2025 IEEE International Conference on Multimedia and Expo (ICME) | 979-8-3315-9495-4/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICME59968.2025.11209892

Federated Open-Set Domain Generalization with Adaptive Adjustment Boundary and Weights

Haoyuan Liang*,1, Shilei Cao*,1, Yushan Lai[1], and Juepeng Zheng†,1
[1] School of Artificial Intelligence, Sun Yat-sen University, China

***Abstract*—Concerns about privacy and the centralized collection of sensitive data have led to the development of Federated Learning, a paradigm enabling collaborative model training without the need to aggregate raw data centrally. However, variations in data distributions between source and target clients, a phenomenon known as domain shift, often lead to degraded model performance. While recent advancements in Federated Domain Generalization address this challenge, they typically operate under a closed-set assumption, disregarding scenarios where target domains introduce entirely new classes, referred to as category shift. This oversight can result in critical misclassifications in real-world applications. To overcome these limitations, we explore Federated Open-Set Domain Generalization (FedOSDG) setting for the first time, which not only preserves data privacy but also identifies new, unseen classes in the unseen target domains. Specifically, we propose the Adaptive Adjustment Boundary and Weights (AABAW) framework, comprising Stronger Classification Boundary (SCB) and Adaptive Adjustment of Weights (AAW). The SCB module reinforces the decision boundaries of the binary classifiers to handle category shift while the AAW module leverages local model diversity to increase the variance of global model, thereby enhancing the model's generalization under domain shifts. Experimental results show that our proposed AABAW achieves state-of-the-art performance in recognizing unknown classes and H-scores in the FedOSDG task with considerable gains and maintains competitive performance in all classes.**

***Index Terms*—federated learning, open-set recognition, domain generalization**

I. Introduction

Due to concerns regarding the centralized collection of sensitive data, Federated Learning (FL) has emerged as a decentralized solution that enables collaborative model training without pooling private data centrally to prioritize data privacy [1]. However, when the data distribution of source clients differs from that of a target client, the resulting domain shift can lead to catastrophic performance degradation [2]–[6]. In real-world settings, such domain shifts are generally inevitable and unforeseeable, given that the training dataset only represents the subset of real-world data.

To tackle domain shift, Federated Domain Generalization (FedDG) has been proposed to learn a model on source clients' data that can generalize to unseen target client data in a privacy-preserving manner [7]. However, existing methods typically assume that the label space in the source domain mirrors that of the target domain. This assumption is often unrealistic in practice: target domains may contain novel classes absent from the source training set, thereby shifting the learning paradigm from a closed-set to an open-set scenario [8]. Moreover, the presence of unknown classes can substantially impede the transfer of knowledge to the target domain. As FedDG does not ensure comprehensive coverage of novel classes in the target domain, conventional closed-set loss functions can inadvertently force these unknown classes into existing known categories, posing severe risks in real-world applications. In response, Federated Open-Set Recognition (FedOSR) methods have been introduced to detect new classes within an FL framework [9]. However, these methods typically overlook the challenges posed by domain shift, hindering their practical applications.

In this work, we formulate and investigate Federated Open-Set Domain Generalization (FedOSDG) for the first time (refer to Fig. 1). FedOSDG aims to safeguard data privacy while simultaneously recognizing new, unseen classes in domains that have not been encountered during training. Although FedOSDG combines the respective advantages of FedDG and FedOSR, it also inherits two principal challenges: **category shift** and **domain shift**. The former arises when the label space

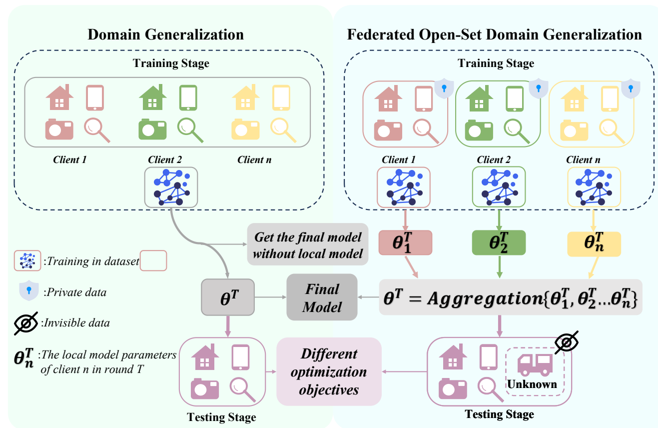

Fig. 1. There are two key differences between DG and FedOSDG: (1) In DG, different source domains are interconnected, whereas in FedOSDG, source clients are isolated from one another to ensure data privacy. (2) In DG, the label space of source and target domains is identical, whereas in FedOSDG, target clients may contain novel unknown categories.

* These authors contributed equally to this work.

