Supplementary Material: When Domain Generalization meets Generalized Category Discovery: An Adaptive Task-Arithmetic Driven Approach

Vaibhav Rathore Shubhranil B Saikat Dutta Sarthak Mehrotra Zsolt Kira
Biplab Banerjee

1. Introduction

This supplementary material provides additional insights and extended analysis to support the main paper. Below, we outline the key sections :

Extended Literature on Class Discovery: (Section 2) A detailed review of prior work in Novel Category Discovery (NCD) and Generalized Category Discovery (GCD), emphasizing advancements relevant to domain shifts.

DG-GCD in Practice: Applications Across Domains: (Section 3) Real-world applications of DG-GCD, including driverless cars, healthcare, and retail.

Dataset Details: (Section 4) Comprehensive information on datasets, their configurations, and class distributions.

Technical Details about Synthetic Data Generation: (Section 5) Description of the synthetic domain generation pipeline and parameter configurations.

Synthetic Domain Utilization in DG-GCD: (Section 6) Integration of synthetic domains into training and validation, with visualizations and statistical insights.

Handling Open-Set Recognition: (Section 7) Methodologies for open-set recognition, including episodic classifiers and adversarial loss.

Pseudocode for Episodic Training Strategy: (Section 8) The pseudocode for our episodic training strategy and its implementation.

Meta-Knowledge Learned in Episodic Training: (Section 9) An exploration of how meta-knowledge is acquired and refined through episodic training, enabling robust domain generalization and task adaptability.

Comprehensive Comparative Analyses Across Datasets: (Section 10.2) Comparative analysis of DG^{2}CD-Net on PACS, Office-Home, and DomainNet, evaluating its robustness and adaptability against benchmarks.

Performance Comparison with Domain Adaptation Methods: (Section 11) A detailed comparison of DG^{2}CD-Net with baseline and upper-bound domain adaptation methods.

Additional Ablation Studies: (Section 12) An analysis of the contributions of individual components in DG^{2}CD-Net across multiple datasets.

Effect of Backbone Initialization: (Section 13) A comparison of results using different backbones.

Effect of LoRA Fine-Tuning: (Section 14) A comparison of results using DG^{2}CD-Net with different LoRA adapters.

Limitations and Future Work: (Section 15) A discussion on the current limitations and potential future directions for DG^{2}CD-Net .

Each section is carefully crafted to provide deeper insights, reproducibility details, and additional context for the results presented in the main paper. We hope this supplementary material enhances the reader's understanding and offers a foundation for future exploration of domain generalization and category discovery.

2. Extended Literature on Class discovery

Category discovery has evolved significantly from Novel Category Discovery (NCD) [6] to GCD [20]. Traditionally, NCD methods [6, 7, 27, 29, 30] have utilized dual-model architectures where separate models are trained on labeled and unlabeled data to facilitate task transfer or employed parametric classifiers on top of generic feature extractors to categorize new classes. Recent advancements in GCD focus on leveraging labeled data to generate pseudo-labels for unlabeled images. DCCL [16] uses InfoMap clustering, while PromptCAL [26] identifies pseudo-positive samples through semi-supervised affinity propagation techniques. PIM [2] improves this by optimizing bi-level mutual information, and SimGCD [23] utilizes knowledge distillation with a parametric classifier to enhance pseudo-label reliability. A key innovation is the contrastive mean-shift clustering from [3], which creates a highly discriminative embedding space for fine-grained class distinctions. Additionally, Gaussian mixture models [28] have been explored for clustering in GCD.

Methods	**$\mathcal{D}_\mathcal{T}$ Available for Training**	**Evaluation on Known/Novel classes**	$\mathcal{P}(\mathcal{S}) = \mathcal{P}(\mathcal{T})$
NCD [5, 9]	✓	Novel	✓
GCD [20]	✓	Both	✓
CD-GCD [18]	✓	Both	✗
DG-GCD	✗	Both	✗

Table 1. Comparison of methods based on target-domain data ($\mathcal{D}_\mathcal{T}$) availability during training, evaluation on known and/or novel classes, and distribution divergence between the source ($\mathcal{S}$) and target ($\mathcal{T}$) domains in class discovery contexts: novel (NCD), generalized (GCD), and cross-domain (CD-GCD). Our proposed DG-GCD setting is different and more challenging than the rest.

3. DG-GCD in Practice: Applications Across Domains

The challenge of *domain generalization for generalized category discovery (DG-GCD)* has numerous real-world applications across various industries. In *autonomous vehicles*, models must adapt to changing environmental conditions (e.g., weather, lighting) and detect novel road objects without access to original training data. In *healthcare*, medical image analysis must generalize across different diagnostic tools while discovering new pathologies, all while respecting patient privacy regulations. *Wildlife monitoring* requires models to generalize across ecosystems while identifying new species, while *surveillance systems* must detect novel threats and adapt to different environments such as airports and public spaces. In *retail and e-commerce*, recommendation systems need to adapt across product categories and discover new items, while *robotics* requires models to generalize in dynamic environments and detect new objects. Lastly, in *precision agriculture*, systems must generalize across different farms and detect novel crop diseases without requiring access to original datasets. These applications highlight the importance of models that can generalize across domains and discover new categories in privacy-preserving, real-world scenarios.

4. Dataset details

Datasets: Our experiments are conducted on three benchmark datasets (i) **PACS** [12] (ii) **Office-Home** [21], and (iii) **DomainNet** [15].

For PACS and Office-Home, each domain was used as the source, with all others as target domains. For DomainNet, a subset of source-target configurations was selected, as summarized in Table 2, ensuring the model was tested on diverse domain pairs.

Source	**Targets**
Sketch	Clipart, Painting, Infograph, Quickdraw, Real World
Painting	Clipart, Sketch, Infograph, Quickdraw, Real World
Clipart	Painting, Sketch, Infograph, Quickdraw, Real World

Table 2. Source-target configurations for DomainNet

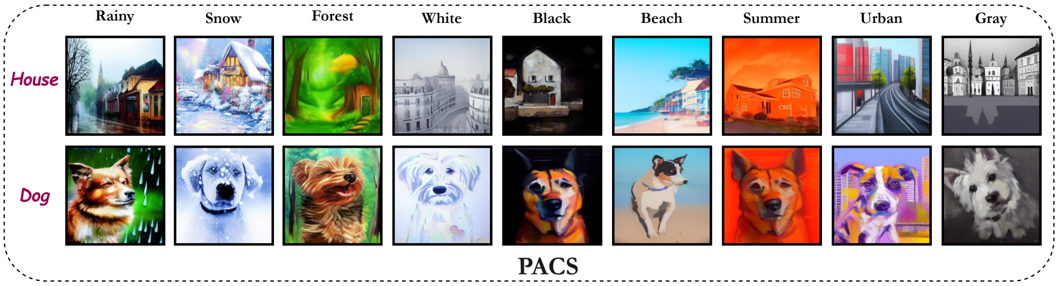

Figure 1. Additional synthetic image samples generated for the PACS dataset, showcasing diverse styles and domains across different categories (House and Dog).

