

A Deep Learning–Based Framework for Hyperspectral Image Classification Using Adversarial Spectral Augmentation

Hiroshi Tanaka^{1*} and Keiko Nakamura^{2 1}

^{1,2}*Department of Computer Science and Engineering Graduate School of Engineering, Kyoto Institute of Technology, Kyoto, Japan*

¹*h.tanaka@kit.ac.jp*, ²*k.nakamura@kit.ac.jp*

**Corresponding author's email: h.tanaka@kit.ac.jp*

DOI: <http://dx.doi.org/10.56828/jsr.2025.4.2.1>

Article Info: Received: (October 1, 2025); Review Result: (November 6, 2025);

Accepted: (December 15, 2025)

Abstract: Hyperspectral Imaging (HSI) provides rich spectral information that is beneficial for a wide range of applications, including remote sensing, environmental monitoring, and medical analysis. However, the high dimensionality of hyperspectral data and the limited availability of labeled samples often degrade classification performance. To address these challenges, this study proposes a deep learning–based framework that integrates Generative Adversarial Networks (GANs) with conventional classifiers for hyperspectral image classification. In the proposed framework, a convolutional neural network–based generator is employed to synthesize representative spectral samples, while a discriminator is trained to distinguish between real and generated spectra through an adversarial learning process. The generated spectra are subsequently utilized to augment training data for classification models, including support vector machines and neural networks. This hybrid learning strategy is intended to mitigate the effects of data scarcity and improve generalization performance. Experimental results obtained from multiple hyperspectral datasets indicate that the proposed framework achieves stable improvements in classification accuracy, recall, and F1-score when compared with models trained using original data only. These findings suggest that adversarial spectral augmentation is an effective approach for enhancing hyperspectral image classification under limited training conditions.

Keywords: Hyperspectral imaging, generative adversarial networks, deep learning, data augmentation, spectral classification

1. Introduction

Hyperspectral Imaging (HSI) acquires dense spectral measurements at each pixel and therefore provides a high-dimensional representation of scene materials that is useful for tasks such as land-cover mapping, mineral identification, agricultural monitoring, and biomedical inspection. However, the high spectral dimensionality of HSI and the frequent scarcity of

¹ORCID: 0000-0002-8451-3921

labeled training samples create two persistent challenges for supervised classification: (i) the curse of dimensionality (Hughes phenomenon), which can degrade classifier generalization when labeled data are limited; and (ii) complex spectral-spatial correlations and domain shifts across sensors and acquisition conditions, which reduce the transferability of learned models. These issues motivate methods that both (a) learn robust spectral-spatial representations from limited labels and (b) expand or adapt the effective training distribution in principled ways [1][2][3].

Recent years have seen three complementary research directions toward these goals. First, deep spectral-spatial networks (including 1D/2D/3D CNNs, residual networks, and hybrid convolution-transformer architectures) have been developed to automatically extract hierarchical features from hyperspectral cubes and often outperform hand-crafted pipelines when sufficient training data are available [2][8]. Second, generative modelling and adversarial augmentation have been proposed to synthetically enlarge scarce hyperspectral datasets and to create samples that reflect plausible spectral variability; conditional or domain-aware GAN variants have been applied for augmentation, domain mapping, and robustness to illumination and sensor variation [3][4]. Third, learning paradigms that reduce dependence on labeled data — including few-shot/meta-learning, self-supervised pretext tasks, and attention-guided domain adaptation — have been introduced to improve generalization under limited labels and cross-scene shifts [5][6][7][9][10][11].

Several recent studies illustrate how these approaches can be combined. For example, ensemble and oversampling strategies have been coupled with CNN backbones to address class imbalance and to stabilize performance on small classes [12][15], while prototype/attention mechanisms have been integrated into domain adaptation pipelines to align class-conditional distributions between source and target scenes [11][14]. Likewise, hybrid architectures that fuse convolutional feature extractors with transformer-style attention modules have produced improved spectral-spatial representations in many benchmarks [8]. These empirical results suggest that a hybrid pipeline — one that uses adversarially generated spectra for augmentation, a spectral-spatial backbone for representation learning, and attention/prototype constructs for domain alignment — can be a practical strategy for improving HSI classification under realistic constraints.

