Received: 19 July 2022 | Revised: 17 October 2022 | Accepted: 3 November 2022
IET Image Processing
DOI: 10.1049/ipr2.12682

ORIGINAL RESEARCH

Compressive sensing reconstruction of hyperspectral images based on codec space-spectrum joint dense residual network

Shuming Xiao[1,2] | Ye Zhang[1,2] | Xuling Chang[1,2] | Jiajia Xu[1,2]

[1]Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, No.3888 Dong Nanhu Road, Erdao, Changchun 130033, China

[2]Daheng College, University of Chinese Academy of Sciences, Beijing 100000, China

Correspondence
Ye Zhang, Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun, China.
Email: yolanda@spirits.ai

Funding information
This work was supported in part by the CIOMP-Fudan University Joint Fund, Grant/Award Number: Y9S333T190; Jilin Provincial Institute Cooperation Program Fund under, Grant/Award Number: 2020SYHZ0031

Abstract
The spatial and spectral information contained in the hyperspectral image (HSI) make it widely used in many fields. However, the sharp increase of HSI data brings enormous pressure to the data storage and real-time transmission. The research shows that hyperspectral compressive sensing (HCS) breaks through the bottleneck of the Nyquist sampling theorem, which can relieve the massive pressure on data storage and real-time transmission. Existing HCS methods try to design advanced compression sampling matrix or reconstruction algorithms, but cannot connect the two through a unified framework. To further improve the image reconstruction quality, a novel codec space-spectrum joint dense residual network (CDS2-DResN) is proposed. The CDS2-DResN is divided into block compression sampling part and reconstruction part. For block compression sampling, coded convolutional layer (CCL) is leveraged to compress and sample HSI. For measurements reconstruction, deconvolution layer is first leveraged to initially reconstruct HSI, and then build a space-spectrum joint network to refine the initial reconstructed HSI. Moreover, the CCL and reconstruction network are optimized via a unified framework, which can simplify the pre-processing and post-processing process of HCS. Extensive experiments have shown that CDS2-DResN has an excellent HCS reconstruction effect at measurement rates 0.25, 0.10, 0.04 and 0.01, respectively.

1 | INTRODUCTION

The hyperspectral image (HSI) contains not only the spatial position information of substance, but also the spectral characteristics of each pixel. Therefore, spatial position information and spectral characteristic information make it widely used in mineral exploration, agricultural production, environmental monitoring, and military reconnaissance [1–4]. However, as the spatial resolution and spectral resolution of HSI have increased significantly, the amount acquired data has increased dramatically. Therefore, to reduce the enormous pressure of data storage and real-time transmission, HSI compression has become a research hotspot in recent years.

The traditional compression technology first obtains HSI data, and then discards redundant information, so as to achieve the effect of convenient storage and real-time transmission. However, the theoretical basis of data acquisition in this technology is Nyquist sampling theorem, which states that the underlying analogue signal must be uniformly sampled at a sampling rate no less than twice the signal bandwidth to preserve the signal information [5]. Therefore, redundant information can only be discarded in the compression stage, and the compression process after data collection is extremely wasteful. Compressive sensing (CS) [6] technology breaks through the bottleneck of Nyquist sampling theorem, which can collect data at a low sampling rate (much lower than the Nyquist sampling rate). It can complete data compression simultaneously as data collection. Moreover, the CS reconstruction algorithm can ideally reconstruct the original data according to the collected sampling data on the premise of sparse original data [7]. Therefore, CS can relieve the massive pressure on data storage and real-time transmission. Therefore, hyperspectral compressive sensing (HCS) has received more and more attention in the field of hyperspectral.

Existing HCS methods try to design advanced compression sampling matrix or reconstruction algorithm, but cannot

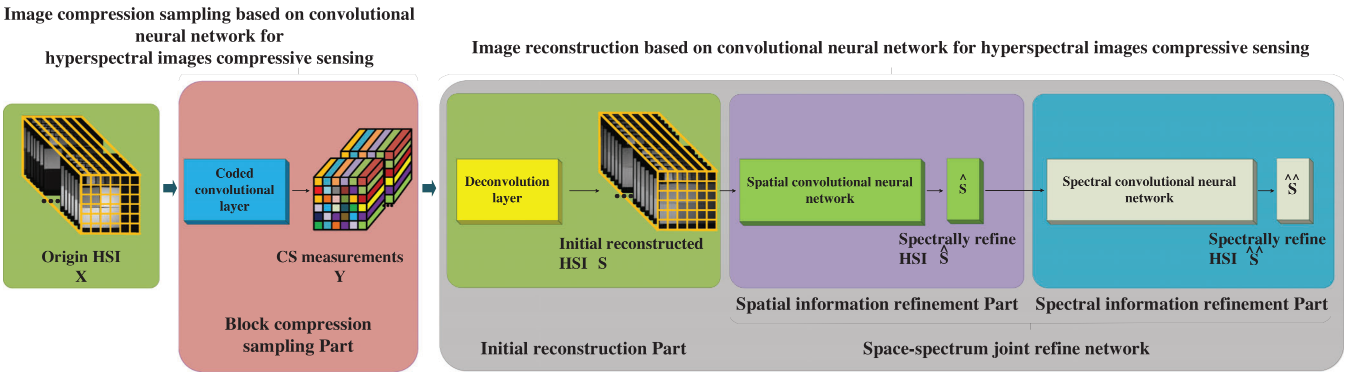


FIGURE 1 Illustration of the overall framework of CDS2-DResN, including four key components: CCL, DeConV, spatial CNN and spectral CNN

connect the two through a unified framework, which limits the further improvement of image reconstruction quality in HCS. Therefore, the joint optimization of compression sampling matrix and reconstruction algorithm can further improve image reconstruction quality.

To further enhances the quality of HSI reconstruction, we propose a convolutional neural networks (CNN)-based image reconstruction network, CDS^2-DResN. Figure 1 shows the overall framework of CDS^2-DResN. Firstly, since the predefined measurement matrix (PMM), for example, an Orthogonal Random Gaussian Matrix, ignores the properties of sampled HSI, coded convolutional layer (CCL) without bias or activation function is introduced for HSI compression and sampling. CCL adaptively generates a learned measurement matrix (LMM) from training HSI, which allows CS measurements to retain more spatial and spectral feature information. Secondly, the deconvolution layer (DeConV) is introduced to initially reconstruct HSI. Thirdly, since the complexity of HSI spatial information, a spatial CNN is designed to refine the reconstructed HSI. Especially, the dense residual block (DRB) is added to spatial CNN, which can accurately extract the spatial feature information in initial reconstructed HSI. Fourthly, since the adjacent spectral bands have strong similarities, a spectral CNN is designed to refine the spatial refinement HSI. Especially, the residual block is added to spatial CNN, which can accurately extract the spectral feature information in spatial refinement HSI.

Our main contributions are as follows:

1. To depend on CNN to implement the block compression sampling part in CS, we leverage the CCL to compress and sample HSI. The LMM, the weight of the learned CCL, can obtain more useful image reconstruction information compared to PMM.
2. Consider the extensive correlation of HSI in spatial and spectral dimensions during the reconstruction part, we design a space-spectrum joint refinement network to refine the initial reconstructed HSI. DRBs in spatial refinement network can better extract multi-layer feature information. The information extraction of adjacent spectral bands in spectral refinement network can better extract spectral feature information.
3. Considering the problem of the difference between adjacent spectral in the HSI reconstruction process, we add a regular term of adjacent spectral differences (RTASpeD) based on Mean Squared Error (MSE), which can significantly improve the reconstruction index SAM.
4. To simplify the HCS pre-processing and post-processing process, CDS^2-DResN jointly trains the compression sampling part and measurements reconstruction part via connecting the CCL and reconstruction network, which can eliminate the block effect of reconstructed hyperspectral images (HSIs).

The experimental results on CAVE and Harvard datasets demonstrate that CDS^2-DResN implements excellent HSI reconstruction performance, proving the feasibility of our model.

2 | RELATED WORKS

In recent years, many HCS reconstruction algorithms have been proposed. The original traditional reconstruction algorithm was to reconstruct each spectrum independently. However, this method only uses the spatial information of HSIs, so the reconstruction quality is low. To improve the reconstruction quality of HSIs, traditional HCS reconstruction algorithms are continuously developed, which can be divided into four categories: HCS based on reference image [8, 9], HCS based on joint optimization [10–13], HCS based on linear mixture model [14–17] and HCS based on tensor-CS [18–21].

The HCS based on reference image is more flexible and the spatial complexity of the method is low. However, it cannot fully leverage the prior information of the HSIs, so the reconstruction quality is limited. The HCS based on joint optimization makes full use of the prior information of HSIs, thus improving the reconstruction quality. However, it needs to restore all band images at the same time, so the computational complexity and spatial complexity of the algorithm are relatively high. The linear mixture model-based HSI reconstruction reduces the degree of freedom of the algorithm itself, so the reconstruction quality is higher. However, if the end member matrix information in the end member matrix is unknown or cannot be accurately estimated, the performance of this method will be significantly reduced. The sampling matrix used in HCS based on tensor-CS has a special structure, so the computational complexity and spatial complexity are very low. However, the compressed measurements corresponding to it has high redundancy. Therefore, the image reconstruction quality is not high.

Although the above-mentioned traditional HSI reconstruction algorithms consider the characteristics of HSI in all aspects, they have common problems. Due to the iterative method to solve the problem, the computational complexity is high and the reconstruction quality is poor at low measurement rates (MRs).

