Hyperspectral image super-resolution reconstruction based on image partition and detail enhancement

Yinghao Xu
Qingdao University of Science and Technology
Yuchao Lv
Qingdao University of Science and Technology
Xijun Zhu
Qingdao University of Science and Technology
Sifan Liu
Qingdao University of Science and Technology
Yuanyuan Sun (yysun@qust.edu.cn)
Qingdao University of Science and Technology
Yimin Wang
Qingdao University of Science and Technology

Research Article

Keywords: Super resolution, Partition reconstruction, Detail enhancement, Hyperspectral image, Convolutional neural network

Posted Date: June 9th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1696328/v1

Abstract

The hyperspectral image (HSI) super-resolution (SR) reconstruction has attracted much attention and been used widely in various study fields due to its low requirements on hardware in practice. However, the distribution of image information is uneven. And HSI is treated equally during the process of super-resolution reconstruction. It is time consuming and many details cannot be extracted specifically. In this paper, a new method named MSDESR (multilevel streams and detail enhancement) is proposed to reconstruct HSI according to its uneven distribution of spatial information. The MSDESR consists of a submap shunt block, a high-low frequency information extraction with detail enhancement block, and a partition image reconstruction block. Firstly, the submap shunk block is designed to pre-classify hyperspectral images. The images are divided into complex and simple parts according to the spatial information distribution of the reconstructed submap. Secondly, the multiscale Retinex with detail enhancement algorithm is constructed to purify high-frequency noise-contaminated and enhance the image details by separating the samples into high and low frequency information. Finally, branching networks of different complexities are designed to reconstruct the images with high credibility and clear content. In this paper, datasets of QUST-1, Pavia University and Chikusei are applied in the experiments. The results show that, the MSDER outperforms state-of-the-art CNN-based methods in terms of quantitative metrics and visual quality, with quantities of 4.18% and 9.35% in the SRE and MPSNR metrics, respectively. Overall, the MSDER performs well in hyperspectral image super-resolution reconstruction, which is time saving and preserves the details of spatial information.

1 Introduction

Hyperspectral images include dozens or even hundreds of tiny spectral channels, making them very spectrally discriminating due to their abundance of spectral information. As a result, HSI has been used for natural disaster warning [1], land cover identification [2], and change detection [3]. However, the spatial resolution of HSI cannot meet the demand for the identification and interpretation of fine features, limiting its application in a variety of fields. Therefore, it is of great importance to improve the spatial resolution of HSI. Obtaining hyperspectral images with great spatial resolution can be difficult. To address this issue, hyperspectral images super-resolution reconstruction has been developed.

Super-resolution reconstruction is a promising image processing technique designed to acquire high-resolution (HR) images from their low-resolution (LR) images to overcome their inherent resolution limitations. With the advancement of image super-resolution reconstruction technology, deep learning has made many achievements in the spatial resolution reconstruction of hyperspectral images [4],[5],[6],[7],[8]. Due to the uneven spatial distribution of images, different strategies for different spatial distribution areas can effectively improve the effect of image reconstruction. For example, RAISR [9] separated the picture patches into clusters and created a matching image enhancement strategy for each cluster. It also employs an efficient hashing approach to decrease the complexity of the clustering procedure. Kong et al. [10] created an accelerated super-resolution network that uses data features to recreate specified areas and produced good results. In addition, previous methods have generated decent

textures; however, the edges of the generated images would inconsistent with the real ones, leading to serious artificial artifacts. To acquire clear and precise edges, researchers have devised edge filters that learn picture edges independently [11],[12],[13],[14],[15]. Compared with natural images, hyperspectral images collected by unmanned aerial vehicles, satellites, and other platforms have a wide imaging range. The inhomogeneity of the image spatial information distribution is more obvious. In the application of hyperspectral remote sensing images, the clarity of the object edge is particularly important. However, few studies have combined partition reconstruction and detail enhancement technologies in super-resolution reconstruction of hyperspectral images.

In this article, we propose a super-resolution reconstruction framework for a hyperspectral image called MSDESR, which provides a reliable solution for the spatial reconstruction of hyperspectral remote sensing images. Hyperspectral remote sensing images are used as the input of the submap shunt network to realize the shunt of spatial information, which is divided into two data streams: simple and difficult. The spatial information of the two data streams is separated by high-low frequency information, and details are enhanced. Finally, the data are input into the regional reconstruction module [16] respectively to achieve spatial hyperspectral remote sensing images super-resolution reconstruction. This method can solve the problem that hyperspectral remote sensing images occupy the same amount of computation due to the uneven distribution of spatial information. In addition, the method can transform redundant computing operations into enhanced processing of complex detailed texture information, thus realizing resource conversion to improve reconstruction efficiency.

The contributions of this article are summarized as follows:

-To encode spatial information effectively and enhance details, a multibranch partition reconstruction network is proposed. The proposed MSDESR is a network with high computational efficiency.

-To fully make use of the local spatial information distribution of the hyper- spectral remote sensing images, the submap shunt network is employed for shunting in the image preprocessing stage.

-To viably exploit the detailed textural features of hyperspectral remote sensing images, the multiscale Retinex with detail enhancement algorithm is proposed. The separation of spatial high-low frequency information from hyperspectral remote sensing images is helpful to extract and strengthen the details of high frequency information.

2 Related Work

2.1 Partition image reconstruction and detail enhancement

For different image regions, researchers are beginning to use different processing strategies. However, the key issue is how to allocate different image areas to different processing strategies. Using multipath CNN reconstruction with pathfinders, Ke et al. [17] combined reinforcement learning and deep learning to find suitable paths to recover each region efficiently. RAISR [9] separated picture patches into clusters and

creates a matching image enhancement strategy for each cluster. It also employs an efficient hashing approach to decrease the complexity of the clustering procedure. Inspired by the above methods, SFTGAN [18] introduced a new spatial feature transformation layer that incorporates a high-level semantic prior before guiding image reconstruction in different regions, which is an implicit way to handle different parameter regions. These deeper networks achieve better image shunt performance but due to limited computing power and memory resources, they are difficult to train.

In addition, in order to obtain clear detail information, the researcher further introduced a detail filter into the model. The EP-GAN [19] model extracted detail information as priori labels from high-quality images to guide the network in inferring detailed details. Kui et al [20] built EESN network, removed noise pollution through mask pro-cessing, extracted detail image contour, and combined the extracted detail contour with the restored intermediate image to realize the enhancement of detail information. Compared with natural images, hyperspectral images are more spatially contaminated. If the details of hyperspectral images are extracted and enhanced directly, noisy results and false image details will appear instead. Therefore, the detail enhancement method for natural images is inapplicable to hyperspectral images.

2.2 Hyperspectral super-resolution methods for remote sensing images

Deep neural networks have been used with great success in computer vision [21],[22],[23],[24]. Inspired by the successful application of RGB image processing, deep learning has also been widely used in HSI [25], [26]. Pan et al [27] presented an SR approach based on a residual dense back projection network (RDBPN), which promotes super-resolution reconstruction of remote sensing images with median and large scale factors. Zou et al [28] combined deep residual convolutional neural networks (DRCNN) with spectral unmixing. Hu [29] et al. proposed an intra fusion network (IFN) hat effectively utilizes spectral information between continuous LR bands. Li et al [30] proposed a band-attentive adversarial learning frame, which introduces a series of spatial spectral constraints to improve the reconstruction results. To achieve efficient exploitation of high multispectral features, MSFMNet [31] uses a multiscale feature generation, a fusion multiscale feature mapping block based on wavelet transform, and a spatial attention mechanism to learn the spectral features between different spectral bands. These extremely deep networks achieve good reconfiguration performance. However, treating all areas of the whole image with only one strategy will reduce the effectiveness and efficiency of reconstruction.