Motivated by these developments, the present study develops and evaluates a hybrid deep-learning framework that combines adversarial spectral synthesis with conventional classifiers and attention-guided adaptation. The framework aims to (i) synthesize spectrally plausible training samples to mitigate label scarcity, (ii) learn discriminative spectral-spatial representations using a compact deep backbone suitable for HSI, and (iii) improve cross-scene robustness via attention-based adaptation or prototype refinement. We evaluate the framework on multiple publicly used hyperspectral benchmarks and report standard classification metrics (accuracy, recall, and F1-score) to quantify improvements obtained from adversarial augmentation and adaptation modules. The approach is framed conservatively: our goal is not to claim universal superiority but to empirically examine when and how adversarial augmentation combined with representation and adaptation components yields stable gains under limited-label and cross-scene conditions [1]-[15].

2. Related Work

Research on hyperspectral image (HSI) classification has evolved through several methodological stages, driven primarily by the challenges of high dimensionality, limited labeled data, and variability across acquisition conditions. This section reviews related work

across four main categories: classical machine-learning methods, deep spectral–spatial representation learning, data augmentation and generative modeling, and domain adaptation techniques.

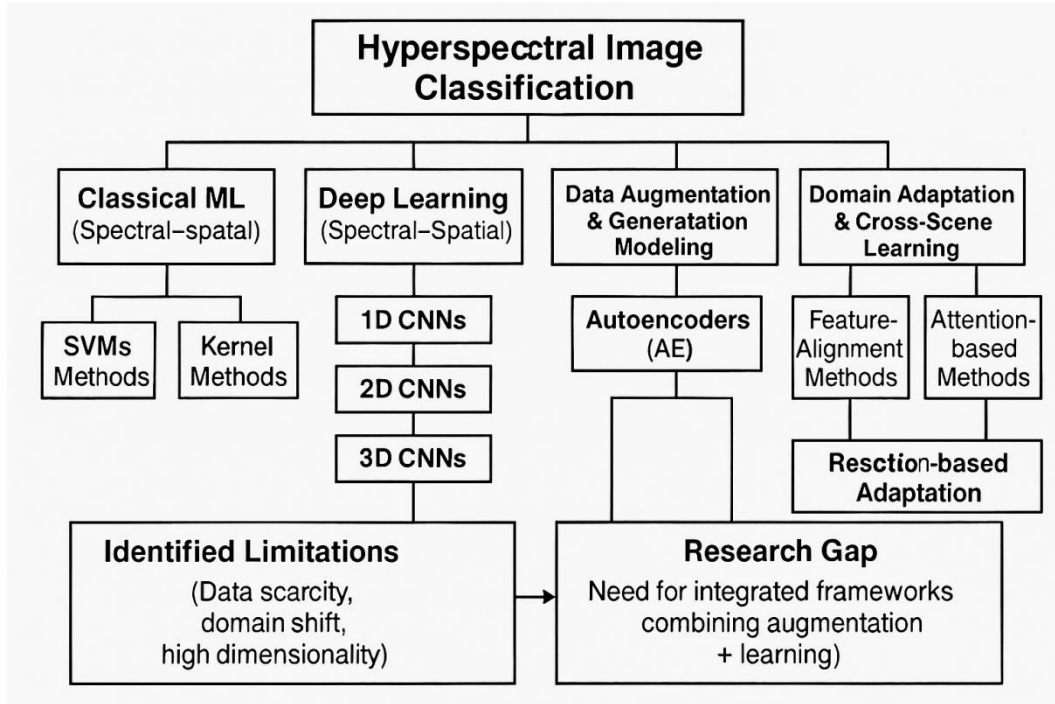


Figure 1: Taxonomy of hyperspectral image classification methods

Figure 1 summarizes the main categories of hyperspectral image classification methods discussed in this section and highlights the limitations that motivate integrated learning frameworks.

2.1. Classical Machine-Learning Approaches

Early studies on hyperspectral image classification predominantly relied on statistical learning and kernel-based classifiers. Support Vector Machines (SVMs), in particular, demonstrated strong generalization capability in high-dimensional spectral spaces and became a de facto baseline for HSI classification tasks [16]. Kernel-based learning further enhanced classification performance by implicitly mapping spectral features into higher-dimensional spaces, enabling improved discrimination of spectrally similar classes [17].

Despite their effectiveness, these methods generally depended on handcrafted spectral features and were sensitive to the availability of labeled training data. As hyperspectral sensors advanced and the number of spectral bands increased, feature redundancy and noise became more pronounced. To address these issues, dimensionality reduction and feature extraction techniques such as low-rank representation, sparse coding, and manifold learning were introduced to suppress redundancy while preserving class-discriminative information [18]. However, these approaches often required careful parameter tuning and did not fully exploit spatial contextual information.

