

Article

A Graph-Regularized Double-Path Interactive Spectral Super-Resolution Network for Hyperspectral Image Reconstruction

Shuo Wang ¹, Ting Hu ^{1,*}, Siyuan Cheng ², Zhe Li ¹, Zhonghua Sun ¹, Kebin Jia ¹ and Jinchao Feng ¹¹ Beijing Key Laboratory of Computational Intelligence and Intelligent System, School of Information Science and Technology, Beijing University of Technology, Beijing 100124, China; wshmz@emails.bjut.edu.cn (S.W.)² Space Star Technology Co., Ltd., Beijing 100089, China

* Correspondence: hut455@bjut.edu.cn

Highlights

What are the main findings?

- A graph-regularized double-path spectral super-resolution network is proposed to enhance spectral fidelity in MSI-to-HSI reconstruction.
- Interactive and enhanced residual modules enable effective spectral–spatial feature interaction and fusion.

What are the implications of the main finding?

- Spectral graph regularization alleviates the ill-posedness of spectral super-resolution in remote sensing.
- The proposed framework achieves robust spectral enhancement on both simulated and real datasets.

Abstract

Deep learning has demonstrated outstanding potential for the spectral super-resolution (S^2R) reconstruction of multispectral images (MSIs). However, it is still a challenge to alleviate spectral distortion during S^2R reconstruction. Given the superiority of a graph, a graph-regularized double-path interactive S^2R network (GDIS²Net) consisting of two parallel branches is proposed to reconstruct hyperspectral images (HSIs) from MSIs. An interactive residual module is carefully schemed as the backbone of the S^2R network to facilitate the feature interaction between the two branches, while an enhanced residual module is constructed for further feature fusion. In addition, a new loss function considering the spectral continuity is proposed to optimize the proposed GDIS²Net. Experimental analyses show that the proposed GDIS²Net outperforms state-of-the-art methods on both simulated and real datasets.



Academic Editor: Akira Iwasaki

Received: 21 January 2026

Revised: 2 March 2026

Accepted: 9 March 2026

Published: 12 March 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 Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: spectral super-resolution; hyperspectral image; multispectral image; deep learning; graph signal

1. Introduction

Hyperspectral images (HSIs) of numerous narrow spectral bands can effectively distinguish materials with subtle differences, making them widely used in remote sensing [1–3], target classification [4–6], agriculture [7–9], and geology [10–12]. Due to the trade-offs among spatial resolution, spectral resolution, and signal-to-noise ratio (SNR) [13,14],

HSIs are often low-resolution, while multispectral images (MSIs) are usually high-resolution (in this paper, high-resolution and low-resolution are only used to describe the spatial resolution of an image) [3,15]. In order to generate high-resolution HSIs (H^2 SI), a promising and cost-effective technology, i.e., spectral super-resolution (S^2 R) for MSIs [16], has been developed recently, aiming to recover high spectral detail from MSIs.

Existing S^2 R methods can be broadly categorized into sparse representation-based [17–20] and deep learning (DL)-based methods [21–23]. S^2 R is inherently a highly ill-posed problem [24], as the available spectral bands in MSIs are much fewer than the spectral requirements when performing H^2 SI reconstruction, causing underdetermined and ambiguous reconstruction [18,25]. To alleviate this ill-posedness, sparse representation-based S^2 R methods [19,24] have been proposed, which assume that H^2 SI share the same set of sparse coefficients and the same high-spectral dictionaries with the corresponding high-resolution MSIs (HMSIs) and low-resolution HSIs (LHSIs), respectively. Thus, H^2 SI can be produced by integrating sparse coefficients into the high-spectral dictionaries learned from pairs of HMSIs/LHSIs, where the coefficients are inferred from HMSIs via the learned dictionaries. The above sparse representation-based methods have achieved success towards solution stabilizing after introducing some effective prior constraints, such as low rankness [20] and sparsity [24–26]. However, the spatial–spectral continuity inherent in spectral images was ignored, as 3D images were usually unfolded into 2D matrices. Moreover, the reconstruction process often relies on an iterative optimization of expensive computation.

In recent years, DL has been introduced into the S^2 R task, which usually executes S^2 R by learning the reconstruction mapping from sufficient pairs of HMSIs/ H^2 SI in an end-to-end manner. Compared to sparse representation-based methods, DL approaches generally consume less time because the computational burden is shifted to the training phase. In addition, DL has been revealed to exhibit superior S^2 R performance because of its outstanding abilities of feature extraction and non-linear expression [27–31]. Considering the difficulty of obtaining high-resolution ground truth(H^2 SI) in real-world scenarios, Chen et al. [32] proposed a spectral–spatial residual network (SSRN). More specifically, the mapping learned from observed LHSIs and simulated low-resolution MSIs (LMSIs) can be generalized to real HMSIs. However, achieving high spectral fidelity when performing the S^2 R task is still a challenge.

To achieve better spectral enhancement for real MSIs, a graph-regularized double-path interactive spectral super-resolution network (GDIS²Net) is proposed. GDIS²Net is designed by connecting an S^2 R subnet and a spectral graph reconstruction (SGR) subnet in parallel, where the latter is used to optimize the S^2 R performance, because graph signal processing has shown superior characterization ability of the spectral continuity in high-dimensional images [33]. An interactive residual module (IRM) is designed to bridge the two subnets for level-by-level feature interaction. In addition, an enhanced residual module (ERM) is proposed for the final feature fusion. Under our spectral graph-assisted strategy, the proposed GDIS²Net is optimized using hyperspectral graph supervision in addition to the hybrid supervision of LHSIs and LMSIs. Correspondingly, a loss function regularized by the spectral graph reconstruction is used to optimize our network.

The main contributions of this article can be summarized as follows:

- (1) A graph-regularized double-path interactive S^2 R network that attempts to improve spectral enhancement by reconstructing a hyperspectral graph from a blurry one is proposed. Furthermore, a spectral graph reconstruction is integrated as a regularization when designing the loss for network training. Through graph signal processing, the spectral continuity inherent in HSIs is explicitly captured, and spectral distortion is effectively mitigated.

- (2) An interactive residual module is designed to achieve level-by-level feature interaction, where the spatial attention mechanism is used to guide feature interaction in a cross way.
- (3) An enhanced residual module is developed to refine the final feature fusion, which executes feature extraction across different receptive fields in a double-path manner, along with spatial attention-based critical feature emphasizing.

The subsequent sections are organized as follows. Section 2 presents the related studies. The proposed GDIS²Net is described in detail in Section 3. In Section 4, the experimental results and analyses are presented. Finally, the conclusion is summarized in Section 5.

2. Related Work

2.1. Deep S²R Methods

Since Galliani et al. [21] first applied a deep neural network for S²R, DL-based spectral resolution enhancement has recently gained increasing attention. For instance, a unified deep learning framework named HSCNN was published [22] to reconstruct H²SI from spectral undersampled projections. Subsequently, a multi-scale deep convolutional neural network (MSCNN) was established on the symmetrical downsampling and upsampling operations, further enhancing the reconstruction performance [23]. To ensure the texture preservation and spectral fidelity as well as improve the robustness and generalization of the S²R algorithms, a variety of different network architectures have been proposed [34,35], such as a deep convolutional neural network (CNN) composed of numerous residual blocks (HSCNN+) [27], a pixel-aware deep function-mixture network (FMNet) dynamically adjusting receptive fields and mapping functions [28], and a four-level hierarchical regression network (HRNet) executing inter-level feature interactions via pixel shuffle [29]. Given that these remarkable S²R deep models involve substantial parameters and require high computational costs during the training phase, some lightweight nets have been developed. A lightweight residue-dense attention network for efficient spectral feature prediction was constructed [30], which utilized dual-path feature extraction and coordination convolutional blocks to markedly reduce model complexity. Exploiting spectral-wise multi-head self-attention mechanism to model the long-range dependencies across different spectral bands, Cai et al. [31] proposed a multi-level transformer-based architecture (MST++). In addition, a reparameterizing coordinate-preserving proximity spectral interaction network (RepCPSI) [36] was reported for lightweight S²R, where the spatial-spectral context features were explored via a multi-channel polymorphic residual context restructuring module.

Obviously, existing studies have demonstrated the effectiveness and superiority of DL when performing spectral reconstruction. However, it remains a significant challenge to apply the above deep S²R networks in real-world scenarios, as the necessary H²SI labels are actually unavailable. Fortunately, a hybrid supervision strategy was published in [32], enabling S²R to be performed without H²SI labels. Based on this, a spectral graph-aided spectral reconstruction network is proposed to ensure a better spectral fidelity by introducing the graph signal theory, for the first time, into spectrum prediction.

2.2. Attention Mechanisms in S²R

The human brain is able to automatically focus on salient regions while processing complex information [37], which enlightens the birth of the attention mechanism. Nowadays, the attention mechanism has been demonstrated with significant benefits for enhancing feature representation, particularly in image recognition and restoration tasks [38–40]. Based on the squeeze-and-excitation module [38] that explicitly modeled the interdependencies among feature channels, a hybrid residual attention module [41] was developed to enhance the spectral reconstruction effect. With this first success of attention mechanism in the S²R

task, various attention-based image spectral enhancement methods have been proposed. For instance, the channel attention module was introduced to characterize the spectral differences between different bands [42]. However, Zheng et al. [16] found that most attention mechanisms focus on the contextual features of one certain dimension, such as channel or space [43]. Consequently, complementary information from other dimensions is ignored, which may compromise spatial–spectral consistency. For this, they designed a spatial–spectral residual attention network (SSRAN), which adopted the 1D convolution to generate attention weights to enhance the correlation between neighboring spectral bands. Then, a coordinate-preserving attention module [36] was further designed to selectively emphasize discriminative spatial–spectral features while suppressing redundant information.

Motivated by the benefits of the attention mechanism in improving spatial and spectral feature representations, two new modules, i.e., IRM and ERM, established on the spatial attention (SA) mechanism, are designed to promote interaction and fusion between two-subnet features in this paper.

3. Proposed Method

3.1. Overview

Referring to Wald’s protocol applied in [32], the GDIS²Net is proposed to achieve spectral enhancement for LMSIs via reconstruction mapping from LMSI to LHSI.