† Author to whom any correspondence should be addressed. E-mail: zhengjp8@mail.sysu.edu.cn

This work was supported by the National Natural Science Foundation of China (42401415) and Jiangsu Innovation Capacity Building Program (Project BM2022028).

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on March 22,2026 at 06:50:51 UTC from IEEE Xplore. Restrictions apply.

TABLE I
COMPARISON OF TARGET DOMAINS UNDER DIFFERENT TASKS. $\mathcal{C} \cap \mathcal{U} = \varnothing$.

Problem Setting	Data Dispersal	Label Space	Participation in Training
Federated Learning	✓	$\mathcal{C}$	✓
Domain Generalization	×	$\mathcal{C}$	×
Open-set Domain Generalization	×	$\mathcal{C} \cup \mathcal{U}$	×
Federated Domain Generalization	✓	$\mathcal{C}$	×
Federated Open-Set Domain Generalization	✓	$\mathcal{C} \cup \mathcal{U}$	×

of the target clients differs from that of the source clients [10], whereas the latter denotes changes in the data distribution from the source clients to the target clients [11].

To tackle these challenges in FedOSDG, we propose Adaptive Adjustment Boundary and Weights (AABAW), comprising two key components: the Stronger Classification Boundary (SCB) and the Adaptive Adjustment of Weights (AAW). Specifically, SCB focuses on sharpening the classification boundaries by reinforcing the confidence in the binary classifier corresponding to the most confident category while penalizing the classifier of closely related category, thus making uncertain class probabilities more decisive (e.g., converting an ambiguous 0.5 probability for “dog” into 0.9 or 0.1). The resulting strengthened decision boundaries help the model better distinguish known classes from unseen ones, thereby addressing **category shift**. Meanwhile, AAW adaptively increases the weight of the local model that deviates from the average model, effectively leveraging local model heterogeneity to increase the variance in the global model—an essential factor for improving generalization under unseen domain distributions, thus addressing **domain shift**. By contrast, most existing methods rely on straightforward averaging that suppresses diversity and can hinder out-of-distribution performance.

Our main contributions are summarized as follows:

- We tackle an emerging and practical issue of Federated Open Set Domain Generalization (FedOSDG). To the best of our knowledge, this is the first work to improve the performance of federated domain generalization models for unknown class recognition in unknown domains.
- We propose Adaptive Adjustment Boundary and Weights (AABAW) to address two main challenges in FedOSDG: **category shift** and **domain shift**. AABAW integrates SCB and AAW modules to achieve robust performance in recognizing unknown classes in unseen domains.
- We conduct comprehensive experiments on two common datasets (PACS, Office-Home) following the setting of FedOSDG. Experimental results show that AABAW achieves state-of-the-art performance on the FedOSDG task with considerable gains.

II. METHOD

A. Preliminaries

Tab. I compares the differences of problem setting between FedOSDG and related tasks. In FedOSDG, we define the set of all domains as $D = \{D^1, D^2, ..., D^C\}$. Among these, E source domains (clients), denoted as $\mathcal{S} = \{\mathcal{S}^e\}_{e=1}^{E}$, is sampled

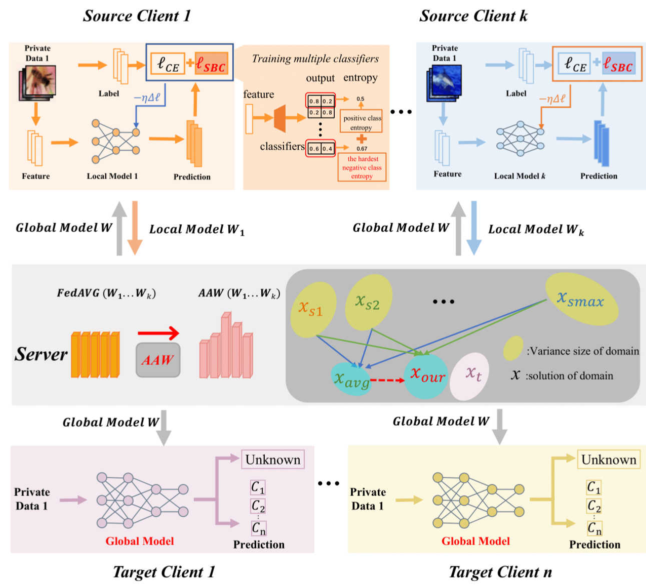

Fig. 2. The framework for Adaptive Adjustment Boundary and Weights (AABAW), designed for Federated Open-Set Domain Generalization (FedOSDG). AABAW consists of two core modules: (1) the Stronger Classification Boundary (SCB) module, which addresses ***category shift*** by reinforcing the decision boundaries of the binary classifiers, and (2) the Adaptive Adjustment of Weights (AAW) module, which improves model variance by redistributing weights during aggregation rounds to tackle ***domain shift***.

correspondingly from $\{D^e\}_{e=1}^{E}$. $(\mathcal{X}, \mathcal{Y})$ represent the image and label space of source domain. Meanwhile, the remaining $C - E$ domains are designated as target domains is sampled from $\{D^e\}_{e=1}^{C-E}$, denoted as $\mathcal{T} = \{\mathcal{T}^e\}_{e=1}^{C-E}$.