2.2. Deep Spectral–Spatial Representation Learning

The introduction of deep learning significantly changed the landscape of HSI classification. Convolutional neural networks (CNNs) were adapted to hyperspectral data by treating spectral signatures as one-dimensional signals or by jointly processing spectral–spatial cubes using two- or three-dimensional convolutions [19]. These models demonstrated superior performance by automatically learning hierarchical representations that integrate spectral and spatial cues.

Subsequent work explored alternative deep architectures, including pixel-pair networks, recurrent neural networks, and residual learning frameworks, aiming to improve contextual modeling and reduce overfitting [20][21]. While these deep models achieved notable accuracy gains, their performance was often constrained by the limited number of labeled hyperspectral samples available in real-world scenarios. This limitation highlighted the need for complementary strategies such as data augmentation, regularization, and semi-supervised learning.

2.3. Data Augmentation and Generative Modeling

To mitigate the scarcity of labeled hyperspectral data, researchers have investigated synthetic data generation techniques. Autoencoder-based models were among the earliest generative approaches used to learn compact latent representations and reconstruct hyperspectral samples [22]. These models contributed to feature learning stability but were limited in their ability to generate diverse and realistic samples.

Generative adversarial networks (GANs) subsequently emerged as a more powerful alternative for hyperspectral data synthesis. By formulating data generation as an adversarial game between a generator and a discriminator, GAN-based methods demonstrated improved capability in modeling complex spectral distributions [23]. These approaches showed promise in enhancing classification performance through adversarial augmentation, particularly when combined with discriminative classifiers. Nevertheless, challenges such as training instability and mode collapse remain active research topics.

2.4. Domain Adaptation and Cross-Scene Learning

Another important line of research focuses on domain adaptation, which addresses performance degradation caused by differences in sensors, atmospheric conditions, and geographical regions. Early domain adaptation techniques for remote sensing relied on feature alignment and statistical distribution matching between source and target domains [24]. These methods reduced domain shift but often assumed shared feature spaces and sufficient target-domain data.

More recent approaches incorporated graph-based learning, manifold alignment, and attention mechanisms to better capture class-specific correspondences across domains [25]. Although these methods improved cross-scene robustness, they were typically studied independently from data augmentation strategies. As a result, the joint exploitation of adversarial data synthesis and domain-aware learning remains insufficiently explored.

2.5. Summary and Research Gap

In summary, prior work demonstrates that no single strategy—whether deep representation learning, generative augmentation, or domain adaptation—fully resolves the challenges

inherent in hyperspectral image classification. Classical methods lack representational capacity, deep models suffer from data scarcity, generative models face stability issues, and domain adaptation techniques often ignore sample imbalance. These observations motivate the development of integrated frameworks that combine adversarial data generation with classification-oriented learning to improve robustness and generalization under limited and heterogeneous training conditions.

3. Methodology

This section presents the proposed adversarial learning framework for hyperspectral image (HSI) classification. The method integrates generative modeling and supervised classification to address the challenges of high spectral dimensionality and limited labeled data.

3.1. Problem Definition and Framework Overview

Let

$$x = \{x_i \in R^B \mid i = 1, \dots, N\}$$

denote a hyperspectral dataset with B spectral bands. Each labeled sample is associated with a class label

$$y_i \in \{1, \dots, C\},$$

where C is the number of classes.

Given a limited labeled training set

$$D_L = \{(x_i, y_i)\}_{i=1}^{N_L} \quad N_L \ll N,$$

the objective is to learn a classifier

$$f = R^B \rightarrow \{1, \dots, C\}$$

that generalizes well to unseen samples.

To mitigate the scarcity of labeled data, the proposed framework introduces an adversarial spectral generation module that synthesizes additional training samples. These synthetic samples are used to augment the original labeled set, enabling more robust classifier learning.

3.2. Adversarial Spectral Data Generation

The generative component is based on a Generative Adversarial Network (GAN) consisting of a generator G and a discriminator D .

The generator maps a latent noise vector

$$z \sim p_z(z), z \in R^Z,$$

to a synthetic hyperspectral spectrum:

$$\bar{x} = G(z; \theta_G) \in [0, 1],$$

indicating whether the input spectrum is real or generated.

