

## 结合注意力机制的双流卷积自编码高光谱解混方法

苏晓通, 郭宝峰\*, 尤靖云, 吴文豪, 许张弛

杭州电子科技大学自动化学院, 浙江 杭州 310018

**摘要** 针对基于卷积自编码进行空-谱联合的高光谱解混方法中, 过度引入像元光谱之间的空间相关性导致丰度过于平滑的现象, 提出一种结合注意力机制的双流卷积自编码高光谱解混方法(DSCU-Net)。首先, 利用双流卷积网络分别提取高光谱图像的空间特征和光谱特征; 其次, 为了确保空间特征和光谱特征之间的平衡性, 引入通道注意力机制对提取到的空间特征进行重加权, 并对光谱特征和重加权后的空间特征进行融合; 最后, 使用融合后的特征进行高光谱图像重构, 并将重构结果送入解混网络的主干网络中进行光谱解混。通过最小化两次重构误差进行解混网络的训练。为了验证所提方法的性能, 在两个真实数据集上进行实验, 并对复杂场景下算法的性能表现进行分析。结果表明, DSCU-Net 能够有效减少过度引入空间相关性造成丰度过于平滑的现象, 具有更好的解混性能。

**关键词** 遥感; 高光谱解混; 卷积自编码器; 通道注意力机制; 双流结构

中图分类号 TP751.1

文献标志码 A

DOI: 10.3788/LOP231022

## Dual-Stream Convolutional Autoencoding Network for Hyperspectral Unmixing using Attention Mechanism

Su Xiaotong, Guo Baofeng\*, You Jingyun, Wu Wenhao, Xu Zhangchi

School of Automation, Hangzhou Dianzi University, Hangzhou 310018, Zhejiang, China

**Abstract** In this paper, a dual-stream convolutional autoencoding network for hyperspectral unmixing with attention mechanism (DSCU-Net) is proposed to address the issue of excessively smooth abundance maps caused by excessive incorporation of spatial correlations during pixel spectra in hyperspectral unmixing using a convolution-based autoencoding network. First, the spatial and spectral features of the hyperspectral images are extracted using a dual-stream convolution network. Second, the extracted spatial features are reweighed using a channel attention mechanism and fused with the spectral features to ensure a balance between the spatial and spectral features. Finally, the fusion features are used to reconstruct the hyperspectral image. Furthermore, these features are sent to the backbone in the unmixing network for hyperspectral unmixing. The entire unmixing network is trained by minimizing the two reconstruction errors. Additionally, experiments were conducted on two real datasets to evaluate the performance of the proposed method. The performance of the methods was also analyzed in complex scenarios. The results show that the proposed DSCU-Net can effectively overcome the fuzziness of abundance details because of the excessive introduction of spatial correlation. Moreover, the proposed method has a better unmixing performance.

**Key words** remote sensing; hyperspectral unmixing; convolutional autoencoder; channel attention mechanism; dual-stream structure

## 1 引言

高光谱遥感影像具有超高的光谱分辨率(一般为 10~20 nm), 能够提供丰富的地表信息<sup>[1]</sup>, 目前已经广泛应用于植被遥感监测<sup>[2]</sup>、模式识别、矿物勘测、图像分类<sup>[3-4]</sup>等领域。然而高光谱成像光谱仪在扫描过程

中需要在空间分辨率与信噪比之间进行权衡, 因此高光谱图像一般有着较低的空间分辨率以达到可以接受的信噪比<sup>[5]</sup>。较低的空间分辨率意味着在高光谱图像中某些像元可能是由几种不同的纯物质组合而成的, 即混合像元<sup>[6]</sup>。混合像元的存在使得高光谱图像只能展示模糊的类别信息, 严重限制了高光谱图像的定量

收稿日期: 2023-04-03; 修回日期: 2023-06-07; 录用日期: 2023-07-24; 网络首发日期: 2023-08-15

基金项目: 国家自然科学基金(61375011)、杭州电子科技大学研究生科研创新基金(GK228810299187)

通信作者: \*gbf@hdu.edu.cn

化使用,这也是传统的像元级遥感分类和面积测量精度难以达到使用要求的主要原因<sup>[7]</sup>。为了克服混合像元产生的不利影响,高光谱解混技术应运而生。高光谱解混技术的目的是提取出混合像元中每种纯物质(也称为端元)并估计其在该像元中所占的比例(也称丰度),混合像元经过解混后可以从中获得更加精准的亚像素级的信息,为对高光谱图像进行更加精准的信息提取提供了可能。

高光谱解混本身属于不适定问题,当数据分布未知时,神经网络能比传统方法取得更加精确的效果<sup>[8]</sup>,以深度学习方式进行光谱解混逐渐成为了解混领域研究的热点问题。自动编码器作为深度学习的一种典型模型,能够无监督地将高维数据转换成低维特征,这与高光谱解混中的丰度估计过程类似,越来越多的研究开始使用自动编码器进行高光谱解混方向的研究<sup>[9-13]</sup>。早期的研究大都使用单层编码器,例如文献[9],但是其编码器输入容易受到噪声干扰。为了提升解混算法对噪声的鲁棒性,文献[10]采用堆叠自编码对原始输入进行去噪处理。一些研究者认为,通过堆叠自编码进行去噪的方式可能会引入额外的重构误差,因此文献[12]使用边缘化去噪自编码算法对输入数据进行去噪,将去噪能力以去噪约束的方式纳入网络优化中,并且在编码器上使用 $L_{1/2}$ 范数,在降低重构误差的同时减少冗余端元。上述研究都只关注高光谱图像中的光谱信息,然而高光谱图像存在“图谱合一”的特征,其图像本身是通过光谱特征和空间特征共同描述的,因此相邻波段的光谱特征以及相邻像元间的空间特征均应与混合像元有紧密联系<sup>[14]</sup>。为了利用遥感图像中的空间信息,文献[13]构建新的损失函数,使得训练过程中需要同时考虑中心以及邻域像元的光谱信息,提升了丰度的平滑性。此外,文献[15]首次将卷积神经网络引入自动编码器中,由于卷积操作会天然地带入邻域像素信息,因此使用卷积自编码进行光谱解混时能够直接在网络结构中引入空间信息,避免了传统算法对空间相关性假设不当导致性能下降等问题。大量的研究<sup>[16-19]</sup>开始将卷积层引入自编码网络中,试图通过引入空间信息来提升解混网络的性能,但是文献[15]中的实验结果却表明,卷积核设置过大,在增加计算复杂度的同时也会使解混精度降低。此外,文献[18]提出了一种利用级联卷积自编码进行光谱解混的方法,虽然网络结构中也引入了卷积层,但是实验结果却表明,当卷积核大小设置为 $1 \times 1$ 时,即没有引入空间信息时,网络的精度最高。这表明,解混过程中不合理地引入空间信息可能会为光谱解混带来负面影响,而仅通过修改卷积核的尺寸不能很好地对引入的空间信息进行平衡,因此需要通过合理的网络结构均衡提取到的空间信息和光谱信息。