3 Porposed Methods

The framework of the MSDESR is shown in Fig. 1. The proposed framework mainly consists of three parts: a submap shunt block, a high-low frequency information extractor with an enhancement block, and a regional image superresolution reconstruction block. First, the input image is spatially decomposed [32]. The data are shunted by a lightweight submap shunt network (based on the DenseNet [33] model) according to the complexity of the image information. Second, mproving the multiscale Retinex [34] with

detail enhancement algorithm to separate the high-low frequency information of the image and enhance the high-frequency information of details in the image. Then, the processed high-low frequency data are fed into large, medium and small networks of different complexity, and the network carries out super resolution reconstruction of high-low frequency information respectively. Finally, the completed super-resolution reconstructed images' high-low frequency information is fused with linear weighting. And the clear images are reconstructed by stitching the submaps into complete images.

3.1 Submap shunts

The submap shunt network classifies images into two categories based on detailed texture information: easy and difficult parts. According to statistics, approximately 36% of the LR subimages (32x32) in the Pavia University dataset fall in the smooth region, while this percentage rises to 56% in the QUST-1 hyperspectral satellite dataset of Gaomi City. Based on this observation, large-scale remote sensing images are spatially decomposed into subimages. Smaller networks are used to deal with smooth regions of the image, and deep networks are used to deal with complex regions [35]. Following spatial decomposition, different networks are used for super-resolution reconstruction of areas containing different morphological information. Subimage decomposition is especially important for large remote sensing images because many areas are easy to reconstruct relatively. Submap shunting not only reduces computational effort and saves memory space in practice but also prepares for the subsequent extraction of high-low frequency information.

As shown in Fig. 2, we create a lightweight submap shunt network inspired by the DenseNet model. The lightweight submap shunt network has four convolutional layers, three LReLU layers, an average pooling layer, and a fully connected layer, all of which are connected in a feed-forward manner. Specific parameters are shown in Table 1. The feature extraction is handled by the convolutional layers, while the pooling and fully connected layers generate probability vectors. Specifically, this classification model allows a probability vector $\mathrm{p}\left(\mathrm{x}_{\mathrm{i}}\right)$ to be generated for a subgraph x_{i} that is decomposed from a large scale image X. The submap is identified by selecting the metric with the highest probability value to determine which class it belongs to. Experiments show that the structure is simple and achieves better performance than DenseNet.

Table 1
Submap diversion network parameter settings

Number of storeys	Convolution kernel size	Input channels	Output channels	Step length
1	4x4	1	256	4
2	1x1	256	256	1
3	1x1	256	256	1
4	1x1	256	32	1

When training with the baseline DenseNet network, it converges to an extreme point, and the resulting images are all classified into complex texture branches. For example, in the case of output vectors, a shunt result vector of [0.90, 0.10] is preferable to [0.54, 0.46]. As the latter appears to be random, the submap classification network loses its functionality. To avoid this issue, we design a lightweight submap shunt network based on the DenseNet model to carry out the submap shunt task, ensuring that both branches have equal chances of being selected. Some of the submap shunt network results are depicted in Fig. 3(a)-(b). Figure 3(a) depicts the resulting map derived from the easily reconstructed shunt data. The result map shows that there is less texture information in the plots, and little difference between textures in terms of thickness, sparsity, and other easily distinguishable information. The results of the difficult reconstructed shunt data in Fig. 3(b) show that the textures are more complex, non-randomly arranged, and densely distributed. By comparing these two images, the submap shunt network can accurately classify the images into two categories: easy and difficult parts, enhancing shunt effectiveness.

3.2 Separation of high-low frequency information

The Retinex algorithm is an image enhancement algorithm based on a human vision system. The traditional multiscale retinex algorithm is used to extract high-low frequency image information, but the local details of high frequency information in the results are very poor. Therefore, we proposed a multiscale Retinex with detail enhancement algorithm to separate high and low frequency information from hyperspectral remote sensing images. It is assumed that the low-frequency information is a spatially smooth image. The improved multiscale Retinex with detail enhancement algorithm can be expressed as:

$$\mathrm{S}(x, y) = \mathrm{R}(x, y)*\mathrm{L}(x, y)$$

1

$$\mathrm{r}(x, y) = \mathrm{D}(x, y) + \sum_{\mathrm{i}=1}^{\mathrm{K}} \mathrm{W_i}\{ \log \mathrm{S}(x, y) - \log \left[\mathrm{F_i}(x, y)*\mathrm{S}(x, y) \right] \}$$

2

Where $S(x, y)$ and $r(x, y)$ represent the original image and the output image. R (x, y) and $L(x, y)$ represent high frequency information and low frequency information respectively. * represents the convolution sign. $\mathrm{F_i}(\mathrm{x}, \mathrm{y})$ is the center surround function. K is the number of Gaussian center surround functions. The K value is usually taken as 3. $\mathrm{W_i}$ is the weighting factor of the scale and $\mathrm{W_1}$ =$\mathrm{W_2}$ =$\mathrm{W_3}$ = 1/3. $D(x, y)$ is the detail recovery section, which is responsible for improving the detail portion of the high-frequency information. The expression for low-frequency information extraction is :

$$\mathrm{L}(x, y) = \mathrm{S}(x, y) / \mathrm{R}(x, y)$$

3

Figure 4 shows some of the results from extracting the image's high-low frequency information from the original remote sensing data. The high-frequency information contains the remote sensing image's main features and detail information, whereas the low-frequency information comprises the remote sensing image with a large amount of smooth information. Therefore, during the reconstruction of super-resolution images, different processing strategies should be used for different frequency information.

The detail recovery section extracts local details in an image by utilizing DoG's multiscale differences [36]. First, the Gaussian kernel function is applied to the image with the high-frequency information $R(x, y)$ to produce three distinct blurred images, and the blurred image of the reflection component is

$$\mathrm{R}_1(\mathrm{x}, \mathrm{y}) = \mathrm{G}_1 \times \mathrm{R}(\mathrm{x}, \mathrm{y}) \quad (4)$$

$$\mathrm{R}_2(\mathrm{x}, \mathrm{y}) = \mathrm{G}_2 \times \mathrm{R}(\mathrm{x}, \mathrm{y}) \quad (5)$$

$$\mathrm{R}_3(\mathrm{x}, \mathrm{y}) = \mathrm{G}_3 \times \mathrm{R}(\mathrm{x}, \mathrm{y}) \quad (6)$$

where G_1, G_2 and G_3 are the Gaussian kernels of the three images with standard deviations of $\sigma_1 = 1.0$, $\sigma_2 = 2.0$ and $\sigma_3 = 4.0$, respectively. Then, high quality details D_{R1}, medium detail D_{R2} and coarse details D_{R3} are extracted for$\mathrm{R}(\mathrm{x}, \mathrm{y})$. The details of high-frequency information can be expressed as

$$D_{R1}(x, y) = R(x, y) - R_1(x, y)$$

7

$$D_{R2}(x, y) = R_1(x, y) - R_2(x, y)$$

8

$$D_{R3}(x, y) = R_2(x, y) - R_3(x, y)$$

9

The detail information D of high frequency information $\mathrm{R}(\mathrm{x}, \mathrm{y})$ is generated by merging three detail images. The expression is as follows

$$\mathrm{D} = [1 - \mathrm{z}_1 \times (\mathrm{D}_{\mathrm{R1}})] \times \mathrm{h}_{\mathrm{R1}} + \mathrm{z}_2 \times \mathrm{D}_{\mathrm{R2}} + \mathrm{z}_3 \times \mathrm{D}_{\mathrm{R3}}$$

10

where z_1, z_2 and z_3 are 0.5, 0.5 and 0.25, respectively. When high-quality detail is added to an image, D will extend the grayscale difference near the edges. But its excessive overshoot may cause grayscale saturation. To overcome this problem, the positive component of D_{R1} is reduced and the negative

component of D_{R1} is enlarged in Eq. (10).Therefore, the detail is increased while the saturation is suppressed in the experiments.

