Orthogonal Knowledge Refreshing for Domain-Incremental Object Detection

Aoting Zhang[1,3], Dongbao Yang[2†], Chang Liu[4], Xiaopeng Hong[5], Can Ma[1,3], and Yu Zhou[2†]

[1] Institute of Information Engineering, Chinese Academy of Sciences
[2] Nankai University
[3] School of Cyber Security, University of Chinese Academy of Sciences
[4] Tsinghua University
[5] Harbin Institute of Technology

Abstract. Domain-incremental object detection (DIOD) requires models to continually adapt to new domains while preserving prior knowledge. Recently, parameter-efficient fine-tuning offers a promising avenue, wherein a pre-trained model is frozen and a small number of learnable parameters are injected for downstream tasks. However, these methods risk overwriting critical past knowledge, triggering inter-domain interference and performance degradation. To address this challenge, we propose Orthogonal Knowledge Refreshing (OKR), a simple yet effective framework for DIOD. OKR incrementally constructs independent domain-specific subspaces via dedicated low-rank branches for each domain, which are seamlessly fused for a holistic decision, enabling conflict-free capacity expansion without domain selection during inference. To minimize knowledge interference during fusion, we present a gradient-based orthogonal refreshing strategy that projects gradient updates of new domains onto the orthogonal complement of the fused historical subspace, supporting continual adaptation without forgetting. Moreover, to mitigate semantic fragmentation across domains, we enforce topology-aware consistency, aligning the semantic structures of old and new domains. Extensive experiments validate the superiority of OKR, outperforming the best exemplar-free method by significant margins of +5.6% and +6.5% mAP on the Pascal VOC and BDD100K series, respectively.

Keywords: Domain-incremental object detection · Gradient-based orthogonal refreshing · Topology-aware consistency

1 Introduction

Deep learning detectors have achieved notable success in object detection [2, 21, 28, 49, 57], largely due to robust representations learned from large-scale datasets. However, most frameworks assume identical training and testing distributions,

† Corresponding author.

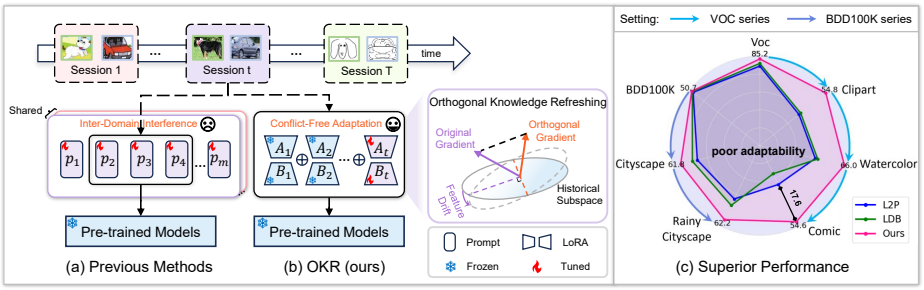

Fig. 1: (a) Previous methods suffer from inter-domain interference due to shared parameter overwriting. (b) OKR enables conflict-free adaptation by expanding dedicated LoRA branches for each session. Through orthogonal knowledge refreshing, it projects gradients orthogonally to the fused historical subspace, preserving prior knowledge while bypassing domain routing. (c) OKR achieves superior performance, delivering a significant +17.6% mAP gain on Comic dataset.

leading to performance drops under domain shifts [52]. Domain adaptation partially addresses this by transferring knowledge from a labeled source to an unlabeled target domain, but it typically supports one-time adaptation and fails under sequential shifts. Real-world applications like autonomous driving demand continual adaptation to changing environments (e.g., from sunny to foggy to rainy) while retaining performance across all domains. This necessitates Domain-Incremental Object Detection (DIOD), a challenging yet underexplored setting requiring both adaptability and long-term knowledge retention.

The central challenge in DIOD is catastrophic forgetting [11,58], where adapting to new domains overwrites previously acquired knowledge. While rehearsal-based methods seek to alleviate this by periodically replaying historical exemplars [29], they are frequently constrained by stringent privacy protocols and memory overhead. Knowledge distillation methods [27, 50, 56] attempt to preserve prior knowledge by mimicking the output behavior of old models, yet they struggle to handle substantial domain shifts and are fundamentally limited by the static capacity of the underlying architecture. Furthermore, architecture-based paradigms [7, 54] expand task-specific sub-networks to accommodate new tasks, incurring computational costs and memory burdens that scale poorly with the number of tasks.

Recently, parameter-efficient fine-tuning (PEFT), particularly prompt learning [26, 60], has emerged as a promising paradigm for domain incremental learning. For instance, L2P [46] dynamically selects prompts from a shared pool, while S-Prompt [43] employs domain-specific prompts coupled with a K-NN-based routing mechanism for domain identification. Similarly, LDB [33] learns domain biases with frozen base models and expands the predictors. However, these approaches overlook inter-domain interference, which undermines the trade-off between stability and plasticity. As illustrated in Figure 1(a), joint optimization within a shared parameter space inevitably distorts representations essential to prior domains, triggering detrimental feature drift. Moreover, prompt-based methods solely alter input embeddings, while bias-based ones apply additive

shifts to neuron outputs, both exerting shallow control over the model's internal mechanisms [44]. This restricted modulation proves insufficient under significant domain shifts, often resulting in suboptimal adaptability and performance degradation (Figure 1(c)). By comparison, low-rank adaptation (LoRA) facilitates intrinsic modulation by directly updating weight matrices, enabling fine-grained control and superior flexibility for DIOD scenarios.

Motivated by these insights, we propose Orthogonal Knowledge Refreshing (OKR), a simple yet effective framework for DIOD. Specifically, OKR incrementally constructs an independent domain-specific subspace for each new domain by introducing a dedicated LoRA branch, enabling conflict-free capacity expansion. During adaptation, only the newly added LoRA is updated. Owing to LoRA's inherent linearity, all branches are seamlessly fused for holistic prediction without redundant forward passes. To minimize knowledge interference during fusion, we devise a gradient-based orthogonal refreshing strategy that constrains gradient updates of the new branch to be orthogonal to historical subspaces. Drawing on the connection between input features and gradients [30], we approximate the composite prior subspaces via aggregated historical adaptation matrices and derive a projection matrix that redirects new gradients toward directions minimally correlated with existing knowledge, promoting stable adaptation without forgetting. Furthermore, to mitigate semantic fragmentation caused by domain transitions, we impose topology-aware consistency to align semantic structures across domains, facilitating semantically coherent representations. As shown in Figure 1(c), OKR consistently outperforms existing methods across all domains, with particularly significant gains of +17.6% mAP on the challenging Comic dataset, establishing new state-of-the-art performance.