5. Technical details about Synthetic Data Generation

We employed the InstructPix2Pix pipeline from Hugging Face's Diffusers library, utilizing the `timbrooks/instruct-pix2pix` model, for image transformation tasks. To achieve a balance between processing time and output quality, we configured the key parameter `num_inference_steps` to 10. The `image_guidance_scale` parameter was set to 1.0, ensuring that the model retained the essential structure of the input image while applying the specified transformations. Furthermore, the `guidance_scale` parameter was adjusted to 7.5, promoting a strong alignment with the transformation prompt. These configurations allow for straightforward replication of our process while maintaining high-quality output.

6. Synthetic Domain Utilization in DG-GCD

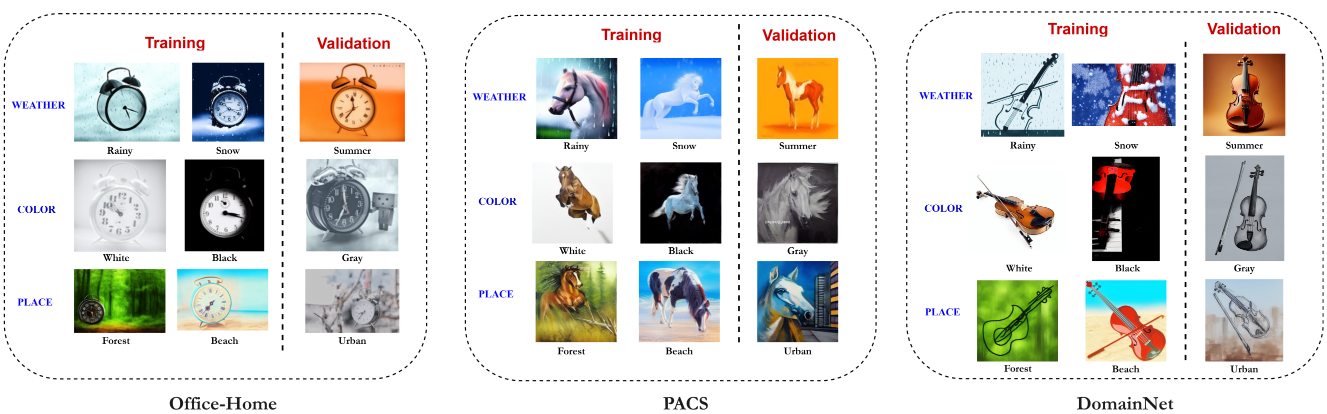

Figure 2. Categorization of synthetic domains utilized in the training and validation phases. Training domains are designed to simulate diverse conditions such as weather, color, and place variations. Validation domains challenge the model's adaptability to new, complex scenarios.

In our study, we generate nine synthetic domains for each of the PACS, Office-Home, and DomainNet datasets to enhance the adaptability of models trained under the Domain Generalization for Generalized Category Discovery (DG-GCD) framework. Table 3 outlines the distribution of these domains, with six utilized in training and three in validation, ensuring comprehensive exposure to varied environmental conditions and rigorous testing of generalization capabilities. This structured approach improves the models' robustness against unseen real-world scenarios.

Figure 2 categorizes the synthetic domains used in our study's training and validation phases. The training domains include variations such as "Rainy" and "Snow" weather, "White" and "Black" colors, and "Forest" and "Beach" settings, broadening the model's exposure to diverse scenarios. The validation domains introduce "Summer" weather, "Gray" color, and "Urban" settings to test the model's ability to generalize across new and complex environments. This strategic use of synthetic domains demonstrates our approach to enhancing robustness and adaptability in models, crucial for effective domain generalization in real-world applications.

Dataset	Total Domains Generated	Used in Training	Used in Validation
PACS	9	6	3
Office-Home	9	6	3
DomainNet	9	6	3

Table 3. Synthetic Domain Generation and Utilization for DG-GCD

Table 4 presents the Average Fréchet Inception Distance (FID) scores for synthetic domains compared to the original domains within the Office-Home dataset. The scores are calculated to evaluate the visual similarity between generated images in environments like "Rainy," "Black," "Urban," "Beach," "White," and "Snow," and the original dataset categories: Art, Clipart, Product, and Real World. Lower FID scores indicate closer visual resemblance to the original domain, suggesting better synthetic image quality and domain adaptation. The data reveal variations in FID scores, with "White" and "Black" environments achieving the lowest scores, indicating higher similarity and potentially more effective domain generalization. This analysis provides insights into which synthetic modifications most accurately reflect the characteristics of their respective real-world counterparts, crucial for training robust models capable of generalizing across diverse visual contexts.

Domain	FID Scores			
	Art	Clipart	Product	Real World
Rainy	140.11	158.81	167.27	154.42
Black	58.69	74.35	86.65	72.28
Urban	97.77	131.80	138.45	119.78
Beach	109.13	133.33	136.54	121.10
White	49.80	71.07	77.34	61.73
Snow	100.99	125.03	132.42	117.50

Table 4. Average Fréchet Inception Distance (FID) comparison between pairs of the generated and original domains over all the classes of the Office-Home dataset.

7. Handling Open-Set Recognition with Episode-Specific Classifiers

In our framework, we employ a $|\mathcal{Y}_s^{e_g}|+1$-class episode-specific classifier to address open-set recognition in episodic training. Here, $|\mathcal{Y}_s^{e_g}|$ denotes the number of known classes in an episode, while the additional class models the open-set category, capturing instances outside the known classes. This design dynamically adapts to episodic data, ensuring robust differentiation between known and unknown classes.

7.1. Adversarial Loss in Open-Set Domain Adaptation

To refine the decision boundary, we use an adversarial loss, $\mathcal{L}_{\text{adv}}$, which separates open-set instances by pushing them further from the known classes in feature space. This adversarial refinement enables the confident classification of unknown samples while maintaining performance on known categories. By tailoring decision boundaries to the episodic data, the classifier improves adaptive learning, generalization to new domains, and detection of novel classes in unseen distributions.

The adversarial loss used in our method is inspired by [19]. It facilitates the separation of known and unknown samples in the target domain by training the classifier $\mathcal{F}_c^{e_g}$ and the generator $\mathcal{F}^{g-1}$ adversarially. The adversarial loss $\mathcal{L}_{\text{adv}}$ is defined as:

$$\mathcal{L}_{\text{adv}}(x^t) = -\alpha \log(p(y = |\mathcal{Y}_s^{e_g}|+1|x^t)) - (1-\alpha)\log(1 - p(y = |\mathcal{Y}_s^{e_g}|+1|x^t)), \tag{1}$$

Where:

- x^t represents a target sample from $\mathcal{D}_{\text{syn}}$,
- $p(y = |\mathcal{Y}_s^{e_g}|+1|x^t)$ is the predicted probability that x^t belongs to the unknown class,
- α is a hyperparameter (set to 0.5 in our experiments) that determines the decision boundary for the unknown class.

