ONE RING TO BRING THEM ALL: MODEL ADAPTATION UNDER DOMAIN AND CATEGORY SHIFT

Anonymous authors
Paper under double-blind review

ABSTRACT

In this paper, we investigate model adaptation under domain and category shift, where the final goal is to achieve *Source-free Open-partial Domain Adaptation* (SF-OPDA), which addresses the situation where there exist both domain and category shifts between source and target domains. Under the SF-OPDA setting, the model cannot access source data anymore during target adaptation, which aims to address data privacy concerns. We propose a novel training scheme to learn a (n+1)-way classifier to predict the n source classes and the unknown class, where samples of only known source categories are available for training. Furthermore, for target adaptation, we simply adopt a weighted entropy minimization to adapt the source pretrained model to the unlabeled target domain without source data. In experiments, we show: **1)** After source training, the resulting source model has strong open classes detection ability for *open-set single domain generalization*; **2)** After target adaptation, our method surpasses current UNDA approaches which demand source data during adaptation. The versatility to several different tasks strongly proves the efficacy and generalization ability of our method. **3)** When augmented with a closed-set domain adaptation approach during target adaptation, our source-free method further outperforms the current state-of-the-art UNDA method by 2.5%, 7.2% and 13% on Office-31, Office-Home and VisDA respectively.

1 INTRODUCTION

Modern deep learning models excel at close-set recognition tasks across various computer vision application areas. However, there are several inevitable obstacles lying on the path to deploying those methods to the challenging real world environments. As there may be 1) some unseen categories in practical scenarios, or 2) distributional shift between training and testing data. The first problem is usually defined as *open-set recognition* (OSR) (Chen et al., 2020; Ge et al., 2017; Neal et al., 2018a; Sun et al., 2020; Zhang et al., 2020; Shu et al., 2020; Vaze et al., 2022) where the model should be able to distinguish samples as coming from unseen categories. The second problem is mostly investigated in the *domain generalization* (DG) (Shi et al., 2022; Robey et al., 2021; Vedantam et al., 2021; Gulrajani & Lopez-Paz, 2021; Wang et al., 2021) and *domain adaptation* (DA) community (Long et al., 2018a; 2015; 2016; Tzeng et al., 2017; Zhang et al., 2019; Cicek & Soatto, 2019; Liang et al., 2021a; Deng et al., 2019; Tang et al., 2020; Cui et al., 2020). DG aims to tackle the domain shift problem in the absence of target domains, while DA seeks to transfer knowledge from labeled source domains to unlabeled target domains with training on them with utilizing both labeled source and unlabeled target data, there is distribution/domain shift between source and target domains. In recent years, several works introduce open-set recognition into DG and DA, which are formalized as *open domain generalization* (ODG) (Shu et al., 2021; Zhu & Li, 2022), *open-set domain adaptation* (OSDA) (Saito et al., 2018b; Bucci et al., 2020; Liu et al., 2019; Pan et al., 2020; Jing et al., 2021; Feng et al., 2019; 2021) and *universal domain adaptation* (UNDA) (Fu et al., 2020; Li et al., 2021a; You et al., 2019; Saito et al., 2020; Saito & Saenko, 2021; Liang et al., 2021b), respectively.

The various settings described above are summarized in Tab. 1. Usually one method tailored for a specific setting in Tab. 1 does not work well under a different setting. Most existing works in *Open-set Recognition* are computationally demanding, either requiring the generation of unknown categories (Neal et al., 2018a) or conducting additional learning (Kong & Ramanan, 2021; Sun et al., 2020; Chen et al., 2020). Additionally, those methods are likely to suffer from performance degradation if test data are from different distributions. The recent CrossMatch (Zhu & Li, 2022) tackles

Table 1: Related setting. $\mathcal{C}_s$ and $\mathcal{C}_t$ denote label set of source and target domain (for evaluation), $\mathcal{P}_s$ and $\mathcal{P}_t$ denote source and target distribution, transductive means model can be trained on target data.

Task	$\mathcal{C}_s = \mathcal{C}_t$	$\mathcal{P}_s = \mathcal{P}_t$	Transductive
Open-set Recognition (OSR)	✗	✓	✗
Domain Generalization (DG)	✓	✗	✗
Open Domain Generalization (ODG)	✗	✗	✗
Domain Adaptation (DA)	✓	✗	✓
Open-partial Domain adaptation (OPDA)	✗	✗	✓

Open-set Single Domain Generalization problem. It proposes to use multiple open class detectors which are put on top of existing single domain generalization methods, and it achieves good results at the expense of introducing multiple open-set detectors and auxiliary unknown sample generation. For *open-partial domain adaptation*, most works are based on an explicitly designed unknown-sample rejection module, which typically requires various hyper-parameters. More importantly, those *UNDA* methods all require access to source data during target adaptation, which is infeasible if having data privacy issues and deployed on devices of low computation capacity.

In this paper we investigate how to detect open classes efficiently under the domain shift. Thus, a question arises, how to build a model training from only known categories aiming to learn to distinguish samples of unknown categories? Since we have no access to unknown class data, we can only use the known class data to train this classifier. We hypothesize that the closest (most similar) class to any known class can be an unknown class. Given the open-endedness of the unknown class this is a reasonable assumption. This hypothesis allows us to train the classifier, enforcing the most probable class to be the ground truth class, and the runner-up class to be the background class for all source data. This is achieved by introducing an extra category in the classifier which represents the unknown classes, during training on samples of known categories (yielding a (n + 1)-way classifier where n is the number of known classes), the classifier is expected to output the largest score for the ground truth class, and the second-largest score for unknown class. This way, the model can learn to reject samples of unknown categories by only training with known classes. The resulting model training on source data can be directly deployed to *open-set single domain generalization*, in other words, it can detect open class efficiently whether there is domain shift or not.

Furthermore, our source model with strong capacity to distinguish unknown categories can be easily adapted to target domain without access to source data under the challenging *source-free open-partial domain adaptation* setting, where both source and target domains have their private classes. We propose to simply use a weighted entropy minimization to achieve the adaptation.

We summarize our contributions as below:

- We propose a simple method called *OneRing*, which excels at recognising open class (even with domain shift) after source training, thus it can be directly deployed to *open-set single domain generalization* (OS-SDG) and *open-set recognition* (OSR).
- We can easily adapt the source model to target domain by using weighted entropy minimization under *source-free open-partial domain adaptation* setting (SF-OPDA).
- In experiments, we show our method is on par with or outperform current state-of-the-art approaches on several benchmarks for various different tasks, which proves the efficacy and generalization ability of our method. Augmented with a close-set DA approach, our source-free method surpasses current open-partial domain adaptation methods by a significant margin.

2 Related Works

Open-set Recognition. *Open-set recognition* (OSR) aims to recognize samples of unknown categories which do not exist in the training set. Several recent methods in OSR do not utilize extra data for training. OpenHybrid (Zhang et al., 2020) introduces a flow-based density estimation module, and ARPL (Chen et al., 2020; 2021) proposes to learn a reciprocal point per category, which is intuitively regarded as the farthest point from the corresponding feature group. More recently (Vaze et al., 2022) shows that actually OSR performance is enhanced when improving the model performance on the training set, for example by using improved data augmentation and other training tricks. In this paper, we propose a simple model training directly with two cross entropy losses without either auxiliary data or an extra learning process. Our proposed OneRing classifier shares similarity with Proser (Zhou et al., 2021a), which aims to assign the second-largest logit to the unknown classes. However, Proser is much more complex compared to ours: it first trains a good $|C_s|$-way close-set classifier and then augment this classifier to $|C_s| + C$-way, and retrain; Further, it needs to synthesize novel samples for training the $|C_s| + C$-way classifier; And they also need to calibrate the output of the dummy classifier over the extra validation set by ensuring 95% of validation data are recognized as known. While in this paper, we directly train the $|C_s| + 1$-way classifier with a simple objective; Another main difference is that they only address open-set recognition, while in our paper we also consider the domain shift, *i.e.*, the challenging source-free open-partial domain adaptation.

Domain Generalization. In *Domain Generalization* (DG), a model is typically trained on multiple labeled source domains. It is expected to have good generalization ability on unseen target domains with which domain shift exists. A typical solution for domain generalization is to learn domain invariant features, which can be achieved by meta learning (Li et al., 2018; 2019; Dou et al., 2019) or additional data generation (Zhou et al., 2021b; 2020). In recent years, there are several DG works that only use a single source domain. This setting is known as *single domain generalization* (SDG) (Qiao et al., 2020; Wang et al., 2021; Fan et al., 2021; Li et al., 2021b). While most of those methods only consider the situation where source and target domains share the same label space, *Open Domain Generalization* (ODG) (Shu et al., 2021) is recently proposed to deal with the problem where the target domain contains open classes. More recently, CrossMatch (Zhu & Li, 2022) introduces an even more challenging setting called *Open-set Single Domain Generalization* (OS-SDG) which only relies on one source and where the target domains contains unknown categories. CrossMatch is built on a complex network model and needs to synthesize samples of unknown categories. It also applies entropy-based unknown class rejection with a manually set threshold. In this paper, our simple source trained model can be directly deployed to OS-SDG task and gets surprisingly decent results.

