

Article

Group-Aware Spectral Perturbation Consistency Network for Cross-Domain Few-Shot Hyperspectral Image Classification

Xun Luo, Wenjie Xu *  and Yueyan Jiang

College of Information Science and Engineering, Hunan Normal University, Changsha 410081, China; luoxun@tsinghua.org.cn (X.L.); 202570294426@hunnu.edu.cn (Y.J.)

* Correspondence: xuwenjie@hunnu.edu.cn

Highlights

What are the main findings?

- G-SPCNet consistently improves the robustness of cross-domain few-shot hyperspectral image classification under spectral-band masking, dropout, noise, and spectral-axis shifts.
- Experiments across four cross-domain tasks demonstrate superior Robust OA and Kappa compared with representative state-of-the-art baselines.

What are the implications of the main findings?

- Group-aware spectral perturbation consistency enables few-shot HSI models to maintain stable predictions and prototype structures under spectral disturbances.
- The proposed lightweight framework provides a practical solution for robust hyperspectral recognition when target-domain annotations are scarce and spectral conditions are unreliable.

Abstract

Cross-domain few-shot hyperspectral image classification is challenged by scarce target-domain annotations, sensor-induced domain shifts, and spectral abnormalities, including band loss, noise, and spectral misalignment. To address these issues, we propose a Group-aware Spectral Perturbation Consistency Network (G-SPCNet) and evaluate both Clean classification and Robust perturbation stability under a unified protocol. G-SPCNet divides the spectrum into contiguous band groups and employs lightweight convolutions and bidirectional spectral-context modeling to construct grouped class prototypes. During source-domain meta-training, contiguous band masking, random band dropout, noisy-band injection, and spectral-axis shifting are applied online to form clean and perturbed views. Classification and predictive-distribution consistency objectives improve global robustness, while group-level prediction consistency and clean-prototype anchoring stabilize local decisions and prototype structures. In five-way five-shot experiments using LongKou as the source domain and Indian Pines, PaviaU, Salinas, and Chikusei as target domains, G-SPCNet achieved average Clean OA/AA, Robust OA/AA, and Robust Kappa values of 87.02%, 79.47%, and 74.34%, respectively. It outperformed seven comparison methods in the main Robust evaluation and maintained the highest average OA among the five evaluated methods at all tested support-set sizes in the shot-sensitivity analysis and at all tested severity levels in the mixed perturbation analysis. With only 0.153 M parameters, G-SPCNet provides a lightweight accuracy–robustness trade-off for cross-domain hyperspectral image classification.



Academic Editor: Salah Bourennane

Received: 26 August 2026

Revised: 26 September 2026

Accepted: 2 October 2026

Published: 7 October 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)

[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: consistency learning; cross-domain few-shot classification; grouped prototype; hyperspectral image; robust classification; spectral perturbation

1. Introduction

Hyperspectral images record dozens to hundreds of contiguous narrow spectral bands at each pixel, enabling discrimination among materials that are difficult to distinguish in multispectral images. They are therefore widely used in applications such as precision agriculture. As remote sensing platforms evolve, practical applications increasingly involve imagery acquired across different regions, times, and sensors. Reliable classification in target scenes with limited annotations and varying imaging conditions is therefore a key challenge in hyperspectral image analysis.

Early hyperspectral image classification mainly relied on shallow models, including support vector machines, random forests, sparse representation, kernel methods, manifold learning, and morphological profiles. These methods generally combined band selection, dimensionality reduction, or manually designed spectral–spatial descriptors to alleviate the curse of dimensionality. They offered certain advantages when only limited annotations were available. Nevertheless, their performance depended heavily on feature engineering, and they often lacked sufficient capacity to represent nonlinear spectral differences and spatial contextual relationships among complex ground objects [1].

The development of deep learning has gradually shifted hyperspectral image classification from manually designed features toward end-to-end representation learning. One-dimensional, two-dimensional, and three-dimensional convolutional networks extract features from spectral sequences, spatial textures, and joint spectral–spatial cubes, respectively. Architectures such as spectral–spatial residual networks and HybridSN further enhance local discriminability through residual connections and hybrid convolutions [2,3]. Multiscale convolution, channel and spatial attention, dual-branch networks, and autoencoder-based reconstruction have subsequently been used to improve classification under blurred boundaries, limited annotation ratios, and spectral redundancy [4].

To overcome the restricted receptive fields of local convolutions, graph convolutional networks organize pixels or samples into graph structures [5]. Transformers employ self-attention to capture long-range spectral–spatial dependencies, whereas state-space models and Mamba-like architectures model long spectral sequences with relatively low computational complexity. Although these methods improve hyperspectral feature representation, they also increase model size and supervision requirements. When only a few labeled samples are available in the target domain, direct fine-tuning is prone to overfitting. Moreover, discriminative patterns learned from the source domain may not transfer effectively to new sensors or imaging scenes.

To address annotation scarcity, metric-learning methods such as Siamese networks, matching networks, relation networks, and prototypical networks employ episodic training to simulate N -way K -shot tasks. These methods learn class references from the support set and classify query samples according to feature distances [6–11]. Local prototypes, multiscale prototypes, covariance-based metrics, task adaptation, and pseudo-label learning have subsequently been introduced to reduce prototype estimation bias in few-shot settings [12–14]. Meanwhile, data augmentation methods, including Mixup [15], Cut-Mix [16], and generative adversarial augmentation [17], enrich the training data and can improve model generalization. Self-supervised pretext tasks [18] and contrastive learning [19] provide supervisory signals from unlabeled data, helping reduce reliance on manual annotations.

When the training and testing scenes differ, research increasingly focuses on cross-domain generalization and domain adaptation [20,21]. Maximum mean discrepancy, correlation alignment, subspace mapping, and domain-adversarial learning have been used to reduce distribution differences between source and target domains [22–24]. Feature disentanglement attempts to separate domain-invariant semantics from domain-specific factors [25]. Relation learning, prototype calibration, and uncertainty estimation are also used to alleviate class-center shifts caused by insufficient target-domain support samples [26–28]. More recently, frequency-domain modeling, vision–language prompting, pretrained foundation models, and parameter-efficient adaptation have been introduced into hyperspectral tasks to improve cross-scene transferability [29,30].

Beyond standard classification accuracy, studies on hyperspectral model reliability have investigated adversarial examples, sensor noise, band selection, bad-band restoration, spectral reconstruction, and missing-data completion. Existing research indicates that denoising, self-supervised reconstruction, and perturbation-based augmentation can reduce the influence of local abnormalities on classifiers [31]. Most studies, however, focus on fully supervised settings within a single domain or consider only one type of noise. Under cross-domain few-shot conditions, the changes in class prototypes, inter-group relationships, and distance boundaries when both support and query samples are affected by spectral-band abnormalities remain insufficiently studied.

In practical applications, differences in sensor spectral responses, residual atmospheric-correction errors, low-signal-to-noise-ratio bands, local detector-element abnormalities, and slight shifts in central wavelengths may appear as missing information in contiguous bands, random band failures, noisy bands, or spectral-axis misalignment. For prototype-based classifiers, these abnormalities affect not only query features but also class centers estimated from a small number of support samples. The resulting errors may therefore be amplified in few-shot settings. Consequently, accuracy measured only on clean target-domain data cannot fully reflect a model’s practical deployment capability.

Based on these considerations, this study conducts both Clean and Robust evaluations under a fixed-source, multiple-target-domain protocol. Clean evaluation measures standard classification performance, whereas Robust evaluation uses controllable compound spectral-band perturbations to examine performance retention under missing bands, noise, and misalignment. The Robust evaluation provides an interpretable and reproducible stress test that approximates possible changes in spectral-band quality during cross-sensor deployment. Reporting both types of results enables a more comprehensive assessment of the balance between classification accuracy and robustness.

To this end, we propose a Group-aware Spectral Perturbation Consistency Network (G-SPCNet). Rather than relying on a large-scale pretrained backbone, G-SPCNet uses the contiguous spectral structure of hyperspectral data and a lightweight grouped-prototype framework to learn discriminative representations for local spectral regions. During source-domain meta-training, clean and perturbed views are generated online so that the model learns feature representations that remain stable under missing bands, noise contamination, and spectral-axis shifts. Perturbation responses are further constrained at the group-level prediction and prototype levels, reducing the adverse effect of locally abnormal spectral regions on overall class distances. The main contributions of this study are summarized as follows:

1. A band-grouped representation (BGR) module is developed by combining contiguous band grouping, lightweight intra-group encoding, and bidirectional spectral-context modeling. The module controls model size while extracting local spectral–spatial features and preserving contextual information among different band groups.

2. A spectral perturbation consistency (SPC) mechanism is developed to address spectral-band abnormalities. Multiple spectral perturbations are simulated online during source-domain meta-training, and robust representations are learned through dual-view classification and consistency constraints on the final predictive distributions.
3. A group-aware consistency (GAC) mechanism is developed to jointly constrain the group-level logits of clean and perturbed views. Clean prototypes processed with stop-gradient serve as stable anchors, reducing perturbation-induced group-level prediction drift and changes in prototype structure.
4. Cross-domain few-shot spectral perturbation experiments are conducted using one fixed source domain and four target domains. Both Clean and Robust performance are reported, and the effectiveness of the proposed method is further examined through shot-sensitivity and mixed-perturbation-severity experiments.

2. Materials and Methods

2.1. Proposed Method

To address the challenges of cross-domain few-shot hyperspectral image classification, this study proposes G-SPCNet. The network is designed to reduce differences in imaging scenes, sensor responses, and spectral distributions between the source and target domains while improving classification stability under missing bands, noise contamination, and spectral-axis misalignment.

G-SPCNet consists of a band-grouped representation module (BGR), a spectral perturbation consistency mechanism (SPC), and a group-aware consistency mechanism (GAC). BGR divides the spectral bands into groups according to their continuity along the spectral axis. It preserves group-level features and constructs class-band-group prototypes, providing structured representations for grouped distance-based classification and subsequent consistency constraints. During source-domain training, SPC generates clean and perturbed views online using contiguous band masking, random band dropout, noisy-band injection, and spectral-axis shifting. Dual-view classification and consistency learning between the final predictive distributions are then used to obtain features that remain relatively stable under spectral-band variations.

On this basis, GAC constrains the two views at both the group-level prediction and class-prototype levels, thereby reducing prediction drift and prototype shifts caused by local spectral perturbations. The three components are jointly optimized in an end-to-end manner, forming a learning process that integrates band-grouped representation, spectral perturbation modeling, and multilevel consistency constraints. After training, only the support samples of the target task are used to construct grouped class prototypes, and the query samples are classified through metric-based inference.

2.1.1. Overall Framework

As illustrated in Figure 1, G-SPCNet employs dual-view learning during training and single-branch inference during testing. The network consists of four main components: spectral-perturbation-based dual-view construction, BGR-based band-grouped representation, grouped prototype classification, and multilevel consistency constraints. Each input hyperspectral sample is first resampled to B spectral bands and then partitioned into G contiguous groups according to the original spectral order. During source-domain meta-training, clean and perturbed views are generated from the same few-shot task. The clean view retains the original support and query sets, whereas the perturbed view is produced through online spectral-band perturbations. Both pathways share the BGR encoder and grouped prototype classifier.

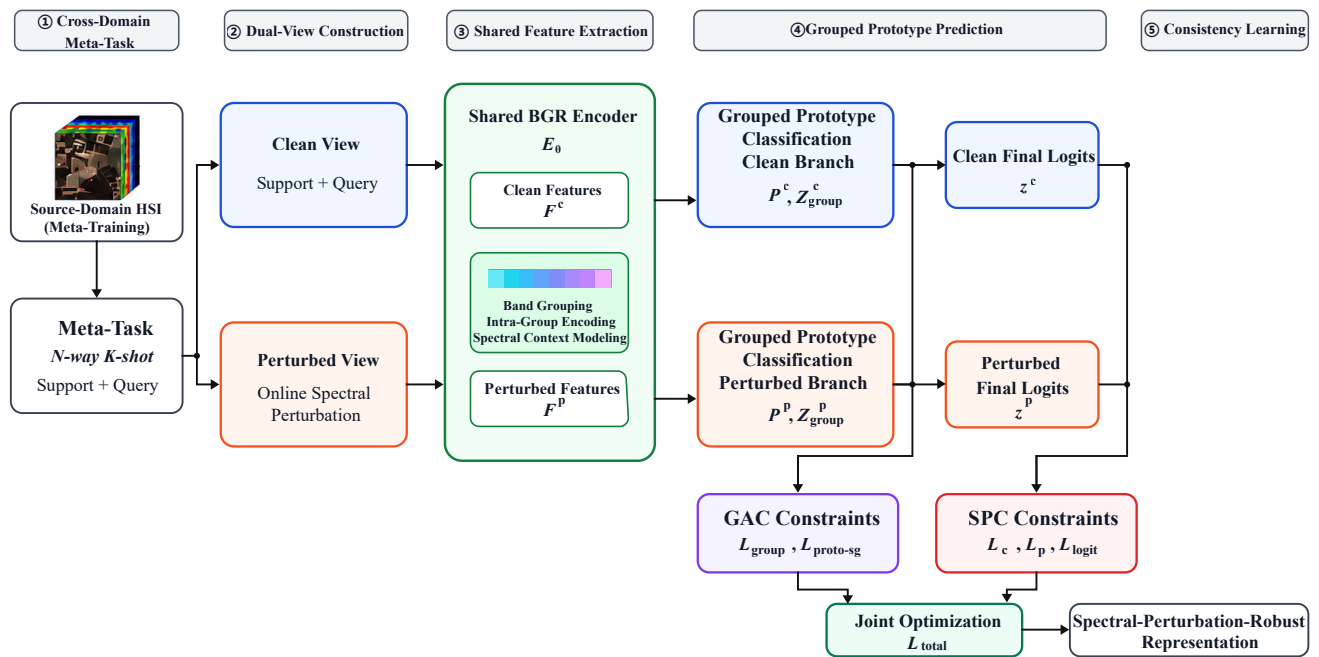


Figure 1. Architecture of the proposed G-SPCNet.

