

Generalizable Decision Boundaries: Dualistic Meta-Learning for Open Set Domain Generalization

Xiran Wang¹ Jian Zhang¹

¹Nanjing University

Lei Qi² Yinghuan Shi^{1*}

²Southeast University

{zzwdx, zhangjian7369}@smail.nju.edu.cn, qilei@seu.edu.cn, syh@nju.edu.cn

Abstract

Domain generalization (DG) is proposed to deal with the issue of domain shift, which occurs when statistical differences exist between source and target domains. However, most current methods do not account for a common realistic scenario where the source and target domains have different classes. To overcome this deficiency, open set domain generalization (OSDG) then emerges as a more practical setting to recognize unseen classes in unseen domains. An intuitive approach is to use multiple one-vs-all classifiers to define decision boundaries for each class and reject the outliers as unknown. However, the significant class imbalance between positive and negative samples often causes the boundaries biased towards positive ones, resulting in misclassification for known samples in the unseen target domain. In this paper, we propose a novel meta-learning-based framework called dualistic MEta-learning with joint DomaIn-Class matching (MEDIC), which considers gradient matching towards inter-domain and inter-class splits simultaneously to find a generalizable boundary balanced for all tasks. Experimental results demonstrate that MEDIC not only outperforms previous methods in open set scenarios, but also maintains competitive close set generalization ability at the same time. Our code is available at <https://github.com/zzwdx/MEDIC>.

1. Introduction

Deep neural networks have achieved enormous success in a wide range of computer vision tasks, usually assuming that the training and test samples are drawn from the same data distribution and label space. However, due to the un-

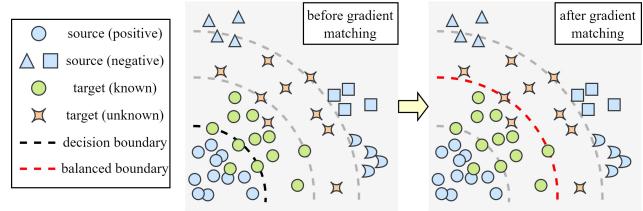


Figure 1. An example for the variation of decision boundaries of a one-vs-all classifier in open set domain generalization.

predictability of real-world application scenarios, the model faces the risk of performance degradation when the above constraints are not satisfied [26]. Domain generalization (DG) [46] is then motivated as a more realistic setting to deal with data distribution shift, which refers to using multiple source domains via data augmentation [29, 53], feature alignment [10, 27], meta-learning [23, 25] and so on, to obtain a model with the generalization ability that can be directly applied to arbitrary unseen target domains.

Most current domain generalization researches are based on the assumption of close set recognition, *i.e.*, the classes in the source domains are consistent with that of the target domains. However, in practice, the deployed model is often exposed to some new classes that have never been encountered in the training phase [40]. For example, in self-driving tasks, thousands of objects are fed into the model that the known label space cannot guarantee complete coverage of them; in medical image processing tasks, some diseases are extremely rare that it is unrealistic to acquire their samples for training. According to the logic of close set classification, these potential objects are forced to be identified into a known class, which lays hidden dangers for the model's robustness and security. To address this problem, it is necessary to investigate a more practical setting of open set domain generalization (OSDG), which focuses on recognizing unknown classes without losing the original classification accuracy.

In open set domain generalization [21, 43], the key is to concurrently deal with the problems of domain shift and category shift. We now turn our attention to a simple yet elegant learning paradigm, meta-learning [19], which has

*Corresponding author: Yinghuan Shi. Xiran Wang, Jian Zhang and Yinghuan Shi are with the State Key Laboratory for Novel Software Technology and National Institute of Healthcare Data Science, Nanjing University, China. Lei Qi is with the School of Computer Science and Engineering, Southeast University, China. This work is supported by NSFC Program (62222604, 62206052, 62192783), China Postdoctoral Science Foundation Project (2023T160100), Jiangsu Natural Science Foundation Project (BK20210224), and CCF-Lenovo Bule Ocean Research Fund.

shown its effectiveness in addressing domain shift. Previous research on meta-learning-based DG [23] and OSDG [43] aims to find an optimal balance between different domains through implicit gradient matching between tasks sampled from them. This domain-wise meta-learning helps prevent the model from becoming overly biased towards individual ones. The rest issue is how to effectively model the category shift. A natural idea is to use a multi-binary classifier [30, 39], which comprises multiple one-vs-all binary classifiers, to define a decision boundary for each known class. If a given sample is classified as negative by all sub-classifiers, it is considered to have a high probability of belonging to the unknown classes. However, we found this issue is non-trivial because the one-vs-all classifier in multi-classification tasks can generate biased predictions. As shown in Fig. 1, due to the limited data distribution of positive samples (*i.e.*, from the corresponding class) and the diverse data distribution of negative samples (*i.e.*, from all of the other classes), the classifier may asymmetrically focus on the negative ones which are easier to optimize. This can cause the decision boundary to cling to the positive samples, leading to inaccurate identification of known class samples in the unseen target domain. Conversely, if the decision boundary is too far away from the positive samples, the model can be prone to predicting an excessive quantity of target samples as known, hindering the recognition performance of unknown classes.

To establish a well-balanced decision boundary between classes, we attempt to fully exploit the gradient matching property of meta-learning. In contrast to the domain-wise strategy, we additionally sample tasks by classes to make the decision boundary rationally lying at the middle zone of each class. As a result, unknown samples are more likely to be distributed near the decision boundary, making them easier to be recognized. Concretely, we mitigate the limitation in previous works that merely select tasks by domains [38, 42] and propose a novel meta-learning strategy called *dualistic MEta-learning with joint DomaiN-Class matching (MEDIC)*. For tasks selected from different domains, we further divide them at category level. By matching the gradients of these divided tasks, we expect the model not only to generalize well across domains, but also to grasp the class-wise relationship more precisely. Furthermore, with the help of an auxiliary multi-binary classifier [30, 39], the model can automatically learn decision boundaries for each individual class.

2. Related Work

Domain Generalization (DG) is intended to learn a model from multiple source domains that can migrate to any unseen target domain with no additional retraining process. Existing domain generalization methods can be roughly divided into three genres: (i) **Feature Representation** [22] aims to extract universal features across source domains, assuming they are also applicable to the target domain, such as

Table 1. Comparison of target domains under different problem settings. The data distribution and label space of source domains are expressed as \mathcal{P} and \mathcal{C} respectively. The data distribution \mathcal{Q} is unseen and the label space \mathcal{U} satisfies $\mathcal{C} \cap \mathcal{U} = \emptyset$.

Problem Setting	Distribution of Data	Label Space	Participation in Training
Domain Adaptation [47]	\mathcal{Q}	\mathcal{C}	✓
Domain Generalization [46]	\mathcal{Q}	\mathcal{C}	✗
Open Set Recognition [14]	\mathcal{P}	$\mathcal{C} \cup \mathcal{U}$	✗
Open Set Domain Generalization [43]	\mathcal{Q}	$\mathcal{C} \cup \mathcal{U}$	✗

domain adversarial learning [12, 27] which trains a domain-irrelevant feature extractor guided by a domain-sensitive discriminator, feature decoupling [6] that distills domain-invariant features from the original ones. (ii) **Data Augmentation**. The goal of data augmentation is to further increase the diversity of training data to enhance the generalization ability of the model. For example, [15, 53] focus on mixing or perturbing statistics at image or feature level, while mapping data distribution [28, 52] and injecting random noise [29] are also proposed to generate artificial domains. (iii) **Learning Strategy**. Meta-learning [2, 10, 50] is aimed at leveraging past knowledge to improve the learning of current tasks. For example, MLDG [23] simulates domain shift by synthesizing meta-train and meta-test domains, which exerts a far-reaching influence on the following meta-learning-based researches. Ensemble learning [4, 54] attempts to reduce the risk of overfitting by combining models learned from different domains or sampled from a series of episodes. Gradient-based strategies perform regularization by shielding some gradients at each iteration [20] or intervening in the direction of gradient update [32, 42] which sometimes can be indirectly achieved through meta-learning.

