

doi: 10.6046/zrzyg.2021115

引用格式: 曲海成,王雅萱,申磊.多感受野特征与空谱注意力结合的高光谱图像超分辨率算法[J].自然资源遥感,2022,34(1):43–52. (Qu H C, Wang Y X, Shen L. Hyperspectral super-resolution combining multi-receptive field features with spectral-spatial attention[J]. Remote Sensing for Natural Resources, 2022, 34(1): 43–52.)

多感受野特征与空谱注意力结合的高光谱图像超分辨率算法

曲海成, 王雅萱, 申磊

(辽宁工程技术大学软件学院, 葫芦岛 125105)

摘要: 针对高光谱图像超分辨率过程中, 图像细节信息容易丢失的问题, 提出多感受野特征与空谱注意力结合的高光谱图像超分辨率算法, 该算法充分利用高光谱图像中的高频信息与低频信息, 减少图像细节信息丢失, 提升了高光谱图像超分辨率效果。首先, 在特征提取阶段采用不同大小卷积核的卷积, 获取到多尺度感受野特征, 更好地提取低分辨率图像中的高频信息与低频信息, 有助于保留原始图像的特征信息; 然后, 把获取到图像特征, 经过“空间-光谱”结合的注意力机制增强, 利用光谱维信息辅助空间维特征重建; 最后, 把每组的特征融合, 通过像素级反卷积层缓解棋盘格效应, 输出清晰的高分辨率图像。实验结果表明: 提出的多感受野特征与空谱注意力结合的超分辨率算法在 Chikusei 和 Pavia center scene 这2个公开数据集上峰值信噪比分别达到了 39.869 7 和 31.942 2, 结构相似度分别达到了 0.937 6 和 0.878 6, 与最新超分辨率算法比较有明显的性能优势。该文提出的算法, 结合了多感受野特征提取模块和空谱结合注意力模块的优势, 超分辨率后的图像细节特征有了明显的改善。

关键词: 高光谱图像; 图像超分辨率; 多感受野特征提取; 注意力机制

中图分类号: TP 751.1 **文献标志码:** A **文章编号:** 2097-034X(2022)01-0043-10

0 引言

高光谱图像是由高光谱成像传感器在多个连续的窄光谱波段上获取同一场景的图像, 通常有几十个到上百个波段, 其中包含了丰富的光谱信息。高光谱图像被广泛应用于生物识别^[1]、植被检测^[2]和医疗诊断^[3]等领域。高光谱图像超分辨率是一种信号后处理技术的研究热点, 旨在从低分辨率图像中获取高分辨率图像, 以克服高光谱图像固有的分辨率限制。

现有高光谱图像超分辨率方法可以分为2类: 基于多源融合的超分辨率方法和基于单幅图像的超分辨率方法。基于多源融合的方法是利用同一场景的辅助高分辨率图像如多光谱图像、全色图像等重建高分辨率高光谱图像。Akhtar等^[4]通过利用辅助图像和低分辨率高光谱图像的稀疏表

示, 提出了一种基于稀疏空间光谱表示的超分辨率方法, 实现了辅助空间信息和光谱信息的融合。Dian等^[5]提出了一种用于低分辨率高光谱图像与高分辨率多光谱图像融合的深度高光谱图像锐化方法, 该方法通过基于深度卷积神经网络的残差学习直接学习图像先验知识, 提高了重建精度。这些融合方法通常都能取得很好的效果, 但这些方法都需要辅助图像, 这极大地限制了实际应用。单幅高光谱图像超分辨率方法不需要辅助图像直接从低分辨率高光谱图像重建高分辨率高光谱图像。主要包括基于插值、基于稀疏表示和基于深度学习的方法。早期 Huang等^[6]提出在傅里叶插值变换域内重构, 从而恢复高频信息, 但插值方法能力有限, 且重构图像细节丢失严重, 导致该方法实用性不强。进一步, 练秋生等^[7]基于稀疏表示提出了一种基于两级字典以及分频带字典的单幅图像超分辨率算法, 该算法利用高低分辨率系数之间的对应关系进行图像重

收稿日期: 2021-04-20; 修订日期: 2021-08-17

基金项目: 国家自然科学基金面上基金项目“改进 BRDF 先验知识耦合策略的 Landsat30 米地表反照率模型研究与验证”(编号: 42071351)、辽宁省教育厅基础科研项目“可见光与红外图像跨域深度行人检测模型研究”(编号: LJ2019JL010)和辽宁工程技术大学学科创新团队资助项目“智慧农业遥感监测创新团队”(编号: LNTU20TD-23)共同资助。