针对基于卷积自编码来联合空间信息进行光谱解

混的研究中,不适当的网络结构中过度引入空间信息可能对光谱解混效果产生的负面影响,本文提出一种结合注意力机制的双流卷积自编码高光谱解混方法(DSCU-Net),尝试通过训练的方式从提取到的空间特征和光谱特征中自行选择出更具有表征性的特征进行光谱解混,以此来提高解混的精度。具体地,本文所做创新点如下。

1)为了兼顾光谱信息在欧几里得空间上的相似性和光谱相似性,在损失函数中同时引入欧氏距离和光谱角距离,将二者加权之和作为重构项的度量。

2)提出一种新的双流卷积自编码解混网络,并引入通道注意力机制对可能存在的冗余信息进行处理,使得解混网络能够关注关键的特征。

3)提出一种多尺度特征融合方法来提取高光谱数据中的特征,通过调整卷积核尺寸来分别提取不同尺寸下的高光谱特征并融合,提高网络对不同尺度特征的表达能。

## 2 基于自动编码器的光谱解混方法

自动编码器一般由编码器和解码器两部分组成。其中编码器将观测到的高光谱像素 $\mathbf{X} \in \mathbf{R}^{L \times N}$ 编码成隐藏层特征 $\mathbf{h}_L$ , $L$ 和 $N$ 分别为高光谱图片的波段数和像素数。而解码层中,解码器根据隐藏层特征 $\mathbf{h}_L$ 进行重构输入,记重构的高光谱数据为 $\hat{\mathbf{X}} \in \mathbf{R}^{L \times N}$ ,则隐藏特征和重构高光谱图片的表达式分别为

$$\begin{cases} \mathbf{h}_L = G_{\text{en}}(\mathbf{W}_{\text{en}}\mathbf{X} + \mathbf{b}_{\text{en}}) \\ \hat{\mathbf{X}} = G_{\text{de}}(\mathbf{W}_{\text{de}}\mathbf{h} + \mathbf{b}_{\text{de}}) \end{cases}, \quad (1)$$

式中: $G_{\text{en}}(\cdot)$ 和 $G_{\text{de}}(\cdot)$ 分别表示编码层和解码层的激活函数,通常为ReLU、Sigmoid、Tanh等; $\mathbf{W}_{\text{en}}$ 和 $\mathbf{W}_{\text{de}}$ 分别代表编码层和解码层权重; $\mathbf{b}_{\text{en}}$ 和 $\mathbf{b}_{\text{de}}$ 为编码器和解码器的偏置项。通过最小化原始数据与重构的高光谱数据之间的重构误差 $L_{\text{re}}$ 训练网络,即

$$(\mathbf{W}_{\text{en}}, \mathbf{W}_{\text{de}}, \mathbf{b}_{\text{en}}, \mathbf{b}_{\text{de}}) = \arg \min_{\mathbf{W}_{\text{en}}, \mathbf{W}_{\text{de}}, \mathbf{b}_{\text{en}}, \mathbf{b}_{\text{de}}} \frac{1}{N} \sum_{i=1}^N (\hat{\mathbf{x}}_i - \mathbf{x}_i)^2, \quad (2)$$

当网络训练完成时,重构出的高光谱数据与原输入的高光谱数据足够小,即认为此时解码器能根据编码器编码出的隐藏特征重构出原高光谱数据。而线性光谱混合模型的数学表达式为

$$\mathbf{X} = \mathbf{MA} + \mathbf{E}, \quad (3)$$

式中: $\mathbf{M} \in \mathbf{R}^{L \times P}$ 是端元数为 $P$ 的端元矩阵; $\mathbf{A} \in \mathbf{R}^{P \times N}$ 是丰度矩阵; $\mathbf{E} \in \mathbf{R}^{L \times N}$ 代表额外噪声。结合式(3),可以将隐藏层特征 $\mathbf{h}_L$ 理解为要估计的丰度向量 $\hat{\mathbf{A}}$ ,解码器的权重 $\mathbf{W}_{\text{de}}$ 代表提取的端元矩阵 $\hat{\mathbf{M}}$ ,因此在基于自动编码的光谱解混方法中,丰度和端元求解过程可以表示为

$$\begin{cases} \hat{\mathbf{A}} \leftarrow \mathbf{h} \\ \hat{\mathbf{M}} \leftarrow \mathbf{W}_{\text{de}} \end{cases}. \quad (4)$$

### 3 基于双流卷积自编码的高光谱解混网络

新的网络结构由特征融合网络 (feature fusion network) 和光谱解混网络 (spectral unmixing network)

两部分组成,具体网络配置如图 1 所示。其中特征融合网络从原始输入数据中提取出光谱特征和空间特征并对它们加以融合,而主干解混网络则将融合后的特征作为输入,实现空-谱联合解混。