7.1.1. Training Objectives

Adversarial loss ($\mathcal{L}_{adv}$) is optimized along with Source Classification loss, which is a standard cross-entropy loss that ensures accurate classification of known source samples:

$$\mathcal{L}_s(x^s, y^s) = -\log(p(y = y^s|x^s)), \tag{2}$$

where (x^s, y^s) represents a source sample and its label from $\mathcal{D}_{\mathcal{S}}$ respectively.

The classifier $\mathcal{F}_c^{e_g}$ and generator $\mathcal{F}^{g-1}$ is trained in the following manner:

- For the classifier $\mathcal{F}_c^{e_g}$, the objective is to minimize the total loss:

$$\min_{\mathcal{F}_c^{e_g}} \mathcal{L}_s(x^s, y^s) + \mathcal{L}_{\text{adv}}(x^t). \tag{3}$$

- For the generator $\mathcal{F}^{g-1}$, the objective is to deceive the classifier by maximizing the adversarial loss:

$$\min_{\mathcal{F}^{g-1}} \mathcal{L}_s(x^s, y^s) - \mathcal{L}_{\text{adv}}(x^t). \tag{4}$$

7.1.2. Implementation Details

To efficiently compute the adversarial loss, we use a Gradient Reversal Layer (GRL), which flips the gradient sign during backpropagation. This allows simultaneous updates to $\mathcal{F}_c^{e_g}$ and $\mathcal{F}^{g-1}$, facilitating stable training of adversarial objectives.

8. Pseudocode of the proposed Episodic Training Strategy

In this section, we present the pseudocode for the proposed episodic training strategy, as detailed in the main paper. This algorithm is used to iteratively update the global model parameters θ_{global} across multiple episodes and global updates. The process involves training task models on synthetic domains and updating the global model based on task vector computations and validation results.

9. Meta-Knowledge learnt in Episodic Training

In our episodic training framework, meta-knowledge encompasses the cumulative insights gained from dynamic adaptation to varying domain conditions. This knowledge is perpetually refined via systematic application and iterative adjustment of task vectors, informed by feedback from domain-specific fine-tuning and rigorous validation processes. Unlike conventional meta-learning, which primarily targets rapid task adaptability, our framework emphasizes robust domain generalization. This approach enhances the model's proficiency in effectively preemptively addressing and adapting to evolving data distributions.

Meta-knowledge is acquired through:

[-] **Cross-Domain Exposure:** By engaging with overlapping features across multiple domains, the model develops a nuanced capability to generalize across diverse training environments. This cross-domain learning is fundamental in enabling the model to abstract and apply domain-invariant patterns to new, unseen scenarios.

[-] **Dynamic Vector Adjustments:** Task vectors are continually updated in response to real-time performance metrics. This dynamic refinement process allows the model to adjust its generalization strategies on the fly, enhancing its responsiveness to changes in domain characteristics.

[-] **Validation-Driven Learning:** The integration of internal validation mechanisms ensures continuous performance feedback. This feedback is instrumental in fine-tuning the model's strategic adjustments, ensuring optimized responses to future domain shifts and data interactions.

Our method enhances the generalizability of pre-trained foundation models by adaptively combining their fine-tuned versions across multiple automatically synthesized domains, eliminating the need for manual annotations. This adaptive strategy leverages the generalization performance of each fine-tuned model, minimizing the impact of poorly generalizable models. Besides, our approach ensures both discriminativeness and domain independence for the DG-GCD task. Consequently, this produces an embedding space predominantly guided by class semantics and suppressing stylistic artifacts, making it highly effective for generalization and clustering.

10. More details on comparisons to literature

10.1. DG-GCD specific adaptations for baselines

We evaluate several state-of-the-art methods for generalized category discovery (GCD) and domain generalization (DG) by adapting them to the DG-GCD setting, as detailed in Table 2 of the main text. In this setting, target domain access

Algorithm 1 Proposed Episodic Training Strategy for Updating θ_{global}

Require: Pre-trained global model parameters θ^0_{global}, number of global updates n_g, number of episodes per global update n_e, source domain $\mathcal{D}^{e_g}_{\mathcal{S}}$, synthetic domain $\mathcal{D}^{e_g}_{\text{syn}}$, validation domain $\mathcal{D}_{\text{valid}}$ in the e^{th}_g episode.
Ensure: Final global model $\theta^{n_g}_{\text{global}}$.

for $g = 1$ to n_g **do** ▷ Global updates
    Randomly shuffle synthetic domains $\mathcal{D}^{e_g}_{\text{syn}}$.
    for each episode $e = 1$ to n_e **do** ▷ Episode training
        Initialize task model parameters $\theta^{e_g}_{\text{local}} \leftarrow \theta^{g-1}_{\text{global}}$.
        Train task model $\theta^{e_g}_{\text{local}}$ on $(\mathcal{D}^{e_g}_{\mathcal{S}}, \mathcal{D}^{e_g}_{\text{syn}})$ for the CD-GCD task.
        Compute task vector δ^e_g (Equ. 1 in the main text) :

$$\delta^e_g = \theta^{g-1}_{\text{global}} - \theta^{e_g}_{\text{local}}$$

        Validate task model on $\mathcal{D}_{\text{valid}}$ and compute accuracies: `All`, `Old`, `New`.
    end for
    Compute weights w^e_g using softmax on `All` accuracies (Equ. 2 in main text) :

$$w^e_g = \frac{\exp(\mathtt{All}^e_g)}{\sum_{e'=1}^{n_e} \exp(\mathtt{All}^{e'}_g)}$$

    Update global model θ^g_{global} (Equ. 3 in main text) :

$$\theta^g_{\text{global}} = \theta^{g-1}_{\text{global}} - \sum_{e=1}^{n_e} w^e_g \delta^e_g$$

end for
Save the final global model: $\theta^{n_g}_{\text{global}}$.

is completely removed during training, and for certain methods, synthetic domain data is incorporated to simulate domain shifts.

For ViT-B/16, pre-trained with DINO, we fine-tuned only the last block using source domain data, following standard GCD practices, and evaluated the model on target domains without any target domain data during training. Similarly, GCD was adapted by fine-tuning the last block of the backbone on the source domain alone, and we introduced a synthetic domain variant to account for domain shifts.

For CMS (Contrastive Mean Shift) and SimGCD, we followed a similar procedure. We fine-tuned the last block using only the source domain and, in addition, created synthetic variants by incorporating synthetic domain data to assess the methods' ability to handle domain shifts effectively.

CDAD-Net, which is designed for cross-domain adaptation, was also adapted for the DG-GCD setting. We ensured that it trained solely on the source domain, without access to target domain data, and created a synthetic variant to evaluate its performance on unseen domains.

