Supplementary Material
Difficulty-Aware Simulator for Open Set Recognition

WonJun Moon, Junho Park, Hyun Seok Seong, Cheol-Ho Cho, and Jae-Pil Heo*

Sungkyunkwan University

1 Split Information

As we elaborated in the main paper, we adopted the protocols from [2] and [4] for evaluations with AUROC and F1-score, respectively. To further encourage the fair comparison, we publicize the split details. Specifically, we enumerate categories that are used for closed-set in Tab. 1 and Tab. 2 for measuring F1-score and AUROC, respectively. Note that for CIFAR+, we show the categories of open-set classes since CIFAR+ experiments utilize the non-animal classes in CIFAR10 dataset, i.e., airplain, automobile, ship, and truck, as the closed-set. We sincerely hope future works use pre-defined standard split information to prevent confusion in understanding the effectiveness of their methods and for a fair comparison.

Table 1. Data splits for Tab. 3 in the main paper. This split information is used for measuring F1-scores. The numbers in the table represent the class indices for closed set except CIFAR+ cases. For CIFAR+ experiments, we provide open-set class indices, since animal classes are utilized for closed set.

F1 Split Information					
	0	**1**	**2**	**3**	**4**
MNIST	2, 3, 4, 6, 7, 8	0, 1, 4, 6, 7, 9	1, 2, 4, 6, 7, 8	1, 3, 4, 6, 7, 8	1, 2, 3, 5, 7, 8
SVHN	2, 3, 4, 6, 7, 8	0, 1, 4, 6, 7, 9	1, 2, 4, 6, 7, 8	1, 3, 4, 6, 7, 8	1, 2, 3, 5, 7, 8
CIFAR10	2, 3, 4, 6, 7, 8	0, 1, 4, 6, 7, 9	1, 2, 4, 6, 7, 8	1, 3, 4, 6, 7, 8	1, 2, 3, 5, 7, 8
CIFAR+10	27, 46, 98, 38, 72, 31, 36, 66, 3, 97	98, 46, 14, 1, 7, 73, 3, 79, 93, 11	79, 98, 67, 7, 77, 42, 36, 65, 26, 64	46, 77, 29, 24, 65, 66, 79, 21, 1, 95	21, 95, 64, 55, 50, 24, 93, 75, 27, 36
CIFAR+50	27, 46, 98, 38, 72, 31, 36, 66, 3, 97, 75, 67, 42, 32, 14, 93, 6, 88, 11, 1, 44, 35, 73, 19, 18, 78, 15, 4, 50, 65, 64, 55, 30, 80, 26, 2, 7, 34, 79, 43, 74, 29, 45, 91, 37, 99, 95, 63, 24, 21	98, 46, 14, 1, 7, 73, 3, 79, 93, 11, 37, 29, 2, 74, 91, 77, 55, 50, 18, 80, 63, 67, 4, 45, 95, 30, 75, 97, 88, 36, 31, 27, 65, 32, 43, 72, 6, 26, 15, 42, 19, 34, 38, 66, 35, 21, 24, 99, 78, 44	79, 98, 67, 7, 77, 42, 36, 65, 26, 64, 66, 73, 75, 3, 32, 14, 35, 6, 24, 21, 55, 34, 30, 43, 93, 38, 19, 99, 72, 97, 78, 18, 31, 63, 29, 74, 91, 4, 27, 46, 2, 88, 45, 15, 11, 1, 95, 50, 80, 44	46, 77, 29, 24, 65, 66, 79, 21, 1, 95, 36, 88, 27, 99, 67, 19, 75, 42, 2, 73, 32, 98, 72, 97, 78, 11, 14, 74, 50, 37, 26, 64, 44, 30, 31, 18, 38, 4, 35, 80, 45, 63, 93, 34, 3, 43, 6, 55, 91, 15	21, 95, 64, 55, 50, 24, 93, 75, 27, 36, 73, 63, 19, 98, 46, 1, 15, 72, 42, 78, 77, 29, 74, 30, 14, 38, 80, 45, 4, 26, 31, 11, 97, 7, 66, 65, 99, 34, 6, 18, 44, 3, 35, 88, 43, 91, 32, 67, 37, 79

* Corresponding author

Table 2. Data splits for Tab. 1 in the main paper. This split information is used for measuring AUROC scores. The numbers in the table represent the class indices for closed set except CIFAR+ cases. For CIFAR+ experiments, we provide open-set class indices, since animal classes are utilized for closed set.