Let H^2SI , $HMSI$, and $LHSI$ be successively denoted as $\mathcal{Z} \in \mathbb{R}^{H \times W \times C}$, $\mathcal{X} \in \mathbb{R}^{H \times W \times c}$, and $\mathcal{Y} \in \mathbb{R}^{h \times w \times C}$, where $H \gg h$, $W \gg w$, and $C \gg c$ are the rows, columns, and bands of the image, respectively. Then, LMSI $\mathcal{X}^\downarrow \in \mathbb{R}^{h \times w \times c}$ is generated by spatially degrading $HMSI$ $\mathcal{X}$ via the point spread function (PSF):

$$\mathcal{X}^\downarrow = \Theta(\mathcal{X} * \mathbf{P}), \quad (1)$$

where $*$ denotes spatial convolution, $\Theta(\cdot)$ is the downsampling operator of interval r , $r = \frac{H}{h} = \frac{W}{w}$, and $\mathbf{P} \in \mathbb{R}^{r \times r}$ represents the PSF. It is theoretically suggested that the S²R mapping learned from $\mathcal{X}^\downarrow \rightarrow \mathcal{Y}$ can be generalized to $\mathcal{X} \rightarrow \mathcal{Z}$.

Attracted by the outstanding ability of graphs to characterize spectral continuity, an SGR subnet (see details in Section 3.2) is further designed to optimize the S²R process, utilizing the spectral graph as an auxiliary reconstruction target. Specifically, a pre-estimated LHSI $\hat{\mathcal{Y}} \in \mathbb{R}^{h \times w \times C}$ is obtained by interpolating $\mathcal{X}^\downarrow$, and its spectral graph $\hat{\mathcal{Y}}_{SG}$ is computed to assist the spectral reconstruction. That is,

$$(\hat{\mathcal{Y}}, \hat{\mathcal{Y}}_{SG}) = \nabla(\mathcal{X}^\downarrow, \hat{\mathcal{Y}}_{SG}), \quad (2)$$

where $\nabla(\cdot)$ describes the trained GDIS²Net, and $\hat{\mathcal{Y}}$ and $\hat{\mathcal{Y}}_{SG}$ denote the reconstructed LHSI and spectral graph, respectively. Such reconstruction is co-supervised via the $(\mathcal{Y}, \mathcal{Y}_{SG})$ pair, where $\mathcal{Y}_{SG}$ is the spectral graph corresponding to $\mathcal{Y}$. In order to preserve spatial information, the S²R process is further self-supervised by the input LMSI $\mathcal{X}^\downarrow$. Therefore, a reconstructed LMSI $\hat{\mathcal{X}}^\downarrow$ is computed from $\hat{\mathcal{Y}}$:

$$\hat{\mathcal{X}}^\downarrow = \hat{\mathcal{Y}} \times_3 \mathbf{R}^T, \quad (3)$$

where $\mathbf{R} \in \mathbb{R}^{C \times c}$ is the spectral response function (SRF). To achieve the above hybrid supervision procedure, a spatial–spectral fidelity loss function is constructed as follows:

$$L = \|\mathcal{Y} - \hat{\mathcal{Y}}\|_1 + \lambda \|\mathcal{X}^\downarrow - \hat{\mathcal{X}}^\downarrow\|_1 + \mu \|\mathcal{Y}_{SG} - \hat{\mathcal{Y}}_{SG}\|_1, \quad (4)$$

where λ and μ are the regularization parameters, and $\|\cdot\|_1$ is the L1-norm. $\|\mathcal{Y} - \hat{\mathcal{Y}}\|_1$ is the spectral fidelity term to ensure spectral enhancement, $\|\mathcal{X}^\downarrow - \widehat{\mathcal{X}^\downarrow}\|_1$ is the spatial fidelity term to ensure spatial preservation, and $\|\mathcal{Y}_{SG} - \hat{\mathcal{Y}}_{SG}\|_1$ is the spectral graph reconstruction term to ensure spectral continuity. Here, the L1-norm but not the common L2-norm is applied, since the former is more sensitive to small errors, so that it has a good edge-preserving ability [44], while the latter usually pays much attention to outliers and then ignores the overall structure [45,46].

Once the GDIS²Net is optimized via the loss function L , a corresponding H²SI $\hat{\mathcal{Z}} \in \mathbb{R}^{H \times W \times C}$ can be restored by inputting any HMSI $\mathcal{X}$ and its corresponding spectral graph $\tilde{\mathcal{Z}}_{SG}$ into the trained GDIS²Net. The spectral reconstruction can be formulated as follows:

$$\hat{\mathcal{Z}} = \nabla(\mathcal{X}, \tilde{\mathcal{Z}}_{SG}). \quad (5)$$

3.2. Spectral Graph

Graphs are widely used in high-dimensional signal processing due to their ability to capture inherent spatial and spectral continuity within signals [47,48]. A weighted undirected graph is defined as $\mathbf{G} = (\mathbf{V}, \mathbf{E}, \mathbf{W})$, where $\mathbf{V}$ and $\mathbf{E}$ denote the vertices and edges, and $\mathbf{W}$ is the symmetric adjacency matrix with element w_{ij} being the weight of the edge e_{ij} between the i th and j th vertices v_i and v_j . There is a diagonal degree matrix $\mathbf{D}$, where its i th diagonal entry d_{ii} equals the sum of weights connected to vertex v_i . The graph Laplacian is then defined as follows:

$$\mathbf{L} = \mathbf{D} - \mathbf{W}, \quad (6)$$

which is a symmetric semi-positive definite matrix.

For a given third-order tensor $\mathcal{A} \in \mathbb{R}^{H \times W \times C}$, a corresponding undirected graph $\mathbf{G} = (\mathbf{V}, \mathbf{E}, \mathbf{W})$ can be constructed by taking each of its bands as a vertex. Then, the spectral continuity of $\mathcal{A}$ can be enforced by the graph Laplacian regularization $\mathcal{J} = \text{tr}(\mathcal{A}_{(3)} \mathbf{L} \mathcal{A}_{(3)}^T)$, where $\mathcal{A}_{(3)}$ denotes the mode-3 unfolding of $\mathcal{A}$ [47]. A smaller value of this term generally indicates better continuity among adjacent vertices; namely, $\mathcal{A} \times_3 \mathbf{L}$ can describe spectral continuity in $\mathcal{A}$. High spectral continuity has already been revealed in HSI [49]. Therefore, in addition to spectral enhancement for LMSI, an SGR is also considered in this paper. Specifically, the spectral graph $\mathcal{A}_{SG}$ of a third-order tensor $\mathcal{A}$ is defined as follows:

$$\mathcal{A}_{SG} = \mathcal{A} \times_3 \mathbf{L}. \quad (7)$$

Following this definition, the spectral graphs of the pre-estimated LHSI $\tilde{\mathcal{Y}}$ and LHSI label $\mathcal{Y}$ are individually calculated as follows:

$$\tilde{\mathcal{Y}}_{SG} = \tilde{\mathcal{Y}} \times_3 \mathbf{L}, \quad (8)$$

$$\mathcal{Y}_{SG} = \mathcal{Y} \times_3 \mathbf{L}, \quad (9)$$

where $\tilde{\mathcal{Y}}_{SG}$ and $\mathcal{Y}_{SG}$ are the spectral graphs of $\tilde{\mathcal{Y}}$ and $\mathcal{Y}$, respectively; $\mathbf{L}$ means that the graphs corresponding to $\tilde{\mathcal{Y}}$ and $\mathcal{Y}$ share the same adjacency matrix. As the common strategy [47], w_{ii} is set to 0, while w_{ij} is set to 1, where j generally denotes n adjacent bands to the i th band, and n is a tunable parameter. As stated in Section 3.1, under the supervision of $\mathcal{Y}_{SG}$, $\tilde{\mathcal{Y}}_{SG}$ is reconstructed as $\hat{\mathcal{Y}}_{SG}$.

As shown in Figure 1, features within SGR interact with the spectral enhancement process in a level-by-level way, and the final SGR and S²R features are fused to improve the spectral fidelity of the super-resolution result. One worth noting is that our approach uses the spectral graph as the reconstruction objective to explicitly facilitate S²R, rather than

simply imposing a graph constraint on HSI reconstruction as existing methodologies [50] have done.

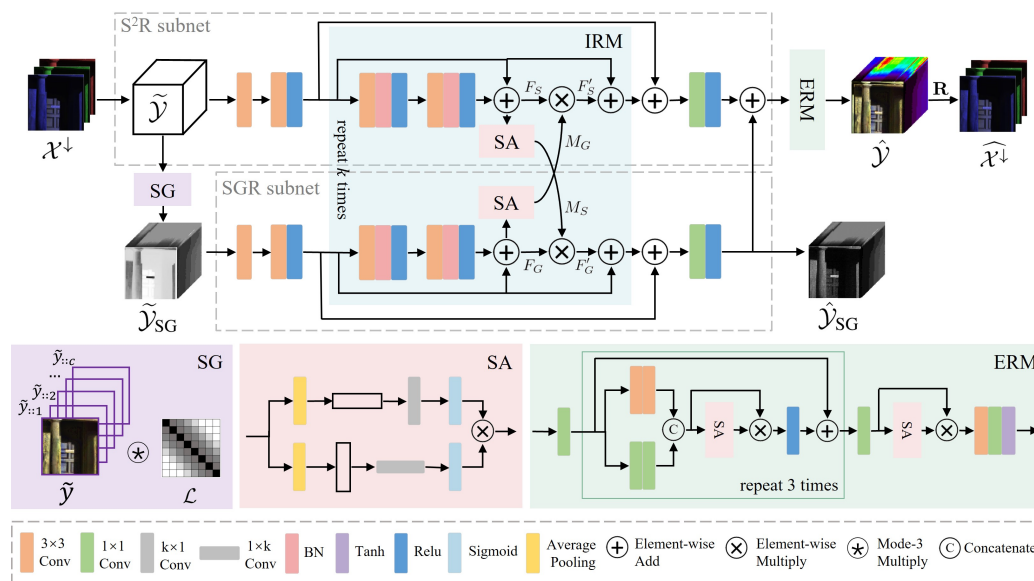


Figure 1. Detailed architecture of our GDIS²Net. SG and SA are the abbreviations for spectral graph and spatial attention, respectively. Our S²R procedure is as follows. The degraded LMSI is interpolated into a coarse LHSI; then, the spectral graph is generated by the SG module. Next, reconstructions for $\mathcal{X}^\downarrow$ and $\tilde{\mathcal{Y}}_{SG}$ are executed via S²R and SGR subnets, respectively, where the feature interaction and fusion between the two subnets are achieved via IRM in a level-by-level way. Finally, ERM further refines the fusion result and generates the S²R target.

