.00	100.00	100.00	100.00	100.00	100.00	96.00	97.33	94.67	98.67	98.67	100.00	100.00	89.33	92.00	97.78
	+MedBN	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**1.33**	**0.00**	**0.00**	**4.00**	**0.00**	**0.00**	**0.00**	**0.36**
	TENT	89.33	84.00	88.00	96.00	94.67	96.00	92.00	96.00	94.67	96.00	90.67	100.00	94.67	80.00	80.00	91.47
	+MedBN	**0.00**	**1.33**	**0.00**	**0.00**	**0.00**	**0.00**	**1.33**	**0.00**	**0.00**	**0.00**	**0.00**	**4.00**	**0.00**	**0.00**	**0.00**	**0.44**
	ETA	100.00	100.00	100.00	96.00	100.00	96.00	90.67	92.00	88.00	98.67	96.00	100.00	97.33	77.33	85.33	94.49
	+MedBN	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**2.67**	**0.00**	**0.00**	**0.00**	**0.00**	**4.00**	**0.00**	**0.00**	**0.00**	**0.44**
	SAR	66.67	66.67	66.67	66.67	66.67	66.67	65.33	64.00	60.00	66.67	64.00	68.00	66.67	53.33	60.00	64.53
	+MedBN	**0.00**	**0.00**	**0.00**	**1.33**	**0.00**	**0.00**	**0.00**	**0.00**	**1.33**	**0.00**	**0.00**	**2.67**	**0.00**	**1.33**	**0.00**	**0.44**
	SoTTA	4.00	5.33	8.00	26.67	14.67	20.00	18.67	14.67	30.67	16.00	12.00	18.67	13.33	16.00	10.67	15.29
	+MedBN	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**1.33**	**2.67**	**4.00**	**0.00**	**0.00**	**4.00**	**0.00**	**0.00**	**0.00**	**0.80**
	sEMA	0.00	8.00	0.00	16.00	8.00	20.00	16.00	8.00	17.33	24.00	4.00	17.33	8.00	12.00	6.67	11.02
	+MedBN	**0.00**	**0.00**	**0.00**	**4.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.27**
	mDIA	24.00	22.67	32.00	37.33	26.67	36.00	36.00	25.33	45.33	40.00	26.67	57.33	28.00	21.33	24.00	32.18
	+MedBN	**8.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**4.00**	**0.00**	**0.00**	**0.00**	**0.00**	**4.00**	**0.00**	**1.07**

Table 16. Extended analysis of Error Rate (%) of indiscriminate and instant attack scenario over all types of corruption (detailed version of Table 2).