For each source domain, $\mathcal{S}^e$ contains data and label pairs of $\{(x_j^e, y_j^e)\}_{j=1}^{N^e}$ where N^e represent the total number of data samples in the source domain. For the target domain, its label space includes not only $\mathcal{Y}$ but also several novel labels, which we uniformly denote as “unknown”. The objective of FedOSDG is to enhance the performance of global model, which is aggregated from decentralized local models trained on source clients, when applied to unseen target clients.

In this work, our proposed Adaptive Adjustment Boundary and Weights (AABAW) for FedOSDG incorporates two types of losses. The overall loss function, $\mathcal{L}$, is defined as:

$$\mathcal{L} = \mathcal{L}_{cls} + \rho \mathcal{L}_{open} \tag{1}$$

where $\mathcal{L}_{cls}$ represents the closed-set loss function, which evaluates the discrepancy between the predicted label $\hat{y} = f(x; w)$ (w is the model parameter) and the ground-truth label y. $\mathcal{L}_{open}$ is the open-set loss function whose specific definition is provided in Section II-B. The parameter ρ controls the trade-off between $\mathcal{L}_{cls}$ and $\mathcal{L}_{open}$.

B. Stronger Classification Boundary (SCB)

The conventional method for handling open-set data typically involves setting a threshold and classifying entities scoring below this threshold as “unknown”. While straightforward, this approach encounters difficulties with similar classes, often resulting in multiple comparably high scores that

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on March 22,2026 at 06:50:51 UTC from IEEE Xplore. Restrictions apply.

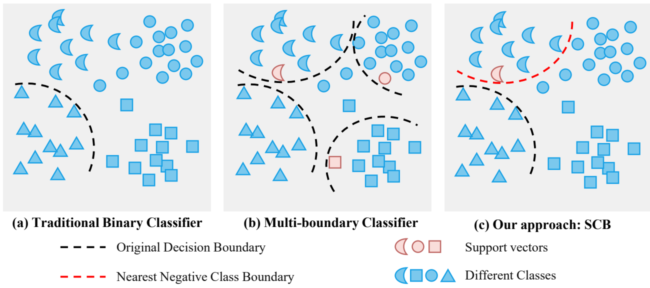

Fig. 3. Utilizing a triangle-recognition classifier as our example, we denote (a) as the traditional binary classifier, (b) as the multi-boundary training approach characterized, and (c) as our proposed method Stronger Classification Boundary (SCB), which focuses on training a pair of boundaries.

still fall below the threshold. For instance, consider a model trained to recognize cats, tables, and dogs. While it can easily differentiate between cats and tables, the significant feature overlap between cats and dogs—both being animals—makes distinguishing between them challenging.

A natural strategy is to reinforce the classification boundary across all classes. As shown in Fig. 3(b), the severe gradient disperse is caused by each class boundary participating in the gradient descent process. Given an input sample pair $(\mathbf{x}^e, y^e)$, this type of open-set loss $\mathcal{L}_{open}$ can be articulated as:

$$\mathcal{L}_{open}(x^e, y^e) = -\log(p(\hat{y}^{y^e}|x^e)) - \sum_{\substack{j=1\\ j\neq y^e}}^{||\mathcal{Y}||} \log(1 - p(\hat{y}^j|x^e)) \tag{2}$$

where $p(\hat{y}^j|\mathbf{x}^e)$ denotes the posterior probability of the classifier assigning the sample with feature $\mathcal{X}$ to the class j.

Hence, we propose the SCB module to maximize the utility of boundary separation and strengthen classification boundaries, as illustrated in Fig. 3(c). The pink samples in Fig. 3 represent the support vectors defined by the decision boundary [12]. Given the input sample is $(\mathbf{x}^e, y^e)$, the SCB loss function is computed as follows:

$$\mathcal{L}_{SCB}(x^e, y^e) = -\log(p(\hat{y}^{y^e}|x^e)) - \min_{j\neq y^e} \log(1 - p(\hat{y}^j|x^e)) \tag{3}$$

As Eq. (3) shown, during gradient descent, the probability of correctly identifying the previous classes is progressively increased, while the fuzzy binary classifier becomes more distinct. This approach subsequently minimize instances of misidentification. We improve the ability of the model to recognize unknown data in this way thereby overcoming the category shift. Consequently, the model's ability to recognize known label from unknown labels is significantly enhanced, effectively mitigating the challenges posed by category shift.