Domain Adaptation. Early methods to tackle *domain adaptation* (DA) conduct feature alignment (Long et al., 2015; Sun et al., 2016; Tzeng et al., 2014) to eliminate the domain shift. DANN (Ganin et al., 2016), CDAN (Long et al., 2018b) and DIRT-T (Shu et al., 2018) further resort to adversarial training to learn domain invariable features. Similarly, (Lee et al., 2019; Lu et al., 2020; Saito et al., 2018a) are based on multiple classifier discrepancy to achieve alignment between domains. Other methods like SRDC (Tang et al., 2020), CST (Liu et al., 2021) address domain shift from the perspective of either clustering or improved pseudo labeling. And there are also methods considering category shift source and target domains. They can be grouped into *partial-set DA* (Cao et al., 2018; 2019; Liang et al., 2020b), *open-set DA* (Panareda Busto & Gall, 2017; Saito et al., 2018b; Liu et al., 2019; Bucci et al., 2020) and *universal DA* (You et al., 2019; Li et al., 2021a; Saito & Saenko, 2021; Fu et al., 2020; Saito et al., 2020) depending on the intersection degree of source and target label space. OVANet (Saito & Saenko, 2021) is a universal DA method. It trains extra n binary classifiers with hard negative classifier sampling to reject unknown samples, OVANet needs to check the normal classifier head and the corresponding binary classifier for the final prediction. While in this paper, we simply train a $n + 1$-way classifier with normal cross entropy, and the final prediction is directly provided by the classifier.

Source-free Domain Adaptation. Recently, several works address *source-free domain adaptation* (SFDA), where a source pretrained model is adapted to target without source data. SHOT (Liang et al., 2020a) proposes to use mutual information maximization along with pseudo labeling. BAIT (Yang et al., 2020) adapts MCD (Saito et al., 2018a) to source-free setting. 3C-GAN (Li et al., 2020) resorts to fake target-style images generation. HCL (Huang et al., 2021) conducts Instance Discrimination (Wu et al., 2018) over different historical models to cluster features, with the companion of pseudo labeling. A^2Net (Xia et al., 2021) learns extra classifier specifically for the target domain and introduce a category-wise matching module for feature clustering. G-SFDA (Yang et al., 2021b) and NRC (Yang et al., 2021a) are all based on neighborhood clustering through local prediction consistency. AaD (Yang et al., 2022) further treats SFDA as a typical unsupervised clustering problem and proposes to optimize an upperbound of a clustering objective. Beyond close-set DA, FS (Kundu et al., 2020b) and USFDA (Kundu et al., 2020a), which are for *source-free open-set and open-partial DA* respectively. However, they both synthesize extra training samples of unknown categories, which help to detect the open classes. OSHT Feng et al. (2021) tackles source-free open-set DA, which adopts pseudo labeling for adaptation and entropy-based metric to reject open classes. UMAD Liang et al. (2021b) is for source-free universal DA, it proposes an informative consistency score to detect open class, then adopts mutual information for source-free adaptation.. In this paper, we show that

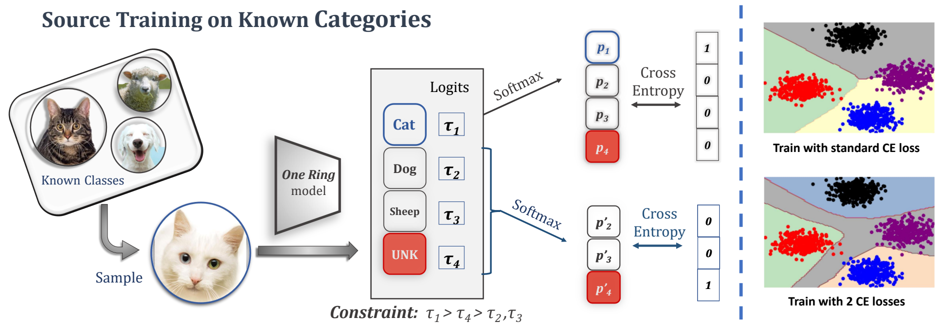

Figure 1: (**Left**) Illustration of training *OneRing* model on source data with only known categories. (**Right**) Toy Example, the decision boundaries and prediction regions (*colorized randomly*) after training on 3 known classes with $(3+1)$-way classifier. Purple points are from unknown category.

our source pretrained model can be adapted to the target domain easily by simply minimizing entropy to achieve *source-free open-partial DA*.

3 Method

3.1 Preliminary

In this paper, we divide data samples into two groups/domains: the labeled source domain with N_s samples as $\mathcal{D}_s = \{(x_i^s, y_i^s)\}_{i=1}^{N_s}$ on which the model will be first trained, and the unlabeled target domain with N_t samples as $\mathcal{D}_t = \{x_i^t\}_{i=1}^{N_t}$. $\mathcal{D}_t$ is used for evaluation. We denote $\mathcal{C}_s$ and $\mathcal{C}_t$ as the label set of the source and target domain, and $\mathcal{P}_s$ and $\mathcal{P}_t$ as the distribution of source and target data respectively. In this paper, we consider three different tasks that vary in the relation between source and target domain data: 1) ***Open-set Recognition***[1] **(OSR)** where the model is only trained on the source domain, and directly tested on the target domain which contains some unknown categories but without domain shift ($\mathcal{C}_s \subset \mathcal{C}_t, \mathcal{P}_s = \mathcal{P}_t$, *inductive*); 2) ***Open-set Single Domain Generalization*** **(Zhu & Li, 2022) (OS-SDG)** which is similarly to OSR trained on a single source domain, however here there exists a domain shift between source and target domains ($\mathcal{C}_s \subset \mathcal{C}_t, \mathcal{P}_s \neq \mathcal{P}_t$, *inductive*); 3) ***Source-free open-partial domain adaptation*** **(SF-OPDA)** which is similar to OS-SDG, here the source model has to adapt to the target domain without access to any source data and both domains have private categories ($\mathcal{C}_s \cap \mathcal{C}_t \neq \emptyset / \mathcal{C}_s / \mathcal{C}_t, \mathcal{P}_s \neq \mathcal{P}_t$, *transductive*). For these settings, we use the same network model containing two parts: a feature extractor f and a classifier head g.

3.2 Source Training: One Ring to Find Unknown Categories

The first stage is to train a model on the labeled source domain which has $|\mathcal{C}_s|$ categories. We expect the resulting model to have the ability to detect unknown categories which do not exist in the source data. To achieve this, we build a classifier head as a $(|\mathcal{C}_s| + \mathbf{1})$-way classifier, where the additional dimension aims to distinguish unknown categories. Then the following problem arises: how to train a $(|\mathcal{C}_s| + 1)$-way classifier without any sample from the last/unknown category? Note, if only training with the normal cross entropy (CE) loss on the source data, the model cannot directly give prediction to unknown categories.

As mentioned in Sec. 1, we hypothesize that any non-ground-truth category could be regarded as unknown categories. This hypothesis gives us a feasible solution to train a open-set classifier without actually accessing open classes. Specifically, we propose to use a simple variant of cross entropy loss with only samples of known categories to train the $(|\mathcal{C}_s| + 1)$-way classifier, which has 2 properties: 1) The largest output logit of the source samples corresponds to the ground truth class and 2) The second-largest output logit of source samples will be the unknown class ($(|\mathcal{C}_s| + 1)$-th class in classifier). This way, the model is expected to detect samples of unknown categories even without training on them. The proposed objective to achieve it is formalized as follows:

$$\mathcal{L}_{source} = \mathbb{E}_{x_i \sim \mathcal{D}_s}[\mathcal{L}_{ce}(p(x_i), y_i) + \mathcal{L}_{ce}(\hat{p}(x_i), \hat{y}_i)] \tag{1}$$

[1]Results for OSR are in the appendix, only aiming to show the generalization ability of our method.

where $p(x_i) = g(f(x_i)) \in \mathbb{R}^{|C_s|+1}$ is the output vector of the $(|\mathcal{C}_s| + 1)$-way classifier, while $\hat{p}(x_i) \in \mathbb{R}^{|C_s|}$ is the output vector removing the dimension corresponding to the ground truth class, and $\hat{y}_i \in \mathbb{R}^{|C_s|}$ is a one-hot label with unknown class as ground truth label. As illustrated in Fig. 1 (*right*), if we have a sample x_i belonging to the ***first*** class, the first CE loss in Eq. 1 is the typical CE loss on $p(x_i)$ with ground truth label, $\hat{p}(x_i)$ is produced by removing the ***first*** dimension and the second CE loss is applied on $\hat{p}(x_i)$ with unknown (last) category as label.