3.3. Graph-Regularized Double-Path Interactive Spectral Super-Resolution Network

To reduce the S²R difficulty, the degraded LMSI $\mathcal{X}^\downarrow$ is first interpolated to a coarse LHSI $\tilde{\mathcal{Y}}$ via the cubic interpolation algorithm. Next, the S²R and SGR subnets in our GDIS²Net focus on reconstructing the image and graph residuals between the interpolated result and LHSI labels, respectively. Therefore, the backbone accordingly adopts the residual architecture.

To promote level-by-level feature interaction between the two subnets, several IRMs, of which the detailed structure is shown in the light red box of Figure 1, are designed to consist of two cascaded dense blocks, two element-wise addition operations, one element-wise multiply operator, and one SA unit. Each dense block contains one 3×3 convolutional layer, one batch normalization (BN) layer, and one rectified linear unit (ReLU), which is used for feature extraction. Its subsequent addition operation is used for high- and low-level feature merging. The SA mechanism [36,39] is introduced into the two subnets to help extract the limited and important spectral features, since the number of linearly independent spectra is usually much less than the space size of an image [51,52]. As illustrated in Figure 1, let F_S^k and F_G^k denote the feature maps extracted from the k th IRM in the S²R and SGR subnets, respectively. Then, the level-by-level feature interaction is accomplished as follows:

$$F_S^{k'} = M_G^k \odot F_S^k, \quad (10)$$

$$F_G^{k'} = M_S^k \odot F_G^k, \quad (11)$$

where $F_S^{k'}$ and $F_G^{k'}$ denote the interacted results in the S²R and SGR subnets, respectively; $\odot$ represents element-wise multiplication; M_S^k and M_G^k are the attention maps separately generated from F_S^k and F_G^k via an SA unit. Specifically, $M_p = \sigma(W_h(\text{Avg}_h(F_p))) \odot$

$\sigma(W_w(\text{Avg}_w(F_p)))$, where $p \in \{S, G\}$; $\text{Avg}_h(\cdot)$ and $\text{Avg}_w(\cdot)$ denote the along-height and -width pooling operations, respectively; $W_h(\cdot)$ and $W_w(\cdot)$ are the corresponding 1D convolutions with kernel sizes of 5×1 and 1×5 , respectively; $\sigma(\cdot)$ denotes the Sigmoid function. During this process, the spectral features from one subnet are injected into the other one as a modulation factor, enabling bidirectional information exchange. Specifically, the attention maps M_S and M_G attempt to act as the modulation factors via element-wise multiplication to dynamically scale features so that informative components are emphasized while the redundant and noisy ones are suppressed. Subsequently, in both subnets, an inner shortcut formed by one 1×1 convolutional layer and the ReLU function is connected behind several residual modules to execute efficient fusion for the high- and low-level features.

A simple element-wise addition is used to initially fuse the features of the two subnets; then, an ERM presented in the light-green box of Figure 1 is carefully schemed to refine the fused features and project them into the image domain. Firstly, the coarse fused result is unified via a 1×1 convolutional layer. Next, 3 attention-based fusion blocks are linked with residual connection [53,54] to promote sufficient fusion. Each fusion block contains double-path feature extraction, a concatenation operation, an SA unit, and a ReLU function. The double-path feature extraction, which consists of two parallel convolutional layers of size 3×3 and 1×1 , respectively, is used to extract features in different receptive fields, as local information could be lost when adopting a single receptive field. Although some different multi-path feature extraction strategies, for example, the multi-branch network in [55], have been published, our ERM is specifically designed to bridge the feature gap between the heterogeneous domains of the image and the spectral graph for a refined feature fusion.

After aggregating the double-path features, an SA unit is used to weaken the unimportant parts and strengthen the important ones. Then, the ReLU function enhances the information representation of the attention result by introducing non-linearity. The fusion output is further refined by a 1×1 convolutional layer and an SA unit. Here, the SA unit is used as the dynamic spatial filter to suppress potential noise or artifacts introduced during the cross-branch fusion process. As further verified in Section 4.5, even if slight noise is mixed during feature fusion, our S²R method maintains stable reconstruction performance. Finally, the refined features are converted into the desired image domain through a dense block cascaded by two convolutional layers with kernel sizes of 3×3 and 1×1 , respectively, and a Tanh function. The stride, padding, and channel size of each convolutional layer are set to 1, 0, and C , respectively.

4. Experiments and Discussion

4.1. Datasets and Metrics

Some simulation experiments on the public Harvard [56], CAVE [57], and Chikusei [58] datasets, as well as some verification experiments on the real GaoFen-5 dataset, are performed to demonstrate the effectiveness, superiority, and practicality of the proposed GDIS²Net.

The Harvard dataset covering the range of 420–720 nm at 10 nm intervals, contains 50 HSIs of size $1392 \times 1040 \times 31$. Subimages of size $512 \times 512 \times 31$ are cut from the original HSIs to simulate the ground truth, i.e., H²SIs. These simulated H²SIs are degraded through the SRF of the Nikon D700 camera to generate HMSIs. Subsequently, LHSIs are produced by applying a Gaussian PSF of $r = 8$ to spatially blur the H²SIs. In total, 40 pairs of HMSI/LHSI are randomly selected for training, with the remaining being applied for testing.

To evaluate the robustness of S²R methods across different data distributions, the CAVE dataset is further utilized here. The dataset comprises 32 HSIs, which consist of 512×512 pixels and 31 spectral bands ranging from 400 nm to 700 nm. Similarly,

the original HSIs are used as H^2SIs , while the corresponding HMSIs and LHSIs are simulated via the Nikon D700 SRF and a Gaussian PSF of $r = 8$. Seven HMSI/LHSI pairs are randomly chosen as the testing set, while the rest serve as the training set.

To examine the influence of different degradation conditions on S^2R methods, the SRF of the IKONOS sensor and a Gaussian PSF of $r = 16$ are employed to perform linear degradation operations on the Chikusei dataset, synthesizing HMSIs and LHSIs. The Chikusei dataset is an HSI of size $2517 \times 2335 \times 128$, which was acquired over Chikusei, Japan, in the wavelength range from 363 nm to 1018 nm. The Chikusei image is divided into 34 simulated H^2SIs of size $512 \times 512 \times 128$. Correspondingly, 34 pairs of HMSI/LHSI are synthesized, which are randomly split into the training and testing dataset, including 25 and 9 pairs, respectively.

Experiments on the real GaoFen-5 dataset are also conducted to verify the applicability of the considered super-resolution methods. The GaoFen-5 dataset, which was captured by the hyperspectral sensor onboard the GaoFen-5 satellite with the wavelength range of [390, 2513] nm, contains an HMSI of size $3555 \times 4026 \times 4$ and an LHSI of size $1185 \times 1342 \times 285$. Both the HMSI and LHSI are cropped into 72 subimages of size $147 \times 147 \times 4$ and $441 \times 441 \times 285$, where 57 pairs of multi- and high-spectral subimages are randomly chosen for training, and the remaining are allocated for testing. All PSFs and SRFs, which are necessary when training S^2R networks, whether on the real or the simulated datasets, are estimated via our previous DiriNet [59].

Five common metrics, including the peak signal-to-noise ratio (PSNR), structural similarity index measure (SSIM), spectral angle mapper (SAM), spectral information divergence (SID), and mean relative absolute error (MRAE), are used to quantitatively evaluate the performance of all considered S^2R methods. PSNR and MRAE are estimates of the overall numerical error, SAM and SID are measures of spectral differences, and SSIM is an evaluation of spatial structure restoration. Generally, the higher the PSNR and SSIM, and the lower the SAM, SID, and MRAE, the better the super-resolution performance.

4.2. Parameter Settings of GDIS²Net

4.2.1. Regularization Parameters

The regularization parameters λ and μ in our loss function, Equation (4), are empirically tuned following the experimental results. Their optimal settings are experientially determined via a grid search, as shown in Table 1.

Table 1. Discussion for λ and μ (PSNR/SAM).

λ/μ	0.1	0.5	0.8	1.0
0.5	36.1865/10.3691	37.6313/6.5482	37.5263/7.2616	35.4165/9.8110
0.8	36.9631/8.6031	37.7579/6.3947	37.1598/7.7920	33.4017/15.6594
1.0	37.4418/7.1760	38.8669/5.0310	37.4843/5.6842	38.3544/6.1862
1.2	36.7791/6.9892	36.8550/6.0217	35.5214/11.2271	36.8120/10.1811

It can be observed that the best trade-offs between spatial and spectral reconstruction are achieved when λ and μ are set to 1.0 and 0.5, respectively. Accordingly, $\lambda = 1$ and $\mu = 0.5$ are adopted for all subsequent experiments.

4.2.2. Network Hyperparameters

The proposed GDIS²Net is optimized using an ADAM optimizer [60] with $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\varepsilon = 10^{-9}$. The number of the adjacent bands considered when constructing the spectral graph and the IRM as the backbone of the proposed net are critical to our super-resolution performance. Certain continuity exists in high spectrums, fewer adjacent bands could not reflect this property well while more might introduce unexpected redundancy.

As to the IRM, the higher its number k is, the better the non-linearity of the GDIS²Net generally. However, the computational cost would be higher with the increasing number of IRM. Hence, settings for the two hyperparameters are discussed on the Harvard dataset, as listed in Table 2. Here, n is manually tuned by fixing k at 9, and k is tuned by fixing n at 3. Experimentally, the numbers of the adjacent bands and IRM are set as 3 and 9, respectively.