Let a training meta-task sampled from the source domain be denoted by $\tau_s = (S_s, y_s^S, Q_s, y_s^Q)$, where S_s and Q_s denote the support and query sets, respectively, and y_s^S and y_s^Q are their corresponding class labels. For a view $v \in \{c, p\}$, the shared BGR encoder E_θ extracts the support features F_s^v and query features F_Q^v , whose dimensions are $F_s^v \in \mathbb{R}^{N_s \times G \times D}$, $F_Q^v \in \mathbb{R}^{N_Q \times G \times D}$. BGR employs independent intra-group encoders to extract local features from individual band groups and uses a bidirectional gated recurrent unit to establish contextual relationships among the groups, thereby preserving feature representations with an explicit grouped structure.

During classification, the grouped prototype classifier constructs a prototype for each class within each band group from the support features. The squared Euclidean distance is then used to calculate the group-level logits of the query samples. The results from the G band groups are equally weighted and combined to obtain the final classification logits. The overall training objective is defined as

$$L = L_c + \lambda_p L_p + \lambda_l L_{logit} + \lambda_g L_{group} + \lambda_r L_{proto-sg}, \quad (1)$$

where L_c and L_p are the classification losses for the clean and perturbed views, respectively; L_{logit} constrains the final predictive distributions before and after perturbation; L_{group} aligns the classification responses of individual band groups; and $L_{proto-sg}$ uses the stop-gradient clean prototypes as anchors to constrain structural changes in the perturbed prototypes. The coefficients λ_p , λ_l , λ_g , and λ_r control the contributions of the perturbed-view classification loss, final-logit consistency loss, group-level logit consistency loss, and stop-gradient prototype consistency loss, respectively. The weights are fixed, non-negative hyperparameters and are not constrained to sum to one. The coefficient of the clean-view classification loss L_c is set to one. The specific weight values used in our experiments are reported in Section 2.2.4. During source-domain meta-training, the encoder parameters θ are optimized end to end using the overall objective.

During target-domain testing, the model neither constructs clean and perturbed views nor performs target-domain self-training or parameter updates. For each target-domain meta-task, the support samples are used to construct grouped class prototypes, and the

query samples are classified by equally averaging their distances to the class prototypes across all band groups. Under the Clean setting, the original target-domain samples are used directly as inputs. Under the Robust setting, the externally perturbed support and query sets are processed as a single input view, without computing any consistency losses during inference. Thus, the multilevel consistency constraints are applied only during source-domain training and incur no additional dual-branch computational cost at test time.

2.1.2. Band-Grouped Representation Module

Figure 2 illustrates the detailed architecture of the band-grouped representation (BGR) module. Adjacent bands in hyperspectral data are often strongly correlated, while spectrally distant bands may also exhibit strong correlations. Different spectral regions may provide complementary information for class discrimination. BGR extracts local spectral-spatial features from contiguous band groups without assuming inter-group independence. A bidirectional gated recurrent unit then models cross-group dependencies to integrate information across spectral regions. Unlike approaches that fuse all bands at the initial encoding stage, this strategy preserves local differences among spectral regions and reduces the direct effect of spectral abnormalities or local domain shifts on the overall feature representation. It thus provides finer-grained features for constructing class prototypes in cross-domain tasks [32].

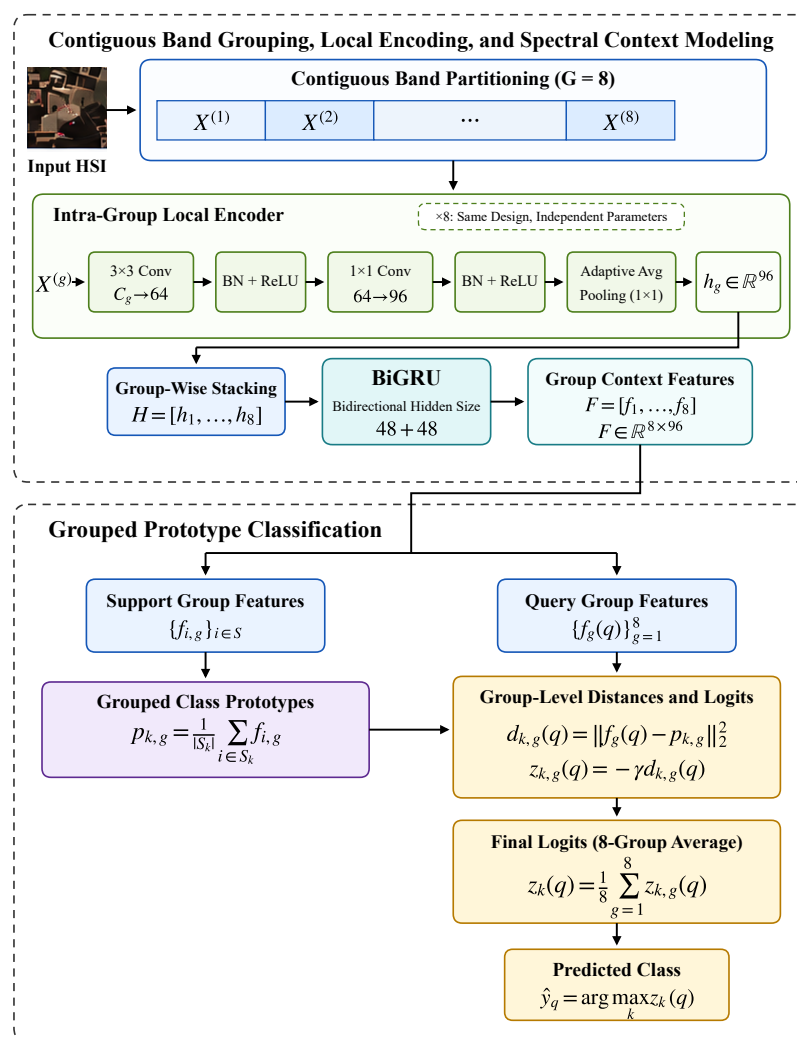


Figure 2. Band-grouped representation module.

Let the input hyperspectral image patch be $X \in \mathbb{R}^{P \times P \times B}$, where P denotes the spatial neighborhood size and B denotes the unified number of spectral bands. The band-group boundaries are defined as $0 = b_0 < b_1 < \dots < b_G = B$. The index set of the g -th contiguous band group is $I_g = \{b_{g-1} + 1, \dots, b_g\}$, and its corresponding subcube is denoted by $X^g = X(:, :, I_g)$.

To provide a common input dimensionality across datasets with different native band counts, each spectrum is resampled to 100 spectral channels before entering BGR. This dimensionality is an experimental configuration rather than a sensor-specific requirement or a theoretically optimal choice. The resampled channels are divided approximately evenly into eight contiguous groups, whose sizes differ by at most one channel. This grouping includes all resampled channels in subsequent feature extraction. For each subcube X^g , the intra-group encoder sequentially applies a 3×3 convolution, batch normalization, and ReLU activation, expanding the number of channels from C_g to 64. A 1×1 convolution, batch normalization, and ReLU activation are then used to expand the channel dimension to 96. Global adaptive average pooling subsequently produces a local vector $h_g \in \mathbb{R}^{96}$. The group encoders share the same architecture but have independent parameters. The local vectors are arranged in spectral order as $H = [h_1, \dots, h_G]$ and are fed into a bidirectional GRU. Both the forward and backward hidden dimensions are set to 48. Concatenating the two directional states produces the group-level contextual features $F = [f_1, \dots, f_G] \in \mathbb{R}^{8 \times 96}$. Therefore, each f_g retains the local information of the current band group while incorporating contextual information provided by the preceding and subsequent groups [33,34].

During grouped prototype classification, this study follows the basic idea of prototypical networks, in which class prototypes are constructed by averaging support features [8]. This principle is further extended to each class k and band group g , and the corresponding group-level class prototype is defined as

$$p_{k,g} = \frac{1}{|S_k|} \sum_{i \in S_k} f_{i,g}, \quad (2)$$

where S_k denotes the index set of support samples belonging to class k , and $f_{i,g}$ is the feature of the i -th support sample in the g -th band group. For a query sample q , its grouped feature is denoted by $f_g(q)$, and the corresponding group-level classification logit is calculated as:

$$z_{k,g} = -\gamma \|f_g(q) - p_{k,g}\|_2^2, \quad (3)$$

where γ is the distance-scaling coefficient. The classification results from the eight band groups are combined with equal weights to obtain the final classification logit:

$$z_k(q) = \frac{1}{8} \sum_{g=1}^8 z_{k,g}(q). \quad (4)$$

A single global prototype jointly depends on all spectral bands. Consequently, a domain shift or quality degradation within a local spectral region may alter the entire class center. In contrast, BGR constructs a separate class reference for each band group, thereby localizing the effects of inter-domain differences and spectral abnormalities. The BiGRU incorporates contextual information from adjacent band groups, while equal-weight fusion limits the contribution of any single abnormal group to the final prediction. BGR is therefore expected to reduce prototype shifts caused by local spectral variations and improve the stability of group-level prototypes and final predictions under cross-domain conditions.

2.1.3. Spectral Perturbation Consistency Mechanism

Figure 3 illustrates the workflow of the spectral perturbation consistency (SPC) mechanism. During source-domain meta-training, clean and perturbed views are constructed from the same support and query sets. The clean view retains the original spectral information, whereas the perturbed view is generated using online compound spectral-band perturbations. Both views share the BGR encoder and grouped prototype classifier and use identical sample labels.

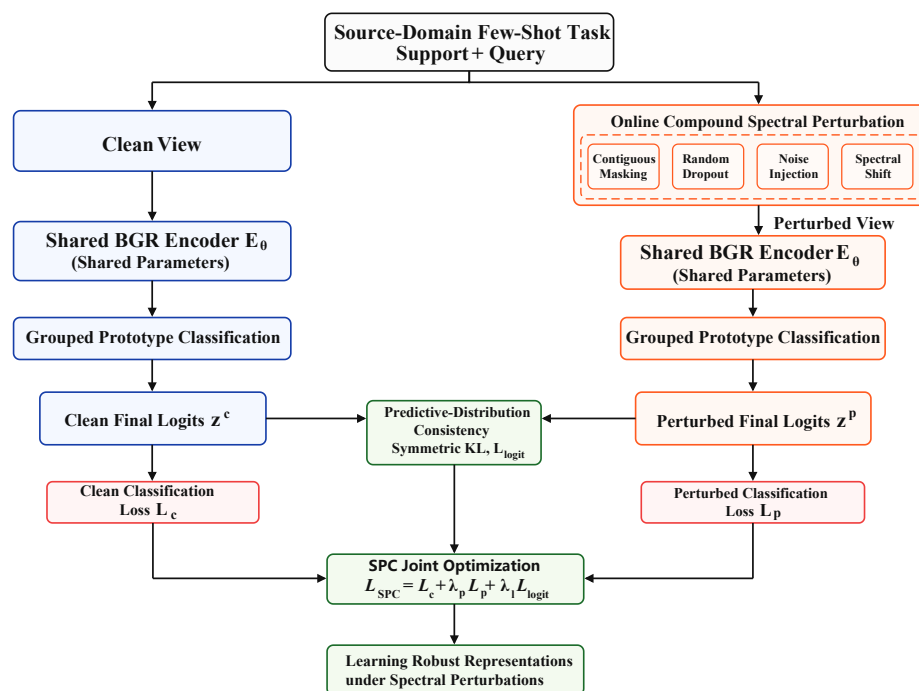


Figure 3. Flowchart of spectral perturbation consistency.

Online compound perturbations comprise contiguous band masking, random band dropout, noisy-band injection, and spectral-axis shifting. Masking zeros a contiguous interval with a uniformly sampled starting position, while dropout zeros bands with a specified per-band probability, simulating contiguous and scattered spectral information loss. Noise injection adds zero-mean Gaussian noise to randomly selected bands, with the standard deviation defined in the standardized input space. Spectral shifting samples an integer offset uniformly from a symmetric range and zero-fills exposed bands, approximating spectral-axis misalignment. These perturbations are generated online during meta-task training without storing augmented samples [31]. Exposure to these variations reduces reliance on fixed spectral regions.

Let the final classification logits of the two views be denoted by z^c and z^p , with their respective classification losses denoted by L_c and L_p . Let $\pi^v = \text{softmax}(z^v/T)$, $v \in \{c, p\}$. The final-logit consistency loss is defined as

$$L_{\text{logit}} = \frac{T^2}{2} [D_{\text{KL}}(\text{sg}(\pi^p) \parallel \pi^c) + D_{\text{KL}}(\text{sg}(\pi^c) \parallel \pi^p)], \quad (5)$$

where $\text{sg}(\cdot)$ denotes the stop-gradient operation. The optimization objective of SPC is

$$L_{\text{SPC}} = L_c + \lambda_p L_p + \lambda_l L_{\text{logit}}, \quad (6)$$

where λ_p and λ_l are the weights assigned to the perturbed-view classification loss and final-logit consistency loss, respectively.