Recently, the CNN-based HCS reconstruction algorithm has become a research hotspot in the field of HCS, which effectively avoids many calculations in the traditional iterative algorithm and achieves excellent HCS reconstruction performance with its strong self-learning ability. Xiong et al. [22] proposed the HSCNN model, which first reconstructed the initial image by an interpolation method, and then enhanced the initial results with CNN to obtain high-quality reconstructed HSI. Choi et al. [23] proposed the AutoEncoder model, which is jointly regularized with the sparsity of gradients in the spatial domain to reconstruct HSI. Wang et al. [24] proposed Hyper-ReconNet model, which improves the reconstruction quality of HSI by jointly training the coding aperture and reconstruction methods. Sun et al. [25] proposed HCS^2-Net model, which utilizes the spatial-spectral attention module to capture the joint spatial-spectral correlation of HSIs to reconstruct HSI. Miao et al. [26] proposed λ-net model, which reconstructs the HSI through a two-stage generative adversarial network. It uses a deeper self-attention U-net in the first stage to obtain the initial reconstruction, and uses another U-net in the first stage to improve the reconstruction. Zhang et al. [27] proposed DEIL model, which mainly utilizes multiple dense residual blocks with residual channel attention to reconstruct HSIs.

However, refs. [22–27] are intended to be applied to coded aperture snapshot spectral imaging (CASSI) system, which first encodes the HSI to obtain measurements, and then leverages an inverse optimization algorithm to reconstruct the original HSI. Therefore, they need to consider the coded aperture problem in the compression measurement process. However, the work in this paper is not for the compression sampling and measurements reconstruction of CASSI system.

Hu et al. [28] proposed the HSI reconstruction algorithm, PRN, based on the residual network, which consists of two residual CNNs. One is a CS measurements reconstruction network, which preserves spectral correlations well. The other is a deblocking network for removing block effect, which is caused by patch-based sampling. Huang et al. [29] proposed a spatial–spectral residual dense network, SSRDNHIR, which consists of two CNNs. One is a residual dense network (RDNHIR) for HSI reconstruction and the other is reconstruction network (SDRN) for spectral disparity reduction.

However, refs. [28, 29] are to independently improve the performance of the reconstruction algorithm, without considering the help of the sampling matrix for image reconstruction, that is, the two are connected through a unified framework to improve the performance of the reconstruction algorithm.

3 | METHODOLOGY

CDS^2-DResN is proposed to complete high-quality HSI reconstruction by connecting to one framework. The overall framework of CDS^2-DResN is shown in Figure 1, where CONV denotes convolution layer and ReLU denotes activation function, including four key components: (1) Block compression sampling part: CCL. (2) Initial reconstruction part: DeConV. (3) Spatial information refinement part: spatial CNN. (4) Spectral information refinement part: spectral CNN.

3.1 | Overview of CDS^2-DResN

To begin with, given a HSI X to obtain the CS measurements Y, CCL is leveraged to block compress sampling the HSI X. CCL without bias or activation function adaptively generates a LMM from training HSI, which allows CS measurements to retain more spatial and spectral feature information. This process can be expressed as:

$$Y = CCL(X). \tag{1}$$

where $CCL(\cdot)$ denotes the block compression sampling process. Secondly, given the CS measurements Y for initial reconstruction, DeConV is leveraged to initially reconstruct HSI. DeConV can infer the mapping matrix of image reconstruction. This process can be expressed as:

$$S = DeConV(Y). \tag{2}$$

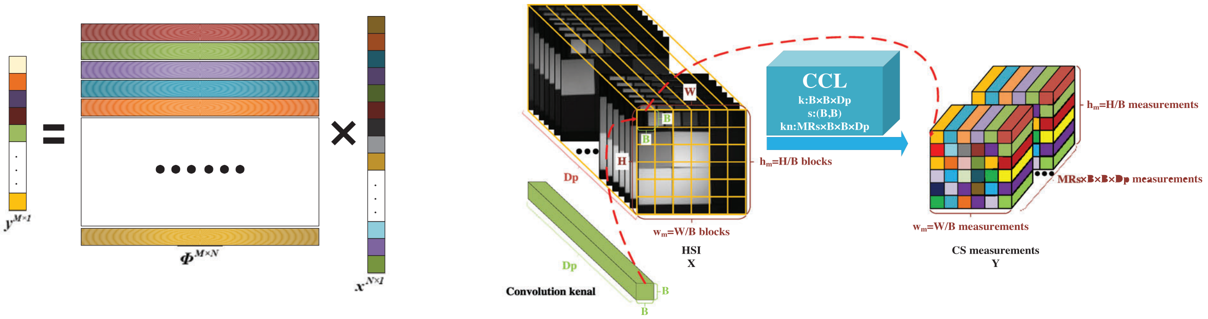
where $DeConV(\cdot)$ denotes the initial reconstruction process. Then, given the initial reconstruction HSI S for spatial information refinement. To further narrow the spatial information difference between original HSI and initial reconstruction HSI, the spatial CNN is leveraged to refine initial reconstruction HSI S. The spatial CNN includes extracting shallow features, extracting spatial information features, and fusion various features. This process can be expressed as:

$$\hat{S} = SpaCNN(S). \tag{3}$$

where $SpaCNN(.)$ denotes spatial information refinement process. Finally, given the spatial refinement HSI $\hat{S}$ for spectral information refinement. To further narrow the spectral information difference between original HSI and spatial refinement HSI, the spectral CNN is leveraged to refine spatial refinement HSI $\hat{S}$. The spectral CNN includes splitting spectral band, combining adjacent spectral bands, extracting features of spectral bands, and fusion various features. This process can be expressed as:

$$\hat{\hat{S}} = SpeCNN(\hat{S}). \tag{4}$$

where $SpeCNN(.)$ denotes spectral information refinement process. Moreover, we jointly train the CCL, DeConV, spatial



(a) The traditional compression sampling process in CS theory.

(b) The block compression sampling process in CDS²-DResN.

FIGURE 2 Illustration of the compression sampling process, where × denotes elementwise multiplication. (a) The traditional compression sampling process in CS theory. (b) The block compression sampling process in CDS²-DResN

CNN and spectral CNN by learning all network parameters Θ in CDS²-DResN. Specifically, the overall framework is trained using the loss function $Loss_{MSE+RTASPeD}$ and all parameters Θ are updated with Equation (5):

$$CNN^{\Theta}_{CDS^2-DResN} = \mathop{\arg\min}_{CNN_{CDS^2-DResN} \in CNN} \max_{\Theta} Loss_{MSE+RTASPeD}(CNN_{CDS^2-DResN}, \Theta). \tag{5}$$

Especially, after the overall framework is jointly trained, the weight value in CCL is the measurement matrix in CS theory. The optimized HSI reconstruction model is composed of DeConV, spatial CNN and spectral CNN.

3.2 | HSI block compression sampling part

The compression sampling process in CS theory is a linear projection process, which linearly projects the original high-dimensional signal to a lower dimension than the original signal. In the traditional compression sampling problem in CS theory, firstly, the image needs to meet the sparse condition, and then the sampling matrix needs to meet the restricted isometry property (RIP). The traditional compression sampling process is shown in Figure 2a. Suppose that $x \in R^{N\times 1}$ is an original high-dimensional signal, $y \in R^{M\times 1}$ is the CS measurements, and $\Phi \in R^{M\times N}$ is a measurement matrix ($M << N$). This process is expressed as:

$$y^{M\times 1} = \Phi^{M\times N} x^{N\times 1}. \tag{6}$$

where M and N denote the dimension size of the CS measurements and original signal, respectively.

The existing sampling matrix are all signal-independent, and do not consider the characteristics of the sampled signal so that more information cannot be retained in measurements. The CNN-based method can solve the compression sampling problem in CS more effectively. Moreover, the LMM, the weight of the learned CCL, can obtain more useful image reconstruction information compared to PMM, which has been verified in Section 4.6 1.

To rely on CNN to implement the compression sampling process in CS, we refer to the compression sampling process in the related work [30] on CS reconstruction, that is, leveraging the CCL to compress and sample HSI. It is worth noting that there is no bias or activation function in CCL. And the weight value of CCL convolution kernel after training is LMM. Figure 2b shows the detailed block compression sampling process in CDS²-DResN. An HSI is divided into $w_m \times h_m$ image blocks of size $B \times B \times Dp$($w_m \times h_m = \frac{W}{B} \times \frac{H}{B}$, W and H are the width and height of the original HSI, respectively. B is the block size of HSI. Dp is the number of HSI channels). Each image block can be denoted as $x^{N\times 1}$ in Figure 2a. Then, the CS measurements $y^{M\times 1}$ of the image block $x^{N\times 1}$ are acquired using a measurement matrix $\Phi^{M\times N}$. Since the number of rows in the measurement matrix $\Phi^{M\times N}$ is M, the size of each convolution kernel in CCL is also M ($M = B \times B \times Dp$), so that each convolution kernel outputs one measurement. Since the number of columns in the measurement matrix $\Phi^{M\times N}$ is N, we need N ($N = MRs \times B \times B \times Dp$, MRs denotes the measurement rates in CS, i.e. the ratio of the data volume of the measurements to the data volume of the original signal) convolution kernels in CCL to obtain N ($N = MRs \times B \times B \times Dp$) measurements. Therefore, the output of each image block from CCL is composed of N ($N = MRs \times B \times B \times Dp$) feature maps.

In conclusion, the block compression sampling process of HSI convolves N ($N = MRs \times B \times B \times Dp$) convolution kernels of size M ($M = B \times B \times Dp$) over HSI using $B \times B$ strides. This process is expressed as:

$$Y = CCL(X, W_{CCL}) = W_{CCL} * X. \tag{7}$$

where $*$ denotes the elementwise convolution. W_{CCL} denotes the weight matrix of the CCL.