		Noise			Blur				Weather				Digital				
	Method	Gauss.	Shot	Impul.	Defoc.	Glass	Motion	Zoom	Snow	Frost	Fog	Brit.	Contr.	Elastic	Pixel.	JPEG	Avg.
CIFAR10-C	TeBN	35.74	34.37	45.60	23.88	37.95	27.86	20.62	29.76	26.83	40.62	20.82	30.97	32.95	23.65	33.70	31.02
	+MedBN	**26.60**	**25.15**	**35.01**	**17.23**	**27.51**	**20.37**	**14.75**	**21.00**	**18.36**	**28.19**	**14.24**	**20.22**	**24.93**	**16.82**	**24.74**	**22.34**
	TENT	32.20	30.72	40.39	23.05	34.92	25.91	20.62	26.91	25.64	28.47	20.66	26.90	31.42	23.93	30.14	28.13
	+MedBN	**24.12**	**22.96**	**32.31**	**15.78**	**27.08**	**17.99**	**14.59**	**19.07**	**17.05**	**20.22**	**13.87**	**18.04**	**23.88**	**15.23**	**22.30**	**20.30**
	ETA	30.67	30.05	38.84	22.44	34.10	25.15	20.76	26.46	25.42	27.99	20.50	26.09	30.65	22.67	29.46	27.42
	+MedBN	**23.42**	**21.93**	**31.02**	**15.04**	**26.37**	**17.62**	**14.20**	**18.37**	**16.93**	**20.13**	**13.60**	**17.78**	**23.12**	**14.89**	**22.70**	**19.81**
	SAR	32.24	30.30	39.30	22.18	33.52	25.44	20.08	25.71	25.18	29.38	20.16	27.57	30.20	22.33	29.85	27.56
	+MedBN	**23.36**	**21.73**	**30.87**	**14.96**	**25.14**	**17.45**	**13.84**	**18.20**	**16.61**	**20.24**	**13.57**	**18.25**	**22.86**	**15.01**	**21.90**	**19.60**
	SoTTA	25.49	23.51	33.12	15.12	26.87	17.30	13.53	19.09	17.34	20.23	14.27	17.56	23.70	15.23	23.62	20.40
	+MedBN	**21.37**	**19.78**	**29.50**	**11.15**	**23.23**	**13.67**	**10.01**	**15.00**	**14.15**	**14.90**	**10.66**	**13.06**	**19.17**	**12.14**	**19.52**	**16.49**
	sEMA	27.22	25.65	37.32	14.40	27.90	18.43	12.09	20.28	17.76	26.50	13.11	19.35	23.60	15.08	26.01	21.65
	+MedBN	**23.68**	**22.09**	**33.07**	**11.00**	**23.16**	**14.42**	**9.18**	**16.32**	**14.10**	**20.85**	**10.58**	**14.73**	**19.32**	**12.35**	**21.76**	**17.77**
	mDIA	38.78	37.39	50.60	16.82	34.20	23.57	14.23	22.99	19.96	38.50	13.61	36.23	27.04	18.00	27.43	27.96
	+MedBN	**26.26**	**25.26**	**34.70**	**12.03**	**24.90**	**15.97**	**9.60**	**16.60**	**13.69**	**22.20**	**9.86**	**20.53**	**19.72**	**13.25**	**21.37**	**19.06**
CIFAR100-C	TeBN	66.56	65.76	73.97	49.61	64.98	55.07	45.74	59.85	57.30	71.08	50.01	64.44	60.07	49.70	62.91	59.80
	+MedBN	**54.39**	**54.82**	**63.47**	**39.26**	**53.64**	**43.44**	**36.08**	**48.36**	**45.71**	**58.51**	**39.71**	**50.15**	**49.33**	**40.22**	**51.15**	**48.55**
	TENT	60.80	59.45	66.79	47.84	60.47	51.82	45.31	55.78	55.21	58.65	47.08	54.01	57.22	48.24	57.75	55.10
	+MedBN	**53.77**	**52.26**	**60.89**	**39.68**	**53.47**	**42.24**	**37.10**	**46.99**	**45.95**	**49.69**	**38.82**	**45.33**	**48.19**	**40.08**	**49.99**	**46.96**
	ETA	60.50	58.90	67.04	47.27	60.60	50.97	44.07	54.27	54.83	57.19	46.33	54.64	56.07	48.23	55.80	54.45
	+MedBN	**53.38**	**52.40**	**61.82**	**39.34**	**52.52**	**42.05**	**36.35**	**45.77**	**45.05**	**49.99**	**38.04**	**44.60**	**48.07**	**39.76**	**49.78**	**46.59**
	SAR	63.51	61.82	69.44	48.43	61.60	52.69	46.33	55.21	55.38	59.57	48.31	57.58	57.68	48.99	59.49	56.40