Conventional perturbation-based augmentation mainly requires perturbed samples to be classified correctly but does not directly constrain the relationships among class probabilities before and after perturbation. In addition to preserving the classification ability of both views, SPC employs bidirectional KL divergence to reduce changes in their final predictive distributions, thereby limiting shifts in the overall classification boundaries caused by spectral abnormalities [35]. SPC is therefore expected to improve Robust performance under compound spectral-band perturbations while largely preserving Clean classification performance and reducing accuracy degradation as perturbation severity increases. The stability of individual band-group predictions and the structure of the class prototypes are further constrained by the subsequent GAC mechanism.

2.1.4. Group-Aware Consistency Mechanism

Figure 4 illustrates the workflow of the group-aware consistency (GAC) mechanism. SPC primarily constrains the final predictive distributions of the clean and perturbed views. However, the final logits are obtained by equally averaging the classification outputs of all band groups, allowing prediction deviations from different groups to cancel each other during fusion. Consequently, constraints applied only to the final output may fail to reveal substantial changes in the classification response of a particular spectral region. To address this limitation, GAC exploits the grouped structure preserved by BGR and aligns model responses before and after perturbation at both the group-level prediction and class-prototype levels. This design enables the model to capture discriminative changes caused by local spectral perturbations.

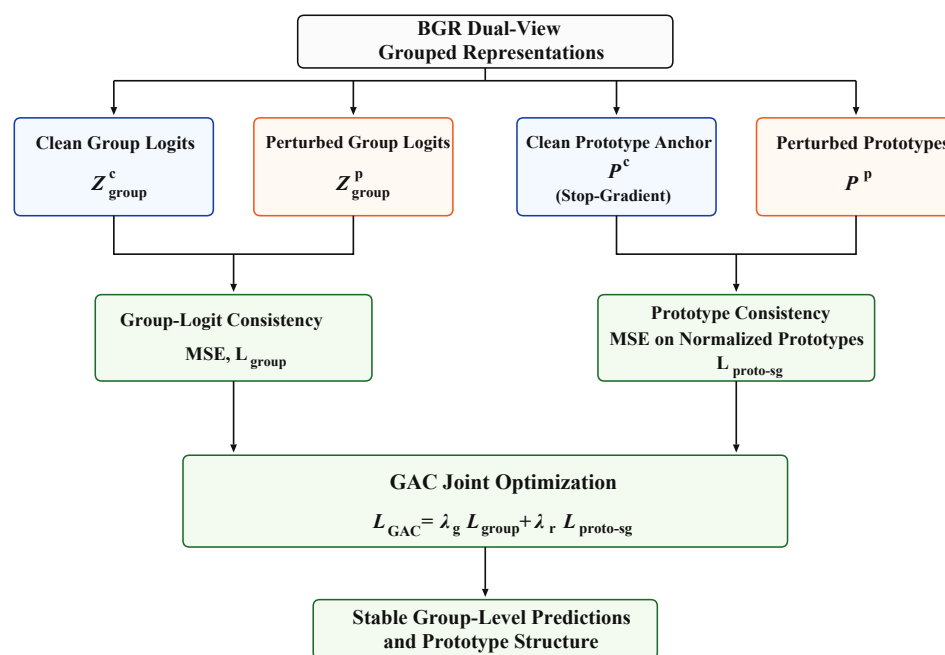


Figure 4. Flowchart of group-aware consistency.

Let Z_{group}^c and Z_{group}^p denote the group-level classification logits of the clean and perturbed views, respectively. The group-level logit consistency loss is defined as

$$L_{group} = \text{MSE}(Z_{group}^c, Z_{group}^p). \quad (7)$$

This loss compares the classification responses of the two views for each band group and class, thereby preventing large prediction shifts in individual groups caused by masking, noise, or spectral-axis shifts. Compared with final-logit consistency, L_{group} preserves

the independent discriminative information of each band group, so local abnormalities are not completely concealed by the equally weighted fusion. In addition to affecting classification outputs, spectral perturbations may alter the class prototypes constructed from a limited number of support samples. Let P^c and P^p denote the prototype sets obtained from the clean and perturbed views, respectively. The group-level prototypes of each class are flattened and L_2 -normalized to obtain $\bar{P}^c$ and $\bar{P}^p$. The stop-gradient prototype consistency loss is defined as

$$L_{proto-sg} = MSE(sg(\bar{P}^c), \bar{P}^p), \quad (8)$$

where $sg(\cdot)$ denotes the stop-gradient operation. Normalization reduces the influence of prototype-magnitude differences, allowing the loss to focus more on consistency in prototype structure and feature direction. The stop-gradient operation treats the clean prototypes as relatively stable anchors that do not receive gradients from this loss, while only encouraging the perturbed prototypes to approach them. This avoids mutual attraction between the two branches during the early stage of training [36]. Therefore, the joint optimization objective of GAC is

$$L_{GAC} = \lambda_g L_{group} + \lambda_r L_{proto-sg}. \quad (9)$$

BGR provides GAC with group-level features, class prototypes, and classification outputs that retain explicit band-group membership. SPC generates perturbed views and stabilizes the final predictive distributions, whereas GAC further constrains variations in local group-level predictions and class prototypes. Although the improvement introduced by GAC is expected to be smaller than that provided by SPC, GAC can further reduce prediction drift and prototype shifts caused by local spectral perturbations, thereby yielding consistent gains in Robust performance.

2.1.5. Algorithmic Procedure

The source-domain meta-training and target-task prototype inference procedures of G-SPCNet are presented in Algorithm 1.

After training, the checkpoint selected using the held-out source-domain validation set is restored, and the network parameters remain fixed during target-domain inference. For each target episode, the support and query samples are passed through the BGR encoder. Grouped class prototypes are constructed from the support features using the target support labels, and the query samples are classified by equally averaging the group-wise distances. The target support set provides class references for the current task without updating the network parameters.

The robustness constraints are applied exclusively during training, while the inference architecture remains as simple as a standard grouped-prototype network. During inference, paired clean and perturbed views are not constructed, and the final-logit consistency, group-level logit consistency, and stop-gradient prototype consistency losses are not computed. This design simplifies the deployment procedure and facilitates practical application.

2.2. Experimental Materials

2.2.1. Source-Domain Dataset

The WHU-Hi-LongKou dataset is used as the source domain in this study. The LK (LongKou) dataset was acquired using a Headwall Nano-Hyperspec sensor mounted on a DJI Matrice 600 Pro unmanned aerial vehicle. It has a spatial size of 550×400 pixels and contains 270 spectral bands, covering farmland, roads, water bodies, and various crop types. Figure 5 presents the pseudo-color image and ground-truth map of LK, while Table 1 lists the classes and corresponding sample numbers.

Algorithm 1: Training and Inference Procedures for G-SPCNet

Input: Source-domain training episode distribution D_s^{tr} ; held-out source-domain validation episode set D_s^{val} ; target test task $\tau_t = (S_t, y_t^S, Q_t)$; number of band groups G , distance scaling factor γ , temperature T , and spectral perturbation operator A_ρ ; training controls E_{max} , $E_{\text{warm}} = 1000$, $E_{\text{val}} = 100$, and $P_{\text{es}} = 10$; loss weights λ_p , λ_l , λ_g , and λ_r .

Output: Network parameters θ^* from the checkpoint selected using the held-out source-domain validation set, and predicted target query labels $\hat{y}_t^Q$.

Procedure:

1. Initialize the shared BGR encoder E_θ and the AdamW optimizer, and set the weight of each band group to $w_g = 1/G$.
2. For $e = 1$ to E_{max} do, where e denotes the training-episode index rather than the conventional epoch index.
3. Sample a source-domain task $\tau_s = (S_s, y_s^S, Q_s, y_s^Q)$ from D_s^{tr} .
4. Set $S_s^c = S_s$ and $Q_s^c = Q_s$, and generate $S_s^p = A_{\rho_S}(S_s)$ and $Q_s^p = A_{\rho_Q}(Q_s)$ online.
5. For $v \in \{c, p\}$, compute $F_S^v = E_\theta(S_s^v)$ and $F_Q^v = E_\theta(Q_s^v)$, whose shapes are $N_S \times G \times D$ and $N_Q \times G \times D$, respectively.
6. Define the class-specific index set $\mathcal{I}_{s,k} = \{i \mid y_{s,i}^S = k\}$, and compute $p_{k,g}^v = \frac{1}{|\mathcal{I}_{s,k}|} \sum_{i \in \mathcal{I}_{s,k}} F_{S,i,g}^v$, where $v \in \{c, p\}$.
7. Compute $z_{q,k,g}^v = -\gamma \|F_{Q,q,g}^v - p_{k,g}^v\|_2^2$, and set $z_{q,k}^v = \frac{1}{G} \sum_{g=1}^G z_{q,k,g}^v$.
8. Compute the classification losses $L_c = \text{CE}(z^c, y_s^Q)$ and $L_p = \text{CE}(z^p, y_s^Q)$.
9. Set $\pi^v = \text{softmax}(z^v/T)$, and compute $L_{\text{logit}} = \frac{T^2}{2} [\text{KL}(\text{sg}(\pi^p) \parallel \pi^c) + \text{KL}(\text{sg}(\pi^c) \parallel \pi^p)]$, $L_{\text{group}} = \text{MSE}(Z_{\text{group}}^c, Z_{\text{group}}^p)$, and $L_{\text{proto-sg}} = \text{MSE}(\text{sg}(\bar{P}^c), \bar{P}^p)$, where $\text{sg}(\cdot)$ denotes the stop-gradient operator, and the prototypes marked with a bar are flattened by class and L_2 -normalized.
10. Compute the total loss: $L = L_c + \lambda_p L_p + \lambda_l L_{\text{logit}} + \lambda_g L_{\text{group}} + \lambda_r L_{\text{proto-sg}}$.
11. Backpropagate the total loss on the source-domain task and update θ using the AdamW optimizer.

Validation-Based Model Selection

12. Starting at training episode E_{warm} , evaluate the model on the held-out source-domain validation set D_s^{val} every E_{val} training episodes and calculate $s_{\text{val}} = \frac{1}{2} (\text{OA}_{\text{Clean}}^{s,\text{val}} + \text{OA}_{\text{Robust}}^{s,\text{val}})$; if the validation score improves, save $\theta^* \leftarrow \theta$, and terminate training early if no improvement is observed in P_{es} consecutive evaluations.
13. End for; restore θ^* from the checkpoint selected using the held-out source-domain validation set.

Target-Task Inference

14. For each target task $\tau_t = (S_t, y_t^S, Q_t)$ do.
15. Compute $F_{S,t} = E_{\theta^*}(S_t)$ and $F_{Q,t} = E_{\theta^*}(Q_t)$.
16. Construct the grouped class prototypes $p_{t,k,g}$ according to the target support labels y_t^S .
17. For each target query sample q , compute its fused classification logit for class k :

$$z_{t,q,k} = -\frac{\gamma}{G} \sum_{g=1}^G \|F_{Q,t,q,g} - p_{t,k,g}\|_2^2$$
 Obtain the corresponding predicted label:

$$\hat{y}_{t,q}^Q = \arg \max_k z_{t,q,k}.$$
18. End for; return θ^* and $\hat{y}_t^Q$.

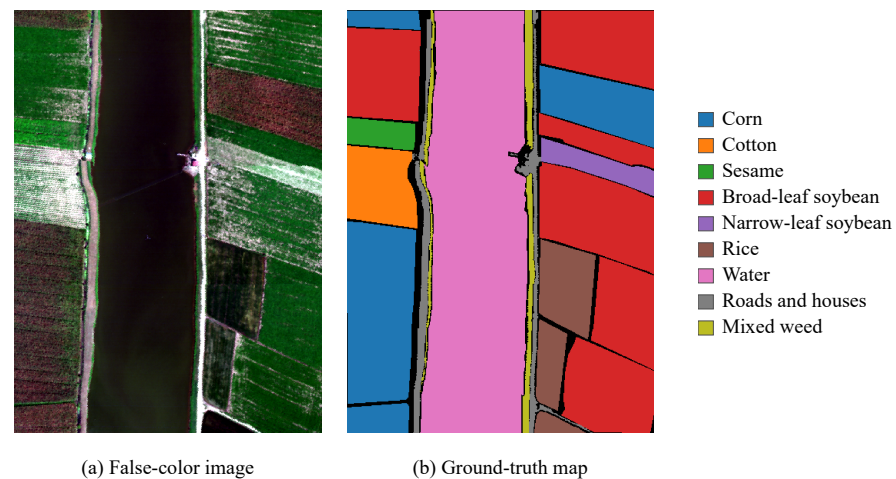


Figure 5. Representative map of the LK dataset.

Table 1. Class names and sample numbers of LK.