Due to our limited computing resources (Titan Xp with memory 24 G), we choose the maximum stride size 16 × 16

TABLE 1 The relationship between $B \times B$ strides and MRs

$B \times B$	MRs
16×16	25%, 10%, 4% and 1%

that the current computing resources can support. Since the number of convolution kernels needs to satisfy the inequalities $MRs \times B \times B \times Dp \geq 1$, the MRs can be any frequency larger than $\frac{1}{7936}$. To avoid the contingency of scene compression sampling at a single MR, MRs will be directly taken as 25%, 10%, 4%, and 1% in the research works [31, 32] of CS. Therefore, the corresponding relationship between $B \times B$ strides and MRs is shown in Table 1. When the data dimension of the image block is 16 and the MRs is 25%, the scale of the LMM is $4 \times 31 \times 16 \times 16$ (The first 4 is obtained by $16 \times 25\%$, the second 31 is the depth of the convolution kernel, and 16×16 is the size of the convolution kernel).

3.3 | CS measurements reconstruction part

The CS theory shows that if the original high-dimensional signal $x \in R^{N\times 1}$ is sparse in a domain Ψ, it is possible for $y \in R^{M\times 1}$ to be restored to $x \in R^{N\times 1}$ accurately. The original signal x can be represented in a sparse domain Ψ, that is,

$$x^{\sim} = \Psi x. \tag{8}$$

where $x^{\sim}$ is a sparse representation of x relative to the sparse domain Ψ.

After the signal $x \in R^{N\times 1}$ is sparsely represented, the simplest formula of CS reconstruction can be expressed as:

$$\min_{\tilde{x}} ||\Psi \tilde{x}||_p,\ s.t.\ y = \Phi \Psi^{-1} \tilde{x}. \tag{9}$$

where the subscript p is usually set to 1 or 0, which represents the sparsity of the vector Ψ. Ψ^{-1} is the inverse sparse transformation.

By solving Equation (9), we can obtain the sparse representation $\tilde{x}$ of x. Therefore, the reconstructed signal can be expressed as:

$$x = \Psi^{-1} \tilde{x}. \tag{10}$$

Since Equations (9) and (10) are over-determined equations, it has no exact solution. To solve this problem, we estimate a reconstruction mapping based on CNN, which minimizes the error between ground HSI and reconstructed HSI. Firstly, we leverage DeConV to solve Equations (9) and (10). Moreover, it is essential to consider the spatial and spectral information of HSI during image reconstruction. Therefore, we introduce the spatial and spectral information refinement network to further optimize Equations (9) and (10).

(1) Initial reconstruction

We leverage DeConV to learn a non-linear function that maps the measurements Y to HSI S. Figure 3 shows the initial reconstruction process in CDS^2-DResN, the output of each measurement block from the DeConV is composed of $B \times B \times Dp$ feature maps. In conclusion, the initial reconstruction process convolves $B \times B \times Dp$ convolution kernels of size $B \times B \times MRsB^2Dp$ over measurements using $B \times B$ strides. The trained DeConV maps the measurements Y into an initial reconstructed HSI S. This process is expressed as:

$$S = DeConV(Y, W_{DeConV}) = W_{DeConV} * Y. \tag{11}$$

where W_{DeConV} denotes the weight parameters for the trained DeConV.

(2) Spatial information refinement

The trained DeConV can only obtain an approximate solution to the ground HSI X. To further narrow the difference between S and X, we design a spatial CNN to predict the spatial information between two HSIs. It is worth noting that spatial CNN referred to the residual dense network for hyperspectral image reconstruction (RDNHIR) network in the related work [26]. Moreover, considering the operation speed and reconstruction quality, we modify the dense residual module to simplify the operation in the dense residual block.

Figure 4 shows the spatial CNN architecture for spatial information refinement. The network consists of three parts: extracting shallow features, extracting spatial features, and fusion various features.

The input of the spatial CNN is the initial reconstructed HSI S. Firstly, we leverage two convolutional layers to extract the shallow features F_{-1} and F_0. Here, the calculation formulas of F_{-1} and F_0 can be expressed as:

$$F_{-1} = W_{-1(3\times3)} * S + B_{-1(3\times3)}. \tag{12}$$

$$F_0 = W_{0(3\times3)} * F_{-1} + B_{0(3\times3)}. \tag{13}$$

where $W_{-1(3\times3)}$ and $B_{-1(3\times3)}$ denote the weight and bias values in the first convolutional layer, respectively. $W_{0(3\times3)}$ and $B_{0(3\times3)}$ denote the weight and bias values in the second convolutional layer, respectively.

Then, we leverage three SpaN units to fully extract the spatial information features F_1, F_2 and F_3. Here, the calculation formulas of F_1, F_2 and F_3 can be expressed as:

$$F_1 = H_{SpaN,1}(F_0). \tag{14}$$

$$F_2 = H_{SpaN,2}(F_1). \tag{15}$$

$$F_3 = H_{SpaN,3}(F_2). \tag{16}$$

where $H_{SpaN,1}(\cdot)$,, $H_{SpaN,2}(\cdot)$ and $H_{SpaN,3}(\cdot)$ denote spatial feature extraction operator of the three SpaN, respectively.

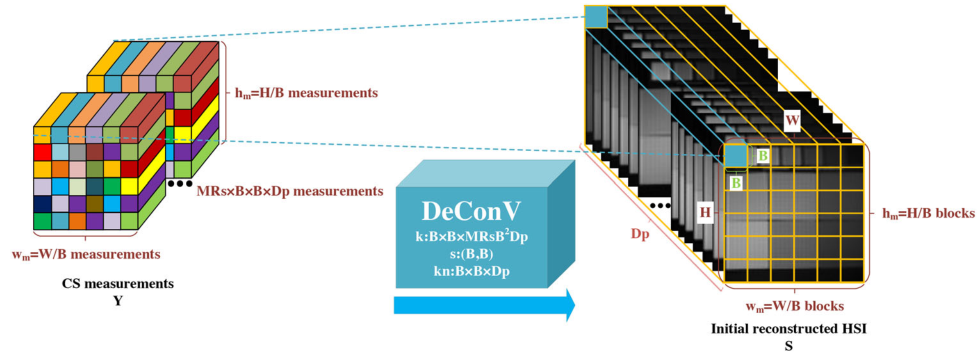


FIGURE 3 Illustration of the initial reconstruction process in CDS^2-DResN

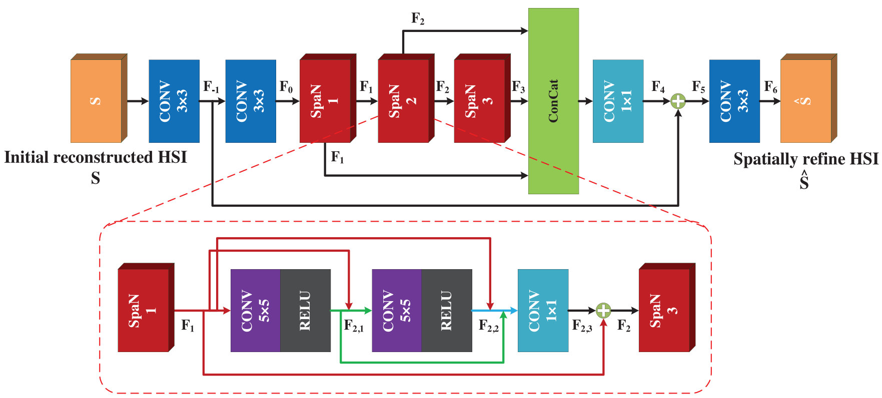


FIGURE 4 Illustration of the spatial CNN architecture for HSI spatial information refinement, where ⊕ denotes elementwise addition, ConCat denotes feature channel concatenate operation

The structure and paraments of the second SpaN unit in Figure 4, which has a memory signal mechanism. Considering the convergence speed and reconstruction accuracy, the spatial CNN adopts three SpaN units, and each SpaN unit contains three convolutional layers. It is worth nothing that the feature maps $F_{2,1}$, $F_{2,2}$,$F_{2,3}$ and F_2 are respectively expressed as:

$$F_{2,1} = \mathrm{cat}(F_1, (\max((W_{1(5\times5)} * F_1 + B_{1(5\times5)}, 0)))). \text{ (17)} F_{2,2}$$

$$= \mathrm{cat}(F_1, F_{2,1}, (\max((W_{2,1(5\times5)} * F_{2,1} + B_{2,1(5\times5)}, 0)))). \quad (18)$$

$$F_{2,3} = W_{2,2(1\times1)} * F_{2,2} + B_{2,2(1\times1)}. \quad (19)$$

$$F_2 = F_1 + F_{2,3}. \quad (20)$$

where $cat(\cdot)$ denotes feature channel concatenate operation. $W_{1(5\times5)}$ and $B_{1(5\times5)}$ denote the weight and bias values in the 5×5 convolutional layer, respectively. $W_{2,1(5\times5)}$ and $B_{2,1(5\times5)}$ denote the weight and bias values in the 5×5 convolutional layer, respectively. $W_{2,2(1\times1)}$ and $B_{2,2(1\times1)}$ denote the weight and bias values in the 1×1 convolutional layer, respectively.

Finally, we fuse the various features to get the spatial information refinement HSI $\hat{S}$. It is worth noting that we leverage the 1×1 convolutional layer to adaptively cascade the feature maps at different levels. Here, the calculation formulas can be expressed as:

$$F_4 = W_{4(1\times1)} * (cat(F_1, F_2, F_3)) + B_{4(1\times1)}. \quad (21)$$

where $W_{4(1\times1)}$ and $B_{4(1\times1)}$ denote the weight and bias values in the 1×1 convolutional layer, respectively. Meanwhile, the global residual feature needs to be utilized to get the final spatial refinement result, and the calculation formulas can be expressed as:

$$\hat{S} = F_6 = W_{5(3\times3)} * (F_4 + F_{-1}) + B_{5(3\times3)}. \quad (22)$$

where $W_{5(3\times3)}$ and $B_{5(3\times3)}$ denote the weight and bias values in the 3×3 convolutional layer, respectively.

