

Supplementary Material

Xiwen Dengxiong

Golisano College of Computing and Information Sciences
Rochester Institute of Technology
sd6384@rit.edu

Yu Kong

Department of Computer Science and Engineering
Michigan State University
yukong@msu.edu

1. Background

1.1. Open Set Recognition

Openness Openness represents the open degree of the dataset in basic open set recognition. The openness is defined by the number of categories in the training process (N_{train}) and the testing process (N_{test}). Let N_{re} , be the number of classes to be recognized. The openness in the open space shows as follows:

$$\text{Openness} = 1 - \sqrt{\frac{2 \times N_{\text{train}}}{N_{\text{re}} + N_{\text{test}}}} \quad (1)$$

In the definition of openness, N_{train} usually equals to N_{re} . Larger openness represents more classes in open space and indicates the difficulty of rejecting the unknown classes.

Open Space Risk Given K known classes $1, 2, \dots, K$, training samples belong to these known classes, and a large amount of testing data belong to unknown classes $K+1, K+2, \dots, K+L$. Here L is the number of unknown classes in the testing set. In the OSR problem, the open space consists of the positive open space $\mathcal{O}_k^{\text{pos}}$ and the negative open space $\mathcal{O}_k^{\text{neg}}$. Since we cannot classify the unknown classes into specific subclasses, the main goal of the open set is to minimize the expected risk \mathcal{R}^K ,

$$\mathcal{R}^k = \mathcal{R}_e(\psi_k, \mathcal{S}_k \cup \mathcal{O}_k^{\text{pos}}) + \alpha \cdot \mathcal{R}_o(\psi_k, \mathcal{O}_k^{\text{neg}}), \quad (2)$$

where ψ_k is the binary measurable prediction function that maps the embedding sample x to class k . α is a positive regularization parameter. \mathcal{R}_e is the classification risk for known classes, while \mathcal{R}_o is the open space risk for unknown classes. Since samples excluded from KCCs and KUCs are regarded as open space, and samples from UUCs are unforeseen in the training process, the risk of open space should be formalized as follows:

$$\mathcal{R}_o(\psi_k, \mathcal{O}_k^{\text{neg}}) = \frac{\int_{\mathcal{O}_k^{\text{neg}}} \psi_k(x) dx}{\int_{\mathcal{S}_k \cup \mathcal{O}_k} \psi_k(x) dx} \quad (3)$$

1.2. Hyperbolic Embedding

The n -dimensional hyperbolic space is defined by Riemannian manifold B^n . Here we define the Riemannian metric as g^B . The Riemannian metric is conformal to the Euclidean metric g^E if they define the same angles. The conformal factor is defined as λ . The Riemannian metric is equivalent to the Euclidean metric when $g^B = \lambda^2 g^E$. Specifically, the n -dimensional unit ball is:

$$\mathbb{B}^n = \{x \in \mathbb{R}^n : \|x\| < 1\} \quad (4)$$

Möbius transformation. Standard operations in hyperbolic space are different from the Euclidean space. Convenient choice of defining basic operations in hyperbolic is based on the Gyrovector space [9].

Möbius addition

$$\mathbf{x} \oplus \mathbf{y} := \frac{(1 + 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2) \mathbf{x} + (1 - \|\mathbf{x}\|^2) \mathbf{y}}{1 + 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{x}\|^2 \|\mathbf{y}\|^2} \quad (5)$$

Möbius scalar multiplication

$$r \otimes x = \begin{cases} \tanh\left(r \operatorname{artanh}\left(\frac{\|x\|}{\|x\|}\right)\right) \frac{x}{\|x\|}, & x \in \mathbb{B}^n \\ 0, & x = 0 \end{cases} \quad (6)$$

where $\langle \mathbf{x}, \mathbf{y} \rangle$ represents Minkowski inner product of input \mathbf{x} and \mathbf{y} , i.e., given $\mathbf{x} = (x_0, x_1, \dots, x_m)$, $\mathbf{y} = (y_0, y_1, \dots, y_m)$, $\langle \mathbf{x}, \mathbf{y} \rangle = -x_0 y_0 + x_1 y_1 + \dots + x_m y_m$.

