

Distance-based Hyperspherical Classification for Multi-source Open-Set Domain Adaptation Supplementary Material

Silvia Bucci* Francesco Cappio Borlino* Barbara Caputo Tatiana Tommasi
Politecnico di Torino, Italy Italian Institute of Technology

{silvia.bucci, francesco.cappio, barbara.caputo, tatiana.tommasi}@polito.it

1. Qualitative Analysis

We visualize the distribution of source and target data in the feature space (output of the contrastive head) with the t-sne [12] plots in Figure 1. In particular, we focus on the Ar,Pr,Rw \rightarrow Cl case of the Office-Home dataset: the red dots represent the source domain, the blue dots are the known samples of the target domain, and the green dots the unknown ones. We take three snapshots of the data on the hyperspherical embedding: at the beginning when the backbone network is inherited from SupClr [6] pre-trained on ImageNet, immediately before the first *break-point* (i.e. before the application of self-training), and at the end of the training process. By observing the intermediate plot we can state that source balancing and style transfer already favor a good alignment of most of the known (blue) target classes with the respective source known clusters (red). The last plot indicates that self-training further improves the alignment while the unknown samples (green) remain in the regions among the clusters.

Randomly zooming on a known sample (the bike) and on an unknown sample (the speaker) we observe how their position change during training. The first moves from an isolated region where its top five neighbors show high class confusion, towards the correct bike class. The second starts from a neighborhood populated by several samples of classes webcam and fan and finally appears in a different region shared mostly by other instances of the class speaker.

2. Further experiments

Complete results with additional metrics In Table 1 we present the same results of the main paper including also additional metrics: the average class accuracy over known classes OS^* , the accuracy on the unknown class UNK and the average accuracy over all classes OS defined as $OS = \frac{|C_s|}{|C_s|+1} \times OS^* + \frac{1}{|C_s|+1} \times UNK$.

*The authors equally contributed to this work.

Robustness to temperature variation The temperature τ in the contrastive loss (main paper Eq. (1)) is kept fixed to the default value 0.07 as suggested in [11]. We verified experimentally that the results are stable even when tuning τ and remain always higher than ROS (65.3) (see Figure 2).

3. Implementation Details

We implemented HyMOS with an architecture composed of the ResNet-50 [4] backbone that corresponds to the *encoder* and two fully connected layers of dimension 2048 and 128 which define the *contrastive head*. The overall network is trained by minimizing the contrastive loss (see the main paper, Eq. (1)), setting $\tau = 0.07$ as in [11]. Our distance-based classifier lives in the hyperspherical space produced by the model, whose dimension is not constrained by the number of classes. As a consequence, the architecture remains exactly the same for all our experiments.

We initialize the backbone network with the ImageNet pre-trained SupClr model [6] and train HyMOS for 40k iterations with a balanced data mini-batch which contains one sample for each class of every source domain. The learning rate grows from 0 to 0.05 (at iteration 2500) with a linear warm-up schedule, to then decrease back to 0 at the end of training (iteration 40k) through a cosine annealing schedule. We use LARS optimizer [13] with momentum 0.9 and weight decay 10^{-6} . For the first 20k iterations we train only on source data, using target data exclusively for the style transfer based data augmentation for the supervised contrastive learning objective. We then perform an eval step that we call self-training *break-point* in order to start including confident known target samples in the learning objective. We perform *break-point* eval steps every 5K iterations till the end of the training.

For style transfer data augmentation we use the standard VGG19-based AdaIN model with default hyperparameters [5], trained with content data from the available source domains and target samples as style data.