Spatial refinement network can better extract multi-layer feature information, which has been verified in Section 4.6 2. In the following, we design a spectral CNN to reduce the spectral error through the spectral correlation between adjacent bands, which can further improve the quality of HSI reconstruction.

(3) Spectral information refinement

The trained spatial CNN can narrow the spatial difference between $\hat{S}$ and X. However, HSI also contains the spectral characteristics of each pixel. To further narrow the spectral difference between $\hat{S}$ and X, we design a spectral CNN to predict the spectral information between two HSIs. It is worth noting that spectral CNN referred to the spectral CNN in the related work [24]. Moreover, since the residual block can effectively retain some information of the previous layers, Therefore, we leverage the residual block to extract features of spectral bands in spectral CNN.

Figure 5 shows the spectral CNN architecture for HSI spectral information refinement. The network consists of

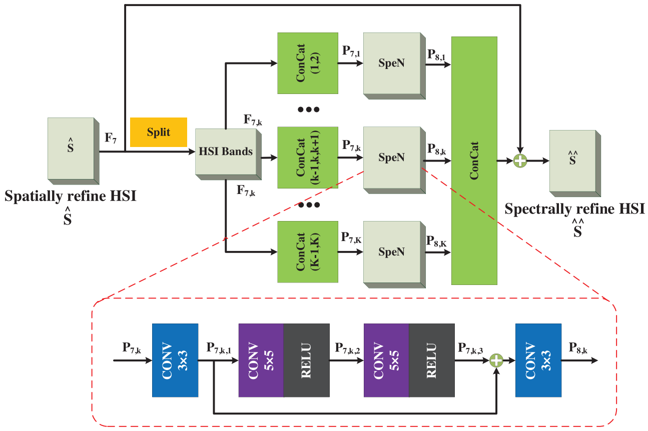


FIGURE 5 Illustration of the spectral CNN architecture for HSI spectral information refinement, where Split denotes slice operator along the spectral dimension

four parts: splitting spectral band, combining adjacent spectral bands, extracting features of spectral bands, and fusion various features.

The input of the spatial CNN is the spatial refinement HSI $\hat{S}$. We first split the $\hat{S}$ into spectral bands $F_{7,k}$ (k $= 1, 2, \ldots, D$p$-1, D$p). Here, spectral bands can be expressed as:

$$F_{7,k} = Split(\hat{S}). \tag{23}$$

where $Split(\cdot)$ denotes the slice operator along the spectral dimension. $F_{7,k}$ (k $= 1, 2, \ldots, D$p$-1, D$p) denotes the $k-$th band in the F_7.

Secondly, we obtain Dp spectral bands $P_{7,k}$ by combining adjacent spectral bands. Two adjacent bands are combined for the first band $P_{7,1}$ and last band $P_{7,Dp}$, and three adjacent bands are combined for all other bands $P_{7,k}$ $(k = 2, 3, \ldots, Dp - 1)$. Here, combined adjacent spectral bands can be expressed as:

$$\begin{aligned} P_{7,1} &= \mathrm{cat}(F_{7,1}, F_{7,2}) \\ P_{7,k} &= \mathrm{cat}(F_{7,k-1}, F_{7,k}, F_{7,k+1}). \\ P_{7,Dp} &= \mathrm{cat}(F_{7,Dp-1}, F_{7,Dp}) \end{aligned} \tag{24}$$

Then, we leverage the designed SpeN to fully extract the features of each spectral band $P_{7,k}$. Here, the calculation formula of $P_{7,k}$ can be expressed as:

$$P_{8,k} = H_{\mathrm{SpeN}}(P_{7,k}). \tag{25}$$

where $H_{\mathrm{SpeN}}(\cdot)$ denote spectral information feature extraction operator of the SpeN.

The structure and paraments of the $k-th$ SpeN unit in Figure 5. It is worth nothing that the feature maps $P_{7,k,1}$, $P_{7,k,2}$, $P_{7,k,3}$, and $P_{8,k}$ are respectively expressed as:

$$P_{7,k,1} = W_{7,k(3\times3)} * P_{7,k} + B_{7,k(3\times3)}. \tag{26}$$

$$P_{7,k,2} = \max(W_{7,k,1(5\times5)} * P_{7,k,1} + B_{7,k,1(5\times5)}, 0). \tag{27}$$

$$P_{7,k,3} = \max(W_{7,k,2(5\times5)} * P_{7,k,2} + B_{7,k,2(5\times5)}, 0). \tag{28}$$

$$P_{8,k} = W_{8,k(3\times3)} * (P_{7,k,1} + P_{7,k,3}) + P_{8,k(3\times3)}. \tag{29}$$

where $W_{7,k(3\times3)}$ and $B_{7,k(3\times3)}$ denote the weight and bias values in the 3×3 convolutional layer, respectively. $W_{7,k,1(5\times5)}$ and $B_{7,k,1(5\times5)}$ denote the weight and bias values in the 5×5 convolutional layer, respectively. $W_{7,k,2(5\times5)}$ and $B_{7,k,2(5\times5)}$ denote the weight and bias values in the 5×5 convolutional layer, respectively. $W_{8,k(3\times3)}$ and $B_{8,k(3\times3)}$ denote the weight and bias values in the 5×5 convolutional layer, respectively.

Finally, we get the final reconstructed HSI $\hat{\hat{S}}$ through global dense feature fusion. Here, the calculation formulas can be expressed as:

$$\hat{\hat{S}} = F_7 + \mathrm{cat}(P_{8,1}, \ldots, P_{8,\mathrm{k}}, \ldots, P_{8,\mathrm{Dp}}). \tag{30}$$

Spectral refinement network can better extract spectral feature information, which has been verified in Section 4.6 2.

3.4 | Joint training of HSI compression sampling part and CS measurements reconstruction part

Traditional CS reconstruction algorithms all leverage a PMM to compress and sample the HSI. However, the PMM ignores the characteristics of the sampled HSI. Therefore, we construct the sampling network, CCL, to adaptively learn a measurement matrix from the training HSIs. Furthermore, we jointly train the CCL, DeConV, spatial CNN, and spectral CNN by learning all the weights and biases in one framework. All parameters in CDS^2-DResN can be expressed as:

$$\Theta = \{\mathrm{W_{CCL}}; \mathrm{W_{DeConV}}; \mathrm{W}_{(s\times s)}; \mathrm{B}_{(s\times s)}\}. \tag{31}$$

where $W_{(s\times s)}$ and $B_{(s\times s)}$ denote the weight and bias values of convolutional layers in the spatial CNN and spectral CNN, respectively.

The loss function of the CS reconstruction network for HSI usually adopts the MSE, which can estimate the model parameter Θ and minimize the loss between the reconstructed HSI $\hat{\hat{S}}$ and the corresponding ground truth HSI X. However, since there are also adjacent spectral differences between the $\hat{\hat{S}}$ and X, we add a regular term of adjacent spectral differences (RTASpeD) based on MSE. In section 4.6 3, we have demonstrated that RTASpeD can significantly improve the reconstruction index SAM. The expression of the loss function is:

$$\begin{aligned} Loss_{MSE+RTASPeD}(\Theta) &= \frac{1}{N}\sum_{i=1}^{N} ||\hat{\hat{S}}_{\mathrm{i}}\ X_i||_2^2 \\ &+ \frac{1}{N}\left(\sum_{i=1}^{N}\lambda\sum_{j=1}^{Dp-1} ||\hat{\hat{S}}_{i,j+1}\ X_{i,j+1} + X_{i,j} - \hat{\hat{S}}_{i,j}||_1\right). \end{aligned} \tag{32}$$

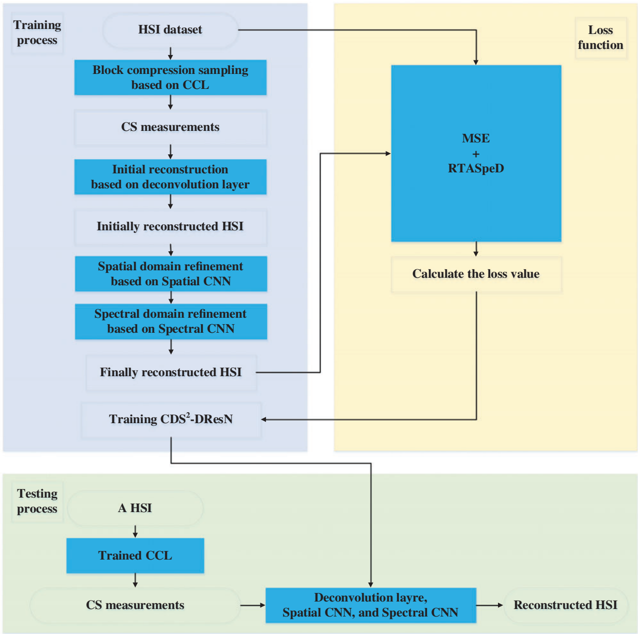


FIGURE 6 Illustration of the flowchart of CDS^2-DResN

where N is the number of training dataset. $S_{i,j}$ is the $j-th$ band image of the $i-th$ ground truth HSI. $\hat{\hat{S}}_{i,j}$ is the $j-th$ band image of the $i-th$ reconstructed HSI. λ is a hyperparameter, we take the value 0.25 [29].

3.5 | CS procedure

The process of HCS is shown in Figure 6. In the testing process, given an HSI, we first leverage the trained CCL to directly perform convolution coding on HSI to obtain CS measurements, which leverages CNN to perform the block compression sampling process in CS. Then, the CS measurements are fed into the DeConV to output the initial reconstructed HSI, which is finally processed with spatial CNN and spectral CNN to refine the input initial reconstructed HSI in turn. It is worth noting that the DeConV, spatial CNN, and spectral CNN constitute the reconstruction model of HCS.