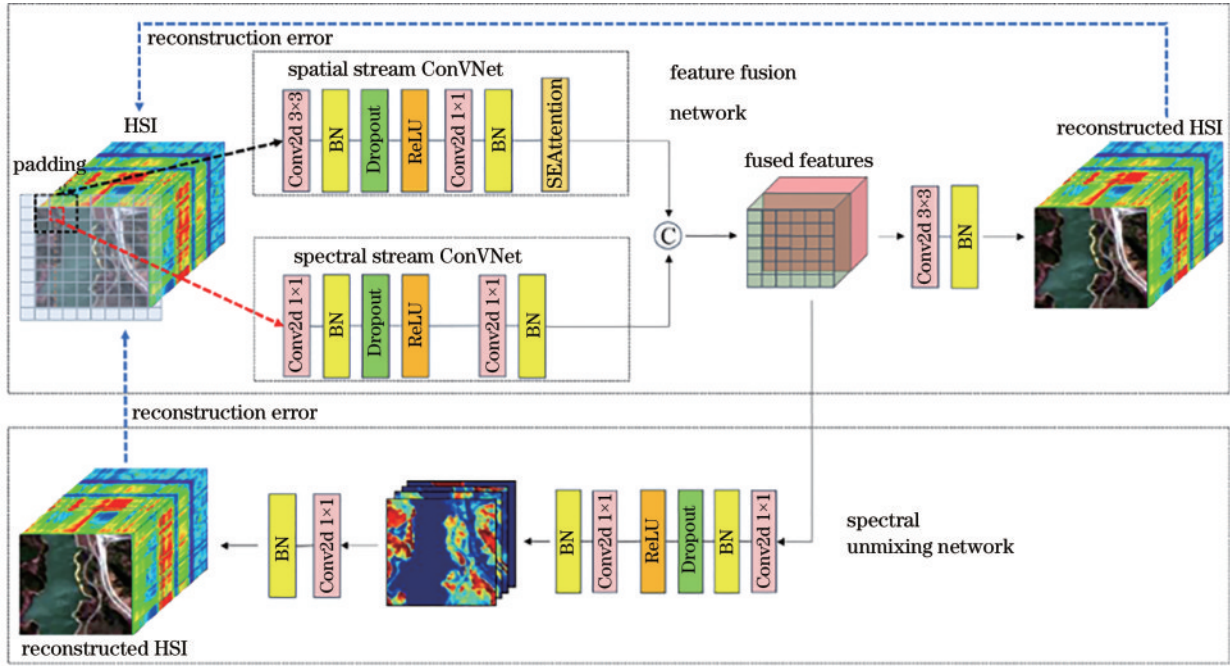


图 1 所提方法的网络框架

Fig. 1 Network architecture of proposed method

#### 3.1 特征融合网络

特征融合网络是一个由双流卷积网络作为编码层的自动编码网络。双流卷积网络由光谱流卷积网络 (spectral stream convolutional network) 和空间流卷积

网络 (spatial stream convolution network) 组成,分别提取高光谱图像的光谱特征和空间特征。网络的具体配置如表 1 所示。

表 1 特征融合网络的配置

Table 1 Configuration of the feature fusion network

Spectral stream convolution network	Spatial stream convolution network
Conv2D, size is $1 \times 1$ , number of channels is 128, padding is 1	Conv2D, size is $3 \times 3$ , number of channels is 128, padding is 1
BatchNorm2d	BatchNorm2d
Dropout	Dropout
ReLU	ReLU
Conv2D, size is $1 \times 1$ , number of channels is 64, padding is 1	Conv2D, size is $1 \times 1$ , number of channels is 64, padding is 1
BatchNorm2d	BatchNorm2d
	SEAttention

##### 1) 光谱流卷积网络

其中光谱流卷积网络有 2 层卷积层。首先通过卷积核尺寸为  $1 \times 1$  的二维卷积来提取中心像素位置的光谱特征,并且设置批正则化来加速学习进程,同时使用 Dropout 层抑制网络过拟合,使用 ReLU 函数作为激活函数。接下来是另外一个卷积层,卷积核大小为  $1 \times 1$ ,同样使用批正则化加速网络收敛。

##### 2) 空间流卷积网络

空间流卷积网络则由 2 层卷积层和通道注意力机制层组成。输入为  $f \times f$  大小的像元块,其中  $f$  为奇数,通过设置一层卷积核尺寸为  $f \times f$  的二维卷积来对像元块的中心像素执行卷积操作,以此来提取出包含邻域像元光谱信息的特征来辅助中心像元解混,与光谱流卷积网络类似,同样采用批正则化和 Dropout 层来加速网络收敛并且防止网络过拟合。

在高光谱解混领域,通常认为混合像元的光谱存



在着一定的空间相关性,即某像元处的光谱信息应该与邻域内其他像元的光谱信息有关,引入空间信息可以增加光谱解混在复杂场景下的适用性,但是这种假设并不总是有效的。混合光谱的形成是由端元物质的光谱、几何分布、在像元中的分布共同决定的<sup>[1]</sup>,因此,并非每个像元处的光谱与其邻域像元都具有空间相关

性。为了进一步挖掘这种局部空间相关性,在卷积层后引入了通道注意力机制对特征进行重加权处理。

文献[20]中提出了通道注意力机制,其目的是通过修改特征通道之间的依赖关系来提高提取特征的质量,如图2所示。本文使用的通道注意力机制主要分为压缩网络(squeeze network)和拓展网络(excitation network)。

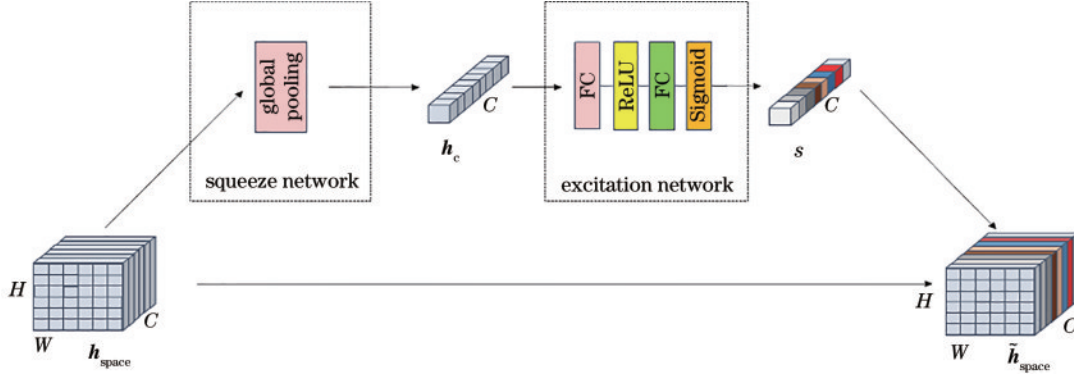


图2 所使用的通道注意力机制框架

Fig. 2 Architecture of proposed channel attention mechanism

**压缩网络:**压缩网络首先通过全局池化实现特征压缩,即将提取出的空间特征的每个通道特征尺寸压缩为 $1 \times 1$ 。

$$\mathbf{h}_c = F_{sq}(\mathbf{h}_{space}) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W \mathbf{h}_{space}(i, j), \quad (5)$$

式中: $\mathbf{h}_{space} \in \mathbf{R}^{C \times H \times W}$ 为经过卷积层提取到的初始空间特征图; $H$ 和 $W$ 分别为空间特征图 $\mathbf{h}_{space}$ 的长和宽; $C$ 为通道数; $\mathbf{h}_c \in \mathbf{R}^{C \times 1 \times 1}$ 表示压缩后的特征图。由于全局池化过程引入了该通道上所有的像素,认为经全局池化后的特征向量具有全局的感受野。

