# Neuro-Modulated Hebbian Learning for Fully Test-Time Adaptation
# Supplemental Materials


Yushun Tang[1,2], Ce Zhang[1], Heng Xu[1], Shuoshuo Chen[1],
Jie Cheng[2], Luziwei Leng[2], Qinghai Guo[2*], Zhihai He[1,3*]

[1]Department of Electronic and Electrical Engineering, Southern University of Science and Technology, Shenzhen, China
[2]Advanced Computing and Storage Laboratory, Huawei Technologies Co., Ltd., Shenzhen, China
[3]Pengcheng Laboratory, Shenzhen, China

{tangys2022, zhangc2019, xuh2022, chenss2021}@mail.sustech.edu.cn
{chengjie8, lengluziwei, guoqinghai}@huawei.com, hezh@sustech.edu.cn


In this Supplemental Material, we provide more details and experimental results for further understanding of the proposed Neuro-Modulated Hebbian learning algorithm.

## 1. More Details on the Hebbian Learning

In this section, we study the convergence of the proposed soft Hebbian learning algorithm. We begin from a generative probabilistic model, as defined in Nessler *et al.* [6], where we denote the observed data as $\boldsymbol{x} = (x_1, x_2, ..., x_n)$ and the hidden cause $\vartheta = \{C_1, ..., C_k\}$. The true distribution $p_t(x)$ is then calculated as

$$p_t(\mathbf{x}) = \prod_{i=1}^{n} (\sum_{k=1}^{K} p(x_i|C_k)p(C_k)). \tag{1}$$

To estimate this true distribution, we design an approximation distribution given weights $\boldsymbol{w}$. That is, $q(\boldsymbol{x}|\boldsymbol{w}) = \sum_{k=1}^{K} q(\boldsymbol{x}|C_k)q_0(C_k)$, with $q(\boldsymbol{x}) = \prod_{i=1}^{n} q(x_i)$ and $\forall k$,

$$q(x_i|C_k) = q(x_i|C_k, w_{ik}) := e^{w_{ik} \cdot x_i}, \tag{2}$$

$$q_0(C_k) = q_0(C_k, w_{0k}) := e^{w_0 k}. \tag{3}$$

According to the above definition, we can find the optimal parameters $\boldsymbol{w}^*$ which minimize the KL-divergence of the generative model and the original input distribution. We then show it is possible to link the generative model to the activation function of a neural network, using the Hebbian rule defined in our main manuscript, the weights of the neural network will eventually converge to the optimal weights $\boldsymbol{w}^*$ that minimize the KL-divergence aforementioned.

To that end, according to Nessler *et al.* [6], it is shown that for each component $k$, the optimal parameter is proportional to the mean of the corresponding component of the input distribution, that is

$$\boldsymbol{w}_k^* = c \cdot \mu_{p_k}(\boldsymbol{x}), \tag{4}$$

where $\mu_{p_k}(\boldsymbol{x})$ is the mean distribution $p_k$ with $p_k = p(\boldsymbol{x}|C_k)$, and $c$ is a positive factor. Given the cause $C_k$, the approximated distribution $q(\boldsymbol{x}|C_k)$ is then

$$q(\boldsymbol{x}|C_k) = \prod_{i=1}^{n} e^{w_{ik}x_i} = e^{u_k}, \tag{5}$$

where $u_k = \sum_{i=1}^{n} w_{ik}x_i$. Using the Bayesian rule, we derive the posterior probability as

$$\begin{aligned} q(C_k|\boldsymbol{x}) &= \frac{q(\boldsymbol{x}|C_k)q_0(C_k)}{q(\boldsymbol{x})} \\ &= \frac{e^{u_k + w_{0k}}}{\sum_l^K e^{u_l + w_{0l}}}. \end{aligned} \tag{6}$$

It can be seen from (6) that the probabilistic model has a neural interpretation, where $q(C_k|\boldsymbol{x})$ can be described as the output neuron's exponential activation divided by its layer's total output, with $w_{0k}$ as the bias term. We show that a soft Hebbian learning rule defined in the main manuscript has the equilibrium point $w_{ik} = w_{ik}^*$ with $c = \frac{\sqrt{R}}{\|\mu_{p_k}(\boldsymbol{x})\|}$, where $R$ is the $L2$-norm of the weight vector $\boldsymbol{w_k}$ defined in the main manuscript. To see this, based on the plasticity rule as we defined in the main manuscript, we have

$$\begin{aligned} E[\Delta w_{ik}] =& \eta \int_{\boldsymbol{x}} y_k(\boldsymbol{x})(Rx_i - u_k(\boldsymbol{x})w_{ik})p(\boldsymbol{x})d\boldsymbol{x} \\ =& \eta \int_{\boldsymbol{x}} y_k(\boldsymbol{x})(Rx_i - u_k(\boldsymbol{x})w_{ik}) \\ & \left(\sum_{l=1}^{K} p_l(\boldsymbol{x})p(C_l)\right) d\boldsymbol{x} \end{aligned}$$

[*] Corresponding authors.