Our contributions can be summarized as follows:

- We propose orthogonal knowledge refreshing (OKR), an effective framework for DIOD that incrementally expands low-rank domain-specific subspaces, enabling conflict-free adaptation without past data or domain selection, while keeping parameters fixed via LoRA's linear additivity.
- A gradient-based orthogonal refreshing strategy is designed to minimize interference with previously learned knowledge by constraining new-domain gradient directions to be orthogonal to historical subspaces.
- To alleviate semantic fragmentation across domains, we enforce topology-aware consistency, encouraging domain-specific subspaces to maintain semantically coherent and transferable structures.
- Extensive experiments on various settings demonstrate that OKR establishes new state-of-the-art performance in DIOD and effectively alleviates catastrophic forgetting.

2 Related Works

2.1 Class Incremental Learning

Class incremental learning aims to learn new classes sequentially over time without forgetting the previously learned classes [55, 59]. Current mainstream meth-

ods for CIL can be divided into three categories. Replay-based methods mitigate catastrophic forgetting by reintroducing representative samples of historical data [1,3,48] to replay or synthesize pseudo samples with a generative model [12] for training. Regularization-based methods impose additional constraints on network parameter updates, preventing drift [34] in weights critical to previous tasks or leveraging knowledge distillation to bridge old and new models [17, 51]. Architecture-based methods [7, 39, 46] achieve strong performance among other competitors by allocating distinct subsets of parameters to various subtasks, facilitating the expansion of network architecture.

2.2 Domain Incremental Learning

In domain incremental learning, the label space remains fixed while domains vary sequentially [32]. The goal is to adapt to new domains without retraining from scratch, while maintaining generalization on previously seen ones. Existing DIL methods can be categorized into replay-based [5] and rehearsal-free approaches. We focus on the latter due to their greater practicality. CIFRCN [13] extends region proposal networks to support incremental domain adaptation while ERD [10] adopts a response-based strategy. Among rehearsal-free methods, prompt-based tuning is particularly prominent. L2P [46] introduces a prompt pool and retrieves top-k prompts to guide feature alignment in the pre-trained model. DualPrompt [45] further separates prompts into task-invariant G-Prompts and task-specific E-Prompts. S-Prompt [43] learns domain-specific prompts via K-means clustering and K-NN-based routing. LDB [33] learns biases and predictors, while SOYO [41] enhances this by training a domain selector, both suffering from redundant computation due to repeated forward passes and limited cross-domain representational capacity. In this work, our orthogonal knowledge refreshing achieves conflict-free adaptation and efficient knowledge accumulation across incremental domains, bypassing the requirement for explicit domain routing.

2.3 Source-Free Domain Adaptation

A setting closely related to DIOD is source-free domain adaptation (SFDA), both aiming to ensure model robustness under evolving data distributions. SFDA adapts a source-trained model to an unlabeled target domain without access to the original source data. Existing SFDA approaches fall into three groups. The first constructs a pseudo source and aligns it with the target domain, treating the task as unsupervised domain adaptation. This is typically achieved via generative modeling [16, 36] or target domain partitioning based on source hypotheses [8]. The second group extracts and transfers domain-invariant factors from the source to the target domain to align feature distributions [6]. The third strategy leverages auxiliary target-domain information, such as geometry or manifold structure [35], beyond conventional pseudo-labeling [23]. However, SFDA is limited to one-time adaptation to a single target domain, whereas our

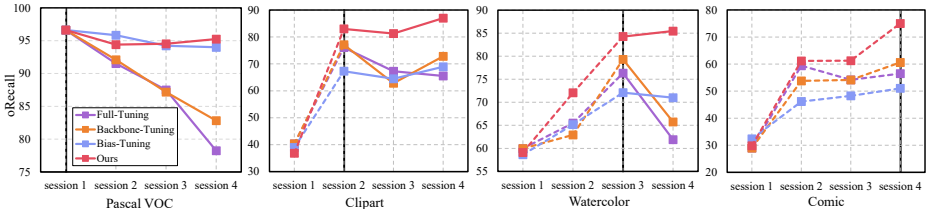

Fig. 2: Class-agnostic object recall (oRecall) on different domains over learning sessions. Solid lines indicate the session in which each domain is learned.

work addresses continual adaptation across sequential domain shifts without forgetting prior domains.

3 Preliminary

3.1 Observations

Performance degradation under domain shift primarily stems from feature distribution misalignment [25, 40, 47]. While mapping source and target features into an aligned representation space is effective for single-step adaptation [20, 31, 37], this principle collapses in domain-incremental object detection. In DIOD, naively forcing sequential domains into a single space triggers a representation conflict, where updates for a new domain inevitably overwrite feature statistics vital to previous ones, leading to catastrophic forgetting. This suggests that effective DIOD requires more than simple alignment, but necessitates non-destructive feature accumulation: the ability to adapt to new distributions while isolating and preserving old-domain representations.

To validate this premise, we conduct a diagnostic experiment measuring class-agnostic object recall (oRecall) across sequential domains. Unlike mAP, oRecall directly quantifies the model's ability to localize foreground objects independent of category-level classification, thus isolating the backbone's representational plasticity. As observed in Figure 2, backbone-tuning attains object recall comparable to full-tuning on new domains, confirming that the representational capacity of the backbone is sufficient to accommodate domain shifts. In contrast, bias-tuning which offers higher stability by updating fewer parameters, fails to capture complex distribution shifts, resulting in poor adaptation. Critically, both full and backbone-tuning suffer from a sharp performance drop on non-current domains. This indicates that joint optimization within a shared parameter space drives forgetting by corrupting the previously learned objectness scoring and feature statistics.

These findings establish the core motivation for our orthogonal knowledge refreshing framework. Instead of shallow bias shifts or risky global updates, OKR performs non-interfering feature accumulation via parameter-efficient subspace learning. By training only 1.8% of the parameters under explicit orthogonal constraints, we decouple the learning process into isolated subspaces, achieving a superior balance between rapid adaptation and long-term knowledge retention.

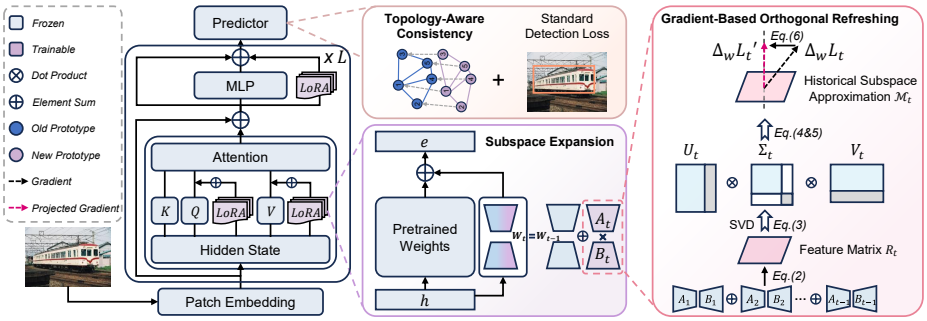

Fig. 3: Overview of OKR. To expand model capacity without conflicts, OKR incrementally constructs independent domain-specific subspaces by injecting dedicated LoRAs into the feature extractor. At session t, only the new LoRA (purple) is trainable, while previously ones (blue) remain frozen. During training and inference, all branches are seamlessly fused for holistic decision-making without redundant forward passes. To further suppress knowledge interference during weight fusion, we devise a gradient-based orthogonal refreshing strategy where A_t and B_t are updated by projecting the gradient $\nabla_w L_t$ onto the orthogonal complement of the approximated historical subspace $\mathcal{M}_t$, yielding the modified gradient $\nabla_w L_t'$. Moreover, we enforce topology-aware consistency to align class prototypes across domains, mitigating semantic fragmentation.