As seen in Table 2 of the main text, the methods incorporating synthetic domain data, such as GCD with synthetic data, generally performed better in handling domain shifts, especially for novel class discovery. Our proposed models, leveraging task merging techniques such as TIES-Arithmetic and LoRA, outperform baseline methods on both benchmark datasets, achieving superior results, particularly on novel classes. The inclusion of synthetic domains proves beneficial, as evidenced by the marked performance improvement across all datasets, with our model consistently achieving the highest or second-highest results.

10.2. Comprehensive comparative analyses of DG^{2}CD-Net across multiple datasets

This section presents a concise comparative analysis of DG^{2}CD-Net on PACS, Office-Home, and DomainNet in Table-5, 6 and 7 respectively. Each dataset challenges DG^{2}CD-Net with unique domain shifts, showcasing its adaptability and robustness. This evaluation aims to validate DG^{2}CD-Net 's performance against established benchmarks, highlighting it's strengths and identifying opportunities for advancement in domain generalization.

PACS																		
Methods	**Art Painting → Sketch**			**Art Painting → Cartoon**			**Art Painting → Photo**			**Photo → Art Painting**			**Photo → Cartoon**			**Photo → Sketch**		
	All	Old	New	All	Old	New	All	Old	New	All	Old	New	All	Old	New	All	Old	New
ViT [4]	37.44	50.73	19.5	47.4	61.3	35.25	76.05	87.13	64.64	53.17	77.31	31.67	47.01	55.54	39.57	31.87	37.57	24.16
GCD [20]	32.02	41.53	19.12	46.78	60.35	28.57	79.16	99.45	48.73	74.73	80.26	67.31	57.53	60.46	53.6	46.23	48.56	43.08
SimGCD [23]	29.35	17.3	62.12	23.08	28.26	16.32	51.98	74.44	33.26	46.29	48.96	43.17	34.26	44.91	20.35	24.84	31.88	5.68
CDAD-Net [18]	46.02	45.95	46.21	51.71	53.43	49.46	99.04	99.21	98.9	76.61	76.97	76.19	56.78	56.67	56.93	46.65	46.15	48.01
GCD With Synthetic	45.78	36.71	58.01	54.84	73.47	38.57	82.6	66.29	99.39	79	86.84	72.02	53.56	67.93	41.01	44.18	47.78	39.32
CDAD-Net with Synthetic	43.09	42.53	44.6	49.45	59.31	36.58	99.16	99.21	99.12	65.38	62.83	68.36	42.92	41.97	44.15	41.51	43.79	35.32
DG^{2}CD-Net (TIES-Merging[24])	41.31	40.56	42.31	45.69	57	35.81	96.11	97.87	94.29	62.87	87.72	40.72	48.98	60.02	39.33	44.1	36.33	54.58
DG^{2}CD-Net [TA[11]]	46.4	51.3	42.8	56.21	58.82	52.7	99.2	99.5	98.8	81.47	91.09	68.57	57.76	56.11	59.99	46.88	48.96	44.06
DG^{2}CD-Net (Ours)	46.79	38.13	58.49	57.96	73.38	44.48	99.34	99.7	98.97	86.67	91.87	82.04	62.97	71.18	55.8	45.72	36.53	58.13
DG^{2}CD-Net * (Ours)[LoRA[10]]	46.83	37.79	59.03	63.82	71.13	57.43	99.46	99.35	99.57	88.89	93.94	84.4	64.19	72.23	57.15	46.45	37.75	58.19

Methods	**Sketch → Art Painting**			**Sketch → Cartoon**			**Sketch → Photo**			**Cartoon → Art Painting**			**Cartoon → Sketch**			**Cartoon → Photo**		
	All	Old	New	All	Old	New	All	Old	New	All	Old	New	All	Old	New	All	Old	New
ViT [4]	23.93	26.53	21.61	40.61	58.92	24.62	33.29	33.88	32.69	38.09	47.36	29.82	33.57	35.67	30.74	41.38	39.08	43.74
GCD [20]	33.25	39.09	25.43	40.89	48.14	31.17	46.86	59.28	28.22	58.15	78.52	30.86	36	44.83	24.04	75.75	85.88	60.55
SimGCD [23]	21.19	31.91	8.67	23.17	36.77	5.4	34.22	27.46	40.8	38.38	42.07	34.07	34.84	33.94	37.31	53.05	45.85	59.06
CDAD-Net [18]	87.99	84.32	92.28	51.88	51.77	52.02	99.04	99.21	98.9	73.05	76.88	68.57	41.84	42.71	39.49	99.22	99.47	99.01
GCD With Synthetic	82.15	85.13	79.5	44.3	48.22	40.89	99.49	99.76	99.21	63.01	63.73	62.37	35.66	29.95	43.36	99.43	99.47	99.39
CDAD-Net with Synthetic	61.91	69.45	53.12	48.59	53.13	42.67	68.44	63.5	72.56	67.24	65.28	69.52	42.05	39.61	48.67	99.34	99.47	99.23
DG^{2}CD-Net (TIES-Merging[24])	80.59	80.78	80.42	58.94	75.71	44.28	99.07	98.64	99.51	87.45	90.93	84.35	40.67	31.39	53.2	98.71	97.99	99.45
DG^{2}CD-Net (TA[11])	73.02	79.37	64.51	55.89	54.84	57.29	99.31	99.5	99.03	90.89	92.75	88.4	46.03	49.67	41.1	99.16	99.35	98.88
DG^{2}CD-Net (Ours)	88.75	93.52	84.49	56.76	72.14	43.33	99.13	98.7	99.57	90.77	93.37	88.46	49.2	43.18	57.33	95.57	91.62	99.64
DG^{2}CD-Net * (Ours)[LoRA[10]]	90.87	95.28	86.93	66.25	78.32	55.72	99.22	98.88	99.57	91.02	93.99	88.37	46.33	38.19	57.33	99.22	98.82	99.64

Table 5. Detailed comparison of our proposed DG^{2}CD-Net on DG-GCD with respect to referred literature for PACS Dataset.