**扩展网络:**压缩网络仅仅能捕获每个通道上的信息,并没有带入通道与通道之间的依赖信息,而扩展网络则将压缩网络得到的特征依次输入全连接层、ReLU激活层、全连接层、Sigmoid激活层,以此来引入更加完整的空间信息。

$$\mathbf{s} = F_{ex}(\mathbf{w}_1, \mathbf{w}_2, \mathbf{h}_c) = \sigma[\mathbf{w}_2 \delta(\mathbf{w}_1 \mathbf{h}_c)], \quad (6)$$

式中: $\sigma$ 代表ReLU激活函数; $\delta$ 代表Sigmoid激活函数; $\mathbf{w}_1$ 和 $\mathbf{w}_2$ 为两个全连接层参数; $\mathbf{s} \in \mathbf{R}^{C \times 1 \times 1}$ 表示经过扩展网络得到的权重矩阵。最后,将得到的权重矩阵 $\mathbf{s}$ 的每个通道的值与原特征矩阵的值相乘,即得到选择后的特征向量 $\tilde{\mathbf{h}}_{space}$ ,并利用Concat操作对提取到的空间特征和光谱特征进行拼接,得到融合特征 $\mathbf{h}_{all} \in \mathbf{R}^{2C \times H \times W}$ ,相关的公式为

$$\begin{cases} \tilde{\mathbf{h}}_{space} = \mathbf{h}_{space} \times \mathbf{s} \\ \mathbf{h}_{all} = \text{Concat}(\tilde{\mathbf{h}}_{space}, \mathbf{h}_{spectral}) \end{cases} \quad (7)$$

为了确保融合特征 $\mathbf{h}_{all} \in \mathbf{R}^{2C \times H \times W}$ 能够表征原高光谱图像的关键信息,将融合特征送入解码器中,对原始高光谱数据 $\hat{\mathbf{X}}_1$ 进行重构。由于文献[21]发现在解码层中引入偏差项后的表现类似于一个端元,可能会

引入不必要的噪声,因此在特征融合网络和光谱解混网络的解码层均不使用偏差项。

### 3.2 光谱解混网络

与特征融合网络类似,光谱解混网络同样采用了自动编码器的网络结构。首先是输入层,特征融合网络提取出的融合特征理论上同时包含了原始高光谱数据的光谱特征和空间特征,更具有表征性,因此将特征融合网络得到的融合特征作为输入,网络的第二层是一个二维卷积层,有64个特征图,卷积核尺寸为 $1 \times 1$ 。与光谱融合网络类似,光谱解混网络也使用了批归一化和Dropout来抑制网络失活,并且加速网络收敛。第三层是一个有 $P$ 个特征图的二维卷积层, $P$ 为原始输入光谱数据的端元数,卷积核尺寸为 $1 \times 1$ 。第二层与第三层联合构成了光谱解混网络的编码层,对融合特征进行编码,得到隐层特征 $\mathbf{h}_l$ 。网络的最后一层是二维卷积层构成的解码层。解码层根据隐层特征 $\mathbf{h}_l$ 重构对应的原始高光谱数据。网络的具体配置如表2所示。

为了约束丰度向量满足丰度非负约束(ANC),参考文献[17]的做法,使用clamp函数将编码器的输出限制在 $[0, 1]$ 。许多研究表明,clamp函数在高光谱图像领域能比其他传统激活函数有更好的性能<sup>[22-23]</sup>。此外,通过最小化每个像元位置丰度之和与1的差异来施加丰度和为一(ASC)约束,表达式为

$$L_{reg} = \sum_{i=1}^N \left\| \mathbf{1} - \sum_{j=1}^P \hat{\mathbf{a}}_{ji} \right\|, \quad (8)$$

式中: $\hat{\mathbf{a}}_{ji}$ 表示预测丰度矩阵 $\hat{\mathbf{A}}$ 的第 $j$ 行第 $i$ 列的值。通过最小化式(8),即可实现丰度合为一约束。

### 3.3 损失函数的构建

通常,基于自编码进行光谱解混的研究中使用均

表 2 光谱解混网络的配置

Table 2 Configuration of the spectral unmixing network

Encoder layer	Conv2D, size is $1 \times 1$ , number of channels is 64, padding is 0
	BatchNorm2d
	Dropout
	ReLU
Decoder layer	Conv2D, size is $1 \times 1$ , number of channels is $P$ , padding is 1
	ReLU
	Conv2D, size is $1 \times 1$ , number of channels is $P$ , padding is 1
	ReLU

方根误差(RMSE)计算重构损失。文献[20]为了减少光照、地形等因素导致的光谱信息放缩现象,使用光谱角距离(SAD)代替均方根误差计算重构误差。相较于均方根误差,光谱角距离能更好地处理光谱可变性,但是却不能提高欧氏空间的相似性。为了均衡均方根误差和光谱角距离对解混效果的影响,提出一种新的

计算重构损失的方法。

$$L_{re} = (1 - \lambda)L_{SAD} + \lambda L_{RMSE}, \quad (9)$$

式中: $\lambda$ 用来控制均方根误差和光谱角距离重构项之间的平衡; $L_{SAD}$ 和 $L_{RMSE}$ 分别表示重构高光谱图与原始图片之间的光谱角距离误差和均方根误差。 $L_{SAD}$ 和 $L_{RMSE}$ 各自的计算公式为

$$\begin{cases} L_{RMSE} = \frac{\alpha}{N} \sum_{i=1}^N (\hat{\mathbf{x}}_{1i} - \mathbf{x}_i)^2 + \frac{(1-\alpha)}{N} \sum_{i=1}^N (\hat{\mathbf{x}}_{2i} - \mathbf{x}_i)^2 \\ L_{SAD} = \beta \arccos\left(\frac{\mathbf{X}^T \hat{\mathbf{X}}_1^T}{\|\mathbf{X}\|_2 \|\hat{\mathbf{X}}_1^T\|_2}\right) + (1-\beta) \arccos\left(\frac{\mathbf{X}^T \hat{\mathbf{X}}_2^T}{\|\mathbf{X}\|_2 \|\hat{\mathbf{X}}_2^T\|_2}\right), \end{cases} \quad (10)$$