Open Set Recognition (OSR) aims to detect novel classes not existing in the training set. From the perspective of whether to use extra data, current open set recognition methods can be classified into two categories. (i) **Artificial Classes**. [9, 18] propose to enrich the training set with classes of no interest to facilitate the feature separation of known classes. However, the effectiveness of these methods heavily relies on the quality of auxiliary samples and it may not be fair to compare them with other methods in a unified framework. Other methods [13, 36] adopt generative adversarial networks (GAN) to automatically generate samples of virtual unknown classes. (ii) **Discriminative Models**. OpenMax [3] abandons the traditional softmax layer and estimates the unknown probability of each sample with extreme value theory (EVT). Based on the assumption that the reconstructions of known class samples are usually more accurate, [37, 49] utilize reconstruction error to find the threshold for determining unknown classes. The idea of metric learning [7, 16] that aggregates intra-class and separate inter-class samples has also been applied to obtain models with higher discrimination ability. However, the above methods tend to treat all

out-of-distribution samples as unknown classes, making it challenging to directly apply them to the DG setting.

Open Set Domain Generalization (OSDG), which is listed with other problem settings in Table 1, is still in its early stages. There are very few related works as far as we know. [43] proposes a DAML framework based on domain augmentation and meta-learning. Since the model is not explicitly or implicitly informed by the existence of potential unknown classes in both training and inference phases, this work is in a way more focused on close set domain generalization instead. [21] borrows metric learning to diffuse the feature representations of unknown classes but relies on existing DG baselines to acquire domain-invariant features, so we think it should be more regarded as an OSR method with DG as a special case. In addition, we find two recent studies [48, 55] in open set single domain generalization (OS-SDG), which differs from OSDG in that only one source domain is available for training. [55] recommends imitating unknown classes via adversarial learning, and [48] forges unknown class samples by masking the labels of original ones. To sum up, the above methods usually have a sense of fragmentation when processing domain and category shifts, while our goal is to integrate them seamlessly.

3. Method

3.1. Problem Analysis

In open set domain generalization, we are provided with source domains $\mathcal{S} = \{\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_S\}$ with a label space \mathcal{C} and unseen target domains $\mathcal{T} = \{\mathcal{D}_{S+1}, \mathcal{D}_{S+2}, \dots, \mathcal{D}_{S+T}\}$ with an extended label space $\mathcal{C} \cup \mathcal{U}$ that ensure $\mathcal{C} \cap \mathcal{U} = \emptyset$. The s -th domain consisting of N_s samples is represented as $\mathcal{D}_s = \{(x_i^s, y_i^s)\}_{i=1}^{N_s}$, where x_i^s denotes the i -th sample from the sample space \mathcal{X} and y_i^s can take values in \mathcal{C} or $\mathcal{C} \cup \mathcal{U}$, signifying the corresponding label in the source or target domains. Our goal is to fully exploit these source domains \mathcal{S} to obtain a model that can directly generalize to any unseen target domain with unknown classes. The meta-learning [23] training scheme requires to split the source domains \mathcal{S} into meta-train set $\mathcal{S}_{\mathcal{F}}$ and meta-test set $\mathcal{S}_{\mathcal{G}}$, where $\mathcal{S}_{\mathcal{F}} \cup \mathcal{S}_{\mathcal{G}} = \mathcal{S}$ and $\mathcal{S}_{\mathcal{F}} \cap \mathcal{S}_{\mathcal{G}} = \emptyset$. We define the loss of the model trained on the data sampled from these two sets $\mathcal{S}_{\mathcal{F}}$ and $\mathcal{S}_{\mathcal{G}}$ as $\mathcal{F}(\Theta)$ and $\mathcal{G}(\Theta)$, respectively, where Θ is the parameter of a training model. First, we update the model's parameter to $\hat{\Theta}$ with the loss of $\mathcal{F}(\Theta)$. Then, we utilize the loss of $\mathcal{G}(\hat{\Theta})$ to update its original parameter Θ .

3.2. The MEDIC Framework

Dualistic Meta-Learning. Given two data splits $\mathcal{S}_{\mathcal{F}}$ and $\mathcal{S}_{\mathcal{G}}$ and their corresponding loss function \mathcal{F} and \mathcal{G} , our objective is to reach a consensus on their gradients $\mathcal{F}'(\Theta)$ and $\mathcal{G}'(\Theta)$ to ensure unbiased optimization direction for the prediction of $\mathcal{S}_{\mathcal{F}}$ and $\mathcal{S}_{\mathcal{G}}$. The rationale behind it is that if the

Algorithm 1 Training process of MEDIC

Input: Domains \mathcal{S} ; classes \mathcal{C} ; model parametrized by Θ ; loss function \mathcal{L} ; learning rates α, β and η

- 1: **while** Θ not converged **do**
- 2: **Random split** $\mathcal{S}_1, \mathcal{S}_2 \leftarrow \mathcal{S}; \mathcal{C}_1, \mathcal{C}_2 \leftarrow \mathcal{C}$
- 3: **Sample** $\mathcal{B}_{\mathcal{F}_1}$ from $(\mathcal{S}_1, \mathcal{C}_1)$; $\mathcal{B}_{\mathcal{F}_2}$ from $(\mathcal{S}_1, \mathcal{C}_2)$; $\mathcal{B}_{\mathcal{G}_1}$ from $(\mathcal{S}_2, \mathcal{C}_1)$; $\mathcal{B}_{\mathcal{G}_2}$ from $(\mathcal{S}_2, \mathcal{C}_2)$
- 4: **Meta-train:** $\mathcal{L}_1 \leftarrow \mathcal{L}(\mathcal{B}_{\mathcal{F}_1}; \Theta) + \mathcal{L}(\mathcal{B}_{\mathcal{G}_1}; \Theta)$
 $\hat{\Theta} \leftarrow \Theta - \alpha \cdot \nabla_{\Theta} \mathcal{L}_1$
- 5: **Meta-test:** $\mathcal{L}_2 \leftarrow \mathcal{L}(\mathcal{B}_{\mathcal{F}_2}; \hat{\Theta}) + \mathcal{L}(\mathcal{B}_{\mathcal{G}_2}; \hat{\Theta})$
 $\Theta \leftarrow \Theta - \eta \cdot (\nabla_{\Theta} \mathcal{L}_1 + \beta \cdot \nabla_{\Theta} \mathcal{L}_2)$
- 6: **end while**

angle between the directions of $\mathcal{F}'(\Theta)$ and $\mathcal{G}'(\Theta)$ is small which means optimizing one task does not hurt the other task, then updating with their combined gradient (*i.e.*, summed gradients in practice) can result in a better performance. Instead, if the angle between them is large which means these two tasks conflict with each other, the optimization procedure will be damaged and lead to an inferior optimization process. The key idea of gradient matching is to find a position in the weight space where the angle between the gradient of $\mathcal{S}_{\mathcal{F}}$ and $\mathcal{S}_{\mathcal{G}}$ is small, which is equivalent to maximize the dot product of $\mathcal{F}'(\Theta)$ and $\mathcal{G}'(\Theta)$ [23]. After moving the model towards this area in the last training step, $\mathcal{S}_{\mathcal{F}}$ and $\mathcal{S}_{\mathcal{G}}$ will reach an agreement on the direction of gradient update.