Table 2. Discussion for parameters n and k .

Parameter	n					k		
	2	3	4	5	7	8	9	10
PSNR	37.7178	38.8669	37.7532	37.2626	37.9261	37.0221	38.8669	36.8870
SAM	6.4463	5.0310	5.6554	6.3619	6.1728	6.3580	5.0310	6.3815

4.3. Experiments on Simulated Datasets

Seven state-of-the-art S²R networks, including MSCNN [23], HSCNN+ [27], FM-Net [28], HRNet [29], MST++ [31], RepCPSI [36], and SSRAN [16], are compared with the proposed GDIS²Net to reveal its effectiveness. These compared nets are originally trained by H²SI. For a fair and comprehensive comparison, the three most advanced nets, i.e., HRNet, RepCPSI, and SSRAN, are trained via pairs of LMSI/LHSI without changing their architectures, where the LMSI is produced from the HMSI via PSF. The three super-resolution approaches that change labels are named as HR-MS, Rep-MS, and SSR-MS, respectively.

The quantitative results of all considered super-resolution methods on the Harvard, CAVE, and Chikusei datasets are listed in Tables 3–5. As Tables 3–5 show, these methods can be divided into super-resolution supervised by H²SI and LHSI. For a convenient analysis, the best and second-best results in both classes are highlighted in bold and underlined. One worth noting is that SID sometimes reveals opposite performance compared with other metrics, especially in Table 5. This might be because SID could be heavily penalized for minute absolute numerical fluctuations in low-intensity or low-reflectance bands, which result from the logarithmic calculation. Therefore, performance discussions are conducted from a comprehensive perspective rather than one specific metric. Among all methods that are supervised via ground truth, it can be seen from the results on the Harvard and CAVE datasets that HRNet almost achieves the best spectral restoration. On the Chikusei dataset, HRNet is inferior to RepCPSI and SSRAN. Therefore, HRNet might be of insufficient robustness to the resolution difference, because compared to Harvard and CAVE datasets, the resolution reduction degree of Chikusei is stronger. However, results on the three datasets all show that our GDIS²Net accomplishes comparable or even better performances to the three better super-resolution methods with H²SI supervision, namely, HRNet, RepCPSI, and SSRAN. For example, GDIS²Net attains the highest PSNR and SSIM values on the Harvard and Chikusei datasets, along with the best SAM results on all datasets, indicating its superior spectral reconstruction accuracy. That is, our super-resolution method achieves competitive and advanced spectral reconstruction when supervised by the accessible LHSI. Moreover, our GDIS²Net appears to have stable super-resolution performance, meaning that it is robust to data distribution and degradation conditions. If SSR-MS, HR-MS, and Rep-MS are separately compared to SSRAN, HRNet, and RepCPSI, it can be found that their reconstruction effects are slightly weaker. This phenomenon illustrates that the adopted linear degradation-based supervision achieves success. Compared with SSR-MS, HR-MS, and Rep-MS, the proposed GDIS²Net executes the best spectral reconstruction. Hence, the proposed graph-regularized double-path interactive net is superior to state-of-the-art networks geared towards spectral resolution enhancement.

Table 3. Quantitative comparisons of all the considered methods on the Harvard dataset. The best results are shown in bold, and the second-best results are underlined.

Label	Methods	PSNR	SSIM	SAM	SID	MRAE
H ² SI	MST++	35.1999	0.9520	7.4706	0.4126	0.1640
	MSCNN	37.4220	0.9642	7.1530	0.3031	0.1801
	FMNet	<u>38.2679</u>	0.9717	7.5996	0.2206	0.1432
	HSCNN+	38.0642	0.9760	<u>6.3277</u>	0.2005	<u>0.1157</u>
	RepCPSI	38.1547	0.9802	6.5840	<u>0.1807</u>	0.1224
	HRNet	38.2461	<u>0.9799</u>	6.1814	0.1777	0.1140
	SSRAN	38.2695	0.9773	6.8105	0.2068	0.1280
LHSI	SSR-MS	36.3486	0.9714	6.8502	0.2467	0.1293
	HR-MS	36.3341	0.9732	6.4564	0.2257	0.1241
	Rep-MS	<u>37.5639</u>	<u>0.9781</u>	<u>6.2561</u>	<u>0.1901</u>	<u>0.1211</u>
	Ours	38.8669	0.9858	5.0310	0.1272	0.0962

Table 4. Quantitative comparisons of all the considered methods on the CAVE dataset. The best results are shown in bold, and the second-best results are underlined.

Label	Methods	PSNR	SSIM	SAM	SID	MRAE
H ² SI	MST++	35.2603	0.9417	<u>14.3400</u>	0.4561	0.3188
	MSCNN	34.8797	0.9261	<u>17.4841</u>	0.5493	0.3468
	FMNet	<u>35.2888</u>	0.9562	15.6942	0.3478	0.2073
	HSCNN+	34.9749	0.9663	19.5772	0.3785	<u>0.1709</u>
	RepCPSI	34.5375	0.9430	16.6376	0.4611	0.2721
	HRNet	35.5354	0.9783	12.0905	0.2660	0.1409
	SSRAN	35.3151	<u>0.9674</u>	14.4301	<u>0.3258</u>	0.1961
LHSI	SSR-MS	<u>34.1259</u>	0.9576	15.3093	0.3914	0.2099
	HR-MS	33.1590	0.9433	17.2188	0.5737	0.2513
	Rep-MS	34.5491	<u>0.9627</u>	<u>14.6407</u>	<u>0.3715</u>	<u>0.1933</u>
	Ours	34.0388	0.9710	11.1062	0.3291	0.1498

Table 5. Quantitative comparisons of all the considered methods on the Chikusei dataset. The best results are shown in bold, and the second-best results are underlined.

Label	Methods	PSNR	SSIM	SAM	SID	MRAE
H ² SI	MST++	36.2656	0.9611	7.3473	0.4576	<u>0.2069</u>
	MSCNN	32.8050	0.9069	6.4380	1.0099	0.2348
	FMNet	37.1867	0.7419	<u>32.6707</u>	1.6767	0.8571
	HSCNN+	32.2415	0.8976	7.2431	0.8723	0.2438
	RepCPSI	<u>38.7702</u>	0.9518	7.3694	<u>0.5280</u>	0.2210
	HRNet	37.5858	0.9343	8.3180	0.5670	0.2455
	SSRAN	39.6178	<u>0.9595</u>	<u>6.7386</u>	0.5718	0.1906
LHSI	SSR-MS	37.4364	<u>0.9509</u>	<u>6.6482</u>	0.5023	0.1819
	HR-MS	38.2603	0.9217	9.5394	0.7827	0.2699
	Rep-MS	<u>37.6962</u>	0.9336	8.1593	<u>0.6481</u>	0.2326
	Ours	39.7644	0.9684	5.8301	0.9079	0.1752

Some visual presentations produced by all of the considered methods on the Harvard, CAVE, and Chikusei datasets are illustrated in Figures 2–4, containing single bands, false-color images, and error heat maps. Error heat maps display the average error distribution between the reconstructed and referenced H²SI in space, where a darker blue color represents lower reconstruction errors. Full-sized error maps in Figure 4 sharing the same color range seem to illustrate that all reconstructions are similar to the reference, which is actually caused by the much worse result of FMNet. Therefore, accurate error analyses for

Figure 4 should refer to the enlarged maps. As can be seen from Figures 2–4, the proposed GDIS²Net produces sharper super-resolution images and darker error maps overall, which indicates its better spatial preservation ability.

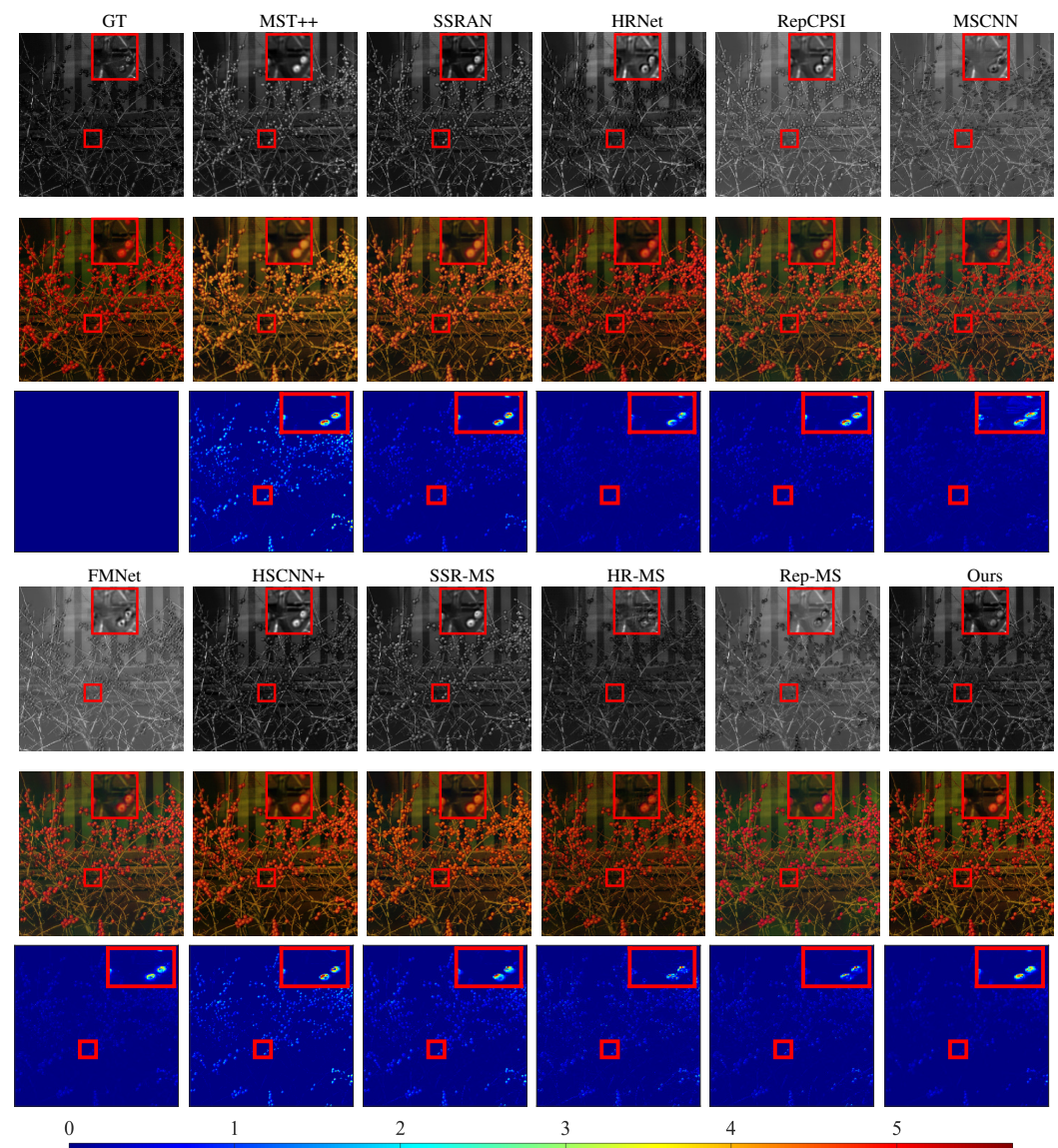


Figure 2. Spectral enhancement results on the Harvard dataset. The 13th band, false-color images (R:29, G:14, B:3), and error heat maps of different methods are placed from top to bottom in sequence. The red squares indicate the zoomed-in regions.