For what concerns the instance transformations, we ap-

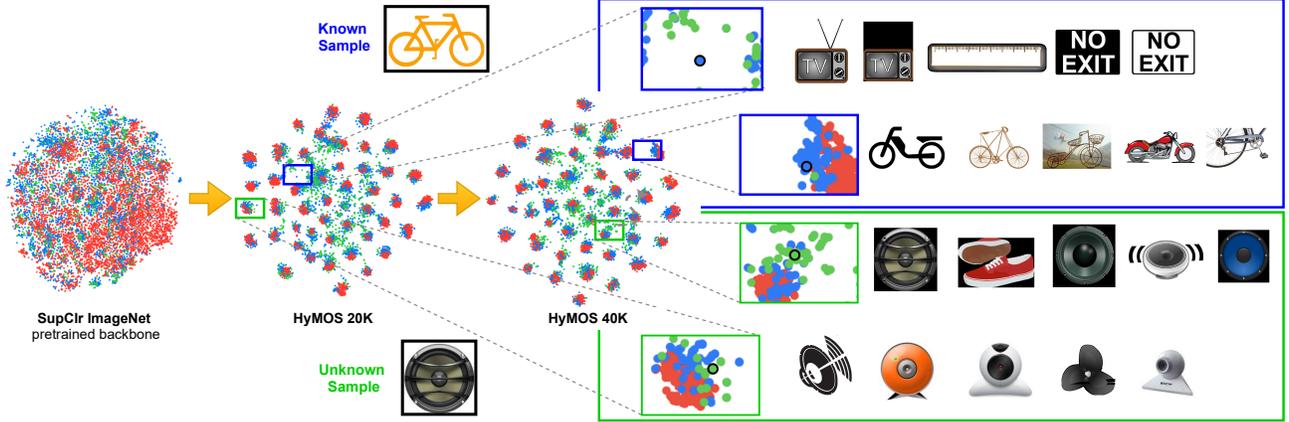


Figure 1. Qualitative analysis on the Ar,Pr,Rw \rightarrow Cl case of the Office-Home dataset. The red dots represent the source domain, the blue dots are the known samples of the target domain, and the green dots are the unknown ones. HyMOS 20k: source balancing and style transfer already favor a good alignment of most of the known target classes with the respective source known cluster. HyMOS 40k: self-training further move the target known samples towards the respective source clusters, while the unknown samples remain in the regions among the clusters. The zooms show how the neighborhood of a known (bike) and unknown (speaker) target samples change during training.

		Office31												DomainNet																											
		D,A \rightarrow W				W,A \rightarrow D				W,D \rightarrow A				Avg.				LP \rightarrow S				LP \rightarrow C				Avg.															
		OS	OS*	UNK	HOS	OS	OS*	UNK	HOS	OS	OS*	UNK	HOS	OS	OS*	UNK	HOS	OS	OS*	UNK	HOS	OS	OS*	UNK	HOS	OS	OS*	UNK	HOS	OS	OS*	UNK	HOS								
Source Combine	Inheritable [7]	69.0	68.1	87.6	76.6	74.7	74.1	85.6	79.5	63.7	62.9	78.9	70.0	69.1	68.4	84.0	75.4	24.9	24.5	60.3	34.8	33.5	33.1	65.6	44.0	29.2	28.8	62.9	39.4	31.7	31.3	77.5	44.5	41.0	40.7	73.6	52.4	36.4	36.0	75.5	48.5
	ROS [11]	82.2	82.3	81.5	81.8	95.3	96.5	68.7	80.1	53.8	52.2	84.9	64.7	77.1	77.0	78.4	75.5	48.0	48.3	26.3	38.1	49.6	49.8	27.6	35.5	48.8	49.1	27.0	36.8	48.0	48.3	26.3	38.1	49.6	49.8	27.6	35.5				
	CMU [3]	96.1	98.7	44.6	61.4	96.2	98.7	47.3	64.0	73.1	74.5	45.4	56.4	88.5	90.6	45.8	60.6	45.6	45.8	22.3	30.0	54.4	54.7	28.7	37.6	50.0	50.3	25.5	33.8	45.6	45.8	22.3	30.0	54.4	54.7	28.7	37.6				
	DANCE [10]	95.9	99.5	23.9	38.5	97.3	100.0	42.6	59.7	78.0	79.6	45.6	58.0	90.4	93.0	37.3	52.0	54.9	55.3	11.1	18.5	59.6	60.1	11.6	19.4	57.3	57.7	11.4	19.0	54.9	55.3	11.1	18.5	59.6	60.1	11.6	19.4				
Multi-Source	MOSDANET [9]	97.7	99.4	43.5	60.5	97.0	99.0	55.9	71.5	80.9	81.5	67.6	73.9	91.9	93.3	55.7	68.6	30.2	29.9	60.2	40.0	31.8	31.6	51.8	39.3	31.0	30.8	56.0	39.6	30.2	29.9	60.2	40.0	31.8	31.6	51.8	39.3				
	HyMOS	96.1	96.6	84.6	90.2	96.7	97.3	83.6	89.9	49.6	48.0	83.1	60.8	80.8	80.6	83.8	80.3	43.6	43.2	86.0	57.5	47.8	47.4	85.5	61.0	45.7	45.3	85.8	59.3	43.6	43.2	86.0	57.5	47.8	47.4	85.5	61.0				