Current gradient-based domain generalization methods often split $\mathcal{S}_{\mathcal{F}}$ and $\mathcal{S}_{\mathcal{G}}$ into different domains to find an optimization direction that is invariant only among domains [23]. However, these methods ignore the crucial relationship between classes which is necessary for open set recognition. Instead of simply adding an extra iteration to avoid biased prediction between classes, we take a step further and propose a novel meta-learning strategy named *dualistic MEdia-learning with joint DomaiIn-Class matching (MEDIC)*, which achieves gradient matching towards inter-domain and inter-class splits simultaneously to learn a more generalizable decision boundary balanced for all tasks.

As illustrated in Fig. 2, for $\mathcal{S}_{\mathcal{F}}$ and $\mathcal{S}_{\mathcal{G}}$ derived from disparate source domains, we further divide them into $\mathcal{S}_{\mathcal{F}_1}, \mathcal{S}_{\mathcal{F}_2}$ and $\mathcal{S}_{\mathcal{G}_1}, \mathcal{S}_{\mathcal{G}_2}$ by classes and define their corresponding loss functions as $\mathcal{F}_1, \mathcal{F}_2$ and $\mathcal{G}_1, \mathcal{G}_2$. The label spaces between $\mathcal{S}_{\mathcal{F}_1}$ and $\mathcal{S}_{\mathcal{F}_2}$ as well as between $\mathcal{S}_{\mathcal{G}_1}$ and $\mathcal{S}_{\mathcal{G}_2}$ are both disjoint. Besides, we require that $\mathcal{S}_{\mathcal{F}_1}$ and $\mathcal{S}_{\mathcal{G}_1}$ share the same labels, and likewise for $\mathcal{S}_{\mathcal{F}_2}$ and $\mathcal{S}_{\mathcal{G}_2}$. Then to implement the gradient matching between domains and classes simultaneously, we employ $(\mathcal{S}_{\mathcal{F}_1}, \mathcal{S}_{\mathcal{G}_2})$ as meta-train set and $(\mathcal{S}_{\mathcal{F}_2}, \mathcal{S}_{\mathcal{G}_1})$ as meta-test set respectively. The final meta-objective function of MEDIC is defined as follows:

$$\operatorname{argmin}_{\Theta} [\mathcal{F}_1(\Theta) + \mathcal{G}_2(\Theta)] + \beta [\mathcal{F}_2(\hat{\Theta}) + \mathcal{G}_1(\hat{\Theta})], \quad (1)$$

where β measures the weight between the two meta sets and

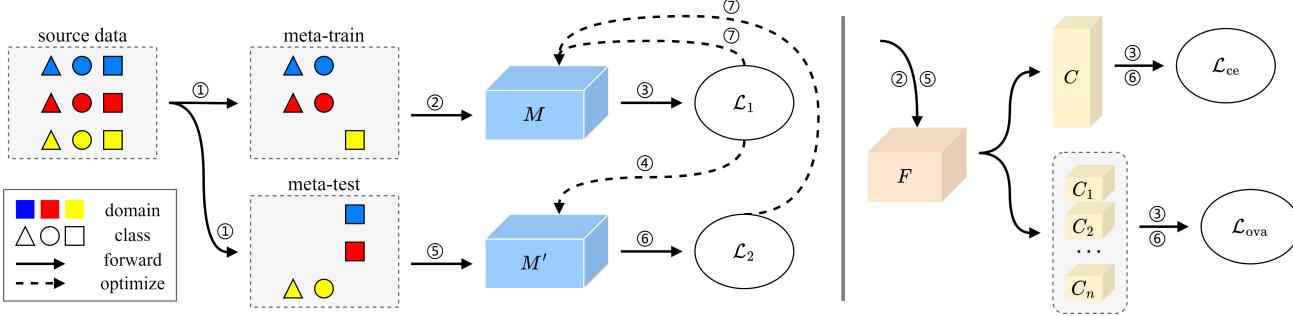


Figure 2. Overview of MEDIC during one training iteration. M is the overall model, and the right figure represents its internal structure. The numbers denote the sequence of data flow (solid arrows) and model updates (dashed arrows) respectively.

$\hat{\Theta}$ is the optimized model parameters on the meta-train set with learning rate α :

$$\hat{\Theta} = \Theta - \alpha(\mathcal{F}'_1(\Theta) + \mathcal{G}'_2(\Theta)). \quad (2)$$

To verify MEDIC can exactly achieve gradient matching between domains and classes at the same time, similar to the analysis in [23], we perform the first order Taylor expansion for the second term in Eq. (1):

$$\mathcal{F}_2(\hat{\Theta}) = \mathcal{F}_2(\Theta) - \alpha(\mathcal{F}'_1(\Theta) + \mathcal{G}'_2(\Theta)) \cdot \mathcal{F}'_2(\Theta), \quad (3)$$

$$\mathcal{G}_1(\hat{\Theta}) = \mathcal{G}_1(\Theta) - \alpha(\mathcal{F}'_1(\Theta) + \mathcal{G}'_2(\Theta)) \cdot \mathcal{G}'_1(\Theta), \quad (4)$$

and the objective function becomes:

$$\begin{aligned} \underset{\Theta}{\operatorname{argmin}} & [\mathcal{F}_1(\Theta) + \mathcal{G}_2(\Theta) + \beta(\mathcal{F}_2(\Theta) + \mathcal{G}_1(\Theta))] \\ & - \beta\alpha[(\mathcal{F}'_1(\Theta) + \mathcal{G}'_2(\Theta)) \cdot (\mathcal{F}'_2(\Theta) + \mathcal{G}'_1(\Theta))]. \end{aligned} \quad (5)$$

The first term of Eq. (5) means optimizing the model with their original losses. The second term is the sum of the product of gradients. By expanding this term, we can obtain the following regularization term:

$$\mathcal{L}_{\text{reg}} = -(\mathcal{F}'_1 \cdot \mathcal{F}'_2 + \mathcal{F}'_1 \cdot \mathcal{G}'_1 + \mathcal{G}'_2 \cdot \mathcal{F}'_2 + \mathcal{G}'_2 \cdot \mathcal{G}'_1). \quad (6)$$

As previously mentioned, maximizing the dot product of gradients can regularize the training procedure that matches the updating directions of different tasks. Since for $\mathcal{S}_{\mathcal{F}_1}$, the counterpart $\mathcal{S}_{\mathcal{G}_1}$ contains the same classes in different domains, while $\mathcal{S}_{\mathcal{F}_2}$ contains different classes in the same domain, the two terms in any one of the above gradient products are either from different domains or different classes, allowing for domain-wise and class-wise matching simultaneously. Compared with conventional methods which only consider inter-domain relationships, the dot product between $\mathcal{S}_{\mathcal{F}_1}$ and $\mathcal{S}_{\mathcal{F}_2}$ fills the gap of class-wise gradient matching inside each domain to learn a more reasonable decision boundary, achieving a more fine-grained model optimization.