$$
\begin{aligned}
=&\eta(\sum_{l=1}^{K}\int_{\boldsymbol{x}} Rx_i y_k(\boldsymbol{x})p_l(\boldsymbol{x})p(C_l)d\boldsymbol{x} \\
&-\sum_{l=1}^{K}\boldsymbol{w}_k w_{ik}\int_{\boldsymbol{x}}\boldsymbol{x}y_k(\boldsymbol{x})p_l(\boldsymbol{x})p(C_l)d\boldsymbol{x}) \\
=&\eta(\int_{\boldsymbol{x}} Rx_i y_k(\boldsymbol{x})p_k(\boldsymbol{x})p(C_k)d\boldsymbol{x} \\
&-\boldsymbol{w}_k w_{ik}\int_{\boldsymbol{x}}\boldsymbol{x}y_k(\boldsymbol{x})p_k(\boldsymbol{x})p(C_k)d\boldsymbol{x}),
\end{aligned} \tag{7}
$$

where we assume the overlap of the support of the measure $p_l$ and $p_k$ can be neglected for $l \neq k$, and $p_l(\boldsymbol{x}) = p(\boldsymbol{x}|C_l)$. The first term can be written as

$$
\int_{\boldsymbol{x}} Rx_i y_k(\boldsymbol{x})p_k(\boldsymbol{x})p(C_k)d\boldsymbol{x} = I_k \cdot R \cdot \mu_{p_k}(x_i), \tag{8}
$$

where $I_k$ is the normalization constant. The second term can be written as

$$
\begin{aligned}
&\boldsymbol{w}_k w_{ik}\int_{\boldsymbol{x}}\boldsymbol{x}y_k(\boldsymbol{x})p_k(\boldsymbol{x})p(C_k)d\boldsymbol{x} \\
&= I_k \cdot \boldsymbol{w}_k w_{ik}\mu_{p_k}(\boldsymbol{x}) \\
&= I_k \cdot R \cdot \mu_{p_k}(x_i).
\end{aligned} \tag{9}
$$

This leads to the conclusion of $E[\Delta w_{ik}] = 0$ given the optimal $\boldsymbol{w}$ defined in (4). The convergence follows the proof we showed in the main manuscript.

## 2. More Implementation Details

For the soft Hebbian layer, the weights are initialized with the source model. The norm $R$ is setting to 1. The learning rate $\eta$ is setting to $10^{-3}$ for the Hebbian layer and $5 \times 10^{-4}$ for the Neuro-Modulation layer. All the results are the average of three times running with different random seeds based on *Pytorch*. All models are tested on a single NVIDIA Tesla A100 GPU. The algorithm pseudo-code is shown in Algorithm 1.

**Algorithm 1** Pseudo code of the proposed algorithm.

**Input**: Source pre-trained model $\boldsymbol{\Gamma}_\theta^s$; target dataset $\mathcal{X}_t$.
**Output**: The prediction of target samples.

1: Initialize the testing model $\boldsymbol{\Gamma}_\theta^t$ with source pre-trained model $\boldsymbol{\Gamma}_\theta^s$ parameter weights;
2: **for** batch $x_t$ **in** $\mathcal{X}_t$ **do**
3: Update the Hebbian layer of $\boldsymbol{\Gamma}_\theta^t$ with Hebbian learning rule: $\Delta w_{ik} = \eta y_k(Rx_i - u_k w_{ik})$;
4: Update the Neuro-Modulation layer of $\boldsymbol{\Gamma}_\theta^t$ with loss function $H$.
5: Output = $\boldsymbol{\Gamma}_\theta^t(x_t)$
6: **end for**

| Corruption Type | Abbreviation |
|---|---|
| Gaussian Noise | gaus |
| Shot Noise | shot |
| Impulse Noise | impul |
| Defocus Blur | defcs |
| Glass Blur | gls |
| Motion Blur | mtn |
| Zoom Blur | zm |
| Snow | snw |
| Frost | frst |
| Fog | fg |
| Brightness | brt |
| Contrast | cnt |
| Elastic | els |
| Pixelate | px |
| JPEG Compression | jpg |

Table 1. Abbreviations of the 15 corruption types in CIFAR-10/100C and ImageNet-C datasets.

## 3. More Experimental Results

Table 1 summarized the abbreviations of the different corruption types. The performance results or classification errors of our algorithm in comparisons with existing methods for different levels of corruption are shown in Table 4, 5 and 6. In the main body of the paper, only the results for Level 5 corruption are reported. Our method outperforms almost all the comparison methods including TTT [9], NORM [8], TENT [10], and DUA [5] in all corruption severity levels and all corruption types. In addition, Figure 1 shows the prediction comparison in some target samples in MNIST [4], MNIST-M [1] and USPS [2] for an SVHN [7] pre-trained source model. Our method recognizes digits right but others are wrong in some cases. Figure 2 shows more feature map visualization after the first convolution layer obtained by different learning methods. We can see that the Hebbian learning is able to generate feature maps that are as good as those from the Oracle supervised by labels.

The parameter amount of our new model is the same as the source model. The testing time cost comparison is shown in Table 2, in which NHL is about 50% slower than TENT, but much faster (10-200 times) than DUA and TTT. We also investigate the effect of the temperature-scaling hyper-parameter $\tau$ in Eq. (6) in the main manuscript. As shown in Table 3, it does not affect the overall performance much in a range.

## References


[1] Yaroslav Ganin and Victor Lempitsky. Unsupervised domain adaptation by backpropagation. In *ICML*, pages 1180–1189. PMLR, 2015. 2

Table 2. Testing time cost comparison for Gaussian corruption of **CIFAR-10C** [3].

| Methods | Time |
|---|---|
| TENT [10] | 7.67s |
| DUA [5] | 187.32s |
| TTT [9] | 2955.04s |
| Ours | 11.60s |

Table 3. Top-1 average Classification Error (%) in **ImageNet-C** at the highest severity (Level 5).

| $\tau$ | Avg. Error |
|---|---|
| 1 | 61.85 |
| 1/2 | 60.82 |
| 1/5 | 60.40 |
| 1/10 | 60.39 |
| 1/15 | 60.36 |
| 1/20 | 60.51 |


[2] Jonathan J. Hull. A database for handwritten text recognition research. *IEEE TPAMI*, 16(5):550–554, 1994. 2

[3] Alex Krizhevsky, Geoffrey Hinton, et al. Learning multiple layers of features from tiny images. *Master's thesis, University of Tront*, 2009. 3, 4