While our proposed SCB does enhance the model to recognize open set classes, it inadvertently impairs the recognition of closed set classes to some degree. This occurs because the open-set loss function disperses the primary gradient descent direction, deviating from the previously exclusive focus on the closed-set loss. To address this, we can adjust the parameter ρ to balance the closed-set loss and open-set loss.

C. Adaptive Adjustment of Weights (AAW)

In some Federated Learning (FL) studies, it has been observed that FedAvg [1] performs suboptimally on heterogeneous data. Adaptive FL methods [13] have emerged as a mainstream approach to address non-iid data. However, in the context of FedDG, the data from the target domain is inaccessible, making it challenging to adapt specifically to the target domain. Moreover, prior adaptive FL methods often aim to closely align with the average model, which results in poor generalization performance on the unseen target domain.

In Fig. 4, when handling iid data, the FedAvg algorithm effectively reduces variance and improves performance on the source domain during aggregation. However, in FedDG, the presence of significant data heterogeneity caused by domain shift hinders the aggregated FL model's applicability to the target domain. To address the **domain shift**, the Adaptive Adjustment of Weights (AAW) module not only increases variance to alleviate overfitting but also shifts the model's focus away from the source domain and closer to the target domain. This approach enables better adaptation and improved generalization across diverse domains. Utilizing the Lagrange multiplier technique and scaling approach, we derive:

$$\frac{M}{m}(n_1^2\sigma_1^2 + n_2^2\sigma_2^2 + ... + n_E^2\sigma_E^2) \geq \frac{1}{E}(\sigma_1^2 + \sigma_2^2 + ... + \sigma_E^2)$$
$$s.t. \sum_{e=1}^{E} n_e = 1. \tag{4}$$

Based on Eq. (4), the minimum variance can be achieved using the averaging method. However, excessively small variance can lead to model overfitting, thereby impairing the model's ability to generalize to unseen domains. When the source and target distributions are identical, FL methods significantly improve model performance. However, these methods perform poorly on the FedOSDG problem as they overfit to **domain-specific knowledge** (e.g., styles) from the source domain, resulting in limited generalization. To address this challenge, we propose the AAW, which shifts the model toward a more distant solution to increase model variance. This mitigates overfitting caused by domain-specific knowledge and enhances the model's ability to generalize to unknown domains. The details of our AAW method are described as follows:

$$W_{i+1} = \sum_{e=1}^{E} \frac{ln(1 + ||\bar{W} - W_{i+1}^e||^2)N^e}{\sum_{e=1}^{E} ln(1 + ||\bar{W} - W_{i+1}^e||^2)N^e} W_{i+1}^e$$
$$s.t. \bar{W} = \sum_{e=1}^{E} \alpha \frac{N^e}{\sum_{e=1}^{E} N^e} W_{i+1}^e \tag{5}$$

where W_{i+1}^e denotes the model parameters of the $i+1$th iteration within domain e and E denotes the number of source domains involved in training. N^e represents the quantity of data in domain e, while α is a hyperparameter that account for the $\frac{M}{m}$. ln introduces non-linearity in the aggregation. Given the typically large scale of model parameters, using ln helps reduce their dominance during aggregation.

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on March 22,2026 at 06:50:51 UTC from IEEE Xplore. Restrictions apply.

TABLE II

RESULTS (%) OF PACS ON RESNET50. THE RATIO OF KNOWN TO UNKNOWN CLASSES IS 6:1. (BEST IN BOLD)

Method	Photo			Art			Cartoon			Sketch			Avg		
	UNK	ALL	H-score	UNK	ALL	H-score	UNK	ALL	H-score	UNK	ALL	H-score	UNK	ALL	H-score
FedAvg	28.01	83.34	43.01	16.72	68.21	27.76	16.79	72.90	27.89	21.25	68.30	33.23	20.69	73.19	32.97
Fedprox	63.19	83.87	73.38	28.06	66.94	40.60	15.06	63.81	24.91	28.75	69.25	41.72	33.76	70.96	45.15
Mixstyle	58.52	84.24	70.46	27.52	66.43	40.12	18.02	65.87	28.98	25.00	71.56	38.02	32.26	72.02	44.40
AM	57.58	86.18	70.73	25.39	72.95	38.65	21.73	68.35	33.81	29.38	65.44	41.63	33.52	73.22	46.20
RSC	58.54	82.94	69.98	29.84	67.20	42.44	17.53	64.39	28.21	16.88	67.20	27.59	30.69	70.42	42.05
Scaffold	48.15	83.10	62.47	22.49	69.14	34.81	16.05	65.01	26.33	26.25	67.23	38.62	28.23	70.90	40.55
FedDG-GA	48.17	86.15	63.35	19.38	71.49	31.21	15.56	66.23	25.75	19.38	69.81	31.06	25.62	73.42	37.84
FedPD	60.88	88.04	73.45	29.84	72.71	43.45	20.74	**73.82**	33.16	25.00	67.11	37.39	34.11	75.41	46.86
AABAW (ours)	**66.44**	**89.39**	**77.58**	**30.29**	**75.68**	**44.19**	**25.00**	70.07	**37.81**	**31.87**	**70.91**	**45.16**	**38.40**	**76.51**	**51.19**