Office-Home																		
Methods	**Art → Clipart**			**Art → Product**			**Art → Real World**			**Clipart → Art**			**Clipart → Real World**			**Clipart → Product**		
	All	Old	New	All	Old	New	All	Old	New	All	Old	New	All	Old	New	All	Old	New
ViT [4]	18.88	20.86	15.79	30.34	35.42	21.83	29.52	32.76	24.85	14.96	15.6	14.12	18.59	20.12	16.4	30.39	32.51	26.84
GCD [20]	31.65	32.11	30.93	63.18	64.35	61.22	63.85	66.56	59.96	51.96	52.7	51	62.62	65.29	58.79	60.59	67.13	49.61
SimGCD [23]	24.54	34.35	8.09	41.95	57.92	13.54	46.78	65.54	14.73	31.11	39.56	11.88	25.66	37.66	5.15	28.88	41.38	12.96
CDAD-Net [18]	30.95	33.65	26.43	64.99	68.04	59.32	67.5	70.89	61.72	53.36	56.05	47.23	64.7	69.4	55.25	67.02	68.8	63.7
GCD With Synthetic	29.86	31.04	28.02	57.92	63.12	49.19	59.47	59.59	59.29	53.3	52.84	53.89	61.46	58.27	66.06	63.84	64.04	63.51
CDAD-Net with Synthetic	31.97	35.1	26.71	65.39	68.94	62.51	67.83	70.87	62.64	53.51	56.65	46.37	66.97	69.76	62.2	61.4	65.55	57.4
DG^{2}CD-Net (TIES-Merging[24])	33.96	37	29.23	59.99	62.93	55.07	66.26	68.42	63.15	52.18	52.3	52.04	58.16	58.62	57.5	65.32	72.33	53.56
DG^{2}CD-Net [TA[11]]	29.52	27.31	33.06	62.42	61.67	63.59	64.46	62.14	67.8	51.24	53.32	47.12	64.23	61.24	68.92	65.28	66.03	64.13
DG^{2}CD-Net (Ours)	31.51	31.96	30.81	67.46	68.73	65.32	64.45	60.25	70.48	50.76	48.76	53.36	64.77	60.58	70.79	65.34	67.48	61.76
DG^{2}CD-Net (Ours)[LoRA [10]]	31.56	31.85	31.1	65.22	65.68	64.45	67.81	65.14	71.66	53.4	48.47	59.81	66.13	61.63	72.61	66.16	67.12	64.57

Methods	**Product → Art**			**Product → Real World**			**Product → Clipart**			**Real World → Art**			**Real World → Product**			**Real World → Clipart**		
	All	Old	New	All	Old	New	All	Old	New	All	Old	New	All	Old	New	All	Old	New
ViT [4]	23.2	24.64	21.33	31.21	35.45	25.13	19.27	20.52	17.31	32.22	35.79	27.58	44.67	52.21	32.03	20.8	23.71	16.26
GCD [20]	50.27	48.18	52.99	65.07	63.09	67.91	29.08	29.22	28.87	54.26	54.05	54.55	69.04	72.76	62.79	31.04	34.93	24.97
SimGCD [23]	38.28	50.42	10.66	48.36	67.07	16.41	22.45	32.37	11.34	48.95	66.79	8.36	57.19	69.23	44.15	21.7	31.46	5.33
CDAD-Net [18]	50.1	52.43	44.67	66.47	72.13	56.81	31.36	34.6	25.94	54.68	58.07	46.96	61.39	64.79	55.06	31.78	36.02	24.69
GCD With Synthetic	49.18	46.54	52.61	63.4	59.67	68.77	28.43	27.72	29.55	51.71	61.55	38.91	61.14	65.34	54.1	26.38	28.11	23.68
CDAD-Net with Synthetic	54.12	57.67	46.04	66.97	70.2	61.46	32.34	35.13	28.68	53.72	56.89	46.5	56.47	62.33	45.62	31.19	33.67	27.02
DG^{2}CD-Net (TIES-Merging[24])	53.28	54.77	51.33	62.74	66.85	56.83	31.82	33.97	28.46	57.11	66.14	45.36	67.04	74.25	54.95	34.41	37.94	28.9
DG^{2}CD-Net [TA[11]]	49.92	52.33	45.17	65.57	67.22	62.99	31.48	30.21	33.51	51.65	55.06	44.92	65.01	63.73	66.99	30.73	28.65	34.08
DG^{2}CD-Net (Ours)	52.45	50.51	54.98	67.87	69.88	64.97	30.71	30.05	31.75	52.31	49.42	56.07	67.37	71.65	60.19	31.28	31.13	31.51
DG^{2}CD-Net * (Ours)[LoRA[10]]	52.66	51.75	53.84	65.48	62	70.48	31.52	31.83	31.04	53.42	51.6	55.78	66.33	68.97	61.91	32.26	30.4	35.15

Table 6. Detailed comparison of our proposed DG^{2}CD-Net on DG-GCD with respect to referred literature for Office-Home Dataset

DomainNet															
Methods	**Sketch → Real**			**Sketch → Quickdraw**			**Sketch → Infograph**			**Sketch → Painting**			**Sketch → Clipart**		
	All	Old	New	All	Old	New	All	Old	New	All	Old	New	All	Old	New
ViT [4]	47.17	47.92	44.95	12.13	12.1	12.21	11.99	12.68	10.28	30.95	33.02	25.75	32.64	34.29	28.64
GCD [20]	51.13	51.88	48.92	16.08	15.65	17.2	12.6	12.57	12.68	35.25	35.96	33.46	31.22	30.85	32.1
SimGCD [23]	3.11	3.47	2.32	2.31	2.4	2.1	3.16	2.27	5.24	4.1	2.57	5.62	3.02	2.3	4.07
CDAD-Net [18]	48.21	47.7	49.77	12.27	11.52	14.24	12.07	12.69	11.34	35.47	36.39	32.86	18.63	17.52	20.39
GCD With Synthetic	53	51.71	47.64	13.71	13.79	13.99	12.24	11.99	11.37	35.43	34.12	30.83	22.49	22.2	21.49
CDAD-Net with Synthetic	47.11	46.09	49.4	12.75	13.1	14.05	12.52	13.04	11.92	35.87	36.73	33.35	18.99	17.68	21.07
DG^{2}CD-Net (TIES-Merging[24])	50.32	52.88	42.8	15.22	15.12	15.49	14.75	16.04	11.53	35.84	38.99	27.93	31.06	33.34	25.53
DG^{2}CD-Net [TA[11]]	51.84	52.58	49.65	13.67	13.44	14.25	12.72	13.05	11.89	33.96	35.32	30.55	21.94	21.8	22.29
DG^{2}CD-Net (Ours)	53.67	55.48	48.35	15.9	16	15.63	14.63	15.66	12.06	37.44	39.53	32.19	30.47	32.89	24.58
DG^{2}CD-Net * (Ours)[LoRA[10]]	53.01	53.75	50.84	13.71	13.38	14.57	13.82	14.23	12.78	36.77	37.9	33.93	24.17	24.46	23.44