The adversarial learning objective is defined as:

$$\frac{\min}{G} \frac{\max}{D} E_{x \sim p_{data}} [\log D(x)] + E_{z \sim p_z} [\log(1 - D(G(z)))].$$

Through this minimax optimization, the generator learns to approximate the real spectral distribution. After training, the generator produces a synthetic dataset

$$D_s = \{(\tilde{x}_j, \tilde{y}_j)\}_{j=1}^{N_s}$$

which is combined with the labeled data to form the augmented training set

$$D_A = D_L \cup D_S$$

3.3. Classification and Learning Algorithm

Using the augmented dataset D_A , a supervised classifier is trained. Two classifiers are considered in this study: a support vector machine (SVM) and a neural network (NN).

For the SVM, the optimization problem is formulated as:

$$\frac{\min}{w, b, e} \frac{1}{2} \|w\|^2 + C \sum_i \xi_i$$

subject to:

$$y_i (w \phi(x_i) + b) \geq 1 - \xi_i, \xi_i \geq 0.$$

For the NN classifier, parameters are optimized by minimizing the cross-entropy loss:

$$L_{cls} = \sum_{i=1}^{|D_A|} \sum_{c=1}^c I(y_i = c) \log p_c(x_i)$$

The overall training procedure is summarized in Algorithm 1.

Algorithm 1: Adversarial Learning Framework for HSI Classification

Input: Labeled dataset D_L , latent dimension z , training epochs T

Output: Trained classifier $f(\cdot)$

1. Initialize generator G and discriminator D
2. for $t=1$ to T
 - a. Sample real spectra $x \sim p_{data}$
 - b. Sample noise vectors $z \sim p_z$
 - c. Generate synthetic spectra $\tilde{x} = G(z)$
 - d. Update discriminator D by maximizing adversarial loss
 - e. Update generator G by minimizing adversarial loss
3. Generate synthetic dataset D_S using trained G
4. Form augmented dataset $D_A = D_L \cup D_S$
5. Train classifier $f(\cdot)$ on D_A
6. Return trained classifier

4. Experimental Setup and Datasets

4.1. Datasets

Experiments are conducted on widely used benchmark hyperspectral datasets to ensure comparability with prior studies. Each dataset consists of a hyperspectral image cube accompanied by ground-truth labels for a subset of pixels. Only labeled pixels are used for training and evaluation, while unlabeled pixels are excluded from the analysis.

All datasets exhibit high spectral dimensionality and class imbalance, which makes them suitable for evaluating learning performance under limited labeled data conditions. Prior to training, noisy or water absorption bands are removed following standard practice, and all spectral vectors are normalized to zero mean and unit variance on a per-band basis.

4.2. Experimental Protocol

To simulate realistic small-sample learning scenarios, a fixed proportion of labeled samples per class is randomly selected for training, while the remaining labeled samples are reserved for testing. This per-class sampling strategy prevents dominance by majority classes and ensures a fair comparison across methods.

The experimental protocol is defined as follows:

- A fixed number of labeled samples is selected from each class for training.
- The remaining labeled samples are used exclusively for testing.
- The random sampling process is repeated multiple times to reduce statistical variance.

Final results are reported as the mean and standard deviation across all runs.

Synthetic samples generated by the adversarial model are used only during training and are not included in the testing phase. This protocol ensures that evaluation strictly reflects generalization to real hyperspectral data.

4.3. Evaluation Metrics

Classification performance is evaluated using standard metrics commonly adopted in hyperspectral image analysis:

$$OA = \frac{1}{N_{\text{test}}} \sum_{i=1}^{N_{\text{test}}} I(\hat{y}_i = y_i),$$

where N_{test} is the number of test samples.

Average Accuracy (AA):

$$AA = \frac{1}{C} \sum_{c=1}^C \frac{N_c^{\text{correct}}}{N_c}$$

mean per-class accuracy and accounts for class imbalance.

$$F1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

which balances precision and recall and is particularly informative for minority classes.

In addition, the confusion matrix is analyzed to qualitatively assess class-wise misclassification patterns.

4.4. Implementation Details

The adversarial network is trained using the Adam optimizer with a fixed learning rate. The generator and discriminator are updated alternately within each training epoch. Training is terminated once the adversarial loss stabilizes.

For the classifier stage, hyperparameters are selected using a validation subset drawn from the training data. The same hyperparameter configuration is applied consistently across all experiments to ensure fairness.

All experiments are conducted on a standard workstation equipped with a GPU. To guarantee reproducibility, random seeds are fixed for data sampling, network initialization, and optimization.

4.5. Summary

The experimental design emphasizes reproducibility, fair comparison, and robustness under limited labeled data conditions. By employing standard benchmark datasets, widely accepted evaluation metrics, and repeated random sampling, the proposed framework is evaluated in a controlled and statistically meaningful manner.