式中: $\hat{\mathbf{X}}_1 = [\hat{\mathbf{x}}_{11}, \dots, \hat{\mathbf{x}}_{1N}] \in \mathbf{R}^{L \times N}$ 和 $\hat{\mathbf{X}}_2 = [\hat{\mathbf{x}}_{21}, \dots, \hat{\mathbf{x}}_{2N}] \in \mathbf{R}^{L \times N}$ 分别表示特征融合网络和光谱解混网络重构出的高光谱数据; $\alpha$ 和 $\beta$ 用来控制不同重构项在总损失中的占比。最后,所提网络中,总的损失函数 $L_{all}$ 的计算公式为

$$L_{all} = L_{re} + \theta L_{reg}, \quad (11)$$

式中: $\theta$ 表示控制正则项与重构损失项之间平衡的权重系数。

## 4 实验内容

为了评估所提方法的性能,在两个真实数据集上进行实验,与 $L_{1/2}$  nonnegative matrix factorization ( $L_{1/2}$ -NMF)<sup>[24]</sup>、total variation regularized reweighted sparse nonnegative matrix factorization (TV-RSNMF)<sup>[25]</sup>两种基于非负矩阵分解的解混方法进行了对比。为了验证所提网络结构能够解决基于自编码网络的解混方法由于过度引入空间信息造成精度下降的问题,还与convolutional neural network autoencoder unmixing (CNNAEU)<sup>[15]</sup>、cycle-consistency unmixing network by learning cascaded autoencoder (CyCU-Net)<sup>[18]</sup>、autoencoder network with adaptive abundance smoothing(AAS)<sup>[13]</sup>3种典型的基于自编码网络进行空-谱联合的解混算法以及一种基于深度学习的非线性解混算法 mode-based deep autoencoder network for nonlinear hyperspectral unmixing (MAC-U)<sup>[26]</sup>进行了对

比实验。在所有的对比方法中,都使用顶点成分分析法进行端元的初始化。所提方法的参数设定如表3所示。

表 3 实验中的参数设置

Table 3 Parameter setting in our experiment

Dataset	$\alpha$	$\beta$	$\lambda$	$\theta$	Learning rate	Epoch
Samson	0.3	0.3	0	$1 \times 10^{-4}$	0.008	520
Jasper Ridge	0.6	0.8	0.4	$1 \times 10^{-7}$	0.008	280

### 4.1 评估指标

实验使用均方根误差用于计算估计丰度矩阵与真实丰度矩阵之间的偏差, RMSE值越小,丰度估计的精度越高。光谱角距离则用来衡量提取端元与真实端元之间的相似性。与均方根误差不同的是,光谱角距离将端元光谱看作一个高维向量,通过计算两向量之间的夹角来度量光谱间的相似性能。光谱角距离的值越小,端元提取精度越高。各自的计算公式为

$$RMSE(\hat{\mathbf{a}}_i, \mathbf{a}_i) = \sqrt{\frac{1}{N} \sum_{i=1}^N \|\hat{\mathbf{a}}_i - \mathbf{a}_i\|_2^2}, \quad (12)$$

$$SAD(\hat{\mathbf{m}}_j, \mathbf{m}_j) = \frac{1}{P} \sum_{j=1}^P \arccos\left(\frac{\hat{\mathbf{m}}_j^T \mathbf{m}_j}{\|\hat{\mathbf{m}}_j\| \|\mathbf{m}_j\|}\right), \quad (13)$$

式中: $\hat{\mathbf{a}}_i$ 和 $\mathbf{a}_i$ 分别是估计的丰度值和真实的丰度值; $\hat{\mathbf{m}}_j$ 和 $\mathbf{m}_j$ 分别是提取到的端元和真实的端元。

### 4.2 实验环境及所用数据集

实验使用的电脑运行内存为 4 GB,处理器为英特

尔 Xeon(至强)E5645,主频 2.4 GHz。为了更好地评估所提方法的性能,利用了 3 个真实数据集(Samson 数据集、Jasper Ridge 数据集和 Cuprite 数据集),数据集具体的介绍如下。

1) Samson 数据集:由 SAMSON 传感器捕获到的原始数据集包含了  $952 \times 952$  个像素,156 个波段,波谱范围为  $0.401 \sim 0.889 \mu\text{m}$ 。由于原始数据较大,本文截取了其中  $95 \times 95$  大小的子图像进行实验,其中包含了“水”、“树木”、“土壤”3 种端元。

2) Jasper Ridge 数据集:该数据集由喷气推进实验室(JPL)的机载可见/红外成像光谱仪(AVIRIS)捕获。原始数据包含了  $512 \times 614$  个像素,224 个波段,波谱范围为  $0.38 \sim 2.50 \mu\text{m}$ 。在去除其中受到密集水蒸气和大气影响的波段后,获得了包含 198 个波段的数据集,本文截取了其中  $100 \times 100$  大小的子图像进行实验,其中包含了“树木”、“土壤”、“水”、“道路”4 种端元。

3) Cuprite 数据集:该数据集由可见/红外光谱仪(AVIRIS)在内华达州 Cuprite 上空捕获,波谱范围为  $0.37 \sim 2.48 \mu\text{m}$ 。在剔除噪声通道和水吸收通道后,获得了包含 188 个波段的数据集。本文截取了其中

$190 \times 250$  大小的子图像进行实验,其中包含 14 种矿物质。由于相似的矿物质变体之间的光谱差异较小,因此将端元数目减少到 12 个。

### 4.3 实验结果与分析

#### 1) Jasper Ridge 数据集上的实验

图 3 展示了在 Jasper Ridge 数据集下不同算法的估计丰度图与真实丰度图的比较。可以看出: $L_{1/2}$ -NMF、TV-RSNMF、CyCU-Net、CNNAEU、MAC-U 都不能很好地对水和道路进行解混;所提方法所估计的丰度图与真实丰度图最为接近;通过与丰度的真值图进行对比可以发现,CNNAEU 算法不能很好地区分开两条较窄的道路,分析可能是 CNNAEU 算法过度地引入了丰度的平滑性导致的;CyCU-Net 虽然也使用了卷积自编码进行光谱解混,但相比 CNNAEU,可以更加清晰地分辨出两条较窄的道路,原因是其卷积核尺寸设置为  $1 \times 1$ ,实际 CyCU-Net 进行光谱解混时没有引入空间信息;所提方法虽然同样使用了卷积自编码进行光谱解混,但可以较为清晰地区分出丰度图中两条较窄的道路,因此从定性的角度分析,DSCU-Net 确实能够有效地减少过度引入空间相关性导致丰度过于平滑的现象。