	+MedBN	**55.16**	**53.88**	**63.38**	**39.98**	**53.59**	**42.99**	**36.49**	**47.42**	**46.59**	**52.34**	**39.22**	**46.52**	**49.95**	**41.02**	**51.51**	**48.00**
	SoTTA	54.94	54.19	63.48	40.10	54.22	44.14	38.49	48.03	48.08	50.45	39.36	46.16	50.12	41.85	51.31	48.33
	+MedBN	**53.52**	**51.94**	**61.14**	**37.24**	**51.32**	**40.71**	**35.23**	**45.16**	**44.68**	**46.67**	**37.09**	**40.82**	**47.57**	**38.51**	**49.05**	**45.38**
	sEMA	53.50	52.80	63.03	37.82	51.65	41.65	34.07	46.85	43.91	56.24	37.94	47.58	47.00	37.32	52.04	46.89
	+MedBN	**50.87**	**51.05**	**59.57**	**33.60**	**48.07**	**37.50**	**30.90**	**42.95**	**40.52**	**52.24**	**34.01**	**42.28**	**43.89**	**34.51**	**48.30**	**43.35**
	mDIA	64.90	63.06	72.89	43.07	59.61	49.40	38.27	50.80	48.32	72.11	40.45	77.57	52.47	44.45	54.01	55.43
	+MedBN	**59.93**	**57.65**	**67.90**	**36.14**	**52.96**	**40.41**	**32.92**	**44.47**	**41.89**	**56.29**	**34.97**	**57.40**	**45.85**	**39.03**	**49.77**	**47.84**
ImageNet-C	TeBN	96.96	94.46	96.52	93.47	93.50	85.97	75.93	80.93	83.13	69.50	49.99	96.58	70.21	63.28	71.45	81.46
	+MedBN	**86.31**	**85.25**	**85.36**	**83.69**	**84.72**	**75.56**	**64.33**	**68.46**	**69.99**	**52.80**	**37.25**	**77.99**	**60.00**	**52.70**	**61.63**	**69.74**
	TENT	86.54	84.05	86.24	86.71	86.08	77.19	66.77	69.92	74.90	60.79	46.15	89.35	60.77	55.25	61.64	72.82
	+MedBN	**84.87**	**83.07**	**84.07**	**82.48**	**83.39**	**73.50**	**62.07**	**66.89**	**68.89**	**51.07**	**35.86**	**77.04**	**57.18**	**50.36**	**59.44**	**68.01**
	ETA	90.81	88.00	90.76	88.37	87.61	77.50	67.86	71.07	75.13	61.63	46.11	87.50	61.45	55.63	62.79	74.15
	+MedBN	**85.90**	**84.41**	**84.86**	**82.99**	**83.92**	**74.37**	**62.18**	**66.47**	**68.69**	**51.73**	**35.98**	**77.03**	**57.88**	**51.08**	**59.63**	**68.47**
	SAR	95.42	92.74	94.83	92.15	91.16	82.32	72.15	75.73	77.79	63.24	46.52	92.92	64.96	58.24	65.93	77.74
	+MedBN	**86.00**	**85.09**	**85.73**	**84.02**	**84.79**	**74.92**	**63.99**	**68.24**	**69.81**	**52.65**	**36.76**	**78.36**	**59.10**	**52.29**	**61.29**	**69.54**
	SoTTA	78.69	78.14	79.30	80.96	80.99	70.53	60.11	63.37	68.14	52.59	39.32	77.96	55.30	49.95	55.41	66.05
	+MedBN	**80.41**	**81.56**	**80.16**	**78.65**	**79.72**	**69.13**	**57.18**	**61.41**	**65.41**	**48.27**	**34.53**	**72.69**	**53.19**	**47.10**	**53.84**	**64.22**
	sEMA	86.24	86.50	85.62	87.26	87.36	78.84	68.13	72.04	74.37	59.49	41.89	86.65	62.88	55.81	65.12	73.21
	+MedBN	**86.44**	**87.33**	**85.73**	**84.87**	**84.66**	**75.85**	**64.19**	**68.71**	**69.62**	**53.84**	**36.90**	**81.81**	**58.58**	**53.39**	**61.42**	**70.22**
	mDIA	93.77	92.77	93.22	87.22	87.50	80.57	73.80	77.39	76.65	67.41	43.81	90.10	65.74	61.76	67.51	77.28
	+MedBN	**87.92**	**86.90**	**86.56**	**79.84**	**81.54**	**73.44**	**64.99**	**68.14**	**67.81**	**52.44**	**36.58**	**78.04**	**58.99**	**55.19**	**60.15**	**69.24**