Open Set Loss Function. For open set recognition, we adopt another multi-binary classifier [39] to serve as a supplement to conventional close set classifier. As illustrated

in Fig. 2, the proposed classifier consists of $|\mathcal{C}|$ one-vs-all classifiers, where each classifier is trained to detect whether a sample belongs to the corresponding class. Let $p(\hat{y}^k|x)$ denote the output probability that an instance x is an inlier of the k -th sub-classifier. For a given sample (x, y) , its loss on the multi-binary classifier can be presented as:

$$\mathcal{L}_{\text{ova}}(x, y) = -\log(p(\hat{y}^y|x)) - \min_{j \neq y} \log(1 - p(\hat{y}^j|x)). \quad (7)$$

The second term denotes that it updates only the most challenging binary classifier when used as a negative sample. We simply adopt this loss and train a close set classifier using cross-entropy loss which is denoted as \mathcal{L}_{ce} . Then the overall open set loss function can be defined as follows:

$$\mathcal{L}_{\text{all}} = \mathcal{L}_{\text{ce}} + \mathcal{L}_{\text{ova}}, \quad (8)$$

which is equivalent to the meta-objective of $\mathcal{F}_1, \mathcal{F}_2, \mathcal{G}_1, \mathcal{G}_2$. The optimization of the multi-binary classifier is facilitated by MEDIC. Although the proposed classifier can learn a boundary between inliers and outliers for each known class, negative training samples for a certain sub-classifier differ a lot among $|\mathcal{C}| - 1$ classes, which may lead to confusion in the direction of gradient descent, causing damage to the gradient of positive ones. The application of inter-class gradient alignment can stabilize both of them, thus seeking a balance between close set generalization and open set recognition.

3.3. Inference

In the test phase, each target sample is first predicted by the close set classifier to obtain a probability distribution $p(\hat{y}|x)$ among known classes. The model either (i) selects the value of its maximum likelihood with the original classifier:

$$\text{conf}_{\text{cls}}(x) = \max_{i=1}^{|\mathcal{C}|} (p(\hat{y}|x)_i), \quad (9)$$

or (ii) refers to the corresponding one-vs-all classifier and choose the value of its positive output channel as the confidence score [39]:

$$\text{conf}_{\text{bcls}}(x) = p(\hat{y}^{\operatorname{argmax}_{i=1}^{|\mathcal{C}|} (p(\hat{y}|x)_i)}|x). \quad (10)$$

Table 2. Results (%) of PACS on ResNet18. The ratio of known to unknown classes is 6:1. (Best in **bold**)

Method	Photo			Art			Cartoon			Sketch			Avg		
	Acc	H-score	OSCR												
OpenMax [†] [3]	95.56	92.48	-	83.68	69.61	-	78.61	64.36	-	70.89	50.67	-	82.19	69.28	-
ERM [35]	96.04	93.40	95.11	84.18	70.54	71.89	77.63	62.80	62.57	70.44	55.81	51.75	82.07	70.64	70.33
ARPL [†] [7]	94.83	95.06	94.63	83.93	67.88	68.82	78.56	62.98	65.30	74.34	61.20	59.80	82.91	71.78	72.14
MMLD [33]	94.83	88.80	92.94	84.43	64.83	69.43	77.11	64.21	65.36	75.14	67.70	64.69	82.88	71.38	73.11
RSC [20]	94.43	88.37	91.38	83.36	70.27	73.55	78.09	65.13	66.15	77.16	52.98	62.31	83.26	69.19	73.35
DAML [†] [43]	91.44	80.87	82.83	83.11	72.05	71.75	79.11	66.26	66.46	82.97	72.63	73.71	84.16	72.95	73.69
MixStyle [53]	95.23	82.02	88.99	86.18	70.62	72.57	78.92	63.23	63.81	80.34	71.90	72.07	85.17	71.94	74.36
SelfReg [22]	95.72	89.34	92.26	86.24	72.45	73.77	80.77	65.75	66.38	78.30	67.06	65.69	85.26	73.65	74.53
MLDG [23]	94.99	91.48	93.70	84.12	69.52	72.15	78.45	61.59	64.32	79.99	69.67	68.60	84.39	73.06	74.69
MVDG [50]	94.43	74.07	88.07	87.62	71.98	75.05	81.18	63.95	66.34	82.41	73.55	73.83	86.41	70.89	75.82
MEDIC-cls	94.83	83.68	90.30	86.20	69.35	74.16	81.94	63.26	67.43	81.84	69.60	70.85	86.20	71.47	75.69
MEDIC-bcls	94.83	89.49	92.40	86.20	73.82	75.58	81.94	66.26	69.04	81.84	74.37	74.52	86.20	75.98	77.89

Table 3. Results (%) of PACS on ResNet50. The ratio of known to unknown classes is 6:1. (Best in **bold**)

Method	Photo			Art			Cartoon			Sketch			Avg		
	Acc	H-score	OSCR												
OpenMax [†] [3]	97.58	93.09	-	88.37	73.91	-	84.38	68.23	-	80.07	68.06	-	87.60	75.82	-
ARPL [†] [7]	97.09	96.81	96.86	88.24	77.48	80.32	82.68	67.19	68.31	78.08	70.04	69.47	86.52	77.88	78.74
MIRO [5]	94.85	92.32	93.27	88.51	65.02	79.01	82.98	63.05	73.72	82.22	69.47	70.61	87.14	72.47	79.15
MLDG [23]	96.77	95.85	96.33	87.99	77.16	79.93	83.45	68.74	71.32	82.25	73.16	72.27	87.61	78.73	79.96
ERM [35]	97.09	96.58	96.68	89.99	76.05	82.44	85.10	65.79	70.59	80.31	70.29	70.16	88.12	77.18	79.97
CIRL [31]	96.53	87.75	95.40	92.06	70.75	77.44	85.71	68.82	73.71	84.35	66.73	77.24	89.66	73.51	80.95
MixStyle [53]	96.53	93.57	95.30	90.87	79.15	83.27	86.80	68.08	74.68	84.88	71.57	73.41	89.77	78.09	81.66
CrossMatch [†] [55]	96.53	96.34	96.12	91.37	75.67	82.32	83.92	67.02	74.55	81.61	72.03	73.99	88.37	77.76	81.75
SWAD [4]	96.37	84.56	93.24	93.75	68.41	85.00	85.57	58.57	75.90	81.90	74.66	74.65	89.40	71.55	82.20
MVDG [50]	97.17	95.02	96.63	92.50	79.47	85.02	86.02	71.05	76.03	83.44	75.24	75.18	89.78	80.20	83.21
MEDIC-cls	96.37	93.80	95.37	91.62	80.80	84.67	86.65	75.85	77.48	84.61	75.80	76.79	89.81	81.56	83.58
MEDIC-bcls	96.37	94.75	95.79	91.62	81.61	85.81	86.65	77.39	78.30	84.61	78.35	79.50	89.81	83.03	84.85

If the score is larger than a preset threshold μ , then detect the sample as known and generate a particular class label, otherwise judge it as unknown. Experimental results are reported on these two inference modes in Sec. 4. We use `cls` and `bcls` to express discriminating unknown classes with close set classifier and multi-binary classifier respectively.