AUROC Split Information					
	0	**1**	**2**	**3**	**4**
MNIST	0, 1, 2, 4, 5, 9	0, 3, 5, 7, 8, 9	0, 1, 5, 6, 7, 8	3, 4, 5, 7, 8, 9	0, 1, 2, 3, 7, 8
SVHN	0, 1, 2, 4, 5, 9	0, 3, 5, 7, 8, 9	0, 1, 5, 6, 7, 8	3, 4, 5, 7, 8, 9	0, 1, 2, 3, 7, 8
CIFAR10	0, 1, 2, 4, 5, 9	0, 3, 5, 7, 8, 9	0, 1, 5, 6, 7, 8	3, 4, 5, 7, 8, 9	0, 1, 2, 3, 7, 8
CIFAR+10	26, 31, 34, 44, 45, 63, 65, 77, 93, 98	7, 11, 66, 75, 77, 93, 95, 97, 98, 99	2, 11, 15, 24, 32, 34, 63, 88, 93, 95	1, 11, 38, 42, 44, 45, 63, 64, 66, 67	3, 15, 19, 21, 42, 46, 66, 72, 78, 98
CIFAR+50	1, 2, 7, 9, 10, 12, 15, 18, 21, 23, 26, 30, 32, 33, 34, 36, 37, 39, 40, 42, 44, 45, 46, 47, 49, 50, 51, 52, 55, 56, 59, 60, 61, 63, 65, 66, 70, 72, 73, 74, 76, 78, 80, 83, 87, 91, 92, 96, 98, 99	0, 2, 4, 5, 9, 12, 14, 17, 18, 20, 21, 23, 24, 25, 31, 32, 33, 35, 39, 43, 45, 49, 50, 51, 52, 54, 55, 56, 60, 64, 65, 66, 68, 70, 71, 73, 74, 77, 78, 79, 80, 82, 83, 86, 91, 93, 94, 96, 97, 98	0, 4, 10, 11, 12, 14, 15, 17, 18, 21, 23, 26, 27, 28, 29, 31, 32, 33, 36, 39, 40, 42, 43, 46, 47, 51, 53, 56, 57, 59, 60, 64, 66, 71, 73, 74, 75, 76, 78, 79, 80, 83, 87, 91, 92, 93, 94, 95, 96, 99	0, 2, 5, 6, 9, 10, 11, 12, 14, 16, 18, 19, 21, 22, 23, 26, 27, 28, 29, 31, 33, 35, 36, 37, 38, 39, 40, 43, 45, 49, 52, 56, 59, 61, 62, 63, 64, 65, 71, 74, 75, 78, 80, 82, 86, 87, 91, 93, 94, 96	0, 1, 4, 6, 7, 12, 15, 16, 17, 19, 20, 21, 22, 23, 26, 27, 28, 32, 39, 40, 42, 43, 44, 47, 49, 50, 52, 53, 54, 55, 56, 59, 61, 62, 63, 65, 66, 67, 68, 73, 74, 77, 82, 83, 86, 87, 93, 94, 97, 98
Tiny-IN	2, 3, 13, 30, 44, 45, 64, 66, 76, 101, 111, 121, 128, 130, 136, 158, 167, 170, 187, 193	4, 11, 32, 42, 51, 53, 67, 84, 87, 104, 116, 140, 144, 145, 148, 149, 155, 168, 185, 193	3, 9, 10, 20, 23, 28, 29, 45, 54, 74, 133, 143, 146, 147, 156, 159, 161, 170, 184, 195	1, 15, 17, 31, 36, 44, 66, 69, 84, 89, 102, 137, 154, 160, 170, 177, 182, 185, 195, 197	4, 14, 16, 33, 34, 39, 59, 69, 77, 92, 101, 103, 130, 133, 147, 161, 166, 168, 172, 173

2 Regularization Loss

As we introduced in the main paper, we simply used cross entropy loss function for regularization loss, $\mathcal{L}_{reg}$. In this section, we simply examine the influence of $\mathcal{L}_{reg}$ with two datasets: CIFAR10 and Tiny-ImageNet. Results in Tab. 3 show that DIAS is not very sensitive to the ratio for $\mathcal{L}_{reg}$.

Loss Ratio	0.1	0.2	1.0	1.5
CIFAR10	$0.852_{\pm 0.02}$	$0.851_{\pm 0.03}$	$0.850_{\pm 0.02}$	$0.851_{\pm 0.03}$
Tiny-ImageNet	$0.713_{\pm 0.02}$	$0.729_{\pm 0.01}$	$0.731_{\pm 0.01}$	$0.726_{\pm 0.01}$

Table 3. AUROC score with varying ratios of $\mathcal{L}_{reg}$.

3 Implementation details

DIAS is an end-to-end framework that all components are learned from the scratch. For the Copycat and the classifier, we use vanilla CNN [3], which is composed of 9 convolution layers. For the subgroups of convolutional layers, each group contains three 3x3 convolution layers. Additionally, the backbone network for the generator and the discriminator each contains 4 convolutional layers. Moreover, we adopt multi-batch normalization layers to process generated images from GAN separately, as we hope to prevent the problem from distribution mismatch, following [1]. Note that features from the Copycat do not need to be processed separately. For scaling parameters, we fix both λ, and β to 0.1.

References

1. Chen, G., Peng, P., Wang, X., Tian, Y.: Adversarial reciprocal points learning for open set recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) (2021). https://doi.org/10.1109/TPAMI.2021.3106743
2. Guo, Y., Camporese, G., Yang, W., Sperduti, A., Ballan, L.: Conditional variational capsule network for open set recognition. Proceedings of the IEEE international conference on computer vision (ICCV) (2021)
3. Neal, L., Olson, M., Fern, X., Wong, W.K., Li, F.: Open set learning with counterfactual images. In: Proceedings of the European Conference on Computer Vision (ECCV) (2018)
4. Yue, Z., Wang, T., Sun, Q., Hua, X.S., Zhang, H.: Counterfactual zero-shot and open-set visual recognition. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2021)