5. Results and Quantitative Analysis

This section presents quantitative results evaluating the proposed adversarial learning framework. Performance is assessed using overall accuracy (OA), average accuracy (AA), and F1-score under limited labeled data conditions. Results are reported as the mean and standard deviation over repeated trials, following the protocol described in Section 4.

5.1. Overall Classification Performance

Table 1 summarizes the classification performance of different methods on the evaluated hyperspectral datasets. The proposed framework is compared against baseline classifiers trained without adversarial data augmentation.

Table 1; Classification Performance (Mean ± Standard Deviation)

Method	OA (%)	AA (%)	F1-score
SVM (Original Data)	82.41 ± 1.97	78.63 ± 2.45	0.801
NN (Original Data)	85.76 ± 1.62	81.92 ± 2.11	0.832
SVM + GAN Augmentation	88.93 ± 1.34	86.15 ± 1.78	0.867
NN + GAN Augmentation (Proposed)	91.27 ± 1.08	89.44 ± 1.36	0.902

The results indicate that adversarial data augmentation consistently improves classification performance across both classifiers. The proposed NN-based framework achieves the highest OA, AA, and F1-score, with reduced variance across trials.

5.2. Class-wise Performance Analysis

To further analyze the impact of adversarial augmentation, Table 2 reports the average per-class F1-scores across all datasets.

Table 2: Average Class-Wise F1-Score

Class Group	SVM	NN	SVM + GAN	NN + GAN
Majority Classes	0.862	0.884	0.901	0.924
Minority Classes	0.711	0.743	0.801	0.836

The improvement is particularly notable for minority classes, where limited training samples typically lead to poor generalization. This suggests that synthetic spectra generated by the adversarial model effectively increase data diversity for underrepresented classes.

5.3. Statistical Significance Analysis

To assess whether the observed performance improvements are statistically significant, paired statistical tests are conducted between the proposed method and baseline approaches across repeated trials.

A paired t-test is performed on OA values, with the null hypothesis that there is no difference in mean performance between methods.

- NN vs. NN + GAN: $p < 0.01$
- SVM vs. SVM + GAN: $p < 0.05$

These results indicate that the performance gains introduced by adversarial augmentation are statistically significant at conventional confidence levels.

5.4. Robustness under Limited Training Samples

To evaluate robustness, experiments are repeated with varying numbers of labeled samples per class. Table 3 reports OA under increasingly constrained training conditions.

Table 3: OA (%) under Different Training Sample Sizes

Samples per Class	NN	NN + GAN
5	72.84	78.91
10	81.35	86.72
20	88.02	91.27

The proposed framework demonstrates consistent performance gains, particularly in extremely low-sample regimes, indicating improved generalization when training data are scarce.

5.5. Discussion of Quantitative Findings

The quantitative results demonstrate three key observations.

First, adversarial data augmentation improves both accuracy and stability, as reflected by higher mean performance and lower standard deviation. Second, the gains are more pronounced for minority classes, suggesting that the generated spectra capture meaningful class-specific variability. Third, the proposed framework maintains robustness across

different training sample sizes, highlighting its suitability for real-world hyperspectral scenarios where labeled data are limited.

It should be noted that while the proposed method consistently outperforms baseline models, performance gains vary across datasets and classes. This suggests that the effectiveness of adversarial augmentation may depend on the intrinsic spectral variability of the scene. Further investigation into dataset-specific characteristics is warranted.

6. Qualitative Analysis

This section provides a qualitative assessment of classification behavior using confusion matrices and error pattern analysis. While quantitative metrics summarize overall performance, qualitative inspection offers insight into class-wise confusion, misclassification tendencies, and the effect of adversarial augmentation on decision boundaries.

6.1. Confusion Matrix Analysis

Figure 2 illustrates representative confusion matrices for the neural network classifier trained (a) without adversarial augmentation and (b) with the proposed GAN-based augmentation. For clarity, results are shown for a typical experimental run using limited training samples.

In the baseline case, misclassifications are predominantly observed among spectrally similar classes. These confusions are particularly evident between vegetation-related classes and between materials exhibiting overlapping reflectance characteristics. Minority classes tend to suffer from both low recall and low precision, indicating insufficient representation during training.

In contrast, the confusion matrix obtained using adversarial augmentation shows a clearer diagonal structure, with reduced off-diagonal entries. Improvements are especially noticeable for minority and transitional classes, where recall increases without a corresponding rise in false positives. This suggests that the synthetic spectra generated by the adversarial model enhance class separability in spectral feature space.