In addition, the average spectral curves over the randomly selected 50×50 cubes (as shown in the red square of Figures 2–4) from the above reconstruction results are shown in Figure 5. Spectral curves could exhibit the average difference between the reconstructed and reference H²SI cubes in space, where a method corresponding to the curve that fits the red curve better is of better spectral fidelity. Compared to others, the proposed S²R method generates results closely aligned to reference curves across the entire spectrum. The curves produced by our GDIS²Net exhibit smaller fluctuations, especially in regions with rapid spectral variations, which means effective preservation for both spectral shape and amplitude. In contrast, spectral curves with stronger fluctuations are created by competing methods, particularly on the Chikusei dataset with a higher spatial–spectral resolution difference. Thus, the proposed GDIS²Net could offer better spectral fidelity and stronger robustness under worse resolution degradation conditions.

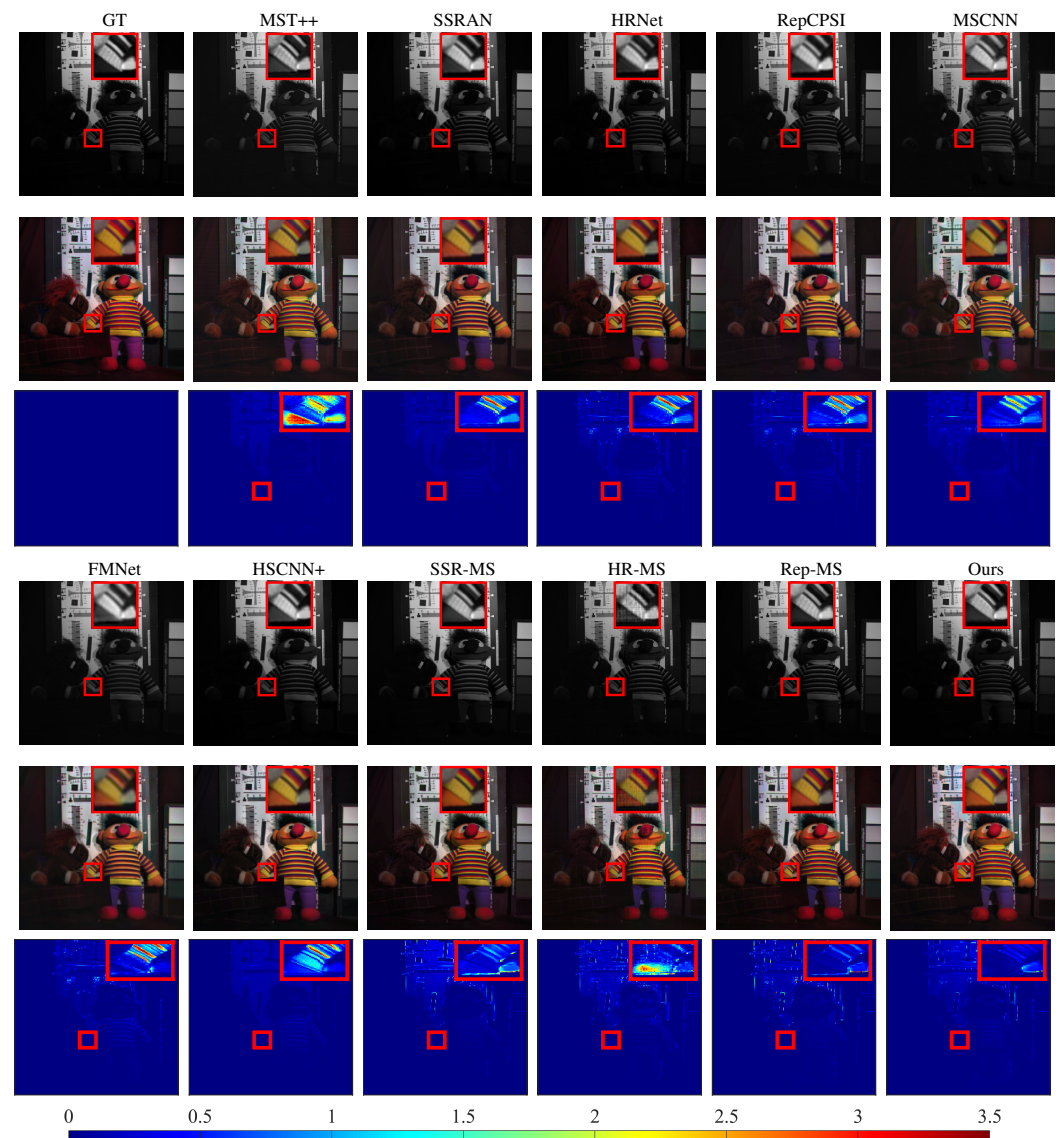


Figure 3. Spectral enhancement results on the CAVE dataset. The 13th band, false-color images (R:31, G:16, B:5), and error heat maps of different methods are placed from top to bottom in sequence. The red squares indicate the zoomed-in regions.

Comprehensively, the proposed GDIS²Net produces an effective and superior spectral enhancement effect, as well as showing strong robustness and generalization to different datasets and degradation conditions.

4.4. Experiments on Real Dataset

To further assess the practical applicability of GDIS²Net, experiments are conducted on the real-world GaoFen-5. As this dataset lacks ground truth (H²SI), existing super-resolution methods that rely on the H²SI label cannot be directly applied. Therefore, in addition to GDIS²Net, three hybrid supervised methods (SSR-MS, HR-MS, and Rep-MS) are evaluated for a fair comparison.

For a quantitative evaluation, the spectral distortion index (D_λ), spatial distortion index (D_s), and the quality with no reference (QNR) are calculated under the Wald protocol as some existing studies [61]. Specifically, the reconstructed H²SI is degraded using the pre-estimated PSF to generate simulated HSI. The degraded results are then compared with the observed HSI to compute D_λ and D_s , and the overall QNR is obtained accordingly.

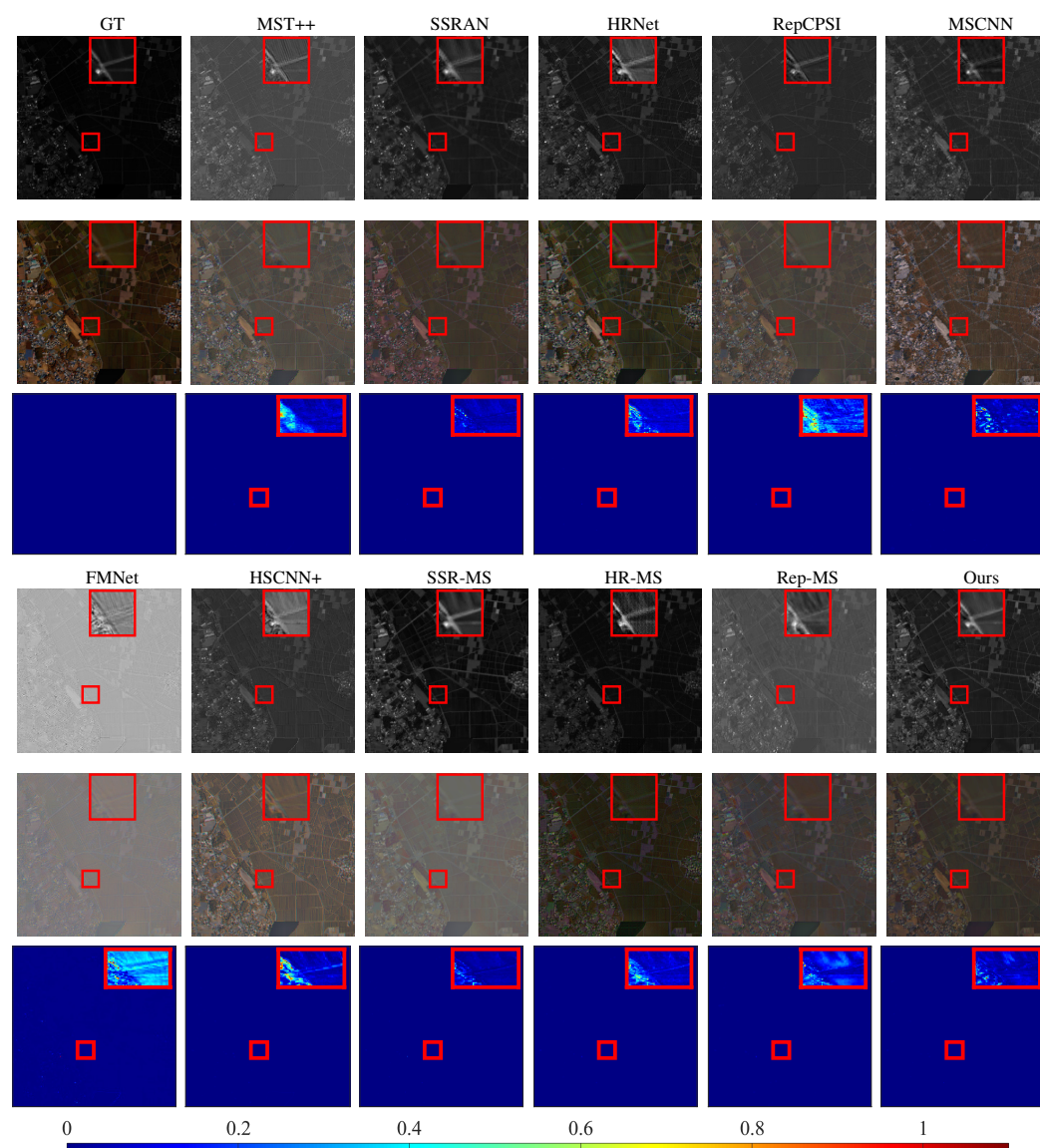


Figure 4. Spectral enhancement results on the Chikusei dataset. The 13th band, false-color images (R:67, G:37, B:15), and error heat maps of different methods are placed from top to bottom in sequence. The red squares indicate the zoomed-in regions.