K. Comprehensive Results of Cumulative Attack Scenario

We provide detailed results of cumulative and targeted attack scenarios in Table 17 and cumulative and indiscriminate attack scenarios in Table 18 across all types of corruptions in the TTA benchmark datasets. The averaged results across all trials are presented in Table 3. Within the scope of the cumulative attack scenario, we use EATA instead of ETA. EATA includes a Fisher regularizer that limits substantial change to important parameters, offering benefits in the cumulative scenario.

Table 17. Attack Success Rate (%) of the targeted and cumulative attack scenario over all types of corruptions (full version of Table 3).

		Noise			Blur				Weather				Digital				
	Method	Gauss.	Shot	Impul.	Defoc.	Glass	Motion	Zoom	Snow	Frost	Fog	Brit.	Contr.	Elastic	Pixel.	JPEG	Avg.
CIFAR10-C	TeBN	82.67	90.00	90.67	76.00	94.67	82.67	78.67	82.67	84.67	88.67	74.67	92.00	87.33	76.67	78.67	84.04
	+MedBN	**26.00**	**24.00**	**22.00**	**16.00**	**28.67**	**16.00**	**15.33**	**20.67**	**16.00**	**16.67**	**12.00**	**17.33**	**20.67**	**10.67**	**26.00**	**19.20**
	TENT	70.67	76.67	76.67	66.67	84.67	72.00	70.67	70.00	72.00	76.00	62.00	84.67	81.33	74.00	74.67	74.18
	+MedBN	**22.67**	**24.67**	**19.33**	**15.33**	**27.33**	**15.33**	**16.67**	**20.00**	**14.67**	**18.00**	**12.67**	**19.33**	**20.00**	**14.00**	**22.00**	**18.80**
	EATA	70.67	80.00	81.33	74.00	85.33	74.67	74.67	72.00	70.67	77.33	67.33	81.33	80.00	68.00	78.67	75.73
	+MedBN	**26.67**	**23.33**	**20.67**	**18.00**	**30.00**	**22.00**	**19.33**	**18.00**	**15.33**	**20.00**	**16.67**	**20.00**	**24.67**	**18.00**	**22.67**	**21.02**
	SAR	72.67	78.00	80.00	70.00	82.00	74.00	78.00	72.67	76.00	83.33	68.67	86.00	84.67	73.33	72.67	76.80
	+MedBN	**25.33**	**23.33**	**20.00**	**15.33**	**26.00**	**16.67**	**10.00**	**20.00**	**14.67**	**14.67**	**15.33**	**18.00**	**22.67**	**15.33**	**24.67**	**18.80**
	SoTTA	24.67	22.67	26.00	16.00	24.67	18.67	18.00	23.33	16.00	20.67	16.67	18.67	24.67	22.67	24.00	21.16
	+MedBN	**10.67**	**16.00**	**12.00**	**6.67**	**8.00**	**3.33**	**6.00**	**11.33**	**6.00**	**5.33**	**6.00**	**8.00**	**10.67**	**7.33**	**14.00**	**8.76**
	sEMA	26.67	23.33	18.67	12.00	26.67	14.00	12.67	12.67	10.00	12.00	8.67	14.67	18.00	10.67	21.33	16.13
	+MedBN	**11.33**	**14.00**	**14.00**	**2.00**	**13.33**	**5.33**	**4.00**	**8.00**	**6.00**	**8.00**	**4.00**	**8.00**	**8.00**	**2.00**	**14.00**	**8.13**
	mDIA	44.00	34.67	52.67	24.00	52.00	30.00	26.00	24.67	20.00	34.67	22.67	36.00	42.00	34.00	34.00	34.09
	+MedBN	**12.00**	**14.00**	**16.00**	**2.00**	**10.00**	**6.00**	**4.00**	**8.00**	**4.00**	**10.00**	**4.00**	**8.00**	**12.00**	**4.00**	**19.33**	**8.89**
CIFAR100-C	TeBN	96.00	96.00	98.00	76.67	90.00	87.33	84.67	98.67	97.33	99.33	93.33	98.00	88.00	84.00	83.33	91.38
	+MedBN	**2.00**	**2.00**	**2.00**	**2.00**	**6.00**	**2.00**	**2.00**	**2.00**	**2.00**	**4.00**	**4.00**	**2.67**	**4.00**	**2.00**	**2.00**	**2.71**