Figure 5 depicts the detail information extracted from the high-frequency information. The majority of the detail information may be extracted from the high-frequency data. Figure 5 shows the results of extracting high-frequency information before and after the Retinex algorithm improvement. In Fig. 5, image1 and image2 are the results of high frequency information extracted from the easy and difficult reconstruction datasets re-spectively. The results show that the details of the edge parts of the features and background in the improved high-frequency information are clearer. This makes the regional image super-resolution reconstruction module more sensitive to the extraction of detailed feature information and better to local feature extraction.

3.3 Partition image reconstruction model

To successfully conduct super-resolution reconstruction, different networks are used to process submaps of various levels of complexity, which is a divide-and-conquer strategy. Deep residual channel attention networks (RCAN) employs a channel attention mechanism in the long and short skip residuals connections to adaptively reclassify the features of the channel to receive more information. In this work, we develop a RCAN -based regionalized image reconstruction method. As shown in Table 2('-O' represents the original network), there is almost no difference in performance between RCAN (32) and RCAN – 0(64) for the "difficult low-frequency and easy low-frequency ” submap, while RCAN (50) can achieve roughly the same performance as RCAN-O(64). This suggests that we can reduce computational costs by using lightweight networks for simple submaps. As shown in Fig. 6, three RCAN models with the same network structure but a different number of channels are employed in this paper. In the first layer, three networks with 32, 50, and 64 channels are used for training of subimage high-low frequency information, specifically difficult low-frequency, easy low-frequency information, easy high-frequency information, and difficult high-frequency information. Because of the small difference in spatial complexity, submap shunt networks classify difficult low-frequency information and easy low-frequency information into one group. As a result, we propose the method for reconstructing regional images that can differentially process different regional images.

Table 2
MPSNR values obtained for the three SR branches of RCAN

Model	FLOPs	Shallow	Medium	Deep
RCAN (32)	8.28G	38.21dB	-	-
RCAN (50)	17.91G	-	28.53dB	-
RCAN (64)	29.34G	-	-	26.33dB
RCAN -O(64)	29.34G	38.20dB	28.56dB	26.32dB

The regional image reconstruction model is divided into three branches: simple, medium, and complex. Each branch is based on RCAN, which is a super-resolution network. The baseline RCAN network is used as the most complex branch, and the other two branches are obtained by reducing the network complexity. Furthermore, controlling network complexity is accomplished by reducing the number of convolutional layers and the number of channels (convolutions) per convolutional layer. Therefore, the smallest number of layers and convolutions for performing all the training tasks can be determined, namely, the maximum branch. Then the minimum and medium branches are determined in the same way. Finally, the completed superresolution reconstructed image's high-low frequency information are fused and linearly weighted. The clear and complete image is reconstructed by stitching the submaps.

$$\mathrm{S} = \mathrm{lR}(\mathrm{x}, \mathrm{y}) + (1 - \mathrm{l})\mathrm{L}(\mathrm{x}, \mathrm{y})$$

11

where $R(x, y)$ is the high-frequency information, $L(x, y)$ is the low-frequency information, S is the final image after the fusion of high-low frequency information using linear weighting. And l is the weighting factor and takes the value 0.5.

To fully demonstrate the ability of the framework to encode spatial spectral information, tests were performed with 2-, 4-, and 8-fold up-sampling factors. According to the data in Table 3, the MSDESR achieves better performance and lower computational cost than the original network, realizing efficient encoding of spatial spectral information, with FLOPs reduced to 52%-70%. Furthermore, the reduction in computation did not lead to a decrease in image quality, indicating that the performance improvement is not at the expense of computational burden. The results certify the significance of diverting the input submaps to their appropriate branches.

Table 3
MPSNR (dB) values for each upsampling factor experiment

	RCAN -O	MSDESR
2Scale Factor	33.31	35.67
FLOPs	29.34G (100%)	21.12G (72%)
4Scale Factor	29.72	30.64
FLOPs	29.34G (100%)	18.48G (63%)
8Scale Factor	25.56	25.99
FLOPs	29.34G (100%)	14.96G (51%)

4 Experimental Results And Analysis

The framework was trained on an ubuntul10. 4 system using an NVIDIA GTX1080Ti GPU device with 28 G of memory. The network was implemented by the open-source PyTorch deep learning framework, and the network parameters were optimized using the Adam correction method, with the initial learning rate set to 0.001. Attempts were made to set the learning rates to 0.1 and 0.01. The experimental results show that when the initial learning rate is 0. 001, the best results are obtained compared to the other.

4.1 Dataset information

In this section, the MSDESR is applied to three hyperspectral image datasets,: the QUST-1 satellite dataset, the Pavia University[37] dataset, and the Chikusei dataset[38]. The QUST-1 satellite data are hyperspectral remote sensing images taken on July 2019 in Weifang City, Shandong Province, China. The spectral range is 400 nm to 1000 nm. The images are equally divided into 32 spectral bands with a spatial resolution of ten meters. The images are cropped to 128×128 pixels from approximately 5000×5000 pixels. The Pavia University data are part of the 2003 hyperspectral data from the city of Pavia, Italy, covering 103 spectral bands from 430 nm to 860 nm, with a spatial resolution of 1.3 meters and 610×340 pixels. The Chikusei dataset was taken on 29 July 2014 by the Headwall hyperspectral imaging sensor in the agricultural and urban areas of Chikusei, Ibaraki Prefecture, Japan. The spectral range was 128 bands from 363 nm to 1018 nm with a spatial resolution of 2. 5 meters. The scene consisted of 2517×2335 pixels, and the images were cropped to a size of 128x 128 pixels with 19 categories.

For each image in each dataset, we randomly chose a 128x128 pixel region to test the performance of our pro-posed hyperspectral image super-resolution reconstruction framework MSDESR. Another 128×128 region was randomly selected for validation. And the remainder was selected for training. To create the super-resolution reconstruction training set, the hyperspectral remote sensing images are used as the base HR image. The LR images are created by simulating the blur method of the remote sensing image, adding Disk blur, and using double triple interpolation. To achieve varying blur levels, various disk blur kernels are performed on the LR images[39].

4.2 Evaluation metric

To evaluate the performance of the proposed method comprehensively, a combination of classical evaluation metrics and visual effects is used for validation. The classical evaluation metrics include signal to reconstruction error (SRE), mean peak signal-to-noise ratio (MPSNR), mean structural similarity (MSSIM), and mean root mean square error (MRMSE) [40]. SRE is a global image quality indicator based on the signal error between the reconstructed images and the scene ground truth images, while MPSNR and MRMSE use the mean square error to estimate the similarity between the generated images and the ground truth image. MSSIM emphasizes structural consistency. Generally, the larger the values of SRE, MPSNR and MSSIM, the better the spatial quality, whereas, the smaller the value of MRMSE, the less the spectral distortion.

4.3 Results and discussion

To assess the performance of the MSDESR, a series of experiments are carried out on three benchmark datasets: QUST-1, Pavia University, and Chikusei. Three of the most advanced deep learning methods with common code were selected as the baseline for comparison on the 4× sampling factor: FSRCNN, LapSRN, and RCAN. Table 4 shows the results of experiment. Figures 7, 8, and 9 show the corresponding visualization results, with the QUST-1 satellite dataset in band 1, the Pavia University dataset in band 49, and the Chikusei dataset in band 93. The red rectangular area in the bottom right corner of the image indicates that the image has been magnified by a factor of three.