We adopt a toy example to illustrate it. As shown in upper part of Fig. 1 (*right*), we generate isotropic Gaussian blobs with 4 categories, where the last one is treated as the unknown category (in *Purple*) and others as known classes (thus $|\mathcal{C}_s| = 3$). We first train the $(|\mathcal{C}_s|+1)$-way classifier which contains 4 linear layers with the normal cross entropy loss on samples of known categories, and then evaluate it on all classes. Upper part of Fig. 1 (*right*) shows that the samples of the unknown category (*Purple*) are misclassified as there are only 3 prediction regions for 3 known categories. As shown in lower part of Fig. 1 (*right*) that there are 4 prediction regions (3 known + 1 unknown categories), after training on 2 CE losses the classifier can detect samples of unknown category which is unseen before. We attach a demo video to show the difference between training the $(|\mathcal{C}_s| + 1)$-way classifier with only standard CE loss and those 2 CE losses.

An intuitive understanding of the proposed method is that, we can split the $(|\mathcal{C}_s|+1)$-way classification into 2 levels: 1) if we check the prediction $p(x_i)$ we would say x_i has to belong to category y_i; 2) if we check the prediction $\hat{p}(x_i)$ we would say that x_i is impossible to belong to all other categories except the potential unknown categories. Since in Eq. 1 the output score of unknown category (last dimension) will always rule other non-ground-truth categories, we call the last dimension of the classifier head as *OneRing* dimension and our model as *OneRing*. In the experimental section, we will show that our *OneRing* model trained on source data can be directly deployed to open-set recognition and open-set single domain generalization.

3.3 Target Adaptation: One Ring to Bind All Categories without the Source

Our source-pretrained *OneRing* model is empowered with the ability to recognition unknown classes in the target domain. We further posit that it can easily be adapted to target domains where domain shift and unknown categories exist. The key part is to rectify the wrong predictions due to the domain shift. We propose to simply use entropy minimization, which is widely used in DA (Shu et al., 2018; Long et al., 2018b; Liang et al., 2020a; Saito et al., 2020; Saito & Saenko, 2021), to achieve adaptation with only a slight but indispensable modification:

$$\mathcal{L}_{target} = \frac{bs}{\hat{n}_{k_{all}}} \mathbb{E}_{\bar{y}_i \in \mathcal{C}_s} \mathcal{L}_{ent}(p(x_i)) + \frac{bs}{\hat{n}_{u_{all}}} \mathbb{E}_{\bar{y}_i \in \mathcal{C}_u} \mathcal{L}_{ent}(p(x_i)) \tag{2}$$

which is computed in the **mini-batch** (bs denotes batch size), and $\bar{y}_i$ is the predicted label, $\hat{n}_{k_{all}}$ is the number of samples in the *whole dataset* which are predicted as *known* category $\mathcal{C}_s$, $\hat{n}_{u_{all}}$ is the number of those predicted as *unknown* category $\mathcal{C}_u$ also in the *whole dataset*. Here $\frac{bs}{\hat{n}_{k_{all}}} = \frac{N_t}{\hat{n}_{k_{all}}} \times \frac{bs}{N_t}$ (similar for $\frac{bs}{\hat{n}_{u_{all}}}$), where $N_t = \hat{n}_{k_{all}} + \hat{n}_{u_{all}}$ and $\frac{N_t}{\hat{n}_{k_{all}}}$ is the reciprocal of the known/unknown category ratio (a prior information according to the predictions). The reason to deploy these weights is to balance the two entropy terms, and $\frac{bs}{N_t}$ is a scale factor. [2] With this simple objective, the source model can be adapted to the target domain under domain and category shift efficiently.

Augmented with Attracting-and-Dispersing. Since our *OneRing* method can equip models to efficiently detect unknown classes, it can be used as a baseline to be combined with methods in close-set source-free DA. Here we integrate our method with a simple state-of-the-art SFDA method Attracting-and-Dispersing (AaD) (Yang et al., 2022), note AaD can not directly tackle the open-partial domain adaptation setting. AaD has an objective with only 2 dot product terms: $\mathcal{L}_{dis}$ for discriminability and $\mathcal{L}_{div}$ for diversity, more details can be found in AaD paper. The resulting objective is:

$$\mathcal{L}_{target+} = \frac{bs}{\hat{n}_{k_{all}}} \mathbb{E}_{\bar{y}_i \in \mathcal{C}_s} [\mathcal{L}_{ent}(p(x_i)) + \mathcal{L}_{dis} + \mathcal{L}_{div}] + \frac{bs}{\hat{n}_{u_{all}}} \mathbb{E}_{\bar{y}_i \in \mathcal{C}_u} [\mathcal{L}_{ent}(p(x_i)) + \mathcal{L}_{dis}] \tag{3}$$

[2] Instead of using the predictions over the whole dataset to compute known-unknown ratio, we can also use prediction of current mini-batch for approximation (thus N_t will be replaced by bs, and similar for $\hat{n}_{u_{all}}$ and $\hat{n}_{k_{all}}$), in the experiment we empirically found these two different estimation manners lead to almost the same results.

Table 2: Accuracy (%) on **Office-31** dataset using ResNet-18. **Open-set Single Domain Generalization** where $|\mathcal{C}_s| = 10$, $|\mathcal{C}_t| = 21$, $|\mathcal{C}_s \cap \mathcal{C}_t| = 10$. All other results are from (Zhu & Li, 2022).

Metric	ERM	+CM (Zhu & Li, 2022)	ADA	+CM (Zhu & Li, 2022)	MEADA	+CM (Zhu & Li, 2022)	***OneRing*-S**
Acc	79.8	78.3	80.1	78.6	**80.3**	79.0	67.3
UNK	27.0	37.6	25.2	34.5	25.1	41.1	**77.0**
OS*	85.1	82.4	85.6	83.0	**85.8**	82.8	66.3
H	40.7	51.1	38.7	48.5	38.6	<u>54.7</u>	**71.3**

where we do not deploy the diversity term for samples predicted as an unknown class since there is only one single unknown class.

4 EXPERIMENTS

Here we provide quantitative results and analyses related to open-set single domain generalization and source-free open-partial domain adaptation, *and also open-set recognition in the appendix*.

4.1 DATASETS

Open-set Single Domain Generalization. For OS-SDG the model is trained on source data and evaluated on target data containing both known and unknown categories, but here domain shift exists between source and target domains. We use the following benchmarks just as CrossMatch (Zhu & Li, 2022): 1) **Office31** (Saenko et al., 2010) has 31 classes with 3 different domains: amazon (A), dslr (D) and webcam (W). The 10 classes shared by Office-31 and Caltech-256 (Gong et al., 2012) will be used as source categories. Then the last 11 classes in alphabetical order along with the 10 source categories will be used as target categories. Following CrossMatch, we only adopt A as the source domain, since D and W contain a relatively small amount of samples. 2) **Office-Home** (Venkateswara et al., 2017) has 4 domains: Artistic (A), Clip Art (C), Product (P), and Real-World (R) with 65 categories. In alphabetic order, the first 15 classes are adopted as source categories. And all classes are used as target categories. 3) **PACS** (Li et al., 2017) has 4 domains: Art Paint, Cartoon, Sketch, and Photo. It has 7 categories. Of these, 4 classes (dog, elephant, giraffe, and guitar) will be used as source categories and all classes will be used as target categories. For Office-Home and PACS, the model will be trained on one domain and evaluated on all remaining domains.

Source-free Univeral Domain Adaptation. For SF-OPDA, the model is trained on the source domain first, then adapted to the target domain without access to any source data. Here both the source and target domains have their private categories and the target domain has some unknown categories. We evaluate our method on several benchmarks following the same setting as previous work in UNDA (You et al., 2019; Saito et al., 2020; Saito & Saenko, 2021): 1) **Office-31** shares 10 classes with Caltech-256 which will be used as the common categories. Then the next 10 classes in alphabetical order will be source private, and the remaining classes will be target private. 2) **Office-Home** The first 10 classes in alphabetical order are shared between domains, and the next 5 categories will be source private, and the remaining classes are target private. 3) **VisDA** (VisDA-C 2017) (Peng et al., 2017) The 6 classes out of 12 classes will be the shared categories, and source and target domain both have 3 private classes. 4) **DomainNet** (Peng et al., 2019) DomainNet is one of the largest domain adaptation benchmarks with around 0.6 million images. Following previous works, we will use 3 domains: Painting (P), Real (R), and Sketch (S). We will use the first 150 classes as shared categories, the next 50 classes are source private and the remaining 145 as target private. The number of source, target and shared categories is described in the title of each Table.

4.2 MODEL DETAILS AND EVALUATION

For all setting, we directly adopt the prediction of our *OneRing* model, without using any extra process for unknown category detection. To ensure fair comparison with previous methods, our method is based on the original code released by UNDA method OVANet (Saito & Saenko, 2021) (modified for OS-SDG and SF-OPDA).

For OS-SDG, we train our *OneRing* model on source with Eq. 1 and directly evaluate on the target. For SF-OPDA, after finishing source training with Eq. 1, we will adapt the source pretrained model to target domain without using source data. Only on the very large DomainNet under SF-OPDA setting we found that our method had difficulties converging. Therefore, we applied a two-phase training on

Table 3: Accuracy (%) on **Office-Home** using ResNet-18. **Open-set Single Domain Generalization** where $|\mathcal{C}_s| = 25$, $|\mathcal{C}_t| = 65$, $|\mathcal{C}_s \cap \mathcal{C}_t| = 25$. Other results are copied from (Zhu & Li, 2022).