TABLE III

RESULTS (%) OF OFFICE-HOME ON RESNET18. THE RATIO OF KNOWN TO UNKNOWN CLASSES IS 30:35. (BEST IN **BOLD**)

Method	Product			Art			Clipart			Real World			Avg		
	UNK	ALL	H-score	UNK	ALL	H-score	UNK	ALL	H-score	UNK	ALL	H-scoe	UNK	ALL	H-score
FedAvg	38.38	57.07	51.64	45.23	53.76	51.75	39.46	49.17	47.77	35.58	58.06	50.04	39.66	54.51	50.30
FedProx	34.78	55.01	48.23	40.70	51.88	50.03	38.56	48.46	46.95	34.96	57.20	49.23	37.25	53.14	48.61
MixStyle	37.79	55.28	50.41	42.13	50.79	49.81	39.07	49.49	47.83	40.02	59.06	53.63	39.75	53.65	50.42
AM	41.46	56.39	53.10	45.71	52.01	51.65	41.16	50.47	49.56	39.44	58.00	52.76	41.94	54.21	51.76
RSC	37.84	55.91	50.73	43.08	52.88	51.61	37.93	48.63	46.81	34.96	57.17	49.22	38.45	53.64	49.59
Scaffold	36.84	55.74	50.00	43.41	50.98	50.31	42.73	47.51	47.35	36.63	57.99	50.82	39.90	53.05	49.62
FedDG-GA	32.92	54.24	46.50	40.90	50.97	49.18	39.81	47.09	46.40	33.77	56.92	48.16	38.90	52.30	49.47
FedPD	30.09	57.28	44.57	38.03	53.57	49.70	34.79	47.07	44.42	30.44	57.43	44.97	38.45	53.83	49.59
AABAW (ours)	**42.23**	**58.33**	**54.66**	**49.84**	**55.23**	**55.06**	**44.39**	**52.29**	**51.36**	**44.41**	**62.39**	**57.95**	**45.22**	**57.06**	**54.76**

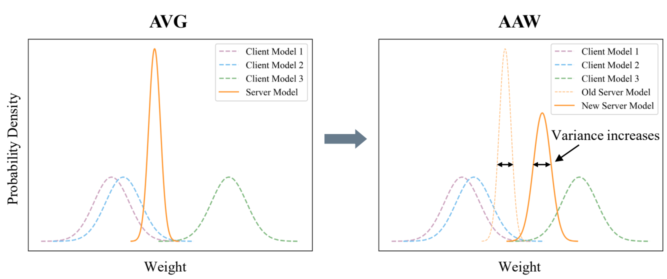

Fig. 4. Taking the Gaussian model as an example, the AAW module shifts the aggregation model in a direction where the weights are far away from AVG. In the process of shifting, it leads to moderate variance of the model and improves the generalization ability of the model.

In FedOSDG, a significant discrepancy between the domain client model w_i^e and the global model w_i often results in poor performance of the global model on domain client e. By adjusting the model closer to e, performance within domain e is enhanced, albeit with a decrease across other domains. The same principle applies to AAW: it sacrifices recognition accuracy in known domains to achieve better recognition in unknown ones. As a result, traditional FL methods can still be applied to known clients, while AAW is employed for unknown clients to enhance the generalization of the global model. From a variance perspective, increasing model variance enhances overall generalization across clients in unseen domains. However, excessively high variance can impede the model's convergence. Below, we present the conclusions regarding the convergence of AAW.