Table 4
MPSNR (dB) values for each upsampling factor experiment

Datasets	Indicators		FSRCNN	LapSRN	RCAN	MSDESR
QUST-1 satellite	SRE↑	39.36	39.88	40.57	**41.78**	
	MPSNR↑	28.87	29.96	33.31	**34.67**	
	MSSIM↑	0.58	0.60	**0.67**	0.64	
	MRMSE↓	8.70	8.05	8.71	**7.78**	
Pavia University	SRE↑	33.77	31.98	35.57	**36.18**	
	MPSNR↑	28.34	29.24	30.29	**31.50**	
	MSSIM↑	0.78	0.80	**0.88**	0.85	
	MRMSE↓	5.49	4.66	5.12	**4.14**	
Chikusei	SRE↑	35.50	35.86	37.58	**38.42**	
	MPSNR↑	35.90	37.06	38.04	**39.08**	
	MSSIM↑	0.86	0.90	**0.94**	0.91	
	MRMSE↓	6.77	5.75	6.47	**5.16**	

The results of experiment from the QUST-1 satellite data in Table 4 show that the proposed hyperspectral image super-resolution reconstruction framework MSDESR achieves the best performance in terms of the SRE, MPSNR, MSSIM, and MRMSE indexes, which are 42.78, 35.67, 0.64, and 7.78, respectively. The SRE, MPSNR, and MSSIM are improved by 5.42, 5.80, and 0.06, respectively, compared with the FSRCNN. The results from the Pavia University data show that MSDESR still has better reconstruction results compared to other methods. In comparison to the FSRCNN, the SRE increased by 2.41, the MPSNR increased by 3.36, the MSSIM decreased by 0.07, and the MRMSE decreased by 1.35. For the Chikusei dataset, compared with FSR- CNN, the proposed model also achieves the optimal reconstruction results, with SRE, MPSNR, and MSSIM improved by 2.52, 3.38, and 0.08 respectively, the MRMSE was reduced by 2.16.

When dealing with relatively simple, irregular natural pattern data, such as the QUST-1 satellite dataset, which is dominated by natural scenes such as plains, fields and small towns, the submap shunt block of the proposed method divides most submap into a category that is easy to reconstruct. The subtle texture and spectral information are preserved properly by the multiscale Retinex structure with detail enhancement, which is transmitted to the simple branch of the regional image reconstruction module for reconstruction. The local image information is smoother and achieves better results with less computation. The Pavia University dataset consists of urban scenes in various frequency bands, which are mainly regular artificial patterns. When dealing with this kind of data, the submap shunt structure divides most of the submaps into the category of difficult-to-reconstruction. Through the structure of multiscale Retinex with detail enhancement and the complex branches of regional image reconstruction module, the edge and texture detail features of different frequency bands are extracted more comprehensively and involved in the reconstruction calculation, so as to avoid the jagged boundary of the image and depict a more realistic shape. Of course, when dealing with moderately complex data, such as the Chikusei suburban-urban combination, the proposed method can also make flexible processing and achieve the optimal computation and performance.

The visualization results are shown in Figs. 7, 8, and 9. And the MPSNR curves for all the band spectra are shown in Fig. 10. These results further demonstrate the superiority of the proposed framework. Figure 7, depicts the FSRCNN effect plot of the experimental results of the QUST-1 satellite dataset. The FSRCNN fails to extract and recover the gaps between small buildings. Although these gaps can be seen in the LapSRN and RCAN reconstruction results, they are not obvious and difficult to detect. The proposed framework preserves the details and makes them more visible by comparison. In the experimental results of Pavia University datasets in Fig. 8, there is a long black outline building in the ground truth image, which is difficult to distinguish from the results of other methods. In contrast, this long black outline can still be observed in the result generated by the proposed method. Figure 9 shows the experimental results of the Chikusei dataset. The two intersecting field paths in the real image are difficult to distinguish in the results of the other methods, or only one of them can be recognized with difficulty. In contrast, these two paths are still visible in our results, with clear details and a consistent structure. This means that the proposed MSDESR can take advantage of the detailed texture features of hyperspectral remote sensing images and reconstruct delicate textures more fully. The results indicate that our network can better preserve spectral information and maintain detailed features.

In all bands for the MPSNR, the proposed method consistently outperforms the other methods. When local spatial details are considered, the results show that the algorithm in this paper can better learn and strengthen spatial features and improve the coherence of structural features. As a result, it can be concluded that the proposed algorithm performs better on datasets acquired by both CMOS (complementary metal-oxide-semiconductor) and ROSIS (reflective optics system imaging spectrometer) sensors.

In the practical application of hyperspectral remote sensing data, in addition to spectral data, corresponding spatial morphological information is also necessary. The spatial resolution of

hyperspectral remote sensing data cannot meet the demand for spatial morphological information of the features. Although researchers [12],[13],[14],[15] investigated the use of cascaded residual methods to improve the spatial resolution, treating data with uneven information distribution equally cannot make symptomatic reconstruction for spatial morphological information with different features. The results of experiment show that the proposed method is not only capable of preserving the true spatial morphological information of features for symptomatic reconstruction without compromising the spectral information, but also effectively makes use of the computational load. It allows for further improvements in spatial resolution and improves the accuracy of feature recognition.

4.4 Discussion on the Proposed Framework: Ablation Study

The importance of each component of the proposed framework is validated through ablation experiments. Without the submap shunt network branch, without the high-low frequency information extraction branch, and without the improved high-low frequency information extraction branch are the settings of the various modules in the ablation experiments. Except for the ablation module, all three comparison methods use the same settings as the proposed model. Table 5 and Fig. 11 depict the results of experiments performed on the QUST-1 datasets. According to the table, the super-resolution reconstruction results using the improved detail enhancement algorithm improved the MPSNR values by approximately 6.4% compared with that of the unimproved algorithm, demonstrating the effectiveness of the detail enhancement algorithm. The super-resolution reconstruction metrics of the proposed framework outperform that of other ablation experimental models, demonstrating the importance of each component.

Table 5
Results on the QUST-1 dataset from different network architectures.

	Without the sub-map shunt network	Without the high and low frequency information extraction	Without the improved high and low frequency information	Complete framework
SRE↑	39.66	35.12	40.19	**41.78**
MPSNR↑	33.32	29.64	32.43	**34.67**
MSSIM↑	**0.68**	0.59	0.6	0.64
MRMSE↓	7.62	7.74	7.55	**7.78**

5 Conclusion

In this study, a hyperspectral image super-resolution reconstruction framework MSDESR is proposedt o improve the spatial resolution of hyperspectral remote sensing images. We design a submap shunt network to minimize the computational load of reconstruction process in light of the uneven spatial information distribution during super-resolution reconstruction and the varying difficulty of image reconstruction in different regions by classify hyperspectral remote sensing images accurately. The high-low frequency information extraction and enhancement modules are designed to address the extraction

of texture detail to fully retain texture detail information and eliminate unnecessary noise. The partition image reconstruction block is designed to reduce the difficulty of HSI reconstruction by targeting different reconstruction difficulties. Compared with other super-resolution reconstruction methods of hyperspectral remote sensing images, this framework can recover finer details and has higher reconstruction efficiency. The focus of our future research will be on how to perform more accurate weighted fusion of high-low frequency information of images under certain constraints.

Declarations

Funding Shandong Province Demonstration Base Project of Postgraduate Joint Training for Industry-Education Integration and Diagnostic study on phosphorus nutrition based on multi-angle and multi-temporal spatial structure characteristics of wheat populations.Grant numbers is (2020-19) and (ZR202102180604)

Conflict of interest All of the authors declare that they have no conflict of interest.