	Artistic			*Clipart*			*Product*			*Real World*			**Average**		
	OS*	UNK	H	OS*	UNK	H	OS*	UNK	H	OS*	UNK	H	OS*	UNK	H
ERM (Koltchinskii, 2011)	68.4	20.5	31.1	66.8	24.7	35.8	62.8	26.3	36.3	69.5	23.2	33.9	66.9	23.7	34.3
ERM+CM (Zhu & Li, 2022)	66.5	48.6	52.9	64.8	42.0	50.5	59.2	40.9	47.3	69.4	43.7	52.6	65.0	43.8	50.8
ADA (Volpi et al., 2018)	**71.4**	22.1	32.9	**67.4**	31.2	42.1	**62.9**	24.6	34.7	**69.9**	23.9	34.9	**67.9**	25.4	36.2
ADA+CM (Zhu & Li, 2022)	67.5	39.6	46.7	64.1	40.7	49.3	59.9	40.7	47.5	68.5	40.8	50.5	65.0	40.4	48.5
MEADA (Zhao et al., 2020)	**71.4**	22.4	33.3	66.5	31.3	42.1	62.8	25.6	35.7	**69.9**	23.7	34.7	67.6	25.7	36.4
MEADA+CM (Zhu & Li, 2022)	66.6	45.3	52.3	64.3	37.8	48.9	59.7	37.7	45.3	68.8	41.3	50.8	64.9	40.5	49.6
***OneRing*-S**	58.9	**68.2**	**63.2**	57.6	**69.9**	**63.2**	52.0	**69.0**	**59.3**	58.9	**69.0**	**63.6**	56.9	**69.0**	**62.3**

Table 4: Accuracy (%) on **PACS** dataset using ResNet-18. **Open-set Single Domain Generalization** where $|\mathcal{C}_s| = 4$, $|\mathcal{C}_t| = 7$, $|\mathcal{C}_s \cap \mathcal{C}_t| = 4$. Other results are copied from (Zhu & Li, 2022).

	Art Paint			*Cartoon*			*Sketch*			*Photo*			**Average**		
	OS*	UNK	H	OS*	UNK	H	OS*	UNK	H	OS*	UNK	H	OS*	UNK	H
ERM (Koltchinskii, 2011)	68.8	24.6	38.9	59.5	33.1	41.0	43.3	20.3	28.9	37.5	30.0	35.7	52.3	27.0	36.1
ERM+CM (Zhu & Li, 2022)	68.7	44.6	44.9	62.3	43.2	48.3	41.0	33.2	30.4	39.9	**54.2**	41.6	53.0	44.5	41.3
ADA (Volpi et al., 2018)	71.0	28.8	39.0	62.1	33.8	41.6	43.2	22.4	26.9	40.7	38.8	38.1	54.2	30.9	36.4
ADA+CM (Zhu & Li, 2022)	**72.9**	**40.1**	42.4	**64.4**	49.1	51.8	**45.0**	40.9	35.2	**43.3**	52.5	**42.8**	**56.4**	45.6	**43.0**
MEADA (Zhao et al., 2020)	70.9	28.7	38.9	62.1	33.6	41.3	43.4	22.9	26.4	39.8	40.3	38.2	54.1	31.4	36.2
MEADA+CM (Zhu & Li, 2022)	70.5	33.4	41.9	63.8	**53.7**	51.4	40.3	48.8	35.8	42.9	50.6	41.6	54.3	46.6	42.7
***OneRing*-S**	57.3	38.4	**46.0**	56.0	50.3	**53.0**	25.9	**86.6**	**39.8**	35.7	22.1	27.1	43.7	**49.4**	41.5

Table 5: Accuracy (%) on **Office-31** and **VisDA** dataset using ResNet-50. **open-partial domain adaptation** where for *Office-31*: $|\mathcal{C}_s| = 20$, $|\mathcal{C}_t| = 21$, $|\mathcal{C}_s \cap \mathcal{C}_t| = 10$; and for *VisDA*: $|\mathcal{C}_s| = 9$, $|\mathcal{C}_t| = 9$, $|\mathcal{C}_s \cap \mathcal{C}_t| = 6$. The second highest H score is underlined. **SF** indicates whether source-free.

Office-31	**SF**	A2W		D2W		W2D		A2D		D2A		W2A		**Avg**		**VisDA**
		OS	H	OS	H	OS	H	OS	H	OS	H	OS	H	OS	H	H
OSBP (Saito et al., 2018b)	✗	66.1	50.2	73.6	55.5	85.6	57.2	72.9	51.1	47.4	49.8	60.5	50.2	67.7	52.3	27.3
UAN (You et al., 2019)	✗	85.6	58.6	94.8	70.6	**98.0**	71.4	86.5	59.7	85.5	60.1	85.1	60.3	89.2	63.5	30.5
ROS (Bucci et al., 2020)	✗	-	71.3	-	94.6	-	95.3	-	71.4	-	81.0	-	81.2	-	82.1	-
CMU (Fu et al., 2020)	✗	86.7	67.3	**96.7**	79.3	**98.0**	80.4	89.1	68.1	88.4	71.4	88.6	72.2	91.1	73.1	34.6
DCC (Li et al., 2021a)	✗	**91.7**	78.5	94.5	79.3	96.2	88.6	**93.7**	**88.5**	**90.4**	70.2	**92.0**	75.9	**93.1**	80.2	43.0
DANCE (Saito et al., 2020)	✗	-	71.5	-	91.4	-	87.9	-	78.6	-	79.9	-	72.2	-	80.3	4.4
OVANet (Saito & Saenko, 2021)	✗	-	79.4	-	**95.4**	-	94.3	-	85.8	-	80.1	-	84.0	-	86.5	53.1
USFDA Kundu et al. (2020a)	✓	-	79.8	-	90.6	-	81.2	-	85.5	-	83.2	-	88.7	-	84.8	-
UMAD Liang et al. (2021b)	✓	-	77.4	-	90.7	-	**97.2**	-	79.1	-	87.4	-	**90.4**	-	87.0	58.3
***OneRing*-S**		69.0	67.9	92.5	90.6	96.5	89.4	81.9	74.9	64.8	74.8	69.9	78.8	79.1	79.4	35.2
OneRing	✓	78.8	83.8	94.7	95.2	97.5	96.0	86.6	85.7	82.0	85.8	81.0	84.7	86.8	88.5	60.7
OneRing+	✓	85.3	**85.4**	94.0	94.2	97.0	93.6	88.4	86.1	88.9	**90.7**	87.3	84.0	90.2	**89.0**	**66.1**

Table 6: **H-score** (%) on **Office-Home** dataset using ResNet-50 as backbone. **open-partial domain adaptation** where $|\mathcal{C}_s| = 15$, $|\mathcal{C}_t| = 60$, $|\mathcal{C}_s \cap \mathcal{C}_t| = 10$. The second highest H score is underlined. **SF** indicates whether source-free.

	SF	A2C	A2P	A2R	C2A	C2P	C2R	P2A	P2C	P2R	R2A	R2C	R2P	**Avg**
OSBP (Saito et al., 2018b)	✗	39.6	45.1	46.2	45.7	45.2	46.8	45.3	40.5	45.8	45.1	41.6	46.9	44.5
UAN (You et al., 2019)	✗	51.6	51.7	54.3	61.7	57.6	61.9	50.4	47.6	61.5	62.9	52.6	65.2	56.6
CMU (Fu et al., 2020)	✗	56.0	56.9	59.1	66.9	64.2	67.8	54.7	51.0	66.3	68.2	57.8	69.7	61.6
DCC (Li et al., 2021a)	✗	58.0	54.1	58.0	**74.6**	70.6	77.5	64.3	**73.6**	74.9	81.0	**75.1**	80.4	70.2
DANCE (Saito et al., 2020)	✗	-	-	-	-	-	-	-	-	-	-	-	-	49.2
OVANet (Saito & Saenko, 2021)	✗	62.8	75.6	78.6	70.7	68.8	75.0	71.3	58.6	80.5	76.1	64.1	78.9	71.8
UMAD (Liang et al., 2021b)	✓	61.1	76.3	82.7	70.7	67.7	75.7	64.4	55.7	76.3	73.2	60.4	77.2	70.1
***OneRing*-S**		55.7	72.4	79.6	64.6	65.3	74.6	65.9	51.5	77.9	72.1	57.8	75.0	67.7
OneRing	✓	63.3	72.4	81.0	68.8	67.2	74.6	73.3	60.8	80.9	78.1	63.9	76.7	71.8
OneRing+	✓	**69.5**	**81.4**	**87.9**	73.2	**77.9**	**82.4**	**81.5**	68.6	**88.1**	**81.1**	70.5	**85.7**	**79.0**

the source data. In the first phase, we train with the standard CE loss. Then after convergence, we add the second CE loss for a few epochs. For all experiments under SF-OPDA setting, the *OneRing* classifier is fixed during target adaptation. When augmented with AaD (Yang et al., 2022), we set the hyperparameter K in $\mathcal{L}_{dis}$ same as AaD, and β in $\mathcal{L}_{div}$ as 1. *We use the predictions in current mini-batch to estimate the known/unknown ratio in Eq. 2, since it does not require access to the whole dataset, and we show it achieves similar results as using the one over whole dataset in Tab. 9 in the appendix.*

Table 7: **H-score** (%) on **DomainNet** using ResNet-50 as backbone. **open-partial domain adaptation** where $|\mathcal{C}_s| = 200$, $|\mathcal{C}_t| = 295$, $|\mathcal{C}_s \cap \mathcal{C}_t| = 150$. The second highest H score is underlined. **SF** indicates whether source-free.