3.2 Problem Formulation

In DIOD, tasks involving T distinct domains are presented sequentially, each accompanied by its own dataset $D = \{D_1, D_2, \cdots, D_T\}$. For the i-th domain, the training set comprises k image-annotation pairs: $D_i = \{(X_i^1, Y_i^1), \cdots, (X_i^k, Y_i^k)\}$, where X_i^k denotes the input image and Y_i^k represents its ground-truth annotation. A key challenge in DIOD lies in the evolving data distributions across domains while maintaining a consistent label space. Our approach follows a sequential training paradigm under the **exemplar-free** constraint. Specifically, at the i-th stage, the model is trained solely on D_i, without accessing historical data from $\{D_1, \cdots, D_{i-1}\}$. The objective is to adapt to new domains while preserving performance on previously learned domains, ensuring knowledge retention across non-stationary data distributions.

4 Proposed Method

4.1 Overall Structure

OKR is built upon the ViTDet [21] detector, which is formulated as $h_\Theta(f_\Phi(\cdot))$. Here, $f_\Phi(\cdot)$ denotes the transformer-based feature extractor with parameters Φ, and $h_\Theta(\cdot)$ represents the linear predictor for object localization and classification with parameters Θ. Following LDB [33], we initialize the feature extractor with ImageNet-1K pretrained weights Φ_0 and train the base model Φ_1 on the first domain D_1 under standard detection loss supervision. Subsequently, OKR employs PEFT for continuous domain adaptation in an exemplar-free manner eliminating the need for storing redundant exemplars.

The architecture of our OKR is illustrated in Figure 3. To enable conflict-free capacity expansion for accommodating multiple domains, we propose a low-rank subspace expansion mechanism that decouples the parameter space into domain-specific subspaces by incrementally injecting dedicated LoRAs into the feature extractor. At session t, only the LoRA parameters A_t and B_t are trainable. During both training and inference, the weights of all LoRAs are aggregated via LoRA's linear properties to support unified decision-making, avoiding redundant forward passes. To further suppress interference with historical knowledge, we introduce gradient-based orthogonal refreshing (GOR) which explicitly constrains the gradient updates of A_t and B_t to be orthogonal to the approximated historical subspaces $\mathcal{M}_t$. Finally, we integrate a topology-aware consistency (TAC) loss into the training objective, which aligns prototype structures across domains to mitigate semantic fragmentation and enhance representational coherence.

4.2 Low-Rank Subspace Expansion

Owing to the inherent diversity and distinct characteristics of disparate domains, reusing shared LoRA parameter space across domains often induces interference and negative transfer of domain-specific knowledge. This challenge becomes more exacerbated in sequential domain-incremental scenarios, where adapting to a new domain may inadvertently distort representations optimized for previous domains. To alleviate such interference, we decouple the parameter space into domain-specific subspaces by incrementally refreshing dedicated LoRA components for new domains on top of a frozen backbone. Although each LoRA is optimized for its corresponding domain, the parameters are ultimately integrated through weight aggregation, maintaining a unified model with a *fixed parameter budget*. As illustrated in Figure 3, at session t, knowledge from the previous $t-1$ sessions is retained in frozen parameters $\{A_{1:t-1}, B_{1:t-1}\}$, while only the newly $\{A_t, B_t\}$ are trainable. The forward computation at session t is:

$$e = W_0 h + \sum_{i=1}^{t} B_i A_i h = W_{t-1} h + B_t A_t h = W_t h, \tag{1}$$

where $W_t = W_0 + \sum_{i=1}^{t} B_i A_i$ denotes the updated projection matrix (e.g., for query, value, or MLP layers), and h is the input hidden state. For clarity, the layer index is omitted and this process is applied at each transformer block.

This design facilitates incremental knowledge acquisition for new domains while preserving historical knowledge, without incurring parameter growth or redundant inference passes. However, freezing previous branches alone is insufficient to eliminate knowledge interference. While weight fusion avoids branch switching and improves inference efficiency, it introduces implicit coupling. Gradient updates for the current domain can distort the latent feature space established by historical branches, causing feature drift and performance degradation on earlier domains. To maintain subspace independence without increasing parameter count, we enforce orthogonal gradient constraints during training, en-

suring that updates for new domains occupy complementary directions in the shared low-rank space.

4.3 Gradient-Based Orthogonal Refreshing

Learning a new task exerts minimal interference on previously acquired knowledge when the gradient direction is orthogonal to the subspace spanned by features of old tasks [24]. We explore the intrinsic relationship between gradient directions and input feature spaces, aiming to identify core gradient subspaces of old domains and enforce orthogonality during updates on new domains. Within the PEFT paradigm, LoRA can effectively capture low-rank gradient subspace [42]. Building upon this capability, we propose an effective gradient-based orthogonal refreshing strategy embedded in a low-rank subspace expansion framework for DIOD. Concretely, we construct a domain-specific LoRA module for each new domain and constrain its gradient updates to be orthogonal to the historical subspace. This design achieves dual objectives: preserving the low-level feature distributions of previous domains to prevent forgetting and enabling efficient, interference-free domain adaptation, a crucial aspect in existing research.