第一作者: 曲海成(1981-), 男, 博士, 副教授, 主要研究方向为遥感影像高性能计算、目标检测识别。Email: quhaicheng@lntu.edu.cn。

通信作者: 王雅萱(1994-), 女, 硕士研究生, 主要研究方向为深度学习、高光谱图像处理。Email: 940556702@qq.com。

构。之后,一些基于稀疏表示的单幅图像高光谱图像超分辨率方法相继被提出^[8],这些方法虽然明显优于基于插值的方法,但由于其浅层启发式模型的表达能力有限,仍然不能很好地恢复复杂的图像细节。近年来,深度学习算法已被广泛应用于各类计算机视觉任务中,由于其神经网络架构,能够拟合复杂的非线性映射函数,研究者开始研究适合单幅高光谱图像超分辨率的映射函数。Li 等^[9]通过低分辨率观测邻近波段与高分辨率观测波段之间的光谱差异学习深度映射,与传统方法相比取得了优异的结果。Qiao 等^[10]提到,深度网络中输入的低分辨率图像与输出的高分辨率图像一定程度上是近似的,即低分辨率图像中的高低频信息与高分辨率图像携带的高低频信息相似,利用图像的低频信息和高频信息,使重建图像的轮廓、细节信息更丰富,完成低分辨率图像向高分辨率图像的映射。另外,基于深度学习的单幅图像超分辨率方法虽然取得了一定的成果,但大多忽略了高光谱图像相邻光谱波段间的强相关性,不能深入研究组间的光谱相关性,便限制了网络提取光谱特征的能力。

为此,本文提出了一种多感受野特征提取与空谱注意力结合的超分辨率算法,命名为 MASR。该算法利用高光谱图像光谱分辨率高、波段覆盖范围广、“图谱合一”等特点,进行空谱结合的注意力增强;同时通过卷积核大小不同的多尺度卷积提取高光谱图像的高频信息和低频信息,进行图像的特征重建;此外还应用像素级反卷积缓解棋盘格效应,改善图像特征重建质量。

1 相关工作

1.1 图像高频信息与低频信息

图像高频信息表示图像中灰度值变化剧烈的区域,对应着图像的边缘、纹理等,反映的是小范围的细节信息。图像低频信息表示图像中灰度值变化缓慢,对应着图像中连续渐变的区域,也就是大范围的轮廓信息,即常说的图像大致概貌和背景,是图像的近似信息。为了获取图像中的高低频信息,使用傅里叶变换实现对图像频率的过滤,过滤方法包括高通滤波和低频滤波,所谓高通滤波便是保留图像中的高频信息,过滤低频成分,高通滤波表达式为:

$$H_h(u, v) = 1 - e^{-D^2(u, v)/2D_0^2}, \quad (1)$$

式中: $H_h(u, v)$ 为高通滤波; (u, v) 为频率矩形中心点坐标; $D(u, v)$ 为距频率矩形中心的距离; D_0 为截止频率。

低通滤波与之相反,只保留图像中的低频信息,低通滤波表达式为:

$$H_l(u, v) = e^{-D^2(u, v)/2\sigma^2}, \quad (2)$$

式中: $H_l(u, v)$ 为低通滤波; σ 为关于中心的扩展度的度量。

为了直观感受图像中高频信息和低频信息的差异,在 Chikusei 和 Pavia Centre scene 数据集上,对图像中的高频信息和低频信息进行可视化,显示结果如图 1 所示。