	TENT	71.33	75.33	68.67	74.00	72.00	79.33	68.00	77.33	77.33	78.67	74.00	90.00	74.00	72.67	63.33	74.40
	+MedBN	**2.00**	**2.00**	**2.67**	**4.67**	**4.00**	**4.00**	**4.67**	**3.33**	**2.00**	**4.67**	**4.67**	**4.00**	**4.67**	**2.67**	**2.00**	**3.47**
	EATA	81.33	78.00	80.00	72.67	75.33	73.33	66.00	78.67	82.00	83.33	68.67	89.33	70.00	68.67	63.33	75.38
	+MedBN	**2.67**	**2.00**	**0.00**	**3.33**	**4.67**	**2.00**	**4.00**	**4.00**	**1.33**	**2.00**	**3.33**	**3.33**	**3.33**	**4.00**	**1.33**	**2.76**
	SAR	84.00	86.00	86.00	82.67	75.33	78.00	73.33	84.67	86.00	92.67	75.33	92.67	77.33	81.33	72.67	81.87
	+MedBN	**2.67**	**3.33**	**2.00**	**5.33**	**5.33**	**2.67**	**3.33**	**1.33**	**2.00**	**6.00**	**3.33**	**2.67**	**3.33**	**3.33**	**2.00**	**3.24**
	SoTTA	6.67	8.67	10.00	6.67	10.67	6.00	6.67	6.67	9.33	11.33	4.00	9.33	8.00	4.67	6.00	7.64
	+MedBN	**2.00**	**2.67**	**2.00**	**2.67**	**4.00**	**4.00**	**2.00**	**2.67**	**1.33**	**2.00**	**3.33**	**3.33**	**4.67**	**2.00**	**2.00**	**2.71**
	sEMA	10.67	14.67	10.00	4.00	10.00	7.33	6.00	6.00	8.67	8.00	6.67	6.67	10.67	6.00	8.67	8.27
	+MedBN	**2.00**	**2.00**	**2.00**	**0.67**	**2.00**	**0.00**	**0.00**	**0.00**	**2.00**	**2.00**	**2.00**	**2.00**	**2.00**	**0.00**	**0.00**	**1.24**
	mDIA	15.33	18.00	22.00	13.33	14.00	17.33	12.00	16.00	18.00	24.00	10.00	24.00	16.00	12.00	17.33	16.62
	+MedBN	**2.00**	**4.00**	**6.00**	**2.00**	**4.00**	**0.00**	**2.00**	**3.33**	**0.00**	**0.00**	**4.00**	**4.00**	**2.00**	**0.00**	**0.00**	**2.22**
ImageNet-C	TeBN	100.00	100.00	100.00	100.00	100.00	100.00	96.00	96.00	94.67	100.00	97.33	100.00	100.00	88.00	92.00	97.60
	+MedBN	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**4.00**	**0.00**	**0.00**	**0.00**	**0.27**
	TENT	84.00	78.67	81.33	88.00	84.00	84.00	85.33	81.33	84.00	94.67	84.00	97.33	94.67	60.00	73.33	83.64
	+MedBN	**0.00**	**0.00**	**0.00**	**0.00**	**1.33**	**0.00**	**1.33**	**0.00**	**0.00**	**0.00**	**0.00**	**4.00**	**0.00**	**0.00**	**0.00**	**0.44**
	EATA	100.00	98.67	100.00	96.00	98.67	92.00	89.33	93.33	88.00	93.33	88.00	100.00	96.00	69.33	82.67	92.36
	+MedBN	**0.00**	**0.00**	**0.00**	**0.00**	**1.33**	**0.00**	**0.00**	**0.00**	**1.33**	**0.00**	**0.00**	**4.00**	**0.00**	**0.00**	**0.00**	**0.44**
	SAR	100.00	100.00	100.00	100.00	100.00	100.00	96.00	97.33	92.00	100.00	96.00	100.00	100.00	80.00	92.00	96.89
	+MedBN	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
	SoTTA	5.33	8.00	6.67	21.33	22.67	18.67	17.33	13.33	28.00	16.00	12.00	21.33	10.67	21.33	13.33	15.73
	+MedBN	**0.00**	**0.00**	**0.00**	**2.67**	**0.00**	**0.00**	**1.33**	**2.67**	**2.67**	**0.00**	**0.00**	**2.67**	**0.00**	**0.00**	**0.00**	**0.80**
	sEMA	8.00	17.33	12.00	18.67	8.00	21.33	16.00	8.00	16.00	24.00	4.00	8.00	8.00	13.33	14.67	13.16
	+MedBN	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
	mDIA	33.33	26.67	40.00	32.00	33.33	33.33	40.00	16.00	44.00	38.67	20.00	57.33	21.33	17.33	20.00	31.56
	+MedBN	**8.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**4.00**	**0.00**	**0.00**	**0.00**	**0.00**	**4.00**	**0.00**	**1.07**