The quantitative results are reported in Table 6. Obviously, our GDIS²Net achieves superior performance across all three metrics. In addition, the single-band and false-color results in Figure 6 illustrate that GDIS²Net achieves better visual texture restoration. Then, the superior net design and better practical potential of our method are validated.

Table 6. Quantitative comparisons on GaoFen-5. The best results are shown in bold, and the second-best results are underlined.

Method	$D_\lambda \downarrow$	$D_s \downarrow$	QNR $\uparrow$
HR-MS	0.2603	0.2277	0.5714
Rep-MS	0.2561	0.2174	0.5824
SSR-MS	0.2434	0.2096	0.5981
Ours	0.2281	0.20469	0.6101

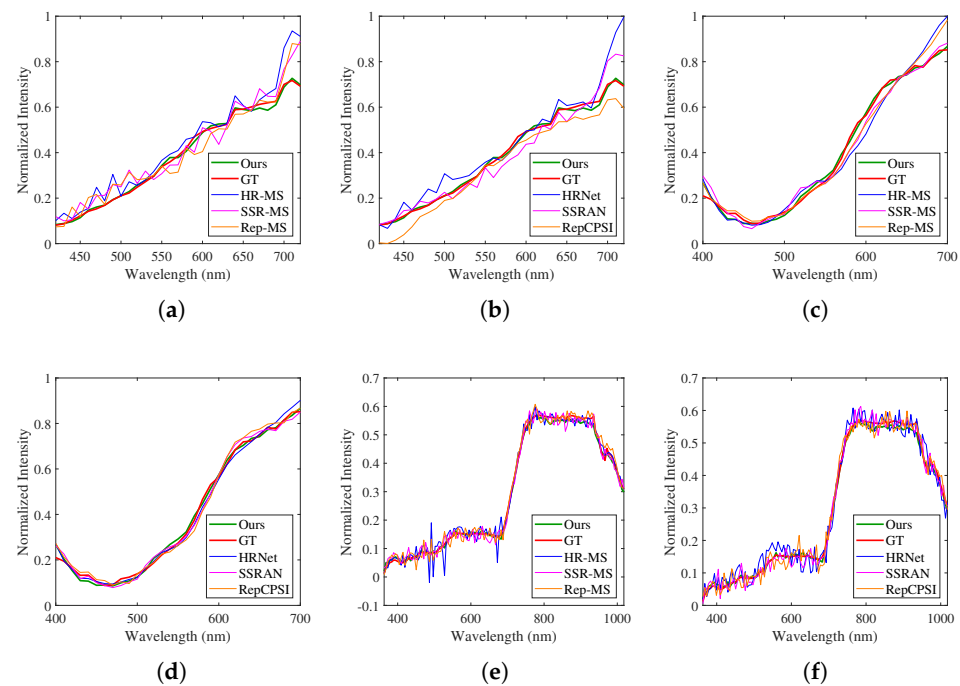


Figure 5. The average spectral curves over the randomly selected 50×50 cubes. (a,b) Harvard. (c,d) CAVE. (e) and (f) Chikusei.



Figure 6. Original images and spectral enhancement results on GaoFen-5. The second band and false-color images (R:50, G:35, B:25 for LHSI and S^2R results; R:3, G:2, B:1 for HMSI) are placed from top to bottom in sequence. The red squares indicate the zoomed-in regions.

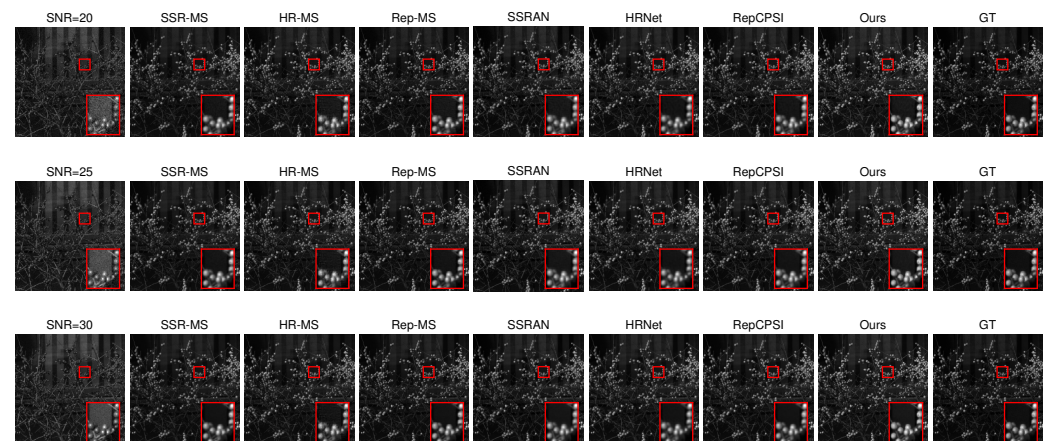
4.5. Noise Robustness Analysis

To evaluate the noise robustness of the proposed GDIS²Net, several comparative S^2R networks are tested on the noisy Harvard dataset by adding Gaussian noise of different signal-to-noise ratios (SNR = 20, 25, and 30 dB) to the inputting MSI.

The quantitative results in Table 7 show that the reconstruction performance of all considered methods is naturally degraded with the increasing in noise intensity. However, our GDIS²Net comprehensively shows stable and better metrics, which could be the result of the task-irrelevant disturbance filtering of the incorporated spectral graph regularization and dual-branch interaction mechanism. Visual comparisons in Figure 7 further confirm that GDIS²Net effectively suppresses noise while preserving structural details.

Table 7. Quantitative noise robustness comparisons of different methods on the Harvard dataset. The best results are shown in bold, and the second-best results are underlined.

Method/Input	SNR = 20	SNR = 25	SNR = 30	Clean
SSRAN	37.5296/6.9622	38.0124/6.8786	38.1857/6.8403	38.2695/6.8105
HRNet	37.7757/6.3401	38.0928/6.2435	38.1937/6.2100	<u>38.2461/6.1814</u>
RepCPSI	<u>37.5995/7.1618</u>	<u>37.9656/6.8096</u>	<u>38.0960/6.6666</u>	38.1547/6.5840
SSR-MS	35.6851/6.9763	35.9821/6.7738	36.0922/6.7029	36.3486/6.8502
HR-MS	35.9552/7.0509	36.1980/6.6627	36.2832/6.5182	36.3341/6.4564
Rep-MS	<u>36.2196/7.6345</u>	<u>37.0318/6.8146</u>	<u>37.3926/6.4623</u>	<u>37.5639/6.2561</u>
Ours	37.3804/6.1379	38.3008/5.5506	38.6633/5.2461	38.8669/5.0310

**Figure 7.** Noise robustness experiments on the Harvard dataset. The 2nd of MSI and 25th bands of H²SI are placed at 1st and 2–9 columns, respectively. The red squares indicate the zoomed-in regions.

4.6. Ablation Studies

A series of ablation studies was conducted on the Harvard dataset for the proposed GDIS²Net to thoroughly evaluate its contributions. Sufficient analysis has been accomplished for the effectiveness of training super-resolution networks with LHSI in the above experiments. Thus, no further discussion about this will be conducted in this section. Ablation experiments were designed by progressively adding components of spectral graph restoration, spectral continuity loss, IRM, and ERM to a baseline model.

Specifically, four network variants are defined. (1) SGLC-Net: GDIS²Net without ERM. (2) SGL-Net: SGLC-Net without the feature interaction of IRM, i.e., the two SA units and cross-feature multiply operations are deleted. (3) SG-Net: SGL-Net without the spectral continuity loss. (4) Base: SG-Net without the SGR subnet but cascading SSRAN. What needs to be explained is that Base would provide insufficient reconstruction without SSRAN, and SSRAN is connected for feature-to-image reconstruction because of its simple architecture.

All quantitative results are summarized in Table 8. Obviously, SG-Net outperforms Base, and SGL-Net achieves further improvements over SG-Net. Thus, both the spectral continuity constraint and the SGR subnet work well. Namely, the proposed spectral graph restoration procedure is quite effective. SGLC-Net surpasses SGL-Net, particularly towards SAM, indicating that attention-based feature interactions are beneficial for precise spectral recovery. The complete GDIS²Net delivers the best performance, which confirms the further refinement of the ERM. In conclusion, the proposed spectral graph restoration and two spatial attention-based residual modules contribute to the superior super-resolution effect of the whole GDIS²Net.

Table 8. Ablation results.

Description	LSS	SGR	SG loss	IRM	ERM	PSNR	SSIM	SAM	SID
GDIS ² Net	✓	✓	✓	✓	✓	38.8669	0.9858	5.0310	0.1272
SGLC-Net	✓	✓	✓	✓	×	38.3362	0.9854	5.4019	0.1445
SGL-Net	✓	✓	✓	×	×	38.0772	0.9804	5.7055	0.1829
SG-Net	✓	✓	×	×	×	37.7825	0.9797	6.7427	0.1929
Base	×	×	×	×	×	35.4598	0.9668	7.3318	0.2827

4.7. Computational Efficiency Analysis

To comprehensively evaluate the computational efficiency of the proposed GDIS²Net, its number of parameters (Params), floating-point operations (FLOPs), model size, and inference time are compared with other state-of-the-art methods on the Chikusei dataset. The inference time is averaged for a 128×128 input over 100 forward passes on a NVIDIA RTX 3090 GPU.