[4] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998. 2

[5] M Jehanzeb Mirza, Jakub Micorek, Horst Possegger, and Horst Bischof. The norm must go on: Dynamic unsupervised domain adaptation by normalization. In *CVPR*, pages 14765–14775, 2022. 2, 3

[6] Bernhard Nessler, Michael Pfeiffer, Lars Buesing, and Wolfgang Maass. Bayesian computation emerges in generic cortical microcircuits through spike-timing-dependent plasticity. *PLoS Computational Biology*, 9(4):e1003037, 2013. 1

[7] Yuval Netzer, Tao Wang, Adam Coates, Alessandro Bissacco, Bo Wu, and Andrew Y Ng. Reading digits in natural images with unsupervised feature learning. 2011. 2

[8] Steffen Schneider, Evgenia Rusak, Luisa Eck, Oliver Bringmann, Wieland Brendel, and Matthias Bethge. Improving robustness against common corruptions by covariate shift adaptation. In *NeurIPS*, volume 33, pages 11539–11551, 2020. 2

[9] Yu Sun, Xiaolong Wang, Zhuang Liu, John Miller, Alexei A. Efros, and Moritz Hardt. Test-time training with self-supervision for generalization under distribution shifts. In *ICML*, volume 119, pages 9229–9248. PMLR, 2020. 2, 3

[10] Dequan Wang, Evan Shelhamer, Shaoteng Liu, Bruno Olshausen, and Trevor Darrell. Tent: Fully test-time adaptation by entropy minimization. In *ICLR*, 2020. 2, 3


Figure 1. Prediction comparison in some target samples for an SVHN source model. The image with a red border means predicting wrong and green means predicting right. It shows that our NML method recognizes digits more accurately than NORM and TENT.

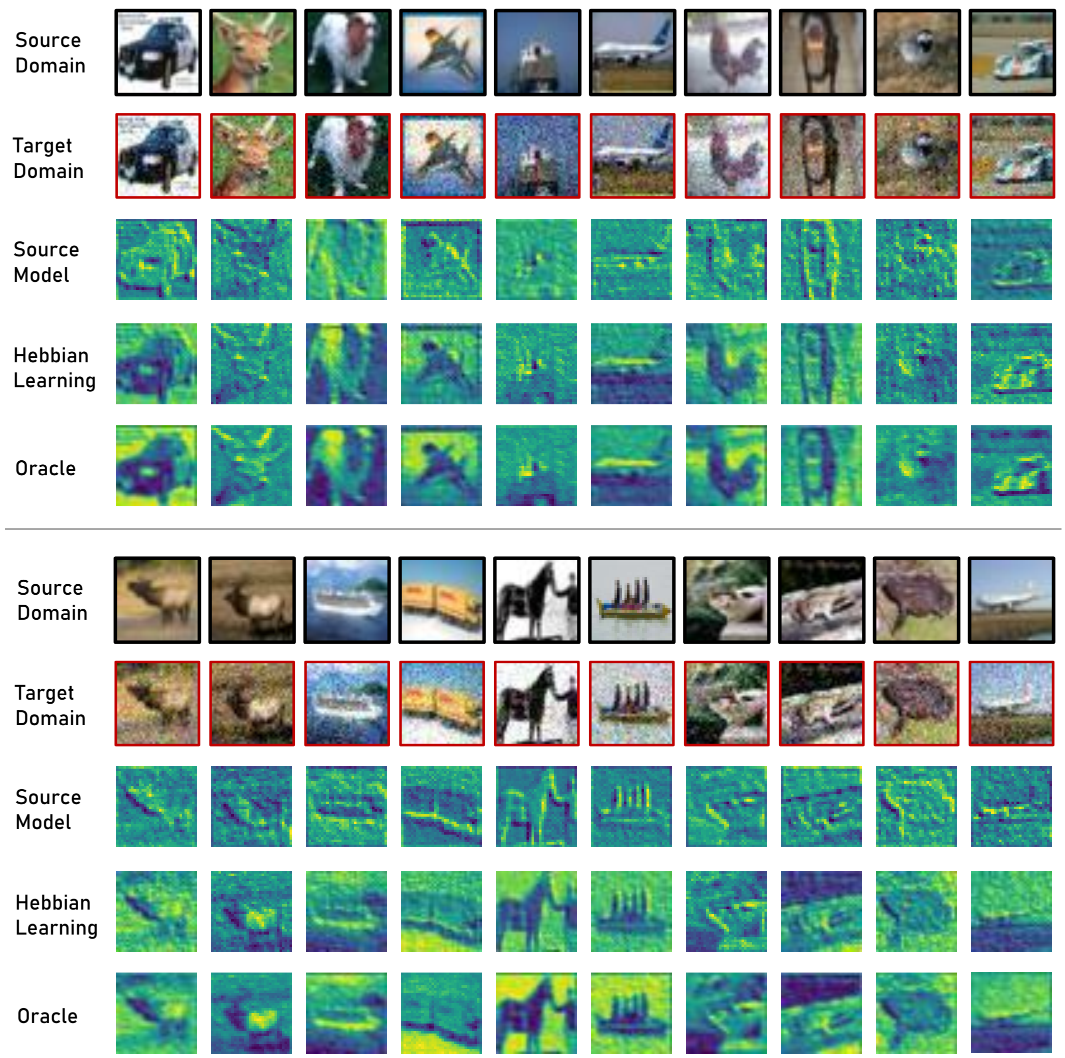


Figure 2. More feature map visualization of CIFAR-10C [3] after the first convolution layer obtained by different learning method. We can see that the unsupervised Hebbian learning is able to generate feature maps which are as good as those from the supervised learning ("Oracle"). It should be noted that the first row only shows the corresponding original image of the testing set but not the training set for the source model.