Table 18. Error Rate (%) of the indiscriminate and cumulative attack scenario over all types of corruptions (full version of Table 3).

		Noise			Blur				Weather				Digital				
	Method	Gauss.	Shot	Impul.	Defoc.	Glass	Motion	Zoom	Snow	Frost	Fog	Brit.	Contr.	Elastic	Pixel.	JPEG	Avg.
CIFAR10-C	TeBN	35.65	34.44	45.67	23.77	37.92	27.74	20.58	29.86	26.85	40.66	20.77	31.01	32.96	23.65	33.69	35.30
	+MedBN	**26.76**	**25.07**	**35.02**	**17.18**	**27.58**	**20.34**	**14.70**	**20.88**	**18.40**	**28.25**	**14.17**	**20.25**	**25.00**	**16.74**	**24.76**	**27.22**
	TENT	34.79	32.67	47.80	25.72	38.49	29.37	23.07	29.47	28.87	32.61	23.32	29.33	34.45	25.52	34.77	35.70
	+MedBN	**24.86**	**23.29**	**34.99**	**15.10**	**27.88**	**17.85**	**13.80**	**18.31**	**17.71**	**22.18**	**13.98**	**19.67**	**24.33**	**15.46**	**22.99**	**25.84**
	EATA	35.84	33.09	46.68	24.61	39.32	27.88	22.51	29.11	27.64	31.70	23.47	28.50	35.94	25.10	34.02	35.30
	+MedBN	**26.52**	**24.66**	**35.86**	**15.72**	**27.53**	**20.50**	**17.15**	**19.23**	**17.31**	**21.63**	**14.96**	**22.56**	**25.34**	**16.00**	**24.16**	**26.84**
	SAR	30.83	29.07	38.36	21.62	33.22	24.34	20.32	25.23	24.15	27.22	19.87	25.09	29.88	21.57	28.77	31.25
	+MedBN	**23.00**	**21.20**	**30.92**	**14.52**	**24.93**	**16.82**	**13.27**	**17.56**	**16.17**	**19.33**	**13.10**	**17.38**	**22.72**	**14.55**	**21.88**	**24.29**
	SoTTA	25.35	24.59	34.17	15.47	28.35	17.77	14.32	19.52	17.75	21.34	14.69	19.55	23.92	16.27	23.71	26.10
	+MedBN	**22.52**	**20.42**	**30.98**	**11.61**	**23.32**	**13.95**	**10.94**	**16.18**	**14.52**	**15.87**	**11.04**	**15.85**	**19.84**	**13.28**	**20.00**	**22.52**
	sEMA	29.10	27.45	39.19	17.51	29.48	20.45	14.63	22.25	18.96	32.03	15.15	23.96	25.52	18.20	27.08	28.79
	+MedBN	**25.35**	**23.49**	**33.90**	**14.97**	**25.62**	**18.51**	**13.31**	**19.12**	**16.77**	**26.26**	**12.47**	**18.20**	**23.09**	**14.97**	**23.64**	**25.62**
	mDIA	38.67	37.37	50.68	16.76	34.20	23.58	14.22	23.04	20.04	38.55	13.62	36.19	26.99	18.03	27.40	32.05
	+MedBN	**26.32**	**25.20**	**34.67**	**12.00**	**24.84**	**16.00**	**9.56**	**16.62**	**13.64**	**22.11**	**9.87**	**20.54**	**19.69**	**13.24**	**21.36**	**23.96**
CIFAR100-C	TeBN	59.67	58.49	61.26	42.36	55.96	49.35	39.85	51.24	48.82	59.46	45.40	59.62	50.70	45.82	52.61	52.04
	+MedBN	**43.64**	**45.12**	**52.97**	**30.25**	**44.35**	**32.25**	**28.77**	**37.01**	**33.22**	**49.61**	**31.11**	**40.82**	**40.94**	**30.05**	**38.15**	**38.55**
	TENT	57.07	54.94	63.90	41.09	56.35	44.32	40.18	51.72	49.54	54.25	42.06	53.19	50.77	41.94	49.65	50.06
	+MedBN	**43.82**	**40.33**	**47.09**	**30.12**	**46.18**	**33.86**	**29.49**	**37.36**	**39.50**	**42.25**	**28.74**	**37.73**	**36.41**	**33.17**	**38.60**	**37.64**
	EATA	52.49	50.88	60.00	39.29	53.75	44.07	39.09	46.35	47.17	51.90	39.11	48.40	48.91	40.27	51.11	47.52
	+MedBN	**40.62**	**41.12**	**48.30**	**31.75**	**41.37**	**31.50**	**26.92**	**35.45**	**34.81**	**40.87**	**28.43**	**38.06**	**39.45**	**31.39**	**40.91**	**36.73**