As reported in Table 9, our model size and parameters are well-controlled and significantly smaller than those of MSCNN and HRNet. Due to the adoption of the dual-branch structure and multi-level interactive residual modules, our method inevitably introduces higher FLOPs and a slightly longer inference time. However, our processing speed remains highly competitive and satisfies most actual requirements.

Table 9. Computational efficiency of different methods on the Chikusei dataset. The best results are shown in bold, and the second-best results are underlined.

Method	Params (M)	FLOPs (G)	Model Size (MB)	Inference Time (ms)
MST++	1.761	6.808	6.721	7.806
MSCNN	26.693	<u>10.472</u>	101.873	2.078
FMNet	2.987	48.920	11.396	6.482
HSCNN+	1.188	19.466	4.532	<u>3.557</u>
RepCPSI	2.148	35.123	8.196	6.291
HRNet	8.382	10.946	31.976	4.397
SSRAN	<u>1.462</u>	23.910	<u>5.575</u>	3.633
Ours	5.638	92.655	26.056	14.085

5. Conclusions

In this paper, a novel residual S²R network aided by spectral graph reconstruction, termed GDIS²Net, was proposed for MSI spectral enhancement. Combining the supervisions of the hyperspectral image and its spectral graph with the self-supervision of the multispectral input, the proposed GDIS²Net co-restored the spectrum and spectral graph for MSI to the corresponding high-spectral-resolution level. Such hybrid supervision was achieved by tailored spatial–spectral fidelity loss, where both spatial–spectral information protection and the spectral continuity guarantee were considered. Furthermore, an interactive residual module and an enhanced residual module are designed to facilitate cross-feature interaction and multi-scale feature fusion, respectively. Experiments on simulated datasets demonstrated that our GDIS²Net achieved competitive performance compared to state-of-the-art methods. Moreover, our super-resolution net was superior to state-of-the-art S²R networks, as shown in the experimental results on both simulated and real datasets. In further studies, we would explore a data-driven Laplacian matrix learning method to improve spectral reconstruction.

Author Contributions: Conceptualization, S.W. and T.H.; methodology, S.W. and T.H.; validation, S.W.; investigation, S.W., S.C., Z.L. and Z.S.; writing, S.W. and T.H.; supervision, T.H., Z.L., Z.S., K.J. and J.F. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the National Natural Foundation of China (No. 62301012).

Data Availability Statement: The Harvard dataset mentioned in this paper is openly and freely available at <http://vision.seas.harvard.edu/hyperspec/> (accessed on 21 January 2026). The CAVE dataset used in this study is openly available at <https://cave.cs.columbia.edu/repository/Multispectral> (accessed on 21 January 2026). The Chikusei dataset used in this paper is freely available at <http://park.its.u-tokyo.ac.jp/sal/hyperdata> (accessed on 21 January 2026). The GaoFen-5 dataset mentioned in this paper was obtained from a third-party data provider and is available at <http://www.sasclouds.com/chinese/home/> (accessed on 21 January 2026).

Conflicts of Interest: Author Siyuan Cheng was employed by the company Space Star Technology Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Akhtar, N.; Mian, A. Nonparametric Coupled Bayesian Dictionary and Classifier Learning for Hyperspectral Classification. *IEEE Trans. Neural Networks Learn. Syst.* **2018**, *29*, 4038–4050. [\[CrossRef\]](#)
2. Sun, X.; Qu, Y.; Gao, L.; Sun, X.; Qi, H.; Zhang, B.; Shen, T. Ensemble-Based Information Retrieval With Mass Estimation for Hyperspectral Target Detection. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5508123. [\[CrossRef\]](#)
3. Lu, R.; Chen, B.; Cheng, Z.; Wang, P. RAFnet: Recurrent attention fusion network of hyperspectral and multispectral images. *Signal Process.* **2020**, *177*, 107737. [\[CrossRef\]](#)
4. Li, Y.; Wang, J.; Liu, X.; Xian, N.; Xie, C. DIM Moving Target Detection using Spatio-Temporal Anomaly Detection for Hyperspectral Image Sequences. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, IEEE: New York, NY, USA, 2018; pp. 7086–7089. [\[CrossRef\]](#)
5. Xing, C.; Wang, M.; Dong, C.; Duan, C.; Wang, Z. Joint sparse-collaborative representation to fuse hyperspectral and multispectral images. *Signal Process.* **2020**, *173*, 107585. [\[CrossRef\]](#)
6. Zhao, J.; Wu, K.; Zhang, L.; Huang, W.; Ruan, C.; Huang, L. Patch-based hierarchical residual spectral-spatial convolutional network for hyperspectral image classification. *Signal Process.* **2025**, *230*, 109850. [\[CrossRef\]](#)
7. Munipalle, V.K.; Nelakuditi, U.R.; Nidamanuri, R.R. Agricultural Crop Hyperspectral Image Classification using Transfer Learning. In *Proceedings of the International Conference on Machine Intelligence for GeoAnalytics and Remote Sensing (MIGARS)*; IEEE: Piscataway, NJ, USA, 2023; Volume 1, pp. 1–4. [\[CrossRef\]](#)
8. Zhu, K.; Sun, Z.; Zhao, F.; Yang, T.; Tian, Z.; Lai, J.; Long, B.; Li, S. Remotely sensed canopy resistance model for analyzing the stomatal behavior of environmentally-stressed winter wheat. *ISPRS J. Photogramm. Remote Sens.* **2020**, *168*, 197–207. [\[CrossRef\]](#)
9. Wang, D.; Hu, M.; Jin, Y.; Miao, Y.; Yang, J.; Xu, Y.; Qin, X.; Ma, J.; Sun, L.; Li, C.; et al. HyperSIGMA: Hyperspectral Intelligence Comprehension Foundation Model. *IEEE Trans. Pattern Anal. Mach. Intell.* **2025**, *47*, 6427–6444. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Wang, P.; Zhang, G.; Wang, L.; Leung, H.; Bi, H. Subpixel Land Cover Mapping Based on Dual Processing Paths for Hyperspectral Image. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 1835–1848. [\[CrossRef\]](#)
11. Liu, R.; Liang, J.; Yang, J.; Hu, M.; He, J.; Zhu, P.; Zhang, L. DHSNet: Dual Classification Head Self-Training Network for Cross-Scene Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2025**, *63*, 5534515. [\[CrossRef\]](#)
12. Yang, J.; Wu, C.; Du, B.; Zhang, L. Enhanced Multiscale Feature Fusion Network for HSI Classification. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 10328–10347. [\[CrossRef\]](#)
13. Jia, J.; Chen, J.; Zheng, X.; Wang, Y.; Guo, S.; Sun, H.; Jiang, C.; Karjalainen, M.; Karila, K.; Duan, Z.; et al. Tradeoffs in the Spatial and Spectral Resolution of Airborne Hyperspectral Imaging Systems: A Crop Identification Case Study. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5510918. [\[CrossRef\]](#)
14. Wang, L.; Xiong, Z.; Gao, D.; Shi, G.; Zeng, W.; Wu, F. High-speed hyperspectral video acquisition with a dual-camera architecture. In *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*; IEEE: New York, NY, USA, 2015; pp. 4942–4950. [\[CrossRef\]](#)
15. Yan, J.; Zhang, K.; Zhang, F.; Ge, C.; Wan, W.; Sun, J. Multispectral and hyperspectral image fusion based on low-rank unfolding network. *Signal Process.* **2023**, *213*, 109223. [\[CrossRef\]](#)
16. Zheng, X.; Chen, W.; Lu, X. Spectral Super-Resolution of Multispectral Images Using Spatial-Spectral Residual Attention Network. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5404114. [\[CrossRef\]](#)
17. Arad, B.; Shahar, O. Sparse Recovery of Hyperspectral Signal from Natural RGB Images. In *Proceedings of the European Conference on Computer Vision (ECCV)*, Amsterdam, The Netherlands, 11–14 October 2016; Springer: Cham, Switzerland, 2016; Volume 9911, pp. 19–34. [\[CrossRef\]](#)