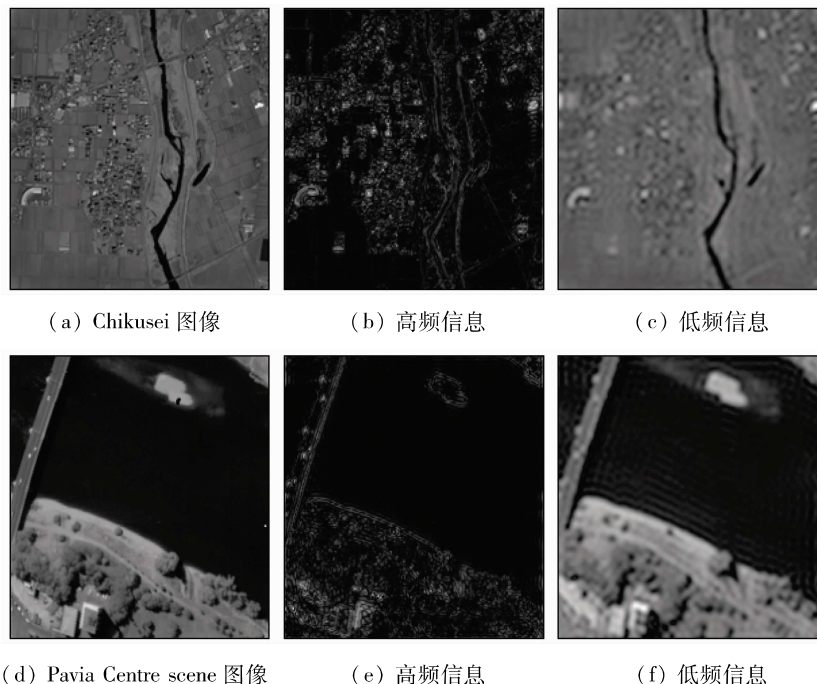


图 1 在 Chikusei 和 Pavia Centre scene 数据集上高低频信息

Fig. 1 High and low frequency information on the Chikusei and the Pavia Centre scene dataset

在卷积过程中,较大的卷积核可以提取图像细节特征,保留原始图像高频信息,较小的卷积核可以提取图像轮廓信息,若在特征提取的过程中使用单一卷积,可能造成提取特征信息的丢失,因此使用不同大小的卷积核构成多感受野卷积,有效地保证了特征图信息的完整性。

1.2 注意力机制

注意力机制 (attention mechanism, AMM) 是机器学习中的数据处理方法,旨在通过深度网络自动学习图像中的局部重要特征,通过对特征信息之间的相互依赖进行建模,利用神经网络的平均池化层生成一个 mask 掩码, mask 上的值代表图像不同区域的特征权重,基于权重值自适应地重新划分特征图中的信息重要度。

空间注意力机制 (spatial attention module, SAM) 可以引导网络聚焦于空间域中最具特征量的区域,通常的做法是将特征图沿通道轴做平均池化处理,

生成二维的特征描述符,该描述符标注了空间域中需要被关注或抑制的特征位置,增强网络对重要信息量的捕获与传递能力。通道注意力机制 (channel attention module, CAM) 是对空间注意力的有效补充,通过平均池化层顺着空间维度进行特征压缩,将每个通道的二维的空间特征变为一个实数,某种程度上,这个实数对特征具有全局的感知能力,代表了该通道对全局特征的响应权重,衡量了每个通道对特征有效提取的贡献能力。

2 MASR 网络模型

不同于对高光谱图像整体进行特征提取,为了更充分利用高光谱图像相邻光谱波段之间的相关性,MASR 网络将高光谱图像沿光谱波段分成多个组。每组主要包括特征提取、上采样、特征融合和图像重建 4 个部分,总体网络结构如图 2 所示。