4 | EXPERIMENT

4.1 | Dataset

We evaluate our model on two HSI datasets, the CAVE dataset [33] and Harvard dataset [34]. The CAVE dataset contains 31 HSIs. Especially, each image has 31 bands, ranging from 400 to 700 nm, with an interval of 10 nm and the image size of each band is 512 × 512 pixels. The Harvard dataset contains 50 HSIs. Especially, each image has 31 bands, ranging from 420 to 720 nm, with an interval of 10 nm and the image size of each band is 1040 × 1392 pixels. Table 2 shows the dataset partitioning for the training set, verification set and test set.

TABLE 2 Dataset partitioning

Dataset	Total number of HSIs	Dataset division	Number of HSIs	Number of block images
CAVE	31	Training set	22	22,528
		Validation set	3	3072
		Test set	6	6144
Harvard	50	Training set	39	39,936
		Validation set	5	5120
		Test set	6	6144

TABLE 3 Experimental environment

System	Ubuntu 18.04
RAM	64.0 GB
CPU	3.60 GHz Intel(R) Core(TM) i7-6850K
GPU	TITAN X, memory 24 G
Framework	Keras (backed as TensorFlow)

TABLE 4 CDS^2-DResN training parameters

Initialize convolution	Gaussian distribution with standard deviation of 0.001 [31]
Optimizer	Adam [35]
Learn rate	10^{-4} [29]
Batch size	2

4.2 | Implementation details

To effectively train and test CDS^2-DResN, we leverage the experimental environment in Table 3. The experimental environment dramatically speeds up the calculation speed of the CNN. Since the size of the image data in the experiment is different, it is uniformly adjusted to 512 × 512 before inputting the model.

4.3 | Implementation parameters

Table 4 shows the training parameter setting for CDS^2-DResN training. It is worth noting that we refer to the research work [29] in the same field to get a reasonable learning rate 10^{-4}. In particular, considering the 24G limitation of GPU memory, we set the batch size to 2.

4.4 | Evaluation metrics

Peak signal-to-noise ratio (PSNR) [36] and structural similarity (SSIM) [37] are calculated on each two-dimensional spatial image, which show the spatial fidelity between the reconstructed HSI and the ground truth. The higher the PSNR and

TABLE 5 Performance comparison on CAVE dataset of different algorithms

Algorithm	MR = 25%	MR = 10%	MR = 4%	MR = 1%
	23.3086	22.8070	22.4747	21.7138
PRN	0.6991	0.6453	0.6081	0.4745
	0.5337	0.5729	0.5693	0.6009
	35.9561	33.7410	30.1976	27.0445
SSRDNHIR	0.9398	0.9185	0.7388	0.6886
	0.1304	0.1475	0.2946	0.3255
	36.6248	**34.5133**	**30.6564**	**27.1520**
CDS^2-DResN	**0.9571**	**0.9360**	**0.8788**	**0.7996**
	0.1039	**0.1205**	**0.1597**	**0.2130**

TABLE 6 Performance comparison on Harvard dataset of different algorithms

Algorithm	MR = 25%	MR = 10%	MR = 4%	MR = 1%
	24.9320	23.7021	22.5562	21.7529
PRN	0.6633	0.5911	0.5449	0.5182
	0.2143	0.2150	0.2112	0.2111
	32.6837	31.0404	29.2457	27.0140
SSRDNHIR	0.8527	0.8331	0.8059	0.7666
	0.1248	0.1341	0.1433	0.1483
	33.0465	**31.0818**	**29.7624**	**27.3830**
CDS^2-DResN	**0.8529**	**0.8335**	**0.8132**	**0.7716**
	0.1225	**0.1326**	**0.1372**	**0.1404**

SSIM values, the better the reconstruction performance of the algorithm. The PSNR and SSIM are respectively expressed as:

$$PSNR(X, \hat{\hat{S}}) = 20 \times \lg\left(\frac{255}{\sqrt{MSE(S, \hat{\hat{S}})}}\right).. \quad (33)$$

$$SSIM(X, \hat{\hat{S}}) = \frac{(2\mu_X \mu_{\hat{\hat{S}}} + C_1)(2\sigma_{X\hat{\hat{S}}} + C_2)}{(\mu_X^2 + \mu_{\hat{\hat{S}}}^2 + C_1)(\sigma_X^2 + \sigma_{\hat{\hat{S}}}^2 + C_2)}. \quad (34)$$

where X denotes the ground truth HSI, $\hat{\hat{S}}$ denotes the reconstructed HSI, $MSE(X, \hat{\hat{S}})$ denotes the mean square error between the X and $\hat{\hat{S}}$, μ_X and $\mu_{\hat{\hat{S}}}$ denote the mean values of the X and $\hat{\hat{S}}$, σ_X and $\sigma_{\hat{\hat{S}}}$ denote the variances of the$X$ and $\hat{\hat{S}}$, and $\sigma_{\hat{\hat{S}}}$ denotes the covariance of the X and $\hat{\hat{S}}$.

Spectral angle mapping (SAM) [38] is calculated on a one-dimensional spectral vector, which shows the spectral fidelity between the reconstructed HSI and the ground truth. The smaller the SAM value, the better the reconstruction performance of the algorithm. The SAM is expressed as:

$$SAM(x, \hat{\hat{S}}) = \cos^{-1}\left(\frac{\hat{\hat{S}}^T x}{||\hat{\hat{s}}||||x||}\right). \quad (35)$$

where $\hat{\hat{s}}$ denotes the reconstructed HSI spectrum vector and x denotes the ground truth HSI spectrum vector.

4.5 | Comparison with state-of-the-art methods

In this section, to verify the performance of the proposed algorithm, we compare it with the best CNN-based HCS algorithm, that is, PRN [28] and SSRDNHIR [29]. The results are summarized in Tables 5 and 6, where the best results are highlighted in bold. It is worth noting that the results are the average of the evaluation metrics of all reconstructed HSIs in the test set. As shown in Table 5, our CDS^2-DResN outperforms PNR and SSRDNHIR on CAVE dataset. The PSNR, SSIM, and SSIM values of the CDS^2-DResN are higher than PNR and SSRDNHIR for the cases that MRs = 0.25, 0.10, 0.04 and 0.01. As shown in Table 6, Our CDS^2-DResN outperforms PNR and SSRDNHIR on Harvard dataset. The PSNR, SSIM and SSIM values of the CDS^2-DResN are also higher than PNR and SSRDNHIR for the cases that MRs = 0.25, 0.10, 0.04 and 0.01.

To verify the visual quality of the proposed algorithm, Figures 7–10 show a representative reconstructed HSI from the test set on the CAVE and Harvard datasets. Figure 7 shows the 16th band image visualization results of an HSI in CAVE test set. Figure 8 shows the 30th band image visualization results of an HSI in Harvard test set. Figure 9 shows the pseudo colour image visualization results of three channels of HSI in the CAVE test set (the three channels are the 25th, 15th and 5th channels respectively). Figure 10 shows the pseudo colour image visualization results of three channels of HSI in the Harvard test set (the three channels are the 25th, 15th and 5th channels respectively). A partial magnification of the reconstructed image is shown in the upper left corner. As shown in Figures 7–10, the reconstructed HSIs of CDS2-DResN have no block effect, and CDS2-DResN has more HSI details compared to PNR and SSRDNHIR at MRs of 0.25, 0.10, 0.04 and 0.01.

5 | DISCUSSION

5.1 | Analysis of PMM or LMM

In this subsection, we analyse the properties of the LMM and PMM in the time and frequency domains. Figure 11 provides the visualize the results of LMM and PMM on CAVE dataset for CDS2-DResN. We select four rows from each of LMM and PMM for visualization, which are the 1th, 4th, 8th and 16th row, respectively. To obtain a better visual effect, the spatial visualization is the result of multiplying each value by 255 to the base 10 logarithm and frequency visualization is the result