Methods	**Clipart → Infograph**			**Clipart → Quickdraw**			**Clipart → Sketch**			**Clipart → Real**			**Clipart → Painting**		
	All	Old	New	All	Old	New	All	Old	New	All	Old	New	All	Old	New
ViT [4]	12.18	12.64	11.03	12.13	12.1	12.21	24.76	26.24	21.27	44.14	45.43	40.34	26.76	28.7	21.91
GCD [20]	14.03	14.64	12.49	14.94	14.67	15.65	25.33	27.68	19.78	53.23	55.48	46.62	34.82	36.82	29.83
SimGCD [23]	2.03	0.4	3.94	0.5	0.3	1	1	0.02	3.842	1.64	1.07	2.42	2.07	2.05	2.13
CDAD-Net [18]	12.79	12.96	12.87	12.06	11.59	12.78	19	19.17	18.76	47.06	44.62	49.2	34.45	36.02	32.85
GCD With Synthetic	11.46	12.03	10.04	12.68	12.57	12.95	18.74	20.54	14.47	50.11	52.26	43.79	32.67	34.91	27.06
CDAD-Net with Synthetic	13	13.37	12.56	12.07	11.76	12.89	17.46	18.03	16.67	48.25	47.51	49.6	33.23	32.79	34.2
DG^{2}CD-Net (TIES-Merging[24])	15.66	17.02	12.28	14.91	14.73	15.39	27.75	30.64	20.89	54.18	56.73	46.72	36.71	38.14	33.16
DG^{2}CD-Net [TA[11]]	15.71	16.88	12.78	14.63	14.18	15.81	27.03	29.89	20.26	53.91	55.14	50.29	36.85	39.69	29.75
DG^{2}CD-Net (Ours)	15.81	17.09	12.63	14.53	14.14	15.58	26.86	29.49	20.64	54.54	56.03	50.17	36.81	38.87	31.67
DG^{2}CD-Net * (Ours)[LoRA[10]]	14.19	14.12	14.35	13.31	13.23	13.53	22.01	22.9	19.91	53.95	54.99	50.91	37.12	37.89	35.21

Methods	**Painting → Infograph**			**Painting → Quickdraw**			**Painting → Sketch**			**Painting → Real**			**Painting → Clipart**		
	All	Old	New	All	Old	New	All	Old	New	All	Old	New	All	Old	New
ViT [4]	12.2	13.1	9.94	12.13	12.1	12.21	23	24.78	18.79	51.53	54.16	43.8	26.57	28.08	22.92
GCD [20]	12.87	12.67	13.37	10.74	10.56	11.21	21.49	22.26	19.68	52.12	51.86	52.86	25.32	24.79	26.6
SimGCD [23]	3.2	2.6	3.8	3.5	2.32	4.65	4.23	3.56	4.86	4.2	3.52	5	4.49	3.6	5.23
CDAD-Net [18]	11.65	12.49	10.66	11.98	11.2	12.44	17.11	17.68	16.32	49.04	48.63	50.27	20.06	19.74	20.57
GCD With Synthetic	10.86	10.56	9.84	11.81	11.8	11.77	17.26	16.25	13.83	49.1	47.3	42.04	19.3	19.45	18.04
CDAD-Net with Synthetic	11.53	12.32	10.59	11.86	10.71	12.32	17.29	18.45	15.7	48.4	50.23	49.7	17.44	15.92	19.86
DG^{2}CD-Net (TIES-Merging[24])	15.34	16.64	12.13	12.89	12.64	13.58	23.45	25.6	18.38	55.16	57.3	47.46	27.5	29.48	22.65
DG^{2}CD-Net [TA[11]]	15.17	16.52	11.8	12.78	12.58	13.29	23.21	25.69	17.34	55.16	57.31	48.87	26.76	27.91	23.96
DG^{2}CD-Net (Ours)	15.71	16.72	13.22	12.9	12.66	13.53	23.14	25.23	18.19	55.07	56.97	49.5	27.6	29.07	24.03
DG^{2}CD-Net * [Ours)[LoRA[10]]	14.41	14.68	13.74	12.9	12.9	12.91	21.39	22.39	19.03	53.83	54.99	50.44	22.94	22.41	24.24

Table 7. Detailed comparison of our proposed DG^{2}CD-Net on DG-GCD with respect to referred literature for DomainNet Dataset

11. Performance comparison with Domain Adaptation (DA) methods

Table 8 presents a performance comparison of our proposed method, DG^{2}CD-Net , against two prominent methods, CROW (DA) and CDAD-Net (DA), across three benchmark datasets: PACS, Office-Home, and DomainNet. Notably, CDAD-Net (DA) represents a strong upper bound as it operates in the Domain Adaptation (DA) setting, leveraging access to target-domain data during training. In contrast, DG^{2}CD-Net is designed for the more challenging Domain Generalization (DG) setting, where no target-domain information is available. While CROW (DA) is included in this supplementary comparison for reference, CDAD-Net (DA) serves as a more appropriate upper bound, given its superior performance.

Our method significantly outperforms CROW (DA) across all datasets in terms of overall accuracy, achieving a margin of +9.34% on PACS (73.30% vs. 63.96%), +3.39% on Office-Home (53.86% vs. 50.47%), and +2.81% on DomainNet (29.01% vs. 26.20%). Additionally, DG^{2}CD-Net demonstrates robust generalization across both old and new classes, underscoring its adaptability in diverse scenarios. While CDAD-Net (DA) achieves higher performance due to its reliance on target domain data, the comparison highlights the inherent trade-off between the DA and DG settings. By including CROW (DA) results here, we provide a holistic view of baseline performance while emphasizing the relevance of CDAD-Net (DA) as the key upper bound in this context. This reinforces the practical value of DG^{2}CD-Net in solving the domain generalization challenge without relying on target domain assumptions.

12. Additional Ablation Studies

Component Impact on Office-Home: Table 9 reveals the significant effects of essential components on the DG^{2}CD-Net's overall performance. Removing synthetic domains leads to a decrease of approximately 3.28% in the "All" metric, un-

Methods	Venue	Target-Domain	PACS			Office-Home			DomainNet		
			All	Old	New	All	Old	New	All	Old	New
CROW (DA) [22]	ICML'24	✓	63.96	61.78	68.65	50.47	54.50	39.71	26.20	26.60	25.80
DG^{2}CD-Net (Ours)	–	×	73.30	75.28	72.56	53.86	53.37	54.33	29.01	30.38	25.46
CDAD-Net (DA) [18] [Upper bound]	CVPR-W'24	✓	83.25	87.58	77.35	67.55	72.42	63.44	70.28	76.46	65.19

Table 8. **Performance comparison** of CROW method with our method, as well as the upper bound CDAD-Net (DA), on all datasets.

derscoring its importance for generalization. The absence of episodic training results in a decrease of 2.33%, highlighting its role in model adaptability. The most considerable impact is observed with a fixed old/novel class split, which shows a reduction of 4.67% compared to the full model's configuration. In contrast, the full implementation of DG^{2}CD-Net achieves a comprehensive performance of 53.79% across all classes, demonstrating the effectiveness of dynamic weighting and the combined utility of all components in enhancing domain generalization and class discovery in the Office-Home dataset.

Model Variant	Office-Home		
	All	Old	New
✓Without Synthetic Domain	50.51	50.58	50.31
✓ Without multi-global updates	51.46	50.77	52.26
✓ Static known/novel class split across episodes	49.11	56.23	38.85
Full DG^{2}CD-Net (Proposed)	**53.79**	**53.83**	**53.66**

Table 9. Ablation study results on the impact of various components of DG^{2}CD-Net for the Office-Home dataset.