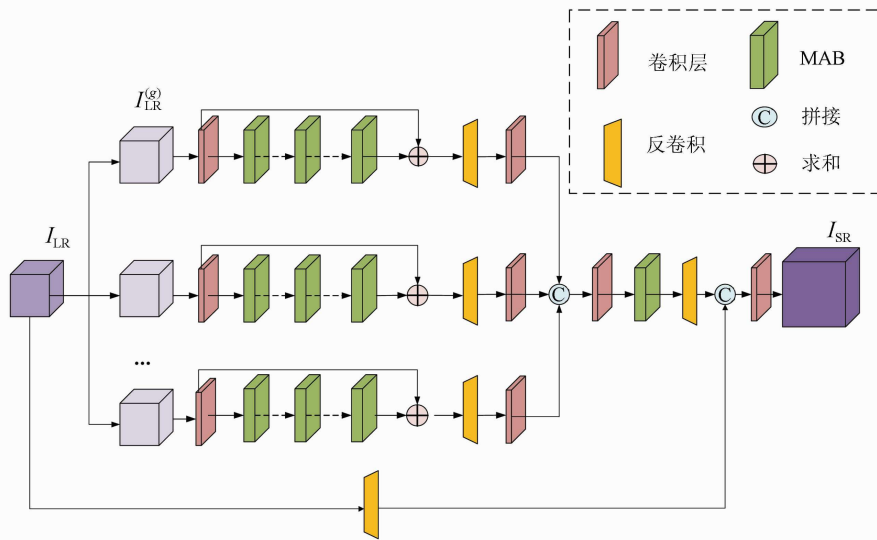


图 2 总体网络结构

Fig. 2 Overall network structure

该网络将多感受野特征与 AMM 串联在一起,封装成多感受野特征提取注意力模块 (multi-receptive field feature extraction attention block, MAB)。首先,输入低分辨率高光谱图像 I_{LR} , $I_{LR} = \{I_{LR}^1, I_{LR}^2, \dots, I_{LR}^G\}$, 其中 G 表示分组的组数。每组 $I_{LR}^{(g)}$ 进行一个 3×3 卷积,其中 g 表示每组光谱通道数,生成的特征图送入到 MAB,进行特征提取,接着为了避免邻近上采样形成的棋盘格效应^[11],本文使用像素反卷积的方式进行上采样,融合各个分支的特征,最后与上采样后的低分辨率图像进行长连接,重建

高分辨率高光谱图像 I_{SR} 。

2.1 多感受野特征提取空谱注意力模块

图像超分辨率是一个不适定问题,需要额外的先验知识 (正则化) 来约束重建过程。对于高光谱图像的超分辨率问题,有效利用高光谱图像的内在特性,即空间上的非局部自相似性和光谱间的高相关性是至关重要的。因此,本文提出了多感受野特征提取和空谱注意力模块来利用空间和光谱先验知识,提取图像特征,该模块结构如图 3 所示。

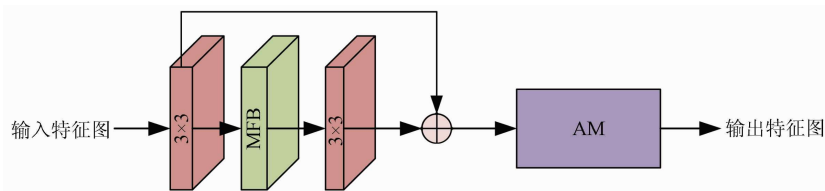


图 3 多感受野特征提取注意力模块

Fig. 3 Multi-receptive field feature extraction attention block

该模块包括多感受野特征提取模块(multi-receptive field feature extraction block, MFB)和空谱结合注意力模块(space spectrum combined with attention module, AM)2部分,该模块可以从输入图像中获得既有低频信息又有高频信息的特征。

2.2 多感受野特征提取模块

卷积核的大小会影响特征提取过程中所获取的信息。小尺寸核对低频信息的特征提取效果较好,对高频信息的特征提取效果较差,而大尺寸核对高频信息的特征提取效果较好,但对小细节的特征提取效果较差。这个问题可以通过在不同的卷积神经网络层中使用不同大小的卷积核来解决。但是,这增加了卷积神经网络架构的深度,会导致更高的复杂性。因此,本文使用同一层内的多感受野卷积,构成 MFB,该模块结构如图 4 所示。

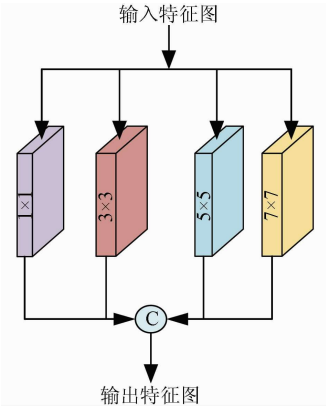


图 4 多感受野特征提取模块

Fig. 4 Multi-receptive field feature extraction module

该模块,同时使用不同的卷积核进行特征提取,该模块采取的卷积核大小分别为 $1 \times 1, 3 \times 3, 5 \times 5, 7 \times 7$,小卷积核可以提取低频信息即图像中比较平滑区域,大卷积核提取高频信息即图像的边缘轮廓信息,同时提取低频信息和高频信息,使输入的特征图有更细的信息。上述过程描述的表达式为:

$$F^{[o]} = \text{Conv}[\text{Concat}(f_1, f_3, f_5, f_7)] \quad (3)$$

式中: $F^{[o]}$ 为输出特征图; Conv 表示卷积操作; Concat 表示特征图通道拼接; f_1, f_3, f_5, f_7 分别对应的是 $1 \times 1, 3 \times 3, 5 \times 5, 7 \times 7$ 卷积操作。具体表达式分别为:

$$F = F_{\text{spa}} + \text{GAP}[\text{Conv}_1(F_{\text{spa}})] \otimes \text{Conv}_1[\text{Conv}_1(F_{\text{spa}})] \quad (4)$$