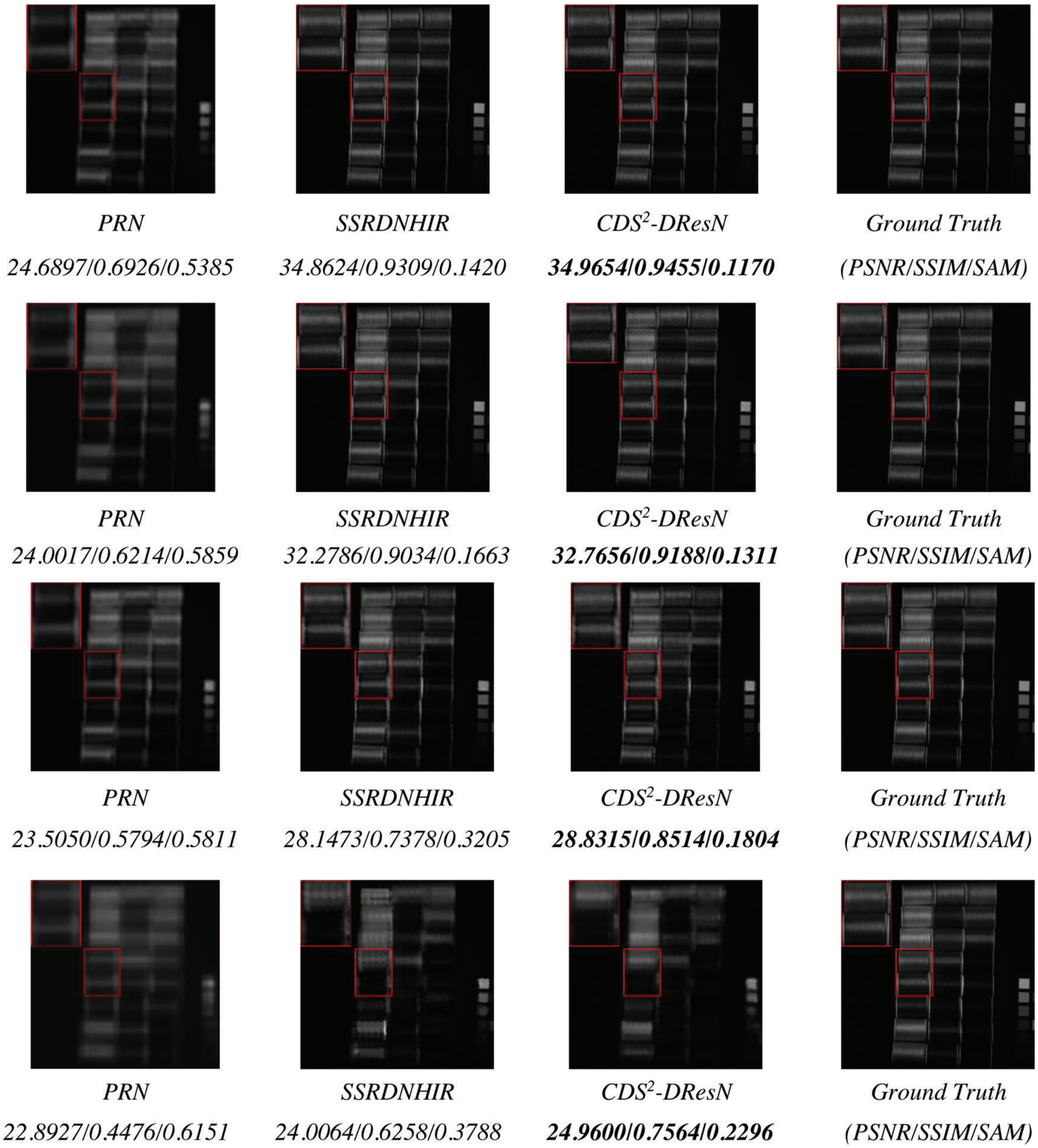


FIGURE 7 Visual quality comparison on CAVE test set (the 16th band image of the HSI)

of the Fourier transform of each row of the measurement matrix. Especially, Figure 11a shows the spatial domain and frequency domain of LMM when MRs = 25%, Figure 11b shows the spatial domain and frequency domain of LMM when MRs = 4%, Figure 11c shows the spatial domain and frequency domain of PMM (Orthogonal Random Gaussian Matrix) when MRs = 25%, and Figure 11d shows the spatial domain and frequency domain of PMM (Orthogonal Random Gaussian Matrix) when MRs = 4%.

From the spatial domain of LMM and PMM, we can find that the LMM is anisotropic, that is, LMM has stronger content adaptive ability. From the frequency domain of LMM and PMM, we can find that the frequency of the PMM is randomly distributed, while the LMM pays more attention to the low-frequency information of the image, which can obtain more useful image reconstruction information. Moreover, the frequency distribution of the LMM at MRs = 4% is narrower than that of the LMM at MRs = 25%. As the MRs increases, the LMM gradually pays attention to more high-frequency information, see Figure 11.

5.2 | Analysis of ablation studies

In this subsection, we conduct ablation studies on the model on the CAVE dataset to verify the effect of spatial CNN and spectral CNN in CDS^2-DResN. The results are summarized in Tables 7–10, with the best results highlighted in bold.

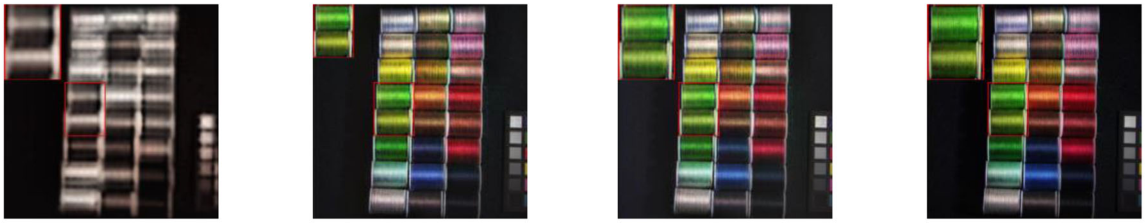
FIGURE 8 Visual quality comparison on Harvard test set (the 30th band image of the HSI)

To study the effect of spatial CNN, we conduct ablation experiments on spatial CNN, and the corresponding experimental results are shown in the first and third rows of Tables 7–10. It can be seen that the values of PSNR and SSIM both increase and the value of SAM decreases when the space CNN is added. Therefore, spatial CNN can achieve better detection accuracy. This is because DRBs in spatial CNN can extract multi-layer feature information, thereby improving the image quality of HCS.

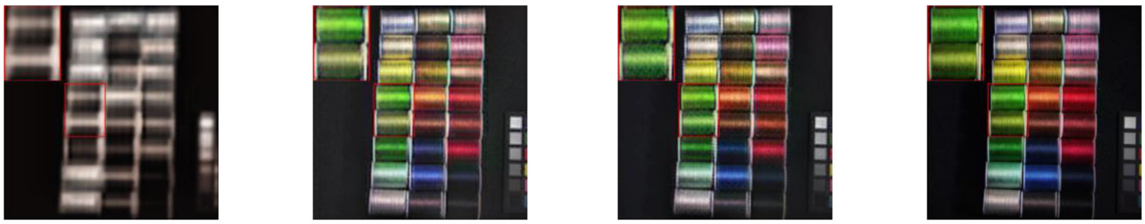
To study the performance of spectral CNN, we conduct ablation experiments on spectral CNN, and the corresponding experimental results are shown in the second and third rows of Tables 7–10. It can be seen that the values of PSNR and SSIM both increase and the value of SAM decreases when the spectral CNN is added. Therefore, spectral CNN can achieve better detection accuracy. This is because the spectral extraction of each spectral band can fully extract the spectral feature information of HSI, thereby improving the image reconstruction quality of HCS.

5.3 | Analysis of MSE or MSE-RTASpeD

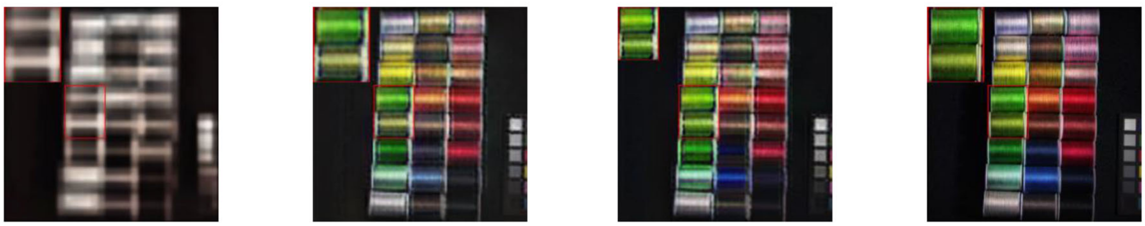
In this subsection, we compare the effects of the MSE and MSE-RTASpeD loss function on the reconstruction quality of HSIs. The results are summarized in Table 11, with the



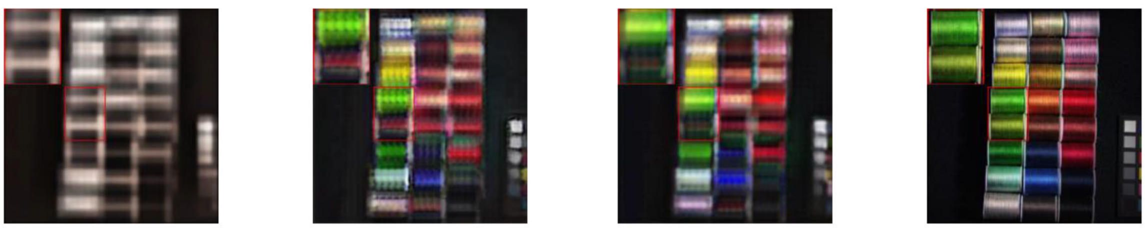
PRN	SSRDNHIR	CDS²-DResN	Ground Truth
24.6897/0.6926/0.5385	34.8624/0.9309/0.1420	**34.9654/0.9455/0.1170**	(PSNR/SSIM/SAM)



PRN	SSRDNHIR	CDS²-DResN	Ground Truth
24.0017/0.6214/0.5859	32.2786/0.9034/0.1663	**32.7656/0.9188/0.1311**	(PSNR/SSIM/SAM)



PRN	SSRDNHIR	CDS²-DResN	Ground Truth
23.5050/0.5794/0.5811	28.1473/0.7378/0.3205	**28.8315/0.8514/0.1804**	(PSNR/SSIM/SAM)



PRN	SSRDNHIR	CDS²-DResN	Ground Truth
22.8927/0.4476/0.6151	24.0064/0.6258/0.3788	**24.9600/0.7564/0.2296**	(PSNR/SSIM/SAM)

FIGURE 9 Visual quality comparison on CAVE test set (Pseudo colour with 25th,15th and 5th channels)

best results highlighted in bold. It can be seen that RTASpeD can improve three reconstruction indicators, especially for the improvement of SAM indicator.

The MSE loss function mainly considers the spatial information between the reconstructed HSI and the ground truth HSI, that is, does not fully consider the adjacent spectral differences between the two. However, the reconstruction method with MSE-RTASpeD loss function solves this problem, which adds an RTASpeD based on MSE. Therefore, it considers both the spatial information between the reconstructed HSI and the ground truth HSI as well as the adjacent spectral differences between the two, which makes the difference between the reconstructed HSI and the ground truth HSI smaller.