Human and animal rights This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Akgun T, Altunbasak Y, Mersereau RM (2005) Super-resolution reconstruction of hyperspectral images. IEEE Trans Image Process 14(11):1860–1875
2. Gou S, Liu S, Yang S, Jiao L (2014) Remote sensing image super-resolution reconstruction based on nonlocal pairwise dictionaries and double regularization. IEEE J Sel Top Appl Earth Observations Remote Sens 7(12):4784–4792
3. Glasner D, Bagon S, Irani M (2009) : Super-resolution from a single image. In: 2009 IEEE 12th International Conference on Computer Vision, pp. 349–356 IEEE
4. He K, Zhang X, Ren S, Sun J (2016) : Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778
5. Kim J, Lee JK, Lee KM (2016) : Accurate image super-resolution using very deep convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1646–1654
6. Luo Y, Zhou L, Wang S, Wang Z (2017) Video satellite imagery super resolution via convolutional neural networks. IEEE Geosci Remote Sens Lett 14(12):2398–2402
7. Lan R, Sun L, Liu Z, Lu H, Su Z, Pang C, Luo X (2020) Cascading and enhanced residual networks for accurate single-image super-resolution. IEEE Trans cybernetics 51(1):115–125
8. Daihong J, Sai Z, Lei D et al (2022) Multi-scale generative adversarial network for image super-resolution[J]. Soft Comput 26(8):3631–3641

9. Romano Y, Isidoro J, Milanfar P (2016) Raisr: rapid and accurate image super resolution. IEEE Trans Comput Imaging 3(1):110–125
10. Kong X, Zhao H, Qiao Y, Dong C (2021) : Classsr: A general framework to accelerate super-resolution networks by data characteristic. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 12016–12025 *Hyperspectral image super-resolution reconstruction…* 19
11. Jiang K, Wang Z, Yi P, Wang G, Lu T, Jiang J (2019) Edge-enhanced gan for remote sensing image superresolution. IEEE Trans Geosci Remote Sens 57(8):5799–5812
12. Mao Q, Wang S, Wang S, Zhang X, Ma S (2018) : Enhanced image decoding via edge-preserving generative adversarial networks. In: 2018 IEEE International Conference on Multimedia and Expo (ICME), pp. 1–6 IEEE
13. Yang W, Feng J, Yang J, Zhao F, Liu J, Guo Z, Yan S (2017) Deep edge guided recurrent residual learning for image super-resolution. IEEE Trans Image Process 26(12):5895–5907
14. Rabbi J, Ray N, Schubert M, Chowdhury S, Chao D (2020) Small-object detection in remote sensing images with end-to-end edge-enhanced gan and object detector network. Remote Sens 12(9):1432
15. Kamgar-Parsi B, Rosenfeld A (1999) Optimally isotropic laplacian operator. IEEE Trans Image Process 8(10):1467–1472
16. Mao Q, Wang S, Wang S, Zhang X, Ma S (2018) : Enhanced image decoding via edge-preserving generative adversarial networks. In: 2018 IEEE International Conference on Multimedia and Expo (ICME), pp. 1–6 IEEE
17. Jiang K, Wang Z, Yi P, Wang G, Lu T, Jiang J (2019) Edge-enhanced gan for remote sensing image superresolution. IEEE Trans Geosci Remote Sens 57(8):5799–5812
18. Zhang Y, Li K, Li K, Wang L, Zhong B, Fu Y (2018) : Image superresolution using very deep residual channel attention networks. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 286–301
19. Yu K, Wang X, Dong C, Tang X, Loy CC (2021) : Path-restore: Learning network path selection for image restoration.IEEE Transactions on Pattern Analysis and Machine Intelligence
20. Zhang Y, Li X, Zhou J (2019) Sftgan: a generative adversarial network for pan-sharpening equipped with spatial feature transform layers. J Appl Remote Sens 13(2):026507
21. Tian C, Xu Y, Fei L, Yan K (2018) : Deep learning for image denoising: A survey. In: International Conference on Genetic and Evolutionary Computing, pp. 563–572 Springer
22. Li Y, Zheng W, Cui Z, Zhang TFace recognition based on recurrent *20 Hyperspectral image super-resolution reconstruction…*regression neural network. Neurocomputing **297**, 50–58
23. Li B, Dai Y, He M (2018) Monocular depth estimation with hierarchical fusion of dilated cnns and soft-weighted-sum inference. Pattern Recogn 83:328–339
24. Li B, Shen C, Dai Y, Van Den Hengel A, He M (2015) : Depth and surface normal estimation from monocular images using regression on deep features and hierarchical crfs. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1119–1127

25. Li W, Wu G, Zhang F, Du Q (2016) Hyperspectral image classification using deep pixel-pair features. IEEE Trans Geosci Remote Sens 55(2):844–853
26. Gao H, Zhang G, Huang M (2021) Hyperspectral image superresolution via structure-tensor-based image matting. IEEE J Sel Top Appl Earth Observations Remote Sens 14:7994–8007
27. Pan Z, Ma W, Guo J, Lei B (2019) Super-resolution of single remote sensing image based on residual dense backprojection networks. IEEE Trans Geosci Remote Sens 57(10):7918–7933
28. Zou C, Huang X (2020) Hyperspectral image super-resolution combining with deep learning and spectral unmixing. Sig Process Image Commun 84:115833
29. Hu J, Chen H, Zhao M, Li Y (2020) : Deep intra fusion for hyperspec- tral image super-resolution. In: IGARSS 2020–2020 IEEE International Geoscience and Remote Sensing Symposium, pp. 2663–2666 IEEE
30. Li J, Cui R, Li B, Song R, Li Y, Dai Y, Du Q (2020) Hyperspectral image super-resolution by band attention through adversarial learning. IEEE Trans Geosci Remote Sens 58(6):4304–4318
31. Zhang J, Shao M, Wan Z, Li Y (2021) Multi-scale feature mapping network for hyperspectral image super-resolution. Remote Sens 13(20):4180
32. Yu K, Dong C, Lin L, Loy CC (2018) : Crafting a toolchain for image restoration by deep reinforcement learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2443–2452 *Hyperspectral image super-resolution reconstruction…* 21
33. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) : Densely connected convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4700–4708
34. Land EH (1964) The retinex. Am Sci 52(2):247–264
35. Veganzones MA, Simoes M, Licciardi G, Yokoya N, Bioucas-Dias JM, Chanussot J (2015) Hyperspectral super-resolution of locally low rank images from complementary multisource data. IEEE Trans Image Process 25(1):274–288
36. Kim Y, Koh YJ, Lee C, Kim S, Kim C-S (2015) : Dark image enhancement based onpairwise target contrast and multi-scale detail boosting. In: 2015 IEEE International Conference on Image Processing (ICIP), pp. 14041408 IEEE
37. Huang X, Zhang L (2009) A comparative study of spatial approaches for urban mapping using hyperspectral rosis images over pavia city, northern italy. Int J Remote Sens 30(12):3205–3221
38. Yokoya N, Iwasaki A (2016) : Airborne hyperspectral data over chikusei.Space Appl. Lab., Univ. Tokyo, Tokyo, Japan, Tech. Rep. SAL-2016-05-27
39. Xie W, Jia X, Li Y, Lei J (2019) Hyperspectral image super-resolution using deep feature matrix factorization. IEEE Trans Geosci Remote Sens 57(8):6055–6067
40. Li J, Cui R, Li B, Song R, Li Y, Du Q (2019) Hyperspectral image superresolution with 1d-2d attentional convolutional neural network. Remote Sens 11(23):2859

