

# Toward Few-Shot Learning in the Open World: A Review and Beyond

Hui Xue<sup>1</sup>, *Member, IEEE*, Yuexuan An<sup>2</sup>, *Member, IEEE*, Yongchun Qin, Wenqian Li<sup>3</sup>, Yixin Wu<sup>4</sup>,  
Yongjuan Che<sup>5</sup>, Pengfei Fang<sup>6</sup>, and Min-Ling Zhang<sup>7</sup>, *Senior Member, IEEE*

(Survey Paper)

## I. INTRODUCTION

**Abstract**—Human intelligence is characterized by our ability to absorb and apply knowledge from the world around us, especially in rapidly acquiring new concepts from minimal examples, underpinned by prior knowledge. Few-shot learning (FSL) aims to mimic this capacity by enabling significant generalizations and transferability. However, traditional FSL frameworks often rely on assumptions of clean, complete, and static data, conditions that are seldom met in real-world environments. Such assumptions falter in the inherently uncertain, incomplete, and dynamic contexts of the open world. This paper presents a comprehensive review of recent advancements designed to adapt FSL to open-world environments. We categorize existing methods into three distinct types of FSL in the open world: those involving varying instances, varying classes, and varying distributions. Each category is discussed in terms of its specific challenges and methods, as well as its strengths and weaknesses. We standardize experimental settings and metric benchmarks across scenarios and provide a comparative analysis of the performance of various methods. In conclusion, we outline potential future research directions for this evolving field. It is our hope that this review will catalyze further development of effective solutions to these complex challenges, thereby advancing the field of artificial intelligence.

**Index Terms**—Few-Shot learning, few-shot learning with varying classes, few-shot learning with varying distributions, few-shot learning with varying instances, open world.

Received 31 July 2024; revised 15 July 2025; accepted 25 July 2025. Date of publication 1 August 2025; date of current version 3 October 2025. This work was supported in part by the National Natural Science Foundation of China under Grant 62476056, Grant 62076062, Grant 62306070, and Grant 62225602, in part by the Social Development Science and Technology Project of Jiangsu Province under Grant BE2022811, and in part by the Fundamental Research Funds for the Central Universities under Grant B250201040 and Grant aiaa-25-03. Furthermore, the work was also supported by the Big Data Computing Center of Southeast University. Recommended for acceptance by H. Shim. (Corresponding author: Hui Xue.)

Hui Xue, Yongchun Qin, Wenqian Li, Yixin Wu, Yongjuan Che, and Pengfei Fang are with the School of Computer Science and Engineering, Southeast University, Nanjing 210096, China, and also with the Key Laboratory of New Generation Artificial Intelligence Technology and Its Interdisciplinary Applications, Ministry of Education, Southeast University, Nanjing 211189, China (e-mail: hxue@seu.edu.cn; ycqin@seu.edu.cn; wenqian.li@seu.edu.cn; wuyx@seu.edu.cn; yjche@seu.edu.cn; fangpengfei@seu.edu.cn).

Yuexuan An is with the School of Computer Science and Engineering, Southeast University, Nanjing 210096, China, and also with the College of Computer Science and Software Engineering, Hohai University, Nanjing 210098, China (e-mail: anyx@hhu.edu.cn).

Min-Ling Zhang is with the School of Computer Science and Engineering, Southeast University, Nanjing 210096, China, and also with the Key Laboratory of Computer Network and Information Integration, Ministry of Education, Southeast University, Nanjing 211189, China (e-mail: zhangml@seu.edu.cn).

Digital Object Identifier 10.1109/TPAMI.2025.3594686

0162-8828 © 2025 IEEE. All rights reserved, including rights for text and data mining, and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See <https://www.ieee.org/publications/rights/index.html> for more information.

IN RECENT years, the explosive growth of annotated data coupled with the rapid advancements in device computing capabilities has propelled deep learning to exhibit remarkable resilience, significantly advancing the field of artificial intelligence. However, these achievements hinge on the availability of vast amounts of labeled training data to construct predictive models endowed with robust generalization capabilities. Nevertheless, in many practical application scenarios, limitations such as manpower, resources, and environmental constraints often hinder the acquisition of a sufficient number of labeled training samples. This limitation severely compromises the effectiveness of machine learning models in scenarios with scarce labeled data. In contrast, human intelligence demonstrates the ability to swiftly establish cognition of new concepts with minimal learning samples, often requiring only one or a few instances [1].

To simulate this powerful learning ability of humans, Few-Shot Learning (FSL) has emerged as an exceptional machine learning paradigm, which learns to recognize new concepts from just a few examples [2], [3], [4]. FSL excels in scenarios without relying on very expensive and time-consuming data collection [5]. Traditional FSL algorithms often operate under the strict assumption that all data observations must be described by a fixed, complete, and clean data space. This assumption necessitates uniformity across instances, classes, and distributions.

In various practical worlds, more and more challenging tasks beyond the setting are present in the FSL community. To wit, consider medical image analysis aided by FSL, in which the collection of scarce medical data is a continuous and laborious process involving contributions from clinical experts from different regions. In this circumstance, challenges abound wrong annotations of ambiguous images serving as examples, unobserved disease types to be predicted, constantly emerging new subspecies of diseases necessitating model updates, and different sources of images for training and testing [6], [7], [8], [9]. Instead of being fixed and known as advanced, the data space used to describe the FSL data is varying and dynamic. Therefore, it is imperative to advance FSL research to equip models with the ability to adapt to the complexities of the open world and to learn in a way that mimics human cognition.

Several existing surveys have made detailed summaries and prospects for FSL and open-world machine learning. In studies [3], [10], [11], authors delve into the significant challenges of

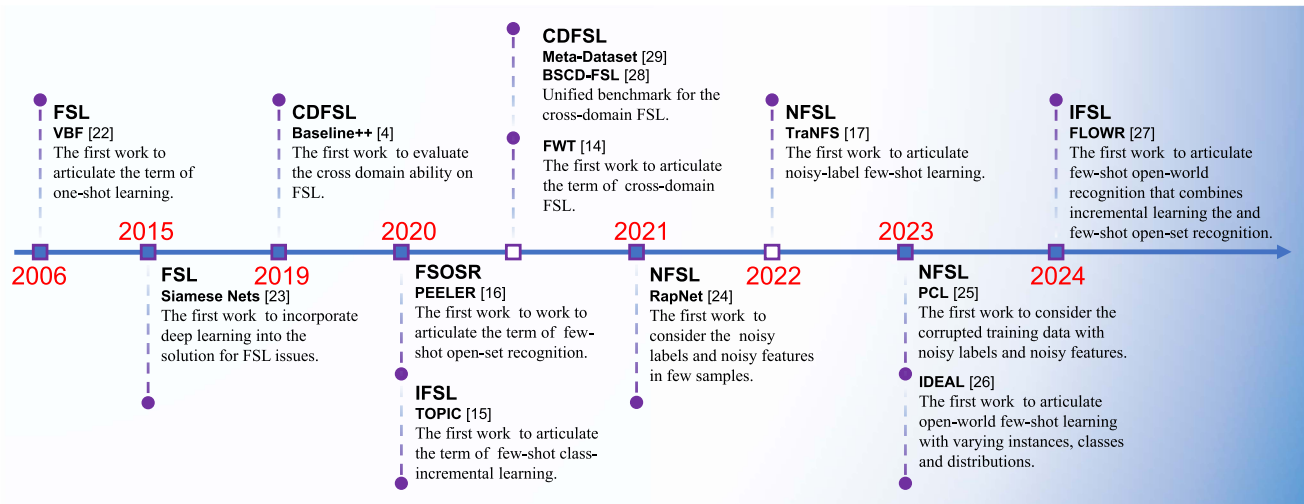


Fig. 1. Chronological milestones on FSL in the open world, including representative FSL approaches and the related benchmarks. FSL was first noticed as a topic in 2003 [22]. Subsequently, it underwent continuous developments [4], [23]. Successively, different open-world scenarios of FSL have been proposed, including FSL with varying instances [16], [17], [24], [25], [26], FSL with varying classes [15], [27] and FSL with varying distributions [14], [28], [29].

FSL and review numerous FSL methodologies within closed-world contexts. Beyond the fixed settings of the closed world, recent attention has also turned toward open-world scenarios. In the study [12], researchers refer to open-world machine learning as cumulative learning and open-world recognition, and compare the relative methods in computer vision and natural language processing. In the study [13], researchers discuss the overall challenges of open-world machine learning and emphasize techniques and relations among unknown rejection, novel class discovery, and class-incremental learning. Recently, we have witnessed a surge of FSL studies that aim to address varying and dynamic scenarios in the open world, including cross-domain FSL (CDFSL)[14], incremental FSL (IFSL)[15], few-shot open-set recognition (FSOSR)[16], noisy FSL (NFSL)[17], and more, as shown in Fig. 1. Moreover, some work also attempts to systematically review FSL in some open scenarios, such as surveys about IFSL and CDFSL [18], [19], [20], [21]. However, existing studies on one specific problem have been divided into multiple communities. Each community possesses its own set of scenario assumptions and design intuitions, leading to disparate and confusing evaluations of model performances. This has made communication increasingly difficult and complex, resulting in a lack of systematic analysis and reviews available for FSL in the open world.

To address this gap and foster future research while providing newcomers with a comprehensive understanding of this challenging problem, this paper presents, for the first time, a comprehensive summary and review of FSL in the open world. Driven by the current state of fragmentation, we present a comprehensive review of prior work and, more importantly, aim to unify and contextualize these efforts within a cohesive framework, which we term Few-Shot Learning in the Open World (OWFSL). We relax the traditional FSL assumption of a clean, complete, static data space to an incomplete and dynamic nature of the data space encountered in the open world. We taxonomize existing OWFSL

into three main categories, including Few-Shot Learning with Varying Instances, Few-Shot Learning with Varying Classes and Few-Shot Learning with Varying Distribution. Each scenario covers various aspects ranging from their related settings and methods, strengths and weaknesses, dataset, evaluation metrics and performance comparisons of representative methods. Furthermore, we describe their relatively few-shot scenarios and highlight this emerging topic's challenges and potential future directions.

#### Specific contributions of our survey are as follows:

- We present a timely and the first survey of recent FSL studies that transitioned from closed environments to open-world environments. By framing and analyzing the prior art, our review provides a unified perspective that can bridge the gap between fragmented research communities and foster communication.
- We introduce a novel taxonomy that classifies FSL in the open world into three scenarios, including varying instances, varying classes and varying distribution. This taxonomy provides a useful framework for understanding the current state of the field and identifying areas for further research.
- We conduct an in-depth illustration of the representative methods in each scenario, analyzing their strengths and weaknesses, the common datasets and evaluation metrics, and comparing the performances of related methods. Our analysis provides valuable insights into the effectiveness of different approaches and the challenges involved in each scenario.
- We highlight the open challenges faced by current FSL studies and strive to illuminate unexplored avenues for future research. Our goal is to stimulate further exploration and discovery in this field to help researchers and practitioners navigate the complexities of FSL in open-world scenarios.

## II. PROBLEM DEFINITION

In FSL, the goal is to build models that generalize effectively from a limited number of labeled examples. Formally, given a  $t$ th  $N$ -way  $K$ -shot task  $\mathcal{T}^t$ , we have a dataset  $\mathcal{D}^t = \{\mathcal{S}^t, \mathcal{Q}^t\}$  including a support dataset  $\mathcal{S}^t = \{(\mathbf{x}_i^t, y_i^t)\}_{i=1}^{N \times K}$  where label  $y_i$  belongs to the label set  $\mathcal{C}^t$  and a query dataset  $\mathcal{Q}^t = \{(\hat{\mathbf{x}}_i^t, \hat{y}_i^t)\}_{i=1}^{M_t}$  where label  $\hat{y}_i^t$  is unknown. Meanwhile,  $\mathcal{S}^t$  and  $\mathcal{Q}^t$  is disjoint, i.e.,  $\mathcal{S}^t \cap \mathcal{Q}^t = \emptyset$ . The FSL task aims to predict the label  $\hat{y}_i^t$  of a query  $\hat{\mathbf{x}}_i^t$  with a limited support dataset  $\mathcal{S}^t$ . FSL contains different FSL tasks from the distribution  $\mathcal{P}_n$  and labels of support datasets of all tasks are from the novel label set  $\mathcal{C}_n$ , i.e.,  $\mathcal{C}^t \subseteq \mathcal{C}_n, t = 1, \dots, T$ . Meanwhile, there usually exists an auxiliary base class datasets  $\mathcal{D}_b = \{(\mathbf{x}_i^b, y_i^b)\}_{i=1}^I$  drawn from a distribution  $\mathcal{P}_b$  with abundant labeled examples belonging to a base label set  $\mathcal{C}_b$ , where  $\mathcal{C}_b \cap \mathcal{C}_n = \emptyset$ . To simulate few-shot conditions, the meta-learning framework divides  $\mathcal{D}_b$  into a series of pseudo-FSL tasks (episodes), each with its own support and query sets. A model learns and updates parameters by optimizing its performance on these pseudo-tasks.

Traditional FSL assumes an ideal scenario in which the dataset is clean, complete, static data space, with fixed classes and consistent distributions. In fact, noisy annotations, emerging novel classes, incrementally expanding classes and distribution shifts between training and testing data are common in the open world.

In contrast, **Few-Shot Learning in the Open World (OWFSL)** relaxes these strict conditions, considering at least one of the following varying conditions:

- *Varying Instances*: Training samples vary with the presence of noise and testing samples vary with the emergence of unknown classes:
  - 1) *NFSL*: some labels of training samples are noisy,  $\hat{y}_i^t \neq \text{GroundTruth}(\hat{\mathbf{x}}_i^t)$  and  $y_i^b \neq \text{GroundTruth}(\mathbf{x}_i^b)$
  - 2) *FSOSR*: some labels of query samples are unknown,  $\hat{y}_i^t \notin \mathcal{C}^t$ .
- *Varying Classes (IFSL)*: Class set expands incrementally: recognize  $\mathcal{C}^1 \cup \dots \cup \mathcal{C}^t$  for  $t$ th task.
- *Varying Distributions (CDFSL)*: Significant domain shift exists between training and novel data:  $\mathcal{P}_n \neq \mathcal{P}_b$ .

## III. FEW-SHOT LEARNING WITH VARYING INSTANCES: NOISY FEW-SHOT LEARNING

In the open world, due to limited knowledge and unintentional impairments, FSL systems often face the challenge of training samples varying with the presence of noise, which is known as noisy few-shot learning (NFSL). For example, in medical image analysis aided by FSL, scarce medical data annotated by clinical experts from different regions frequently contain labeling errors resulting from inherent ambiguity, inter-observer variability, or annotator fatigue. These noisy labels adversely affect the accuracy of diagnostic models trained with limited examples, posing significant challenges for developing reliable clinical decision-support systems.

*Definition*: NFSL considers the scenario that the training samples are no longer unreliable and some samples may be polluted by noise. The noisy samples may exist in  $\mathcal{D}_b$  or  $\mathcal{S}^t$ .

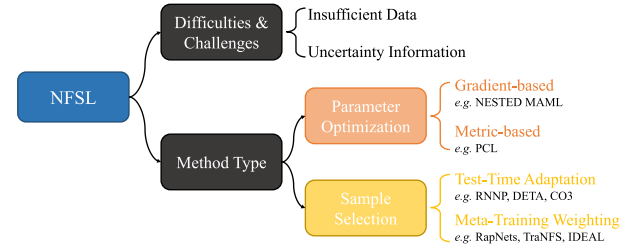


Fig. 2. The structural relationships of the existing two types of methods and their subclasses, as well as the difficulties and challenges faced by NFSL problems.

The noise commonly comes from label space, i.e.,  $\hat{y}_i^t$  and  $y_i^b$  are not the ground truth label of  $\hat{\mathbf{x}}_i^t$  and  $\mathbf{x}_i^b$ .

NFSL requires mitigating the degradation of model performance when training with noisy samples. The key challenges of NFSL come from two aspects: 1) insufficient data: due to the rarity of samples, the potential noise is hard to detect with insufficient data, and 2) uncertainty information: each sample might contain important category information, incorrect recognition of noise and training can lead to devastating uncertainty on model training. Therefore, FSL models should eliminate the uncertainty and disturbance from noisy samples. To address this issue, several works have been devoted to adapting to the noisy environment in FSL. We categorize existing methods into two classes: parameter optimization and sample selection, as shown in Fig. 3. The methods based on parameter optimization are devoted to optimizing the model by assigning proper weights to the losses of the tasks or the instances. Common approaches include gradient-based weighting (e.g. NESTED MAML [30], Eigen-Reptile [31]) and metric-based weighting (e.g. PCL [25]). The methods based on sample selection aim to assign lower weight to the sample that is considered as noise. Common approaches include test-time adaptation (e.g. RNNP [32], DETA [33], CO3 [34]) and meta-training weighting (e.g. RapNets [24], TraNFS [17], IDEAL [26]).