Theorem 1: Utilizing common assumptions, we arrive at the anticipated reduction in the global loss function across two successive rounds adhering to the following relationship:

$$\mathcal{L}(w_{i+1}) \leq \mathcal{L}(w_i) - \eta[1 - \frac{\beta}{2}\eta]min_{e\in E}\{||\nabla\mathcal{L}_e(w_i)||^2\} \quad (6)$$

Typically, the learning rate η is set to a quite low value to avoid the model from drifting the optimal solution during the gradient descent process. Observation of Eq.(6) underlines that, with an adequately minimal learning rate, the loss function of the global model demonstrates a consistent decline. Attributed to the loss function $\mathcal{L}_e$ for each domain client e remains above 0 and the global model being linear weighting of domain client model, this implies that the loss function of the global model $\mathcal{L}$ possesses a minimal boundary of 0. It affirms the convergence of the AAW method. The upper bound of convergence is obtained through the accumulation method:

$$\begin{gathered}\mathcal{L}(w_{R+n}) \leq \mathcal{L}(w_0) - \eta[1 - \frac{\beta}{2}\eta]A + B \\ s.t. A = \sum_{i=1}^{R} min_{e\in E}\{||\nabla\mathcal{L}_e(w_i)||^2\}\end{gathered} \quad (7)$$

where the model gradient $\nabla\mathcal{L}_e(w_i)$ for each domain e approaches 0 after the R-th iteration. Consequently, we introduce a constant B to denote the upper of the loss function limit over the subsequent n rounds, while A signifies the cumulative gradient over the initial R rounds. Our analysis confirms its convergence and its applicability across part of linear weighting schemes in adaptive weight FL, encompassing FedAvg and so on. While AAW does not guarantee superior performance in every unknown domain, Fig. 4 demonstrates that AAW has a statistically significant advantage. Empirical evidence suggests

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on March 22,2026 at 06:50:51 UTC from IEEE Xplore. Restrictions apply.

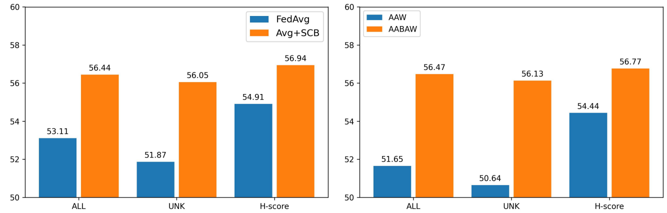

Fig. 5. Comparison average of the performance of FedAvg and FedAvg+SCB, FedAvg+AAW and AABAW on PACS using ResNet18.

promising performance in extensive evaluations across many unseen domains. Compared to FedAvg, AAW increases the generalization of the model by increasing the variance, which prevents overfitting and thus overcomes **domain shift**.

III. Experiments

A. Setup

In this section, we introduce the experimental setup. Without an explicit statement, we conduct all experiments under the following conditions. The default settings are as follows:

Dataset. We evaluate our method on two FedDG benchmarks. (i) The **PACS** [14] dataset comprises four domains (*photo, art-painting, cartoon and sketch*), with 9,991 images in total. (ii) **Office-Home** [15] encompasses approximately 15,500 images across 65 classes from four domains (*art, clipart, product and real-world*). Additionally, we perform a FedOSDG task on DomainNet-126 [16], a subset of the DomainNet dataset. DomainNet-126 features six diverse domains with roughly 600,000 images across 345 classes.

Comparing Methods. Given the lack of methods directly tailored to our problem setting, we have refined existing FL and DG approaches, adapting them to align with the FedOSDG paradigm. We compare our proposed AABAW with existing FL methods such as FedAvg [1], FedProx [17], MixStyle [18], AM [19], and RSC [20], with the addition of a threshold μ. Additionally, we compare AABAW against state-of-the-art FedDG and FedOSR methods, including FedDG-GA [21] for FedDG and FedPD [22] for FedOSR.

Evaluation Metrics. To effectively evaluate AABAW, we employ the following three metrics for comparison: (i) **ALL** [23], the accuracy of all instances. (ii) **UNK** [24], indicating open-set accuracy, measure the model in correctly labeling unknown classes as "unknown". (iii) **H-score** [25], the harmonic mean of accuracy in known classes and UNK is widely accepted in current open-set recognition studies, to evaluate the performance in handling both known and unknown classes.

B. Results

In our study, we evaluate AABAW within the FedOSDG paradigm. The results for the PACS and Office-Home datasets are presented in Tab. II and Tab. III, respectively. While traditional FL and FedOSR methods demonstrate strong performance on source domains, their efficacy diminishes when faced with unknown domains. For instance, as shown in Tab. II, AABAW achieves a 1.10% improvement in overall accuracy on the PACS dataset compared to FedPD. Furthermore, it delivers a 4.29% increase in identifying unknown data and improves the H-score by 4.33%. These results highlight that AABAW is better suited to address the challenges of FedOSDG compared to methods from FL, FedDG, and FedOSR.

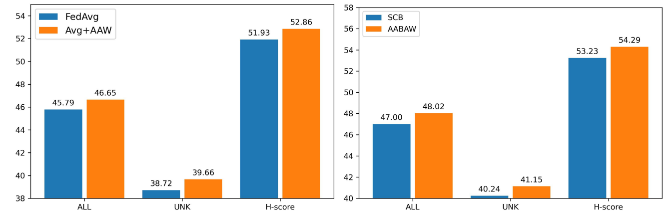

Fig. 6. Comparison of the average performance of FedAvg and FedAvg+AAW, FedAvg+SCB and AABAW on PACS using ResNet 18.