Method	SF	P2R	R2P	P2S	S2P	R2S	S2R	Avg
OSBP (Saito et al., 2018b)	✗	33.6	33.0	30.6	30.5	30.6	33.7	32.0
DANCE (Saito et al., 2020)	✗	21.0	47.3	37.0	27.7	**46.7**	21.0	33.5
UAN (You et al., 2019)	✗	41.9	43.6	39.1	38.9	38.7	43.7	41.0
CMU (Fu et al., 2020)	✗	50.8	**52.2**	45.1	44.8	45.6	51.0	48.3
DCC (Li et al., 2021a)	✗	56.9	50.3	43.7	44.9	43.3	56.2	49.2
OVANet (Saito & Saenko, 2021)	✗	56.0	51.7	**47.1**	47.4	44.9	57.2	<u>50.7</u>
UMAD Liang et al. (2021b)	✓	59.0	50.1	44.3	32.0	42.1	55.3	47.1
***OneRing*-S**		**59.1**	42.9	43.8	35.5	39.5	52.9	45.6
OneRing	✓	57.9	52.0	46.5	**49.6**	44.1	**57.8**	**51.3**

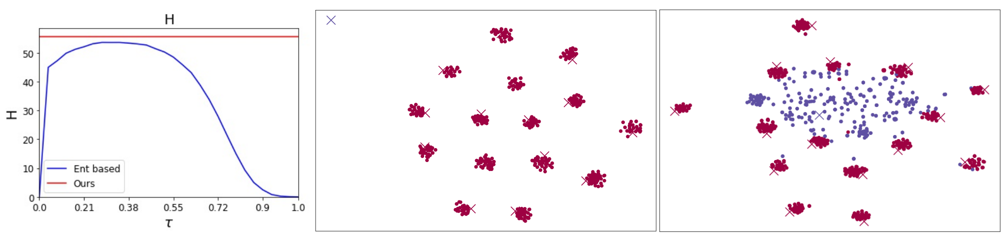

Figure 3: (**Left**) H value of our source model and entropy based rejection on A2C of Office-Home. t-SNE visualization of features with either only source known categories (**Middle**) or also with 10 source extra unknown categories (**Right**) from source model on *Artistic* of Office-Home, where the cross is the class prototype. The red denotes known classes while other for unknown class.

For OS-SDG, we will report average per-class accuracy over known categories (OS^*), unknown class accuracy (UNK) and harmonic mean (H) between OS^* and UNK. For SF-OPDA, we will mainly report the harmonic mean, as all previous methods did, and also the average per-class accuracy over all categories (OS) on Office-31. Note for OS-SDG and SF-OPDA, the model is expected to have high performance on both known and unknown accuracy, which should result in a high harmonic mean (H). As pointed out by ROS (Bucci et al., 2020), OS is not a reasonable evaluation metric and can be quite high even when UNK is 0, since $OS = \frac{|\mathcal{C}_s|}{|\mathcal{C}_s|+1} \times OS^* + \frac{1}{|\mathcal{C}_s|+1} \times UNK$. In the following tables, we will denote our model trained with only source data as ***OneRing*-S**, model after target adaptation as ***OneRing***, and model augmented with AaD after target adaptation as ***One Ring*+**.

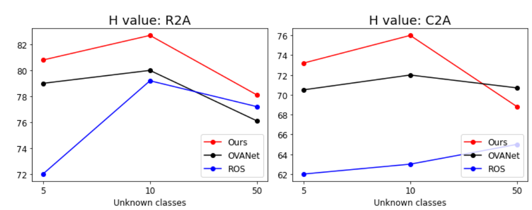

Figure 2: **H** value of open-partial domain adaptation on Office-Home. We vary the number of unknown classes as shown in the x axis. Here 'ours' denotes OneRing without being augmented with AaD, OVANet and ROS demand source data.

4.3 Quantitative Results

Open-set Single Domain Generalization. In Tab. 2-4, we show the results of our source model *OneRing*-S on Office-31, Office-Home and PACS. ERM (Koltchinskii, 2011), ADA (Volpi et al., 2018) and MEADA (Zhao et al., 2020) are methods originally designed for typical domain generalization, CrossMatch (CM) (Zhu & Li, 2022) is plugged into these methods which empower them with the ability to detect unknown classes in the target domain with several complex modules, as well as generating unknown samples. While our *OneRing*-S is elegantly simple, the results show it can better detect open classes under domain shift compared to CM. Note, we have no module specifically for DG in *OneRing*-S. The fact that *OneRing*-S has better performance proves the efficacy of our method.

Source-free open-partial domain adaptation In Tab. 5-7, we show the results under open-partial DA setting where **SF** column indicates whether source-free. Note that our method does not need source data during target adaptation. As shown in the tables, our source model (*One Ring-S*) already achieves decent H performance. The simple *OneRing* with only entropy minimization already outperforms all other methods on all 4 benchmarks, adding AaD (Yang et al., 2022) into method as shown in Eq. 3 (*OneRing*+) can further improve the results significantly, leading to 0.5%, 5.4% and 7.2% improvement on Office-31, VisDA and Office-Home respectively, and it surpasses the current state-of-the-art OVANet by by 2.5%, 7.2% and 13% on these 3 benchmarks respectively.

4.4 Analysis

Compare One Ring with entropy based unknown rejection. We also show the results with entropy based unknown rejection, where a sample is predicted as unknown if the entropy (maximal normalized) of the prediction (*with normal classifier head*) is higher than a manually set threshold. Fig. 3 (*left*) shows the H value of *source pretrained model* on A2C task of Office-Home under open-partial DA setting, where the *x axis* denotes the threshold. Our source model gets better results without any extra effort.

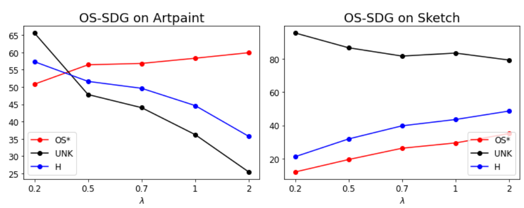

Figure 4: Results on PACS (**OS-SDG**), with different weight factors applied to the normal CE loss.

Trade-off between 2 CE losses. In this paper, we show results where the two CE losses have equal weight, and hence our method does not have any hyperparameter. However, in Eq. 1, we can also multiply a weight factor to the standard CE loss as a trade-off. Intuitively, a smaller factor to the standard CE loss gives more weight to unknown-class recognition and vice verse. The results under OS-SDG setting in Fig. 4 verify this , where the *x axis* denotes the weight factor multiplied to the standard CE loss. As can be seen, this trade-off can be used to further improve results. However, for the sake of simplicity, and given the already good results, we choose not to optimize this parameter.

Visualization of features and class prototypes. In Fig. 3 (*Middle*), we visualize the source features and class prototypes (weights of *OneRing* classifier) from *source model* with t-SNE. The prototype of the unknown category is in the corner with no source features around it. In Fig. 3 (*Right*), we further visualize 10 extra unknown classes. It shows that those features of unknown categories will not cluster around any of the known classes, but they are close to the unknown prototype. This implies that the *OneRing* model can efficiently distinguish known and unknown categories.

Importance of weight in entropy minimization. We ablate the weights in entropy minimization in Eq. 2. If removing weights, the *OS**, *UNK* and *H* on R2C (Office-Home) will decrease from 57.8/71.6/63.9 to 19.2/97.8/32.1 respectively, showing the deployed weights are important and effective to balance the two terms in Eq. 2.

Robustness to amount of unknown categories. In Fig. 2, we compare our source-free *OneRing* (without being augmented with AaD) to ROS (Bucci et al., 2020) and OVANet (Saito & Saenko, 2021) under UNDA setting with different amount of unknown categories from target domain. The results show that our method is robust to the amount of unknown categories.

5 Conclusion

In this paper, we first introduce a simple method with the proposed *OneRing* classifier head, it possesses strong ability to detect unknown categories from target data even no matter without or with domain shift after training with two simple cross entropy losses. Then, we further adapt the model to the target domain which contains unknown categories, with only weighted entropy minimization and no access to source data. In the experiment, we show that our method achieves good performance on open-set single domain generalization and source-free open-partial domain adaptation, which proves the effectiveness of our method.