Before training on the t-th task, we approximate the gradient subspace $\mathcal{M}_t$ spanned by previous tasks. Let $\mathcal{M}_t^{\perp}$ denote the residual gradient subspace orthogonal to $\mathcal{M}_t$, where gradient updates along $\mathcal{M}_t$ cause high interference to old tasks, while those along $\mathcal{M}_t^{\perp}$ induce minimum or no interference. To suppress interference from the new task, we constrain its updates to lie within $\mathcal{M}_t^{\perp}$. Our objective is to identify the basis of $\mathcal{M}_t$ or $\mathcal{M}_t^{\perp}$ and ensure gradient steps are orthogonal to $\mathcal{M}_t$ during training. Prior work [30] shows that the gradient update of a linear layer lies in the span of its inputs. Leveraging the linear superposition property of LoRA decompositions, we define the gradient space of task t using its input matrix and the aggregated domain-specific parameters from previous tasks, denoted by $W_{t-1} = \sum_{i=1}^{t-1} B_i A_i$. Specifically, we feed training samples from D_t into the model with linear layer parameterized by W_{t-1} to obtain layer-wise feature matrix R_t, followed by Singular Value Decomposition (SVD):

$$R_t = f_{\Phi}(W_{t-1}, D_t), \tag{2}$$

$$SVD(R_t) = U_t \Sigma_t (V_t)^T, \tag{3}$$

where $R_t \in \mathbb{R}^{m \times n}$, $U_t \in \mathbb{R}^{m \times m}$ and $V_t \in \mathbb{R}^{n \times n}$ are orthogonal, and Σ_t is a diagonal matrix with singular values sorted in descending order. To obtain a more compact representation, we apply a norm-based criterion to compute the k-rank approximation $(R_t)_k$ of R_t as follows:

$$\|(R_t)_k\|_F^2 \geq \epsilon \|R_t\|_F^2, \tag{4}$$

where $\|\cdot\|_F^2$ denotes the Frobenius norm, and $(R_t)_k$ retains the top-k singular values along the diagonal. The threshold ϵ controls the number of selected singular values and basis vectors. A larger ϵ yields a higher-dimension feature space. The

historical subspace is then spanned by the first k columns of U_t, which can be presented as:

$$\mathcal{M}_t = span\left\{u_t^1, u_t^2, \cdots, u_t^k\right\}. \tag{5}$$

After approximating the gradient space, the new gradient is first projected onto $\mathcal{M}_t$, and the projected component is subtracted to ensure the remaining gradient lies in the orthogonal complement $\mathcal{M}_t^{\perp}$. Specifically, during training on task t, only new LoRA parameters B_t and A_t in each layer are updated. Let $\nabla_w L_t$ denote the gradient generated at each iteration, where $w = \{A_t, B_t\}$. Then gradients are updated as follows:

$$\nabla_w L_t = \nabla_w L_t - (\nabla_w L_t)\mathcal{M}_t(\mathcal{M}_t)^T. \tag{6}$$

Beyond mitigating interference, our gradient refreshing strategy redirects learning toward residual subspaces beyond historical knowledge. This not only enhances the model's capacity to effectively acquire new knowledge but also establishes an optimal stability-plasticity balance.

4.4 Topology-Aware Consistency

While subspace expansion and orthogonal refreshing effectively capture cross-domain representations and alleviate inter-domain interference, the independently optimized domain-specific parameters introduce an inherent limitation. Isolated subspaces can fragment the global semantic coherence across domains. Although the label space remains shared, each subspace adapts solely to its respective domain, causing features of the same class to diverge in latent space. This disrupts intrinsic semantic alignment and gradually degrades cross-domain generalization. Such divergence accumulates with incremental steps, weakening the structural consistency of shared categories, which is crucial for robust multi-domain detection. To address this, we design topology-aware consistency, which explicitly regularizes semantic structures and bridges domain shifts.

We construct the topological structure in the feature space using class prototypes from the base domain, serving as a global reference for subsequent domains. Specifically, for each class c, we compute the prototype as: $\mu^c = \frac{1}{N^c}\sum_{i=1}^{N^c} p_i^c \cdot f_{\Phi}(x_i^c)$, where f_{Φ} is the feature extractor, p_i^c is the classification confidence for object i, and N^c is the number of instances in class c. The confidence-weighted average ensures robust prototype representation. The resulting prototype set $\mu_1 = [\mu_1^1, \cdots, \mu_1^N]$ reflects the fundamental semantic topology and remains invariant throughout incremental training process. For a new domain D_t, we obtain updated prototypes $\mu_t = [\mu_t^1, \cdots, \mu_t^N]$ and encourage them to align with their corresponding base prototypes with the same class, thereby mitigating fragmentation caused by parameter-isolated subspace learning:

$$L_{tac} = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N} w_t^{ij} \cdot (1 - cos(\mu_t^i, \mu_1^j)), 2 \le t \le T, \tag{7}$$

where $w_{ij} = \frac{exp(\mu_t^i \cdot \mu_1^j)}{\sum_{k=1}^{N} exp(\mu_t^i \cdot \mu_1^k)}$ denotes the matching confidence between new and old prototypes. Higher semantic similarity yields greater alignment weights to reduce intra-class divergence. $cos(\cdot, \cdot)$ represents the cosine similarity function. Minimizing L_{tac} promotes closer alignment between each new prototype and its corresponding base prototype, thereby enhancing intra-class compactness and maintaining structural consistency across domains.

4.5 Optimization

Combining the detection loss L_{det} and topology-aware consistency L_{tac}, the final objective when learning the new task t ($2 \leq t \leq T$) can be expressed as follows:

$$L_{final} = L_{det} + L_{tac}, \tag{8}$$

Note that when learning the base domain ($t = 1$), the model is optimized solely using the detection loss.

5 Experiments

5.1 Experimental Setups

Datasets and Evaluation Metrics. We conduct extensive experiments on three DIOD benchmarks with a broad spectrum of domain shifts, including style, scene and corruption variation. The *Pascal VOC series* comprises four domains with diverse styles: VOC 2007 [9], Clipart, Watercolor and Comic [18], with six shared object categories across all datasets. The *BDD100K series* focuses on scene-level shifts in autonomous driving, including BDD100k [53], Cityscape [4] and Rainy Cityscape [15]. BDD100k is among the largest and most diverse driving datasets publicly available. The *VOC-Corruption series* includes 15 domains with low-level visual corruptions [14] applied to VOC images, serving as a fine-grained benchmark for robustness under distributional perturbations. We adopt the mean average precision (mAP) at a 0.5 IoU threshold as the evaluation metric. At each session t, after training on data X^t, we evaluate the model on the test sets of all previously seen domains.

Implementation Details. We follow the implementation setup of previous work [33] to ensure consistency and comparability. Our framework is built on ViTDet, using ViT-Base as the backbone, initialized with ImageNet-1K pre-trained weights. All experiments are conducted under a strict rehearsal-free setting and performed using four NVIDIA RTX 3090 GPUs with a total batch size of 8. We adopt the AdamW optimizer with a learning rate of 2e-4 and a weight decay of 0.1. The LoRA module is configured with a rank of 16. After training on the first domain, all parameters are frozen except for a small subset introduced during parameter-efficient fine-tuning.

Table 1: Comparison results on Pascal VOC and BDD100K series. The best and second exemplar-free methods are in **bold** and underline.