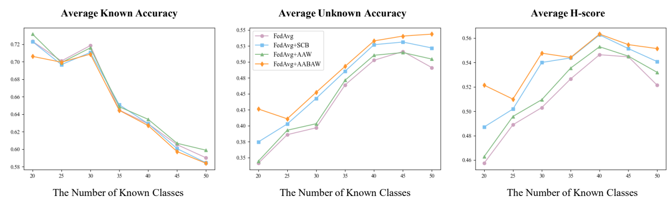

Fig. 7. The average performance metrics on the Office-Home dataset across varying numbers of known classes using ResNet18.

C. Ablation Studies

We conduct ablation studies to evaluate the impact of the SCB module, as illustrated in Fig. 5. The SCB module enhances the average UNK by 3.18% and 5.49%. Regarding the average ALL, AABAW achieves improvements of 3.33% and 4.78%. Additionally, the average H-score increases by 2.03% and 2.33%, highlighting the SCB module's effectiveness in improving recognition performance, particularly for unknown classes. Our experiments also demonstrate that a moderate increase in variance improves the model's generalization ability. As shown in Fig. 6, the AAW module contributes to consistent performance gains: ALL increases by 0.86% and 1.02%, UNK by 0.94% and 0.91%, and the H-score by 0.93% and 1.06%. These results confirm the complementary roles of the SCB and AAW modules in enhancing the overall performance.

Changing the threshold μ for identifying unknown classes. Raising the rejection threshold reduces closed-set recognition accuracy, as higher thresholds cause some previously recognized instances to fall below the threshold, incorrectly categorizing them as unknown, as shown in Tab. IV. However, this adjustment improves open-set recognition and increases the overall H-score, confirming the SCB method's efficacy in enhancing predictions by reinforcing classification boundaries. Thus, increasing the rejection threshold generally benefits model performance. Moreover, our experiments reveal that the optimal threshold varies across domains, indicating that dynamic threshold adjustments could further enhance the performance of FedOSDG.

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on March 22,2026 at 06:50:51 UTC from IEEE Xplore. Restrictions apply.

TABLE IV
AVERAGE PERFORMANCE METRICS WITH DIFFERENT THRESHOLDS ON OFFICE-HOME AND PACS USING AABAW(WITH RESNET18).

Threshold	OfficeHome			PACS		
	UNK	ALL	H-score	UNK	ALL	H-score
0.5	31.20	**71.86**	42.86	47.53	**60.30**	52.79
0.6	45.22	70.87	54.76	56.05	58.54	56.94
0.7	**51.71**	66.99	**57.36**	**64.50**	56.11	**59.72**

Changing the ratio of known classes. As shown in Fig. 7, increasing the number of known classes causes the recognition performance of AABAW to decline due to diluted feature representation arising from greater class diversity. However, its ability to identify unknown classes improves significantly.

IV. CONCLUSION

In this paper, we tackle the practical problem of Federated Open Set Domain Generalization (FedOSDG), which involves both **category shift** and **domain shift** in target domains while safeguarding the privacy of source domain data. To address these challenges, we propose a novel method, Adaptive Adjustment Boundaries and Weights (AABAW), comprising two main modules: Stronger Classification Boundary (SCB) and Adaptive Adjustment of Weights (AAW). The SCB module leverages multiple binary classifiers to enhance the recognition of unknown classes, effectively addressing **category shift**. The AAW module improves model generalization by increasing model variance to mitigate **domain shift**. The experiment results show that our method achieves state-of-the-art performance on metrics such as ALL, UNK, and H-score, validating its effectiveness. We posit that the FedOSDG framework will advance the development of computer vision in privacy protection and user promotion.

REFERENCES

[1] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*. PMLR, 2017, pp. 1273–1282.

[2] Shilei Cao, Yan Liu, Juepeng Zheng, Weijia Li, Runmin Dong, and Haohuan Fu, "Exploring test-time adaptation for object detection in continually changing environments," *arXiv preprint arXiv:2406.16439*, 2024.

[3] Qingmei Li, Yibin Wen, Juepeng Zheng, Yuxiang Zhang, and Haohuan Fu, "Hyunida: Breaking label set constraints for universal domain adaptation in cross-scene hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, 2024.

[4] Jiaxu Guo, Yushan Lai, Jinxiao Zhang, Juepeng Zheng, Haohuan Fu, Lin Gan, Liang Hu, Gaochao Xu, and Xilong Che, "C 3 da: A universal domain adaptation method for scene classification from remote sensing imagery," *IEEE Geoscience and Remote Sensing Letters*, 2024.