Table 4. Error (%) for each corruption in **CIFAR-10C** severity (Level 1–4) is reported. For TENT and DUA, we use the **ResNet-26** (top), **WRN-28-10** (middle) and **WRN-40-2** (bottom) from their official implementation. Lowest error is highlighted for each corruption.

| Methods | gaus | shot | impul | defcs | gls | mtn | zm | snw | frst | fg | brt | cnt | els | px | jpg | Avg. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Level 4** | | | | | | | | | | | | | | | | |
| Source | 63.9 | 53.7 | 57.0 | 28.9 | 58.9 | 32.4 | 38.1 | 25.9 | 33.9 | 17.5 | 10.4 | 33.7 | 26.7 | 40.7 | 27.2 | 36.6 |
| TTT | 41.5 | 35.4 | 39.8 | 15.0 | 47.8 | 19.1 | 18.4 | **20.1** | 24.0 | 13.5 | **10.0** | **14.1** | **17.7** | 29.4 | **24.5** | 24.7 |
| NORM | 40.7 | 37.4 | 43.2 | 16.7 | 47.4 | 21.8 | 20.2 | 29.9 | 30.3 | 19.0 | 16.1 | 20.5 | 26.5 | 26.6 | 35.7 | 28.8 |
| TENT | 35.3 | 32.6 | 39.0 | 14.9 | 44.1 | 20.1 | 17.8 | 26.8 | 27.8 | 17.1 | 13.8 | 21.3 | 23.9 | 22.5 | 30.4 | 25.8 |
| DUA | 31.0 | 27.6 | 35.8 | 13.2 | 40.7 | 20.3 | 15.4 | 22.2 | **20.6** | **12.7** | 10.1 | 14.8 | 20.5 | **18.6** | 24.6 | 21.9 |
| Ours | **28.9** | **26.3** | **32.1** | **13.1** | **40.6** | 16.9 | 15.7 | 22.5 | 22.0 | 13.7 | 12.4 | 16.6 | 20.9 | 18.9 | 24.7 | **21.7** |
| Source | 67.4 | 54.7 | 59.9 | 22.6 | 56.8 | 25.2 | 29.7 | 19.5 | 29.5 | 10.4 | 7.2 | 16.4 | 21.1 | 39.7 | 25.9 | 32.4 |
| NORM | 24.9 | 21.0 | 29.5 | 8.9 | 34.5 | 11.6 | 9.7 | 16.4 | 14.5 | 8.7 | 7.1 | 9.0 | 15.6 | 13.9 | 23.9 | 16.6 |
| TENT | 22.2 | 18.1 | 26.0 | 8.5 | 31.0 | 11.1 | 9.6 | 15.8 | 13.9 | 8.5 | 7.2 | 8.0 | 14.8 | 12.4 | 21.3 | 15.2 |
| DUA | 24.4 | 20.8 | 28.9 | 9.3 | 35.6 | 12.3 | 10.5 | 16.4 | 13.7 | **8.0** | **6.4** | 8.8 | 15.1 | 13.7 | 23.0 | 16.5 |
| Ours | **21.2** | **17.6** | **24.3** | **8.0** | **30.5** | **10.9** | **9.5** | **15.1** | **13.1** | 8.4 | 6.6 | **7.7** | **14.5** | **12.3** | **20.7** | **14.7** |
| Source | 24.1 | 17.1 | 16.4 | 6.6 | 23.5 | 8.4 | 7.4 | 12.2 | 11.5 | 8.3 | 6.2 | 9.2 | 10.6 | 19.4 | 13.1 | 12.9 |
| TENT | 13.8 | 11.7 | 14.3 | 6.7 | 18.6 | 8.2 | 7.1 | 10.6 | 9.7 | 7.5 | 6.1 | 8.4 | 10.9 | 8.5 | 13.2 | 10.3 |
| DUA | 13.7 | 11.8 | 13.5 | **5.9** | 18.3 | 7.6 | **6.6** | 10.3 | 9.0 | 7.4 | **5.8** | 7.2 | **9.9** | 9.3 | 13.0 | 10.0 |
| Ours | **12.1** | **10.2** | **12.4** | 6.2 | **15.9** | **7.5** | 7.1 | **9.9** | **8.5** | **7.1** | **5.8** | **6.7** | 10.6 | **7.8** | **12.2** | **9.3** |
| **Level 3** | | | | | | | | | | | | | | | | |
| Source | 58.0 | 47.5 | 38.5 | 17.7 | 46.2 | 32.8 | 30.6 | 22.7 | 31.8 | 12.6 | 9.5 | 19.3 | 20.7 | 23.7 | 24.7 | 29.1 |
| TTT | 37.2 | 31.6 | 28.6 | 11.5 | 35.8 | 19.1 | 15.8 | **17.8** | 23.3 | **11.0** | **9.1** | **11.6** | **14.3** | 18.9 | **22.3** | 20.5 |
| NORM | 37.8 | 35.1 | 34.7 | 14.1 | 38.2 | 21.7 | 18.2 | 27.5 | 29.0 | 16.6 | 15.2 | 18.6 | 19.6 | 21.1 | 33.3 | 25.4 |
| TENT | 33.1 | 29.7 | 30.7 | 12.4 | 35.1 | 20.3 | 15.9 | 24.3 | 25.7 | 15.2 | 13.0 | 18.1 | 17.5 | 17.5 | 29.8 | 22.5 |
| DUA | 28.3 | 24.6 | 27.0 | **10.4** | 30.7 | 20.2 | 14.4 | 20.4 | **19.3** | **11.0** | 9.2 | 12.3 | 14.6 | **15.1** | 23.1 | 18.7 |