In this section, we will introduce the latest progress of NFSL from two aspects. The overview of the problem definition and method strategy is depicted in Fig. 2.

### A. Parameter Optimization

Parameter optimization learns the optimal weights of the losses or the tasks and the parameter update direction to improve the model performance under noisy scenarios. According to the way of parameter update, the parameter optimization can further be divided into gradient-based methods and metric-based methods.

Gradient-based methods, as depicted in Fig. 2(b), fine-tune model parameters on a few support examples. NESTED MAML [30] modifies MAML by assigning weights to the losses of both tasks and solving the weights in a nested bi-level optimization. It studies the general form of task and instance-weighted meta-learning, where the optimal weights and model initialization parameters are learned by optimizing a nested bi-level objective function. Eigen-Reptile [31] constructs prior

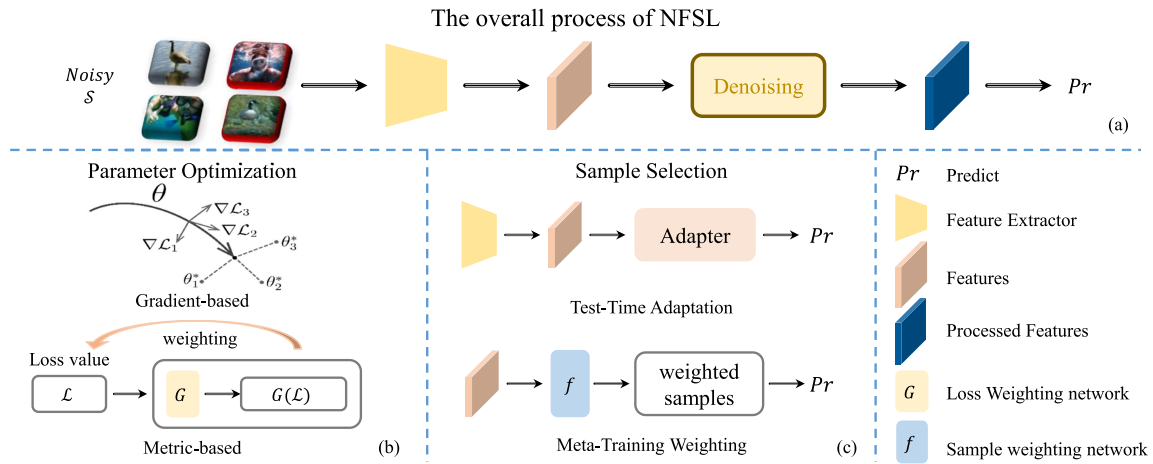


Fig. 3. An overall diagram of existing NFSL research. (a) Overview of the NFSL setting. (b) Parameter optimization learns the optimal weights of tasks and the parameter update direction by Gradient-based and Metric-based methods. (c) Sample Selection Learn the weights of few-shot examples from noisy datasets, which are mainly divided into Test-Time Adaptation and Meta-Training Weighting.

models to decide which sample should be discarded with self-paced learning to learn a more accurate main direction. The main direction is computed efficiently, with the scale of the calculated matrix related to the number of gradient steps rather than the number of parameters. DKT [35] presents a Bayesian treatment for the meta-learning inner loop through the use of deep kernels, which can learn a kernel that transfers to new tasks. It is straightforward to implement as a single optimizer, provides uncertainty quantification, and does not require estimation of task-specific parameters.

Metric-based methods classify samples based on their similarities. PCL [25] adopts a collaborative mechanism to reweight the loss according to the global ranks of losses and local intra-class correlation information. This mechanism extracts the sample importance information from both global and local aspects by encoding the feature representations and the corresponding loss values of the training samples.

*Summary and Discussion:* Parameter optimization can be applied to various models, including deep learning models and traditional machine learning models, without additional data preprocessing. However, these methods depend too much on the choice of weighting strategy that inappropriate strategy may lead to performance degradation. Besides, they are not sensitive to hard samples and outliers which could be assigned higher weights, adversely affecting model training.

## B. Sample Selection

Sample Selection, as depicted in Fig. 2(c), learns the weights of the few-shot examples to alleviate the impact of the noisy samples. According to the way of denoising, the sample selection can further be divided into test-time adaptation methods and meta-training weighting methods.

Test-time adaptation methods mainly focus on dealing with noise in base classes, rather than in the few-shot task. RNNP [32] adopts the Mixup [36] strategy to synthetic samples with samples with the same label and cluster all features to generate new classes. It generates mixed features of samples and

adopts k-means to produce refined prototypes. DETA [33] is a meta-training-free method that focuses on filtering out task-irrelevant, noisy representations. It crops each image into different regions to extract more relevance in images and weights each region according to the cosine similarity of paired regions to filter image noise and label noise. CO3 [34] effectively leverages the prior knowledge of foundation models to enhance model performance. It introduces TeFu-Adapter to reduce the negative impact of noisy labels during the adapter updating stage.

Meta-training weight methods mainly focus on dealing with the noise in the current few-shot task. Dr.k-NN [37] modifies the classifier with a weighted k-NN classifier by weighting k neighbor samples. It first transforms the weighted k-NN problems to finite-dimensional convex programming. Then, it solves the convex optimization by jointly learning the feature embedding and the minimax optimal classifier. RapNets [24] regenerates sample features in each class according to the similarity between the paired samples within a class and encodes sample features to BiLSTM to achieve the weight of each sample. It proposed an attentive module to learn the weight of samples in the same class. TraNFS [17] achieves the weight of each sample according to the attention value in transformer by concatenating samples as a sequence to filter label noise. It leverages a modified version of the self-attention mechanism to downweight samples considered to be mislabeled. IDEAL [26] proposes a unified framework to implement comprehensive calibration from instance to metric. It leverages dual networks, where one extracts intra-class similarity and the other extracts inter-class divergence, to jointly weight samples. IDEAL reevaluates each sample by aggregating the information of dual networks while considering the consistent constraints. DCML [38] proposes a curriculum meta-learning model with a dual-level class-example sampling strategy that formulates a robust curriculum to adaptively adjust the task distribution for robust model training.

*Summary and Discussion:* Sample selection can delicately remove noisy samples, which easily enhances the model's generalization to unseen data. However, these methods involve additional data preprocessing steps, including identification and

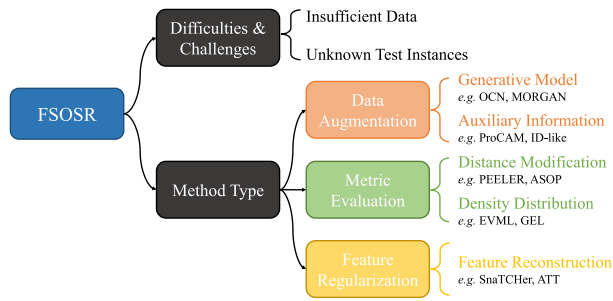


Fig. 4. The structural relationships of the existing three types of methods and their subclasses, as well as the difficulties and challenges faced by FSOSR problems.

removal of noisy labels, which may increase implementation complexity. Besides, sample selection may introduce bias about how to select and identify noisy samples.

#### IV. FEW-SHOT LEARNING WITH VARYING INSTANCES: FEW-SHOT OPEN-SET RECOGNITION

In the open world, due to unpredictable, unknown situations, FSL systems often face the challenge of testing samples that do not belong to known classes, which is known as few-shot open-set recognition (FSOSR). For example, medical diagnostics frequently confront unknown or previously unobserved diseases that were absent in training datasets. FSOSR requires diagnostic models to maintain accurate classification of known diseases while robustly detecting and flagging these unfamiliar disease instances for further expert evaluation.

*Definition:* FSOSR focuses on the scenario that the query dataset  $\mathcal{Q}^t$  includes unknown class queries. That is, apart from the novel label set  $\mathcal{C}_n^t$  in FSL, the labels of queries are also from an unknown label set  $\mathcal{C}_u^t$ , i.e.,  $\hat{y}^t \in \mathcal{C}_n^t \cup \mathcal{C}_u^t$ . Therefore, the objective of few-shot open-set recognition is to predict the class label of the query accurately if  $\hat{y}^t \in \mathcal{C}_n^t$  or to identify the unknown query sample if  $\hat{y}^t \in \mathcal{C}_u^t$ .

FSOSR requires maintaining classification performance on known classes and rejecting the unknown ones. The key challenges of FSOSR come from two aspects: 1) insufficient data: due to the scarcity of samples, it is difficult to construct precise decision boundaries for known classes and 2) unknown instances: the model needs to detect unknown test instances while simultaneously classifying known classes. Therefore, FSL models should fully capture the distribution information of known classes with limited samples and learn the true distribution of unknown classes to simulate the test environment. To address this issue, several works have focused on adapting to the open-set environment, which can be categorized into a unified taxonomy including Data Augmentation, Metric Evaluation, and Feature Processing, as shown in Fig. 4. The data-based approaches generate pseudo-samples to simulate the distribution of unknown classes. Common approaches include generative models (e.g., OCN [39], MORGAN [40] and DA-FSOS [41]) and auxiliary information (e.g. ProCAM [42] and ID-like prompts [43]). The metric-based approaches construct a highly discriminative feature space. The mainstream technological paths include

distance modification (e.g., PEELER [16], FSOSTC [44], openFEAT [45], CTR [46], MRM [47], ASOP [48], MSCL [49] and G-FOOD [50]) and density distribution (e.g. EVML [51], OSLO [52] and GEL [53]). Feature-based methods compare the differences between the reconstructed sets before and after (e.g., SnaTCHer [54], ATT [55], ReFOCS [56], FreqDiMFT [57] and RFDNet [58]).

In this section, we will introduce the latest progress of few-shot learning with unknown instances from three aspects. The overview of the problem definition and method strategy is depicted in Fig. 5.

##### A. Data Augmentation

Data augmentation, as depicted in Fig. 5(b), enriches the original training set to tackle the issue of insufficient labeled training samples. According to the source of pseudo samples, data augmentation can further be divided into generative model methods and auxiliary information methods.

Generative model methods generate pseudo-samples using probabilistic statistical models or machine learning models. OCN [39] proposes using samples generated by a Variational Autoencoder (VAE) to augment the query set features, applying it to hyperspectral image recognition. It generates pseudo-known samples by optimizing the VAE using the known class support set and query sets, allowing it to recognize samples generated by the VAE as known classes and detect samples from unknown classes. Building on this, MORGAN [40] combines FSOSR with Generative Adversarial Networks (GAN) and uses two different noise variances to generate pseudo-known and pseudo-unknown samples, including a decaying Gaussian distribution with lower noise variance to generate pseudo-known samples and a generator with higher noise variance to simulate features of unknown classes. DA-FSOS [41] involves augmenting both known and pseudo-unknown classes from source and target domains using a pair of class and domain conditional adversarial networks.

Generative models trained with limited samples may produce low-quality pseudo-samples, resulting in a decline in model performance. To solve this problem, auxiliary information is adopted in FSOSR. ProCAM [42] suggests using the background information of the support set as the unknown class, proposing that the background of support samples significantly differs from the foreground known class which can be known as an extra background unknown class. ID-like prompts [43] leverages the CLIP model to construct ID-like outliers from ID samples through random cropping and filtering based on cosine similarity with ID prompts. The adaptive prompts are learned by aligning OOD prompts with outliers and maximizing the dissimilarity between prompts to improve classification probabilities.

*Summary and Discussion:* Data augmentation methods can directly tackle the challenges of limited samples and the unpredictable nature of unknown classes by utilizing pseudo-samples, which enhances the model's ability to generalize to unknown classes. However, generated pseudo-unknown samples might not accurately represent the true distribution of unknown classes. Besides, due to the small size of the support set, the amount of auxiliary information is limited.

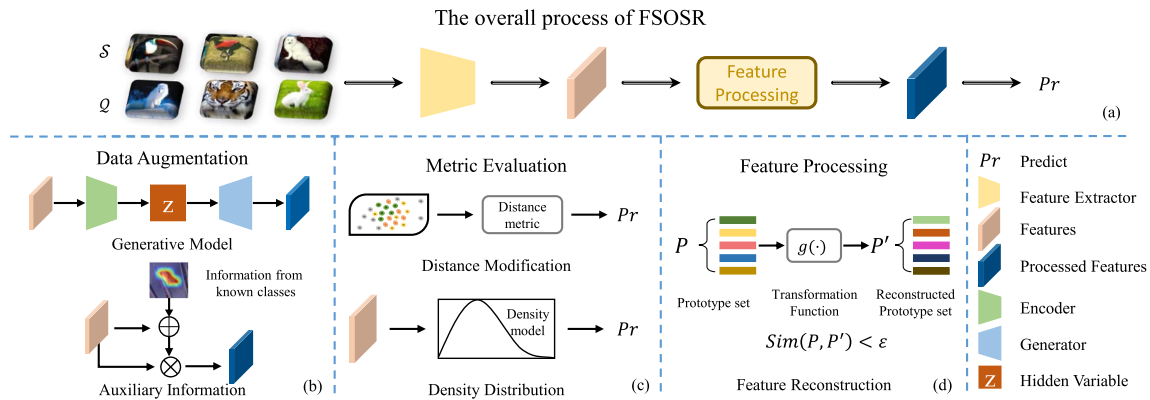


Fig. 5. An overall diagram of existing FSOSR research. (a) Overview of the FSOSR setting. (b) Data Augmentation enriches the insufficient labeled samples by Generative Model and Auxiliary Information. (c) Metric Evaluation constructs a feature space with appropriate metric functions with Distance Modification and Density Distribution. (d) Feature Processing compares the differences between the reconstructed sets before and after using transformation functions.

### B. Metric Evaluation

Metric evaluation, as depicted in Fig. 5(c), constructs an appropriate feature space and evaluates the similarity between samples using a metric function. According to the measurement of unknown classes, the metric evaluation can further be divided into distance modification and density distribution.

Distance modification evaluates the distance margin between known and unknown classes. PEELER [16] integrates a meta-learning framework with Mahalanobis distance and introduces an open-set loss through negative entropy loss, aimed at maximizing the entropy of known classes. FSOSTC [44] employs self-supervised learning to assist pre-training for the few-shot open-set traffic classification. openFEAT [45] collaboration with FEAT [59], utilizing a set-to-set function, adapts speaker representations from a predefined universal embedding space to a family-specific embedding space, significantly enhancing the performance of household member identification. CTR [46] proposes a unified framework for few-shot open-world recognition utilizing contrastive learning and employs negative training through distribution boundary contrastive loss. MRM [47] proposes learning a distributional hypersphere for each class while simultaneously increasing the margin between different categories, implicitly delineating the distribution space of unknown classes. ASOP [48] proposes a task-agnostic open-set class with distance scaling factors, which is the nearest class for every sample except its corresponding ground-truth class. MSCL [49] combines supervised contrastive learning and meta-learning principles to boost the open-set automatic modulation classification. G-FOOD [50] proposes two innovative modules: the Class Weight Sparsification Classifier (CWSC) and the Unknown Decoupling Learner (UDL). CWSC reduces the co-adaptability between classes by randomly sparsifying the normalized weights for class logit prediction, while UDL decouples the training of the unknown class, enabling the model to form a compact unknown decision boundary.

Density distribution is based on the assumption that unknown samples are located in low-density regions of the distribution, while known samples are in dense neighborhoods. EVML [51]

fits Weibull distributions for each known class based on distances from the limited support set to their respective prototypes. OSLO [52] applies a refined Maximum Likelihood principle and recalibrates likelihood values through a strategic weighting of samples, ensuring that samples categorized to known classes are assigned higher likelihoods. GEL [53] leverages features at both the class and pixel levels to learn energy scores. Specifically, global energy scores are computed using class-level features, capturing overall distinctions among samples and their predicted classes, while local energy scores concentrate on pixel-level details, uncovering fine-grained differences in features.

*Summary and Discussion:* Metric evaluation methods optimize the feature space and enlarge the differences between known classes and unknown ones, enabling the distinction of unknown class samples even with a small support set. However, these methods struggle to establish clear distribution boundaries when the feature distributions of different classes are similar or overlapping. This ambiguity leads to confusion in distinguishing samples and ultimately reduces the model's recognition capability.