Reproducibility Statement

We will release our code if accepted. The code is based on pytorch 1.3 with cuda 10.0, and the results of the submission can be reproduced with the code, with the random seed as 2021 or 2022.

References

Silvia Bucci, Mohammad Reza Loghmani, and Tatiana Tommasi. On the effectiveness of image rotation for open set domain adaptation. In *ECCV*, pp. 422–438. Springer, 2020.

Zhangjie Cao, Lijia Ma, Mingsheng Long, and Jianmin Wang. Partial adversarial domain adaptation. In *ECCV*, pp. 135–150, 2018.

Zhangjie Cao, Kaichao You, Mingsheng Long, Jianmin Wang, and Qiang Yang. Learning to transfer examples for partial domain adaptation. In *CVPR*, pp. 2985–2994, 2019.

Guangyao Chen, Limeng Qiao, Yemin Shi, Peixi Peng, Jia Li, Tiejun Huang, Shiliang Pu, and Yonghong Tian. Learning open set network with discriminative reciprocal points. In *ECCV*, 2020.

Guangyao Chen, Peixi Peng, Xiangqian Wang, and Yonghong Tian. Adversarial reciprocal points learning for open set recognition. *IEEE TPAMI*, 2021.

Safa Cicek and Stefano Soatto. Unsupervised domain adaptation via regularized conditional alignment. In *ICCV*, pp. 1416–1425, 2019.

Shuhao Cui, Shuhui Wang, Junbao Zhuo, Liang Li, Qingming Huang, and Qi Tian. Towards discriminability and diversity: Batch nuclear-norm maximization under label insufficient situations. *CVPR*, 2020.

Zhijie Deng, Yucen Luo, and Jun Zhu. Cluster alignment with a teacher for unsupervised domain adaptation. In *ICCV*, pp. 9944–9953, 2019.

Qi Dou, Daniel Coelho de Castro, Konstantinos Kamnitsas, and Ben Glocker. Domain generalization via model-agnostic learning of semantic features. *NeurIPS*, 32, 2019.

Xinjie Fan, Qifei Wang, Junjie Ke, Feng Yang, Boqing Gong, and Mingyuan Zhou. Adversarially adaptive normalization for single domain generalization. In *CVPR*, pp. 8208–8217, 2021.

Qianyu Feng, Guoliang Kang, Hehe Fan, and Yi Yang. Attract or distract: Exploit the margin of open set. In *ICCV*, pp. 7990–7999, 2019.

Zeyu Feng, Chang Xu, and Dacheng Tao. Open-set hypothesis transfer with semantic consistency. *IEEE Transactions on Image Processing*, 30:6473–6484, 2021.

Bo Fu, Zhangjie Cao, Mingsheng Long, and Jianmin Wang. Learning to detect open classes for universal domain adaptation. In *ECCV*, pp. 567–583. Springer, 2020.

Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario Marchand, and Victor Lempitsky. Domain-adversarial training of neural networks. *JMLR*, 17(1):2096–2030, 2016.

Zongyuan Ge, Sergey Demyanov, and Rahil Garnavi. Generative openmax for multi-class open set classification. In *BMVC*, 2017.

Boqing Gong, Yuan Shi, Fei Sha, and Kristen Grauman. Geodesic flow kernel for unsupervised domain adaptation. In *CVPR*, pp. 2066–2073. IEEE, 2012.

Ishaan Gulrajani and David Lopez-Paz. In search of lost domain generalization. *ICLR*, 2021.

Jiaxing Huang, Dayan Guan, Aoran Xiao, and Shijian Lu. Model adaptation: Historical contrastive learning for unsupervised domain adaptation without source data. *NeurIPS*, 34, 2021.

Taotao Jing, Hongfu Liu, and Zhengming Ding. Towards novel target discovery through open-set domain adaptation. In *ICCV*, pp. 9322–9331, 2021.

Vladimir Koltchinskii. *Oracle inequalities in empirical risk minimization and sparse recovery problems: École D'Été de Probabilités de Saint-Flour XXXVIII-2008*, volume 2033. Springer Science & Business Media, 2011.

Shu Kong and Deva Ramanan. Opengan: Open-set recognition via open data generation. *ICCV*, 2021.

Alex Krizhevsky, Geoffrey Hinton, et al. Learning multiple layers of features from tiny images. Technical report, Citeseer, 2009.

Jogendra Nath Kundu, Naveen Venkat, and R Venkatesh Babu. Universal source-free domain adaptation. *CVPR*, 2020a.

Jogendra Nath Kundu, Naveen Venkat, Ambareesh Revanur, R Venkatesh Babu, et al. Towards inheritable models for open-set domain adaptation. In *CVPR*, pp. 12376–12385, 2020b.

Chen-Yu Lee, Tanmay Batra, Mohammad Haris Baig, and Daniel Ulbricht. Sliced wasserstein discrepancy for unsupervised domain adaptation. In *CVPR*, pp. 10285–10295, 2019.

Da Li, Yongxin Yang, Yi-Zhe Song, and Timothy M Hospedales. Deeper, broader and artier domain generalization. In *ICCV*, pp. 5542–5550, 2017.

Da Li, Yongxin Yang, Yi-Zhe Song, and Timothy M Hospedales. Learning to generalize: Meta-learning for domain generalization. In *AAAI*, 2018.

Da Li, Jianshu Zhang, Yongxin Yang, Cong Liu, Yi-Zhe Song, and Timothy M Hospedales. Episodic training for domain generalization. In *ICCV*, pp. 1446–1455, 2019.

Guangrui Li, Guoliang Kang, Yi Zhu, Yunchao Wei, and Yi Yang. Domain consensus clustering for universal domain adaptation. In *CVPR*, pp. 9757–9766, 2021a.

Lei Li, Ke Gao, Juan Cao, Ziyao Huang, Yepeng Weng, Xiaoyue Mi, Zhengze Yu, Xiaoya Li, and Boyang Xia. Progressive domain expansion network for single domain generalization. In *CVPR*, pp. 224–233, 2021b.

Rui Li, Qianfen Jiao, Wenming Cao, Hau-San Wong, and Si Wu. Model adaptation: Unsupervised domain adaptation without source data. In *CVPR*, pp. 9641–9650, 2020.

Jian Liang, Dapeng Hu, and Jiashi Feng. Do we really need to access the source data? source hypothesis transfer for unsupervised domain adaptation. *ICML*, 2020a.

Jian Liang, Yunbo Wang, Dapeng Hu, Ran He, and Jiashi Feng. A balanced and uncertainty-aware approach for partial domain adaptation. In *ECCV*, pp. 123–140. Springer, 2020b.

Jian Liang, Dapeng Hu, and Jiashi Feng. Domain adaptation with auxiliary target domain-oriented classifier. In *CVPR*, pp. 16632–16642, 2021a.

Jian Liang, Dapeng Hu, Jiashi Feng, and Ran He. Umad: Universal model adaptation under domain and category shift. *arXiv preprint arXiv:2112.08553*, 2021b.

Hong Liu, Zhangjie Cao, Mingsheng Long, Jianmin Wang, and Qiang Yang. Separate to adapt: Open set domain adaptation via progressive separation. In *CVPR*, pp. 2927–2936, 2019.

Hong Liu, Jianmin Wang, and Mingsheng Long. Cycle self-training for domain adaptation. In *NeurIPS*, 2021.

Mingsheng Long, Yue Cao, Jianmin Wang, and Michael I Jordan. Learning transferable features with deep adaptation networks. *ICML*, 2015.

Mingsheng Long, Han Zhu, Jianmin Wang, and Michael I Jordan. Unsupervised domain adaptation with residual transfer networks. In *NIPS*, pp. 136–144, 2016.

Mingsheng Long, Yue Cao, Zhangjie Cao, Jianmin Wang, and Michael I Jordan. Transferable representation learning with deep adaptation networks. *TPAMI*, 41(12):3071–3085, 2018a.

Mingsheng Long, Zhangjie Cao, Jianmin Wang, and Michael I Jordan. Conditional adversarial domain adaptation. In *NIPS*, pp. 1647–1657, 2018b.

Zhihe Lu, Yongxin Yang, Xiatian Zhu, Cong Liu, Yi-Zhe Song, and Tao Xiang. Stochastic classifiers for unsupervised domain adaptation. In *CVPR*, pp. 9111–9120, 2020.

Lawrence Neal, Matthew Olson, Xiaoli Fern, Weng-Keen Wong, and Fuxin Li. Open set learning with counterfactual images. In *ECCV*, 2018a.

Lawrence Neal, Matthew Olson, Xiaoli Fern, Weng-Keen Wong, and Fuxin Li. Open set learning with counterfactual images. In *ECCV*, pp. 613–628, 2018b.

Yuval Netzer, Tao Wang, Adam Coates, Alessandro Bissacco, Bo Wu, and Andrew Y. Ng. Reading digits in natural images with unsupervised feature learning. In *NIPS Workshop on Deep Learning and Unsupervised Feature Learning 2011*, 2011.