| Ours | **27.0** | **23.3** | **23.7** | 10.7 | **29.5** | **17.1** | **13.4** | 20.2 | 22.3 | 12.3 | 11.1 | 13.3 | 15.4 | 15.3 | 23.2 | **18.5** |
| Source | 60.8 | 46.5 | 42.6 | 11.0 | 43.1 | 25.6 | 22.5 | 16.4 | 27.4 | 7.7 | 6.2 | 10.3 | 14.1 | 20.0 | 22.0 | 25.1 |
| NORM | 22.2 | 18.7 | 20.7 | **6.7** | 23.9 | 11.5 | 8.6 | 13.9 | 14.0 | 7.1 | 6.5 | 7.8 | 9.9 | 9.9 | 21.1 | 13.5 |
| TENT | 20.1 | 16.9 | 18.1 | **6.7** | 22.0 | 11.2 | **8.3** | 12.7 | 13.5 | 6.9 | 6.5 | 7.1 | 9.5 | **9.4** | 18.2 | 12.5 |
| DUA | 20.8 | 17.7 | 20.1 | 6.9 | 23.3 | 12.4 | 9.5 | 13.9 | 14.0 | **6.6** | **6.1** | 7.7 | 10.0 | 10.1 | 19.1 | 13.2 |
| Ours | **19.3** | **16.3** | **17.1** | 6.8 | **21.1** | **10.8** | **8.3** | **12.3** | **12.9** | 7.1 | 6.3 | **7.0** | **9.3** | 9.7 | **17.6** | **12.1** |
| Source | 20.4 | 14.6 | 9.7 | **5.4** | 12.9 | 8.6 | 6.5 | 9.9 | 11.4 | 6.3 | 5.5 | 7.2 | 7.4 | 9.6 | 12.1 | 9.8 |
| TENT | 12.6 | 10.4 | 10.4 | 6.0 | 12.7 | 8.1 | 6.7 | 9.5 | 9.1 | 6.7 | 5.9 | 7.5 | 8.4 | 7.5 | 12.7 | 8.9 |
| DUA | 12.2 | 10.5 | 9.3 | 5.5 | 11.9 | 7.8 | **6.1** | 9.1 | 9.1 | **6.1** | **5.4** | 6.3 | **7.0** | 7.4 | 11.7 | 8.4 |
| Ours | **11.2** | **9.5** | **9.2** | 5.5 | **11.1** | **7.5** | 6.6 | **8.3** | **8.4** | 6.4 | 5.6 | **6.1** | 7.8 | **7.3** | **11.5** | **8.1** |
| **Level 2** | | | | | | | | | | | | | | | | |
| Source | 43.1 | 27.8 | 29.3 | 10.2 | 49.5 | 23.4 | 22.4 | 26.4 | 21.3 | 10.3 | 8.7 | 13.4 | 14.7 | 17.9 | 22.3 | 22.7 |
| TTT | 28.8 | 20.7 | 23.0 | **9.0** | 36.6 | 15.4 | 13.1 | 20.2 | 16.9 | **9.2** | **8.3** | **10.2** | **12.5** | 14.8 | **19.7** | 17.2 |
| NORM | 31.0 | 25.3 | 28.7 | 13.5 | 38.8 | 18.8 | 16.3 | 27.8 | 23.9 | 15.4 | 14.6 | 17.1 | 18.7 | 19.6 | 30.6 | 22.7 |
| TENT | 26.5 | 21.2 | 25.1 | 11.9 | 34.1 | 16.5 | 14.1 | 24.3 | 20.7 | 13.4 | 12.2 | 16.4 | 16.2 | 16.3 | 26.7 | 19.7 |
| DUA | 22.3 | 16.8 | 22.9 | 9.2 | 30.3 | 16.0 | 12.7 | 21.5 | **15.7** | 9.6 | 8.7 | 11.1 | 12.7 | **13.3** | 20.8 | **16.2** |
| Ours | **22.1** | **16.6** | **19.5** | 10.6 | **29.1** | **14.5** | **12.4** | **19.3** | 17.2 | 11.3 | 10.7 | 12.2 | 14.6 | 14.1 | 21.8 | 16.4 |
| Source | 42.1 | 24.3 | 31.0 | 6.6 | 44.3 | 16.2 | 15.5 | 19.6 | 16.5 | 6.4 | **5.6** | 7.8 | 9.8 | 14.0 | 20.3 | 18.7 |
| NORM | 17.0 | 12.3 | 15.9 | 6.2 | 24.4 | 9.5 | 7.6 | 14.2 | 11.0 | 6.5 | 6.2 | 7.1 | 8.9 | 9.1 | 19.3 | 11.7 |
| TENT | 14.9 | 11.0 | 14.7 | 6.5 | 22.1 | 9.8 | **7.5** | 12.9 | **10.6** | **6.5** | 6.3 | 6.5 | 8.6 | 8.6 | 17.3 | 10.9 |
| DUA | 15.4 | 11.9 | 15.7 | **6.1** | 24.9 | 10.5 | 8.3 | 13.7 | **10.6** | **5.9** | 5.7 | 6.7 | 8.6 | 9.3 | 18.4 | 11.4 |
| Ours | **14.7** | **10.4** | **13.9** | 6.3 | **20.8** | **9.4** | **7.5** | **12.2** | **10.6** | 6.4 | 6.1 | 6.7 | **8.5** | **8.5** | **16.7** | **10.6** |
| Source | 13.4 | 8.8 | 8.0 | **5.1** | 14.2 | 6.5 | 5.8 | 9.2 | 8.5 | 5.3 | 5.3 | 6.1 | 6.5 | 7.8 | 10.9 | 8.1 |
| TENT | 10.2 | 7.6 | 8.6 | 5.9 | 13.0 | 7.2 | 6.2 | 8.1 | 7.8 | 6.3 | 5.8 | 6.9 | 7.5 | 7.0 | 11.8 | 8.0 |