### C. Feature Processing

Feature processing, as depicted in Fig. 5(d), reconstructs the feature and compares the differences between the reconstructed sets before and after to identify whether the sample belongs to an unknown class. SnaTCHer [54] proposes that the features of a sample will exhibit a high degree of consistency after undergoing specific transformation functions. SnaTCHer replaces the feature vector of each query sample with the predicted class prototype. The modified class prototype set is then reconstructed, and the sample is judged to match the predicted class prototype by comparing the differences of the set before and after reconstruction. Building on this, ATT [55] proposes using transformation functions and other reconstruction techniques to generate prototypes representing unknown classes, which aims to provide a flexible rejection boundary for various few-shot open-set recognition tasks. Specifically, during the meta-testing, if a test query sample's predicted scores for all known classes

are lower than its corresponding negative class prototype score, the classifier will reject the sample, deeming it to belong to an unknown class. ReFOCS [56] introduces reconstruction as an auxiliary task into the few-shot learning framework, guiding the classifier to develop a self-awareness ability. This unique self-awareness allows the model to naturally detect samples that likely belong to an unknown class when it struggles to accurately reconstruct the input. FreqDiMFT [57] introduces a module utilizing Fast Fourier Transform (FFT) to create a frequency distribution map from the original image and feature maps. This module captures domain-specific features by integrating frequency distribution information, addressing the high inter-class similarity and intra-class diversity in remote sensing images. RFDNet [58] learns to generate feature displacements for both the support set and the query adaptively based on information about the whole task, which is generated by transferring the knowledge extracted from the auxiliary dataset, i.e., linearly combining the base-class centroids followed by attentive modulation.

*Summary and Discussion:* Feature processing methods introduce additional information through feature reconstruction, establishing reference points with a limited number of known support samples, which can lead to more accurate measurements between known and unknown classes. However, these methods are hard to select the appropriate transformation function to ensure that the reconstructed features effectively represent the known class.

## V. FEW-SHOT LEARNING WITH VARYING CLASSES: INCREMENTAL FEW-SHOT LEARNING

In the open world, few-shot learning systems often face the challenge that classes of tasks as inputs vary frequently. The problem of few-shot learning with varying classes, also known as incremental few-shot learning (IFSL). For example, in medical discoveries, new diseases or disease subtypes emerge frequently, requiring model updates. In practice, stringent patient privacy regulations restrict access to historical medical images, compelling diagnostic systems to incrementally learn new categories from very few labeled examples without revisiting prior datasets.

*Definition:* IFSL focuses on the scenario that training data is presented as a sequence of tasks. The data sequence is defined as  $\{\mathcal{D}^0, \mathcal{D}^1, \dots, \mathcal{D}^T\}$ , with  $T$  representing the latest session.  $\mathcal{D}^0 = \mathcal{D}_b$  typically contains sufficient training data to provide the model with a good initialization. When session  $t$  is greater than 0,  $\mathcal{D}^t$  follows the N-way-K-shot setting, which is a few-shot learning scenario. The tasks in different sessions have disjoint class sets, i.e.,  $\mathcal{C}^0 \cap \mathcal{C}^1 \dots \cap \mathcal{C}^T = \emptyset$ , with the test set in session  $t$  containing all previously seen classes, i.e.,  $C_{\text{test}} = \mathcal{C}^0 \cup \mathcal{C}^1 \cup \dots \cup \mathcal{C}^t$ . During session  $t$ , the model can only access the current task's data  $\mathcal{D}^t$  and a memory set  $\mathcal{D}^m$  for data replay.

The key challenges of IFSL arise from three aspects: 1) catastrophic forgetting: the significant degradation of performance on previously learned tasks when incorporating new knowledge, largely due to the highly coupled parameters in deep neural networks; 2) the stability-plasticity dilemma: fixed parameters help retain old knowledge but simultaneously

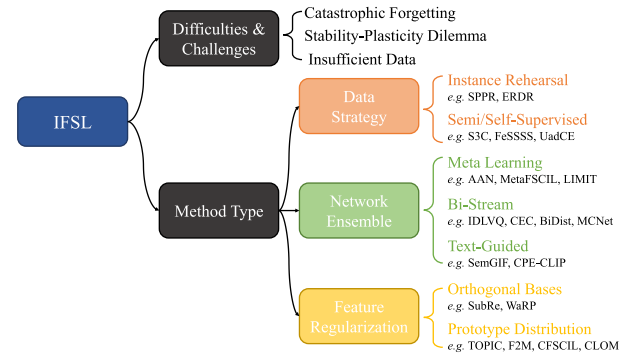


Fig. 6. The structural relationships of the existing three types of methods and their subclasses, as well as the difficulties and challenges faced by IFSL problems.

hinder effective adaptation to new tasks, thereby compromising the model's learning capacity, and 3) insufficient data: each newly encountered class typically comes with very limited labeled examples, while historical data is usually inaccessible due to privacy or storage constraints, significantly increasing the difficulty of achieving robust performance across both old and new tasks. To address the problem, existing works can be categorized into a unified taxonomy including Data Strategy, Network Ensemble, and Feature Regularization as shown in Fig. 6. The data-based approach provides the system with the ability to backtrack knowledge through data manipulation. Common approaches include instance rehearsal (SPPR [60] and ERDR [61]) and the introduction of unlabeled data for self-supervised learning and semi-supervised learning ([62], [63], [64]). Network-based methods aim to aggregate models with different performances. The mainstream technological paths include meta-learning (AAN [65], MetaFSCIL [66] and LIMIT [67]), bi-stream networks (IDLVQ [68], CEC [69], BiDist [70], MCNet [71]), M2SD [72] and BH-RCNN [73], and text guidance (SemGIF [74] and CPE-CLIP [75]). Feature-based methods alleviate catastrophic forgetting by measuring the distribution in the feature space. They commonly including orthogonal basis (SubRe [76], WaRP [77]) and OrCo [78] and prototype distribution (TOPIC [15], F2M [79], CFSCIL [80], CLOM [81], FACT [82], ALFSCIL [83] and NCFSCIL [84]).

In this section, we will introduce the latest progress of few-shot learning with varying classes from three aspects. The overview of the problem definition and method strategy is depicted in Fig. 7.

### A. Data Strategy

Data strategy, as depicted in Fig. 7(b), focuses on how to efficiently utilize limited labeled data, maintaining a dynamic memory set, and introducing unlabeled data are mainstream methods. According to the data usage method, data strategy can further be divided into instance rehearsal and unlabeled data.

Instance rehearsal preserves representative samples in incremental learning for knowledge playback by maintaining a memory set. SPPR [60] proposes a random episode selection strategy for sampling and obtaining pseudo-incremental data to enhance

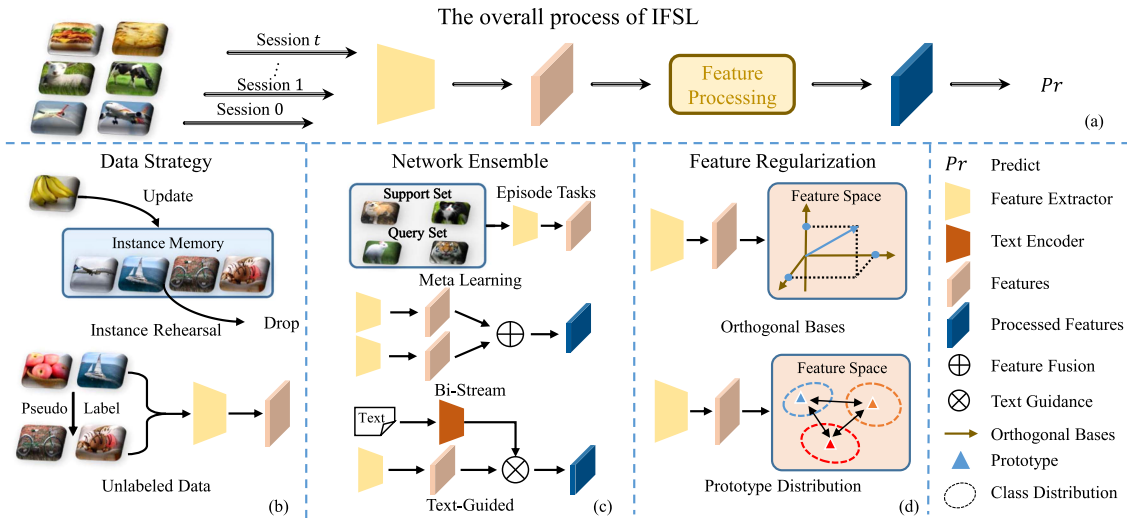


Fig. 7. An overall diagram of existing IFSL research. (a) Overview of the IFSL setting. (b) The data strategy can be categorized into instance rehearsal and Unlabeled Data utilization based on the manner of data utilization. (c) Network Ensemble methods focus on combining networks with different functions, including Meta Learning networks, Bi-Stream networks, and Text-Guidance. (d) Feature Regularization methods enable incremental learning by constraining the distribution of instances in the feature space with Orthogonal Basis and Prototype Distribution.

the adaptability of the model. In addition, a dynamic relationship projection module is proposed based on the similarity of features before and after the model to ensure the stationarity of the feature space before and after updates. Unlike most previous studies, ERDR [61] believes that data replay can also be applied to IFSL tasks. Considering the issue of data privacy, the author proposes a rehearsal-free method that introduces entropy regularization to generate several pseudo samples near the decision boundary. This uncertainty can alleviate overfitting and forgetting problems in incremental learning. Agarwal et al. [85] believe that it is not feasible to directly sample the memory set from the training set in IFSL, because the obtained examples cannot reflect the sample characteristics well due to data scarcity. They propose a method of instance rehearsal based on adversarial generation networks (GAN) to generate representative pseudo-samples.

Unlabeled data introduces unlabeled data in IFSL through unsupervised and semi-supervised learning, thus alleviating the problem of data scarcity. S3C [62] proposes that the scarcity of data caused by few-shot is an important reason for forgetting, which can be mitigated by introducing self-supervised learning with unlabeled data. By randomly enhancing unlabeled images and proposing a random classifier to fuse the results of multiple classifiers, self-supervised learning is effectively applied to IFSL. FeSSSS [63] proposes that self-supervised learning models can be used to assist IFSL. It believes that the main reason for the few-shot problem is the high labeling cost. Therefore, from a practical perspective, self-supervised models can be used to improve model performance. At the same time, the self-supervised data used does not need to be strictly distributed with the incremental learning data, and only needs to be roughly kept from the same domain. UaD-CE [86] explores the application of semi-supervised learning in IFSL tasks, and its designed UaD module alleviates the interference of few-shot new tasks on pseudo labels by retaining the pseudo label information of the previous round, improving the quality

of pseudo labels in semi-supervised learning. At the same time, UaD-CE designed the CIE module, which ensures the diversity of pseudo labels in the incremental learning process by applying class-balanced self-training (CB-ST). Cui et al. [64] propose a semi-supervised few-shot class-incremental learning framework named Us-KD, which is based on knowledge distillation and uncertainty guidance. By integrating semi-supervised learning into the few-shot class-incremental learning process, this framework utilizes an uncertainty-guided module to select unlabeled samples, combines knowledge distillation to alleviate catastrophic forgetting

*Summary and Discussion:* Data strategy can mitigate catastrophic forgetting by using real samples for knowledge retention. However, there are some limitations to this approach. First, regarding the selection of the memory set, inappropriate sampling techniques may result in a collection of low quality samples within the memory. These sub-optimal samples can then misdirect the model's learning process, leading to subpar performance. Second, in scenarios where privacy sensitive information is involved or in edge computing environments, maintaining a memory set is often deemed unfeasible or inappropriate. This constraint significantly restricts the applicability and versatility of such methods.

### B. Network Ensemble

Network ensemble, as depicted in Fig. 7(c), aims to combine different networks and parameters to adapt to constantly increasing new tasks. According to the difference of functions, network ensemble can further be divided into meta learning, bi-stream and text-guide methods.

Meta learning methods employ episode training to integrate the ability to handle different tasks on a network. Unlike general ensemble learning, this is the integration at the level of network representation. AAN [65] first proposes the concept of IFSL

and solves this problem through meta-learning. AAN consists of three stages which are pre-training, attention attractor training, and meta incremental learning separately, proving that the current backpropagation effectively updates the parameters of incremental models. MetaIFSL [66] introduces meta-learning and proposes that task sequences can be obtained by sampling in the base class to simulate the incremental learning process. Meanwhile, a bidirectional guidance mechanism (BGM) is proposed to automatically calculate modulation maps on the support set to assist in query set classification. By sampling in the base class for pseudo incremental learning, LIMIT [67] is prepared in advance for tasks that have not yet appeared. In addition, a transformer-based meta calibration method is proposed to align new prototypes and old classifiers before and after incremental learning.

Bi-Stream divides the network into two parallel branches, obtains the classification results from two perspectives respectively, and achieves the final output results through decision fusion. IDLVQ [68] only updates the correct prototype and the most recent incorrect prototype, alleviating the risk of overfitting caused by the few-shot setting. The proposed Learning Vector Quantization (LVQ) can be seen as a dynamic network architecture, as it continuously adds new reference vectors as incremental learning progresses. CEC [69] proposes a three-stage network that considers new features as a combination of old features and calculates weighted weights through self-attention. It saves the features learned in each round as a memory for calculating Graph Attention Network (GAT), which means that the features have a contextual relationship and can consider the relationships between different features as the training rounds increase, then it is possible to learn this weighting method through meta-learning. BiDist [70] proposes that both the base model and the previous model can be used as teachers for bidirectional distillation, which can effectively alleviate the risk of overfitting during current task learning. At the same time, a bilateral network was introduced to predict the base class and new class, respectively, and the branch prediction results were merged through the attention aggregation module. MCNet [71] obtains an ensemble learning model suitable for IFSL by decision fusion of feature extractors with different performances. CNN and Transformer are used to train two independent feature extractors with different preferences, and an attention regularization is used to fuse them. M2SD [72] uses a dual branch structure during the base class training session, which promotes the base class to learn more robust feature representations through Multiple Mixing Self-Distillation mechanism, which is beneficial for subsequent incremental learning. BH-RCNN [73] separates the models of the base class and the novel class, and applies a method similar to ensemble learning to obtain a unified prediction result.

Text-guide methods introduce text information to guide visual learning. SemGIF [74] proposes that generating instances can replace instance memory. Conditional-GAN is applied in the instance generation process, which generates base class pseudo instances by sampling the support set to achieve data rehearsal. In addition, SemGIF uses label text embedding as guiding information for image semantics, thereby assisting in feature

extraction. Furthermore, Alessandro et al. [75] design the Continuous Parameter Efficient CLIP (CPE-CLIP) for text-vision multimodal IFSL based on CLIP. Leveraging multimodal information, CPE can significantly reduce the number of parameters to be updated while maintaining competitive performance.

*Summary and Discussion:* Network ensemble methods have good scalability with multimodal information and can support both CIL and TIL settings. However, these methods are debatable whether the ensemble of networks can lead to the integration of knowledge. Training on few-shot tasks does not achieve comprehensive and sufficient training, which may result in insufficient knowledge integration and poor model generalization. Besides, due to the introduction of multi-model aggregation, such methods face fairness problems in comparison.

### C. Feature Regularization

Feature regularization, as depicted in Fig. 7(d), aims to avoid mutual influence between incremental tasks by applying appropriate regularization. According to different types of regularization, this approach can be further divided into prototype distribution and orthogonal basis.

Prototype distribution methods quantify the distance between prototypes from different classes to enhance class separability. These methods typically aim to minimize the intra-class variance while maximizing the inter-class distance. Existing approaches address this by optimizing the distribution of prototypes in the feature space, using metrics like Euclidean distance or cosine similarity. By doing so, they improve the model's ability to distinguish between classes, especially in few-shot settings where class information is limited. TOPIC [15] is proposed to learn the topological structure of prototype vectors. TOPIC uses anchor loss when learning new tasks and minimizes the movement of old categories, which limits the distance of node movement during each update. F2M [79] proposes to find a flat area in the base class model that is not sensitive to parameters, and learn new tasks in this area to slow down forgetting. CLOM [81] finds that class level overfitting is the main cause of catastrophic forgetting, and models often overfit to base classes rather than specific samples. By introducing edges, an integrated model containing different edges was constructed, effectively improving the performance of IFSL. C-IFSL [80] proposes that pairwise orthogonality can be imposed on prototypes corresponding to distinct classes. This approach is designed to safeguard the stability of the feature space associated with old tasks during the process of learning new tasks. FACT [82] proposes that during training, preparations can be made in advance for future category expansion. It is assumed that there are several virtual classes, and when training the base classes, the target is specified as both real and virtual classes. By analyzing models trained based on CE loss, Song et al. [87] finds that such models often have poor inter-class separability. Therefore, it is proposed to improve generalization performance by introducing supervised contrastive learning (SCL). Based on the MoCo [88] framework, Song et al. [87] proposes the SAVC model, which introduces virtual categories during contrastive learning and reserves feature space positions

in advance for possible new categories, thereby improving generalization, which is similar to FACT [82].