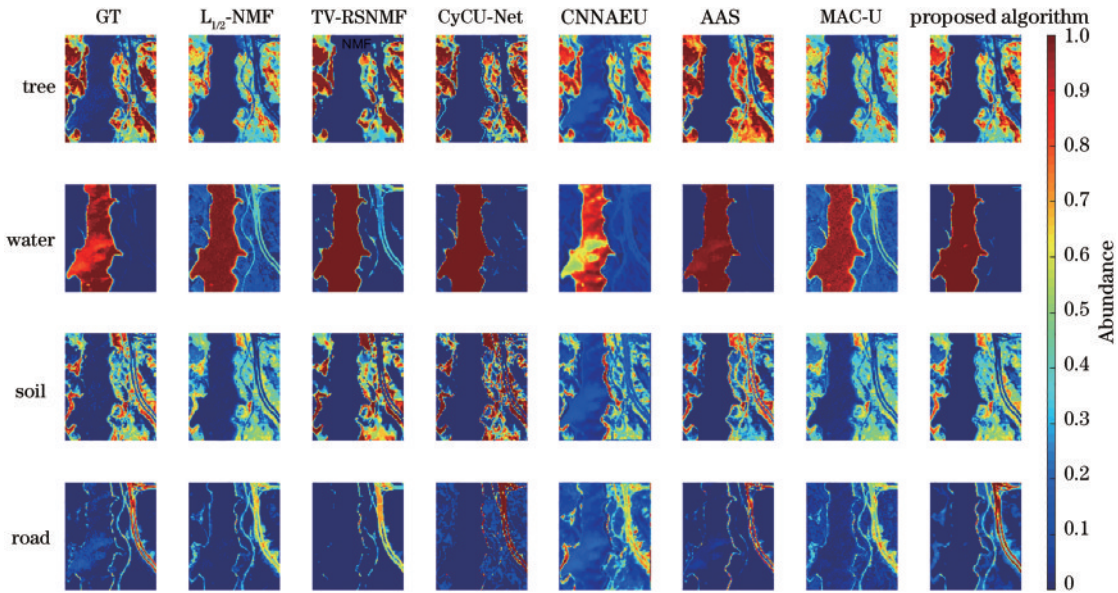


图 3 不同算法获得的 Jasper Ridge 数据集中的树木、水、土壤和道路的丰度图

Fig. 3 Abundance maps of trees, water, soil, and roads on Jasper Ridge dataset obtained by different algorithms

表 4 展示了在 Jasper Ridge 数据集下不同算法得出的 SAD 值和 RMSE 值,最优值用加粗字体表示。从表 4 可以看出,所提方法提取到的端元平均 SAD 和 RMSE 都小于其他方法,其中水和道路这两种端元与真实端元最为接近,因此从定量的角度分析,所提方法的效果相比其他对比算法是最优的。

#### 2) Samson 数据集上的实验

图 4 展示了 Samson 数据集下不同算法的估计丰度图与真实丰度图的比较。从图 4 可以看出: $L_{1/2}$ -NMF、

TV-RSNMF、CyCU-Net、AAS、CNNAEU、MAC-U 和 DSCU-Net 都对水对应的丰度图存在不同程度的错误估计;其中 MAC-U 得到的水对应的估计结果有着大面积的错误估计,分析可能是所选择场景中端元数量较少,且场景较为简单,非线性解混算法的端元之间的非线性关系过度拟合导致的;DSCU-Net 所提取到的丰度图与真值丰度图最为接近;此外,从 CNNAEU 和 AAS 算法对应的丰度图中可以明显观察到一些区域存在模糊的现象,分析可能是过度引入了空间信息导



表 4 Jasper Ridge 数据集下不同算法的定量结果

Table 4 Quantitative results of different algorithms on Jasper Ridge dataset

Parameter	$L_{1/2}$ -NMF	TV-RSNMF	CNNAEU	CyCU-Net	AAS	MAC-U	Proposed algorithm
SAD	Tree	0.2046	<b>0.0340</b>	0.1420	0.0399	0.0387	0.1546
	Water	0.2729	0.1583	0.4159	0.1527	0.1330	0.0900
	Soil	0.0799	<b>0.0306</b>	0.0526	0.0306	0.1554	0.1175
	Road	0.0691	0.0692	0.1005	0.0392	0.0456	0.0913
Mean SAD	0.1566	0.0730	0.1778	0.0656	0.0932	0.1134	<b>0.0580</b>
RMSE	0.1420	0.1131	0.1790	0.1163	0.1233	0.1599	<b>0.0811</b>

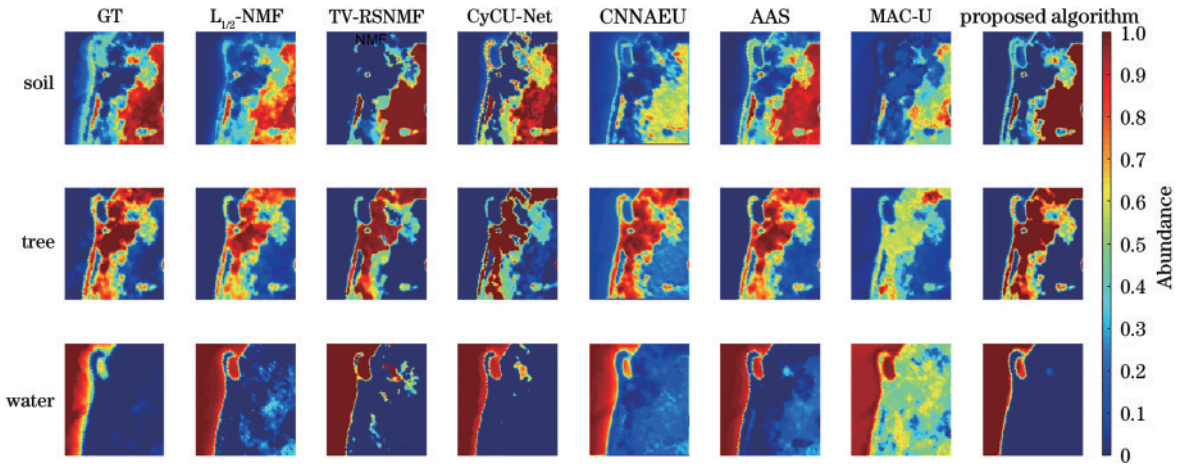


图 4 不同算法获得的 Samson 数据集中的土壤、树木和水的丰度图

Fig. 4 Abundance maps of soil, trees, and water on Samson dataset obtained by different algorithms

致的丰度的平滑性过高。所提方法却很好地避免此类情况的发生,再一次证明所提网络结构能够避免过度引入空间信息而造成丰度图像过于平滑的现象。

表 5 展示了 Samson 数据集下不同算法得出的 SAD 值和 RMSE 值。从表 5 可以得出:所提算法提取

到的端元平均 SAD 和 RMSE 都小于其他方法,其中水对应的端元与真实端元最为接近,土壤、树木两种物质对应端元提取结果在所对比方法中也均处于较优的水平。因此从定量的角度分析,所提算法在 Samson 数据集中的结果相比其他对比算法是最优的。