Class	Class Name	Number of Samples
1	Corn	34,511
2	Cotton	8374
3	Sesame	3031
4	Broad-leaf soybean	63,212
5	Narrow-leaf soybean	4151
6	Rice	11,854
7	Water	67,056
8	Roads and houses	7124
9	Mixed weed	5229
Total Number of Samples		204,542

2.2.2. Target-Domain Datasets

The target domains comprise four publicly available datasets: IP (Indian Pines), PU (Pavia University), SA (Salinas), and CK (Chikusei). IP and SA were acquired using the AVIRIS sensor, PU using the ROSIS sensor, and CK using a Headwall Hyperspec-VNIR-C sensor. These datasets differ from LK in spatial resolution, land-cover classes, imaging scenes, and spectral responses, making them suitable for evaluating the robust generalization ability of the model under cross-domain conditions. Figures 6–9 show the pseudo-color images and ground-truth maps of the four target domains, and Tables 2–5 present their corresponding sample distributions.

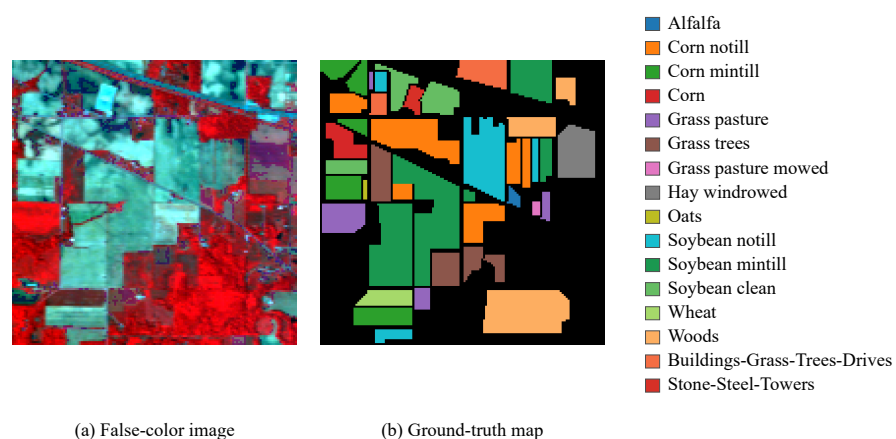
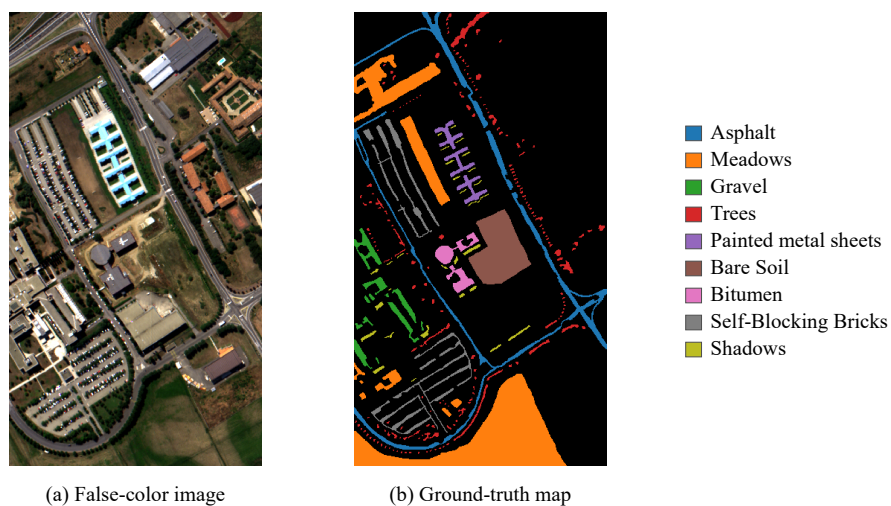


Figure 6. Representative map of the IP dataset.

Table 2. Class names and sample numbers of IP.

Class	Class Name	Number of Samples
1	Alfalfa	46
2	Corn-notill	1428
3	Corn-mintill	830
4	Corn	237
5	Grass-pasture	483
6	Grass-trees	730
7	Grass-pasture-mowed	28
8	Hay-windrowed	478
9	Oats	20
10	Soybean-notill	972
11	Soybean-mintill	2455
12	Soybean-clean	593
13	Wheat	205
14	Woods	1265
15	Buildings-Grass-Trees-Drives	386
16	Stone-Steel-Towers	93
Total Number of Samples		10,249

**Figure 7.** Representative map of the PU dataset.**Table 3.** Class names and sample numbers of PU.

Class	Class Name	Number of Samples
1	Asphalt	6631
2	Meadows	18,649
3	Gravel	2099
4	Trees	3064
5	Painted metal sheets	1345
6	Bare Soil	5029
7	Bitumen	1330
8	Self-Blocking Bricks	3682
9	Shadows	947
Total Number of Samples		42,776

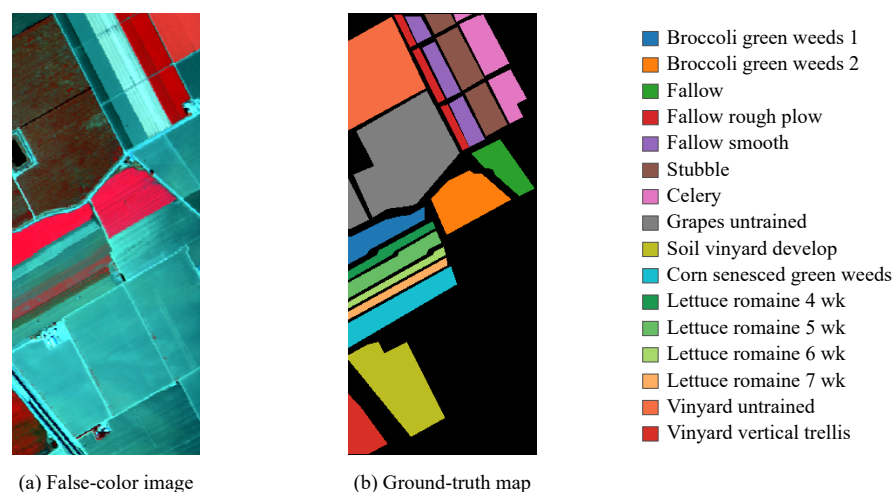


Figure 8. Representative map of the SA dataset.

Table 4. Class names and sample numbers of SA.

Class	Class Name	Number of Samples
1	Broccoli_green_weeds_1	2009
2	Broccoli_green_weeds_2	3726
3	Fallow	1976
4	Fallow_rough_plow	1394
5	Fallow_smooth	2678
6	Stubble	3959
7	Celery	3579
8	Grapes_untrained	11,271
9	Soil_vinyard_develop	6203
10	Corn_senesced_green_weeds	3278
11	Lettuce_romaine_4wk	1068
12	Lettuce_romaine_5wk	1927
13	Lettuce_romaine_6wk	916
14	Lettuce_romaine_7wk	1070
15	Vinyard_untrained	7268
16	Vinyard_vertical_trellis	1807
Total Number of Samples		54,129

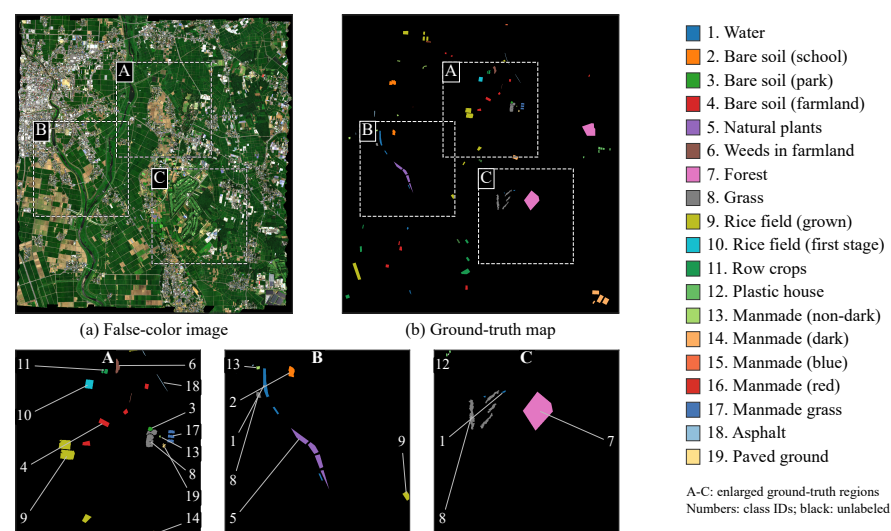


Figure 9. Representative map of the CK dataset.

Table 5. Class names and sample numbers of CK.

Class	Class Name	Number of Samples
1	Water	2845
2	Bare soil (school)	2859
3	Bare soil (park)	286
4	Bare soil (farmland)	4852
5	Natural plants	4297
6	Weeds in farmland	1108
7	Forest	20,516
8	Grass	6515
9	Rice field (grown)	13,369
10	Rice field (first stage)	1268
11	Row crops	5961
12	Plastic house	2193
13	Manmade (non-dark)	1220
14	Manmade (dark)	7664
15	Manmade (blue)	431
16	Manmade (red)	222
17	Manmade grass	1040
18	Asphalt	801
19	Paved ground	145
Total Number of Samples		77,592

2.2.3. Comparison Methods

To systematically evaluate G-SPCNet for cross-domain few-shot hyperspectral image classification under spectral-band perturbations, seven representative methods are selected for comparison. These methods include classical cross-domain few-shot learning frameworks and recent approaches based on robust prototypes and frequency-domain modeling, enabling the effectiveness and robustness of G-SPCNet to be assessed from different perspectives.

Early studies mainly focused on learning transferable features, modeling relationships among samples, and aligning cross-domain distributions. Deep Cross-Domain Few-Shot Learning (DCFSL) combines deep feature transfer with episodic metric learning to learn transferable representations from source-domain meta-tasks [20]. Graph Information Aggregation Cross-Domain Few-Shot Learning (Gia-CFSL) represents relationships between support and query samples as a graph and enhances class-structure modeling through node-level information propagation [21]. Adaptive Domain-Adversarial Few-Shot Learning (ADA-FSL) employs domain discrimination and adversarial training to reduce feature-distribution differences between the source and target domains [22].

Subsequent studies explored more refined feature decomposition and relation modeling. Cross-domain Few-shot Learning based on Feature Disentanglement (FDFSL) decomposes features into domain-invariant and domain-related components, thereby reducing the influence of domain-specific factors on class discrimination [25]. Multi-level Relation Learning for Cross-Domain Few-Shot Hyperspectral Image Classification (MLRL-FSL) models relationships between support and query samples at the sample, local-feature, and task levels [26].

More recent studies have emphasized frequency-domain representations and prototype stability. Spectral Context-aware Frequency Alignment (SCFA), referred to as SCFDA hereafter following the authors' official code repository, decomposes and aligns spectral contextual information in the frequency domain [29]. The Momentum-Enhanced Dual-Prototype Learning Framework (MEDPL) uses dual prototypes and momentum updates to reduce fluctuations in prototype estimates caused by limited support samples [28].

2.2.4. Experimental Settings

All methods are trained using LK as the sole source domain under a common episodic evaluation protocol. The labeled source samples are divided into spatially separated training and validation pools, with no shared center pixels or overlapping input patches between the two pools. Training episodes are sampled exclusively from the source training pool. Checkpoint selection uses 200 fixed held-out source-domain validation episodes and maximizes the arithmetic mean of Clean and Robust OA. No target-domain validation set is used for checkpoint selection, and target-domain samples are not used for gradient-based parameter updates. Checkpoint selection begins at training episode 1000 and is performed every 100 episodes thereafter. Training stops after 10 consecutive validation evaluations without improvement, subject to a maximum of 10,000 training episodes.

Each pixel spectrum is linearly interpolated to 100 equally spaced positions along its original band-index axis, unifying input dimensionality without aligning physical wavelengths across sensors. Each resampled channel is standardized as $(x_b - \mu_b) / (\sigma_b + 10^{-6})$, with μ_b and σ_b computed separately for each image over all pixels without using class labels. Spatial patches are then extracted. All methods use the same preprocessing, while checkpoint selection relies exclusively on source-domain validation episodes.

Each input is represented by a 9×9 spatial patch with 100 resampled spectral bands. G-SPCNet is optimized using AdamW with a learning rate of 10^{-3} and weight decay of 10^{-4} . The loss weights are $\lambda_p = 0.50$, $\lambda_l = 0.40$, $\lambda_g = 0.10$, and $\lambda_r = 0.15$, corresponding to perturbed-view classification, final predictive-distribution consistency, group-level prediction consistency, and prototype consistency, respectively. The consistency temperature is $T = 2$, and the distance-scaling coefficient is $\gamma = 10$. All comparison methods use the Adam optimizer.

For each target dataset, performance is evaluated using 600 fixed test episodes per run under both Clean and Robust conditions. The same test episodes are shared across methods. Each main-experiment episode randomly selects five classes and samples five labeled support pixels and 15 query pixels per class. Target support labels provide task-specific class references, whereas query labels are used to compute evaluation metrics. Six independent training runs are conducted with fixed random seeds 7, 17, 27, 1007, 1017, and 1027. For each run, the metrics are first averaged over the 600 test episodes; the reported results are then expressed as the mean and sample standard deviation across the six runs, rather than across individual episodes or independently generated data splits. Because every episode contains the same number of query samples per class, OA and AA are numerically identical and are reported jointly as OA/AA. Kappa is reported as a supplementary agreement measure, not as evidence of improved performance on particular class-frequency groups.