式中: F 为输出特征图; Conv_1 为 1×1 卷积操作。

2.4 损失函数

为了更好地训练 MASR 网络模型,使得超分辨率结果接近真实的高分辨率图像,本文使用的总体损失函数^[12] L 定义为:

$$L(\theta) = L_1 + \beta L_{\text{SSTV}} \quad (7)$$

$$\begin{cases} f_1 = w_1 F^{[i]} + b_1 \\ f_3 = w_3 F^{[i]} + b_3 \\ f_5 = w_5 F^{[i]} + b_5 \\ f_7 = w_7 F^{[i]} + b_7 \end{cases} \quad (4)$$

式中: $F^{[i]}$ 为输入特征图; w 和 b 分别代表卷积的权重和偏置,特征图从 4 个卷积操作中获得,它们可以通过多感受野特征的连接操作组合起来。为了对连接后的特征图进行降维,将其输入至一个 3×3 的卷积层,降低通道维度。

2.3 空谱结合注意力模块

该模块有助于有效地提高网络的表示能力,文中使用的 AM 设计如图 5 所示。

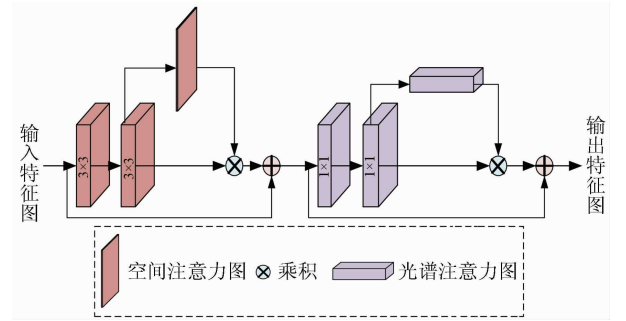


图 5 空谱结合注意力模块

Fig. 5 Space spectrum combined with attention module

该模块前半部分是空间残差注意力,后半部分是光谱残差注意力。该模块采用 2 个 3×3 卷积构建更深层网络,然后沿通道维度做全局空间池化,形成空间注意力图。将注意力图和卷积后的特征图相乘,增强高光谱图像中信息的传递,将空间注意力残差模块和光谱残差模块串联,将输出的特征图经过 2 个 1×1 卷积块,做光谱通道缩放,学习光谱维度的权重值,将其与特征图相乘,增强光谱信息的有效表达,通过光谱特征之间的相互依赖进行建模,自适应地重新分配每个通道特征。前半部分过程描述的表达式为:

$$F_{\text{spa}} = F' + \text{GAP}[\text{Conv}_3(F')] \otimes \text{Conv}_3[\text{Conv}_3(F')] \quad (5)$$

式中: F_{spa} 为空间注意力; F' 为输入的特征图; $\text{GAP}()$ 为全局平均池化函数; Conv_3 为 3×3 卷积。后半部分过程描述的表达式为:

$$F = F_{\text{spa}} + \text{GAP}[\text{Conv}_1(F_{\text{spa}})] \otimes \text{Conv}_1[\text{Conv}_1(F_{\text{spa}})] \quad (6)$$

式中: β 为平衡不同损失贡献度的权重参数,通常使用常数去衡量,所以在实验中设定 $\beta = 1\text{E} - 3$; L_1 损失被称为最小绝对值偏差,它衡量了目标值与估计值之间的绝对差值,在整个训练阶段都能保持较好的收敛性,所以选定 L_1 损失来衡量模型重构精度。该 L_1 损失函数定义为:

$$L_1(\theta) = \frac{1}{M} \sum_{m=1}^M ||I_{\text{HR}}^m - F_{\text{Net}}(I_{\text{LR}}^m)||_1, \quad (8)$$