5.4 | Analysis of computational complexity

In this subsection, we analyse the computational complexity of the model. The results are summarized in Table 12. It can be seen that the parameter size of the model and the HSI recon-

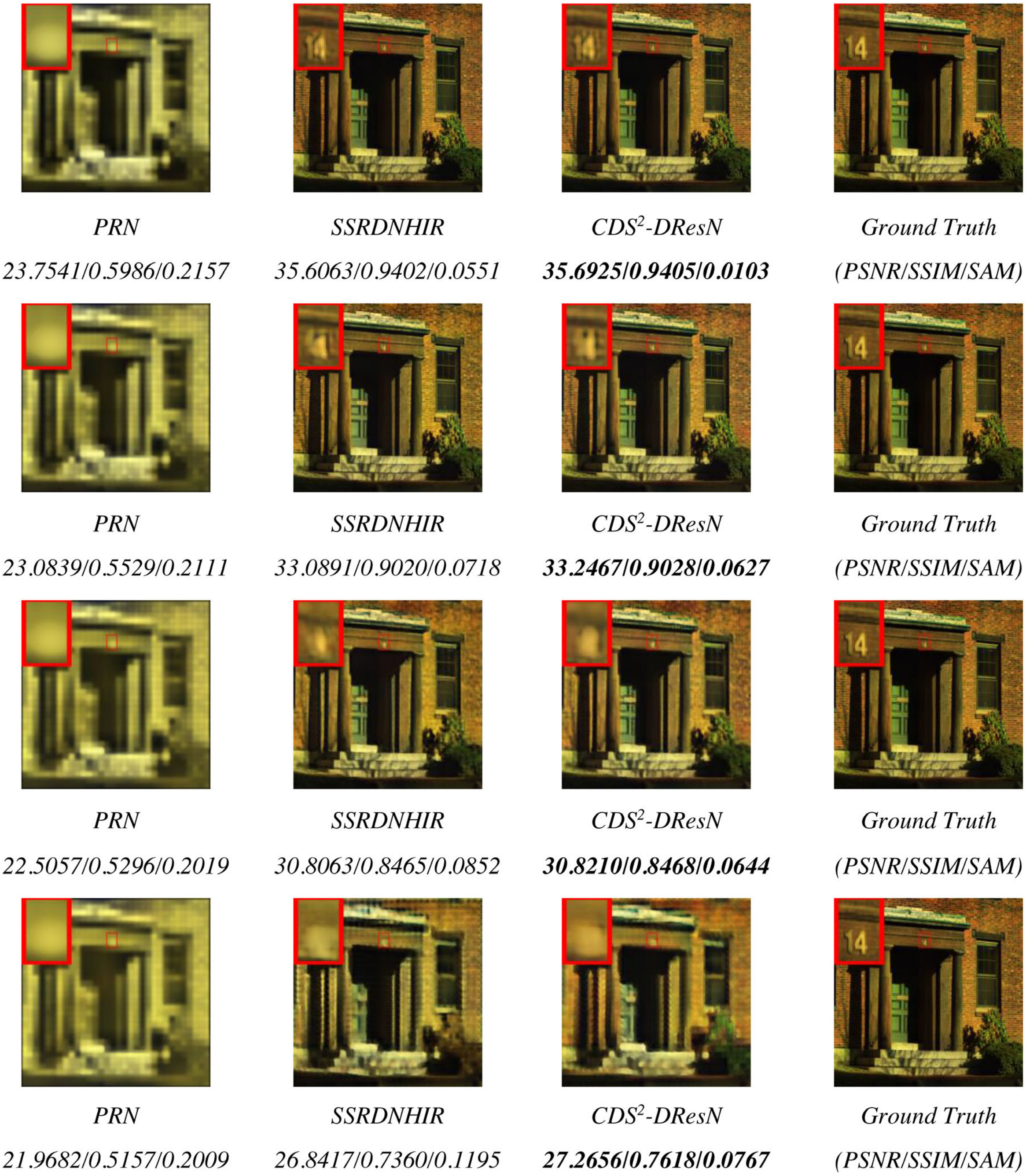


FIGURE 10 Visual quality comparison on Harvard test set (Pseudo colour with 25th,15th and 5th channels)

struction time are decreasing with the decrease of MRs. This is because the data volume of HCS measurements is decreasing. However, the reconstruction time of the model is relatively long under the four MRs. This is because the space-spectrum joint refinement network has a complex network structure.

6 | CONCLUSION

In this paper, we leverage the self-learning ability of CNN to connect HSI compression sampling and measurements reconstruction are connected in one framework, which can improve HSI reconstruction quality. From compression sampling part, the entities of the compression sampling matrix in CS are treated as the CCL weights for optimization training. From measurements reconstruction part, the DeConV is used to initially reconstruct HSI. Additionally, spatial CNN and spectral CNN are designed to refine spatial information and spectral information in the initial HSI, respectively. The experimental results show that CDS^2-DResN has excellent HCS reconstruction effect in terms of quantitative index and visual quality.

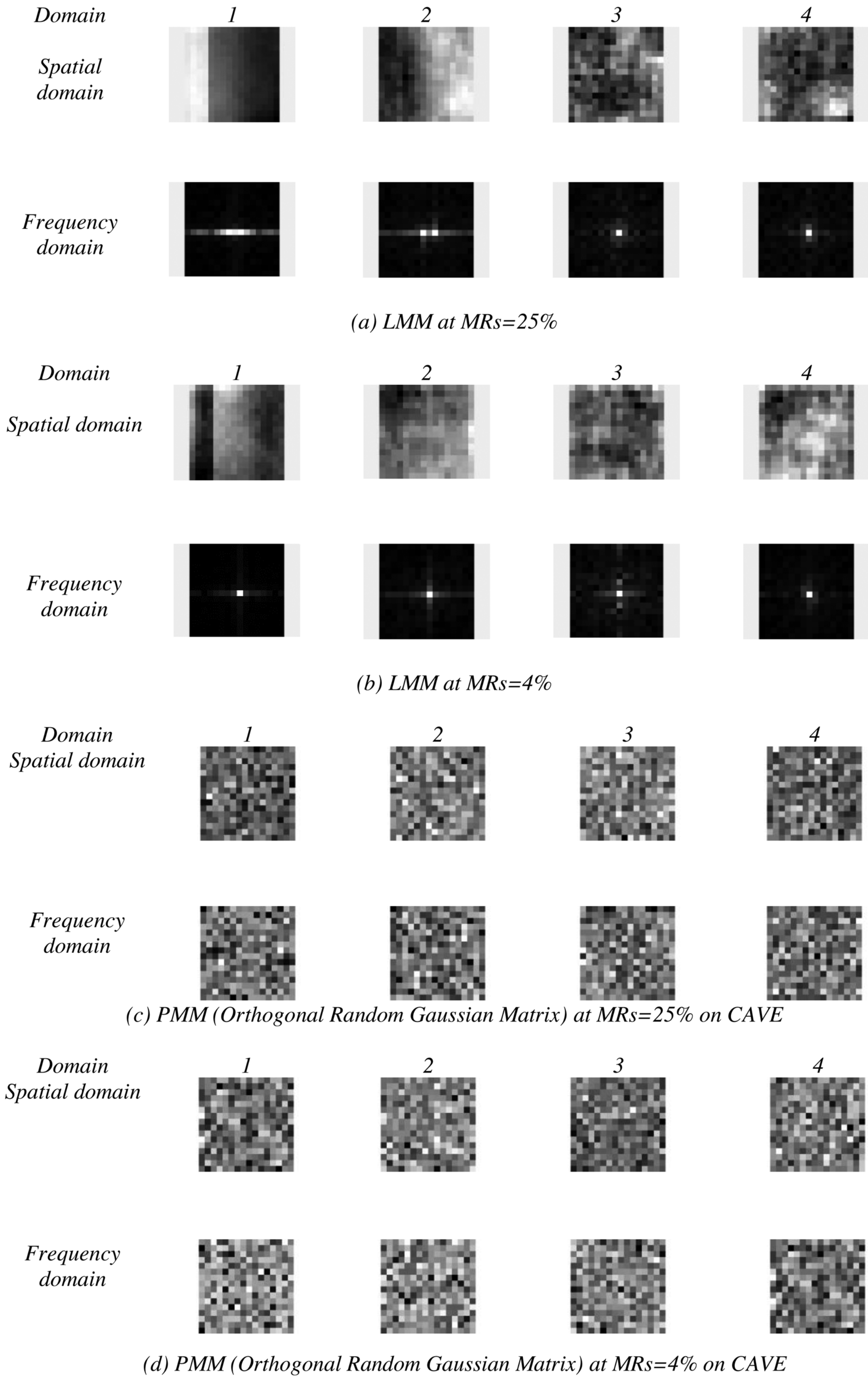


FIGURE 11 The time domain and frequency domain of the measurement matrix

TABLE 7 Reconstruction results with different modification models on the CAVE dataset at MRs = 25%

Model	Spatial CNN	Spectral CNN	PSNR	SSIM	SAM
CDS^2-DResN	×	√	30.0355	0.8460	0.1801
	√	×	34.3008	0.9149	0.1581
	√	√	**36.6248**	**0.9571**	**0.1039**

TABLE 8 Reconstruction results with different modification models on the CAVE dataset at MRs = 10%

Model	Spatial CNN	Spectral CNN	PSNR	SSIM	SAM
CDS^2-DResN	×	√	30.2947	0.8685	0.1546
	√	×	33.4419	0.9103	0.1501
	√	√	**34.5133**	**0.9360**	**0.1205**

TABLE 9 Reconstruction results with different modification models on the CAVE dataset at MRs = 4%

Model	Spatial CNN	Spectral CNN	PSNR	SSIM	SAM
CDS^2-DResN	×	√	27.3378	0.8417	0.2544
	√	×	29.9044	0.8004	0.2529
	√	√	**30.6564**	**0.8788**	**0.1597**