Support and query pixels are sampled without replacement within each target episode, ensuring disjoint center-pixel sets. The shot-sensitivity experiments additionally consider matched 1-shot and 10-shot training, source-validation, and target-testing settings. G-SPCNet training uses the Medium perturbation configuration in Table 11. The main Robust evaluation uses the same configuration except that the random band-dropout probability is set to zero. Table 11 also specifies the masking fractions, per-band dropout and noise probabilities, noise standard deviations, and maximum spectral shifts for the Weak, Medium, Strong, and Severe settings.

3. Results

3.1. Performance Comparison

Tables 6–9 report the classification performance on the IP, PU, SA, and CK datasets, respectively, using five labeled support samples per class. As shown in Table 6, G-SPCNet achieves a Clean OA/AA of 79.07%, which is 0.07 percentage points lower than the best-

performing method under the Clean condition, ADA-FSL. Under spectral perturbations, however, G-SPCNet achieves the highest Robust OA/AA of 69.36%, outperforming the strongest external baseline, FDFSL, by 0.71 percentage points. This result indicates that the group-aware perturbation consistency design maintains more stable prototype distances under target-domain spectral-band perturbations. G-SPCNet also achieves the highest Robust Kappa of 61.70%, compared with 60.81% for FDFSL. Kappa is reported as a supplementary agreement measure rather than evidence of improvements in particular class-frequency groups. Because IP contains several crop classes with similar spectral characteristics, these results demonstrate that the group-level consistency constraint helps the model use the remaining band groups for effective discrimination when local spectral regions are corrupted.

Table 6. Classification performance comparison on IP.

Metric	DCFSL	Gia-CFSL	ADA-FSL	FDFSL	MLRL-FSL	SCFDA	MEDPL	G-SPCNet
Clean OA/AA (%)	78.76 ± 1.22	77.54 ± 1.16	79.14 ± 1.85	78.28 ± 1.02	63.43 ± 3.03	78.63 ± 1.11	78.71 ± 1.39	79.07 ± 0.55
Clean Kappa (%)	73.46 ± 1.52	71.92 ± 1.45	73.93 ± 2.31	72.85 ± 1.28	54.28 ± 3.79	73.29 ± 1.39	73.38 ± 1.73	73.84 ± 0.69
Robust OA/AA (%)	55.37 ± 2.15	60.83 ± 3.10	59.66 ± 3.34	68.65 ± 1.03	45.03 ± 4.04	46.05 ± 2.35	41.87 ± 2.17	69.36 ± 1.12
Robust Kappa (%)	44.21 ± 2.69	51.03 ± 3.87	49.58 ± 4.17	60.81 ± 1.29	31.28 ± 5.05	32.56 ± 2.94	27.34 ± 2.71	61.70 ± 1.40

Note: The best results are highlighted in bold.

Table 7. Classification performance comparison on PU.

Metric	DCFSL	Gia-CFSL	ADA-FSL	FDFSL	MLRL-FSL	SCFDA	MEDPL	G-SPCNet
Clean OA/AA (%)	82.55 ± 0.81	81.52 ± 0.36	80.17 ± 1.34	80.50 ± 1.29	67.24 ± 3.54	82.04 ± 0.73	81.87 ± 1.27	82.75 ± 0.65
Clean Kappa (%)	78.19 ± 1.01	76.90 ± 0.45	75.21 ± 1.67	75.62 ± 1.62	59.05 ± 4.42	77.55 ± 0.91	77.34 ± 1.58	78.44 ± 0.82
Robust OA/AA (%)	63.19 ± 1.84	64.30 ± 3.76	63.87 ± 2.96	72.54 ± 1.23	51.40 ± 2.14	48.62 ± 1.71	46.69 ± 3.22	77.56 ± 0.59
Robust Kappa (%)	53.99 ± 2.30	55.38 ± 4.70	54.84 ± 3.70	65.68 ± 1.54	39.25 ± 2.68	35.77 ± 2.14	33.37 ± 4.02	71.95 ± 0.74

Note: The best results are highlighted in bold.

Table 7 shows that G-SPCNet achieves the highest Clean OA/AA of 82.75% and Robust OA/AA of 77.56% on the PU target domain. Under the Robust condition, it outperforms the strongest external baseline, FDFSL, by 5.02 percentage points, indicating that the proposed method retains strong perturbation resistance in urban scenes. Different ground objects in PU have more distinctive spatial shapes and boundary structures, but their spectral distributions still exhibit substantial differences during cross-domain transfer. The simultaneous improvements in the Clean and Robust metrics show that the local spectral-spatial representations extracted by BGR do not sacrifice standard classification performance because of robustness-oriented training. Instead, more stable grouped prototypes improve the metric space of the target domain.

Table 8. Classification performance comparison on SA.

Metric	DCFSL	Gia-CFSL	ADA-FSL	FDFSL	MLRL-FSL	SCFDA	MEDPL	G-SPCNet
Clean OA/AA (%)	93.40 ± 0.56	92.49 ± 0.79	94.05 ± 0.69	93.57 ± 0.24	75.65 ± 3.99	94.58 ± 0.44	95.03 ± 0.65	94.94 ± 0.17
Clean Kappa (%)	91.75 ± 0.70	90.61 ± 0.99	92.56 ± 0.86	91.96 ± 0.30	69.57 ± 4.99	93.23 ± 0.55	93.79 ± 0.81	93.67 ± 0.21
Robust OA/AA (%)	59.95 ± 1.67	61.58 ± 2.17	68.93 ± 2.31	80.35 ± 1.01	44.33 ± 4.23	55.48 ± 3.61	52.09 ± 3.17	87.21 ± 0.87
Robust Kappa (%)	49.94 ± 2.09	51.97 ± 2.71	61.16 ± 2.89	75.44 ± 1.26	30.41 ± 5.29	44.35 ± 4.51	40.11 ± 3.96	84.02 ± 1.08

Note: The best results are highlighted in bold.

As shown in Table 8, most methods perform well on Salinas data under the Clean setting but undergo substantial degradation under the Robust setting. G-SPCNet achieves a Robust OA/AA of 87.21%, outperforming the strongest external baseline, FDFSL, by 6.86 percentage points and demonstrating the stability provided by SPC and GAC under complex spectral perturbations in agricultural scenes. This result also shows that robustness mechanisms remain valuable even when the target-domain classes are already highly separable under the Clean setting. Conventional models may appear close to saturation under the Clean setting, but their advantages diminish rapidly when contiguous bands are missing or the spectral axis is shifted. G-SPCNet explicitly constrains these changes through perturbation consistency during training and therefore achieves better Robust performance despite the high Clean baseline.

Table 9. Classification performance comparison on CK.

Metric	DCFSL	Gia-CFSL	ADA-FSL	FDFSL	MLRL-FSL	SCFDA	MEDPL	G-SPCNet
Clean OA/AA (%)	92.54 ± 0.82	90.32 ± 1.55	90.52 ± 1.07	90.35 ± 0.23	77.79 ± 4.13	91.04 ± 0.47	91.37 ± 0.76	91.30 ± 0.28
Clean Kappa (%)	90.68 ± 1.02	87.89 ± 1.94	88.15 ± 1.33	87.93 ± 0.29	72.24 ± 5.16	88.80 ± 0.59	89.21 ± 0.95	89.12 ± 0.36
Robust OA/AA (%)	55.59 ± 0.80	58.96 ± 1.23	67.51 ± 2.41	75.66 ± 0.72	42.11 ± 1.95	51.54 ± 1.38	47.72 ± 3.18	83.75 ± 0.99
Robust Kappa (%)	44.49 ± 1.00	48.70 ± 1.53	59.39 ± 3.01	69.57 ± 0.90	27.64 ± 2.44	39.42 ± 1.73	34.65 ± 3.97	79.68 ± 1.23

Note: The best results are highlighted in bold.

In Table 9, DCFSL achieves the highest Clean OA/AA of 92.54% on CK under the Clean setting, whereas G-SPCNet achieves 91.30%, which is 1.24 percentage points lower. Under the Robust setting, however, G-SPCNet achieves 83.75%, outperforming the strongest external baseline, FDFSL, which achieves 75.66%, by 8.09 percentage points. Compared with currently popular cross-domain few-shot methods such as SCFDA and MEDPL, G-SPCNet achieves comparable Clean performance while substantially outperforming them in the Robust metrics. These results indicate that when target-domain classification under the Clean setting is relatively easy, G-SPCNet can obtain considerable robustness gains at only a small cost in Clean performance.

Although some comparison methods perform slightly better under the Clean setting, their performance decreases rapidly under perturbations. G-SPCNet does not always achieve the highest Clean results, but it shows a clear advantage under the Robust setting.

This demonstrates that G-SPCNet provides strong cross-domain generalization and stable robustness across different datasets, giving it favorable potential for practical applications.

Figures 10–13 show the ground truth and predictions of all eight methods on IP, PU, SA, and CK under the Clean condition. For each dataset, the five largest labeled classes are selected, with five support pixels per class sampled using seed 7 and shared across methods. Checkpoints from training seed 7, selected by source-domain validation, are used without target-domain parameter updates. The remaining labeled pixels of these classes are classified in row-major batches of 75 queries using the fixed support set. These are fixed-task five-class visualizations, not aggregated episodic results or all-class maps. Support pixels retain their known labels; unlabeled pixels and unselected classes appear in black. For CK, the same regions are enlarged at an identical scale across methods.

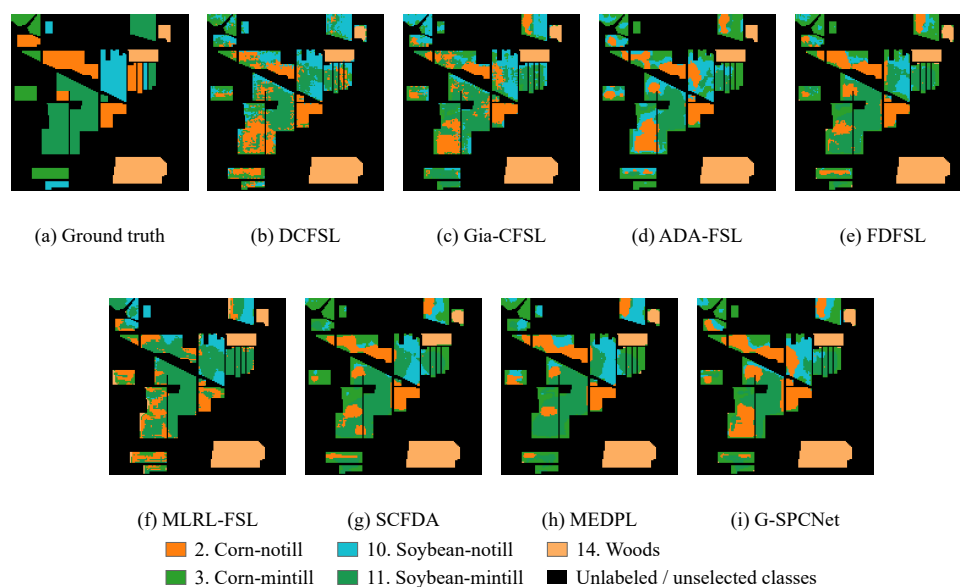


Figure 10. Classification maps on the IP dataset.

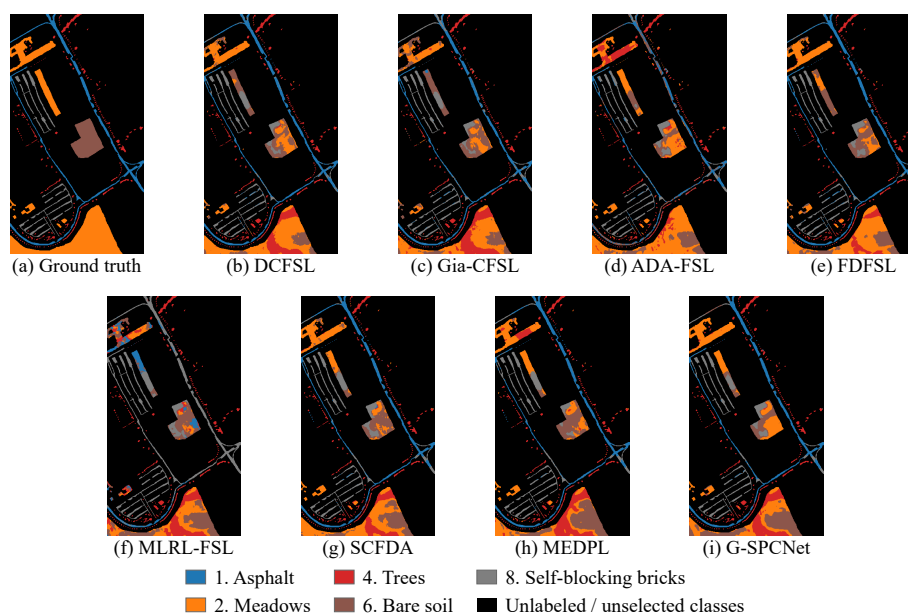


Figure 11. Classification maps on the PU dataset.