18. Yokoya, N.; Heiden, U.; Bachmann, M. Spectral Enhancement of Multispectral Imagery Using Partially Overlapped Hyperspectral Data and Sparse Signal Representation. In *Proceedings of the Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Amsterdam, The Netherlands, 23–26 September 2018*; IEEE: Piscataway, NJ, USA, 2018; pp. 1–6.
19. Yi, C.; Zhao, Y.Q.; Chan, J.C.W. Spectral Super-Resolution for Multispectral Image Based on Spectral Improvement Strategy and Spatial Preservation Strategy. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 9010–9024. [\[CrossRef\]](#)
20. Gao, L.; Hong, D.; Yao, J.; Zhang, B.; Gamba, P.; Chanussot, J. Spectral Superresolution of Multispectral Imagery With Joint Sparse and Low-Rank Learning. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 2269–2280. [\[CrossRef\]](#)
21. Galliani, S.; Lanaras, C.; Marmanis, D.; Baltsavias, E.; Schindler, K. Learned Spectral Super-Resolution. *arXiv* **2017**, arXiv:1703.09470. [\[CrossRef\]](#)
22. Xiong, Z.; Shi, Z.; Li, H.; Wang, L.; Liu, D.; Wu, F. HSCNN: CNN-Based Hyperspectral Image Recovery from Spectrally Undersampled Projections. In *Proceedings of the 2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*; IEEE: New York, NY, USA, 2017; pp. 518–525. [\[CrossRef\]](#)
23. Yan, Y.; Zhang, L.; Li, J.; Wei, W.; Zhang, Y. Accurate Spectral Super-resolution from Single RGB Image Using Multi-scale CNN. *arXiv* **2018**, arXiv:1806.03575.
24. Han, X.; Yu, J.; Luo, J.; Sun, W. Reconstruction From Multispectral to Hyperspectral Image Using Spectral Library-Based Dictionary Learning. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 1325–1335. [\[CrossRef\]](#)
25. Fotiadou, K.; Tsagkatakis, G.; Tsakalides, P. Spectral Super Resolution of Hyperspectral Images via Coupled Dictionary Learning. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 2777–2797. [\[CrossRef\]](#)
26. Barman, T.; Deka, B.; Prasad, A. GPU-Accelerated Adaptive Dictionary Learning and Sparse Representations for Multispectral Image Super-resolution. In *Proceedings of the 2021 IEEE 18th India Council International Conference (INDICON)*; IEEE: New York, NY, USA, 2021; pp. 1–7. [\[CrossRef\]](#)
27. Shi, Z.; Chen, C.; Xiong, Z.; Liu, D.; Wu, F. HSCNN+: Advanced CNN-Based Hyperspectral Recovery from RGB Images. In *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*; IEEE: New York, NY, USA, 2018; pp. 1052–1060. [\[CrossRef\]](#)
28. Zhang, L.; Lang, Z.; Wang, P.; Wei, W.; Liao, S.; Shao, L.; Zhang, Y. Pixel-aware Deep Function-mixture Network for Spectral Super-Resolution. *arXiv* **2019**, arXiv:1903.10501. [\[CrossRef\]](#)
29. Zhao, Y.; Po, L.M.; Yan, Q.; Liu, W.; Lin, T. Hierarchical Regression Network for Spectral Reconstruction from RGB Images. In *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*; IEEE: New York, NY, USA, 2020; pp. 1695–1704. [\[CrossRef\]](#)
30. Nathan, D.S.; Uma, K.; Vinothini, D.S.; Bama, B.S.; Roomi, S.M.M.M. Light Weight Residual Dense Attention Net for Spectral Reconstruction from RGB Images. *arXiv* **2020**, arXiv:2004.06930. [\[CrossRef\]](#)
31. Cai, Y.; Lin, J.; Lin, Z.; Wang, H.; Zhang, Y.; Pfister, H.; Timofte, R.; Gool, L.V. MST++: Multi-stage Spectral-wise Transformer for Efficient Spectral Reconstruction. In *Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*; IEEE: New York, NY, USA, 2022; pp. 744–754. [\[CrossRef\]](#)
32. Chen, W.; Zheng, X.; Lu, X. Hyperspectral Image Super-Resolution with Self-Supervised Spectral-Spatial Residual Network. *Remote Sens.* **2021**, *13*, 1260. [\[CrossRef\]](#)
33. Ma, X.; Liu, W.; Li, S.; Tao, D.; Zhou, Y. Hypergraph p -Laplacian Regularization for Remotely Sensed Image Recognition. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 1585–1595. [\[CrossRef\]](#)
34. Arad, B.; Ben-Shahar, O.; Timofte, R.; Van Gool, L.; Zhang, L.; Yang, M.H. NTIRE 2018 Challenge on Spectral Reconstruction from RGB Images. In *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*; IEEE: New York, NY, USA, 2018; pp. 1042–1051. [\[CrossRef\]](#)
35. Arad, B.; Timofte, R.; Ben-Shahar, O.; Lin, Y.T.; Finlayson, G.; Givati, S. NTIRE 2020 Challenge on Spectral Reconstruction from an RGB Image. In *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*; IEEE: New York, NY, USA, 2020; pp. 1806–1822. [\[CrossRef\]](#)
36. Wu, C.; Li, J.; Song, R.; Li, Y.; Du, Q. RepCPSI: Coordinate-Preserving Proximity Spectral Interaction Network With Reparameterization for Lightweight Spectral Super-Resolution. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5508313. [\[CrossRef\]](#)
37. Chun, M.; Golomb, J.; Turk-Browne, N. A Taxonomy of External and Internal Attention. *Annu. Rev. Psychol.* **2009**, *62*, 73–101. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Hu, J.; Shen, L.; Sun, G. Squeeze-and-Excitation Networks. In *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*; IEEE: New York, NY, USA, 2018; pp. 7132–7141. [\[CrossRef\]](#)
39. Woo, S.; Park, J.; Lee, J.Y.; Kweon, I.S. CBAM: Convolutional Block Attention Module. In *Proceedings of the Computer Vision–ECCV 2018*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 3–19.
40. Cao, Y.; Xu, J.; Lin, S.; Wei, F.; Hu, H. GCNet: Non-Local Networks Meet Squeeze-Excitation Networks and Beyond. In *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*; IEEE: New York, NY, USA, 2019; pp. 1971–1980. [\[CrossRef\]](#)

41. Li, J.; Wu, C.; Song, R.; Xie, W.; Ge, C.; Li, B.; Li, Y. Hybrid 2-D–3-D Deep Residual Attentional Network With Structure Tensor Constraints for Spectral Super-Resolution of RGB Images. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 2321–2335. [[CrossRef](#)]
42. He, J.; Li, J.; Yuan, Q.; Shen, H.; Zhang, L. Spectral Response Function-Guided Deep Optimization-Driven Network for Spectral Super-Resolution. *IEEE Trans. Neural Networks Learn. Syst.* **2022**, *33*, 4213–4227. [[CrossRef](#)]
43. Liu, S.; Liu, S.; Zhang, S.; Li, B.; Hu, W.; Zhang, Y.D. SSAU-Net: A Spectral–Spatial Attention-Based U-Net for Hyperspectral Image Fusion. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5542116. [[CrossRef](#)]
44. Zhao, H.; Gallo, O.; Frosio, I.; Kautz, J. Loss Functions for Image Restoration With Neural Networks. *IEEE Trans. Comput. Imaging* **2017**, *3*, 47–57. [[CrossRef](#)]
45. Dong, W.; Qu, J.; Zhang, T.; Li, Y.; Du, Q. Context-Aware Guided Attention Based Cross-Feedback Dense Network for Hyperspectral Image Super-Resolution. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5530814. [[CrossRef](#)]
46. Zhu, Q.; Zhang, M.; Chen, Y.; Zheng, G.; Luo, J. Spectral Correlation-Based Fusion Network for Hyperspectral Image Super-Resolution. *IEEE Trans. Geosci. Remote Sens.* **2025**, *63*, 5500314. [[CrossRef](#)]
47. Liu, N.; Li, W.; Tao, R.; Du, Q.; Chanussot, J. Multigraph-Based Low-Rank Tensor Approximation for Hyperspectral Image Restoration. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5530314. [[CrossRef](#)]
48. Su, X.; Zhang, Z.; Yang, F. Fast Hyperspectral Image Denoising and Destriping Method Based on Graph Laplacian Regularization. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5511214. [[CrossRef](#)]
49. Zhu, Z.Y.; Huang, T.Z.; Huang, J.; Wu, L. Tensor singular value decomposition and low-rank representation for hyperspectral image unmixing. *Signal Process.* **2025**, *230*, 109799. [[CrossRef](#)]
50. Wang, H.; Xu, Y.; Wu, Z.; Wei, Z. Unsupervised Hyperspectral and Multispectral Image Blind Fusion Based on Deep Tucker Decomposition Network With Spatial–Spectral Manifold Learning. *IEEE Trans. Neural Networks Learn. Syst.* **2025**, *36*, 12721–12735. [[CrossRef](#)] [[PubMed](#)]
51. Qu, Y.; Qi, H.; Kwan, C. Unsupervised Sparse Dirichlet-Net for Hyperspectral Image Super-Resolution. In *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*; IEEE: New York, NY, USA, 2018; pp. 2511–2520. [[CrossRef](#)]
52. Wang, L.; Zhou, J.; Li, Z.; Zhao, X.; Wu, C.; Xu, M. Adversarial MixUp with implicit semantic preservation for semi-supervised hyperspectral image classification. *Signal Process.* **2023**, *211*, 109116. [[CrossRef](#)]
53. Li, Y.; Fu, M.; Zhang, H.; Xu, H.; Zhang, Q. Hyperspectral Image Fusion Algorithm Based on Improved Deep Residual Network. *Signal Process.* **2023**, *210*, 109058. [[CrossRef](#)]
54. Chhapariya, K.; Ientilucci, E.J.; Buddhiraju, K.M.; Kumar, A. A Spectral-Spatial Classification Network for Hyperspectral Images using a Residual Attention Network. In *Proceedings of the 2023 13th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, Athens, Greece, 31 October–2 November 2023; IEEE: New York, NY, USA, 2023; pp. 1–5.
55. Yang, J.; Du, B.; Xu, Y.; Zhang, L. Can Spectral Information Work While Extracting Spatial Distribution?—An Online Spectral Information Compensation Network for HSI Classification. *IEEE Trans. Image Process.* **2023**, *32*, 2360–2373. [[CrossRef](#)]
56. Chakrabarti, A.; Zickler, T. Statistics of Real-World Hyperspectral Images. In *Proceedings of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*; IEEE: New York, NY, USA, 2011; pp. 193–200.
57. Yasuma, F.; Mitsunaga, T.; Iso, D.; Nayar, S.K. Generalized Assorted Pixel Camera: Postcapture Control of Resolution, Dynamic Range, and Spectrum. *IEEE Trans. Image Process.* **2010**, *19*, 2241–2253. [[CrossRef](#)]
58. Yokoya, N.; Iwasaki, A. *Airborne Hyperspectral Data over Chikusei*. Technical Report; The University of Tokyo: Tokyo, Japan, 2016.
59. Hu, T.; Cheng, S.; Liu, C. DiriNet: An Estimation Network for Spectral Response Function and Point Spread Function. *J. Beijing Inst. Technol.* **2024**, *33*, 287–297.
60. Kingma, D.; Ba, J. Adam: A Method for Stochastic Optimization. In *Proceedings of the International Conference on Learning Representations*, Banff, AB, Canada, 14–16 April 2014; pp. 1–15.
61. He, W.; Cai, Y.; Ren, Q.; Ruze, A.; Jia, S. Adaptive Expert Learning for Hyperspectral and Multispectral Image Fusion. *IEEE Trans. Geosci. Remote Sens.* **2025**, *63*, 5528615. [[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.