Inspired by the analogical learning mechanism of the human brain, ALFSCIL [83] proposes the idea of using analogical learning for classifier generation and model update to address the issues of overfitting and catastrophic forgetting in IFSL. Meanwhile, the optimization is formulated as a bi-level optimization problem to demonstrate the stability and plasticity of the model.

Wang et al. [89] propose a training-free prototype calibration strategy. By integrating new prototypes and weighted base prototypes, TEEN calibrates the biased prototypes of new classes using the rich semantic information in the base classes, thereby enhancing the discriminability of new classes. In IFSL, a training-free strategy is efficient and can calibrate prototypes using semantic similarity, making them suitable for scenarios with few new samples and similar data distributions. However, there exist limitations in generalization and they rely on parameter tuning. Traditional training methods offer high flexibility and great potential for accuracy optimization, making them suitable for scenarios with large amounts of data and complex and variable tasks. Nevertheless, they suffer from problems such as overfitting, forgetting old knowledge, and high training costs.

MICS [90] presents a mixup-based [36] IFSL method named MICS (Midpoint Interpolation for Compact and Separated Representation). It constructs mixup samples, establishes midpoint classifiers, and trains the model, enabling effective learning of compact and separated feature representations during both the base training and incremental learning stages. The method theoretically guarantees an increase in the thickness of the margin space between classes.

Orthogonal basis methods decompose the feature vector into a set of orthogonal basis vectors, thereby avoiding interfering with old knowledge when learning new tasks. SubRe [76] proposes to perform orthogonal decomposition on the base class features to obtain a set of bases, and use the convex combination of bases as the new class feature representation, thereby ensuring that the model learns semantic features. Similar to [80], WaRP [77] finds that the catastrophic forgetting in incremental learning is due to the high coupling of neural network parameters, and even fine-tuning very few parameters can change most of the previous knowledge. Therefore, the weight space rotation method WaRP is proposed, which improves the forgetting of previous knowledge by finding a set of orthogonal bases in the weight space and only fine-tuning them in their smoothest direction. OrCo [78] also utilized a method similar to FACT [82] by preparing virtual class prototypes for future categories in advance. Additionally, OrCo also aligns instance features and pseudo targets by applying orthogonal loss and perturbed supervised contrastive loss, improving the separability of instances in the feature space.

*Summary and Discussion:* The methods focus on the distribution of feature space and alleviate catastrophic forgetting through regularization. They take into account the global relationship of feature distributions, making the feature representation consistent among different tasks. However, it is difficult to disentangle the effects of large numbers of coupling parameters during regularization, which makes knowledge retention very difficult.

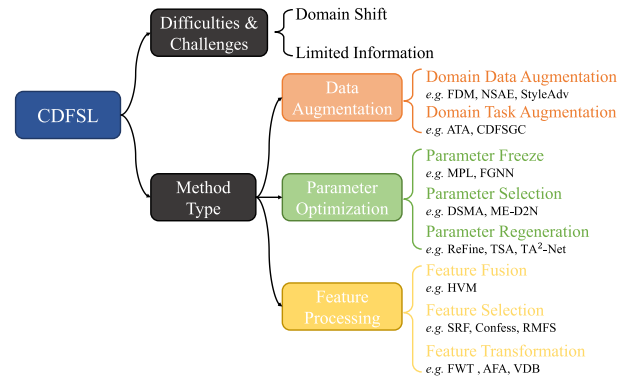


Fig. 8. The structural relationships of the existing three types of methods and their subclasses, as well as the difficulties and challenges faced by CDFSL problems.

Besides, feature regularization requires clear distribution boundaries between categories, but this is often difficult to achieve under few-shot conditions.

## VI. FEW-SHOT LEARNING WITH VARYING DISTRIBUTION: CROSS-DOMAIN FEW-SHOT LEARNING

In the open world, few-shot learning systems often face the challenge that distributions from source domains to target domains vary frequently. The problem of few-shot learning with varying distribution requires to transfer the knowledge learned on the source domain to the target domain, which is known as Cross-Domain Few-shot Learning (CDFSL). For example, medical images collected across different clinical institutions or imaging devices exhibit notable differences in image quality, resolution, and acquisition protocols. Such domain discrepancies cause significant performance degradation when diagnostic models trained in one domain are applied to another, demanding methods that can generalize effectively with limited labeled examples from diverse medical domains.

*Definition:* CDFSL focuses on the scenario that domain gaps between novel class datasets  $\mathcal{D}_n$  (a.k.a target domain) and base class datasets  $\mathcal{D}_b$  (a.k.a source domain) are significant, i.e.,  $\mathcal{C}_n \cap \mathcal{C}_b = \emptyset, \mathcal{P}_n \neq \mathcal{P}_b$ . The goal is to classify the images of the query set  $\mathcal{Q}^t$  by training a feature extractor and a classification head on the support set  $\mathcal{S}^t$ .

The key challenges of CDFSL come from two aspects: 1) domain shift: there are considerable gaps between source and target domains, and 2) limited information: the scarcity of sample sizes in the target domain makes the cross-domain problems more challenging. Therefore, models trained under the CDFSL problem need to have robust and strong generalization capabilities for source domains adapting to novel target domains with limited examples. To address this issue, several works have focused on adapting to the cross-domain environment, which can be categorized into a unified taxonomy including Data Augmentation, Parameter Optimization and Feature Processing, as shown in Fig. 8. The data-based approach focuses on augmenting data as well as tasks. Common approaches include data

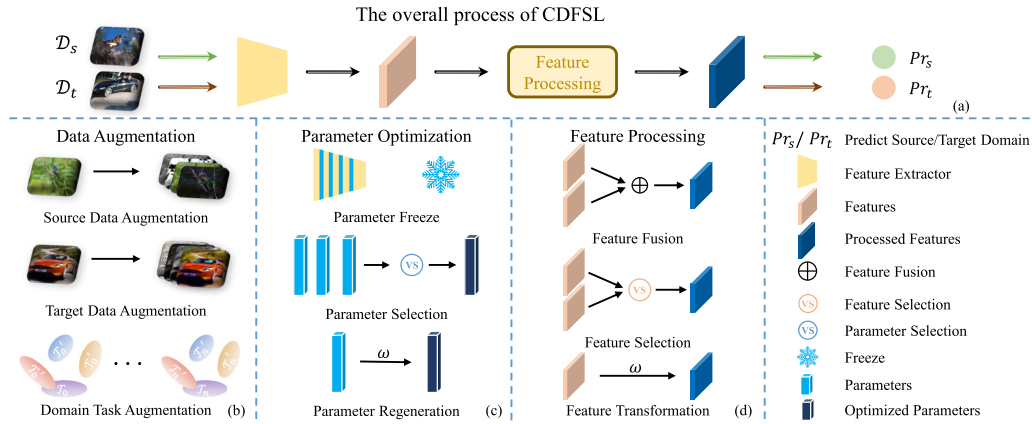


Fig. 9. An overall diagram of existing CDFSL research. (a) Overview of the CDFSL setting. (b) Data Augmentation enhances the quality and size of the samples by auxiliary information including Source Data, Target Data and Domain Task. (c) Parameter Optimization performs operations on parameters including Parameter Freeze, Parameter Selection and Regeneration. (d) Feature Processing optimizes features by Feature Fusion, Selection and Transformation.

enhancement (e.g. FDM [91], NSAE [92] and StyleAdv [93]) and task enhancement (e.g. ATA [94] and CDFSGC [95]). The parameter-based approach is widely used in CDFSL. The mainstream methods include parameter freeze (e.g. MPL [96], and FGNN [97]), parameter selection (e.g. DSMA [98] and ME-D2N [99]), and parameter regeneration (e.g. ReFine [100], TSA [101] and TA<sup>2</sup>-Net [102]). Feature-based methods perform different operations on features such as feature fusion (e.g. HVM [103]), feature selection (e.g. SRF [104], Confess [105] and RMFS [106]), and feature transformation (e.g. FWT [14], AFA [107] and VDB [108]).

In this section, we will introduce the latest progress of few-shot learning with unknown distributions from three aspects. The overview of the problem definition and method strategy is depicted in Fig. 9.

### A. Data Augmentation

Data enrichment, as depicted in Fig. 9(b), improves the size and quality of training datasets for deep learning models, especially those facing Few-Shot Learning (FSL) problems [109]. Depending on different types of data, data augmentation can be divided into domain data augmentation and domain task data augmentation.

Domain data augmentation aims to boost the training samples to extract more general knowledge to improve the generalization performance of the model over the target domain. The augmentation data includes source and target data. Feature Distribution Matching (FDM) [91] proposes a feature-wise domain adaptation module to augment the query set samples by MixUp, which encourages the model to generate more synthetic samples for better domain adaptation. NSAE [92] trains the model by jointly reconstructing the inputs and predicting the labels of the inputs as well as their reconstructed pairs. StyleAdv [93] is proposed to augment source domain samples by a novel style adversarial attack method. SVasP [110] simulates more diverse potential target domain adversarial styles via diversifying input patterns and aggregating localized crop style gradients. SET-RCL [111] simulates the style distributions of unknown target domains.

TACDFSL [112] applies rotation transformations to augment images and predict the rotation angle in the pretrain phase. Similarly, CDTA [113] corrupts the semantic information of the benign image by scrambling the outputs of both the intermediate feature layers and the final layer of the feature extractor.

Domain task augmentation takes the task as a whole and tries to generate more tasks. For example, ATA [94] considers the worst-case problem around source domain tasks and proposes adversarial methods to augment the source domain task distribution to address the domain gap in few-shot learning. In CDFSGC [95], by augmenting the graphs from sampled tasks into three views, the representation of graphs is obtained for fast adaptation as well as knowledge transfer. Furthermore, by comparing predictions of weakly-augmented unlabeled target data from a teacher network to strongly-augmented versions of the same images from a student network [114], consistency regularization is imposed to learn a representation that can be easily adapted to the target domain. KT [115] obtains two views of the support set by stochastic image augmentation.

*Summary and Discussion:* Data augmentation is a simple but effective technique that can be flexibly used in CDFSL to expand the amount of data. However, the improvement of their model performance is not significant since the fundamental issue of domain shift has not been addressed. The domain gap between the source and target domains is still significant. Therefore, compared to the other two methods that are mentioned below, data augmentation can probably be seen more as a trick than as a core.

### B. Parameter Optimization

Parameter optimization, as depicted in Fig. 9(c), improves the model performance by fine-tuning model parameters. According to different parameter handling methods, parameter optimization can further be divided into parameter freeze, parameter selection and parameter regeneration.

Parameter freeze methods fix some model parameters to limit the complexity of the hypothesis space. MPL [96] proposes meta-prototypical learning to alternately freeze the

meta-encoder's parameters to learn the general features. Moreover, FGNN [97] proposes two new strategies, respectively for the encoder and the metric function of metric-based network.

Parameter selection strategy tries to identify the most suitable set of parameters for better adaptation to the target domains. For example, DSMA [98] samples several sub-networks by dropping neurons (or feature maps) to construct a bunch of models with diverse features for the target domain and chooses the most suitable sub-networks to construct the ensemble for the target domain learning. ME-D2N [99] introduces few labeled target domain data and transfers the knowledge from two teachers to a unified student model by the knowledge distillation technique.

The parameter regeneration technique adjusts some model parameters to further optimize the performance of the model. Various studies have been explored to boost cross-domain ability. For instance, ReFine [100] re-randomizes the parameters fitted on the source domain before adapting to the target data, which resets source-specific parameters of the source pre-trained model and thus facilitates fine-tuning on the target domain. TSA [101] proposes to learn task-specific weights from scratch directly on a small support set, in contrast to dynamically estimating them. A Feature-Wise Transformation (FWT) layer [14] is added after the BN layer in convolutional neural networks to simulate feature distributions in different domains. ProLAD [116] utilizes two separate adapters with one devoid of a normalization layer, which is more effective for similar domains, and another embedded with a normalization layer, designed to leverage the batch statistics of the target domain, thus proving effective for dissimilar domains. TA<sup>2</sup>-Net [102] is trained by reinforcement learning to adaptively estimate the optimal task-specific parameter configuration for each test task. UCDFSL [117] investigates which pre-training is preferred based on domain similarity and few-shot difficulty of the target domain and designs two pre-training schemes, mixed-supervised and two-stage learning. MetaPrompt [118] jointly exploits prompt learning and the parameter generation framework to flexibly produce a task-adaptive prompt with arbitrary length for unseen tasks. SWP [119] combines the soft weight pruning strategy and regularization to effectively restrict redundant weights while simultaneously learning crucial features for both source and target tasks.

*Summary and Discussion:* Parameter Optimization is critical to improve the model's performance which can find the most suitable set of parameters. However, most methods utilize simple structures or have a limited selection of parameters, resulting in limited improvement of the model's performance.

### C. Feature Processing

Feature processing aims at achieving the most transferable feature representation by processing the original features. According to different feature processing methods, feature processing can further be divided into feature fusion, feature selection and feature transformation.

Feature fusion focuses on fusing features from different depths of the neural network or from different domains or different dimensions within a single feature. HVM [103] uses

a hierarchical prototype model to flexibly rely on features at different semantic levels which enables the model to adaptively choose the most generalizable features. Likewise, Zou et al. [104] fuse mid-level features to learn discriminative information of each sample and thus boost the discriminability of the model. FDFSL [120] employs an orthogonal low-rank feature disentanglement method to acquire the desired features of source and target pipelines.

Feature selection strategy often selects domain-general features to aid target domain generation. For example, In order to obtain the more relevant domain feature, SRF [104] uses a multi-domain feature bank to automatically select the most relevant representations. Confess [105] utilizes a contrastive learning and feature selection system to address domain gaps between base and novel categories. In addition, the RMFS [106] algorithm together with multi-domain feature selection is proposed to optimize the multi-domain features. MMT [121] bridges the domain gap between source and target domains, leveraging style-memory and content-memory components. LRP [122] develops a model-agnostic explanation-guided training strategy that dynamically finds and emphasizes the features which are important for the predictions.

At the same time, feature transformation reweights features when extra network and multi-domain data are not available. Simple CNAPS [123] learns adaptive feature extractors that allow useful estimation of the high-dimensional feature covariances from surprisingly few samples. RDC-FT [124] constructs a non-linear subspace to minimize task-irrelevant features there-within while keeping more transferrable and discriminative information by a hyperbolic tangent transformation. Wang et al. [125] and Xu et al. [126] both normalize features to alleviate feature dissimilarity. VDB [108] transfers the intermediate feature using the source and target domain statistical elements. IM-DCL [126] employs a weighted distance calculation among features to establish a soft classification of the positive and negative sets for the entire feature set. ProD [127] uses the prompting mechanism in the transformer to disentangle the domain-general (DG) and domain-specific (DS) knowledge from the backbone feature. MFIE [128] uses LMMs to automatically translate class-specific images into class-agnostic natural language descriptions for various key semantic attributes in target domains and edit origin images based on class-agnostic natural language descriptions. zxVAD [129] uses a novel Normalcy Classifier module to learn the features of normal event videos by learning how such features are different 'relatively' to features in pseudo-abnormal examples. STARTUP [130] proposes to train feature representations on the novel target domain to replicate the induced grouping. PMNet [131] extracts the domain-agnostic pixel-level affinity matching with a frozen backbone and captures both the pixel-to-pixel and pixel-to-patch relations in each support-query pair with the bidirectional 3D convolutions. Another work [107] introduces a plug-and-play adversarial feature augmentation (AFA) method for CDFSL. FAP [132] designs a mutual attention module to learn a generalizable inductive bias under CD-FSL settings. URL [133] learns a single set of universal visual representations by distilling knowledge of multiple domain-specific networks after co-aligning their features with the help of adapters

TABLE I  
THE NFSL PERFORMANCE OF THE PROPOSED METHODS ON *miniImageNet* AND *tieredImageNet* BENCHMARKS

Type	Methods	Venue	Backbone	K-shot	<i>miniImageNet</i>			<i>tieredImageNet</i>		
					0%	20%	40%	0%	20%	40%
Parameter Optimization	Nested-MAML [30]	AAAI22	ConvNet4	5	61.20±0.21	59.60±0.54	56.46±0.74	63.42±0.86	62.31±0.86	56.49±0.85
	PCL [25]	IJCAI23	ConvNet4	5	70.31±0.69	68.70±0.72	67.64±0.68	69.94±0.81	68.95±0.75	66.69±0.77
Sample Selection	RNNP [32]	WACV21	ConvNet4	5	68.38±0.24	62.43±0.25	51.62±0.23	71.36±0.18	65.95±0.19	54.86±0.21
	RapNets [24]	TNNLS21	ConvNet4	5	66.45±0.25	63.10±0.26	55.08±0.25	68.02±0.25	64.61±0.26	56.85±0.25
	TraNFS [17]	CVPR22	ConvNet4	5	68.13±0.23	64.45±0.25	56.76±0.25	69.63±0.23	65.76±0.25	56.96±0.25
	IDEAL [26]	TPAMI23	ConvNet4	5	67.93±0.22	64.27±0.24	56.45±0.22	71.07±0.22	66.97±0.22	60.13±0.23
	DETA [33]	ICCV23	ConvNet4	5	70.21±0.22	65.93±0.21	57.04±0.22	72.41±0.23	68.77±0.22	59.57±0.22

*K* means 5-way *K*-shot.

and centered kernel alignment. MvSA [134] makes use of labeled source domain data and easy target domain data to perform target-to-source and target-to-target alignment. CDSTN [135] introduces a set of domain classifiers behind each residual block for the purpose of domain alignment by extracting more domain information at different depths of the network. Dara [136] proposes a prototypical feature alignment to recalibrate support instances as prototypes and reproject these prototypes with a differentiable closed-form solution. AFGR [137] enhances different semantic feature information to help the metric function better locate the fine-grained feature information of the image.