	SAR	51.39	50.32	55.06	40.06	51.30	41.36	36.46	43.92	47.64	50.51	39.70	45.07	46.44	39.18	47.05	45.70
	+MedBN	**42.29**	**41.47**	**47.93**	**27.83**	**40.45**	**33.24**	**25.26**	**35.55**	**35.15**	**42.26**	**29.75**	**39.21**	**36.61**	**29.75**	**41.14**	**36.53**
	SoTTA	44.10	42.92	50.09	28.53	43.43	34.78	29.47	41.13	38.85	42.48	27.67	40.34	38.16	32.75	38.37	38.21
	+MedBN	**41.07**	**38.67**	**45.63**	**26.71**	**37.17**	**29.54**	**25.07**	**33.97**	**34.90**	**33.85**	**27.05**	**32.43**	**35.52**	**29.23**	**36.71**	**33.84**
	sEMA	41.95	44.95	47.57	28.43	46.33	35.59	28.93	36.75	37.20	46.40	29.37	42.28	37.73	28.75	41.43	38.24
	+MedBN	**36.40**	**39.76**	**46.11**	**27.47**	**40.11**	**29.85**	**25.90**	**33.25**	**31.49**	**43.04**	**27.56**	**36.48**	**35.65**	**27.17**	**36.34**	**34.44**
	mDIA	52.12	52.81	60.54	37.40	50.35	41.74	29.54	40.60	39.53	64.64	29.67	72.37	43.61	37.01	41.84	46.25
	+MedBN	**45.33**	**42.47**	**53.33**	**27.42**	**41.30**	**30.27**	**25.97**	**32.96**	**31.94**	**43.51**	**24.66**	**50.98**	**32.61**	**30.66**	**35.82**	**36.62**
ImageNet-C	TeBN	97.10	94.53	96.50	93.42	93.36	86.07	76.11	80.95	83.20	69.33	50.19	96.64	69.92	63.12	71.59	81.47
	+MedBN	**86.18**	**85.34**	**85.22**	**83.63**	**84.72**	**75.52**	**64.36**	**68.40**	**69.96**	**52.86**	**37.36**	**77.96**	**59.88**	**52.73**	**61.42**	**69.70**
	TENT	85.81	85.02	85.58	86.80	86.71	77.39	65.66	71.18	74.67	58.31	43.58	90.76	60.86	53.75	60.09	72.41
	+MedBN	**84.72**	**83.66**	**83.88**	**82.21**	**83.55**	**74.88**	**61.96**	**66.28**	**68.89**	**50.84**	**36.43**	**76.70**	**57.30**	**50.40**	**59.96**	**68.11**
	EATA	95.64	92.59	94.10	90.44	91.20	80.29	68.27	74.44	76.41	61.84	44.06	91.86	62.01	54.25	63.00	76.03
	+MedBN	**85.45**	**84.16**	**84.89**	**82.79**	**83.64**	**74.59**	**61.71**	**66.32**	**68.85**	**51.13**	**35.81**	**76.21**	**57.08**	**50.96**	**59.61**	**68.21**
	SAR	96.27	93.21	95.61	91.90	90.68	81.10	70.97	74.60	77.28	62.92	46.45	93.22	64.24	57.85	65.52	77.46
	+MedBN	**86.36**	**85.60**	**85.58**	**83.55**	**84.67**	**75.50**	**63.93**	**68.05**	**70.22**	**52.67**	**36.93**	**78.27**	**58.96**	**52.05**	**61.13**	**69.56**
	SoTTA	81.20	80.61	81.10	81.95	83.69	71.89	61.56	65.08	68.78	53.28	39.38	77.87	56.13	50.47	56.33	67.29
	+MedBN	**82.59**	**81.93**	**81.49**	**78.26**	**79.83**	**69.16**	**57.51**	**61.71**	**64.97**	**48.00**	**34.70**	**71.45**	**53.00**	**46.96**	**53.94**	**64.37**
	sEMA	88.99	87.84	88.90	88.24	88.35	80.55	70.83	74.97	76.28	63.30	44.18	86.87	65.04	57.49	66.37	75.21
	+MedBN	**89.27**	**87.18**	**87.41**	**84.06**	**84.59**	**75.55**	**64.47**	**69.00**	**69.98**	**53.85**	**37.89**	**79.76**	**59.85**	**53.29**	**61.88**	**70.54**
	mDIA	93.83	92.73	93.24	87.04	87.59	80.50	73.73	77.39	76.62	67.27	43.72	90.02	65.67	61.64	67.53	77.24
	+MedBN	**87.90**	**86.92**	**86.53**	**79.83**	**81.59**	**73.44**	**65.02**	**68.13**	**67.81**	**52.41**	**36.44**	**78.04**	**58.90**	**55.15**	**60.13**	**69.22**