*continued on next page*

| Methods | gaus | shot | impul | defcs | gls | mtn | zm | snw | frst | fg | brt | cnt | els | px | jpg | Avg. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DUA | 10.0 | **7.5** | **7.6** | **5.1** | 12.4 | **6.4** | **5.7** | 8.3 | **7.3** | **5.2** | **5.2** | **5.7** | **6.4** | **6.8** | 10.9 | **7.4** |
| Ours | **9.1** | **7.5** | **7.6** | 5.5 | **11.3** | 7.0 | 6.3 | **7.6** | **7.3** | 5.9 | 5.4 | **5.7** | 7.1 | 6.9 | **10.4** | **7.4** |
| **Level 1** | | | | | | | | | | | | | | | | |
| Source | 25.8 | 18.4 | 19.0 | 8.5 | 51.1 | 14.7 | 18.2 | 15.0 | 13.8 | 8.3 | 8.3 | 8.7 | 14.4 | 11.3 | 16.5 | 16.8 |
| TTT | 19.1 | 15.8 | 16.5 | **8.0** | 37.9 | **11.7** | **12.2** | **12.8** | **11.9** | **8.2** | **8.0** | **8.3** | **12.6** | 11.1 | **15.5** | 14.0 |
| NORM | 24.0 | 20.9 | 22.5 | 13.4 | 38.1 | 16.5 | 15.5 | 20.5 | 18.8 | 14.9 | 14.0 | 15.3 | 19.1 | 16.9 | 24.7 | 19.7 |
| TENT | 20.7 | 18.0 | 19.1 | 11.8 | 35.0 | 14.6 | 13.7 | 17.5 | 16.4 | 12.6 | 11.8 | 13.8 | 16.7 | 14.5 | 20.9 | 17.1 |
| DUA | 16.5 | **13.8** | 16.6 | 8.3 | 30.4 | 12.4 | 12.6 | 14.5 | 12.2 | 8.4 | 8.4 | 8.8 | 13.4 | **11.0** | 15.9 | **13.6** |
| Ours | **16.4** | 14.1 | **15.2** | 10.4 | **28.9** | 12.8 | 12.3 | 14.8 | 13.8 | 10.9 | 10.5 | 11.1 | 14.3 | 12.4 | 17.4 | 14.4 |
| Source | 22.2 | 15.0 | 17.1 | **5.4** | 46.6 | 9.7 | 12.3 | 10.1 | 10.5 | **5.5** | **5.3** | 5.7 | 9.5 | 8.1 | 13.6 | 13.1 |
| NORM | 11.7 | 9.9 | 11.2 | 6.0 | 23.9 | 7.7 | 7.9 | 9.4 | 8.3 | 6.0 | 6.0 | 6.2 | 9.2 | 7.8 | 13.2 | 9.6 |
| TENT | 11.0 | 8.8 | 10.7 | 6.2 | 21.4 | **7.5** | **7.6** | 8.8 | 8.2 | 5.8 | 6.3 | 5.9 | **8.7** | 7.8 | 12.7 | 9.2 |
| DUA | 11.8 | **8.7** | 11.4 | 5.5 | 23.2 | 7.9 | 8.5 | 9.3 | 8.2 | **5.5** | 5.4 | **5.6** | 9.0 | **7.5** | 12.8 | 9.4 |
| Ours | **10.5** | **8.7** | **9.8** | 6.0 | **20.7** | 7.6 | **7.6** | **8.5** | **7.8** | 5.8 | 6.0 | 6.1 | 8.9 | **7.5** | **12.2** | **8.9** |
| Source | 8.7 | 6.5 | **6.2** | **4.9** | 14.1 | **5.5** | 5.9 | 6.4 | 6.5 | **4.9** | **5.0** | **5.0** | **6.9** | **5.8** | 8.7 | 6.7 |
| TENT | 7.5 | 6.9 | 7.2 | 5.7 | 12.4 | 6.2 | 6.3 | 6.8 | 6.5 | 5.9 | 5.7 | 6.0 | 7.9 | 6.5 | 9.1 | 7.1 |
| DUA | 7.3 | **6.2** | **6.2** | 5.1 | 11.9 | **5.5** | **5.8** | 6.2 | 6.1 | 5.1 | 5.1 | 5.1 | 7.0 | **5.8** | 8.5 | **6.5** |
| Ours | **6.8** | 6.4 | 6.5 | 5.4 | **11.3** | 5.9 | **5.8** | **6.1** | **6.0** | 5.2 | 5.3 | 5.3 | 7.3 | 6.0 | **8.3** | **6.5** |

Table 5. Error (%) for each corruption in **CIFAR-100C** severity (Level 1–4) is reported. Source refers to results obtained from a model trained on clean train set and tested on corrupted test sets. For TENT and DUA, we use the **WRN-40-2** from their official implementation. Lowest error is highlighted for each corruption.