*Summary and Discussion.* Feature processing strategy is beneficial to preserve or generate the most suitable features for the target domain. However, feature processing may introduce additional computational costs due to additional parameters. Besides, feature processing may only allow limited exploration for most process features from the final layer.

## VII. PERFORMANCES

### A. Few-Shot Learning With Noisy Instances

1) *Datasets and Settings:* To ensure a fair comparison of the NFSL approaches, we select some representative NFSL methods and compare them with the same implementation requirements and configuration. We conduct two widely used few-shot image recognition benchmarks: *miniImageNet* [138] and *tieredImageNet* [139].

- *miniImageNet* [138] contains 64, 16, and 20 classes for training, validation, and testing set, with 60 K images in total.
- *tieredImageNet* [139] contains 351, 97, and 160 classes for training, validation, and testing sets, respectively, with a total of approximately 0.78 million images.

In this paper, we focus on exploring symmetric label swap noise, which means the noise comes from the current task, it is produced by replacing the raw sample with a sample from the classes of the episode. Following previous work [17], we add noise to each class uniformly, in other words, meaning that the same proportion of noise is added for all classes. Parameter optimization methods are devoted to alleviating the noise in the meta-training datasets instead of the meta-testing datasets. Sample selection methods focus on training the meta-learner to adapt to the noisy support sets in the test stage.

2) *Metric:* In NFSL, the metric is the accuracy under different noise rates. Specifically, we explore three types of noise rates: 0%, 20%, 40%. In the support set, the number of noise

samples per class is equal to the noise ratio multiplied by the number of samples per class.

3) *Performance Comparison and Analysis:* Table I summarizes the performance of selected NFSL approaches evaluated on the two benchmark datasets, *miniImageNet* and *tieredImageNet*. The results of the methods based on parameter optimization are obtained from the original papers, while the results of the methods based on sample selection are from our version of implementation.

Table I presents the performance of compared methods faced with various noise rates of symmetric noise on two datasets. From Table I, it can be observed that, for the parameter optimization methods, compared to PCL [25], Nested-MAML [30] shows inferior performance because its model is based on gradient-based methods and PCL is the model based on metric learning. gradient-based methods commonly require multiple gradient updates, resulting in high computational costs and a tendency to get stuck in local optima. In contrast, metric-based methods are more efficient in training and possess strong generalization capabilities, leading to superior performance. Therefore, PCL achieves better performance. Meanwhile, in the sample selection methods, among the meta-training weighting methods, IDEAL achieves the best performance. For the test-time adaptation methods, DETA [33] achieves the best performance. Comparing DETA and IDEAL [26], we observe that DETA generally outperforms IDEAL in most cases, indicating that test-time adaptation methods with powerful backbones currently hold a certain advantage over meta-training methods.

### B. Few-Shot Learning With Unknown Instances

1) *Datasets and Settings:* To compare the FSOSR approaches in a fair manner, We select some representative IFSL methods and compare them in the same implementation requirements and configurations. FSOSR experiments are commonly conducted on two datasets, *miniImageNet* [138] and *tieredImageNet* [139].

In terms of open-set settings, following existing research [16], we set  $N = 5$  and  $K = 1, 5$  during meta-training and meta-testing. For each episode, we sample 5 known classes and 5 unknown classes, each containing 15 queries. All support samples are from the known classes. The query set samples are from both known and unknown classes.

2) *Metrics:* In FSOSR, two metrics are applied to measure the performance of the model, including closed-set classification accuracy (Acc) and Area Under the Receiver Operating Characteristics curve (AUROC) for unknown class detection. The Acc measures the classification capacity via known samples,

TABLE II  
THE FSOSR PERFORMANCE OF THE PROPOSED METHODS ON *miniImageNet* AND *tieredImageNet* BENCHMARKS

Type	Methods	Venue	Backbone	<i>miniImageNet</i>				<i>tieredImageNet</i>			
				1-shot		5-shot		1-shot		5-shot	
				Acc	AUROC	Acc	AUROC	Acc	AUROC	Acc	AUROC
Data Augmentation	ProCAM [42]	ACM MM22	ResNet12	67.89±0.85	71.52±1.02	82.82±0.53	81.12±0.60	69.70±0.76	74.71±0.99	85.12±0.79	82.31±0.61
	ProCAMsm [42]	ACM MM22	ResNet12	67.84±0.85	70.98±0.86	83.39±0.53	79.52±0.66	69.43±1.01	74.79±0.94	85.23±0.77	82.43±0.62
Metric Evaluation	PEELER [16]	CVPR20	ResNet12	65.86±0.85	60.57±0.83	80.61±0.59	67.35±0.80	69.51±0.92	65.20±0.76	84.10±0.66	73.27±0.71
	MRCM-ProtoNet [47]	IJCAI23	ResNet12	64.05±0.82	68.11±0.73	84.73±0.51	77.07±0.63	68.35±0.94	72.99±0.74	84.75±0.62	78.18±0.60
	MRRMT-SnaTcheF [47]	IJCAI23	ResNet12	67.03±0.83	71.20±0.80	82.00±0.55	80.39±0.59	71.13±0.91	75.59±0.77	85.27±0.62	83.03±0.63
	ASOP [48]	ICIP23	ResNet12	66.30±0.60	71.72±0.50	82.68±0.20	80.92±0.20	68.68±0.70	74.92±0.40	84.84±0.40	82.37±0.30
	ASOP-L [48]	ICIP23	ResNet12	67.85±0.20	71.91±0.70	82.81±0.30	81.04±0.20	71.49±0.40	75.04±0.40	85.15±0.10	81.51±0.10
	GEL [53]	CVPR23	ResNet12	68.26±0.85	73.70±0.82	83.05±0.55	82.29±0.60	70.50±0.93	75.86±0.81	84.60±0.65	81.95±0.72
Feature Processing	SnaTche-F [54]	CVPR21	ResNet12	67.02±0.85	68.27±0.96	82.02±0.53	77.42±0.73	70.52±0.96	74.28±0.80	84.74±0.69	82.02±0.64
	SnaTche-T [54]	CVPR21	ResNet12	66.60±0.80	70.17±0.88	81.77±0.53	76.66±0.78	70.45±0.95	74.84±0.79	84.42±0.68	82.03±0.66
	SnaTche-L [54]	CVPR21	ResNet12	67.60±0.83	69.40±0.92	82.36±0.58	76.15±0.83	70.85±0.99	74.95±0.83	85.23±0.64	80.81±0.68
	ATT [55]	CVPR22	ResNet12	67.64±0.81	71.35±0.68	82.31±0.49	79.85±0.58	69.34±0.95	72.74±0.78	83.82±0.63	78.66±0.65
	ATT-G [55]	CVPR22	ResNet12	68.11±0.81	72.41±0.72	83.12±0.48	79.85±0.57	70.58±0.93	73.43±0.78	85.38±0.61	81.64±0.63
	RFDNet [58]	TMM22	ResNet12	66.23±0.80	71.37±0.80	82.44±0.54	80.31±0.59	66.84±0.89	72.68±0.76	82.64±0.63	81.95±0.72

1-shot, 5-shot means 5-way 1-shot and 5-way 5-shot.

and the AUROC evaluates the unknown detection capacity via both known and unknown samples.

3) *Performance Comparison and Analysis*: The experimental results on the *miniImageNet* and *tieredImageNet* datasets are counted as shown in Table II. All the results come from the original papers.

Among the three types of methods, metric-based approaches consistently exhibit superior stability and performance compared to the other two types. Specifically, GEL [53] stands out as the top performer across all methods. GEL learns global open scores for detecting unknown samples by analyzing class-wise and pixel-wise scales. It assigns high energy scores to samples that significantly differ from the few-shot examples in either class-wise features or pixel-wise features, and assigns low energy scores otherwise. In the data-based method, proCAM [42] and its variants exhibit remarkable performance by leveraging background regions of images from known classes, thus adding auxiliary information from known samples. It reveals that generating appropriate virtual samples can simulate the open environment during testing, thereby indicating the ability of few-shot models to detect unknown class samples. In the metric-based methods, apart from GEL, MRM [47] and ASOP [48] also demonstrate strong performance, which might be attributed to their effective optimization of the feature space. The excellent performances of these approaches demonstrate the efficient generalization capability of metric-based methods. In the feature-based methods, ATT [55] and its variant ATT-G showcase good performance, which generates negative prototypes to represent unknown classes through the transformation function, thus providing a flexible and concise rejection boundary. Meanwhile, ATT-G utilizes class semantics to model the unknown class and achieve better performance than other methods.

### C. Few-Shot Learning With Varying Classes

1) *Datasets and Settings*: In order to compare the IFSL approaches in a fair manner, we select some representative IFSL methods and compare them with the same implementation requirements and configurations. IFSL experiments are commonly conducted on three datasets, CIFAR-100 [140], *miniImageNet* [138] and CUB-200 [141]. Apart from *miniImageNet* which has been described, we describe CIFAR-100 and CUB-200 as follows.

- CIFAR-100 [140] has 100 categories and each category contains 500 training instances and 100 test instances.
- CUB-200 [141] contains 200 categories of bird images, with 5994 images for training and 5794 images for testing.

In terms of incremental learning settings, CIFAR-100 and *miniImageNet* take 60 categories as base classes, and the remaining 40 categories are continually learned in a 5-way 5-shot manner, for a total of 8 sessions. CUB-200 uses the first 100 categories as base classes and the remaining 100 categories as a 10-way 5-shot input to the model.

2) *Metrics*: For comprehensive performance comparisons, four commonly used metrics are applied to measure the performance of methods, including Start Accuracy (SA), End Accuracy (EA), Average Accuracy (AA), and Performance Drop (PD). SA is the accuracy of the first sessions  $Acc_1$ . EA is the accuracy of the last sessions  $Acc_T$ , where  $T$  is the number of all sessions. AA is accuracy across all incremental sessions:

$$AA = \frac{1}{T} \sum_{t=1}^T Acc_t \quad (1)$$

PD measures the absolute accuracy drops in the last session:

$$PD = Acc_T - Acc_1 \quad (2)$$

3) *Performance Comparison and Analysis*: In Table III, we summarize the performance of selected IFSL approaches evaluated on the commonly used benchmarks, CIFAR-100, *miniImageNet*, and CUB-200. Due to significant differences in the initial accuracy of these methods, the direct comparison of AA does not reflect the true capability of the model. We evaluate these methods mainly considering the PD after incremental learning.

Among the three types of methods, network ensemble methods are consistently more stable and superior than the other two types of methods. It is worth noting that CPE-CLIP [75], with the language-image pre-training model, performs best in all methods. It reveals that the language-image pairs with abundant semantic information effectively guide the learning of incremental new tasks. In the meantime, CPE-CLIP, which applies the CLIP model [142] pre-trained on out-of-distribution datasets, is unfair to directly compare with other methods. In the data strategy methods, the UadCE with unlabeled samples almost outperforms other data strategy methods with the least PD, especially in CIFAR-100 and CUB-200. In *miniImageNet*, the superiority

TABLE III  
THE PERFORMANCE OF SELECTED IFSL METHODS ON CIFAR-100, *mini*ImageNet AND CUB-100

Type	Methods	Venue	BackBone	CIFAR-100				BackBone	<i>mini</i> ImageNet				CUB-200			
				SA	EA	AA $\uparrow$	PD $\downarrow$		SA	EA	AA $\uparrow$	PD $\downarrow$	SA	EA	AA $\uparrow$	PD $\downarrow$
Data Strategy	SPPR [60]	CVPR21	ResNet20	64.10	43.25	54.51	20.85	ResNet18	78.68	48.12	60.84	30.56	68.68	37.33	49.32	31.35
	FeSSSS [63]	CVPRW21	ResNet20	75.35	50.23	60.92	25.12	ResNet18	81.50	58.87	68.23	22.63	79.60	52.98	62.85	26.62
	FSIL-GAN [85]	ACM MM22	ResNet20	70.14	46.61	55.31	23.53	ResNet18	69.87	46.14	56.39	23.73	81.25	59.21	69.34	22.04
	ERDR [61]	ECCV22	ResNet20	74.40	50.14	60.77	24.26	ResNet18	71.84	48.21	58.02	23.63	75.90	53.39	61.52	22.51
	Us-KD [64]	TMM22	ResNet20	76.85	54.42	61.80	22.43	ResNet18	72.35	50.47	58.79	21.88	74.69	55.54	62.89	19.15
UadCE [86]	TNNLS23	ResNet20	75.55	54.50	63.93	21.05	ResNet18	72.35	50.52	58.82	21.83	75.17	60.72	65.70	14.45	
Network Ensemble	CEC [69]	CVPR21	ResNet20	73.07	49.14	59.53	23.93	ResNet18	72.00	47.63	57.75	24.37	75.85	52.28	61.33	23.57
	MetaFSCL [66]	CVPR22	ResNet20	74.50	49.97	60.79	24.53	ResNet18	72.04	49.19	58.85	22.85	75.90	52.64	61.92	23.26
	BiDist [70]	CVPR23	ResNet20	79.45	56.56	66.55	22.89	ResNet18	74.65	53.39	62.06	21.26	79.12	63.81	68.61	15.31
	MCNet [71]	TIP23	ResNet20	73.30	50.72	60.39	22.58	ResNet18	72.33	49.08	58.64	23.25	77.57	59.08	65.57	18.49
	LIMIT [67]	TPAMI23	ResNet20	73.81	51.23	61.84	22.58	ResNet18	72.32	49.19	59.06	23.13	75.89	57.41	65.45	18.48
	M2SD [72]	AAAI24	ResNet20	79.48	53.55	64.74	25.93	ResNet18	82.11	56.51	68.63	25.60	81.49	60.96	68.14	20.53
	CPE-CLIP [75]	ICCVW23	CLIP-ViT-b	87.83	80.52	83.42	7.31	CLIP-ViT-b	90.23	82.77	86.13	7.46	81.58	64.60	70.79	16.98
Feature Regularization	TOPIC [15]	CVPR20	ResNet20	64.10	29.37	38.36	34.73	ResNet18	61.31	24.42	35.68	36.89	68.68	26.28	43.92	42.40
	F2M [79]	NeurIPS21	ResNet20	71.45	49.35	59.14	22.10	ResNet18	67.28	44.65	54.89	22.63	81.07	60.26	67.24	20.81
	CLOM [81]	NeurIPS22	ResNet20	74.20	50.25	60.56	23.95	ResNet18	73.08	48.00	58.48	25.08	79.57	59.58	67.17	19.99
	FACT [82]	CVPR22	ResNet20	74.60	52.10	62.24	22.50	ResNet18	72.56	50.49	60.70	22.07	75.90	56.94	64.42	18.96
	WaRP [77]	ICLR23	ResNet20	80.31	54.74	65.82	25.57	ResNet18	72.99	50.65	59.69	22.34	77.74	57.01	62.15	20.73
	TEEN [89]	NeurIPS23	ResNet20	74.92	52.64	63.10	22.28	ResNet18	73.53	52.08	61.43	21.45	77.26	59.31	66.62	17.95
	SAVC [87]	CVPR23	ResNet20	78.77	53.12	63.63	25.65	ResNet18	81.12	57.11	67.05	24.01	81.85	62.50	69.35	19.35
	ALFSCL [143]	TCSVT24	ResNet20	80.75	55.17	66.74	25.58	ResNet18	81.27	53.31	64.92	27.96	79.79	59.30	66.99	20.49
	MCS [90]	WACV24	ResNet20	78.18	52.94	63.62	25.24	ResNet18	84.40	60.74	70.17	23.66	78.77	61.37	67.39	17.40

The “ $\downarrow$ ” after the distance measures indicates “the smaller the better”, and the “ $\uparrow$ ” after the similarity measures indicates “the larger the better”. In the comparison, for all methods except for CPE-CLIP, ResNet20 was used on CIFAR-100, and ResNet18 for *mini*ImageNet and CUB-200. CPE-CLIP applies the pre-trained model from openai clip and is not suitable for comparison with other methods.

of UadCE [86] is relatively weakened. In the network ensemble methods, apart from CPE-CLIP, BiDist [70] showcases good performance as well, which might be attributed to the attention fusion mechanism to leverage the discrepancy between base classes and novel classes. Among the feature-based methods, F2M [79], a representative of the early feature-based approaches, still exhibits robust competitiveness. TEEN [89] and EHS [143] outperform other methods to a relatively greater extent.