Figures

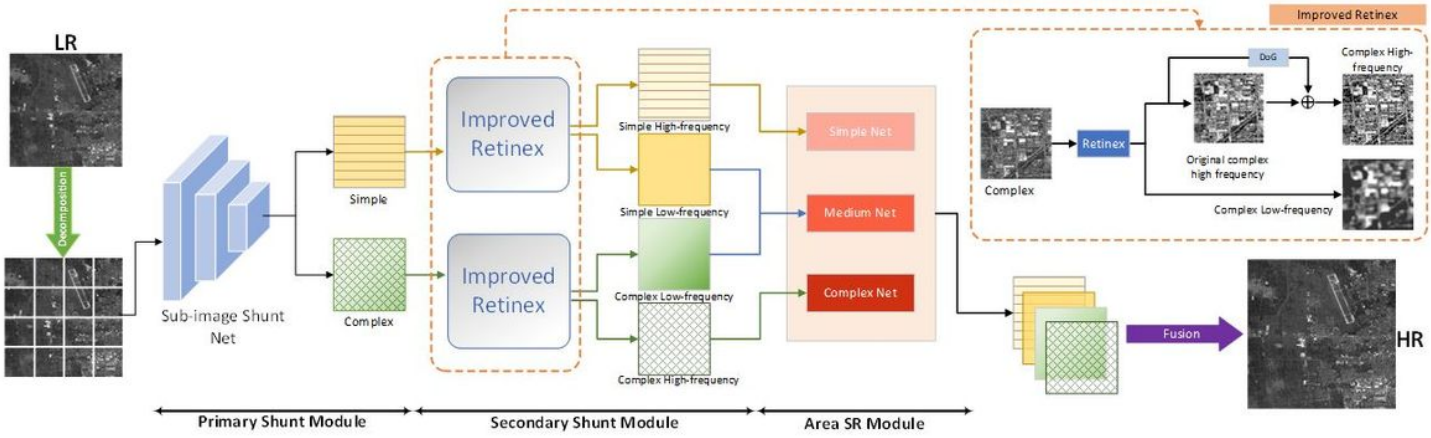

Figure 1

MSDESR Framework

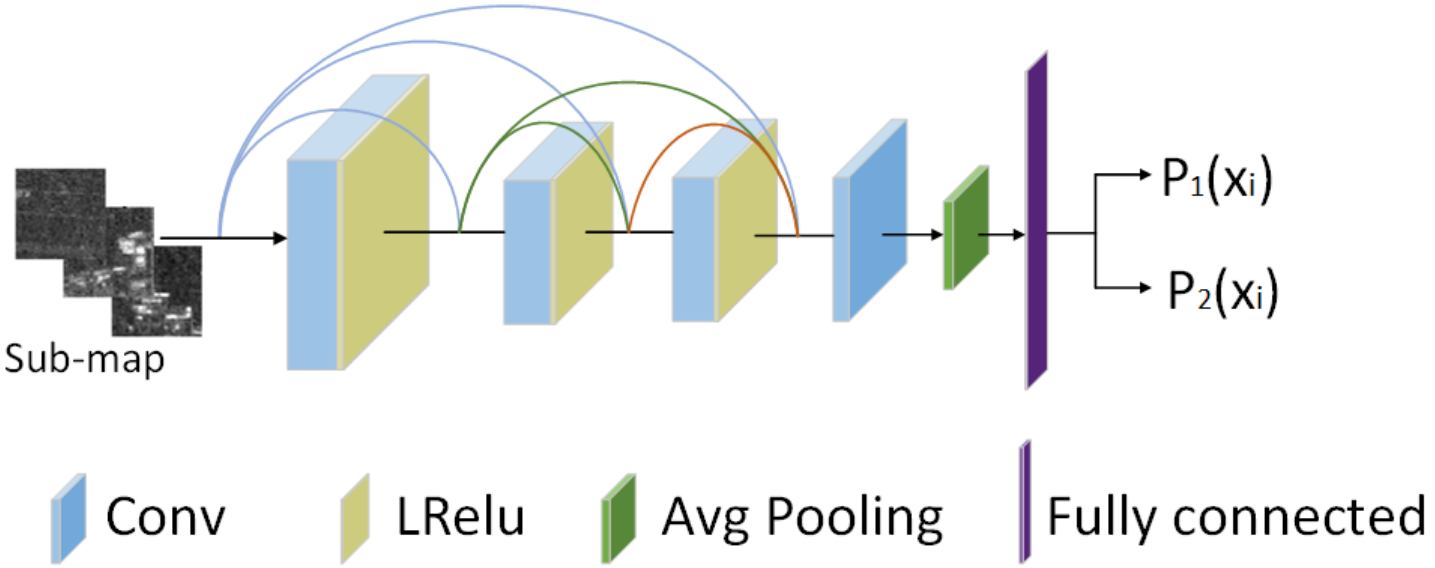

Figure 2

Submap shunt network

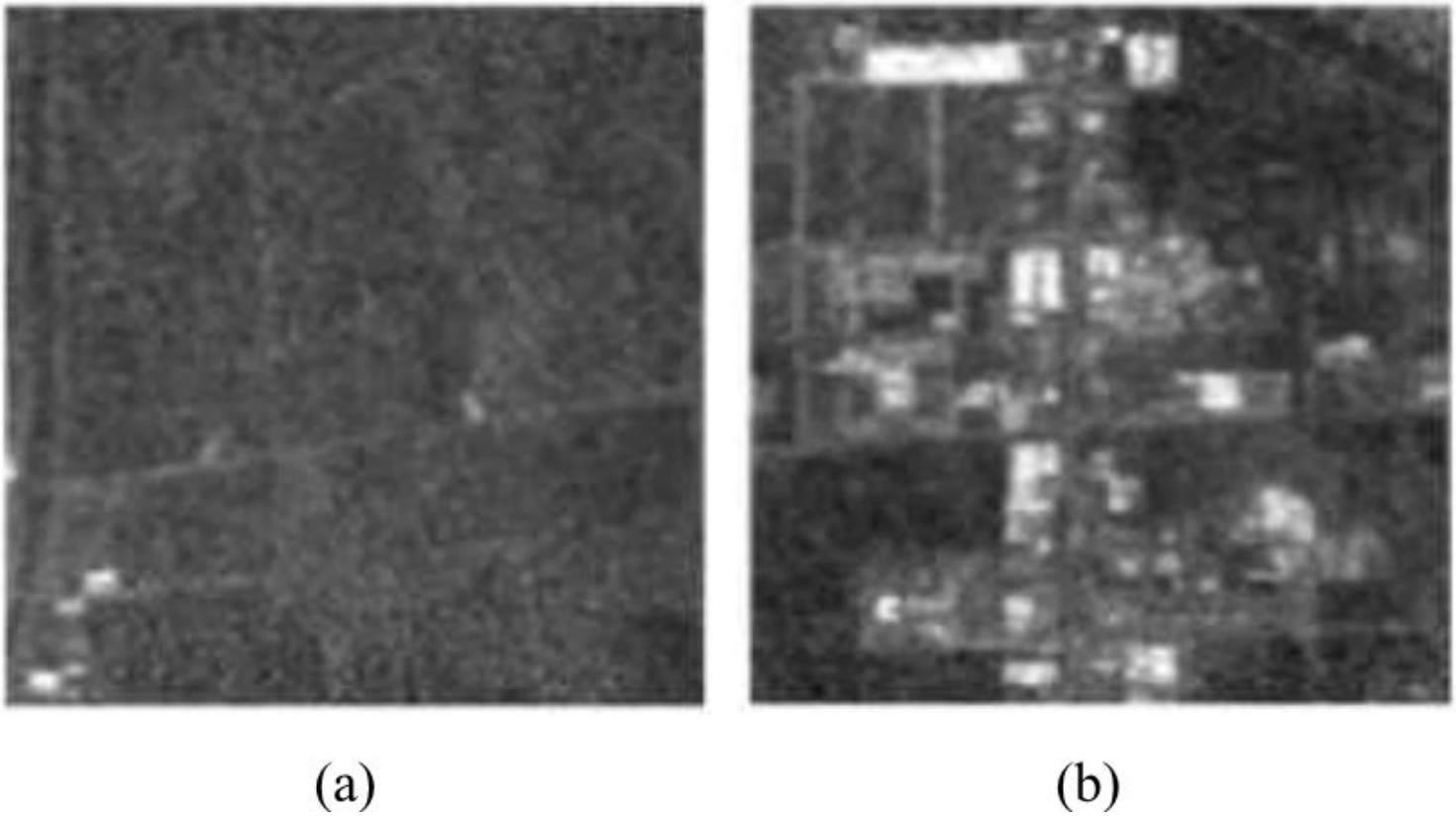

Figure 3

Schematic diagram of shunt results. figures (a) and (b) show the result of the experiment with the easy and difficult reconstructed shunt data.