| Methods | gaus | shot | impul | defcs | gls | mtn | zm | snw | frst | fg | brt | cnt | els | px | jpg | Avg. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Level 4** | | | | | | | | | | | | | | | | |
| Source | 60.7 | 51.6 | 47.9 | 27.1 | 54.4 | 30.3 | 28.9 | 37.4 | 39.0 | 35.4 | 27.2 | 35.9 | 34.4 | 39.0 | 40.1 | 39.3 |
| NORM | 42.6 | 39.9 | 41.6 | 29.2 | 45.9 | 31.4 | 30.9 | 38.0 | 35.2 | 36.0 | 28.1 | 31.7 | 35.7 | 32.6 | 41.9 | 36.1 |
| TENT | 38.9 | 36.3 | 36.6 | 27.3 | 42.0 | 28.9 | 28.4 | 34.8 | 32.8 | 32.1 | 26.3 | 29.8 | 33.3 | 29.9 | 38.4 | 33.1 |
| DUA | 43.0 | 39.0 | 37.9 | 26.9 | 44.7 | 29.5 | 28.3 | 36.5 | 34.3 | 33.8 | 26.6 | 31.0 | 33.8 | 31.0 | 38.9 | 34.3 |
| Ours | **36.8** | **33.9** | **32.4** | **26.3** | **39.8** | **28.1** | **27.8** | **32.4** | **30.9** | **30.4** | **25.4** | **27.6** | **33.0** | **27.8** | **36.8** | **31.3** |
| **Level 3** | | | | | | | | | | | | | | | | |
| Source | 55.2 | 45.9 | 36.9 | 25.7 | 39.9 | 30.5 | 27.4 | 33.3 | 38.1 | 29.5 | 25.5 | 30.5 | 28.6 | 30.3 | 38.0 | 34.4 |
| NORM | 40.7 | 37.7 | 35.6 | 28.2 | 38.7 | 31.3 | 29.9 | 35.1 | 35.4 | 32.4 | 27.6 | 30.5 | 31.7 | 30.5 | 40.4 | 33.7 |
| TENT | 37.1 | 34.3 | 32.0 | 26.1 | 34.6 | 29.2 | 27.7 | 32.3 | 32.3 | 29.0 | 25.6 | 28.4 | 29.2 | 28.3 | 37.4 | 30.9 |
| DUA | 40.8 | 36.5 | 32.2 | 25.3 | 37.0 | 29.6 | 27.1 | 32.7 | 34.4 | 29.0 | 25.2 | 28.6 | **28.4** | 28.7 | 37.3 | 31.5 |
| Ours | **35.3** | **32.5** | **28.9** | **25.2** | **33.5** | **27.8** | **26.4** | **30.0** | **30.6** | **27.9** | **24.8** | **26.5** | 28.8 | **26.8** | **35.7** | **29.4** |
| **Level 2** | | | | | | | | | | | | | | | | |
| Source | 44.6 | 34.5 | 30.7 | **24.3** | 41.5 | 27.7 | 26.2 | 32.7 | 31.8 | 26.8 | 24.4 | 27.5 | 27.9 | 28.0 | 36.5 | 31.0 |
| NORM | 36.2 | 32.3 | 32.2 | 27.6 | 38.0 | 29.5 | 29.2 | 33.2 | 32.6 | 29.6 | 27.5 | 29.5 | 31.4 | 30.5 | 38.7 | 31.9 |
| TENT | 33.3 | 29.5 | 28.8 | 25.7 | 34.9 | 27.4 | 27.0 | 30.6 | 29.6 | 27.0 | 25.4 | 27.3 | 29.1 | 27.6 | 36.3 | 29.3 |
| DUA | 35.9 | 31.2 | 28.8 | 24.1 | 36.9 | 27.2 | **26.0** | 32.2 | 30.9 | **26.1** | **24.0** | 26.6 | **27.9** | 27.6 | 35.9 | 29.4 |
| Ours | **31.9** | **28.3** | **27.3** | 24.9 | **33.3** | **26.3** | 26.1 | **28.8** | **28.4** | 26.2 | 24.5 | **25.9** | 28.3 | **26.5** | **34.5** | **28.1** |
| **Level 1** | | | | | | | | | | | | | | | | |
| Source | 34.4 | 29.6 | 26.9 | **23.8** | 42.9 | 25.6 | 26.1 | 26.1 | 27.4 | **24.0** | **23.8** | 24.3 | **28.4** | **25.2** | 32.4 | 28.1 |
| NORM | 32.2 | 30.5 | 29.3 | 27.3 | 37.7 | 28.3 | 28.7 | 29.4 | 29.8 | 27.6 | 27.5 | 27.9 | 32.3 | 29.0 | 35.2 | 30.2 |
| TENT | 29.3 | 27.6 | 26.9 | 25.5 | 34.7 | 26.6 | 26.7 | 27.0 | 27.3 | 25.6 | 25.2 | 26.0 | 29.8 | 26.7 | 32.9 | 27.9 |
| DUA | 31.2 | 28.5 | 26.4 | **23.8** | 36.9 | **25.3** | 25.9 | 26.2 | 27.1 | 24.1 | 23.9 | **23.9** | 28.6 | 25.4 | 32.0 | 27.3 |
| Ours | **28.1** | **26.9** | **25.9** | 24.6 | **33.1** | 25.8 | **25.6** | **25.9** | **26.5** | 24.6 | 24.5 | 25.2 | 29.4 | 25.9 | **31.2** | **26.9** |

Table 6. Error (%) for each corruption in **ImageNet-C** severity (Level 1–4) is reported. Source refers to results obtained from a model pre-trained on ImageNet and tested on corrupted test sets. For TENT and DUA, we use the **ResNet-18** from the RobustBench. Lowest error is highlighted for each corruption.