Methods	Source	Pascal VOC series					BDD100K series			
		Exemplar	Session 2	Session 3	Session 4	ΔmAP	Exemplar	Session 2	Session 3	ΔmAP
Upper-bound	-	-	72.6	69.4	67.7	-	-	58.7	59.1	-
TP-DIOD-B [5]	TCSVT 23	150/domain	65.8	62.1	57.5	-7.7	200/domain	53.4	51.5	-6.7
FT-Seq	-	0/domain	57.5	52.6	49.5	-15.7	0/domain	51.6	43.6	-14.6
FT-FC	-		59.1	54.4	44.2	-21.0		51.3	48.0	-10.2
MCC [19]	ECCV 20		47.6	34.4	23.7	-41.5		44.3	36.1	-22.1
IRG [38]	CVPR 23		51.5	43.7	33.2	-32.0		49.3	38.7	-19.5
LwF [22]	TPAMI 17		60.4	53.6	53.2	-12.0		52.1	44.1	-14.1
PASS [61]	CVPR 21		61.7	51.4	49.8	-12.7		51.7	43.3	-14.9
L2P [46]	CVPR 22		59.9	55.2	45.5	-19.7		51.5	47.7	-10.5
S-Prompt [43]	NeurIPS 22		59.4	54.3	45.0	-20.2		51.6	49.4	-8.8
CIFRCN [13]	ICME 19		65.3	57.7	53.5	-11.7		51.8	48.9	-9.3
ERD [10]	CVPR 22		58.9	50.7	48.7	-16.5		51.1	48.1	-10.1
LDB [33]	AAAI 25		68.1	64.2	56.8	-8.4		52.3	51.1	-7.1
LDB+SOYO [41]	CVPR 25		69.2	65.3	59.6	-5.6		52.4	51.7	-6.5
OKR (ours)	-	0/domain	**73.0**	**67.3**	**65.2**	-0.0	0/domain	**57.6**	**58.2**	-0.0

Table 2: Comparison with other PEFT-based methods at the final session on VOC series and BDD100K series.

Method	VOC	Clipart	Watercolor	Comic	mAP	BDD100K	Cityscape	Rainy	mAP
L2P [46]	78.2	32.5	43.3	27.8	45.6	49.7	48.8	44.8	47.7
S-Prompt [43]	80.8	33.9	45.2	20.1	45.0	51.6	52.0	44.7	49.4
LDB [33]	82.4	50.1	57.5	37.0	56.8	50.3	52.7	50.2	51.1
OKR (ours)	**85.2**	**54.8**	**66.0**	**54.6**	**65.2**	**50.7**	**61.8**	**62.2**	**58.2**

5.2 Comparison Results

Results on Pascal VOC Series. The left section of Table 1 presents the experimental results on Pascal VOC series, where the model undergoes sequential adaptation across four domains: Pascal VOC 2007 → Clipart → Watercolor → Comic, evaluating the model's ability to retain prior knowledge and address escalating visual style variations. OKR consistently outperforms the best exemplar-free methods LDB and LDB+SOYO, achieving significant gains of +8.4% and +5.6% mAP in the final session. Notably, even when compared with the exemplar-based method TP-DIOD-B [5], our method demonstrates superior generalization by exceeding its performance by +7.7% at Session 4. Furthermore, OKR exhibits significant advantages over recent prompt-based and SFDA paradigms. For instance, at Session 4, it surpasses IRG [38] by +32.0% and L2P by +19.7%, underscoring the effectiveness of our orthogonal knowledge refreshing strategy in adapting to new domain distributions.

Results on BDD100K Series. As displayed in the right section of Table 1, OKR achieves consistent improvements on the BDD100K series, where domains progress in a realistic autonomous driving sequence: BDD100K → Cityscape → Rainy Cityscape. OKR surpasses all non-exemplar-based baselines, including LDB (+5.3% and +7.1%) and LDB+SOYO (+5.2% and +6.5%). Figure 4 visualizes the curves of average precision with the session number increasing. We observe that OKR maintains a clear lead with performance margins steadily

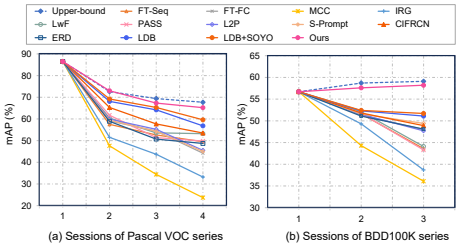

Fig. 4: Performance across sessions on VOC and BDD100K series. OKR consistently outperforms PEFT-based methods.

Table 3: Performance evolution on VOC-Corruption series.

Methods	Average precision in each session (%)			
	Session 2	Session 3	Session 4	Session 5
Upper-bound	65.8	63.3	78.8	74.3
FT-Seq	**70.5**	49.7	55.3	52.7
FT-FC	32.3	28.6	38.3	35.2
L2P [46]	40.2	28.4	37.5	32.3
S-Prompt [43]	39.5	27.9	36.2	32.6
ERD [10]	38.9	23.8	31.3	30.3
LDB [33]	52.3	49.8	54.8	50.7
OKR (ours)	67.3	**54.6**	**62.9**	**61.0**

Table 4: Detailed performance comparison after the final session across 16 corruption domains in VOC-C series. Our method achieves superior balance between stability on previously learned domains and plasticity on new corruptions.

Methods	Org	Noise			Blur				Weather				Digital				Avg.	Old	New
	Clean	Gau	Sht	Imp	Def	Gls	Mtn	Zm	Snw	Frs	Fog	Brt	Cnt	Els	Px	Jpg			
L2P [46]	79.6	13.4	18.1	13.5	16.1	18.3	19.0	23.1	36.8	37.2	49.3	45.8	33.1	37.9	37.2	37.5	32.3	30.9	36.4
S-Prompt [43]	81.8	16.1	22.4	17.1	20.9	23.8	25.1	30.9	38.6	38.4	53.8	39.9	28.4	29.2	25.7	29.6	32.6	34.1	28.2
ERD [10]	52.0	13.6	18.3	13.7	15.8	18.0	18.7	22.7	33.8	34.2	45.3	42.1	35.4	40.6	39.8	40.1	30.3	27.4	39.0
LDB [33]	83.9	41.3	43.3	43.1	43.3	**45.9**	**46.5**	37.3	62.8	65.0	58.7	75.7	39.5	64.5	23.5	37.4	50.7	53.9	41.2
OKR (ours)	**87.9**	**49.4**	**55.0**	**53.4**	**43.9**	44.2	45.9	**39.9**	**69.4**	**68.5**	**72.3**	**81.3**	**51.2**	**77.8**	**68.8**	**66.8**	**61.0**	**59.3**	**66.5**

widening. Notably, our method approaches the upper-bound performance, highlighting OKR's exceptional resilience to domain shifts.

Table 2 provides a more granular comparison against other parameter-efficient fine-tuning approaches in the final session. OKR a superior balance of knowledge retention and adaptability. Specifically, on the VOC series, it maintains high mAP scores on previously learned domains–85.2% (VOC), 54.8% (Clipart), and 66.0% (Watercolor), while achieving 54.6% on the newly introduced Comic domain. These results, echoed by similar trends on the BDD100K series, underscore OKR's enhanced cross-domain representational capacity and its unique ability to integrate knowledge across highly diverse visual contexts without explicit domain routing.