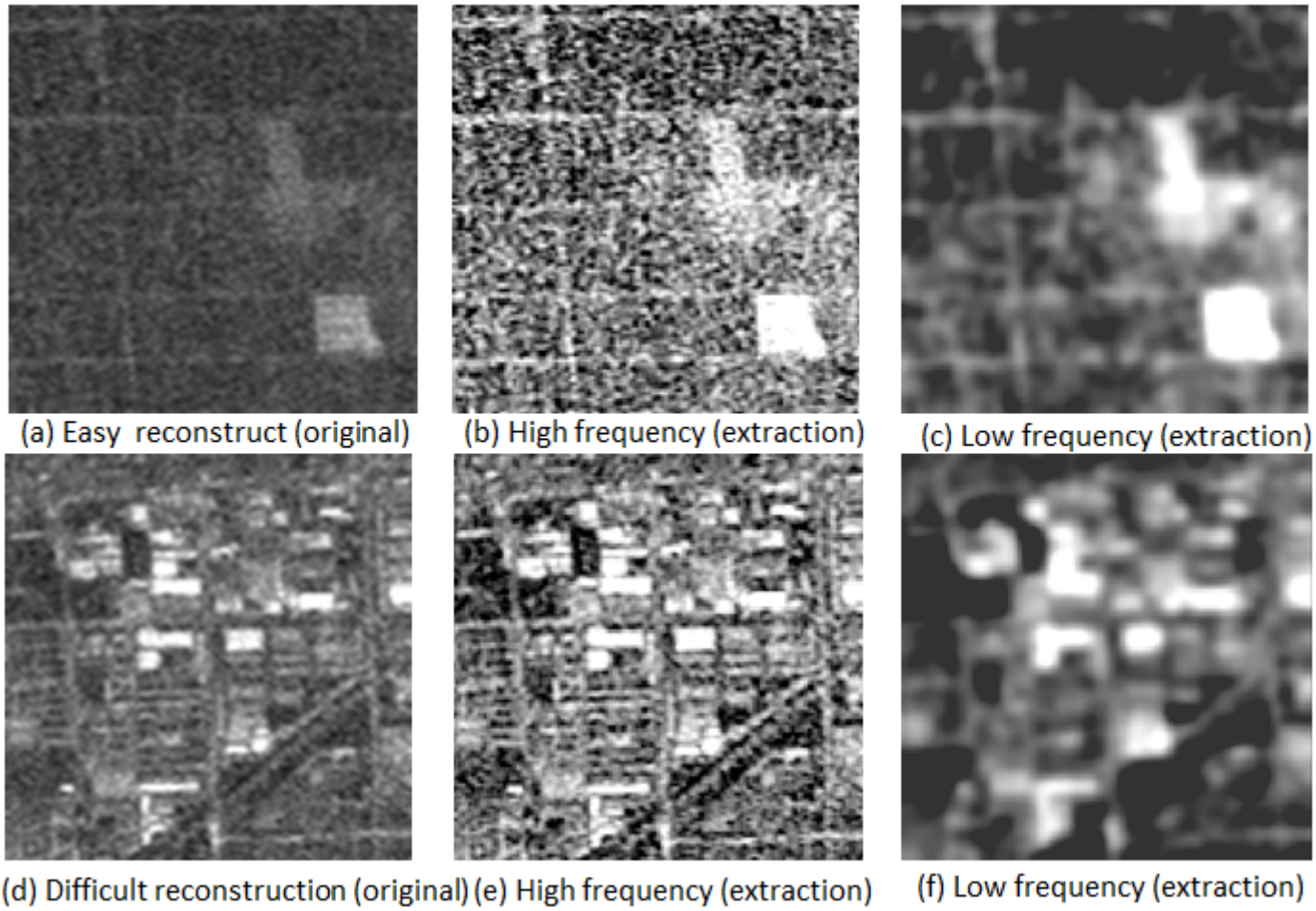

Figure 4

High-low frequency information separation effect

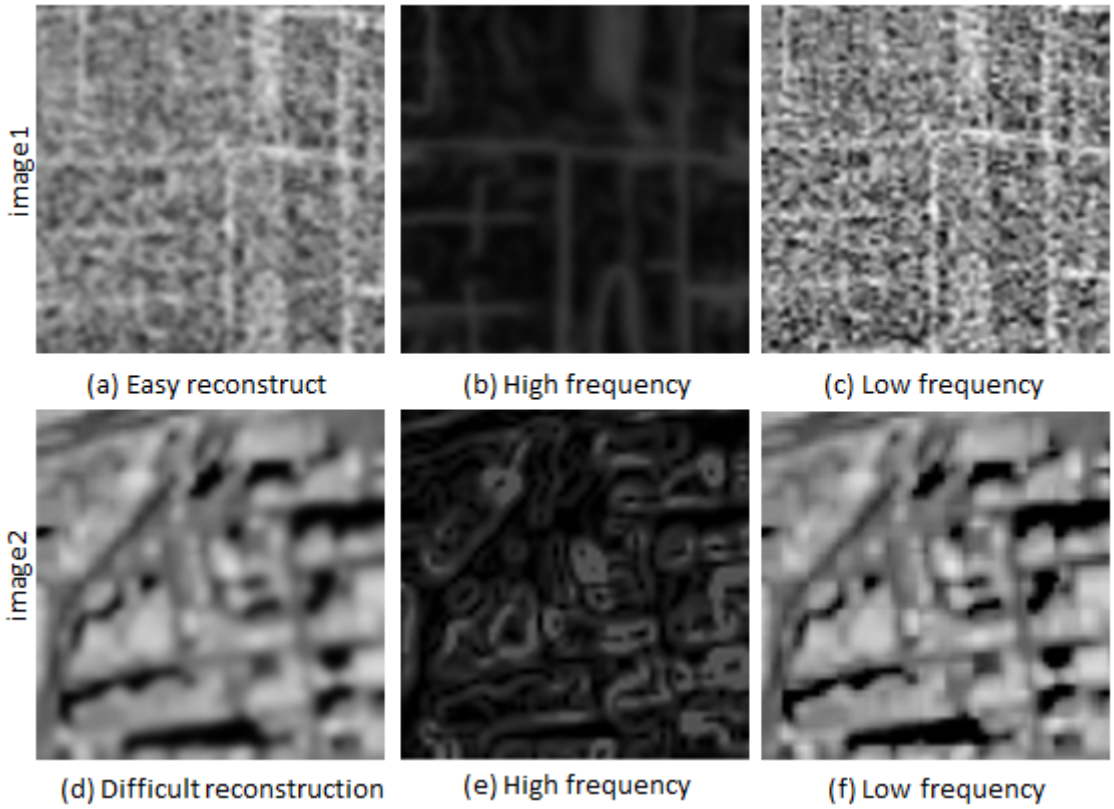

Figure 5

High-frequency information comparison.(a) and (d) are the original algorithm extraction results, (b) and (e) are the extracted detail information, and (c) and (f) are the improved algorithm extraction results.