TABLE 10 Reconstruction results with different modification models on the CAVE dataset at MRs = 1%

Model	Spatial CNN	Spectral CNN	PSNR	SSIM	SAM
CDS^2-DResN	×	√	24.4720	0.7389	0.2771
	√	×	25.9110	0.7515	0.2801
	√	√	**27.1520**	**0.7996**	**0.2130**

TABLE 11 Performance comparison on CAVE dataset of different algorithms

Algorithm	MR = 25%	MR = 10%	MR = 4%	MR = 1%
	36.5479	33.6902	30.3275	26.5537
CDS^2-DResN+MSE	0.9521	0.9273	0.8773	0.7809
	0.1294	0.1426	0.2313	0.2759
	36.6248	**34.5133**	**30.6564**	**27.1520**
CDS^2-DResN +MSE+RTASpeD	**0.9571**	**0.9360**	**0.8788**	**0.7996**
	0.1039	**0.1205**	**0.1597**	**0.2130**

TABLE 12 The parameter size of the model and GPU runtime of reconstruction results on the CAVE dataset

MRs	25%	10%	4%	1%
Total params	2,657,023	2,053,887	1,799,935	1,688,831
Time (s)	1.1053	1.0563	1.0442	1.0266

ACKNOWLEDGEMENTS

This work was supported in part by the CIOMP-Fudan University Joint Fund under Grant No.Y9S333T190, and Jilin Provincial Institute Cooperation Program Fund under Grant No.2020SYHZ0031.

CONFLICT OF INTEREST

The authors declare that they do not have any commercial or associative interest that represents a conflict of interest in connection with the work.

AUTHOR CONTRIBUTIONS

S.X.: Methodology; Writing – original draft. Y.Z.: Funding acquisition; Resources. X.C.: Project administration; Writing – review & editing

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in the CAVE dataset at DOI: 10.1109/TIP.2010.2046811, 'Generalized Assorted Pixel Camera: Postcapture Control of Resolution, Dynamic Range, and Spectrum', and in the Harvard dataset at DOI: 10.1109/CVPR.2011.5995660, 'Statistics of Real-World Hyperspectral Images'.

REFERENCES

1. Wang, J., Jiao, L.: Application of a homogenous patch mean kernel with within-class collaborative representation for hyperspectral imagery classification. Remote Sens. Lett. 8(1), 11–20 (2017)
2. Zhang, Y., He, K., Dong, Y., et al.: Joint Sparse Representation and Multitask Learning for Hyperspectral Anomaly Detection. *IGARSS 2020–2020 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, pp. 2424–2427 (2020)
3. Hu, X., Zhong, Y., Luo, C., Wang, X.: WHU-Hi: Uav-borne hyperspectral with high spatial resolution (H2) benchmark datasets for hyperspectral image classification. Remote Sens. Environ. 250, 112012 (2020)
4. Xu, Y., Wu, Z., Li, J., et al.: Anomaly detection in hyperspectral images based on low-rank and sparse representation. IEEE Trans. Geosci. Electron. 54(4), 1990–2000 (2015)
5. Hero, A., Nyquist Sampling Theorem, University of Michigan. (2006)
6. Donoho, D.L.: Compressed sensing. IEEE Trans. Inf. Theory 52(4), 1289–1306 (2006)
7. Baraniuk, R. A lecture on compressive sensing. IEEE Signal processing magazine 24(4), 1–9 (2007)
8. Coluccia, G., Kuiteing, S.K., Abrardo, A., et al.: Progressive compressed sensing and reconstruction of multidimensional signals using hybrid transform/prediction sparsity model. IEEE Journal on Emerging and Selected Topics in Circuits and Systems 2(3), 340–352 (2012)
9. Kuiteing, S.K., Coluccia, G., Barducci, A., Barni, M., Magli, E.: Compressive hyperspectral imaging using progressive total variation. In: *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, pp. 7794–7798. (2014)
10. Lv, P., Liu, H., Chen, R., Wang, Y.: Fast OMP reconstruction for compressive hyperspectral imaging using joint spatial-spectral sparsity model. In: 10th International Conference on Information Optics and Photonics. (2018)
11. Gogna, A., Shukla, A., Agarwal, H.K., Majumdar, A.: Split Bregman algorithms for sparse /joint-sparse and low-rank signal recovery: Application in compressive hyperspectral imaging. In: *IEEE International Conference on Image Processing*. (2015)
12. Jia, Y., Feng, Y., Wang, Z.: Reconstructing hyperspectral images from compressive sensors via exploiting multiple priors. Spectrosc. Lett. 48(1), 22–26 (2015)

13. Zhang, X., Zhang, X.: Compressive hyperspectral imaging with spatial and spectral priors. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 11(11), 4156–4169 (2018)
14. Golbabaee, M., Vandergheynst, P.: Compressive source separation for hyperspectral imaging. (2013)
15. Bioucas-Dias, J.M., Plaza, A., Martin, G.: Hyca: A new technique for hyperspectral compressive sensing. IEEE Trans. Geosci. Remote Sens. 53(5), 2819–2831 (2015)
16. Martin, G., Bioucas-Dias, J.M., Plaza, A.: B-Hyca: Blind hyperspectral compressive sensing. In: *IGARSS 2015 - 2015 IEEE International Geoscience and Remote Sensing Symposium*. (2015)
17. Martín, G., Bioucas-Dias, J.M.: Hyperspectral blind reconstruction from random spectral projections. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9(6), 2390–2399 (2016)
18. Caiafa, C.F., Cichocki, A.: Computing sparse representations of multidimensional signals using Kronecker bases. Neural Comput. 25(1), 186 (2013)
19. Qiu, W., Zhou, J., Zhao, H.Z., Fu, Q.: Fast sparse reconstruction algorithm for multidimensional signals. Electron. Lett. 50(22), 1583–1585 (2014)
20. Zhao, R., Wang, Q., Shen, Y., Li, J.: Multidimensional dictionary learning algorithm for compressive sensing-based hyperspectral imaging. J. Electron. Imaging 25(6), 063013 (2016)
21. Wang, M., Yang, S., Jiao, L., Bin, P., Jin, L.: Compressive hyperspectral imaging via sparse tensor and nonlinear compressed sensing. IEEE Trans. Geosci. Remote Sens.. 53(11), 5943–5957 (2015)
22. Xiong, Z.W., Shi, Z., Li, H.Q., Wang, L.Z., Liu, D., Wu, F.: Hscnn: Cnn-based hyperspectral image recovery from spectrally undersampled projections. In: 2017 IEEE International Conference on Computer Vision Workshops. (2017)
23. Choi, I., Jeon, D.S., Nam, G., Gutierrez, D., Kim, M.H.: High-quality hyperspectral reconstruction using a spectral prior. ACM Trans. Graph 36(6), 13 (2017)
24. Wang, L.Z., Zhang, T., Fu, Y., Huang, H.: Hyperreconnet: Joint coded aperture optimization and image reconstruction for compressive hyperspectral imaging. IEEE Trans. Image Process. 28(5), 2257–2270 (2019)
25. Sun, Y., Yang, Y., Liu, Q., et al.: Unsupervised spatial-spectral network learning for hyperspectral compressive snapshot reconstruction. IEEE Trans. Geosci. Electron. 60, 1–14 (2021)
26. Miao, X., Yuan, X., Pu, Y., Athitsos, V.: Lambda-Net: Reconstruct hyperspectral images from a snapshot measurement. In: *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. (2019)
27. Zhang, T., Fu, Y., Wang, L., Huang, H.: Hyperspectral image reconstruction using deep external and internal learning. In: *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. (2020)
28. Hu, X., Xu, Y., Wei, Z., Liu, H., Qian, L.: Patch-based residual networks for compressively sensed hyperspectral images restruction. In: *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*. (2018)
29. Huang, W., Xu, Y., Hu, X., Wei, Z.: Compressive hyperspectral image reconstruction based on spatial–spectral residual dense network. IEEE Geosci. Remote Sens. Lett. 17(5), 884–888 (2020)
30. Shi, W., Jiang, F., Liu, S., Zhao, D.: Image compressed sensing using convolutional neural network. IEEE Transactions on Image Processing 29, 375–388 (2019)
31. Yao, H.T., Dai, F., Zhang, S.L., Zhang, Y.D., Tian, Q., Xu, C.S.: Dr2-Net: Deep residual reconstruction network for image compressive sensing. Neurocomputing. 359, 483–493 (2019)
32. Kulkarni, K., Lohit, S., Turaga, P., Kerviche, R., Ashok, A.: Reconnet: Non-iterative reconstruction of images from compressively sensed measurements. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. (2016)
33. 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. 19(9), 2241–2253 (2010)
34. Chakrabarti, A., Zickler, T.: Statistics of real-world hyperspectral images. In: *IEEE Conference on Computer Vision and Pattern Recognition*. (2011)
35. Kingma, D.P., Ba, J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, (2014)
36. Korhonen, J., You, J.: Peak signal-to-noise ratio revisited: Is simple beautiful? In: Fourth International Workshop on Quality of Multimedia Experience. (2012)
37. Wang, Z., Bovik, A.C., Sheikh, H.R., et al. Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing 13(4), 600–612 (2004)
38. Kruse, F.A., Lefkoff, A.B., Boardman, J.W., et al. The spectral image processing system (SIPS)interactive visualization and analysis of imaging spectrometer data. Remote sensing of environment 44(2–3), 145–163 (1993)

How to cite this article: Xiao, S., Zhang, Y., Chang, X., Xu, J.: Compressive sensing reconstruction of hyperspectral images based on codec space-spectrum joint dense residual network. IET Image Process. 1–16 (2022). https://doi.org/10.1049/ipr2.12682