6.2. Error Pattern Characterization

A closer examination of misclassified samples reveals consistent error patterns across datasets. In the baseline setting, errors frequently occur at class boundaries and in regions affected by mixed pixels or spectral variability caused by illumination and atmospheric effects. Such errors indicate limited robustness of classifiers trained solely on scarce real samples.

With adversarial augmentation, many of these boundary-related errors are reduced. Generated samples appear to regularize the classifier by exposing it to a wider range of plausible spectral variations during training. As a result, the classifier exhibits smoother decision boundaries and reduced sensitivity to local spectral fluctuations.

However, certain misclassification patterns persist even after augmentation. Classes with extremely high intra-class variability or strong spectral overlap remain challenging. This observation indicates that while adversarial augmentation improves robustness, it does not fully resolve ambiguities inherent in the spectral signatures themselves.

6.3. Impact on Minority Classes

Qualitative inspection further confirms the quantitative findings regarding minority classes. In baseline confusion matrices, minority classes are often absorbed into spectrally dominant neighboring classes. After augmentation, these classes exhibit improved recognition, with fewer samples misassigned to majority categories.

This behavior suggests that adversarially generated samples help balance class representation and prevent classifier bias toward majority classes. Nevertheless, improvements are not uniform across all minority classes, implying that the effectiveness of augmentation depends on the degree to which synthetic spectra reflect true class variability.

6.4. Discussion

The qualitative analysis supports the quantitative results by demonstrating that adversarial augmentation not only improves aggregate performance metrics but also leads to more interpretable and stable classification behavior. Reduced confusion among spectrally similar classes and improved minority-class recognition indicate enhanced generalization.

At the same time, the persistence of certain error patterns highlights inherent limitations of purely spectral-based approaches. Incorporating additional spatial context or domain adaptation mechanisms may further reduce these residual errors.

7. Discussion and Limitations

This section discusses the experimental findings in light of both quantitative metrics and qualitative observations, and outlines the limitations of the proposed framework.

7.1. Discussion of Results

The quantitative results presented in Section 5 demonstrate that the proposed adversarial learning framework consistently improves classification performance under limited labeled data conditions. Increases in overall accuracy, average accuracy, and F1-score indicate that adversarial spectral augmentation contributes positively to both global and class-wise performance. These gains are further supported by the reduced variance observed across repeated trials, suggesting improved training stability.

The qualitative analysis in Section 6 provides additional insight into the mechanisms underlying these improvements. Confusion matrix inspection reveals a clearer diagonal structure when adversarial augmentation is applied, indicating enhanced class separability. This observation aligns with the quantitative improvements in average accuracy and F1-score, particularly for minority classes. The reduction of off-diagonal errors suggests that synthetic spectra help the classifier better model intra-class variability and inter-class boundaries.

Moreover, robustness experiments under varying training sample sizes show that the proposed framework maintains performance advantages in extremely low-sample regimes. Qualitatively, this robustness is reflected in smoother decision boundaries and fewer boundary-related misclassifications. These findings collectively suggest that adversarial augmentation acts as an effective regularization mechanism, exposing the classifier to a broader range of plausible spectral variations during training.

7.2. Interpretation of Minority-Class Improvements

A notable outcome of the experiments is the improved recognition of minority classes. Quantitatively, this is reflected in higher per-class F1-scores, while qualitatively, confusion matrices show fewer minority samples being absorbed into dominant classes. This behavior indicates that adversarially generated samples partially mitigate class imbalance by increasing effective sample diversity.

However, the degree of improvement varies across classes. Classes with high intra-class spectral variability or strong overlap with neighboring classes remain challenging, even with augmentation. This observation suggests that while adversarial synthesis improves data representation, it does not fully compensate for intrinsic spectral ambiguity.

7.3. Limitations

Despite the demonstrated improvements, several limitations of the proposed framework should be acknowledged.

First, the quality of synthetic spectra is inherently dependent on the stability of adversarial training. Although training convergence was observed in the conducted experiments, generative adversarial networks are known to be sensitive to hyperparameter selection and network architecture. In practice, careful tuning may be required to avoid mode collapse or the generation of unrealistic spectra.

Second, the current framework focuses primarily on spectral information and does not explicitly incorporate spatial context during generation. As a result, certain spatially induced misclassifications persist, as observed in the qualitative analysis. Integrating spatial constraints or spectral-spatial generative models may further improve performance.