4. Experiment

4.1. Setup

We experiment on three standard DG datasets whose details are described as follows: (i) **PACS** [24] contains four domains (*photo*, *art-painting*, *cartoon*, *sketch*) with a total number of 9,991 pictures. All of the domains share the same label space of 7 classes but differ in image style. Making use of the official split adopted from [24] for training and validation, we evaluate the generalization ability of our model in both close set and open set scenarios. (ii) **Office-Home** [45] comprises around 15,500 images of 65 classes from four domains (*art*, *clipart*, *product*, *real-world*). The richness of classes allows us to select the ratio of known to unknown in a relatively flexible way. With the train-validation split provided by [43], we conduct experiments in various known-unknown divisions on this dataset. (iii) **Digits-DG** [51] is an aggre-

gation of four classic benchmarks including *mnist*, *mnist-m*, *svhn* and *syn*. Digit datasets are popular in OSR since the close set accuracy of the model tends to saturate, making it purer to verify the capability of unknown class recognition. We utilize the original train-validation split in [51] for our open set domain generalization tasks.

The leave-one-domain-out evaluation protocol is carried out on all benchmarks, *i.e.*, picking one target domain for testing and using the remaining ones for training and validation. The classes are then rearranged in alphabetical order, with the former part designated as known classes and the latter part as unknown classes. The split rate of known and unknown classes on each dataset are explained in the caption of the corresponding table. For PACS and Office-Home, we employ ResNet18 and ResNet50 [17] pretrained on ImageNet [8] as our backbone networks. For Digits-DG which is relatively simple to handle, we borrow a lightweight convolutional architecture called ConvNet from [51]. The validation set does not contain any unseen classes, and we use close set accuracy for model selection.

4.2. Evaluation Metrics

Considering the trade-off between the accuracy of known and unknown classes, we choose three evaluation metrics to verify the performance of our model: (i) **Acc** is used to represent the typical close set accuracy. (ii) **H-score** [11] is

[†]Open-set-based methods.

Table 4. Results (%) of Digits-DG on ConvNet. The ratio of known to unknown classes is 6:4. (Best in **bold**)

Method	MNIST			MNIST-M			SVHN			SYN			Avg		
	Acc	H-score	OSCR												
OpenMax [†] [3]	97.33	52.03	-	71.03	57.26	-	72.00	49.46	-	84.83	54.78	-	81.30	53.38	-
MixStyle [53]	97.86	73.25	89.36	74.50	59.30	56.95	69.28	53.24	48.43	85.06	60.22	65.44	81.68	61.50	65.05
ERM [35]	97.47	80.90	92.60	71.03	53.92	54.04	71.08	54.37	49.86	85.67	51.57	67.63	81.31	60.19	66.03
ARPL [†] [7]	97.75	85.74	91.86	69.78	58.08	54.21	71.78	56.98	53.63	85.31	64.04	65.89	81.16	66.21	66.40
MLDG [23]	97.83	80.36	94.28	71.11	46.84	55.17	73.64	53.54	53.64	86.08	63.56	70.34	82.16	61.08	68.36
SWAD [4]	97.71	84.44	92.65	73.09	53.35	55.94	76.08	59.18	56.25	87.95	51.27	69.03	83.71	62.06	68.47
MEDIC-cls	97.89	67.37	96.17	71.14	48.44	55.37	76.00	51.20	55.58	88.11	64.90	73.62	83.28	57.98	70.19
MEDIC-bcls	97.89	83.20	95.81	71.14	60.98	58.28	76.00	58.77	57.60	88.11	62.24	72.91	83.28	66.30	71.15

the harmonic mean of known class accuracy acc_k and unknown class accuracy acc_u , which is widely used in current OSDA and OSDG studies. When $\text{acc}_k + \text{acc}_u$ is constant, the smaller the difference between acc_k and acc_u is, the larger the H-score will be. Compared with the weighted average, H-score puts more emphasis on the balance between close set classification and open set recognition. However, due to the uncertainty of domain shift, the manually designed threshold to reject unknown classes is not applicable for a random target domain with arbitrary image styles. As shown in Fig. 3, there is a gap between the optimal threshold of the source and target domains. To mitigate the unreliability of artificial selection, we propose to import a threshold-independent metric (iii) open set classification rate (**OSCR**) [9] which plots the true positive rate against the false positive rate through an ever-moving threshold. Different from area under the receiver operating characteristic (AUROC) [36] that ignores the accuracy of known classes, in OSCR only the correctly labeled samples are considered as true positive ones.

4.3. Results

We carry out experiments premised on the existence of unknown target classes. The results on PACS and Digits-DG are presented in Table 2, Table 3, and Table 4. We first compare our method with DG and OSDG methods such as MLDG [23], MVDG [50] and DAML [43]. Although the accuracy of the known classes is relatively similar, the open set performance of MEDIC with multi-binary classifier is significantly superior to other methods. For example, in Table 2, MEDIC exceeds the previous OSDG method DAML [43] by 4.20% on the average of OSCR with even fewer model parameters and a higher close set accuracy, while well ahead of the second best method MVDG [50] by 2.07%, demonstrating that our method is capable of producing more generalizable and discriminative representations which is beneficial for both DG and OSR tasks.

We also compare with several open set recognition methods such as OpenMax [3] and ARPL [7]. Note that we do not calculate the OSCR of OpenMax [3], because it already possesses a threshold-independent property, *i.e.*, the classifier is designed with $|\mathcal{C}| + 1$ output channels, one of which is corresponding to the probability of the unknown classes.

However, the H-score of OpenMax is still below average, further proving that the way of hard inference derived from source domains is not suitable for the unseen target domains. It can be observed that ARPL [7], which is one of the state-of-the-art approaches for open set recognition, fails to perform well than the methods that are tailored for standard domain generalization. This may indicate that deep learning models possess a natural inclination to identify unknown classes to some extent and the issue of close set classification under distribution shift remains a top priority in OSDG.

4.4. Ablation Study

On the effect of different learning paradigms. We conduct ablation studies to reveal the importance of the dualistic meta-learning scheme in our method. To this end, we compare MEDIC with the baseline ERM [35] and the most related method MLDG [23]. The core idea of MLDG is to simulate the virtual target domain at each learning iteration, regardless of the relationship among classes. When using the same loss function and model architecture, the training strategy becomes the only variable between MEDIC and other methods. As shown in Table 5, MEDIC outperforms the above methods no matter which option is uniformly appointed, proving that dualistic gradient matching plays a critical role in improving open set performance. Moreover, after switching from cross-entropy loss (*i.e.*, training with close set classifier only) to open set loss function (*i.e.*, training with the two classifiers), our method ushers in the largest performance gain on both H-score and OSCR by the same average value of 2.9%, indicating that the proposed strategy has better compatibility with the multi-binary classifier to learn a generalizable boundary for each known class.

Verify the unbiased decision boundaries. In practice, the shape and position of decision boundaries are influenced by the feature extractor. An unbiased decision boundary implies a clear separation between known and unknown classes, while a biased one indicates that the features of known classes tend to resemble those of unknown, making it difficult to differentiate between them. We provide t-SNE results of feature representations in Fig. 4. It can be observed that unknown classes are typically distributed around the central region. In MEDIC, the overlap between known and

Table 5. Ablation studies (%) of classifiers and parameter sharing on PACS using ResNet50. D_w and C_w refer to domain-wise and class-wise gradient matching that can be achieved by the corresponding method. The first three rows of each method denote (i) training and inference with close set classifier only, (ii) training with two classifiers and inference with close set classifier, (iii) training with two classifiers and inference with multi-binary classifier. The option *share* means sharing parameters between the two classifiers.