Observation on Old-New class splits: Table 10 illustrates the impact of varying base (old) and novel (new) class splits on the performance of DG^{2}CD-Net on the PACS dataset. We tested five different splits, ranging from 2 old classes with five novel classes to 6 old classes with one novel class. The results show that as the proportion of novel classes increases, the model's performance on novel classes improves, but there is a slight decline in accuracy for base classes. This behavior highlights the challenge of maintaining a balance between recognizing old and novel categories. The best overall performance is observed with the 5-2 split, indicating that DG^{2}CD-Net is more effective when the distribution of old and novel classes is moderately balanced.

Splits (Old-New)	PACS		
	All	Old	New
2 - 5	74.94	73.29	76.77
3 - 4	75.11	76.28	74.69
4 - 3	73.3	75.28	72.56
5 - 2	75.85	74.01	80.72
6 - 1	74.89	74.90	74.96

Table 10. Sensitivity on different `Old-New` class splits.

Effect of $\mathcal{L}_{\text{margin}}$: As shown in Figure 3, adding $\mathcal{L}_{\text{margin}}$ improves accuracy across all categories. `Old` classes benefit the most with a 2.36% increase, indicating enhanced feature separation for well-known classes. `New` classes see a 2.23% improvement, suggesting better discrimination for novel classes. Overall, accuracy increases by 1.85%, demonstrating that $\mathcal{L}_{\text{margin}}$ enhances class separability and generalization across both familiar and unseen categories.

Effect of m in $\mathcal{L}_{\text{margin}}$: Table 11 illustrates the average accuracy of our model across different margin values (m) in $\mathcal{L}_{margin}$. We experimented with different values of m ranging from 0.3 to 0.9. The highest scores were observed for $m = 0.7$, indicating that this margin setting provides better separation between known and novel classes.

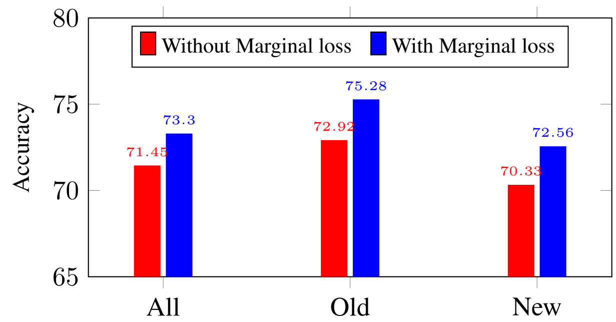

Figure 3. Accuracy comparison of Old, All, and New categories with and without $\mathcal{L}_{\text{margin}}$ loss.

m	PACS		
	All	**Old**	**New**
0.3	71.50	74.48	69.05
0.4	71.27	75.58	67.92
0.5	69.82	74.00	66.23
0.6	63.13	65.42	61.99
0.7	73.30	75.28	72.56
0.8	72.24	75.77	69.92
0.9	71.91	75.69	68.98

Table 11. Average accuracy for different sensitivity of the hyper-parameter m.

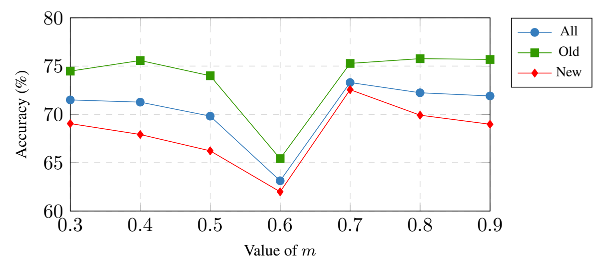

Figure 4. The relationship between margin m and accuracy (All, Old, New). The margin m in the loss function influences class separation, with peak accuracy observed at $m = 0.7$.

13. Effect of initialization of backbone

Backbone selection significantly impacts the performance of DG^{2}CD-Net in domain generalization. Initially, we employed ViT-B/16 with DINO initialization for comparability with prior work, but inspired by DINO-v2's advancements, we expanded our study to include ViT-B/16 with DINO-v2 initialization. Additionally, we have experimented with ResNet-50 with CLIP and ImageNet initializations and ViT-B/16 with CLIP initialization. Table 12 summarizes the results on PACS.

Model	Backbone	PACS		
		All	**Old**	**New**
ResNet-50	**CLIP** [17]	25.39	20.97	29.33
ResNet-50	**ImageNet** [8]	54.98	64.18	45.33
ViT-B/16	**DINO** [1]	73.30	75.28	72.56
ViT-B/16	**DINO-v2** [14]	87.71	90.67	84.91
ViT-B/16	**CLIP** [17]	90.07	92.25	87.72

Table 12. Performance Comparison of DG^{2}CD-Net with different backbones on the PACS Dataset

The results highlight the superiority of the DINO-v2 and CLIP-based models, with CLIP achieving the highest performance. The strong results of ViT-B/16 (CLIP) suggest that pre-training with vision-language data improves generalization. Given these findings, we recommend DINO-v2 and CLIP-based ViT as strong baselines for future domain generalization studies.

14. Effect of LoRA Fine-Tuning

In our experiments, we incorporated LoRA (Low-Rank Adaptation) into DG^2 CD-Net to improve the efficiency of fine-tuning while minimizing memory overhead. Unlike full fine-tuning, LoRA updates a small subset of parameters while keeping pre-trained weights frozen, thereby reducing catastrophic forgetting and enhancing adaptation and generalization.

Method	Trainable Parameters (K)	Total Parameters (K)	Percentage (%)
DG^{2}CD-Net (Vanilla)	7,088	85,799	8.261
DG^{2}CD-Net* (LoRA)	98	85,799	0.115

Table 13. Comparison of model parameters *with* and *without* LoRA fine-tuning.

Table 14 presents a performance comparison of DG^{2}CD-Net using different LoRA-based adapters, including *LoRA* [10], *DoRA* [13], and *AdaLoRA* [25]. These adapters aim to improve the model efficiency while maintaining high accuracy, with LoRA achieving the best performance.

Adapters	All (%)
LoRA [10]	75.21
DoRA [13]	74.11
AdaLoRA [25]	74.20

Table 14. Performance comparison of DG^{2}CD-Net with different LoRA Adapters

These findings emphasize the efficiency of LoRA in reducing the computational burden of fine-tuning while maintaining the model's overall capacity. Given the benefits, we recommend adopting LoRA-based fine-tuning in future research for domain generalization tasks to optimize memory usage and training speed without sacrificing performance.

15. Limitations and Future Work

While DG^{2}CD-Net demonstrates strong performance in domain generalization and category discovery, there are areas for future enhancement. One aspect that can be improved is the reliance on synthetic domain generation, which, while effective, can be optimized to reduce computational costs. Exploring more streamlined approaches to synthetic domain creation or alternative techniques that do not require synthetic data could further improve scalability and efficiency. Additionally, the episodic training framework, though beneficial for adaptation, demands considerable computational resources, especially when applied to large-scale datasets like DomainNet. Optimizing this process could make the method more feasible for real-world, large-scale applications.