Generating Loss. Details about generating loss.

$$\mathcal{L}_{\text{gen}} = - \sum_{\hat{\mathbf{f}} \in \mathcal{T}} \log \frac{e^{c \cdot \rho(\hat{\mathbf{f}}; i, j)}}{\sum_{i, j \in Y_1} e^{c \cdot \rho(\hat{\mathbf{f}}; i, j)}} \quad (7)$$

where $\rho(\hat{\mathbf{f}}; i, j)$ defines the total distance of $\hat{\mathbf{f}}$ to the attribute feature $\mathbf{f}_a^{(i)}$ and $\mathbf{f}_b^{(j)}$ in class i and j , i.e., $\rho(\hat{\mathbf{f}}, \mathbf{f}_a^{(i)}) + \rho(\hat{\mathbf{f}}, \mathbf{f}_b^{(j)})$. Here, the distance measure $\rho(\cdot, \cdot)$ is defined as the Euclidean distance in Euclidean space. The c is the

temperature hyperparameter. \mathcal{T} is the set of novel attribute features.

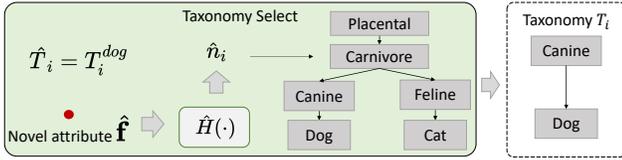


Figure 1. Example of Taxonomy select process. Green block illustrate how taxonomy T_i is selected. White block denotes the selected taxonomy.

Ancestor Search This section illustrates details about taxonomy select process. If the initial potential ancestor node \hat{n}_i is in taxonomy \hat{T}_i the selected taxonomy $T_i = \hat{T}_i - \{\hat{n}_i, \dots, \hat{n}_{root}\}$. For example in Figure 1, $\hat{n}_i = Carnivore$ is the initial predicted ancestor of novel attribute $\hat{\mathbf{f}}$. $\hat{T}_i = T_i^{dog}$, i.e., $\hat{T}_i = \{Placental, Carnivore, Canine, Dog\}$. Because $\hat{n}_i \in \hat{T}_i$, selected taxonomy $T_i = \{Placental, Carnivore, Canine, Dog\} - \{Carnivore, Placental\}$. Taxonomy $T_i = \{Canine, Dog\}$. If the initial potential ancestor node \hat{n}_i is the leaf node of \hat{T}_i , the framework will skip the ancestor node remove process.

2. Experiment Details

2.1. Datasets

Datasets with attribute annotations. The following two datasets contain attribute annotations. These two datasets are used for generalized open set recognition.

- **CUB-200**[10]: Caltech-UCSD bird dataset consists of 11788 images with accurate quantized attribute annotation from 200 bird species. There are 312 image-level attributes in the annotation of the dataset introducing some key features of the bird species. The dataset also provides the confidence score for these features, which can be used in the Möbius transformation. The maximum depth of the dataset is two.
- **AWA 2**[3]: Animal with attributes 2 dataset contains 85 attributes for 50 animal species. We use the taxonomy reported in previous hierarchical class detection work[5]. The taxonomy of AWA 2 dataset is more complex than CUB-200. The maximum depth of the dataset is five.

Datasets without attribute annotations. The following five datasets are used for open set recognition. Details about datasets and dataset split are as follows.

- **MNIST**[4] MNIST is a Handwritten digits dataset with six known classes and four unknown classes. The openness of this task is 13.39%.
- **SVHN**[7] SVHN is a Street View House Numbers dataset with colorful images of digits. We use the same known and unknown class setting as MNIST. Therefore the openness in this dataset is 13.39%
- **CIFAR-10**[2] CIFAR-10 is a colored image dataset with 10 categories. It contains six known classes and four unknown classes.
- **CIFAR+50**[2] CIFAR-100 is a colored image dataset with 100 categories. Specifically, four known classes are sampled from the CIFAR-10 dataset, and N non-overlapping classes from CIFAR-100 are randomly sampled as unknown classes. The openness in this dataset is 61.53%
- **Tiny-Imagenet**[8] The Tiny-Imagenet dataset contains 200 categories without the semantic label. Here we randomly choose 20 known classes and 180 unknown classes. The openness in this dataset is 57.36%.

2.2. Evaluation

In open set recognition, we use the Area Under the Receiver Operating Characteristic (AUROC) [6] curve and the Open Set Classification Rate (OSCR) [1] as evaluation metrics. A higher AUROC denotes better classification ability in known and unknown samples. OSCR further distinguishes the accuracy of the known classes in the open set recognition. However, these two evaluation metrics only evaluate the distinction about known classes and do not consider the accuracy of the unknown classes. If we evaluate the generalized open set recognition using the same evaluation metrics, the accuracy of unknown classes will be hidden, which is unreasonable in the real-world scenario. In the experiment, we will use AUROC and OSCR to evaluate the open set recognition task. For the generalized open set recognition task, we use the hierarchical similarity index to evaluate the hierarchical distance.