Third, the experimental evaluation is conducted on commonly used benchmark datasets. While these datasets are representative, they may not fully capture the diversity of real-world hyperspectral acquisition conditions. Generalization to different sensors or extreme environmental conditions requires further investigation.

Finally, computational overhead associated with adversarial training, although manageable in offline settings, may limit applicability in resource-constrained or real-time scenarios.

7.4. Future Directions

The identified limitations suggest several avenues for future research. Incorporating spatial information into the generative process, exploring more stable adversarial objectives, and extending the framework to domain-adaptive or self-supervised settings may further enhance robustness. Additionally, evaluation on large-scale and cross-sensor datasets would provide deeper insight into real-world applicability.

By jointly considering quantitative metrics and qualitative evidence, this discussion clarifies how adversarial spectral augmentation improves hyperspectral image classification while also highlighting its limitations. The results indicate that the proposed framework is a practical and effective approach under limited labeled data conditions, while leaving room for further methodological refinement.

8. Conclusion and Future Work

This paper presented an adversarial learning framework for hyperspectral image classification that integrates generative spectral augmentation with supervised classifiers. The

proposed approach was designed to address the challenges of high spectral dimensionality and limited labeled training data, which commonly constrain hyperspectral analysis in practical applications.

Experimental results demonstrated that adversarial spectral augmentation consistently improves classification performance across multiple evaluation metrics. Quantitative gains in overall accuracy, average accuracy, and F1-score were observed, accompanied by reduced performance variance across repeated trials. Qualitative analysis further revealed improved class separability, reduced confusion among spectrally similar classes, and enhanced recognition of minority classes. These findings indicate that the generated spectra effectively enrich the training distribution and contribute to more stable and robust classifier learning.

Despite these improvements, the proposed framework does not fully resolve misclassification issues arising from strong spectral overlap or high intra-class variability. In addition, adversarial training introduces additional computational cost and requires careful parameter tuning to ensure stable convergence.

Future work will focus on incorporating spatiol context into the generative process, exploring more stable adversarial objectives, and extending the framework to domain-adaptive and cross-sensor classification scenarios. Evaluation on larger-scale and more diverse hyperspectral datasets is also planned to further assess real-world applicability.

References

- [1] Wambugu, N., Chen, Y., Xiao, Z., Tan, K., Wei, M., Liu, X., & Li, J. (2021). Hyperspectral image classification on insufficient-sample and feature learning using deep neural networks: A review. *International Journal of Applied Earth Observation and Geoinformation*, 105, 102603. DOI:10.1016/j.jag.2021.102603.
- [2] Kumar, V., Singh, R. S., Rambabu, M., & Dua, Y. (2024). Deep learning for hyperspectral image classification: A survey. *Computer Science Review*, 53, 100658. DOI:10.1016/j.cosrev.2024.100658.
- [3] Yu, Z., & Cui, W. (2024). Robust hyperspectral image classification using generative adversarial networks. *Information Sciences*, 666, 120452. DOI:10.1016/j.ins.2024.120452.
- [4] Lyu, H., Grafton, M., Ramilan, T., Irwin, M., & Sandoval, E. (2025). Synthetic hyperspectral reflectance data augmentation by generative adversarial network to enhance grape maturity determination. *Computers and Electronics in Agriculture*, 235, 110341. DOI:10.1016/j.compag.2025.110341.
- [5] Zhang, X., Gao, K., Wang, J., Hu, Z., Wang, H., Wang, P., Zhao, X., & Li, W. (2023). Self-supervised learning with deep clustering for target detection in hyperspectral images with insufficient spectral variation prior. *International Journal of Applied Earth Observation and Geoinformation*, 122, 103405. DOI:10.1016/j.jag.2023.103405.
- [6] Xia, M., Yuan, G., Yang, L., Xia, K., Ren, Y., Shi, Z., & Zhou, H. (2022). Few-shot hyperspectral image classification based on convolutional residuals and SAM Siamese networks. *Electronics*, 12(16), 3415. DOI:10.3390/electronics12163415.
- [7] Fang, X., Zhang, G., Zhang, G., Zhou, X., Wu, J., & Zhao, L. (2023). A hybrid self-supervised learning framework for hyperspectral image classification. In *Proceedings of the International Conference on Computer, Vision and Intelligent Technology (ICCVIT 2023)* (pp. 1–11). ACM, New York, NY, USA. DOI:10.1145/3627341.3631083.