表 5 Samson 数据集下不同算法的定量结果

Table 5 Quantitative results of different algorithms on Samson dataset

Parameter	$L_{1/2}$ -NMF	TV-RSNMF	CNNAEU	CyCU-Net	AAS	MAC-U	Proposed algorithm
SAD	Soil	0.0403	0.0166	0.1954	<b>0.0102</b>	0.0342	0.0195
	Tree	0.0848	0.0291	0.1100	<b>0.0250</b>	0.0697	0.0391
	Water	0.2884	0.2462	0.2455	0.0520	0.1760	0.1562
Mean SAD	0.1378	0.0973	0.1836	0.0290	0.0933	0.0716	<b>0.0285</b>
RMSE	0.1668	0.1904	0.1531	0.1779	0.1228	0.3293	<b>0.07404</b>

### 3) 在复杂场景下的性能表现

对于 Jasper Ridge 和 Samson 数据集,其端元个数较少,无法验证算法在复杂场景下的性能,因此在 Cuprite 数据集上进行了实验。表 6 展示了在 Cuprite 数据集下 CNNAEU、MAC-U 及所提方法进行端元提取的结果。由表 6 结果可知:CNNAEU 算法取得了最佳的端元提取效果;所提算法与 MAC-U 次之,分析可能是在 Cuprite 数据集中端元的数量较多,且端元之间的光谱差异较小,像元光谱容易受到邻域像元光谱的影响。CNNAEU 算法在进行端元估计时,卷积核尺寸设置为  $11 \times 11$ ,即端元估计过程中,某处像元的光

谱信息是由该像元与其邻域像元共同决定的,因此 CNNAEU 所取得的端元估计精度较高。

图 5 展示了不同算法对应的丰度估计结果,实验选取了编号分别为 #1Alunite、#3Buddingtonite、#5KaoLinite1、#7Muscovite、#9Nontronite、#11 Sphere 的端元对应的丰度图。由于 Cuprite 数据集没有提供先验的丰度信息,无法通过直接计算估计丰度与丰度真值之间的差异来判定算法的性能。从整体看:CNNAEU 算法所估计丰度图更加均匀,即丰度相对平滑,但可能会丢失一些细节信息;与 Cuprite 数据伪彩色图片进行比对发现,对于 #11Sphene 端元所估计

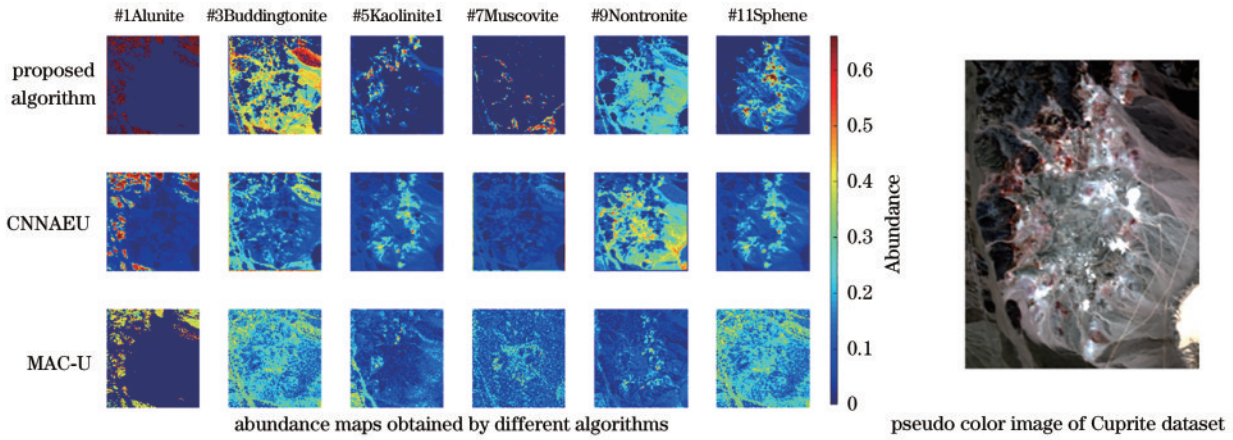


图 5 不同算法获得的 Cuprite 数据集中不同端元对应的丰度图和 Cuprite 数据对应的伪彩色图片

Fig. 5 Abundance maps corresponding to different endmembers in Cuprite dataset obtained by different algorithms and the pseudo color image of Cuprite dataset

表 6 Cuprite 数据集下各种算法的 SAD 值

Table 6 SAD values of different algorithms on Cuprite dataset

Sequence	MAC-U	CNNAEU	Proposed algorithm
#1Alunite	<b>0.0963</b>	0.4665	1.0038
#2Andradite	<b>0.0729</b>	0.0809	0.1001
#3Buddingtonite	0.1177	<b>0.1114</b>	0.1584
#4Dumortierite	0.1134	0.1241	<b>0.1004</b>
#5Kaolinite1	0.2312	0.1744	<b>0.0819</b>
#6Kaolinite2	0.0966	<b>0.0709</b>	0.0882
#7Muscovite	0.1140	0.1768	<b>0.1084</b>
#8Montmorillonite	0.0874	0.0873	<b>0.0719</b>
#9Nontronite	0.0960	<b>0.0789</b>	0.0805
#10Pyrope	0.1160	0.0838	<b>0.0588</b>
#11Sphene	0.8368	0.0720	<b>0.0703</b>
#12Chalcedony	0.1142	0.1380	<b>0.1188</b>
Mean	0.1743	<b>0.1395</b>	0.1701

的丰度图,所提算法能够较为完整地还原出原始数据右下角存在的浅白色直线,而 CNNAEU 和 MAC-U 算法所估计的丰度图中则较为模糊。这表明所提算法确实能够抑制过度引入空间相关性导致丰度过于平滑的现象。

## 5 结 论

提出了一种结合注意力机制的双流卷积自编码高光谱解混方法(DSCU-Net),相比以往的研究(直接使用卷积提取出的光谱特征进行光谱解混),所提网络使用双流结构分别提取空间特征和光谱特征,同时引入通道注意力机制对提取出的空间特征进行重加权,并且对得到的空间特征和光谱特征进行融合,使用融合后的特征进行空-谱联合高光谱解混。在两个真实数据集上进行实验,并且对比分析了复杂场景下算法的性能表现。结果表明 DSCU-Net 有效地减少了过度引