2.3. Parameters

For CUB-200 [10] and AWA 2 [3] datasets, we add the confidence score α_i and α_j in the Möbius transformation. Confidence scores are available from the dataset. We normalized two confidence scores to make sure that $\alpha_i + \alpha_j = 1$. In Eq. 5, α_i and α_j represent the confidence scores that are provided in the datasets. The portion of relabeling r is 0.1. For other datasets without attribute annotations, we use λ to replace the confidence score. λ_i is sampled from a Beta distribution influenced by the parameter α . When $\alpha = 2$, sampled λ leads to best performance. During training, we set the hyperparameter c_r to 0.5.

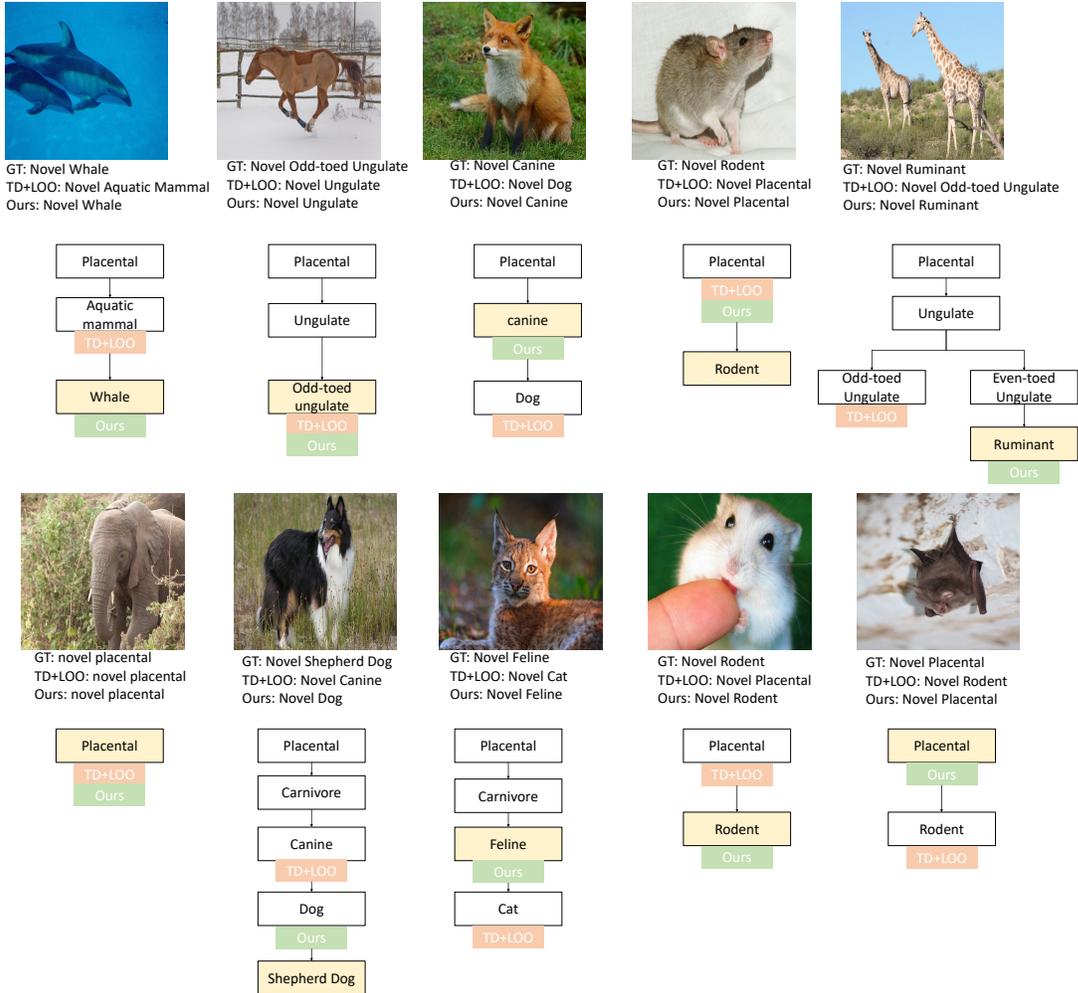


Figure 2. Visualization results of generalized open set recognition in CUB-200 dataset. Yellow boxes are the ground truth. Orange and green boxes denote the prediction of the TD-LOO method [5] and our proposed method, respectively. GT denotes the ground truth node of the image.

Table 1. AUROC results of different classifiers and generating methods in Tiny ImageNet (TI) and SVHN datasets. Mixup denotes random generate.