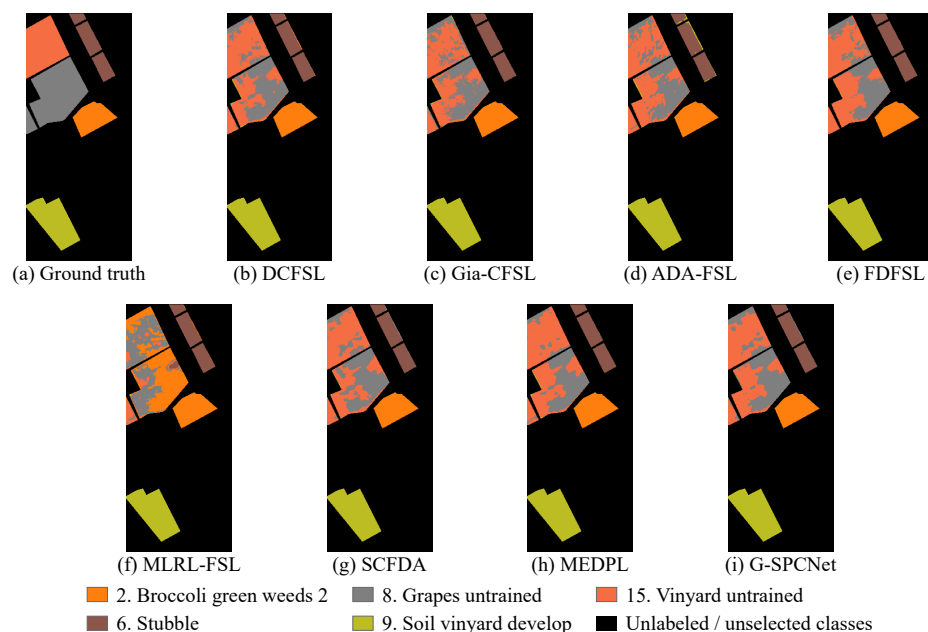


Figure 12. Classification maps on the SA dataset.

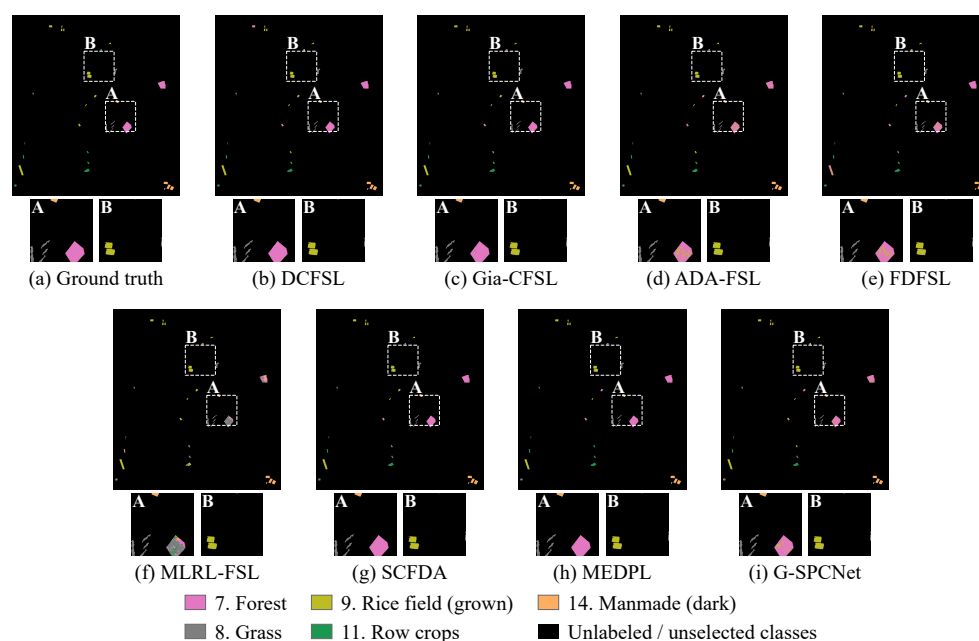


Figure 13. Classification maps on the CK dataset.

The maps illustrate differences among methods in the spatial distribution of predictions within the selected classes. G-SPCNet produces coherent predictions in several regions, although local misclassifications remain and its visual advantage is not uniform across all regions or datasets. These qualitative results complement the quantitative comparisons in Tables 6–9.

To complement the aggregate results and classification maps, Table 10 compares the class-wise query accuracy of FDFSL and G-SPCNet under Clean and Robust conditions across the four target datasets. These results are computed from the 600 fixed five-way test episodes per run rather than from the full-scene classification maps and are averaged over six seeds.

Table 10. Class-wise query accuracy (%) of FDFSL and G-SPCNet under Clean and Robust conditions.

IP					SA				
Class	Clean		Robust		Class	Clean		Robust	
	FDFSL	G-SPCNet	FDFSL	G-SPCNet		FDFSL	G-SPCNet	FDFSL	G-SPCNet
1	80.52	92.18	65.60	81.52	1	97.22	98.18	55.55	77.67
2	59.51	61.88	46.94	53.45	2	97.06	98.72	69.74	83.85
3	62.25	62.47	60.85	52.98	3	84.35	89.13	78.23	83.44
4	72.78	72.51	66.96	62.75	4	99.14	99.79	81.37	91.09
5	61.94	57.26	57.75	57.95	5	90.27	96.04	80.66	92.28
6	82.36	83.82	81.02	77.49	6	99.15	97.96	87.46	94.64
7	99.49	98.65	90.13	88.04	7	99.10	99.54	76.45	85.91
8	90.68	86.50	75.44	83.11	8	83.97	85.11	81.95	79.92
9	97.22	99.91	96.15	88.99	9	99.26	99.63	73.01	91.14
10	64.15	62.24	51.87	50.48	10	69.41	75.03	71.68	70.17
11	63.16	62.21	46.54	52.13	11	96.48	96.84	81.51	83.48
12	50.47	55.38	46.06	50.31	12	98.32	96.98	89.53	93.43
13	95.35	95.63	79.34	79.55	13	99.47	99.58	99.49	99.66
14	88.66	93.17	71.82	74.73	14	99.77	99.43	94.67	93.80
15	83.05	81.92	78.83	67.14	15	89.18	90.19	83.61	83.65
16	99.39	98.95	81.31	87.71	16	92.31	95.40	77.24	87.12

PU					CK				
Class	Clean		Robust		Class	Clean		Robust	
	FDFSL	G-SPCNet	FDFSL	G-SPCNet		FDFSL	G-SPCNet	FDFSL	G-SPCNet
1	76.87	80.51	75.98	75.41	1	92.64	94.42	72.86	89.27
2	78.16	71.18	76.85	71.01	2	94.85	94.38	58.32	82.22
3	68.42	68.61	54.99	65.87	3	97.58	99.41	81.37	90.12
4	92.47	97.38	88.90	87.79	4	69.31	60.54	66.40	61.01
5	98.04	99.61	74.55	91.74	5	99.19	99.57	80.23	94.36
6	65.67	67.20	67.65	69.46	6	96.62	96.07	84.60	88.45
7	80.60	85.46	67.87	75.10	7	95.57	97.50	90.42	90.02
8	71.60	77.19	62.69	70.02	8	95.15	96.50	79.15	90.96
9	93.40	98.84	83.11	92.71	9	92.68	93.43	85.12	85.80
10					10	96.71	99.68	78.05	95.53
11					11	71.05	75.75	77.05	76.55
12					12	77.55	82.17	63.22	75.51
13					13	94.07	95.47	73.90	90.02
14					14	79.31	79.50	67.36	66.86
15					15	97.98	97.83	74.61	77.98
16					16	84.47	82.64	67.39	76.23
17					17	93.97	94.44	66.54	84.97
18					18	94.23	94.13	81.78	80.58
19					19	92.93	99.81	93.33	94.26

Note: Class IDs correspond to the dataset class tables. The higher accuracy between the two methods under each condition is highlighted in bold.

Under the Robust condition, G-SPCNet achieves higher class-wise accuracy than FDFSL for 9 of 16 classes on IP, 6 of 9 on PU, 13 of 16 on SA, and 14 of 19 on CK. These results show that its robustness gains extend across multiple classes, although FDFSL remains more accurate for some categories.

3.2. Shot-Sensitivity Analysis

To investigate the influence of the number of support samples on robust classification, five representative methods—DCFSL, FDFSL, SCFDA, MEDPL, and G-SPCNet—are evaluated under 1-shot, 5-shot, and 10-shot settings, with matching shot settings used during training, source-domain validation, and target-domain testing. The average Robust OA across the four target domains is calculated over six random seeds. As shown in Figure 14, G-SPCNet ranks first under all three shot settings, achieving 71.42%, 79.47%, and 80.28%, respectively. Compared with FDFSL, the strongest external baseline, G-SPCNet provides improvements of 3.82, 5.17, and 5.38 percentage points, respectively. These results indicate that the robustness advantage of G-SPCNet is maintained across the evaluated support-set sizes.

Figure 14 also shows that the average Robust OA of all methods generally increases as the number of shots rises from 1 to 10, indicating that additional support samples contribute to more reliable target-domain class prototypes. However, the methods differ in their dependence on the number of support samples. Several comparison methods exhibit relatively low robustness under the 1-shot setting, suggesting that they require more labeled target-domain samples to construct reliable prototypes. In contrast, G-SPCNet maintains relatively high Robust OA even under the 1-shot setting, demonstrating that the spectral perturbation consistency learned in the source domain alleviates the instability caused by insufficient target-domain support data.

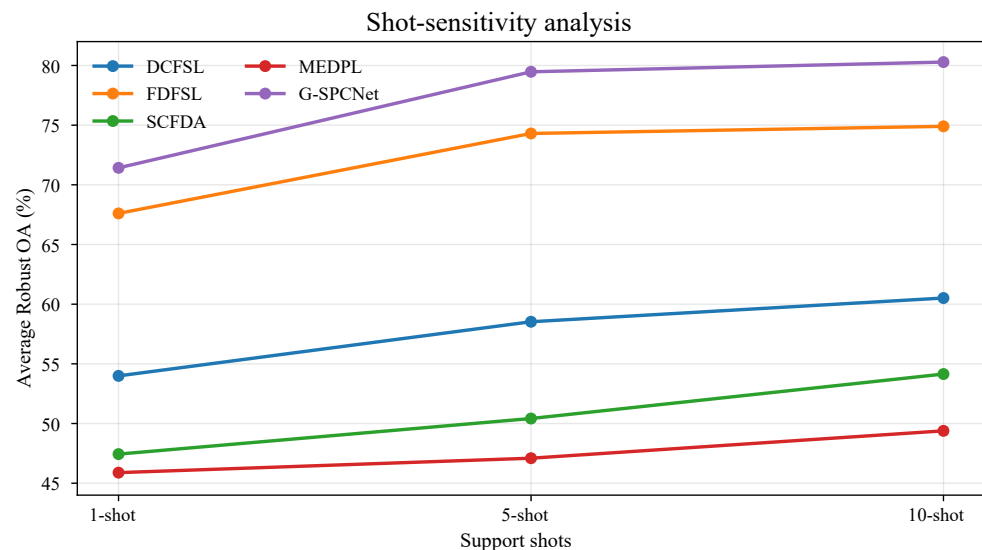


Figure 14. Shot-sensitivity analysis.

G-SPCNet remains the leading method under the 5-shot and 10-shot settings, indicating that the proposed mechanisms are not effective only when extremely few samples are available. As the number of support samples increases, the estimation of grouped prototypes becomes more stable, allowing the advantages of the group-level consistency constraint in GAC to be more fully realized. These results demonstrate that G-SPCNet adapts well to different annotation budgets.

3.3. Mixed Perturbation Severity Analysis

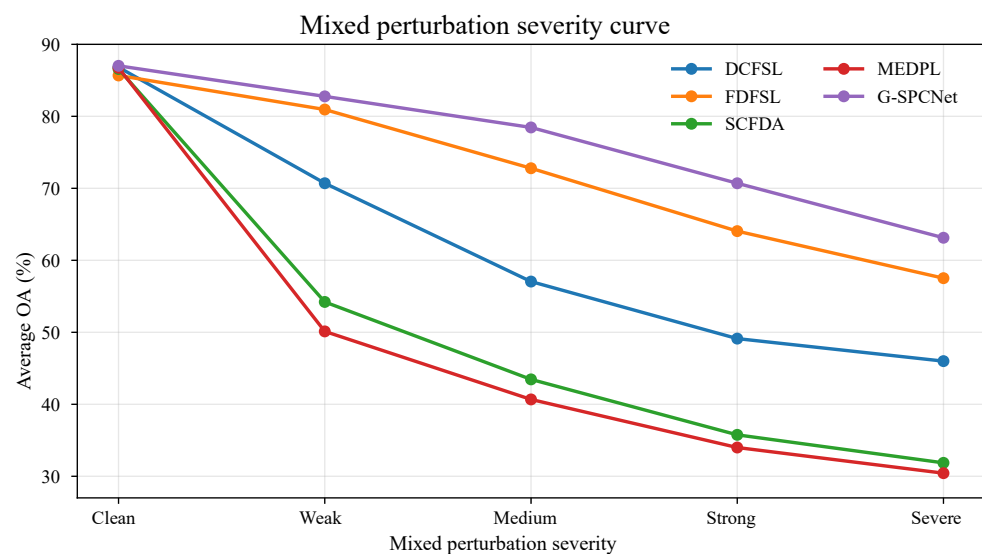
To further assess the stability of G-SPCNet under increasingly severe compound perturbations, all models are first trained using the standard protocol. Their selected checkpoints are then fixed, and only the perturbation severity during target-domain testing is varied. The Clean, Weak, Medium, Strong, and Severe settings represent five levels ranging from no perturbation to severe compound perturbations. Table 11 specifies the perturbation parameters for the four non-Clean levels. Compound perturbations are applied to both support and query samples, while Clean uses unperturbed inputs. Medium matches the perturbation settings used during G-SPCNet training and differs from the main-table Robust setting only in its random band dropout probability (0.05 versus 0). All methods use the same perturbation configuration at each severity level.