Method	D_w	C_w	Option	P	A	C	S	Avg
H-Score								
ERM [35]	-	-	-	96.6	76.1	65.8	70.3	77.2
			cls	97.1	77.3	64.5	69.9	77.2
			bcls	97.3	79.5	71.1	71.7	79.9
MixStyle [53]	-	-	-	93.6	79.2	68.1	71.6	78.1
			cls	90.9	74.8	67.6	74.8	77.0
			bcls	91.2	78.8	71.5	77.6	79.8
-	-	✓	-	93.2	79.2	66.5	71.5	77.6
			cls	95.0	79.3	69.0	76.3	79.9
			bcls	95.4	79.1	69.9	76.6	80.3
MLDG [23]	✓	-	-	95.8	77.2	68.7	73.2	78.7
			cls	95.8	77.9	67.2	69.4	77.6
			bcls	95.6	80.1	72.2	71.6	79.9
MEDIC	✓	✓	-	94.1	80.5	71.3	74.5	80.1
			cls	93.8	80.8	75.9	75.8	81.6
			bcls	94.8	81.6	77.4	78.3	83.0
			share	95.6	81.0	76.0	77.3	82.5
OSCR								
ERM [35]	-	-	-	96.7	82.4	70.6	70.2	80.0
			cls	97.3	83.9	70.7	70.9	80.7
			bcls	97.1	83.8	71.1	72.0	81.0
MixStyle [53]	-	-	-	95.3	83.3	74.7	73.4	81.7
			cls	94.5	83.3	74.9	76.6	82.3
			bcls	94.4	83.3	74.1	77.4	82.3
-	-	✓	-	95.2	81.8	72.6	71.5	80.3
			cls	95.7	83.9	73.7	75.8	82.3
			bcls	95.6	84.1	75.0	77.0	82.9
MLDG [23]	✓	-	-	96.3	79.9	71.3	72.3	80.0
			cls	96.7	83.1	74.6	73.5	82.0
			bcls	96.8	83.3	75.4	74.3	82.5
MEDIC	✓	✓	-	95.1	83.7	73.7	75.5	82.0
			cls	95.4	84.7	77.5	76.8	83.6
			bcls	95.8	85.8	78.3	79.5	84.9
			share	96.2	85.2	77.6	79.4	84.6

unknown classes seems smaller, while reserving a wider space for other potential unknown classes.

Varying the ratio of known to unknown classes. The openness reflects the proportion of unknown target classes at test time. We conduct experiments on Office-Home to verify the adaptability of our model at different levels of openness. The results of OSCR are visually displayed in Fig. 5. It is obvious that increasing the number of classes poses greater challenges to the classification task, resulting in lower accuracy of the model. On the other hand, MEDIC continues to outperform others with noticeable improvement, indicating the robustness of our method in various scenarios.

Varying the threshold of rejecting unknown classes. Although OSCR provides a relatively fair way of model

Table 6. Close set accuracy (%) of PACS on ResNet50.

Method	P	A	C	S	Avg
IRM [1]	96.7	84.8	76.4	76.1	83.5
MetaReg [2]	97.6	87.2	79.2	70.3	83.6
ERM [35]	97.4	85.7	77.1	76.6	84.2
MMD [27]	96.6	86.1	79.4	76.5	84.7
MLDG [23]	97.4	85.5	80.1	76.6	84.9
RSC [20]	97.6	85.4	79.7	78.2	85.2
Mixstyle [53]	96.6	86.8	79.0	78.5	85.2
CORAL [44]	97.5	88.3	80.0	78.8	86.2
DSON [41]	96.0	87.0	80.6	82.9	86.6
SWAD [4]	97.3	89.3	83.4	82.5	88.1
MEDIC	97.3	88.5	84.4	83.0	88.3

Table 7. Close set accuracy (%) of Office-Home on ResNet50.

Method	A	C	P	R	Avg
Mixstyle [53]	51.1	53.2	68.2	69.2	60.4
IRM [1]	58.9	52.2	72.1	74.0	64.3
RSC [20]	60.7	51.4	74.8	75.1	65.5
DANN [12]	59.9	53.0	73.6	76.9	65.9
MMD [27]	60.4	53.3	74.3	77.4	66.4
MLDG [23]	61.5	53.2	75.0	77.5	66.8
ERM [35]	63.1	51.9	77.2	78.1	67.6
SagNet [34]	63.4	54.8	75.8	78.3	68.1
CORAL [44]	65.3	54.4	76.5	78.4	68.7
SWAD [4]	66.1	57.7	78.4	80.2	70.6
MEDIC	68.1	58.1	78.7	79.9	71.2

comparison, in practice, we have to select a specific threshold to tell knowns apart from unknowns. The variation of H-score with thresholds on PACS is presented in Fig. 3. We can observe that the proposed MEDIC consistently precedes MLDG with a large margin on two target domains: *cartoon*, *sketch*, and also shows a later advantage on *art-painting* when the threshold is around 0.9, illustrating that MEDIC can maintain stable performance under different thresholds. It can be also noticed that results on *photo* are usually higher than other domains and the vanilla ERM outperforms competing methods on this domain. This may indicate that when served as a target domain, *photo* shows a higher inclination towards specific source domains, limiting the efficacy of domain-agnostic DG methods to leverage their inherent advantages. The ideal threshold for source domains is not always reliable. Take *photo* and *sketch* for example, when the model reaches its maximum H-score on the source domains, the corresponding value on the target domain is less than satisfactory, which indicates that determining the optimal threshold remains a challenging task to be addressed.

Close set domain generalization. We evaluate the close set classification accuracy on PACS and Office-Home. As shown in Table 6 and Table 7, we make a list of representative approaches in a variety of DG genres to confirm the superiority of our proposed method. In particular, MEDIC achieves evenly matched or better performance on all tar-

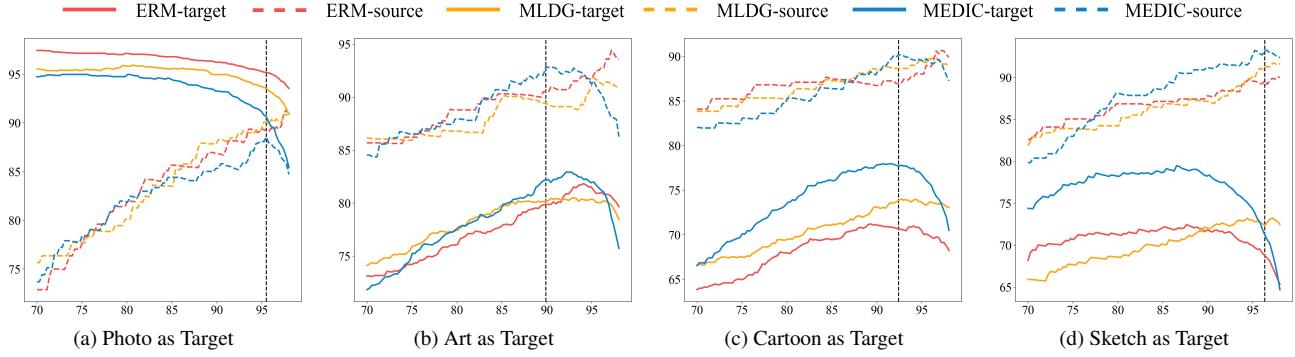


Figure 3. The H-score (%) with the varying threshold (%) on PACS using ResNet50. The solid and dashed lines represent results on the target domain and the validation set of source domains respectively. The vertical lines mark the optimal threshold on the source domains.