		Ar,Pr,Cl \rightarrow Rw				Ar,Pr,Rw \rightarrow Cl				Cl,Pr,Rw \rightarrow Ar				Cl,Ar,Rw \rightarrow Pr				Avg.							
		OS	OS*	UNK	HOS	OS	OS*	UNK	HOS	OS	OS*	UNK	HOS												
Source Combine	Inheritable [7]	58.6	58.4	68.9	63.2	44.3	43.7	66.5	52.6	36.4	35.5	77.6	48.7	58.6	58.5	63.3	60.7	49.5	49.1	69.1	56.3	58.6	58.4	68.9	63.2
	ROS [1]	69.9	69.8	76.9	73.0	57.1	57.1	57.6	57.3	57.5	57.2	66.7	61.6	70.3	70.3	68.0	69.1	63.7	63.6	67.3	65.3	69.9	69.8	76.9	73.0
	CMU [3]	62.9	62.5	81.5	70.8	35.8	34.6	89.9	50.0	44.6	43.7	87.0	58.1	60.6	60.1	81.7	69.3	51.0	50.2	85.0	62.1	62.9	62.5	81.5	70.8
	DANCE [10]	83.9	85.6	4.5	12.4	66.8	68.0	9.2	16.1	72.7	74.1	10.7	18.6	85.1	86.7	13.4	22.9	77.1	78.6	9.4	17.5	83.9	85.6	4.5	12.4
Multi-Source	PGL [8]	83.4	84.6	26.2	40.0	62.0	63.0	21.0	31.5	69.5	70.6	20.5	31.8	82.6	83.8	28.2	42.2	74.4	75.5	24.0	36.4	83.4	84.6	26.2	40.0
	MOSDANET [9]	78.4	79.4	55.0	65.0	67.5	68.1	40.9	51.1	61.0	61.3	48.7	54.3	81.1	82.2	55.0	65.9	72.0	72.8	49.9	59.1	78.4	79.4	55.0	65.0
	HyMOS	69.5	69.4	72.7	71.0	52.5	51.7	86.0	64.6	50.1	49.4	84.1	62.2	71.5	71.5	70.6	71.1	60.9	60.5	78.4	67.2	69.5	69.4	72.7	71.0

Table 1. Accuracy (%) averaged over three runs for each method on the Office31, DomainNet and Office-Home datasets.

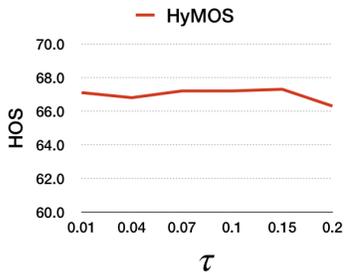


Figure 2. Sensitivity analysis for the temperature value τ on Office-Home.

plied the same data augmentations originally proposed for SimClr [2], extending them with style transfer. Specifically, we used random resized crop with scale in $\{0.08, 1\}$ and random horizontal flip. The style transfer is applied with probability $p = 0.5$ on the source images, while the remaining not-stylized images are transformed via color jittering with probability $p = 0.8$ and grayscale with probability $p = 0.2$.

The final evaluation procedure of HyMOS is summa-

Algorithm 1 HyMOS evaluation procedure

Input: \mathcal{T} ; trained Enc and $Proj$

Output: Predictions on \mathcal{T}

procedure FINALEVAL()

$\alpha \leftarrow$ (main paper Eq. (4))

for each x_t **in** \mathcal{T} **do**

$z^t = Proj(Enc(x^t))$

$h_{y^s} \leftarrow$ nearest prototype to z^t

if $d_{h_{y^s}}(z^t) < \alpha$ **then**

$\hat{y}^t = y^s$

else

$\hat{y}^t = \text{unknown}$

procedure MAIN()

$finalEval()$

rized in Algorithm 1.

References

- [1] Silvia Bucci, Mohammad Reza Loghmani, and Tatiana Tommasi. On the effectiveness of image rotation for open set

- domain adaptation. In *ECCV*, 2020.
- [2] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *ICML*, 2020.
 - [3] Bo Fu, Zhangjie Cao, Mingsheng Long, and Jianmin Wang. Learning to detect open classes for universal domain adaptation. In *ECCV*, 2020.
 - [4] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *CVPR*, 2016.
 - [5] Xun Huang and Serge Belongie. Arbitrary style transfer in real-time with adaptive instance normalization. In *ICCV*, 2017.
 - [6] Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschiot, Ce Liu, and Dilip Krishnan. Supervised contrastive learning. In *NeurIPS*, 2020.
 - [7] Jogendra Nath Kundu, Naveen Venkat, Ambareesh Revanur, R Venkatesh Babu, et al. Towards inheritable models for open-set domain adaptation. In *CVPR*, 2020.
 - [8] Yadan Luo, Zijian Wang, Zi Huang, and Mahsa Baktashmotlagh. Progressive graph learning for open-set domain adaptation. In *ICML*, 2020.
 - [9] Sayan Rakshit, Dipesh Tamboli, Pragati Shuddhodhan Meshram, Biplob Banerjee, Gemma Roig, and Subhasis Chaudhuri. Multi-source open-set deep adversarial domain adaptation. In *ECCV*, 2020.
 - [10] Kuniaki Saito, Donghyun Kim, Stan Sclaroff, and Kate Saenko. Universal domain adaptation through self-supervision. In *NeurIPS*, 2020.
 - [11] Jihoon Tack, Sangwoo Mo, Jongheon Jeong, and Jinwoo Shin. Csi: Novelty detection via contrastive learning on distributionally shifted instances. In *NeurIPS*, 2020.
 - [12] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008.
 - [13] Yang You, Igor Gitman, and Boris Ginsburg. Large batch training of convolutional networks. *arXiv preprint arXiv:1708.03888*, 2017.