Results on VOC-Corruption Series. To assess the robustness under real-world conditions involving exposure to multiple domains within one session, we evaluate on a challenging 16-dataset sequence split into 5 incremental sessions based on corruption types: Clean → Noise (Gaussian, Shot, Impulse) → Blur (Defocus, Glass, Motion, Zoom) → Weather (Snow, Frost, Fog, Brightness) → Digital (Contrast, Elastic, Pixelate, Jpeg), simulating escalating environmental challenges from low-level noise to complex weather and digital distortions. Table 3 summarizes the results on all learned datasets after each session, while Table 4 reports the detailed performance across 16 domains after the final session, highlighting both stability (Old) and plasticity (New). LDB suffers from poor plasticity (41.2% vs. 66.4%), in which learning biases cannot overcome the inherent complexities of learning multiple domains concurrently. ERD experiences severe forgetting. OKR outperforms all baselines, achieving 61.0% mAP in Ses-

Table 5: Ablation of components where FT-back means fine-tuning backbone, SE denotes low-rank Subspace Expansion, GOR is used for Gradient-based Orthogonal Refreshing, and TAC for Topology-Aware Consistency.

Exp.	FT-Back	SE	GOR	TAC	Session 2	Session 3	Session 4
#1	-	-	-	-	59.1	54.4	44.2
#2	✓	-	-	-	67.7	57.9	49.5
#3	✓	✓	-	-	72.4	64.4	60.9
#4	✓	✓	✓	-	72.4	66.9	64.9
#5	✓	✓	-	✓	72.6	66.0	63.3
#6	✓	-	-	✓	67.7	59.8	60.2
#7	✓	✓	✓	✓	**73.0**	**67.3**	**65.2**

sion 5, +10.3% higher than LDB. Delved into individual domains, our method maintains strong performance, striking a better stability-plasticity trade-off.

5.3 Ablation Study

Effect of Different Components. We conduct extensive ablation studies on the Pascal VOC series to evaluate the individual contributions of each component in OKR, as summarized in Table 5. The baseline (Exp.#1) trained without anti-forgetting constraints, suffers from catastrophic forgetting, reaching only 44.2% mAP in the final session. In Exp.#2, updating only the backbone parameters improves performance to 49.5%, suggesting that the predictor is largely domain-agnostic and the primary challenge of domain shifts lies in feature-level adaptation. With this insight, our proposed subspace expansion (SE, Exp.#3) enables conflict-free adaptation without overwriting previous parameters, yielding a +16.7% gain in Session 4. Gradient-based orthogonal refreshing (GOR) in Exp.#4 constrains gradient update to be orthogonal to past subspaces, acquiring unique domain-specific knowledge and preventing interference, which improves performance by +20.7%. Separately, topology-aware consistency (TAC, Exp.#6) independently boosts performance by 10.7% by mitigating semantic fragmentation. The synergy between SE and TAC (Exp.#5) confirms the necessity of structural alignment even across isolated subspaces. Finally, the complete OKR framework achieves the best result, establishing a powerful pipeline for DIOD.

Effect of Different Ranks. We investigate the impact of different ranks ($r = 8, 16, 64$) for the expandable LoRA subspaces. As shown in Figure 5(a), smaller ranks lead to a performance decline, due to insufficient capacity to represent complex domain-specific information. Interestingly, performance plateaus as the rank increases beyond 16, indicating that the model's gradient space possesses a relatively low intrinsic dimensionality. We set r=16 as it provides an optimal balance between model adaptability and computational efficiency.

Analysis of Topology-Aware Consistency. To further validate the efficacy of L_{tac}, we examine the topological evolution of class prototypes during incremental learning sessions. Specifically, we construct semantic structures by calculating cosine similarity matrices between class prototypes within both base and target domains. The resulting aggregated difference matrix, averaged over

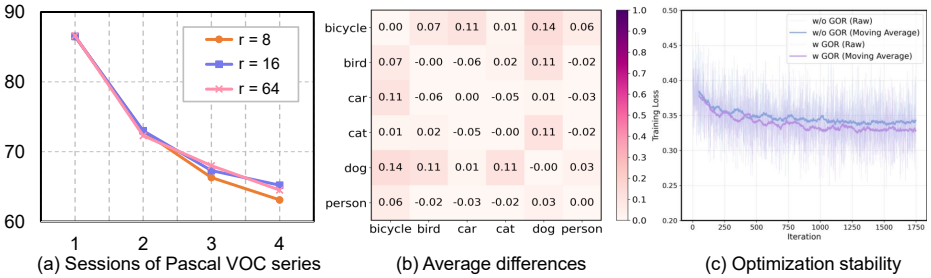

Fig. 5: Further analysis of OKR. (a) OKR remains robust under different ranks of the expandable LoRA subspaces. (b) Topology-aware consistency preserves semantic topology across domain increments. (c) GOR maintains stable optimization, with similar training loss curves with and without orthogonal refreshing.

all sessions, quantifies the cumulative structural drift throughout the learning process. As illustrated in Figure 5(b), the values remain predominantly near zero, indicating that our framework preserves the underlying semantic topology with negligible distortion. This confirms that L_{tac} effectively anchors emerging domain representations to the foundational semantic structure, successfully mitigating semantic fragmentation across sequential tasks.

Effect of Orthogonality on Stability. Figure 5(c) demonstrates that the proposed orthogonality constraint ensures stable and consistent optimization throughout the training process. The training loss curves with and without GOR exhibit nearly identical convergence, indicating that the orthogonal projection does not introduce additional optimization difficulties or instability. This stability aligns with the theoretical design of OKR, where GOR restricts new-domain updates to the residual gradient subspace orthogonal to historical ones. By isolating these updates, the model effectively minimizes interference with prior knowledge while maintaining subspace independence. Rather than disrupting the general convergence pattern, orthogonality enhances continual adaptation by directing the model to acquire complementary, domain-specific features, which accounts for the significant performance gains observed in the ablation study.

Computational Efficiency. The proposed subspace expansion strategy introduces a marginal 1.8% increase in total parameters, which is negligible relative to the substantial performance gains achieved. By leveraging the linear additivity of LoRA, all domain-specific branches are seamlessly merged into a unified parameter set after training, *ensuring a fixed model size and eliminating the need for domain routing during inference*. Regarding training efficiency, the projection overhead is nearly negligible: SVD is performed only once per session, while per-step orthogonal refreshing involves solely low-rank matrix multiplications. This results in a mere 0.02% increase in training time with only 1.8% memory overhead. Empirically, OKR achieves an inference speed of 0.1348 s/sample with a 5.1 GB memory footprint on an RTX-3090, significantly outperforming LDB (0.2836 s/sample, 6.0 GB) due to OKR's efficient single-stage architecture.