Yingwei Pan, Ting Yao, Yehao Li, Chong-Wah Ngo, and Tao Mei. Exploring category-agnostic clusters for open-set domain adaptation. In *CVPR*, pp. 13867–13875, 2020.

Pau Panareda Busto and Juergen Gall. Open set domain adaptation. In *ICCV*, pp. 754–763, 2017.

Xingchao Peng, Ben Usman, Neela Kaushik, Judy Hoffman, Dequan Wang, and Kate Saenko. Visda: The visual domain adaptation challenge. *arXiv preprint arXiv:1710.06924*, 2017.

Xingchao Peng, Qinxun Bai, Xide Xia, Zijun Huang, Kate Saenko, and Bo Wang. Moment matching for multi-source domain adaptation. In *ICCV*, pp. 1406–1415, 2019.

Fengchun Qiao, Long Zhao, and Xi Peng. Learning to learn single domain generalization. In *CVPR*, pp. 12556–12565, 2020.

Alexander Robey, George J Pappas, and Hamed Hassani. Model-based domain generalization. *NeurIPS*, 34:20210–20229, 2021.

Kate Saenko, Brian Kulis, Mario Fritz, and Trevor Darrell. Adapting visual category models to new domains. In *ECCV*, pp. 213–226. Springer, 2010.

Kuniaki Saito and Kate Saenko. Ovanet: One-vs-all network for universal domain adaptation. In *ICCV*, pp. 9000–9009, 2021.

Kuniaki Saito, Kohei Watanabe, Yoshitaka Ushiku, and Tatsuya Harada. Maximum classifier discrepancy for unsupervised domain adaptation. In *CVPR*, pp. 3723–3732, 2018a.

Kuniaki Saito, Shohei Yamamoto, Yoshitaka Ushiku, and Tatsuya Harada. Open set domain adaptation by backpropagation. In *ECCV*, pp. 153–168, 2018b.

Kuniaki Saito, Donghyun Kim, Stan Sclaroff, and Kate Saenko. Universal domain adaptation through self supervision. *NeurIPS*, 33, 2020.

Yuge Shi, Jeffrey Seely, Philip HS Torr, N Siddharth, Awni Hannun, Nicolas Usunier, and Gabriel Synnaeve. Gradient matching for domain generalization. *ICLR*, 2022.

Rui Shu, Hung H Bui, Hirokazu Narui, and Stefano Ermon. A dirt-t approach to unsupervised domain adaptation. *ICLR*, 2018.

Yang Shu, Zhangjie Cao, Chenyu Wang, Jianmin Wang, and Mingsheng Long. Open domain generalization with domain-augmented meta-learning. In *CVPR*, pp. 9624–9633, 2021.

Yu Shu, Yemin Shi, Yaowei Wang, Tiejun Huang, and Yonghong Tian. P-odn: Prototype-based open deep network for open set recognition. *Scientific Reports*, 2020.

Baochen Sun, Jiashi Feng, and Kate Saenko. Return of frustratingly easy domain adaptation. In *AAAI*, 2016.

Xin Sun, Zhenning Yang, Chi Zhang, Guohao Peng, and Keck-Voon Ling. Conditional gaussian distribution learning for open set recognition. In *CVPR*, 2020.

Hui Tang, Ke Chen, and Kui Jia. Unsupervised domain adaptation via structurally regularized deep clustering. In *CVPR*, pp. 8725–8735, 2020.

Eric Tzeng, Judy Hoffman, Ning Zhang, Kate Saenko, and Trevor Darrell. Deep domain confusion: Maximizing for domain invariance. *arXiv preprint arXiv:1412.3474*, 2014.

Eric Tzeng, Judy Hoffman, Kate Saenko, and Trevor Darrell. Adversarial discriminative domain adaptation. In *CVPR*, pp. 7167–7176, 2017.

Sagar Vaze, Kai Han, Andrea Vedaldi, and Andrew Zisserman. Open-set recognition: A good closed-set classifier is all you need. In *ICLR*, 2022. URL `https://openreview.net/forum?id=5hLP5JY9S2d`.

Ramakrishna Vedantam, David Lopez-Paz, and David J Schwab. An empirical investigation of domain generalization with empirical risk minimizers. *NeurIPS*, 34, 2021.

Hemanth Venkateswara, Jose Eusebio, Shayok Chakraborty, and Sethuraman Panchanathan. Deep hashing network for unsupervised domain adaptation. In *CVPR*, pp. 5018–5027, 2017.

Riccardo Volpi, Hongseok Namkoong, Ozan Sener, John C Duchi, Vittorio Murino, and Silvio Savarese. Generalizing to unseen domains via adversarial data augmentation. *NeurIPS*, 31, 2018.

Zijian Wang, Yadan Luo, Ruihong Qiu, Zi Huang, and Mahsa Baktashmotlagh. Learning to diversify for single domain generalization. In *ICCV*, pp. 834–843, 2021.

Zhirong Wu, Yuanjun Xiong, Stella X Yu, and Dahua Lin. Unsupervised feature learning via non-parametric instance discrimination. In *CVPR*, pp. 3733–3742, 2018.

Haifeng Xia, Handong Zhao, and Zhengming Ding. Adaptive adversarial network for source-free domain adaptation. In *ICCV*, pp. 9010–9019, 2021.

Shiqi Yang, Yaxing Wang, Joost van de Weijer, Luis Herranz, and Shangling Jui. Unsupervised domain adaptation without source data by casting a bait. *arXiv preprint arXiv:2010.12427*, 2020.

Shiqi Yang, Joost van de Weijer, Luis Herranz, Shangling Jui, et al. Exploiting the intrinsic neighborhood structure for source-free domain adaptation. *NeurIPS*, 34, 2021a.

Shiqi Yang, Yaxing Wang, Joost van de Weijer, Luis Herranz, and Shangling Jui. Generalized source-free domain adaptation. In *ICCV*, pp. 8978–8987, 2021b.

Shiqi Yang, Yaxing Wang, Kai Wang, Shangling Jui, and Joost van de Weijer. Attracting and dispersing: A simple approach for source-free domain adaptation. *arXiv preprint arXiv:2205.04183*, 2022.

Kaichao You, Mingsheng Long, Zhangjie Cao, Jianmin Wang, and Michael I Jordan. Universal domain adaptation. In *CVPR*, pp. 2720–2729, 2019.

Hongjie Zhang, Ang Li, Jie Guo, and Yanwen Guo. Hybrid models for open set recognition. In *ECCV*, pp. 102–117. Springer, 2020.

Yabin Zhang, Hui Tang, Kui Jia, and Mingkui Tan. Domain-symmetric networks for adversarial domain adaptation. In *CVPR*, pp. 5031–5040, 2019.

Long Zhao, Ting Liu, Xi Peng, and Dimitris Metaxas. Maximum-entropy adversarial data augmentation for improved generalization and robustness. *NeurIPS*, 33:14435–14447, 2020.

Da-Wei Zhou, Han-Jia Ye, and De-Chuan Zhan. Learning placeholders for open-set recognition. In *CVPR*, pp. 4401–4410, 2021a.

Kaiyang Zhou, Yongxin Yang, Timothy Hospedales, and Tao Xiang. Learning to generate novel domains for domain generalization. In *ECCV*, pp. 561–578. Springer, 2020.

Kaiyang Zhou, Yongxin Yang, Yu Qiao, and Tao Xiang. Domain generalization with mixstyle. *ICLR*, 2021b.

Ronghang Zhu and Sheng Li. Crossmatch: Cross-classifier consistency regularization for open-set single domain generalization. In *ICLR*, 2022.

Table 8: Results of **Open-set Recognition** task. All results indicate the area under the Receiver-Operator curve (**AUROC**) averaged over five 'known/unknown' class splits. All methods are augmented with improved optimization strategies from (Vaze et al., 2022). Results are taken from (Vaze et al., 2022).

Method	SVHN	CIFAR10	CIFAR + 10	CIFAR + 50
OSRCI (Neal et al., 2018b)	89.9	87.2	91.1	90.3
(ARPL + CS) (Chen et al., 2021)	96.8	**93.9**	**98.1**	**96.7**
MSP (Vaze et al., 2022)	96.0	90.1	95.6	94.0
MLS (Vaze et al., 2022)	97.1	93.6	97.9	96.5
***OneRing*-S**	**97.3**	93.7	97.8	96.2

A Appendix

Open-set Recognition. We also evaluate on open-set recognition to further show the generalization ability of our method, where the model is trained on the source data and directly tested on the target data containing unknown classes. We use the following benchmarks to evaluate our method with the same setting as (Vaze et al., 2022): 1) **SVHN** (Netzer et al., 2011) contains 10 street-view house numbers respectively. 2) **CIFAR10** (Krizhevsky et al., 2009) consists of natural images of 10 classes covering animals and vehicles. For these benchmarks, the model will be trained on 6 out of 10 categories and evaluated on the remaining 4 classes . 3) **CIFAR + N** (Krizhevsky et al., 2009) is an extension of CIFAR10. Here, methods are trained on 4 classes from CIFAR10 and evaluated on N classes from CIFAR100, where N is set to 10 or 50 classes . We will report AUROC to quantify the open class detection performance. All results are the average of three random runs, except for OSR which we only run once under 5 different splitting.