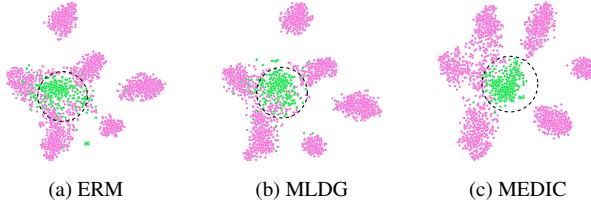


Figure 4. T-SNE results of features in the target domain, where pink and green represents known and unknown classes respectively.

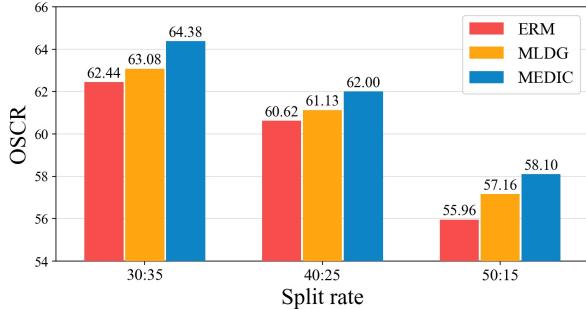


Figure 5. The values of OSCR (%) with the varying known-unknown divisions on Office-Home using ResNet18.

get domains compared with one of the state-of-the-art algorithms SWAD [4], which highlights the competitiveness of our method in the field of classic domain generalization.

4.5. Analysis & Discussion

Sharing parameters between classifiers. It can be observed that each distinct known class is associated with a unique one-vs-all classifier and a single output channel in the close set classifier. This channel's functioning mirrors the positive output channel of the corresponding binary classifier, implying activation by samples from the same class. This naturally prompts us to consider the feasibility of parameter sharing between these channels. In the MEDIC part of Table 5, the third and fourth rows demonstrate that parameter sharing yields comparable performance to the original architecture while reducing the total output channels from $3|\mathcal{C}|$ to

$2|\mathcal{C}|$. If we assume linearity for all classifiers, the number of reduced parameters in the improved model is equivalent to that of a complete close set classifier.

Intuition for choosing datasets. Since there are no widely acknowledged datasets for open set domain generalization, we experiment on standard domain generalization benchmarks. Our criteria for selecting datasets are as follows: (i) The dataset should have been well settled in the conventional DG setting because when an unknown sample is given a low confidence score, we prefer it to be a reasonable rejection based on sufficient understanding of known classes rather than a random guess due to the lack of adequate knowledge at all. (ii) Ideally, the semantics of each sample should correspond to an individual label. Suppose there are *dog* and *cat* in the label space and we have an image labeled as *dog* that possesses the primary semantic of *dog* and the secondary semantic of *cat*. In close set classification, it is reasonable to treat it as *dog* because the dog holds the principal status in this image. However, in open set experiments, some labels are converted into unknowns such as *dog*. We thus get a wrong sample with semantic of *cat* but label of unknown, which can interfere with the fairness in evaluation. Therefore, the semantics of the samples in the label space should be as pure as possible. We hope the above two thoughts could provide a clue for future studies.

5. Conclusion

In this paper, we introduce the practical problem setting of open set domain generalization, which aims to tackle the challenges of domain shift and category shift concurrently in the unseen target domain. To address this, we propose a simple yet effective meta-learning-based framework, which considers both domain-wise and class-wise gradient matching simultaneously, along with a multi-binary classifier to learn a balanced decision boundary for each known class. We conduct experiments using comprehensive evaluation metrics on multiple benchmarks and demonstrate its superior performance in both close set and open set scenarios.

References

- [1] Martin Arjovsky, Léon Bottou, Ishaaan Gulrajani, and David Lopez-Paz. Invariant risk minimization. *Stat*, 1050:27, 2020. [7](#)
- [2] Yogesh Balaji, Swami Sankaranarayanan, and Rama Chellappa. Metareg: Towards domain generalization using meta-regularization. *Advances in Neural Information Processing Systems*, 31, 2018. [2, 7](#)
- [3] Abhijit Bendale and Terrance E Boult. Towards open set deep networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1563–1572, 2016. [2, 5, 6](#)
- [4] Junbum Cha, Sanghyuk Chun, Kyungjae Lee, Han-Cheol Cho, Seunghyun Park, Yunsung Lee, and Sungrae Park. Swad: Domain generalization by seeking flat minima. *Advances in Neural Information Processing Systems*, 34:22405–22418, 2021. [2, 5, 6, 7, 8](#)
- [5] Junbum Cha, Kyungjae Lee, Sungrae Park, and Sanghyuk Chun. Domain generalization by mutual-information regularization with pre-trained models. In *European Conference on Computer Vision*, pages 440–457, 2022. [5](#)
- [6] Prithvijit Chattopadhyay, Yogesh Balaji, and Judy Hoffman. Learning to balance specificity and invariance for in and out of domain generalization. In *European Conference on Computer Vision*, pages 301–318, 2020. [2](#)
- [7] Guangyao Chen, Peixi Peng, Xiangqian Wang, and Yonghong Tian. Adversarial reciprocal points learning for open set recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021. [2, 5, 6](#)
- [8] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, 2009. [5](#)
- [9] Akshay Raj Dhamija, Manuel Günther, and Terrance Boult. Reducing network agnostophobia. *Advances in Neural Information Processing Systems*, 31, 2018. [2, 6](#)
- [10] Qi Dou, Daniel Coelho de Castro, Konstantinos Kamnitsas, and Ben Glocker. Domain generalization via model-agnostic learning of semantic features. *Advances in Neural Information Processing Systems*, 32, 2019. [1, 2](#)
- [11] Bo Fu, Zhangjie Cao, Mingsheng Long, and Jianmin Wang. Learning to detect open classes for universal domain adaptation. In *European Conference on Computer Vision*, pages 567–583, 2020. [5](#)
- [12] Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario Marchand, and Victor Lempitsky. Domain-adversarial training of neural networks. *The Journal of Machine Learning Research*, 17(1):2096–2030, 2016. [2, 7](#)
- [13] Zongyuan Ge, Sergey Demyanov, Zetao Chen, and Rahil Garnavi. Generative openmax for multi-class open set classification. In *British Machine Vision Conference 2017*, 2017. [2](#)
- [14] Chuanxing Geng, Sheng-jun Huang, and Songcan Chen. Recent advances in open set recognition: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(10):3614–3631, 2020. [2](#)
- [15] Jintao Guo, Na Wang, Lei Qi, and Yinghuan Shi. Aloft: A lightweight mlp-like architecture with dynamic low-frequency transform for domain generalization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 24132–24141, 2023. [2](#)
- [16] Yunrui Guo, Guglielmo Camporese, Wenjing Yang, Alessandro Sperduti, and Lamberto Ballan. Conditional variational capsule network for open set recognition. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 103–111, 2021. [2](#)
- [17] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016. [5](#)
- [18] Dan Hendrycks, Mantas Mazeika, and Thomas Dietterich. Deep anomaly detection with outlier exposure. In *International Conference on Learning Representations*, 2018. [2](#)
- [19] Timothy Hospedales, Antreas Antoniou, Paul Micaelli, and Amos Storkey. Meta-learning in neural networks: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(9):5149–5169, 2021. [1](#)
- [20] Zeyi Huang, Haohan Wang, Eric P Xing, and Dong Huang. Self-challenging improves cross-domain generalization. In *European Conference on Computer Vision*, pages 124–140, 2020. [2, 5, 7](#)
- [21] Kai Katsumata, Ikki Kishida, Ayako Amma, and Hideki Nakayama. Open-set domain generalization via metric learning. In *2021 IEEE International Conference on Image Processing*, pages 459–463, 2021. [1, 3](#)
- [22] Daehee Kim, Youngjun Yoo, Seunghyun Park, Jinkyu Kim, and Jaekoo Lee. Selfreg: Self-supervised contrastive regularization for domain generalization. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9619–9628, 2021. [2, 5](#)
- [23] Da Li, Yongxin Yang, Yi-Zhe Song, and Timothy Hospedales. Learning to generalize: Meta-learning for domain generalization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018. [1, 2, 3, 4, 5, 6, 7](#)
- [24] Da Li, Yongxin Yang, Yi-Zhe Song, and Timothy M Hospedales. Deeper, broader and artier domain generalization. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 5542–5550, 2017. [5](#)
- [25] Da Li, Jianshu Zhang, Yongxin Yang, Cong Liu, Yi-Zhe Song, and Timothy M Hospedales. Episodic training for domain generalization. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 1446–1455, 2019. [1](#)
- [26] Haidong Li, Jiongcheng Li, Xiaoming Guan, Binghao Liang, Yuting Lai, and Xinglong Luo. Research on overfitting of deep learning. In *2019 15th International Conference on Computational Intelligence and Security*, pages 78–81, 2019. [1](#)
- [27] Haoliang Li, Sinno Jialin Pan, Shiqi Wang, and Alex C Kot. Domain generalization with adversarial feature learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5400–5409, 2018. [1, 2, 7](#)
- [28] Lei Li, Ke Gao, Juan Cao, Ziyao Huang, Yepeng Weng, Xiaoyue Mi, Zhengze Yu, Xiaoya Li, and Boyang Xia. Pro-