Figure 15 presents the average OA curves across the four target domains under the Clean and four mixed-perturbation settings. G-SPCNet achieves the highest mean OA at every severity level. Under the Medium and Strong settings, it outperforms the strongest external baseline, FDFS, by 5.67 and 6.65 percentage points, respectively. These results indicate that its performance degrades more slowly as perturbation severity increases.

Table 11. Parameter settings for mixed perturbation severity analysis.

Severity	Contiguous Masking	Random Dropout	Noise Probability	Noise Std.	Max. Shift (Bands)
Weak	0.10	0.03	0.03	0.02	1
Medium	0.20	0.05	0.05	0.05	2
Strong	0.30	0.10	0.05	0.08	3
Severe	0.40	0.15	0.10	0.10	4

Note: Contiguous masking denotes the fraction of bands masked. Random dropout and noise probability denote the per-band probabilities of removal and noise addition, respectively. Noise standard deviations are specified in the standardized input space. No perturbations are applied in the Clean setting.

**Figure 15.** Mixed perturbation severity curve.

As the perturbation level increases from Clean to Severe, the average OA of all methods declines, although their rates of degradation differ substantially. SCFDA and MEDPL show pronounced declines from Clean to Weak, followed by further degradation as perturbation severity increases. In contrast, G-SPCNet maintains the highest average OA across all severity levels, with a more pronounced advantage under the Medium, Strong, and Severe settings. These results show that its performance advantage extends beyond the training perturbation setting.

The curves show that the advantage of G-SPCNet is already evident between the Clean and Weak settings and becomes more substantial under Medium and stronger perturbations. This trend is consistent with the intended design of the method. Under mild perturbations, most original features remain available, allowing many models to retain satisfactory classification performance. However, when contiguous band masking, random band dropout, noise injection, and spectral-axis shifting are simultaneously intensified, models explicitly trained with perturbation-consistency constraints maintain a more stable prototype space.

3.4. Ablation Study

To evaluate the effectiveness of the band-grouped representation module, spectral perturbation consistency mechanism, and group-aware consistency mechanism, six ablation configurations are considered. The first is a prototype-classification baseline. The second incorporates BGR to introduce grouped representations and grouped prototypes;

both configurations use the same basic augmentation. The third, BGR + augmentation only, computes classification losses separately on the original and perturbed views without consistency constraints. The fourth, BGR + SPC, further introduces final predictive-distribution consistency between the two views. The fifth, BGR + SPC + prototype consistency, adds only the prototype-consistency loss of GAC. The complete G-SPCNet further incorporates the group-level predictive-consistency loss. Table 12 reports the Robust OA/AA results across the four target domains.

Table 12. Ablation study of key components in terms of Robust OA/AA (%).

Configuration	IP	PU	SA	CK
Prototype baseline	56.46	63.73	68.97	64.09
BGR only	58.45	64.86	75.43	70.57
BGR + augmentation only	63.50	70.95	80.90	76.37
BGR + SPC	67.39	74.63	85.73	80.07
BGR + SPC + prototype consistency	67.90	74.86	85.73	80.09
G-SPCNet	69.36	77.56	87.21	83.75

Note: Values are averaged over six seeds. The best results among the listed configurations are highlighted in bold.

As shown in Table 12, BGR increases the average Robust OA/AA across the four target domains from 63.31% to 67.33% under the same basic augmentation setting. The BGR + augmentation only configuration further raises it to 72.93% by computing classification losses separately on the original and perturbed views. On this basis, introducing final predictive-distribution consistency through SPC increases the average Robust OA/AA to 76.96%, indicating that predictive-consistency constraints provide additional benefits beyond data augmentation alone. Further incorporating prototype consistency yields the BGR + SPC + prototype consistency configuration, which increases the average Robust OA/AA to 77.15%. This constraint uses clean prototypes as references to anchor perturbed prototypes and yields a modest additional improvement in average Robust OA/AA. Finally, the complete G-SPCNet achieves an average Robust OA/AA of 79.47%, the highest among the configurations listed in the table.

The ablation results indicate that BGR preserves local discriminative information from different spectral regions and improves the stability of class prototypes under cross-domain conditions. The largest improvement in Robust performance occurs after introducing SPC, confirming that online compound spectral perturbations, perturbed-view classification, and final predictive-distribution consistency jointly constitute the primary source of robustness gains. GAC provides a smaller but consistent improvement by constraining group-level classification outputs and using clean prototypes as references to calibrate perturbed prototypes. This process reduces group-level prediction drift and class-prototype shifts. Combining BGR, SPC, and GAC achieves the highest average Robust performance, demonstrating their complementary roles in grouped representation, perturbation learning, and local structural stabilization.

3.5. Computational Efficiency and Model Complexity

Table 13 reports the mean training and testing times over six seeds for each target domain, together with the parameter count of each method. Here, Params (M) denotes the total number of model parameters expressed in millions, where 1 M equals 10^6 parameters. The counts include retained auxiliary modules, even when these modules are not used during inference. Because the model configurations and implementation details differ from those in the original publications, the parameter counts of some comparison methods may

also differ. The reported values therefore characterize the model sizes of the implementations used in this study. G-SPCNet contains approximately 0.153 million parameters, the fewest among the evaluated methods, while achieving the highest mean Robust OA/AA on all four target domains. Training time includes source-domain validation, and testing time covers both the Clean and Robust conditions. Overall, the parameter counts and classification results demonstrate that G-SPCNet combines a compact model with strong robustness to spectral-band perturbations.

Table 13. Training time, testing time, and parameter counts of different methods.

Dataset	Metric	DCFSL	Gia-CFSL	ADA-FSL	FDFSL	MLRL-FSL	SCFDA	MEDPL	G-SPCNet
IP	Training Time (s)	294.00	517.22	1008.63	458.55	223.28	338.49	295.93	399.59
	Testing Time (s)	12.44	20.02	51.71	15.34	14.14	20.97	12.00	9.17
	Params (M)	4.245	1.268	4.245	4.273	6.052	0.619	0.256	0.153
PU	Training Time (s)	294.00	517.22	1008.63	458.55	223.28	338.49	295.93	399.59
	Testing Time (s)	12.63	20.18	51.53	14.70	14.46	20.85	12.31	9.48
	Params (M)	4.245	1.268	4.245	4.273	6.052	0.619	0.256	0.153
SA	Training Time (s)	294.00	517.22	1008.63	458.55	223.28	338.49	295.93	399.59
	Testing Time (s)	12.65	19.10	50.69	14.63	14.39	21.60	12.43	9.36
	Params (M)	4.245	1.268	4.245	4.273	6.052	0.619	0.256	0.153
CK	Training Time (s)	294.00	517.22	1008.63	458.55	223.28	338.49	295.93	399.59
	Testing Time (s)	12.44	20.43	52.24	15.42	15.40	21.34	12.49	9.46
	Params (M)	4.245	1.268	4.245	4.273	6.052	0.619	0.256	0.153

Note: Times are averaged over six seeds. Training time includes source-domain validation. The same source-trained models are evaluated on all four target domains. Params (M) denotes the total number of model parameters in millions, including retained auxiliary modules.

G-SPCNet does not always require the shortest training time because both clean and perturbed views are processed during training, together with multiple levels of consistency loss. During target-task inference, however, the consistency branches are no longer used. The model retains only single-view BGR encoding, grouped prototype construction, and distance-based classification. Therefore, its testing cost does not increase proportionally with the additional computations required for dual-view training. Moreover, the use of lightweight convolutional operations and GRU-based contextual modeling avoids the parameter overhead associated with large-scale pretrained backbones and high-dimensional relation modules.

Unlike models with larger parameter counts, such as DCFSL, ADA-FSL, FDFSL, and MLRL-FSL, G-SPCNet does not improve its representational ability by simply increasing network size. Instead, it enhances adaptability to spectral abnormalities through band grouping and multilevel consistency constraints. These results suggest that greater model capacity does not necessarily yield more stable cross-domain classification when only limited labeled samples are available in the target domain. Architectures and training constraints specifically designed for spectral perturbations can achieve strong robustness with relatively low model complexity.

4. Discussion

The experimental results support the central hypothesis of this study that explicitly modeling spectral-band perturbations during source-domain meta-training improves the stability of grouped prototypes under cross-domain conditions. Across the four target domains, G-SPCNet achieves competitive Clean performance and consistently outperforms the comparison methods in Robust OA/AA and Robust Kappa. Existing methods mainly focus on transferable feature learning, inter-domain distribution alignment, relational mod-

eling, or prototype estimation [20–22,25,26,28,29]. Although these methods can perform well under the Clean setting, they do not directly constrain changes in class prototypes and predictive distributions caused by band loss, noise, or spectral-axis shifts. G-SPCNet instead treats spectral abnormalities as a specific source of cross-domain uncertainty and addresses them during both representation learning and prototype classification.

The ablation results further clarify the roles of the three components. Introducing BGR increases the average Robust OA/AA from 63.31% to 67.33%, indicating that retaining band-group-level representations is more robust than immediately compressing the full spectrum into a single global representation. Grouped prototypes restrict the direct effect of a locally abnormal spectral region on the overall class center, while the bidirectional GRU provides contextual information from adjacent groups. The BGR + augmentation only variant raises the average Robust OA/AA to 72.93% by applying classification losses to both clean and perturbed views. Adding predictive-distribution consistency to complete SPC further increases it to 76.96%, demonstrating an additional benefit beyond augmentation alone. Together, the SPC operations yield the largest component-level improvement. This result is consistent with the design of SPC, which exposes the model to compound spectral perturbations and jointly optimizes clean-view classification, perturbed-view classification, and predictive-distribution consistency.

GAC further increases the average Robust OA/AA from 76.96% to 79.47%. The intermediate BGR + SPC + prototype consistency variant achieves 77.15%, whereas the complete model additionally incorporates group-level prediction consistency. Although the overall gain from GAC is smaller than that produced by SPC, it remains consistent across the target domains. Final-logit consistency alone may conceal unstable responses from individual band groups because deviations from different groups can offset one another during equal-weight fusion. GAC addresses this limitation by aligning group-level logits and using clean prototypes as stop-gradient anchors for the perturbed prototypes. Its main contribution is therefore to refine local prediction stability and preserve prototype structure after SPC has stabilized the final predictive distribution. This interpretation is consistent with the modest but stable improvement observed in the ablation results.

The shot-sensitivity experiment shows that G-SPCNet maintains consistent performance under different annotation budgets. Among the compared methods, it records the highest average Robust OA at 1, 5, and 10 shots, indicating that its robustness advantage does not depend on a particular support-set size. When only a few labeled target-domain samples are available, the grouped representations and perturbation-consistency objectives learned from the source domain still support reliable target-domain classification. As the number of support samples increases, the grouped class prototypes can be estimated more reliably, while G-SPCNet continues to outperform the comparison methods. These results show that the proposed method remains effective across the tested support-set sizes.

The Mixed Severity experiment provides complementary evidence regarding robustness under different perturbation intensities. The average OA of all methods decreases as the compound perturbations become stronger, whereas G-SPCNet maintains the highest performance across the evaluated severity levels. Its advantage becomes more pronounced under the Medium, Strong, and Severe settings. These results show that its performance advantage extends beyond the training perturbation setting. When perturbations are weak, most methods can still rely on the remaining original spectral information. As contiguous band masking, random band dropout, noisy-band injection, and spectral-axis shifting become stronger, the benefits of explicitly learning perturbation-consistent representations become more apparent.

The low model complexity of G-SPCNet is another practical advantage. The model contains approximately 0.153 M parameters, the lowest number among the compared

methods. Dual-view processing and multilevel consistency objectives increase the training cost because both clean and perturbed samples must be encoded during source-domain meta-training. However, these consistency-related computations are not required during target-domain inference. Target prediction involves only single-view BGR encoding, grouped prototype construction, and distance-based classification. The robustness improvement is therefore mainly obtained through the grouped representation and training objectives rather than through a large backbone or increased model capacity. This design may be useful for hyperspectral applications with limited computational resources and scarce target-domain annotations.

Several limitations remain. First, the synthetic perturbations provide controlled approximations of spectral-quality changes but do not fully capture sensor-specific spectral responses, calibration errors, or band-dependent noise characteristics. In particular, noisy-band injection uses the same selection probability and noise standard deviation across bands in the standardized input space, rather than sensor-specific SNR profiles. Generalization to hyperspectral data with substantially different spectral coverage or imaging characteristics therefore requires further validation. Second, support and query center pixels are distinct within each episode, but no spatial exclusion buffer is imposed, allowing neighboring patches to overlap. This spatial-context sharing may yield optimistic performance estimates. Strict separation would reduce the available samples, particularly for small or spatially concentrated classes under fixed few-shot support/query requirements. All methods follow the same sampling protocol, ensuring a consistent comparison setting, although the overlap limitation remains. Finally, equal weighting of all band groups may be suboptimal when their reliability and discriminative contributions differ considerably.