Even though in this paper our focus is on open-set recognition under domain shift, we also include results for OSR (which does not include any domain shift) in Tab. 8. All methods in the table use the same training tricks to improve the source performance including learning rate decay, warmup, label smoothing and more data augmentations, which are proposed in (Vaze et al., 2022). Note OSRCI (Neal et al., 2018b) and ARPL+CS (Chen et al., 2021) are complex methods which either need to generate open-set samples or learn extra reciprocal points. The results show that our source model *OneRing*-S can work quite well on OSR task without any extra learning process, indicating good generalization ability. It is important to observe that we do not have any hyperparameter.

Table 9: **H-score** (%) on **Office-Home** dataset using ResNet-50 as backbone. **open-partial domain adaptation** where $|\mathcal{C}_s| = 15$, $|\mathcal{C}_t| = 60$, $|\mathcal{C}_s \cap \mathcal{C}_t| = 10$. The second highest H score is underlined. **SF** indicates whether source-free. * indicates using predictions over the *whole dataset* instead of mini-batch in Eq. 2.

	SF	A2C	A2P	A2R	C2A	C2P	C2R	P2A	P2C	P2R	R2A	R2C	R2P	**Avg**
OVANet (Saito & Saenko, 2021)	✗	62.8	75.6	78.6	70.7	68.8	75.0	71.3	58.6	80.5	76.1	64.1	78.9	71.8
***OneRing*-S**		55.7	72.4	79.6	64.6	65.3	74.6	65.9	51.5	77.9	72.1	57.8	75.0	67.7
OneRing	✓	63.3	72.4	81.0	68.8	67.2	74.6	73.3	60.8	80.9	78.1	63.9	76.7	71.8
OneRing*	✓	60.9	72.1	80.9	67.7	66.0	73.7	73.1	60.4	81.4	77.7	63.4	78.2	71.3
OneRing+	✓	69.5	81.4	87.9	73.2	77.9	82.4	81.5	68.6	88.1	81.1	70.5	85.7	79.0
OneRing*+	✓	70.1	82.5	88.9	75.1	80.1	83.0	82.5	64.6	89.3	81.0	66.4	86.0	**79.1**

Known/unknown ratio estimation through mini-batch or whole dataset. In Eq. 2, we have two choice to estimate the known/unknown ratio, which will be utilized to balance the 2 entropy terms. In Tab. 9, we show that these 2 different manners lead to almost the same results. Though there may exist some imbalance mini-batches which only contain few samples predicted as known or unknown, the results imply that the known/unknown ratio estimated by the mini-batch is enough to achieve decent performance. Note the Office-Home here is not a well balance (amount of samples per category) dataset, and also in the target domain the unknown categories (50) are much more than known (10).

Ablation study of the weight in entropy minimization. In Tab. 10, we show OS*, UNK, OS and H to ablate the effectiveness of the weight in Eq. 2. The results show that the weight it important to achieve high known accuracy OS*.

Table 10: Ablation study (R2C of Office-Home) on the proposed weight in the weighted entropy minimization. **Results of OVANet are from our running based on their official code.**

R2C	**OS***	**UNK**	**OS**	**H**
OVANet Saito & Saenko (2021)	55.1	70.0	56.5	61.7
w/o weight in Eq.2	19.2	97.8	26.3	32.1
w/ weight in Eq.2 (OneRing)	57.8	71.6	59.1	63.9
+ AaD (OneRing+)	61.5	82.7	63.4	70.5

Table 11: Open-partial DA on VisDA, results of OVANet are from our running based on their code.

VisDA	**source-free**	**OS***	**UNK**	**OS**	**H**
OVANet Saito & Saenko (2021)	✗	60.5	46.4	58.5	52.5
OneRing-S		25.7	55.9	30.0	35.2
OneRing	✓	57.2	64.6	58.3	60.7
OneRing+	✓	65.5	66.8	65.7	66.1

Table 12: Open-set DA on Office-31 (VGG19), results (H) except ours are from OVANet.

ODA/UnDA methods	***source-free***	**A2D**	**A2W**	**D2A**	**D2W**	**W2D**	**W2A**	**Avg**
OSBP	✗	81.0	77.5	78.2	95.0	91.0	72.9	82.6
ROS	✗	79.0	81.0	78.1	94.4	99.7	74.1	84.4
OVANet	✗	89.5	84.9	89.7	93.7	85.8	88.5	88.7
OneRing	✓	91.0	84.5	90.1	96.0	93.7	90.1	**90.9**

Table 13: Closed-set DA on Office-31 (ResNet50), results (accuracy) except ours are from DANCE, and **results of OVANet are from our running based on their official code**.

UNDA methods	***source-free***	**A2W**	**D2W**	**W2D**	**A2D**	**D2A**	**W2A**	**Avg**
ETN	✗	87.9	99.2	100	88.4	68.7	66.8	85.2
STA	✗	77.1	90.7	98.1	75.5	51.4	48.9	73.6
UAN	✗	86.5	97.0	100	84.5	69.6	68.7	84.4
DANCE	✗	88.6	97.5	100	89.4	69.5	68.2	85.5
OVANet	✗	88.1	97.0	99.1	88.6	68.8	67.0	84.8
OneRing	✓	89.0	97.3	100	89.0	70.1	68.5	**85.7**

Table 14: Accuracy (%) on open-partial DA.

Office-Home																		
	Ar → Cl			Ar → Pr			Ar → Rw			Cl→ Ar			Cl → Pr			Cl → Rw		
	OS*	UNK	**HOS**	OS*	UNK	**HOS**	OS*	UNK	**HOS**	OS*	UNK	**HOS**	OS*	UNK	**HOS**	OS*	UNK	**HOS**
OneRing-S	42.9	79.3	55.7	75.7	69.5	72.4	91.7	70.3	79.6	52.9	82.1	64.4	60.0	71.7	65.3	75.2	74.0	74.6
OneRing	54.1	73.9	62.5	78.5	69.8	73.9	93.3	72.5	81.6	65.9	73.0	69.3	67.5	66.1	66.8	80.0	69.0	74.1
OneRing+	58.5	84.4	69.1	78.3	84.8	81.4	92.6	84.4	88.3	62.7	88.2	73.3	72.1	86.3	78.6	80.4	86.0	83.1
	Pr → Ar			Pr → Cl			Pr → Rw			Rw→ Ar			Rw → Cl			Rw → Pr		
	OS*	UNK	**HOS**	OS*	UNK	**HOS**	OS*	UNK	**HOS**	OS*	UNK	**HOS**	OS*	UNK	**HOS**	OS*	UNK	**HOS**
OneRing-S	55.9	80.2	65.9	38.6	77.1	51.5	86.9	70.5	77.9	70.4	73.9	72.1	46.6	76.0	57.8	82.6	68.7	75.0
OneRing	73.1	73.2	73.1	52.8	70.2	60.3	91.6	73.4	81.5	77.9	78.1	78.0	57.7	70.2	63.4	88.2	70.3	78.2
OneRing+	77.4	86.7	82.2	58.3	82.3	68.3	92.2	84.7	88.3	76.5	86.2	81.1	61.5	82.7	70.5	86.9	85.4	86.1

Table 15: Results of OneRing with different number of unknown categories on R2A of Office-Home.

R2A, shared classes = 10 (OS*/UNK/H)	**OneRing**
unknown classes = 5	77.3/84.4/80.7
unknown classes = 10	84.4/81.1/82.7
unknown classes = 50	77.9/78.1/78.0

Table 16: Results of OneRing with different number of unknown categories on C2A of Office-Home.

C2A, shared classes = 10 (OS*/UNK/H)	**OneRing**
unknown classes = 5	65.6/82.8/73.2
unknown classes = 10	72.1/80.3/76.0
unknown classes = 50	65.9/73.0/69.3

Results with OS*, UNK and H on open-partial DA. In Tab. 11, we report OS*, UNK, OS and H on VisDA under open-partial DA, we outperform OVANet on the metrics of both OS and H.

Results on open-set DA. In Tab. 12, we report the results of several open-set or open-partial DA methods under open-set DA. Our method still get the best performance.

Results on closed-set DA In Tab. 13, we report the results of several UNDA methods under closed-set DA, our method is still superior to other methods.

OS*, UNK, H on Office-Home. In Tab. 14, we report the detailed results of our method under open-partial DA (the H may not be exactly the same as Tab. 6 in the submission, which is the average over three random run as mentioned in the paper details). As shown in the table, we can conclude that OneRing clearly has higher known accuracy compared to the source model (OneRing-S). After augmented with AaD, OneRing+ can get even higher known and unknown accuracy.

Detailed results with different number of unknown categories. In Tab. 15 and Tab. 16, we show the detailed results when the amount of unknown categories vary, the results are corresponding to Fig. 2.