- gressive domain expansion network for single domain generalization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 224–233, 2021. 2
- [29] Pan Li, Da Li, Wei Li, Shaogang Gong, Yanwei Fu, and Timothy M Hospedales. A simple feature augmentation for domain generalization. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 8886–8895, 2021. 1, 2
- [30] Hong Liu, Zhangjie Cao, Mingsheng Long, Jianmin Wang, and Qiang Yang. Separate to adapt: Open set domain adaptation via progressive separation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2927–2936, 2019. 2
- [31] Fangrui Lv, Jian Liang, Shuang Li, Bin Zang, Chi Harold Liu, Ziteng Wang, and Di Liu. Causality inspired representation learning for domain generalization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8046–8056, 2022. 5
- [32] Lucas Mansilla, Rodrigo Echeveste, Diego H Milone, and Enzo Ferrante. Domain generalization via gradient surgery. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 6630–6638, 2021. 2
- [33] Toshihiko Matsuura and Tatsuya Harada. Domain generalization using a mixture of multiple latent domains. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 11749–11756, 2020. 5
- [34] Hyeonseob Nam, HyunJae Lee, Jongchan Park, Wonjun Yoon, and Donggeun Yoo. Reducing domain gap by reducing style bias. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8690–8699, 2021. 7
- [35] Vapnik Vladimir Naumovich and Vapnik Vlaimir. Statistical learning theory, 1998. 5, 6, 7
- [36] Lawrence Neal, Matthew Olson, Xiaoli Fern, Weng-Keen Wong, and Fuxin Li. Open set learning with counterfactual images. In *Proceedings of the European Conference on Computer Vision*, pages 613–628, 2018. 2, 6
- [37] Poojan Oza and Vishal M Patel. C2ae: Class conditioned auto-encoder for open-set recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2307–2316, 2019. 2
- [38] Alexandre Rame, Corentin Dancette, and Matthieu Cord. Fishr: Invariant gradient variances for out-of-distribution generalization. In *International Conference on Machine Learning*, pages 18347–18377. PMLR, 2022. 2
- [39] Kuniaki Saito and Kate Saenko. Ovanet: One-vs-all network for universal domain adaptation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9000–9009, 2021. 2, 4
- [40] Walter J Scheirer, Anderson de Rezende Rocha, Archana Sapkota, and Terrance E Boult. Toward open set recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(7):1757–1772, 2012. 1
- [41] Seonguk Seo, Yumin Suh, Dongwan Kim, Geeho Kim, Jongwoo Han, and Bohyung Han. Learning to optimize domain specific normalization for domain generalization. In *Euro-*
pean Conference on Computer Vision, pages 68–83, 2020. 7
- [42] Yuge Shi, Jeffrey Seely, Philip Torr, N Siddharth, Awni Hannun, Nicolas Usunier, and Gabriel Synnaeve. Gradient matching for domain generalization. In *International Conference on Learning Representations*, 2021. 2
- [43] Yang Shu, Zhangjie Cao, Chenyu Wang, Jianmin Wang, and Mingsheng Long. Open domain generalization with domain-augmented meta-learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9624–9633, 2021. 1, 2, 3, 5, 6
- [44] Baochen Sun and Kate Saenko. Deep coral: Correlation alignment for deep domain adaptation. In *European Conference on Computer Vision*, pages 443–450, 2016. 7
- [45] Hemanth Venkateswara, Jose Eusebio, Shayok Chakraborty, and Sethuraman Panchanathan. Deep hashing network for unsupervised domain adaptation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5018–5027, 2017. 5
- [46] Jindong Wang, Cuiling Lan, Chang Liu, Yidong Ouyang, Tao Qin, Wang Lu, Yiqiang Chen, Wenjun Zeng, and Philip Yu. Generalizing to unseen domains: A survey on domain generalization. *IEEE Transactions on Knowledge and Data Engineering*, 2022. 1, 2
- [47] Mei Wang and Weihong Deng. Deep visual domain adaptation: A survey. *Neurocomputing*, 312:135–153, 2018. 2
- [48] Shiqi Yang, Yaxing Wang, Kai Wang, Shangling Jui, and Joost van de Weijer. One ring to bring them all: Towards open-set recognition under domain shift. *ArXiv Preprint ArXiv:2206.03600*, 2022. 3
- [49] Ryota Yoshihashi, Wen Shao, Rei Kawakami, Shaodi You, Makoto Iida, and Takeshi Naemura. Classification-reconstruction learning for open-set recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4016–4025, 2019. 2
- [50] Jian Zhang, Lei Qi, Yinghuan Shi, and Yang Gao. Mvdg: A unified multi-view framework for domain generalization. In *European Conference on Computer Vision*, pages 161–177, 2022. 2, 5, 6
- [51] Kaiyang Zhou, Yongxin Yang, Timothy Hospedales, and Tao Xiang. Deep domain-adversarial image generation for domain generalisation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 13025–13032, 2020. 5
- [52] Kaiyang Zhou, Yongxin Yang, Timothy Hospedales, and Tao Xiang. Learning to generate novel domains for domain generalization. In *European Conference on Computer Vision*, pages 561–578, 2020. 2
- [53] Kaiyang Zhou, Yongxin Yang, Yu Qiao, and Tao Xiang. Domain generalization with mixstyle. In *International Conference on Learning Representations*, 2020. 1, 2, 5, 6, 7
- [54] Kaiyang Zhou, Yongxin Yang, Yu Qiao, and Tao Xiang. Domain adaptive ensemble learning. *IEEE Transactions on Image Processing*, 30:8008–8018, 2021. 2
- [55] Ronghang Zhu and Sheng Li. Crossmatch: Cross-classifier consistency regularization for open-set single domain generalization. In *International Conference on Learning Representations*, 2021. 3, 5