Method	AUROC @ TI	AUROC @ SVHN
Softmax	57.71	88.62
Euclidean classifier	68.84	93.77
PROSER [11]	69.31	94.30
Hyperbolic classifier	71.71	93.82
Hyperbolic+Mixup	72.38	93.96
Ours	78.19	95.33

2.4. Results

More results for open set recognition. In these experiments, all results are averaged by five random trials. We report ablation study results of open-set recognition. Table 1 illustrates detailed results of open set recognition using dif-

ferent classifiers and generating methods. Our method outperforms comparing methods in SVHN and Tiny ImageNet datasets. Euclidean classifier and hyperbolic classifier do not contain feature generating processes. The PROSER [11] denotes placeholder learning and random feature generating method in the Euclidean space. Mixup means randomly selecting two attribute features and generating new attribute features. Our method uses hyperbolic placeholder learning process and the novel attribute feature generator.

Additional ablation study Table 2 shows ablation study results for novel attribute feature generator. Hyperbolic similarity constrain and Möbius transformation are two components in the generating processes. AUROC results in Tiny ImageNet illustrate that both components work in open set recognition.

In Figure 4, we compare the recognition result using dif-

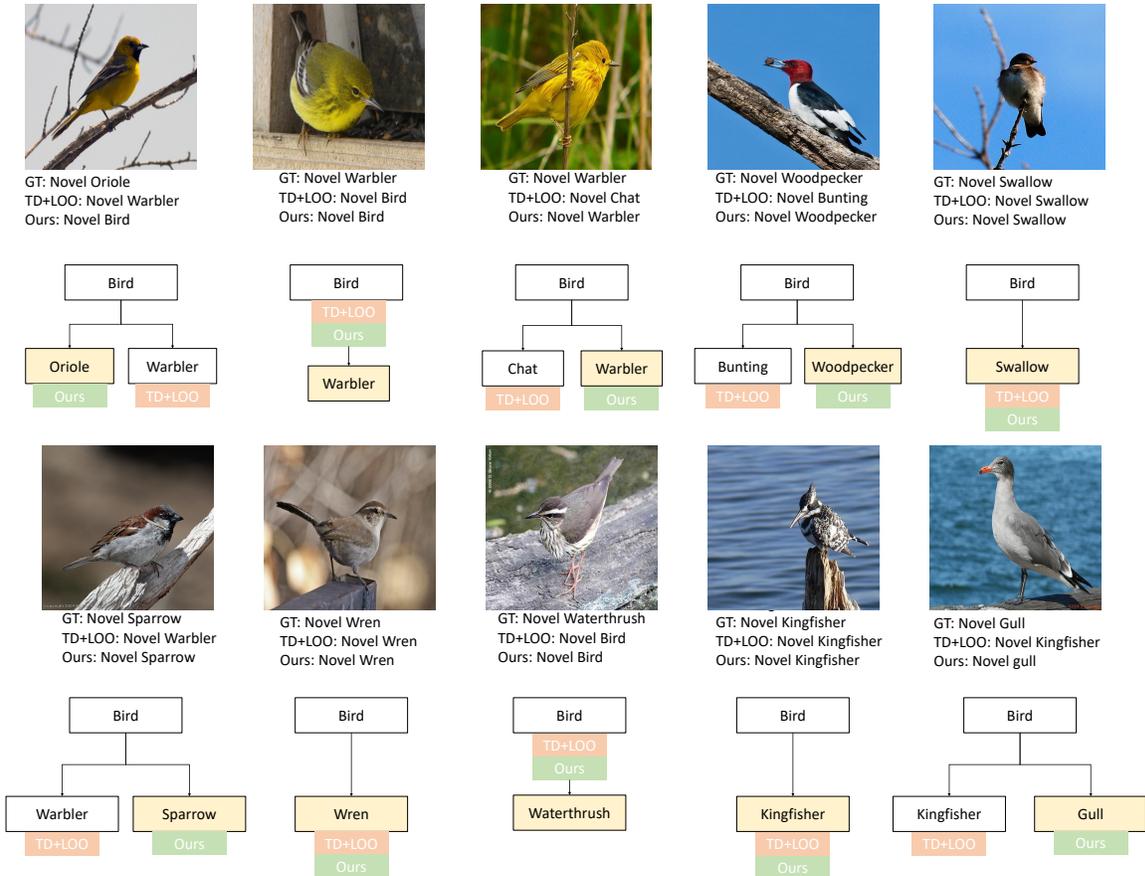


Figure 3. Visualization results of generalized open set recognition in AWA 2 dataset. Yellow boxes are the ground truth. Orange and green boxes denote the prediction of the TD+LOO method [5] and our proposed method, respectively. GT denotes the ground truth node of the image. The LCA node and other ancestor nodes are included in the taxonomy.