#### D. Few-Shot Learning With Varying Distributions

1) *Datasets and Settings*: CDFSL has various experimental setups including different datasets, backbones and configurations, we compare some representative CDFSL methods in a unified and fair manner. There are three main benchmarks of the CDFSL problem, including BSCD-FSL [28], mini-CUB [14] and Meta-Dataset [29]. BSCD-FSL has five image datasets from a diverse assortment of image acquisition methods, including *mini*ImageNet [138], ChestX [144], ISIC [145], [146], EuroSAT [147], CropDiseases [148]. mini-CUB [14] has five datasets including CUB [141], Cars [149], Places [150], Plantae [151]. Meta-dataset benchmark consists of multiple datasets of different data distributions including Omniglot [152], Aircraft [153], Textures [154], Quick Draw [155], Fungi [156], VGG Flower [157], Traffic Signs [158] and MSCOCO [159]. In the CDFSL, we mainly leverage the BSCD-FSL for comparison.

The standard evaluation metric used in CDFSL is prediction accuracy and there are various settings, including 5-way 1-shot, 5-way 5-shot, 5-way 20-shot, and 5-way 50-shot. Moreover, backbones mainly include ResNet10 and ResNet18.

2) *Performance Comparison and Analysis*: In Table IV, we summarize the performance of the representative CDFSL methods evaluated on the BSCD-FSL and mini-CUB benchmarks. All the results come from the original papers.

Among the three types of methods, parameter optimization methods achieve optimal results on most datasets. The reason behind this may be that the introduction of additional parameters specific to the target domain allows for increased generalization of the model. The results show that the performance decreases when the distance between the source domain and the

target domain increases. Data-based methods for CDFSL are relatively simple as they mostly enhance the generalization of the model by augmenting the data or introducing supplementary information. Among data-based methods, StyleAdv [93] achieves optimal results due to the stronger generalization of the adversarial samples generated by the introduction of stylistic noise. Parameter-based methods are more stable. For example, using ResNet10 as the backbone, the results of SWP [119] on BSCD-FSL and mini-CUB (5-way 5-shot) demonstrate this trend, with scores of 31.30% on ChestX, 93.52% on EuroSAT, 97.84% on CropDisease and 76.08% on Plantae. Moreover, in the feature-based methods, the results of Dara [136] show that introducing target domain information in the training process can effectively improve cross-domain classification results.

#### VIII. OPEN CHALLENGES AND FUTURE DETECTIONS

Despite significant progress in OWFSL, it continues to present unique challenges that require attention. As such, we outline several promising research directions for the future.

*Few-Shot Learning with Different Concepts*: Conventional FSL models are trained with an assumption of a single problem, i.e., each instance can only encapsulate one visual concept even though in a multi-class setting. In the open world, such as fashion recognition [160], bioinformatics [161], and drug discovery [162], each instance in few-shot scenarios may be interpreted by different concepts. To address this challenge, recent advancements in multi-label learning have been adopted in FSL to combat multi-label issues [163], [164], [165]. For instance, LaSO [163] employs a data augmentation strategy that generates synthetic feature vectors through label-set operations. KGGR [164] adopts a GCN where labels are modeled as nodes and statistical label co-occurrences are modeled as edges to exploit the label dependencies. The above methods assume that the importance degree of different instances for a certain label is the same and that of different labels in an instance is also the same. BCR [165] explores diverse important degrees and effectively leverages the underlying label correlations, along with the varying importance from both instance-to-label and label-to-instance perspectives.

TABLE IV  
THE CDFSL PERFORMANCE OF THE PROPOSED METHODS ON BSCD-FSL BENCHMARK

Type	methods	Venue	Train Set	BackBone	$K$ -shot	ChestX	ISIC	EuroSAT	CropDisease
Data Augmentation	ATA [94]	IJCAI21	<i>miniImageNet</i>	ResNet10	1	22.10±0.20	33.21±0.40	61.35±0.50	67.47±0.50
					5	24.32±0.40	44.91±0.40	83.75±0.40	90.59±0.30
	SET-RCL [111]	ACM MM22	<i>miniImageNet</i>	ResNet10	1	22.74±0.20	33.33±0.40	65.53±0.60	68.43±0.50
					5	25.65±0.20	44.93±0.40	83.84±0.40	88.11±0.30
	KT [115]	PR23	<i>miniImageNet</i>	ResNet10	1	22.68±0.60	34.06±0.77	66.43±0.93	73.10±0.87
					5	26.79±0.61	46.37±0.77	82.53±0.66	89.53±0.58
StyleAdv [93]	CVPR23	<i>miniImageNet</i>	ResNet10	1	22.64±0.35	33.96±0.57	70.94±0.82	74.13±0.78	
				5	26.07±0.37	45.77±0.51	86.58±0.54	93.65±0.39	
Parameter Optimization	FWT [14]	ICLR20	<i>miniImageNet</i>	ResNet10	1	22.04±0.44	31.58±0.67	62.36±1.05	66.36±1.04
					5	25.18±0.45	43.17±0.70	83.01±0.79	87.11±0.67
	UCDFSL [117]	Neurips22	<i>miniImageNet</i>	ResNet10	1	21.97±0.41	34.90±0.64	79.50±0.78	89.49±0.74
					5	25.62±0.43	46.76±0.61	92.36±0.37	97.24±0.33
	ReFine [100]	CIKM22	<i>miniImageNet</i>	ResNet10	1	22.48±0.41	35.30±0.59	64.14±0.82	68.93±0.84
					5	26.76±0.42	51.68±0.63	82.36±0.57	90.75±0.49
SWP [119]	TMM24	<i>miniImageNet</i> target data	ResNet10	1	24.16±0.43	36.07±0.66	77.72±0.76	91.96±0.64	
				5	31.30±0.42	51.36±0.63	93.52±0.32	97.84±0.42	
STARTUP [130]	ICLR21	<i>miniImageNet</i> target data	ResNet10	1	23.09±0.43	32.66±0.60	63.88±0.84	75.93±0.80	
				5	26.94±0.44	47.22±0.61	82.29±0.60	93.02±0.45	
Feature Processing	LRP [122]	ICPR21	<i>miniImageNet</i>	ResNet10	1	22.11±0.20	30.94±0.30	54.99±0.50	59.23±0.50
					5	24.53±0.30	44.14±0.40	77.14±0.40	86.15±0.40
	RDC-FT [124]	CVPR22	<i>miniImageNet</i>	ResNet10	1	22.27±0.2	35.84±0.4	71.57±0.5	86.33±0.5
					5	25.48±0.2	49.06±0.3	84.67±0.3	93.55±0.3
	VDB [108]	CVPR22	<i>miniImageNet</i>	ResNet10	1	22.99±0.44	35.32±0.65	63.60±0.87	71.98±0.82
					5	26.62±0.45	48.72±0.65	82.06±0.63	90.77±0.49
IM-DCL [126]	TIP24	target data	ResNet10	1	23.98±0.79	38.13±0.57	77.14±0.71	84.37±0.99	
				5	28.93±0.41	52.74±0.69	89.47±0.42	95.73±0.38	
Dara [136]	ITPAMI23	<i>miniImageNet</i> target data	ResNet10	1	22.92±0.40	36.42±0.64	67.42±0.80	80.74±0.76	
				5	27.54±0.42	56.28±0.66	85.84±0.54	95.32±0.34	
FAP [132]	IJCAI24	<i>miniImageNet</i>	ResNet10	1	22.36±0.20	35.63±0.40	62.96±0.50	69.97±0.50	
				5	25.31±0.20	47.60±0.40	82.52±0.40	91.79±0.30	

$K$  means 5-way  $K$ -shot.

*Few-Shot Learning with Adversarial Attacks:* Conventional FSL models are crafted based on original clean examples without imperceptible perturbations [166], [167]. In the open world, these models are vulnerable to adversarial examples, which easily misclassify these adversarial examples [168]. Therefore, it is crucial to learn of robust few-shot models in the open world that can effectively defend against adversarial attacks [166]. Recent works on few-shot learning methods defensive against adversarial attacks are proposed [169], [170], [171]. ADML[169] finds that retraining earlier layers of the network during fine-tuning impairs the robustness of the network, while only retraining later layers can largely preserve robustness. AQ [170] adopts projected gradient descent attack (PGD) [172] by crafting adversarial perturbations and minimizing the losses of perturbed query data. DFSL [171] performs episode-based adversarial training at the task level. In the meantime, DFSL assumes distribution consistency between the clean and adversarial examples and enforces the distribution consistency between clean and adversarial examples from the feature-wise or prediction-wise perspectives within each task.

*Few-Shot Learning with Imbalanced Classes:* Conventional FSL models are trained with the assumption of the same amount of data in each class. In the open world, it is unrealistic in aerial scenes [173] when the number of samples can vary and follow a wide range of imbalance distributions [174]. To address this challenge, recent advancements adopt random-shot tasks with randomly selected classes and samples [173]. RF-MML implements random episodic training on the majority of classes to learn generalized knowledge across data distribution and then fine-tunes all classes to improve the performance for multi-category classification further. Bayesian Task-Adaptive Meta-Learning [175] can adaptively adjust the parameters that guide

the scale of the meta-learner and the task-specific learner for each task and class. CIFSL [176] formalizes the class-imbalanced FSL setting at three levels of imbalance: task-level, dataset-level, and combined imbalance. Then it adapts rebalancing methods such as random oversampling and various reweighting losses to the CIFSL setting. Adaptive Manifold [177] hypothesizes that class centroid approaches will benefit from exploiting the data manifold to obtain more representative centroids and iteratively updates both the class centroids and the manifold-specific parameters with label propagation.

## IX. CONCLUSION

The open world poses a critical challenge in the realm of artificial intelligence, spanning applications such as autonomous driving and medical analysis. In autonomous Driving, unreliable noisy sensor data affects the system's ability to accurately perceive the environment. Unseen vehicles and traffic signs make it impossible for driving systems to recognize or handle them correctly. Inconsistencies in online model updates impact real-time decision-making. Heavy snow and other extreme weather result in significant performance degradation. Few-Shot Learning in the Open World (OWFSL) emerges as a generalization of few-shot learning (FSL), aiming to enhance the safety, robustness, and adaptability of autonomous driving systems within the diverse and dynamic open-world context. The core of OWFSL lies in the capacity to detect and adapt to new, varied, and ever-changing scenarios. This paper delineates OWFSL settings, encompassing varying instances, classes, and distributions. We comprehensively review the evolution of related methods and conduct comparative analyses of representative approaches within each OWFSL scenario. Additionally, we

outline several challenging open problems that merit substantial research endeavors and efforts in the future.