In future work, efforts can focus on enhancing the efficiency of both synthetic domain generation and the episodic training process. Advanced techniques for model merging can also be explored to further improve performance. Addressing challenges like data imbalance, which is common in real-world scenarios, will strengthen the model's robustness and adaptability. Overall, extending DG^{2}CD-Net in these directions holds great promise for developing even more scalable and effective solutions for complex tasks in domain generalization.

References

[1] Mathilde Caron, Hugo Touvron, Ishan Misra, Hervé Jégou, Julien Mairal, Piotr Bojanowski, and Armand Joulin. Emerging properties in self-supervised vision transformers. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 9650–9660, 2021. 10

[2] Florent Chiaroni, Jose Dolz, Ziko Imtiaz Masud, Amar Mitiche, and Ismail Ben Ayed. Parametric information maximization for generalized category discovery. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 1729–1739, 2023. 2

[3] Sua Choi, Dahyun Kang, and Minsu Cho. Contrastive mean-shift learning for generalized category discovery. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 23094–23104, 2024. 2

[4] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale. In *International Conference on Learning Representations*, 2021. 7, 8

[5] Kai Han, Andrea Vedaldi, and Andrew Zisserman. Learning to discover novel visual categories via deep transfer clustering. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 8401–8409, 2019. 2

[6] Kai Han, Sylvestre-Alvise Rebuffi, Sebastien Ehrhardt, Andrea Vedaldi, and Andrew Zisserman. Automatically discovering and learning new visual categories with ranking statistics. In *International Conference on Learning Representations*, 2020. 2

[7] Kai Han, Sylvestre-Alvise Rebuffi, Sebastien Ehrhardt, Andrea Vedaldi, and Andrew Zisserman. Autonovel: Automatically discovering and learning novel visual categories. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(10):6767–6781, 2021. 2

[8] 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. 10

[9] Yen-Chang Hsu, Zhaoyang Lv, and Zsolt Kira. Learning to cluster in order to transfer across domains and tasks. In *International Conference on Learning Representations*, 2018. 2

[10] Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*, 2021. 7, 8, 11

[11] Gabriel Ilharco, Marco Tulio Ribeiro, Mitchell Wortsman, Ludwig Schmidt, Hannaneh Hajishirzi, and Ali Farhadi. Editing models with task arithmetic. In *The Eleventh International Conference on Learning Representations*, 2023. 7, 8

[12] 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. 2

[13] Shih-Yang Liu, Chien-Yi Wang, Hongxu Yin, Pavlo Molchanov, Yu-Chiang Frank Wang, Kwang-Ting Cheng, and Min-Hung Chen. DoRA: Weight-decomposed low-rank adaptation. In *Proceedings of the 41st International Conference on Machine Learning*, pages 32100–32121. PMLR, 2024. 11

[14] Maxime Oquab, Timothée Darcet, Théo Moutakanni, Huy V. Vo, Marc Szafraniec, Vasil Khalidov, Pierre Fernandez, Daniel HAZIZA, Francisco Massa, Alaaeldin El-Nouby, Mido Assran, Nicolas Ballas, Wojciech Galuba, Russell Howes, Po-Yao Huang, Shang-Wen Li, Ishan Misra, Michael Rabbat, Vasu Sharma, Gabriel Synnaeve, Hu Xu, Herve Jegou, Julien Mairal, Patrick Labatut, Armand Joulin, and Piotr Bojanowski. DINOv2: Learning robust visual features without supervision. *Transactions on Machine Learning Research*, 2024. Featured Certification. 10

[15] Xingchao Peng, Qinxun Bai, Xide Xia, Zijun Huang, Kate Saenko, and Bo Wang. Moment matching for multi-source domain adaptation. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 1406–1415, 2019. 2

[16] Nan Pu, Zhun Zhong, and Nicu Sebe. Dynamic conceptional contrastive learning for generalized category discovery. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 7579–7588, 2023. 2

[17] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In *International conference on machine learning*, pages 8748–8763. PmLR, 2021. 10

[18] Sai Bhargav Rongali, Sarthak Mehrotra, Ankit Jha, Shirsha Bose, Tanisha Gupta, Mainak Singha, Biplab Banerjee, et al. Cdad-net: Bridging domain gaps in generalized category discovery. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2616–2626, 2024. 2, 7, 8, 9

[19] Kuniaki Saito, Shohei Yamamoto, Yoshitaka Ushiku, and Tatsuya Harada. Open set domain adaptation by backpropagation. In *Proceedings of the European conference on computer vision (ECCV)*, pages 153–168, 2018. 4

[20] Sagar Vaze, Kai Han, Andrea Vedaldi, and Andrew Zisserman. Generalized category discovery. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7492–7501, 2022. 2, 7, 8

[21] 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. 2

[22] Shuo Wen and Maria Brbic. Cross-domain open-world discovery. *arXiv preprint arXiv:2406.11422*, 2024. 9

[23] Xin Wen, Bingchen Zhao, and Xiaojuan Qi. Parametric classification for generalized category discovery: A baseline study. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 16590–16600, 2023. 2, 7, 8

[24] Prateek Yadav, Derek Tam, Leshem Choshen, Colin A Raffel, and Mohit Bansal. Ties-merging: Resolving interference when merging models. *Advances in Neural Information Processing Systems*, 36, 2024. 7, 8

[25] Qingru Zhang, Minshuo Chen, Alexander Bukharin, Pengcheng He, Yu Cheng, Weizhu Chen, and Tuo Zhao. Adaptive budget allocation for parameter-efficient fine-tuning. In *The Eleventh International Conference on Learning Representations*, 2023. 11

[26] Sheng Zhang, Salman Khan, Zhiqiang Shen, Muzammal Naseer, Guangyi Chen, and Fahad Shahbaz Khan. Promptcal: Contrastive affinity learning via auxiliary prompts for generalized novel category discovery. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3479–3488, 2023. 2

[27] Bingchen Zhao and Kai Han. Novel visual category discovery with dual ranking statistics and mutual knowledge distillation. *Advances in Neural Information Processing Systems*, 34:22982–22994, 2021. 2

[28] Bingchen Zhao, Xin Wen, and Kai Han. Learning semi-supervised gaussian mixture models for generalized category discovery. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 16623–16633, 2023. 2

[29] Zhun Zhong, Enrico Fini, Subhankar Roy, Zhiming Luo, Elisa Ricci, and Nicu Sebe. Neighborhood contrastive learning for novel class discovery. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10867–10875, 2021. 2

[30] Zhun Zhong, Linchao Zhu, Zhiming Luo, Shaozi Li, Yi Yang, and Nicu Sebe. Openmix: Reviving known knowledge for discovering novel visual categories in an open world. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9462–9470, 2021. 2