[5] Juepeng Zheng, Wenzhao Wu, Shuai Yuan, Haohuan Fu, Weijia Li, and Le Yu, "Multisource-domain generalization-based oil palm tree detection using very-high-resolution (vhr) satellite images," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2021.

[6] Wentang Chen, Yibin Wen, Juepeng Zheng, Jianxi Huang, and Haohuan Fu, "Ban: A universal paradigm for cross-scene classification under noisy annotations from rgb and hyperspectral remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, 2025.

[7] Khiem Le, Long Ho, Cuong Do, Danh Le-Phuoc, and Kok-Seng Wong, "Efficiently assemble normalization layers and regularization for federated domain generalization," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 6027–6036.

[8] Juepeng Zheng, Yibin Wen, Mengxuan Chen, Shuai Yuan, Weijia Li, Yi Zhao, Wenzhao Wu, Lixian Zhang, Runmin Dong, and Haohuan Fu, "Open-set domain adaptation for scene classification using multi-adversarial learning," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 208, pp. 245–260, 2024.

[9] Chen Yang, Meilu Zhu, Yifan Liu, and Yixuan Yuan, "Fedpd: Federated open set recognition with parameter disentanglement," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 4882–4891.

[10] Ruijia Xu, Ziliang Chen, Wangmeng Zuo, Junjie Yan, and Liang Lin, "Deep cocktail network: Multi-source unsupervised domain adaptation with category shift," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 3964–3973.

[11] Yu Liang, Shilei Cao, Juepeng Zheng, Xiucheng Zhang, Jianxi Huang, and Haohuan Fu, "Low saturation confidence distribution-based test-time adaptation for cross-domain remote sensing image classification," *International Journal of Applied Earth Observation and Geoinformation*, vol. 139, pp. 104463, 2025.

[12] Xiran Wang, Jian Zhang, Lei Qi, and Yinghuan Shi, "Generalizable decision boundaries: Dualistic meta-learning for open set domain generalization," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 11564–11573.

[13] Hongda Wu and Ping Wang, "Fast-convergent federated learning with adaptive weighting," *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 4, pp. 1078–1088, 2021.

[14] 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*, 2017, pp. 5542–5550.

[15] 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*, 2017, pp. 5018–5027.

[16] 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*, 2019, pp. 1406–1415.

[17] Tian Li, Anit Kumar Sahu, Manzil Zaheer, Maziar Sanjabi, Ameet Talwalkar, and Virginia Smith, "Federated optimization in heterogeneous networks," *Proceedings of Machine learning and systems*, vol. 2, pp. 429–450, 2020.

[18] Kaiyang Zhou, Yongxin Yang, Yu Qiao, and Tao Xiang, "Domain generalization with mixstyle," *arXiv preprint arXiv:2104.02008*, 2021.

[19] Zhe Qu, Xingyu Li, Rui Duan, Yao Liu, Bo Tang, and Zhuo Lu, "Generalized federated learning via sharpness aware minimization," in *International conference on machine learning*. PMLR, 2022, pp. 18250–18280.

[20] Zeyi Huang, Haohan Wang, Eric P Xing, and Dong Huang, "Self-challenging improves cross-domain generalization," in *Computer vision–ECCV 2020: 16th European conference, Glasgow, UK, August 23–28, 2020, proceedings, part II 16*. Springer, 2020, pp. 124–140.

[21] Ruipeng Zhang, Qinwei Xu, Jiangchao Yao, Ya Zhang, Qi Tian, and Yanfeng Wang, "Federated domain generalization with generalization adjustment," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 3954–3963.

[22] Chen Yang, Meilu Zhu, Yifan Liu, and Yixuan Yuan, "Fedpd: Federated open set recognition with parameter disentanglement," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, October 2023, pp. 4882–4891.

[23] Yadan Luo, Zijian Wang, Zi Huang, and Mahsa Baktashmotlagh, "Progressive graph learning for open-set domain adaptation," in *Proceedings of the 37th International Conference on Machine Learning*, Hal Daumé III and Aarti Singh, Eds. 13–18 Jul 2020, vol. 119 of *Proceedings of Machine Learning Research*, pp. 6468–6478, PMLR.

[24] Hong Liu, Zhangjie Cao, Mingsheng Long, Jianmin Wang, and Qiang Yang, "Separate to adapt: Open set domain adaptation via progressive separation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[25] Bo Fu, Zhangjie Cao, Mingsheng Long, and Jianmin Wang, "Learning to detect open classes for universal domain adaptation," in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XV 16*. Springer, 2020, pp. 567–583.

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on March 22,2026 at 06:50:51 UTC from IEEE Xplore. Restrictions apply.