## REFERENCES

- [1] B. M. Lake, R. Salakhutdinov, and J. B. Tenenbaum, "Human-level concept learning through probabilistic program induction," *Science*, vol. 350, no. 6266, pp. 1332–1338, 2015.
- [2] F. Li, R. Fergus, and P. Perona, "One-shot learning of object categories," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 4, pp. 594–611, Apr. 2006.
- [3] Y. Wang, Q. Yao, J. T. Kwok, and L. M. Ni, "Generalizing from a few examples: A survey on few-shot learning," *ACM Comput. Surv.*, vol. 53, pp. 63:1–63:34, 2020.
- [4] W. Chen, Y. Liu, Z. Kira, Y. F. Wang, and J. Huang, "A closer look at few-shot classification," in *Proc. Int. Conf. Learn. Representations*, 2019.
- [5] Z. Yu, L. Chen, Z. Cheng, and J. Luo, "TransMatch: A transfer-learning scheme for semi-supervised few-shot learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 12853–12861.
- [6] C. Xue, Q. Dou, X. Shi, H. Chen, and P. Heng, "Robust learning at noisy labeled medical images: Applied to skin lesion classification," in *Proc. IEEE 16th Int. Symp. Biomed. Imag.*, 2019, pp. 1280–1283.
- [7] C. Xue, L. Yu, P. Chen, Q. Dou, and P. Heng, "Robust medical image classification from noisy labeled data with global and local representation guided co-training," *IEEE Trans. Med. Imag.*, vol. 41, no. 6, pp. 1371–1382, Jun. 2022.
- [8] L. Sun, M. Zhang, B. Wang, and P. Tiwari, "Few-shot class-incremental learning for medical time series classification," *IEEE J. Biomed. Health Inform.*, vol. 28, no. 4, pp. 1872–1882, Apr. 2024.
- [9] G.-X. Xu et al., "Cross-site severity assessment of COVID-19 from CT images via domain adaptation," *IEEE Trans. Med. Imag.*, vol. 41, no. 1, pp. 88–102, Jan. 2022.
- [10] W. Li, "LibFewShot: A comprehensive library for few-shot learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 12, pp. 14938–14955, Dec. 2023.
- [11] J. Lu, P. Gong, J. Ye, J. Zhang, and C. Zhang, "A survey on machine learning from few samples," *Pattern Recognit.*, vol. 139, 2023, Art. no. 109480.
- [12] J. Parmar, S. S. Chouhan, V. Raychoudhury, and S. S. Rathore, "Open-world machine learning: Applications, challenges, and opportunities," *ACM Comput. Surv.*, vol. 55, no. 10, pp. 205:1–205:37, 2023.
- [13] F. Zhu, S. Ma, Z. Cheng, X. Zhang, Z. Zhang, and C. Liu, "Open-world machine learning: A review and new outlooks," 2024, arXiv: 2403.01759.
- [14] H. Tseng, H. Lee, J. Huang, and M. Yang, "Cross-domain few-shot classification via learned feature-wise transformation," in *Proc. Int. Conf. Learn. Representations*, 2020.
- [15] X. Tao, X. Hong, X. Chang, S. Dong, X. Wei, and Y. Gong, "Few-shot class-incremental learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 12180–12189.
- [16] B. Liu, H. Kang, H. Li, G. Hua, and N. Vasconcelos, "Few-shot open-set recognition using meta-learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 8795–8804.
- [17] K. J. Liang, S. B. Rangrej, V. Petrovic, and T. Hassner, "Few-shot learning with noisy labels," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2022, pp. 9079–9088.
- [18] H. Xu, S. Zhi, S. Sun, V. M. Patel, and L. Liu, "Deep learning for cross-domain few-shot visual recognition: A survey," 2023, arXiv: 2303.08557.
- [19] W. Wang, L. Duan, Y. Wang, J. Fan, Z. Gong, and Z. Zhang, "A survey of deep visual cross-domain few-shot learning," 2023, arXiv: 2303.09253.
- [20] S. Tian, L. Li, W. Li, H. Ran, X. Ning, and P. Tiwari, "A survey on few-shot class-incremental learning," *Neural Netw.*, vol. 169, pp. 307–324, 2024.
- [21] J. Zhang, L. Liu, O. Silvén, M. Pietikäinen, and D. Hu, "Few-shot class-incremental learning: A survey," 2023, arXiv: 2308.06764.
- [22] L. Fei-Fei, R. Fergus, and P. Perona, "A Bayesian approach to unsupervised one-shot learning of object categories," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2003, pp. 1134–1141.
- [23] G. Koch et al., "Siamese neural networks for one-shot image recognition," in *Proc. Int. Conf. Mach. Learn. Appl. Workshops*, 2015, pp. 1–30.
- [24] J. Lu, S. Jin, J. Liang, and C. Zhang, "Robust few-shot learning for user-provided data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 4, pp. 1433–1447, Apr. 2021.
- [25] Y. An, X. Zhao, and H. Xue, "Learning to learn from corrupted data for few-shot learning," in *Proc. Int. Joint Conf. Artif. Intell.*, 2023, pp. 3423–3431.
- [26] Y. An, H. Xue, X. Zhao, and J. Wang, "From instance to metric calibration: A unified framework for open-world few-shot learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 8, pp. 9757–9773, Aug. 2023.
- [27] J. Willes, J. Harrison, A. Harakeh, C. Finn, M. Pavone, and S. L. Waslander, "Bayesian embeddings for few-shot open world recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1513–1529, Mar. 2024.
- [28] Y. Guo et al., "A broader study of cross-domain few-shot learning," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 124–141.
- [29] E. Triantafillou et al., "Meta-dataset: A dataset of datasets for learning to learn from few examples," in *Proc. Int. Conf. Learn. Representations*, 2020.
- [30] K. Killamsetty, C. Li, C. Zhao, F. Chen, and R. K. Iyer, "A nested bi-level optimization framework for robust few shot learning," in *Proc. Conf. Assoc. Advance. Artif. Intell.*, 2022, pp. 7176–7184.
- [31] D. Chen, L. Wu, S. Tang, X. Yun, B. Long, and Y. Zhuang, "Robust meta-learning with sampling noise and label noise via eigen-reptile," in *Proc. Int. Conf. Mach. Learn.*, 2022, pp. 3662–3678.
- [32] P. Mazumder, P. Singh, and V. P. Namboodiri, "RNNP: A robust few-shot learning approach," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, 2021, pp. 2663–2672.
- [33] J. Zhang, L. Gao, X. Luo, H. Shen, and J. Song, "Deta: Denoised task adaptation for few-shot learning," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2023, pp. 11507–11517.
- [34] S. Shao, Y. Bai, Y. Wang, B. Liu, and B. Liu, "Collaborative consortium of foundation models for open-world few-shot learning," in *Proc. Conf. Assoc. Advance. Artif. Intell.*, 2024, pp. 4740–4747.
- [35] M. Patacchiola, J. Turner, E. J. Crowley, M. O'Boyle, and A. J. Storkey, "Bayesian meta-learning for the few-shot setting via deep kernels," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2020, pp. 16108–16118.
- [36] H. Zhang, M. Cissé, Y. N. Dauphin, and D. Lopez-Paz, "Mixup: Beyond empirical risk minimization," in *Proc. Int. Conf. Learn. Representations*, 2018.
- [37] S. Zhu, L. Xie, M. Zhang, R. Gao, and Y. Xie, "Distributionally robust weighted k-nearest neighbors," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2022, pp. 29088–29100.
- [38] X. Que and Q. Yu, "Dual-level curriculum meta-learning for noisy few-shot learning tasks," in *Proc. Conf. Assoc. Advance. Artif. Intell.*, 2024, pp. 14740–14748.
- [39] D. Pal, V. Bundeale, R. Sharma, B. Banerjee, and Y. Jeppu, "Few-shot open-set recognition of hyperspectral images with outlier calibration network," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, 2022, pp. 2091–2100.
- [40] D. Pal, S. Bose, B. Banerjee, and Y. Jeppu, "MORGAN: Meta-learning-based few-shot open-set recognition via generative adversarial network," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, 2023, pp. 6284–6293.
- [41] D. Pal, D. More, S. Bhargav, D. Tamboli, V. Aggarwal, and B. Banerjee, "Domain adaptive few-shot open-set learning," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2023, pp. 18785–18794.
- [42] N. Song, C. Zhang, and G. Lin, "Few-shot open-set recognition using background as unknowns," in *Proc. ACM Int. Conf. Multimedia*, 2022, pp. 5970–5979.
- [43] Y. Bai, Z. Han, C. Zhang, B. Cao, X. Jiang, and Q. Hu, "ID-like prompt learning for few-shot out-of-distribution detection," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, 2024, pp. 17480–17489.
- [44] J. Li, C. Gu, L. Luan, F. Wei, and W. Liu, "Few-shot open-set traffic classification based on self-supervised learning," in *Proc. IEEE 47th Conf. Local Comput. Netw.*, 2022, pp. 371–374.
- [45] K. C. Kishan et al., "OpenFEAT: Improving speaker identification by open-set few-shot embedding adaptation with transformer," in *Proc. Int. Conf. Acoust. Speech Signal Process.*, 2022, pp. 7062–7066.
- [46] N. Dionelis, S. A. Tsaftaris, and M. Yaghoobi, "CTR: Contrastive training recognition classifier for few-shot open-world recognition," in *Proc. Int. Conf. Pattern Recognit.*, 2022, pp. 1792–1799.
- [47] Y. Che, Y. An, and H. Xue, "Boosting few-shot open-set recognition with multi-relation margin loss," in *Proc. Int. Joint Conf. Artif. Intell.*, 2023, pp. 3505–3513.
- [48] B. Kim, J.-T. Lee, K. Shim, and S. Chang, "Task-agnostic open-set prototype for few-shot open-set recognition," in *Proc. Int. Conf. Image Process.*, 2023, pp. 31–35.
- [49] J. Zhao, H. Wang, S. Peng, and Y. Yao, "Meta supervised contrastive learning for few-shot open-set modulation classification with signal constellation," *IEEE Commun. Lett.*, vol. 28, no. 4, pp. 837–841, Apr. 2024.

- [50] B. Su, H. Zhang, J. Li, and Z. Zhou, "Toward generalized few-shot open-set object detection," *IEEE Trans. Image Process.*, vol. 33, pp. 1389–1402, 2024.
- [51] D. Pal, S. Bose, B. Banerjee, and Y. Jeppu, "Extreme value meta-learning for few-shot open-set recognition of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5512516.
- [52] M. Boudiaf et al., "Open-set likelihood maximization for few-shot learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2023, pp. 24007–24016.
- [53] H. Wang, G. Pang, P. Wang, L. Zhang, W. Wei, and Y. Zhang, "Glocal energy-based learning for few-shot open-set recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2023, pp. 7507–7516.
- [54] M. Jeong, S. Choi, and C. Kim, "Few-shot open-set recognition by transformation consistency," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 12566–12575.
- [55] S. Huang, J. Ma, G. Han, and S. Chang, "Task-adaptive negative envision for few-shot open-set recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2022, pp. 7161–7170.
- [56] S. Nag, D. S. Raychaudhuri, S. Paul, and A. K. Roy-Chowdhury, "Reconstruction guided meta-learning for few shot open set recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 12, pp. 15394–15405, Dec. 2023.
- [57] J. Zhang, Y. Rao, X. Huang, G. Li, X. Zhou, and D. Zeng, "Frequency-aware multi-modal fine-tuning for few-shot open-set remote sensing scene classification," *IEEE Trans. Multimedia*, vol. 26, pp. 7823–7837, 2024.
- [58] S. Deng, J.-G. Yu, Z. Wu, H. Gao, Y. Li, and Y. Yang, "Learning relative feature displacement for few-shot open-set recognition," *IEEE Trans. Multimedia*, vol. 25, pp. 5763–5774, 2023.
- [59] H.-J. Ye, H. Hu, D.-C. Zhan, and F. Sha, "Few-shot learning via embedding adaptation with set-to-set functions," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 8805–8814.
- [60] K. Zhu, Y. Cao, W. Zhai, J. Cheng, and Z.-J. Zha, "Self-promoted prototype refinement for few-shot class-incremental learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 6801–6810.
- [61] H. Liu et al., "Few-shot class-incremental learning via entropy-regularized data-free replay," in *Proc. Eur. Conf. Comput. Vis.*, 2022, pp. 146–162.
- [62] J. Kalla and S. Biswas, "S3C: Self-supervised stochastic classifiers for few-shot class-incremental learning," in *Proc. Eur. Conf. Comput. Vis.*, 2022, pp. 432–448.
- [63] T. Ahmad et al., "Few-shot class incremental learning leveraging self-supervised features," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2022, pp. 3900–3910.
- [64] Y. Cui et al., "Uncertainty-guided semi-supervised few-shot class-incremental learning with knowledge distillation," *IEEE Trans. Multimedia*, vol. 25, pp. 6422–6435, 2022.
- [65] M. Ren, R. Liao, E. Fetaya, and R. S. Zemel, "Incremental few-shot learning with attention attractor networks," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2019, pp. 5276–5286.
- [66] Z. Chi, L. Gu, H. Liu, Y. Wang, Y. Yu, and J. Tang, "MetaFS-CIL: A meta-learning approach for few-shot class incremental learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2022, pp. 14166–14175.
- [67] D.-W. Zhou, H.-J. Ye, L. Ma, D. Xie, S. Pu, and D.-C. Zhan, "Few-shot class-incremental learning by sampling multi-phase tasks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 11, pp. 12816–12831, Nov. 2023.
- [68] K. Chen and C.-G. Lee, "Incremental few-shot learning via vector quantization in deep embedded space," in *Proc. Int. Conf. Learn. Representations*, 2020.
- [69] C. Zhang, N. Song, G. Lin, Y. Zheng, P. Pan, and Y. Xu, "Few-shot incremental learning with continually evolved classifiers," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 12455–12464.
- [70] L. Zhao et al., "Few-shot class-incremental learning via class-aware bilateral distillation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2023, pp. 11838–11847.
- [71] Z. Ji, Z. Hou, X. Liu, Y. Pang, and X. Li, "Memorizing complementation network for few-shot class-incremental learning," *IEEE Trans. Image Process.*, vol. 32, pp. 937–948, 2023.
- [72] J. Lin, Z. Wu, W. Lin, J. Huang, and R. Luo, "M2SD: Multiple mixing self-distillation for few-shot class-incremental learning," in *Proc. Conf. Assoc. Advance. Artif. Intell.*, 2024, pp. 3422–3431.
- [73] Y. Li, H. Zhu, S. Tian, J. Ma, C. Xiang, and P. Vadakkepat, "Bilateral-head region-based convolutional neural networks: A unified approach for incremental few-shot object detection," *IEEE Trans. Artif. Intell.*, vol. 5, no. 9, pp. 4376–4390, Sep. 2024.
- [74] S. D. Bhat, B. Banerjee, and S. Chaudhuri, "SemGIF: A semantics guided incremental few-shot learning framework with generative replay," in *Proc. Brit. Mach. Vis. Conf.*, 2021, Art. no. 169.
- [75] M. D'Alessandro, A. Alonso, E. Calabrés, and M. Galar, "Multimodal parameter-efficient few-shot class incremental learning," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2023, pp. 3393–3403.
- [76] A. F. Akyürek, E. Akyürek, D. Wijaya, and J. Andreas, "Subspace regularizers for few-shot class incremental learning," in *Proc. Int. Conf. Learn. Representations*, 2021.
- [77] D.-Y. Kim, D.-J. Han, J. Seo, and J. Moon, "Warping the space: Weight space rotation for class-incremental few-shot learning," in *Proc. Int. Conf. Learn. Representations*, 2022.
- [78] N. Ahmed, A. Kukleva, and B. Schiele, "OrCo: Towards better generalization via orthogonality and contrast for few-shot class-incremental learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2024, pp. 28762–28771.
- [79] G. Shi, J. Chen, W. Zhang, L. Zhan, and X. Wu, "Overcoming catastrophic forgetting in incremental few-shot learning by finding flat minima," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2021, pp. 6747–6761.
- [80] M. Hersche, G. Karunaratne, G. Cherubini, L. Benini, A. Sebastian, and A. Rahimi, "Constrained few-shot class-incremental learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2022, pp. 9057–9067.
- [81] Y. Zou, S. Zhang, Y. Li, and R. Li, "Margin-based few-shot class-incremental learning with class-level overfitting mitigation," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2022, pp. 27267–27279.
- [82] D.-W. Zhou, F.-Y. Wang, H.-J. Ye, L. Ma, S. Pu, and D.-C. Zhan, "Forward compatible few-shot class-incremental learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2022, pp. 9046–9056.
- [83] J. Li, S. Dong, Y. Gong, Y. He, and X. Wei, "Analogical learning-based few-shot class-incremental learning," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 34, no. 7, pp. 5493–5504, Jul. 2024.
- [84] Y. Yang, H. Yuan, X. Li, Z. Lin, P. Torr, and D. Tao, "Neural collapse inspired feature-aligner for few-shot class-incremental learning," in *Proc. Int. Conf. Learn. Representations*, 2022.
- [85] A. Agarwal, B. Banerjee, F. Cuzzolin, and S. Chaudhuri, "Semantics-driven generative replay for few-shot class incremental learning," in *Proc. ACM Int. Conf. Multimedia*, 2022, pp. 5246–5254.
- [86] Y. Cui, W. Deng, H. Chen, and L. Liu, "Uncertainty-aware distillation for semi-supervised few-shot class-incremental learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 10, pp. 14259–14272, Oct. 2024.
- [87] Z. Song, Y. Zhao, Y. Shi, P. Peng, L. Yuan, and Y. Tian, "Learning with fantasy: Semantic-aware virtual contrastive constraint for few-shot class-incremental learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2023, pp. 24183–24192.
- [88] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, "Momentum contrast for unsupervised visual representation learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 9729–9738.
- [89] Q. Wang, D. Zhou, Y. Zhang, D. Zhan, and H. Ye, "Few-shot class-incremental learning via training-free prototype calibration," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2023, pp. 15060–15076.
- [90] S. Kim, Y. Jeong, J. S. Park, and S. W. Yoon, "MICS: Midpoint interpolation to learn compact and separated representations for few-shot class-incremental learning," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, 2024, pp. 2236–2245.
- [91] Y. Fu, Y. Fu, and Y.-G. Jiang, "Meta-FDMixup: Cross-domain few-shot learning guided by labeled target data," in *Proc. ACM Int. Conf. Multimedia*, 2021, pp. 5326–5334.
- [92] H. Liang, Q. Zhang, P. Dai, and J. Lu, "Boosting the generalization capability in cross-domain few-shot learning via noise-enhanced supervised autoencoder," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2021, pp. 9424–9434.
- [93] Y. Fu, Y. Xie, Y. Fu, and Y.-G. Jiang, "StyleAdv: Meta style adversarial training for cross-domain few-shot learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2023, pp. 24575–24584.
- [94] H. Wang and Z.-H. Deng, "Cross-domain few-shot classification via adversarial task augmentation," 2021, arXiv: 2104.14385.
- [95] K. Hassani, "Cross-domain few-shot graph classification," in *Proc. Conf. Assoc. Advance. Artif. Intell.*, 2022, pp. 6856–6864.
- [96] R.-Q. Wang, X.-Y. Zhang, and C.-L. Liu, "Meta-prototypical learning for domain-agnostic few-shot recognition," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 11, pp. 6990–6996, Nov. 2022.