- [8] Arshad, T., Zhang, J., & Ullah, I. (2024). A hybrid convolution transformer for hyperspectral image classification. *European Journal of Remote Sensing*, 57(1). DOI:10.1080/22797254.2024.2330979.
- [9] Li, Z., Hong, D., Gao, L., Yao, J., Zhang, B., & Chanussot, J. (2024). Attention-based subdomain adaptation network for hyperspectral image classification. *IEEE Transactions on Neural Networks and Learning Systems*, 35(4), 4872–4886. DOI:10.1109/TNNLS.2023.3284421.
- [10] Xin, Z., Li, Z., Xu, M., Wang, L., Ren, G., Wang, J., & Hu, Y. (2024). Feature disentanglement based domain adaptation network for cross-scene coastal wetland hyperspectral image classification. *International Journal of Applied Earth Observation and Geoinformation*, 129, 103850. DOI:10.1016/j.jag.2024.103850.
- [11] Zhang, Z., Gao, D., Liu, D., & Shi, G. (2023). Spectral–spatial domain attention network for hyperspectral image few-shot classification. *Remote Sensing*, 16(3), 592. DOI:10.3390/rs16030592.
- [12] Feng, W., Long, Y., Dauphin, G., Quan, Y., Huang, W., & Xing, M. (2024). Ensemble synthetic oversampling with pixel pair for class-imbalanced and small-sized hyperspectral data classification. *International Journal of Applied Earth Observation and Geoinformation*, 128, 103697. DOI:10.1016/j.jag.2024.103697.
- [13] Quan, D., Feng, W., Dauphin, G., Wang, X., Huang, W., & Xing, M. (2021). A novel double ensemble algorithm for the classification of multi-class imbalanced hyperspectral data. *Remote Sensing*, 14(15), 3765. DOI:10.3390/rs14153765.
- [14] Zhang, W., Hu, R., Wang, J., Zhang, L., & Zhu, C. (2024). Spectral prototype attention domain adaptation for hyperspectral image classification. *Remote Sensing*, 17(23), 3901. DOI:10.3390/rs17233901.
- [15] Lv, Q., Feng, W., Quan, Y., Li, H., & Chanussot, J. (2021). Ensemble CNN with enhanced feature subspaces for imbalanced hyperspectral image classification. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2021)* (pp. 3669–3672). DOI:10.1109/IGARSS47720.2021.9554677.
- [16] Melgani, F., & Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 42(8), 1778–1790. DOI:10.1109/TGRS.2004.831865.
- [17] Camps-Valls, G., & Bruzzone, L. (2005). Kernel-based methods for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 43(6), 1351–1362. DOI:10.1109/TGRS.2005.846154.
- [18] Zhang, H., He, W., Zhang, L., Shen, H., & Yuan, Q. (2014). Hyperspectral image restoration using low-rank matrix recovery. *IEEE Transactions on Geoscience and Remote Sensing*, 52(8), 4729–4743. DOI:10.1109/TGRS.2013.2284280.
- [19] Chen, Y., Jiang, H., Li, C., Jia, X., & Ghamisi, P. (2016). Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 54(10), 6232–6251. DOI:10.1109/TGRS.2016.2584107.
- [20] Li, W., Wu, G., Zhang, F., & Du, Q. (2017). Hyperspectral image classification using deep pixel-pair features. *IEEE Transactions on Geoscience and Remote Sensing*, 55(2), 844–853. DOI:10.1109/TGRS.2016.2616355.
- [21] Mou, L., Ghamisi, P., & Zhu, X. X. (2017). Deep recurrent neural networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(7), 3639–3655. DOI:10.1109/TGRS.2016.2636241.

- [22] Ma, X., Wang, H., & Geng, J. (2016). Spectral–spatial classification of hyperspectral image based on deep auto-encoder. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(9), 4073–4085. DOI:10.1109/JSTARS.2016.2517204.
- [23] Zhu, L., Chen, Y., Ghamisi, P., & Benediktsson, J. A. (2018). Generative adversarial networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 56(9), 5046–5063. DOI:10.1109/TGRS.2018.2805286.
- [24] Tuia, D., Persello, C., & Bruzzone, L. (2016). Domain adaptation for the classification of remote sensing data: An overview of recent advances. *IEEE Geoscience and Remote Sensing Magazine*, 4(2), 41–57. DOI:10.1109/MGRS.2016.2548504.
- [25] Hong, D., Yokoya, N., Ge, N., Chanussot, J., & Zhu, X. X. (2018). Learnable manifold alignment (LeMA): A semi-supervised cross-modality learning framework for land cover and land use classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 147, 193–205. DOI:10.1016/j.isprsjprs.2018.10.006