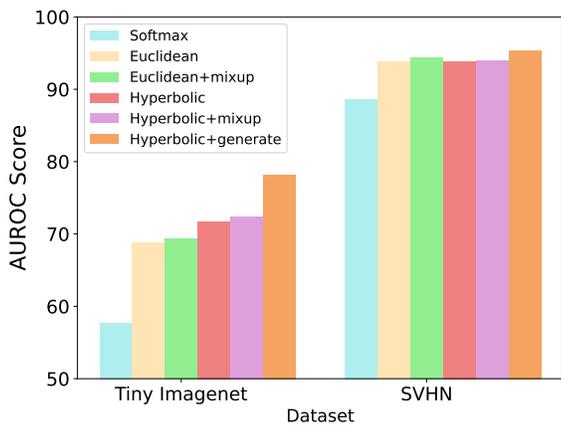


Figure 4. AUROC results (%) v.s. generation methods.

ferent attribute generating strategies. The results show that our generating method surpasses all comparing generating

Table 2. Ablation study of novel attribute feature generator in Tiny Imagenet dataset. Similarity denotes Hyperbolic similarity constrain, and Möbius represents Möbius transformation.

Method	Similarity	Möbius	AUROC@Tiny Imagenet(↑)
Case1			71.71
Case2		✓	73.58
Case3	✓		72.47
Case4	✓	✓	74.34

Table 3. Ablation study for hyperbolic feature mapping

Method	AUROC@Tiny Imagenet(↑)
With mapping	73.34
Without mapping	71.58

strategies. From the result of pure hyperbolic and pure Euclidean, embedding the feature in hyperbolic space also improves the overall performance of the open set recognition.

We provide the ablation study results to demonstrate the effectiveness of hyperbolic feature mapping in Table 3. Hyperbolic feature mapping enhances the AUROC score at Ting Imagenet dataset.

Generalized Open Set Recognition. In the generalized open set recognition experiment, we use attribute annotations and tree-structured side information from CUB-200 and AWA 2 datasets. Figure 2 shows some detailed results about bird taxonomy in CUB-200 dataset. Figure 3 shows some detailed predictions of AWA 2 dataset. Figure 2 and Figure 3 visualize the generalized open set recognition result and the taxonomy. The openness of CUB-200 is greater than the AWA 2 dataset. We find that our prediction result is closer to the ground truth node. In some scenarios, the TD-LOO method misrecognizes samples as the sibling node of the ground truth, which influences the result under the hierarchical similarity index measurement.

References

- [1] Akshay Raj Dhamija, Manuel Günther, and Terrance E Boulton. Reducing network agnostophobia. *arXiv preprint arXiv:1811.04110*, 2018.
- [2] Alex Krizhevsky, Geoffrey Hinton, et al. Learning multiple layers of features from tiny images. 2009.
- [3] Christoph H Lampert, Hannes Nickisch, and Stefan Harmeling. Attribute-based classification for zero-shot visual object categorization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(3):453–465, 2013.
- [4] Yann LeCun. The mnist database of handwritten digits. <http://yann.lecun.com/exdb/mnist/>, 1998.
- [5] Kibok Lee, Kimin Lee, Kyle Min, Yuting Zhang, Jinwoo Shin, and Honglak Lee. Hierarchical novelty detection for visual object recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1034–1042, 2018.
- [6] Lawrence Neal, Matthew Olson, Xiaoli Fern, Weng-Keen Wong, and Fuxin Li. Open set learning with counterfactual images. In *European Conference on Computer Vision*, pages 613–628, 2018.
- [7] Yuval Netzer, Tao Wang, Adam Coates, Alessandro Bisaccho, Bo Wu, and Andrew Y Ng. Reading digits in natural images with unsupervised feature learning. 2011.
- [8] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3):211–252, 2015.
- [9] Abraham A Ungar. Hyperbolic trigonometry and its application in the poincaré ball model of hyperbolic geometry. *Computers & Mathematics with Applications*, 41(1-2):135–147, 2001.
- [10] Catherine Wah, Steve Branson, Peter Welinder, Pietro Perona, and Serge Belongie. The caltech-ucsd birds-200-2011 dataset. 2011.
- [11] Da-Wei Zhou, Han-Jia Ye, and De-Chuan Zhan. Learning placeholders for open-set recognition. In *Conference on Computer Vision and Pattern Recognition*, pages 4401–4410, 2021.