Future work will evaluate G-SPCNet using sensor-aware perturbations informed by spectral-response functions and band-dependent noise characteristics, and develop spatially disjoint evaluation protocols tailored to few-shot settings to further assess its performance. Adaptive group weighting and interpretable band-group importance analysis may also improve the integration of spectral regions with different reliability and discriminative value.

5. Conclusions

This study proposed G-SPCNet for cross-domain few-shot hyperspectral image classification under spectral-band perturbations. The method integrates band-grouped representation, online spectral perturbation consistency, and group-aware consistency to stabilize group-level predictions and class prototypes. Experiments across four target domains showed that G-SPCNet maintains competitive Clean performance while achieving clear improvements in Robust OA/AA and Robust Kappa. With only 0.153 M parameters, it also performs consistently across different support-set sizes and perturbation severities. These findings demonstrate that G-SPCNet provides a lightweight and effective solution for improving robustness to spectral perturbations in cross-domain few-shot hyperspectral image classification.

6. Patents

The work reported in this manuscript did not result in any patents.

Author Contributions: Conceptualization, X.L. and W.X.; methodology, W.X.; software, W.X.; data curation, X.L. and W.X.; validation, W.X. and Y.J.; visualization, X.L. and W.X.; writing—original draft preparation, W.X.; writing—review and editing, X.L., W.X. and Y.J.; supervision, X.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The WHU-Hi-LongKou dataset is publicly available from the WHU-Hi repository at https://rsidea.whu.edu.cn/e-resource_WHUHi_sharing.htm (accessed on 25 August 2026). The Indian Pines and Pavia University benchmark datasets used in this study, including their ground-truth maps, are available from accessible mirrors at https://huggingface.co/datasets/danaroth/indian_pines/tree/main (accessed on 25 August 2026) and <https://huggingface.co/datasets/danaroth/pavia/tree/main> (accessed on 25 August 2026), respectively. The Salinas benchmark dataset is archived on Zenodo at <https://doi.org/10.5281/zenodo.15771735> (accessed on 25 August 2026). The Chikusei dataset is available from the dataset author's webpage at <https://naotoyokoya.com/Download.html> (accessed on 25 August 2026). The original Indian Pines and Salinas imagery was acquired by the AVIRIS sensor, whereas the Pavia University imagery was acquired by the ROSIS-03 sensor. No new datasets were created in this study.

Acknowledgments: The authors gratefully acknowledge the institutions and research teams that made the WHU-Hi-LongKou, Indian Pines, Pavia University, Salinas, and Chikusei hyperspectral datasets publicly available. During the preparation of this manuscript, the authors used ChatGPT (OpenAI, GPT-5.6 Sol) for Chinese-to-English translation, language polishing, and formatting assistance. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AA	Average accuracy
ADA-FSL	Adaptive domain-adversarial few-shot learning
BGR	Band-grouped representation
BiGRU	Bidirectional gated recurrent unit
CE	Cross-entropy
CK	Chikusei
CNN	Convolutional neural network
DCFSL	Deep cross-domain few-shot learning
FDFSL	Cross-domain few-shot learning based on feature disentanglement
GAC	Group-aware consistency
Gia-CFSL	Graph information aggregation cross-domain few-shot learning
GRU	Gated recurrent unit
G-SPCNet	Group-aware spectral perturbation consistency network
HSI	Hyperspectral image
IP	Indian Pines
KL	Kullback–Leibler
LK	LongKou
MEDPL	Momentum-enhanced dual-prototype learning framework
MLRL-FSL	Multi-level relation learning for cross-domain few-shot hyperspectral image classification
MSE	Mean squared error
OA	Overall accuracy
PU	Pavia University
ReLU	Rectified linear unit
SA	Salinas
SCFDA	Spectral context-aware frequency domain alignment
SNR	Signal-to-noise ratio
SPC	Spectral perturbation consistency

References

- Melgani, F.; Bruzzone, L. Classification of Hyperspectral Remote Sensing Images with Support Vector Machines. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 1778–1790. [\[CrossRef\]](#)
- Li, Y.; Zhang, H.; Shen, Q. Spectral–Spatial Classification of Hyperspectral Imagery with 3D Convolutional Neural Network. *Remote Sens.* **2017**, *9*, 67. [\[CrossRef\]](#)
- Zhong, Z.; Li, J.; Luo, Z.; Chapman, M. Spectral–Spatial Residual Network for Hyperspectral Image Classification: A 3-D Deep Learning Framework. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 847–858. [\[CrossRef\]](#)
- Hang, R.; Li, Z.; Liu, Q.; Ghamisi, P.; Bhattacharyya, S.S. Hyperspectral Image Classification with Attention-Aided CNNs. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 2281–2293. [\[CrossRef\]](#)
- Hong, D.; Gao, L.; Yao, J.; Zhang, B.; Plaza, A.; Chanussot, J. Graph Convolutional Networks for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 5966–5978. [\[CrossRef\]](#)
- Vinyals, O.; Blundell, C.; Lillicrap, T.; Kavukcuoglu, K.; Wierstra, D. Matching Networks for One Shot Learning. In *Proceedings of the Advances in Neural Information Processing Systems*; American Institute of Physics: College Park, MD, USA, 2016; Volume 29, pp. 3630–3638.
- Sung, F.; Yang, Y.; Zhang, L.; Xiang, T.; Torr, P.H.; Hospedales, T.M. Learning to Compare: Relation Network for Few-Shot Learning. In *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA*; IEEE: Piscataway, NJ, USA, 2018; pp. 1199–1208. [\[CrossRef\]](#)
- Snell, J.; Swersky, K.; Zemel, R. Prototypical Networks for Few-shot Learning. In *Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA*; American Institute of Physics: College Park, MD, USA, 2017; Volume 30, pp. 4077–4087.
- Liu, B.; Yu, X.; Yu, A.; Zhang, P.; Wan, G.; Wang, R. Deep Few-Shot Learning for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 2290–2304. [\[CrossRef\]](#)
- Gao, K.; Liu, B.; Yu, X.; Qin, J.; Zhang, P.; Tan, X. Deep Relation Network for Hyperspectral Image Few-Shot Classification. *Remote Sens.* **2020**, *12*, 923. [\[CrossRef\]](#)
- Tang, H.; Li, Y.; Han, X.; Huang, Q.; Xie, W. A Spatial–Spectral Prototypical Network for Hyperspectral Remote Sensing Image. *IEEE Geosci. Remote Sens. Lett.* **2020**, *17*, 167–171. [\[CrossRef\]](#)
- Finn, C.; Abbeel, P.; Levine, S. Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks. In *Proceedings of the 34th International Conference on Machine Learning, Proceedings of Machine Learning Research, Sydney, Australia*; PMLR: New York, NY, USA, 2017; Volume 70, pp. 1126–1135.
- Sun, Q.; Liu, Y.; Chua, T.S.; Schiele, B. Meta-Transfer Learning for Few-Shot Learning. In *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA*; IEEE: Piscataway, NJ, USA, 2019; pp. 403–412. [\[CrossRef\]](#)
- Wang, Y.; Yao, Q.; Kwok, J.T.; Ni, L.M. Generalizing from a Few Examples: A Survey on Few-shot Learning. *ACM Comput. Surv.* **2021**, *53*, 63. [\[CrossRef\]](#)
- Zhang, H.; Cisse, M.; Dauphin, Y.N.; Lopez-Paz, D. Mixup: Beyond Empirical Risk Minimization. In *Proceedings of the International Conference on Learning Representations*; ICLR: Appleton, WI, USA, 2018.
- Yun, S.; Han, D.; Chun, S.; Oh, S.J.; Yoo, Y.; Choe, J. CutMix: Regularization Strategy to Train Strong Classifiers with Localizable Features. In *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea*; IEEE: Piscataway, NJ, USA, 2019; pp. 6022–6031. [\[CrossRef\]](#)
- Antoniou, A.; Storkey, A.; Edwards, H. Augmenting Image Classifiers Using Data Augmentation Generative Adversarial Networks. In *Artificial Neural Networks and Machine Learning—ICANN 2018*; Lecture Notes in Computer Science; Kůrková, V., Manolopoulos, Y., Hammer, B., Iliadis, L., Maglogiannis, I., Eds.; Springer International Publishing: Cham, Switzerland, 2018; Volume 11141, pp. 594–603. [\[CrossRef\]](#)
- Doersch, C.; Gupta, A.; Efros, A.A. Unsupervised Visual Representation Learning by Context Prediction. In *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile*; IEEE: Piscataway, NJ, USA, 2015; pp. 1422–1430. [\[CrossRef\]](#)
- Chen, T.; Kornblith, S.; Norouzi, M.; Hinton, G. A Simple Framework for Contrastive Learning of Visual Representations. In *Proceedings of the 37th International Conference on Machine Learning, Proceedings of Machine Learning Research*; Daumé, H., III, Singh, A., Eds.; PMLR: New York, NY, USA, 2020; Volume 119, pp. 1597–1607.
- Li, Z.; Liu, M.; Chen, Y.; Xu, Y.; Li, W.; Du, Q. Deep Cross-Domain Few-Shot Learning for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5501618. [\[CrossRef\]](#)
- Zhang, Y.; Li, W.; Zhang, M.; Wang, S.; Tao, R.; Du, Q. Graph Information Aggregation Cross-Domain Few-Shot Learning for Hyperspectral Image Classification. *IEEE Trans. Neural Netw. Learn. Syst.* **2024**, *35*, 1912–1925. [\[CrossRef\]](#)
- Ye, Z.; Wang, J.; Liu, H.; Zhang, Y.; Li, W. Adaptive Domain-Adversarial Few-Shot Learning for Cross-Domain Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5532017. [\[CrossRef\]](#)

23. Ganin, Y.; Ustinova, E.; Ajakan, H.; Germain, P.; Larochelle, H.; Laviolette, F.; March, M.; Lempitsky, V. Domain-Adversarial Training of Neural Networks. *J. Mach. Learn. Res.* **2016**, *17*, 1–35.
24. Long, M.; Cao, Z.; Wang, J.; Jordan, M.I. Conditional Adversarial Domain Adaptation. In *Proceedings of the Advances in Neural Information Processing Systems*; American Institute of Physics: College Park, MD, USA, 2018; Volume 31, pp. 1640–1650.
25. Qin, B.; Feng, S.; Zhao, C.; Li, W.; Tao, R.; Xiang, W. Cross-Domain Few-Shot Learning Based on Feature Disentanglement for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 5514215. [[CrossRef](#)]
26. Liu, C.; Yang, L.; Li, Z.; Yang, W.; Han, Z.; Guo, J.; Yu, J. Multi-Level Relation Learning for Cross-Domain Few-Shot Hyperspectral Image Classification. *Appl. Intell.* **2024**, *54*, 4392–4410. [[CrossRef](#)]
27. Wang, X.; Jiang, W.; Kong, Y.; Cheng, Y. Hyperspectral Image Cross-Domain Few-Shot Classification Based on Disentangled Confidence Prototype Network. *Natl. Remote Sens. Bull.* **2026**, *30*, 156–169. (In Chinese) [[CrossRef](#)]
28. Liu, H.; He, J.; Zhang, X.; Li, Z. Momentum-Enhanced Dual-Prototype Learning Framework for Robust Few-Shot Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2025**, *63*, 5534816. [[CrossRef](#)]
29. Ma, S.; He, J.; Gao, Y.; Li, Z. Spectral Context-Aware Frequency Alignment for Few-Shot Hyperspectral Image Classification. *Knowl.-Based Syst.* **2026**, *332*, 114908. [[CrossRef](#)]
30. Li, Z.; Zhang, C.; Wang, Y.; Li, W.; Du, Q.; Fang, Z.; Chen, Y. Cross-Domain Few-Shot Hyperspectral Image Classification with Cross-Modal Alignment and Supervised Contrastive Learning. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 5519319. [[CrossRef](#)]
31. Haut, J.M.; Paoletti, M.E.; Plaza, J.; Plaza, A.; Plaza, L. Hyperspectral Image Classification Using Random Occlusion Data Augmentation. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 1751–1755. [[CrossRef](#)]
32. Hong, D.; Han, Z.; Yao, J.; Gao, L.; Zhang, B.; Plaza, A.; Chanussot, J. SpectralFormer: Rethinking Hyperspectral Image Classification with Transformers. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5518615. [[CrossRef](#)]
33. Mou, L.; Ghamisi, P.; Zhu, X.X. Deep Recurrent Neural Networks for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 3639–3655. [[CrossRef](#)]
34. Wu, G.; Ning, X.; Hou, L.; He, F.; Zhang, H.; Shankar, A. Three-Dimensional Softmax Mechanism Guided Bidirectional GRU Networks for Hyperspectral Remote Sensing Image Classification. *Signal Process.* **2023**, *212*, 109151. [[CrossRef](#)]
35. Miyato, T.; Maeda, S.I.; Koyama, M.; Ishii, S. Virtual Adversarial Training: A Regularization Method for Supervised and Semi-Supervised Learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *41*, 1979–1993. [[CrossRef](#)]
36. Chen, X.; He, K. Exploring Simple Siamese Representation Learning. In *Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN, USA; IEEE: Piscataway, NJ, USA, 2021; pp. 15750–15758. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.