| Methods | gaus | shot | impul | defcs | gls | mtn | zm | snw | frst | fg | brt | cnt | els | px | jpg | Avg. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | **Level 4** | | | | | | | | |
| Source | 92.8 | 93.5 | 96.0 | 90.2 | 82.0 | 85.9 | 78.6 | 86.8 | 87.8 | 99.1 | 65.7 | 99.5 | 56.2 | 57.2 | 52.5 | 81.6 |
| TTT | 64.5 | 68.2 | 70.6 | 84.5 | 68.7 | 69.7 | 64.7 | 81.4 | 72.8 | 94.8 | **48.5** | 98.8 | **49.5** | **49.1** | **47.1** | 68.9 |
| NORM | 59.3 | 61.2 | 60.0 | 65.6 | 59.1 | 61.1 | 58.9 | 66.5 | 62.8 | 69.2 | 53.4 | 72.3 | 51.1 | 51.3 | 51.2 | 60.2 |
| TENT | 58.2 | 58.8 | 58.5 | 66.0 | 59.4 | 59.4 | 60.6 | 60.2 | 61.8 | **59.7** | 52.7 | 91.4 | 51.9 | 51.9 | 51.9 | 60.2 |
| DUA | 66.0 | 67.8 | 66.2 | 86.1 | 71.3 | 78.0 | 71.2 | 78.7 | 73.7 | 80.8 | 57.1 | 99.2 | 53.7 | 55.8 | 51.8 | 70.5 |
| Ours | **53.8** | **54.0** | **53.7** | **64.1** | **58.4** | **58.0** | **59.0** | **59.1** | **60.6** | 60.4 | 51.2 | **67.3** | 50.3 | 51.1 | 50.7 | **56.8** |
| | | | | | | | | **Level 3** | | | | | | | | |
| Source | 78.9 | 80.5 | 85.2 | 85.4 | 76.5 | 77.8 | 75.7 | 77.3 | 86.2 | 98.8 | 56.7 | 98.9 | 53.9 | 54.1 | 51.4 | 75.8 |
| TTT | 57.2 | 55.8 | 58.0 | 76.8 | 60.9 | 61.9 | 61.1 | 70.8 | 70.3 | 93.7 | **43.4** | 96.9 | **48.4** | **46.3** | **43.8** | 63.0 |
| NORM | 55.2 | 55.7 | 56.9 | 59.4 | 56.1 | 55.3 | **56.4** | 60.5 | 62.1 | 67.2 | 51.2 | 60.2 | 50.5 | 50.5 | 50.8 | 56.5 |
| TENT | 54.6 | 54.5 | 55.5 | 59.8 | 57.0 | 56.2 | 58.5 | 56.5 | 61.1 | 59.3 | 51.2 | 71.6 | 51.2 | 50.9 | 51.3 | 56.6 |
| DUA | 60.6 | 61.5 | 61.7 | 78.5 | 67.1 | 69.4 | 68.9 | 70.4 | 72.4 | 79.5 | 53.1 | 94.4 | 52.2 | 53.4 | 51.0 | 66.3 |
| Ours | **51.3** | **51.0** | **52.2** | **57.3** | **56.0** | **54.8** | 56.9 | **55.5** | **60.3** | **59.2** | 49.4 | **59.2** | 49.3 | 49.9 | 50.0 | **54.2** |
| | | | | | | | | **Level 2** | | | | | | | | |
| Source | 61.9 | 66.5 | 75.1 | 76.2 | 68.8 | 68.3 | 70.4 | 73.7 | 77.0 | 96.7 | 51.6 | 95.4 | 71.5 | 52.4 | 51.0 | 70.4 |
| TTT | 51.8 | **49.6** | 52.7 | 66.5 | 56.3 | 54.0 | 56.8 | 65.4 | 67.0 | 87.6 | **41.4** | 87.8 | **60.7** | **44.2** | **43.5** | 59.0 |
| NORM | 52.0 | 52.6 | 55.7 | 54.0 | 52.3 | 52.4 | 54.1 | 56,4 | 57.4 | 62.0 | 50.0 | 55.9 | 68.7 | 50.0 | 50.3 | 54.9 |
| TENT | 52.1 | 52.7 | 54.1 | 54.6 | 53.5 | 52.6 | 55.3 | 54.7 | 57.3 | 57.4 | 50.7 | 59.5 | 70.4 | 50.9 | 51.4 | 55.2 |
| DUA | 54.6 | 56.6 | 58.6 | 67.6 | 61.3 | 62.4 | 64.3 | 66.3 | 65.7 | 73.6 | 50.3 | 83.1 | 70.8 | 51.6 | 50.4 | 62.5 |
| Ours | **49.5** | 50.1 | **51.4** | **52.8** | **51.8** | **51.7** | **53.8** | **53.2** | **55.8** | **56.8** | 48.9 | **54.8** | 66.3 | 48.8 | 49.3 | **53.0** |
| | | | | | | | | **Level 1** | | | | | | | | |
| Source | 51.8 | 53.7 | 60.4 | 69.3 | 61.1 | 60.1 | 65.0 | 60.1 | 59.4 | 91.0 | 49.1 | 84.9 | 56.3 | 51.6 | 50.5 | 61.6 |
| TTT | **44.4** | **45.1** | **49.0** | 58.2 | **47.2** | 52.7 | 53.5 | 53.1 | **48.8** | 78.0 | **40.9** | 72.1 | **49.0** | 49.2 | 49.6 | 52.7 |
| NORM | 50.6 | 50.6 | 53.6 | 52.1 | 50.9 | 50.6 | 52.9 | 53.0 | 52.3 | 57.3 | 49.8 | 53.3 | 54.4 | 49.7 | 50.1 | 52.1 |
| TENT | 51.5 | 51.2 | 52.4 | 52.7 | 51.6 | 51.3 | 53.9 | 52.4 | 53.0 | 55.4 | 50.0 | 56.1 | 55.1 | 50.3 | 50.4 | 52.5 |
| DUA | 50.3 | 51.6 | 54.1 | 61.5 | 57.7 | 57.0 | 61.5 | 56.6 | 54.5 | 67.6 | 48.9 | 74.0 | 55.2 | 51.1 | 50.2 | 56.8 |
| Ours | 49.0 | 49.1 | 50.5 | **51.5** | 49.9 | **50.2** | **52.0** | **51.0** | 50.8 | **54.7** | 48.4 | **52.8** | 52.4 | **48.9** | **49.0** | **50.7** |