- [97] Y. Chen et al., "Cross-domain few-shot classification based on lightweight Res2Net and flexible GNN," *Knowl.-Based Syst.*, vol. 247, 2022, Art. no. 108623.
- [98] P.-C. Tu and H.-K. Pao, "A dropout style model augmentation for cross domain few-shot learning," in *Proc. IEEE Int. Conf. Big Data*, 2021, pp. 1138–1147.
- [99] Y. Fu, Y. Xie, Y. Fu, J. Chen, and Y.-G. Jiang, "ME-D2N: Multi-expert domain decomposition network for cross-domain few-shot learning," in *Proc. ACM Int. Conf. Multimedia*, 2022, pp. 6609–6617.
- [100] J. Oh, S. Kim, N. Ho, J.-H. Kim, H. Song, and S.-Y. Yun, "Refine: Re-randomization before fine-tuning for cross-domain few-shot learning," in *Proc. Conf. Inf. Knowl. Manage.*, 2022, pp. 4359–4363.
- [101] W.-H. Li, X. Liu, and H. Bilen, "Cross-domain few-shot learning with task-specific adapters," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2022, pp. 7161–7170.
- [102] Y. Guo, R. Du, Y. Dong, T. Hospedales, Y.-Z. Song, and Z. Ma, "Task-aware adaptive learning for cross-domain few-shot learning," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2023, pp. 1590–1599.
- [103] Y. Du, X. Zhen, L. Shao, and C. G. M. Snoek, "Hierarchical variational memory for few-shot learning across domains," in *Proc. Int. Conf. Learn. Representations*, 2022.
- [104] Y. Zou, S. Zhang, J. Yu, Y. Tian, and J. M. Moura, "Revisiting mid-level patterns for cross-domain few-shot recognition," in *Proc. ACM Int. Conf. Multimedia*, 2021, pp. 741–749.
- [105] D. Das, S. Yun, and F. Porikli, "ConFeSS: A framework for single source cross-domain few-shot learning," in *Proc. Int. Conf. Learn. Representations*, 2021.
- [106] Z. Weng, C. Feng, T. Zhang, Y. Zhu, and Z. Chen, "Representative multi-domain feature selection based cross-domain few-shot classification," in *Proc. 7th IEEE Int. Conf. Netw. Intell. Digit. Content*, 2021, pp. 86–90.
- [107] Y. Hu and A. J. Ma, "Adversarial feature augmentation for cross-domain few-shot classification," in *Proc. Eur. Conf. Comput. Vis.*, 2022, pp. 20–37.
- [108] M. Yazdanpanah and P. Moradi, "Visual domain bridge: A source-free domain adaptation for cross-domain few-shot learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2022, pp. 2868–2877.
- [109] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, no. 1, pp. 1–48, 2019.
- [110] W. Li, P. Fang, and H. Xue, "SVasP: Self-versatility adversarial style perturbation for cross-domain few-shot learning," 2024, arXiv: 2412.09073.
- [111] J. Zhang, J. Song, L. Gao, and H. Shen, "Free-lunch for cross-domain few-shot learning: Style-aware episodic training with robust contrastive learning," in *Proc. ACM Int. Conf. Multimedia*, 2022, pp. 2586–2594.
- [112] Q. Zhang, Y. Jiang, and Z. Wen, "TACDFSL: Task adaptive cross domain few-shot learning," *Symmetry*, vol. 14, no. 6, 2022, Art. no. 1097.
- [113] Z. Li, W. Wu, Y. Su, Z. Zheng, and M. R. Lyu, "CDTA: A cross-domain transfer-based attack with contrastive learning," in *Proc. Conf. Assoc. Advance. Artif. Intell.*, 2023, pp. 1530–1538.
- [114] A. Islam, C. R. Chen, R. Panda, L. Karlinsky, R. Feris, and R. J. Radke, "Dynamic distillation network for cross-domain few-shot recognition with unlabeled data," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2021, pp. 3584–3595.
- [115] P. Li et al., "Knowledge transduction for cross-domain few-shot learning," *Pattern Recognit.*, vol. 141, 2023, Art. no. 109652.
- [116] Y. Yang, T. Kim, and S.-Y. Yun, "Leveraging normalization layer in adapters with progressive learning and adaptive distillation for cross-domain few-shot learning," in *Proc. Conf. Assoc. Advance. Artif. Intell.*, 2024, pp. 16370–16378.
- [117] J. Oh, S. Kim, N. Ho, J. Kim, H. Song, and S. Yun, "Understanding cross-domain few-shot learning based on domain similarity and few-shot difficulty," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2022, pp. 2622–2636.
- [118] J. Wu, X. Liu, X. Yin, T. Zhang, and Y. Zhang, "Task-adaptive prompted transformer for cross-domain few-shot learning," in *Proc. Conf. Assoc. Advance. Artif. Intell.*, 2024, pp. 6012–6020.
- [119] F. Ji, X. Yuan, and Q. Liu, "Soft weight pruning for cross-domain few-shot learning with unlabeled target data," *IEEE Trans. Multimedia*, vol. 26, pp. 6759–6769, 2024.
- [120] B. Qin, S. Feng, C. Zhao, W. Li, R. Tao, and W. Xiang, "Cross-domain few-shot learning based on feature disentanglement for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5514215.
- [121] W. Wang, L. Duan, Y. Wang, J. Fan, and Z. Zhang, "MMT: Cross domain few-shot learning via meta-memory transfer," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 12, pp. 15018–15035, Dec. 2023.
- [122] J. Sun, S. Lapuschkin, W. Samek, Y. Zhao, N.-M. Cheung, and A. Binder, "Explanation-guided training for cross-domain few-shot classification," in *Proc. Int. Conf. Pattern Recognit.*, 2021, pp. 7609–7616.
- [123] P. Bateni, R. Goyal, V. Masrani, F. Wood, and L. Sigal, "Improved few-shot visual classification," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 14493–14502.
- [124] P. Li, S. Gong, C. Wang, and Y. Fu, "Ranking distance calibration for cross-domain few-shot learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2022, pp. 9099–9108.
- [125] H. Wang et al., "Experiments in cross-domain few-shot learning for image classification," in *Proc. Workshops Eur. Conf. Mach. Learn.*, 2022, pp. 81–83.
- [126] H. Xu et al., "Enhancing information maximization with distance-aware contrastive learning for source-free cross-domain few-shot learning," *IEEE Trans. Image Process.*, vol. 33, pp. 2058–2073, 2024.
- [127] T. Ma, Y. Sun, Z. Yang, and Y. Yang, "ProD: Prompting-to-disentangle domain knowledge for cross-domain few-shot image classification," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2023, pp. 19754–19763.
- [128] Z. Lin et al., "Diversifying cross-domain few-shot learning via multi-modal image editing," in *Proc. Int. Conf. Acoust. Speech Signal Process.*, 2024, pp. 7375–7379.
- [129] A. Aich, K.-C. Peng, and A. K. Roy-Chowdhury, "Cross-domain video anomaly detection without target domain adaptation," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, 2023, pp. 2579–2591.
- [130] C. P. Phoo and B. Hariharan, "Self-training for few-shot transfer across extreme task differences," in *Proc. Int. Conf. Learn. Representations*, 2021.
- [131] H. Chen, Y. Dong, Z. Lu, Y. Yu, and J. Han, "Pixel matching network for cross-domain few-shot segmentation," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, 2024, pp. 978–987.
- [132] T. Zhang, Q. Cai, F. Gao, L. Qi, and J. Dong, "Exploring cross-domain few-shot classification via frequency-aware prompting," 2024, arXiv: 2406.16422.
- [133] W.-H. Li, X. Liu, and H. Bilen, "Universal representation learning from multiple domains for few-shot classification," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2021, pp. 9526–9535.
- [134] Y. Zhao, H. Han, S. Shan, and X. Chen, "Deep subdomain alignment for cross-domain image classification," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, 2024, pp. 2820–2829.
- [135] M. Zhang, H. Liu, M. Gong, H. Li, Y. Wu, and X. Jiang, "Cross-domain self-taught network for few-shot hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 4501719.
- [136] Y. Zhao, T. Zhang, J. Li, and Y. Tian, "Dual adaptive representation alignment for cross-domain few-shot learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 10, pp. 11720–11732, Oct. 2023.
- [137] L. Sa, C. Yu, X. Ma, X. Zhao, and T. Xie, "Attentive fine-grained recognition for cross-domain few-shot classification," *Neural Comput. Appl.*, vol. 34, no. 6, pp. 4733–4746, 2022.
- [138] O. Vinyals, C. Blundell, T. Lillicrap, K. Kavukcuoglu, and D. Wierstra, "Matching networks for one shot learning," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2016, pp. 3630–3638.
- [139] M. Ren et al., "Meta-learning for semi-supervised few-shot classification," in *Proc. Int. Conf. Learn. Representations*, 2018.
- [140] A. Krizhevsky et al., "Learning multiple layers of features from tiny images," Univ. Toronto, Tech. Rep., 2009.
- [141] C. Wah, S. Branson, P. Welinder, P. Perona, and S. Belongie, "The caltech-UCSD birds-200–2011 dataset," Tech. Rep. CNS-TR-2010-001, 2011.
- [142] A. Radford et al., "Learning transferable visual models from natural language supervision," in *Proc. Int. Conf. Mach. Learn.*, 2021, pp. 8748–8763.
- [143] Y. Deng and X. Xiang, "Expanding hyperspherical space for few-shot class-incremental learning," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, 2024, pp. 1967–1976.
- [144] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "ChestX-Ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 2097–2106.
- [145] P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Sci. Data*, vol. 5, no. 1, pp. 1–9, 2018.
- [146] N. Codella et al., "Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (ISIC)," 2019, arXiv: 1902.03368.

- [147] P. Helber, B. Bischke, A. Dengel, and D. Borth, "EuroSAT: A novel dataset and deep learning benchmark for land use and land cover classification," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 12, no. 7, pp. 2217–2226, Jul. 2019.
- [148] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Front. Plant Sci.*, vol. 7, 2016, Art. no. 215232.
- [149] J. Krause, M. Stark, J. Deng, and L. Fei-Fei, "3D object representations for fine-grained categorization," in *Proc. Int. Conf. Comput. Vis. Workshops*, 2013, pp. 554–561.
- [150] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, and A. Torralba, "Places: A 10 million image database for scene recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 6, pp. 1452–1464, Jun. 2018.
- [151] G. Van Horn et al., "The inaturalist species classification and detection dataset," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 8769–8778.
- [152] B. Lake, R. Salakhutdinov, J. Gross, and J. Tenenbaum, "One shot learning of simple visual concepts," in *Proc. Annu. Meeting Cogn. Sci. Soc.*, 2011, pp. 2568–2573.
- [153] S. Maji, E. Rahtu, J. Kannala, M. Blaschko, and A. Vedaldi, "Fine-grained visual classification of aircraft," 2013, arXiv: 1306.5151.
- [154] M. Cimpoi, S. Maji, I. Kokkinos, S. Mohamed, and A. Vedaldi, "Describing textures in the wild," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 3606–3613.
- [155] R. Fernandez-Fernandez, J. G. Victores, D. Estevez, and C. Balaguer, "Quick, stat!: A statistical analysis of the quick, draw! dataset," 2019, arXiv: 1907.06417.
- [156] Visipedia, "FGVCx fungi classification challenge," 2018. [Online]. Available: [https://github.com/visipedia/fgvcx\\_fungi\\_comp](https://github.com/visipedia/fgvcx_fungi_comp)
- [157] M.-E. Nilsback and A. Zisserman, "Automated flower classification over a large number of classes," in *Proc. 6th Indian Conf. Comput. Vis. Graph. Image Process.*, 2008, pp. 722–729.
- [158] S. Houben, J. Stallkamp, J. Salmen, M. Schlipsing, and C. Igel, "Detection of traffic signs in real-world images: The german traffic sign detection benchmark," in *Proc. Int. Joint Conf. Neural Netw.*, 2013, pp. 1–8.
- [159] T.-Y. Lin et al., "Microsoft COCO: Common objects in context," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 740–755.
- [160] N. Inoue, E. Simo-Serra, T. Yamasaki, and H. Ishikawa, "Multi-label fashion image classification with minimal human supervision," in *Proc. Int. Conf. Comput. Vis. Workshops*, 2017, pp. 2261–2267.
- [161] Z. Barutcuoglu, R. E. Schapire, and O. G. Troyanskaya, "Hierarchical multi-label prediction of gene function," *Bioinformatics*, vol. 22, no. 7, pp. 830–836, 2006.
- [162] D. Heider, R. Senge, W. Cheng, and E. Hüllermeier, "Multilabel classification for exploiting cross-resistance information in HIV-1 drug resistance prediction," *Bioinformatics*, vol. 29, no. 16, pp. 1946–1952, 2013.
- [163] A. Alfassy et al., "LaSO: Label-set operations networks for multi-label few-shot learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 6548–6557.
- [164] T. Chen, L. Lin, R. Chen, X. Hui, and H. Wu, "Knowledge-guided multi-label few-shot learning for general image recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 3, pp. 1371–1384, Mar. 2022.
- [165] Y. An, H. Xue, X. Zhao, N. Xu, P. Fang, and X. Geng, "Leveraging bilateral correlations for multi-label few-shot learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 36, no. 4, pp. 6816–6828, Apr. 2025.
- [166] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," in *Proc. Int. Conf. Learn. Representations*, 2015.
- [167] A. Chakraborty, M. Alam, V. Dey, A. Chattopadhyay, and D. Mukhopadhyay, "A survey on adversarial attacks and defences," *CAAI Trans. Intell. Technol.*, vol. 6, no. 1, pp. 25–45, 2021.
- [168] L. Schmidt, S. Santurkar, D. Tsipras, K. Talwar, and A. Madry, "Adversarially robust generalization requires more data," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2018, pp. 5019–5031.
- [169] A. Shafahi et al., "Adversarially robust transfer learning," in *Proc. Int. Conf. Learn. Representations*, 2020.
- [170] M. Goldblum, L. Fowl, and T. Goldstein, "Adversarially robust few-shot learning: A meta-learning approach," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2020, pp. 17886–17895.
- [171] W. Li et al., "Defensive few-shot learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 5, pp. 5649–5667, May 2023.
- [172] A. Madry, A. Makelov, L. Schmidt, D. Tsipras, and A. Vladu, "Towards deep learning models resistant to adversarial attacks," in *Proc. Int. Conf. Learn. Representations*, 2018.
- [173] J. Guan, J. Liu, J. Sun, P. Feng, T. Shuai, and W. Wang, "Meta metric learning for highly imbalanced aerial scene classification," in *Proc. Int. Conf. Acoust. Speech Signal Process.*, 2020, pp. 4047–4051.
- [174] M. Buda, A. Maki, and M. A. Mazurowski, "A systematic study of the class imbalance problem in convolutional neural networks," *Neural Netw.*, vol. 106, pp. 249–259, 2018.
- [175] H. B. Lee et al., "Learning to balance: Bayesian meta-learning for imbalanced and out-of-distribution tasks," in *Proc. Int. Conf. Learn. Representations*, 2020.
- [176] M. Ochal, M. Patacchiola, J. Vazquez, A. J. Storkey, and S. Wang, "Few-shot learning with class imbalance," *IEEE Trans. Artif. Intell.*, vol. 4, no. 5, pp. 1348–1358, Oct. 2023.
- [177] M. Lazarou, Y. Avrithis, and T. Stathaki, "Adaptive manifold for imbalanced transductive few-shot learning," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, 2024, pp. 2286–295.



**Hui Xue** (Member, IEEE) received the BSc degree in mathematics from Nanjing Normal University in 2002, the MSc degree in mathematics from Nanjing University of Aeronautics & Astronautics (NUAA), in 2005, and the PhD degree in computer application technology from NUAU in 2008. She is currently a professor with the School of Computer Science and Engineering, Southeast University, China. Her research interests include pattern recognition and machine learning.



**Yuexuan An** (Member, IEEE) received the BSc degree from Jiangsu Normal University, Xuzhou, China, in 2015, the MSc degree from the China University of Mining and Technology, Xuzhou, in 2019, and the PhD degree from Southeast University, Nanjing, China, in 2024. She is currently an assistant professor with the College of Computer Science and Software Engineering, Hohai University, China. Her research interest includes machine learning and pattern recognition.



**Yongchun Qin** received the BSc in artificial intelligence from Southeast University, in 2023. He is currently working toward the MSc degree with the School of Computer Science and Engineering at Southeast University. His research interest includes machine learning and pattern recognition.



**Wenqian Li** received the BE degree in information science and engineering technology from Southeast University in 2023. She is currently working toward the PhD degree with the School of Computer Science and Engineering, Southeast University. Her research interest includes machine learning and pattern recognition.



**Yixin Wu** received the BSc degree in computer science and technology from the Nanjing University of Information Science & Technology, in 2022. He is currently working toward the MSc degree with the School of Computer Science and Engineering, Southeast University. His research interest includes machine learning and pattern recognition.



**Yongjuan Che** received the BSc degree in computer science and technology from Hohai University, in 2021 and the MSc degree in software engineering from Southeast University, in 2024. Her research interest includes machine learning and pattern recognition.



**Pengfei Fang** received the PhD degree from the Australian National University and DATA61-CSIRO in 2022, and the ME degree from the Australian National University in 2017. He is a professor with the School of Computer Science and Engineering, Southeast University (SEU), China. Before joining SEU, he was a postdoctoral fellow with Monash University in 2022. His research interests include computer vision and machine learning.



**Min-Ling Zhang** (Senior Member, IEEE) received the BSc, MSc, and PhD degrees in computer science from Nanjing University, China, in 2001, 2004, and 2007, respectively. Currently, he is a professor with the School of Computer Science and Engineering, Southeast University, China. His main research interests include machine learning and data mining. In recent years, he has served as the general co-chairs of ACML'18, Program co-chairs of CCML'25, PAKDD'19, CCF-ICAI'19, ACML'17, CCF-ICAI'17, PRICAI'16, Senior PC member or area chair of KDD 2021-2024, AAAI 2022-2025, IJCAI 2017-2024, ICML 2024, ICLR 2024, etc. He is also on the editorial board of *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *Science China Information Sciences*, *ACM Transactions on Intelligent Systems and Technology*, *Frontiers of Computer Science*, *Machine Intelligence Research*, etc. He is the Steering Committee member of ACML and PAKDD, vice-chair of the CAAI (Chinese Association of Artificial Intelligence) Machine Learning Society. He is a distinguished member of CCF, CAAI, and senior member of AAAI and ACM.