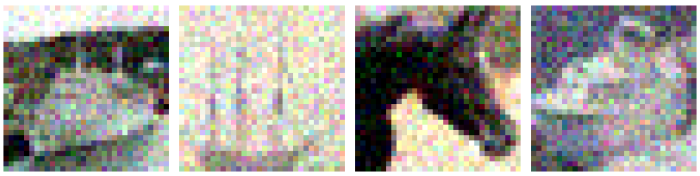
L. Error Rates without Attacks

To evaluate the performance of the model under a normal TTA setup, we utilize ER on benign samples without attacks in Table 19. It provides an understanding of how the model behaves in a non-adversarial environment, i.e., the model's baseline effectiveness, establishing a fundamental metric for comparison against scenarios involving attacks.

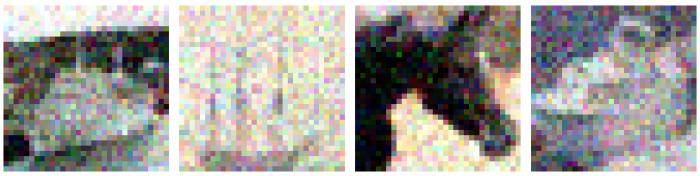
Table 19. Error Rate (%) on benign samples without attacks.

ER (%)	B / m	Normalization	Method						
			TeBN	TENT	ETA	SAR	SoTTA	sEMA	mDIA
CIFAR10-C	200 / 40 (20%)	BatchNorm	14.92	13.68	13.14	13.28	13.73	14.87	15.31
		Ours (MedBN)	15.19	14.12	13.67	13.35	14.06	15.14	15.20
CIFAR100-C	200 / 40 (20%)	BatchNorm	40.08	37.74	37.44	39.30	41.22	39.72	41.72
		Ours (MedBN)	40.77	39.66	39.62	41.32	42.26	40.47	41.79
ImageNet-C	200 / 20 (10%)	BatchNorm	66.62	61.08	59.13	62.13	60.87	68.35	66.62
		Ours (MedBN)	69.55	68.38	66.20	66.65	64.39	70.18	68.27

M. Examples of Malicious Samples



(a) Visualization of benign samples.



(b) Visualization of malicious samples ($\varepsilon = 8/255$).

Figure 10. Visualization of test samples from CIFAR10-C benchmark with Gaussian noise and severity level 5. Malicious samples are hardly distinguished from benign samples.