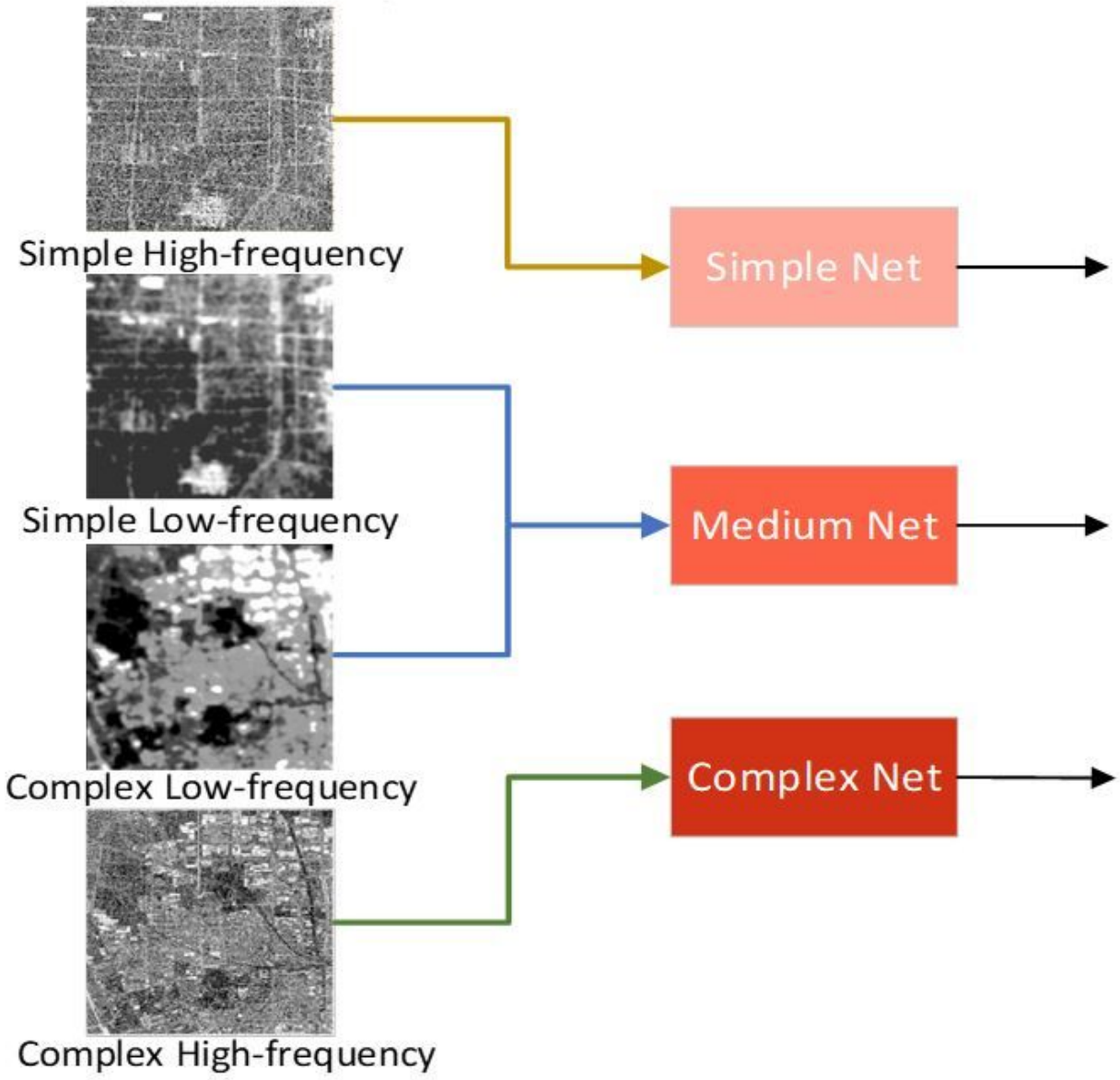

Figure 6

MSDESR Framework

Figure 7

Experimental results of the QUST-1 satellite dataset

Figure 8

Experimental results for the Pavia University dataset

Figure 9

Experimental results for the Chikusei dataset

Figure 10

MPSNR as a function of spectral band wavelength for the methods tested on the hyperspectral image dataset.

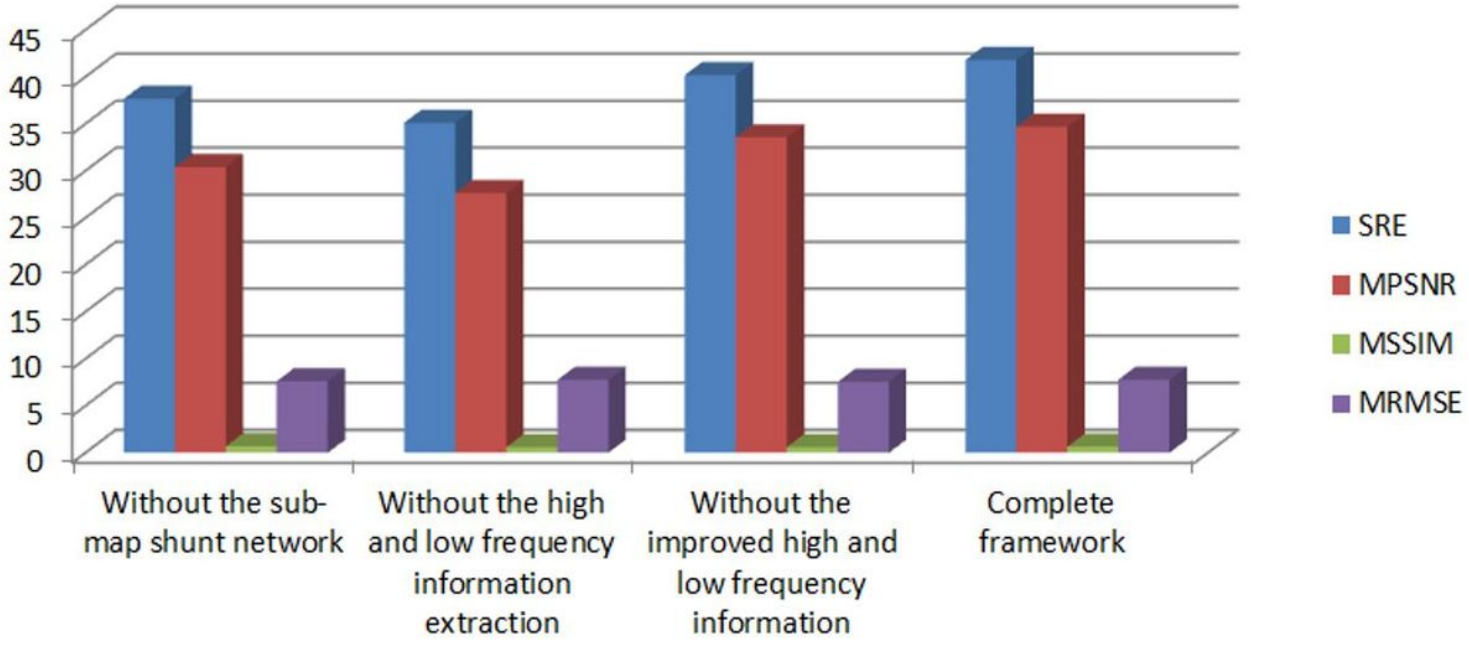

Figure 11

Results of ablation experiments on QUST-1 datasets