Appendix for
"A Unified Framework for Robustness on Diverse Sampling Errors"

Myeongho Jeon Myungjoo Kang Joonseok Lee
Seoul National University
{andyjeon, mkang, joonseok}@snu.ac.kr

A. Data Configuration

We report the overall data configuration for SDG (DG-CIFAR, PACS) and UBL (B-CIFAR, IMDB).

- DG-CIFAR: Following the original data split, the train set in CIFAR10 (50,000 images) is used for training and validation set is augmented to be 12 domain sets, {*fog, snow, frost, zoom blur, defocus blur, glass blur, speckle noise, shot noise, impulse noise, jpeg compression, pixelate, spatter*}, each of which is composed of 10,000 images.
- PACS: the train set of 'photo' (1,499) domain is used for training, and the model is evaluated on test sets of 'art painting' (2,048), 'cartoon' (2,344), 'sketch' (3,929) (Fig. II (b)).
- B-CIFAR: The augmented training set of CIFAR10 containing $\{(airplane, fog), \ldots, (truck, saturate)\}$ with 0.5% unbiased instances is used for training (50,000), and uniformly distributed test set (10,000) is employed for evaluation (Fig. I).
- IMDB: First, EB1 (16,800) is set as training set and EB2 (36,004), TEST are exploited for evaluation. Second, EB2 for training and EB1, TEST (22,468) for evaluation (Fig. II (a)).
    1. EB1: women aged 0-29, men aged 40+
    2. EB2: women aged 40+, men aged 0-29
    3. TEST: 0-29 or 40+ women and men

B. Implementation Details

In all experiments, we initialize the backbone network F (ResNet18) by ImageNet-pretrained weights and set the batch size as 64. During training, we use stochastic gradient descent (SGD) optimizer with momentum (0.9), weight decay (0.0005). K is set as 5 for all the datasets, with a different kernel size of `Conv` and $\texttt{Conv}^\top$. We consider a residual block in ResNet18 as a layer l below.

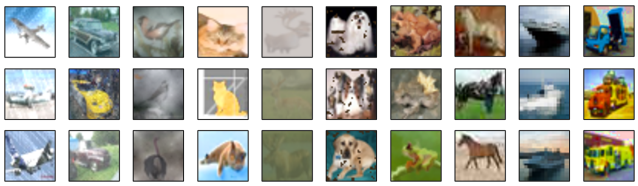

Figure I. B-CIFAR. To present corrupted CIFAR-10 image samples, we display B-CIFAR. Note that the images for DG-CIFAR are augmented in exactly the same way, only with different distributions. The figure is referenced by [1].

DG-CIFAR.

- Input image: resized to (32,32), scaled to [-1,1], normalized, and applied random horizontal flip.
- Epoch: 50.
- Architectural parameters: $l = 4$, $M = 2$
- Learning rate: initialized as 0.005 and decayed to 0.0005 when epoch is 40.
- Balancing loss: $\alpha = 0.0001$, $\beta_1 = 5$, $\beta_2 = 5$, and $\beta_3 = 0.1$.

PACS.

- Input image: resized to (224,224), scaled to [-1,1], and normalized.
- Epoch: 50.
- Architectural parameters: $l = 4$, $M = 3$
- Learning rate: 0.005 and decayed to 0.0005 when epoch is 40.
- balancing loss: $\alpha = 0.1$, $\beta_1 = 3.33$, $\beta_2 = 3.33$, and $\beta_3 = 0.1$.

B-CIFAR.

- Input image: resized to (32,32), scaled to [-1,1], and normalized.
- Epoch: 100.
- Architectural parameters: $l = 4$, $M = 3$
- Learning rate: 0.005 and decayed to 0.0005 when epoch is 40.

(a) (b)

Figure II. Image samples of real-world datasets. (a) IMDB, (b) PACS. For PACS, we sample 3 classes, dog, guitar, and house, out of 7 classes

- Balancing loss: $\alpha = 0.1$, $\beta_1 = 3.33$, $\beta_2 = 3.33$, and $\beta_3 = 0.1$.

IMDB.

- Input image: resized to (128,128).
- Epoch: 30.
- Architectural parameters: $l = 4$, $M = 3$
- Learning rate: 0.005 and decayed to 0.0005 when epoch is 24.
- Balancing loss: $\alpha = 0.1$, $\beta_1 = 1$, $\beta_2 = 1$, and $\beta_3 = 0.1$.

References

[1] Myeongho Jeon, Hyoje Lee, Yedarm Seong, and Myungjoo Kang. Learning without prejudices: Continual unbiased learning via benign and malignant forgetting. In *International Conference on Learning Representations*, 2023. i