入空间特征导致丰度过于平滑的现象,提升了光谱解混的精度。但所提方法仍有不足之处,复杂场景下不能取得很好的解混效果;此外,虽然引入了光谱角距离作为重构损失的度量,但是由于解码层中对所有像素都使用了同一端元矩阵,不能很好地处理光谱可变性带来的影响。因此如何处理同一场景下光谱可变性导致的光谱矩阵不唯一是接下来要研究的重点问题。

## 参 考 文 献

- [1] 童庆禧,张兵,郑兰芬. 高光谱遥感:原理、技术与应用[M]. 北京:高等教育出版社,2006.  
Tong Q X, Zhang B, Zheng L F. Hyperspectral remote sensing: theory, technology and application[M]. Beijing: Higher Education Press, 2006.
- [2] 龚威,史硕,陈博文,等. 机载高光谱激光雷达成像技术发展与应用[J]. 光学学报, 2022, 42(12): 1200002.  
Gong W, Shi S, Chen B W, et al. Development and application of airborne hyperspectral LiDAR imaging technology[J]. Acta Optica Sinica, 2022, 42(12): 1200002.
- [3] 陈善学,何宇峰. 基于空谱字典的加权联合稀疏表示高光谱图像分类[J]. 光学学报, 2023, 43(1): 0110002.  
Chen S X, He Y F. Weighted joint sparse representation hyperspectral image classification based on spatial spectrum dictionary[J]. Acta Optica Sinica, 2023, 43(1): 0110002.
- [4] 汪丽,崔新月,索建平,等. 空气温度绝对探测激光雷达系统标定技术[J]. 光学学报, 2022, 42(18): 1828001.  
Wang L, Cui X Y, Suo J P, et al. Calibration technology of lidar system for absolute detection of air temperature[J]. Acta Optica Sinica, 2022, 42(18): 1828001.
- [5] Shaw G A, Burke H K. Spectral imaging for remote sensing[J]. Lincoln Laboratory Journal, 2003, 14(1): 3-28.
- [6] 贾响响,郭宝峰,丁繁昌,等. 基于约束非负矩阵分解的高光谱解混[J]. 光子学报, 2021, 50(7): 0710005.  
Jia X X, Guo B F, Ding F C, et al. Hyperspectral unmixing based on constrained nonnegative matrix



- factorization[J]. *Acta Photonica Sinica*, 2021, 50(7): 0710005.
- [7] 伍小明. 基于核方法的高光谱遥感图像解混技术研究[D]. 杭州: 浙江大学, 2021.  
Wu X M. Research on unmixing technology of hyperspectral remote sensing images based on kernel method[D]. Hangzhou: Zhejiang University, 2021.
- [8] Zhang X R, Sun Y J, Zhang J Y, et al. Hyperspectral unmixing via deep convolutional neural networks[J]. *IEEE Geoscience and Remote Sensing Letters*, 2018, 15(11): 1755-1759.
- [9] Guo R, Wang W, Qi H R. Hyperspectral image unmixing using autoencoder cascade[C]//2015 7th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), June 2-5, 2015, Tokyo, Japan. New York: IEEE Press, 2017.
- [10] Su Y C, Li J, Plaza A, et al. DAEN: deep autoencoder networks for hyperspectral unmixing[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(7): 4309-4321.
- [11] Qu Y, Qi H R. uDAS: an untied denoising autoencoder with sparsity for spectral unmixing[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(3): 1698-1712.
- [12] Chen M, Weinberger K, Sha F, et al. Marginalized denoising auto-encoders for nonlinear representations [C]//Proceedings of the 31th International Conference on Machine Learning, ICML 2014, June 21-26, 2014, Beijing, China. Cambridge: JMLR, 2014: 1476-1484.
- [13] Hua Z Q, Li X R, Qiu Q H, et al. Autoencoder network for hyperspectral unmixing with adaptive abundance smoothing[J]. *IEEE Geoscience and Remote Sensing Letters*, 2021, 18(9): 1640-1644.
- [14] 袁博. 空间与谱间相关性分析的 NMF 高光谱解混[J]. *遥感学报*, 2018, 22(2): 265-276.  
Yuan B. NMF hyperspectral unmixing algorithm combined with spatial and spectral correlation analysis[J]. *Journal of Remote Sensing*, 2018, 22(2): 265-276.
- [15] Palsson B, Ulfarsson M O, Sveinsson J R. Convolutional autoencoder for spectral-spatial hyperspectral unmixing[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(1): 535-549.
- [16] Ranasinghe Y, Herath S, Weerasooriya K, et al. Convolutional autoencoder for blind hyperspectral image unmixing[C]//2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS), November 26-28, 2020, RUPNAGAR, India. New York: IEEE Press, 2021: 174-179.
- [17] Tulczyjew L, Kawulok M, Longépé N, et al. A multibranch convolutional neural network for hyperspectral unmixing[J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 6011105.
- [18] Gao L R, Han Z, Hong D F, et al. CyCU-net: cycle-consistency unmixing network by learning cascaded autoencoders[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5503914.
- [19] Hua Z Q, Li X R, Jiang J F, et al. Gated autoencoder network for spectral - spatial hyperspectral unmixing[J]. *Remote Sensing*, 2021, 13(16): 3147.
- [20] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 7132-7141.
- [21] Ozkan S, Kaya B, Akar G B. EndNet: sparse AutoEncoder network for endmember extraction and hyperspectral unmixing[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(1): 482-496.
- [22] Zheng K, Gao L R, Liao W Z, et al. Coupled convolutional neural network with adaptive response function learning for unsupervised hyperspectral super resolution[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(3): 2487-2502.
- [23] Yao J, Hong D, Chanussot J, et al. Cross-attention in coupled unmixing nets for unsupervised hyperspectral super-resolution[M]//Vedaldi A, Bischof H, Brox T, et al. *Computer vision-ECCV 2020. Lecture notes in computer science*. Cham: Springer, 2020, 12374: 208-224.
- [24] Qian Y T, Jia S, Zhou J, et al. Hyperspectral unmixing via  $L_{1/2}$  sparsity-constrained nonnegative matrix factorization[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2011, 49(11): 4282-4297.
- [25] He W, Zhang H Y, Zhang L P. Total variation regularized reweighted sparse nonnegative matrix factorization for hyperspectral unmixing[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 55(7): 3909-3921.
- [26] Li H Q, Borsoi R A, Imbiriba T, et al. Model-based deep autoencoder networks for nonlinear hyperspectral unmixing[J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 5506105.