6 Conclusion

In this work, we propose orthogonal knowledge refreshing (OKR), a simple yet effective framework for domain-incremental object detection. Grounded in the insight that backbone tuning provides sufficient plasticity but lacks stability due to representation overwriting, OKR leverages isolated domain-specific LoRA branches to facilitate conflict-free capacity expansion. By exploiting the linear additivity of low-rank matrices, all branches are seamlessly fused into a unified model, ensuring *zero architectural overhead* and eliminating the need for complex domain routing during inference. To eliminate knowledge interference during this fusion, we introduce a gradient-based orthogonal refreshing strategy that constrains new updates to the orthogonal complement of historical subspaces. This approach effectively redirects learning toward novel information while safeguarding the integrity of prior features. Moreover, we enforce topology-aware consistency to maintain semantic coherence across domains. Extensive experiments demonstrate OKR's superiority, establishing a new state-of-the-art of DIOD.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (Grant NO 62406318, 62376266, 62406167, U24B6012, 62376070 and 62076195).

References

1. Bang, J., Kim, H., Yoo, Y., Ha, J.W., Choi, J.: Rainbow Memory: Continual learning with a memory of diverse samples. In: CVPR. pp. 8218–8227 (2021)
2. Cao, T., Lyu, J., Zeng, W., Mu, W., Zhou, Y.: The devil is in fine-tuning and long-tailed problems: a new benchmark for scene text detection. arXiv preprint arXiv:2505.15649 (2025)
3. Chaudhry, A., Ranzato, M., Rohrbach, M., Elhoseiny, M.: Efficient lifelong learning with a-gem. arXiv preprint arXiv:1812.00420 (2018)
4. Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., Schiele, B.: The cityscapes dataset for semantic urban scene understanding. In: CVPR. pp. 3213–3223 (2016)
5. Ding, L., Song, X., He, Y., Wang, C., Dong, S., Wei, X., Gong, Y.: Domain incremental object detection based on feature space topology preserving strategy. In: IEEE TCSVT. vol. 34, pp. 424–437 (2023)
6. Ding, N., Xu, Y., Tang, Y., Xu, C., Wang, Y., Tao, D.: Source-free domain adaptation via distribution estimation. In: CVPR. pp. 7212–7222 (2022)
7. Douillard, A., Ramé, A., Couairon, G., Cord, M.: Dytox: Transformers for continual learning with dynamic token expansion. In: CVPR. pp. 9285–9295 (2022)
8. Du, Y., Yang, H., Chen, M., Luo, H., Jiang, J., Xin, Y., Wang, C.: Generation, augmentation, and alignment: A pseudo-source domain based method for source-free domain adaptation. In: ML. vol. 113, pp. 3611–3631 (2024)
9. Everingham, M., Van Gool, L., Williams, C.K., Winn, J., Zisserman, A.: The pascal visual object classes (voc) challenge. In: IJCV. vol. 88, pp. 303–338 (2010)

10. Feng, T., Wang, M., Yuan, H.: Overcoming catastrophic forgetting in incremental object detection via elastic response distillation. In: CVPR. pp. 9427–9436 (2022)
11. French, R.M.: Catastrophic forgetting in connectionist networks. In: TCS. vol. 3, pp. 128–135 (1999)
12. Gao, R., Liu, W.: DDGR: Continual learning with deep diffusion-based generative replay. In: ICML. pp. 10744–10763 (2023)
13. Hao, Y., Fu, Y., Jiang, Y.G., Tian, Q.: An end-to-end architecture for class-incremental object detection with knowledge distillation. In: ICME. pp. 1–6 (2019)
14. Hendrycks, D., Dietterich, T.: Benchmarking neural network robustness to common corruptions and perturbations. arXiv preprint arXiv:1903.12261 (2019)
15. Hu, X., Fu, C.W., Zhu, L., Heng, P.A.: Depth-attentional features for single-image rain removal. In: CVPR. pp. 8022–8031 (2019)
16. Huang, J., Guan, D., Xiao, A., Lu, S.: Model Adaptation: Historical contrastive learning for unsupervised domain adaptation without source data. In: NeurIPS. vol. 34, pp. 3635–3649 (2021)
17. Huang, L., Zeng, Y., Yang, C., An, Z., Diao, B., Xu, Y.: eTag: Class-incremental learning via embedding distillation and task-oriented generation. In: AAAI. vol. 38, pp. 12591–12599 (2024)
18. Inoue, N., Furuta, R., Yamasaki, T., Aizawa, K.: Cross-domain weakly-supervised object detection through progressive domain adaptation. In: CVPR. pp. 5001–5009 (2018)
19. Jin, Y., Wang, X., Long, M., Wang, J.: Minimum class confusion for versatile domain adaptation. In: ECCV. pp. 464–480 (2020)
20. Li, W., Liu, X., Yuan, Y.: SIGMA++: improved semantic-complete graph matching for domain adaptive object detection. IEEE TPAMI **45**(7), 9022–9040 (2023)
21. Li, Y., Mao, H., Girshick, R., He, K.: Exploring plain vision transformer backbones for object detection. In: ECCV. pp. 280–296 (2022)
22. Li, Z., Hoiem, D.: Learning without forgetting. In: IEEE TPAMI. vol. 40, pp. 2935–2947 (2017)
23. Liang, J., Hu, D., Feng, J.: Do we really need to access the source data? source hypothesis transfer for unsupervised domain adaptation. In: ICML. pp. 6028–6039 (2020)
24. Liang, Y.S., Li, W.J.: Adaptive plasticity improvement for continual learning. In: CVPR. pp. 7816–7825 (2023)
25. Long, M., Cao, Y., Wang, J., Jordan, M.: Learning transferable features with deep adaptation networks. In: ICML. vol. 37, pp. 97–105 (July 2015)
26. Lu, Y., Liu, J., Zhang, Y., Liu, Y., Tian, X.: Prompt distribution learning. In: CVPR. pp. 5206–5215 (June 2022)
27. Peng, C., Zhao, K., Maksoud, S., Li, M., Lovell, B.C.: SID: Incremental learning for anchor-free object detection via selective and inter-related distillation. In: CVIU. vol. 210, p. 103229 (2021)
28. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: Towards real-time object detection with region proposal networks. In: NeurIPS. vol. 28, pp. 91–99 (2015)
29. Rolnick, D., Ahuja, A., Schwarz, J., Lillicrap, T., Wayne, G.: Experience replay for continual learning. In: NeurIPS. vol. 32, pp. 350–360 (2019)
30. Saha, G., Garg, I., Roy, K.: Gradient projection memory for continual learning. arXiv preprint arXiv:2103.09762 (2021)
31. Saito, K., Ushiku, Y., Harada, T., Saenko, K.: Strong-weak distribution alignment for adaptive object detection. In: CVPR (June 2019)
32. Shi, H., Wang, H.: A unified approach to domain incremental learning with memory: Theory and algorithm. In: NeurIPS. vol. 36, pp. 15027–15059 (2023)

33. Song, X., He, Y., Dong, S., Gong, Y.: Non-exemplar domain incremental object detection via learning domain bias. In: AAAI. vol. 38, pp. 15056–15065 (2024)
34. Tang, S., Chen, D., Zhu, J., Yu, S., Ouyang, W.: Layerwise optimization by gradient decomposition for continual learning. In: CVPR. pp. 9634–9643 (2021)
35. Tang, S., Yang, Y., Ma, Z., Hendrich, N., Zeng, F., Ge, S.S., Zhang, C., Zhang, J.: Nearest neighborhood-based deep clustering for source data-absent unsupervised domain adaptation. arXiv preprint arXiv:2107.12585 (2021)
36. Tian, J., Zhang, J., Li, W., Xu, D.: VDM-DA: Virtual domain modeling for source data-free domain adaptation. In: IEEE TCSVT. vol. 32, pp. 3749–3760 (2021)
37. VS, V., Gupta, V., Oza, P., Sindagi, V.A., Patel, V.M.: MeGA-CDA: Memory guided attention for category-aware unsupervised domain adaptive object detection. In: CVPR. pp. 4516–4526 (June 2021)
38. VS, V., Oza, P., Patel, V.M.: Instance relation graph guided source-free domain adaptive object detection. In: CVPR. pp. 3520–3530 (2023)
39. Wang, F.Y., Zhou, D.W., Liu, L., Ye, H.J., Bian, Y., Zhan, D.C., Zhao, P.: BEEF: Bi-compatible class-incremental learning via energy-based expansion and fusion. In: ICLR (2022)
40. Wang, M., Wang, S., Wang, W., Shen, L., Zhang, X., Lan, L., Luo, Z.: Reducing bi-level feature redundancy for unsupervised domain adaptation. PR **137**, 109319 (2023)
41. Wang, Q., Song, X., He, Y., Han, J., Ding, C., Gao, X., Gong, Y.: Boosting domain incremental learning: Selecting the optimal parameters is all you need. In: CVPR. pp. 4839–4849 (June 2025)
42. Wang, X., Chen, T., Ge, Q., Xia, H., Bao, R., Zheng, R., Zhang, Q., Gui, T., Huang, X.: Orthogonal subspace learning for language model continual learning. arXiv preprint arXiv:2310.14152 (2023)
43. Wang, Y., Huang, Z., Hong, X.: S-Prompts learning with pre-trained transformers: An occam's razor for domain incremental learning. In: NeurIPS. vol. 35, pp. 5682–5695 (2022)
44. Wang, Y., Chauhan, J., Wang, W., Hsieh, C.J.: Universality and limitations of prompt tuning. In: NeurIPS. vol. 36, pp. 75623–75643 (2023)
45. Wang, Z., Zhang, Z., Ebrahimi, S., Sun, R., Zhang, H., Lee, C.Y., Ren, X., Su, G., Perot, V., Dy, J., et al.: DualPrompt: Complementary prompting for rehearsal-free continual learning. In: ECCV. pp. 631–648 (2022)
46. Wang, Z., Zhang, Z., Lee, C.Y., Zhang, H., Sun, R., Ren, X., Su, G., Perot, V., Dy, J., Pfister, T.: Learning to prompt for continual learning. In: CVPR. pp. 139–149 (2022)
47. Xie, B., Yuan, L., Li, S., Liu, C.H., Cheng, X., Wang, G.: Active learning for domain adaptation: An energy-based approach. In: AAAI. vol. 36, pp. 8708–8716 (2022)
48. Yang, D., Zhou, Y., Hong, X., Zhang, A., Wang, W.: One-shot replay: Boosting incremental object detection via retrospecting one object. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 37, pp. 3127–3135 (2023)
49. Yang, D., Zhou, Y., Hong, X., Zhang, A., Wei, X., Zeng, L., Qiao, Z., Wang, W.: Pseudo object replay and mining for incremental object detection. In: Proceedings of the 31st ACM International Conference on Multimedia. pp. 153–162 (2023)
50. Yang, D., Zhou, Y., Shi, W., Wu, D., Wang, W.: RD-IOD: Two-level residual-distillation-based triple-network for incremental object detection. ACM Transactions on Multimedia Computing, Communications, and Applications **18**(1), 1–23 (2022)

51. Yang, D., Zhou, Y., Zhang, A., Sun, X., Wu, D., Wang, W., Ye, Q.: Multi-view correlation distillation for incremental object detection. In: PR. vol. 131, p. 108863 (2022)
52. Yao, X., Zhao, S., Xu, P., Yang, J.: Multi-source domain adaptation for object detection. In: ICCV. pp. 3273–3282 (2021)
53. Yu, F., Chen, H., Wang, X., Xian, W., Chen, Y., Liu, F., Madhavan, V., Darrell, T.: Bdd100k: A diverse driving dataset for heterogeneous multitask learning. In: CVPR. pp. 2636–2645 (2020)
54. Zhang, A., Yang, D., Liu, C., Hong, X., Shang, M., Zhou, Y.: DCA: Dividing and conquering amnesia in incremental object detection. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 39, pp. 9851–9859 (2025)
55. Zhang, A., Yang, D., Liu, C., Hong, X., Zhou, Y.: Specifying what you know or not for multi-label class-incremental learning. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 39, pp. 22345–22353 (2025)
56. Zhang, A., Yang, D., Liu, C., Hong, X., Zhou, Y.: Focus, Align, and Sustain: Counteracting gradient dilution in incremental object detection. arXiv preprint arXiv:2606.15253 (2026)
57. Zhang, D., Lyu, J., Shen, Z., Zhou, Y.: Class-agnostic region-of-interest matching in document images. In: International Conference on Document Analysis and Recognition. pp. 446–464 (2025)
58. Zhang, L., Qian, L., Li, R., Li, T., Zhang, W.: Leveraging multi-modal and historical knowledge graphs for continual robot navigation. Visual Intelligence **4**(1), 15 (2026)
59. Zhou, D.W., Wang, Q.W., Qi, Z.H., Ye, H.J., Zhan, D.C., Liu, Z.: Class-incremental learning: A survey. IEEE TPAMI (2024)
60. Zhou, K., Yang, J., Loy, C.C., Liu, Z.: Conditional prompt learning for vision-language models. In: CVPR. pp. 16816–16825 (2022)
61. Zhu, F., Zhang, X.Y., Wang, C., Yin, F., Liu, C.L.: Prototype augmentation and self-supervision for incremental learning. In: CVPR. pp. 5871–5880 (2021)