式中: I_{HR}^m 为第 m 幅高光谱图像真值图; $F_{\text{Net}}(I_{\text{LR}}^m)$ 为第 m 幅高光谱图像重建图; $||\cdot||_1$ 表示 1 范式;

$$L_{\text{SSTV}}(\theta) = \frac{1}{M} \sum_{m=1}^M (||\nabla_h I_{\text{SR}}^m||_1 + ||\nabla_w I_{\text{SR}}^m||_1 + ||\nabla_c I_{\text{SR}}^m||_1), \quad (9)$$

式中: I_{SR}^m 为第 m 幅高光谱图像超分辨率结果; ∇_h , ∇_w 和 ∇_c 分别为用来计算 I_{SR}^m 的长、宽和光谱的梯度函数。

3 实验与结果分析

3.1 实现数据集与参数设置

实验用 3 个高光谱图像数据集来测试,分别为 Chikusei, Pavia Centre scene 和 CAVE 数据集。

第一个高光谱数据集是 Chikusei 数据集,有 128 个波段,光谱范围为 363 ~ 1 018 nm。场景由 2 517 × 2 335 像素组成,地面采样距离为 2.5 m。首先对图像的中心区域进行裁剪,得到 2 304 × 2 048 × 128 像素的子图像,再将其分为训练数据和测试数据。提取该图像顶部区域作为测试数据,测试数据为 4 幅无重叠的高光谱图像,图像大小为 512 × 512 × 128 像素。

第二个数据集是 Pavia Centre scene 数据集,它有 102 个光谱波段(从最初的 115 个光谱波段中去除了水汽吸收和噪声光谱波段)和 1 096 × 1 096 像素。需要注意的是,在中心场景中,不包含信息的区域被删除,只留下一个有意义的 1 096 × 715 像素区域。对图像的中心区域进行裁剪,提取图像的左半部分作为测试数据,测试数据为 4 幅无重叠的高光谱图像,图像大小为 224 × 224 像素。

为了证明本文算法对于高光谱图像的普适性,本文不仅选择遥感高光谱数据集进行实验,还在普通高光谱数据集 CAVE 上进行验证。CAVE 数据集是由 32 个日常物体场景组成,图像大小为 512 × 512 像素,有 31 个光谱波段,光谱范围为 400 ~ 700 nm。为了准备训练样本,从数据集中随机抽取 20 幅高光谱图像,随机抽取 10% 的样本作为验证集。数据集剩余的 12 幅高光谱图像用于测试,其中原始图像作为高分辨率高光谱图像的真值图。输入的低分辨率高光谱图像是真值图空间分辨率的 1/4 大小。

实验采用 Pytorch 框架实现了多感受野特征与空谱结合的卷积神经网络,所使用计算机的 CPU 型号为 Intel Core i7 - 6700 @ 3.4 GHz, GPU 型号为 NVIDIA TITAN RTX,内存为 24 GB。

M 为一次训练批中图像的数量; θ 为网络的参数集。为了更好地捕获空间和光谱信息的相关性,本文使用 Aggarwal 等^[13]提出的空间 - 光谱总变差损失函数, L_{SSTV} 损失函数定义为:

3.2 评价指标

实验中使用了 6 个指标评估本文的算法,它们分别是均方根误差 $RMSE$, 峰值信噪比 $PSNR$, CC (cross correlation)^[14], 光谱角映射 SAM , 相对全局合成误差 $ERGAS$ ^[15], 结构相似度 $SSIM$ ^[16]。对于重建高光谱图像的 $PSNR$ 和 $SSIM$, 取它们在所有波段的平均值,即平均峰值信噪比 $MPSNR$ ^[17] 和平均结构相似度 $MSSIM$ ^[17]。 CC , SAM 和 $ERGAS$ 是高光谱融合任务中被广泛采用的 3 个质量指标^[12], 其余 3 个指标是常用于定量评价图像恢复质量的指标。这些指标的最佳值分别是 0, $+\infty$, 1, 0, 0 和 1。

3.3 消融实验

3.3.1 多感受野特征提取与注意力模块

为了验证 MFB 和 AM 的有效性,实验对 Chikusei 数据集采用 6 个指标进行实验比较,结果见表 1。

表 1 Chikusei 数据集在不同模块实验对比
Tab. 1 Experimental comparison of Chikusei dataset in different modules

指标	无 MFB	MFB	MFB + AM
$SAM \downarrow$	3.234 8	2.981 9	2.444 6
$RMSE \downarrow$	0.014 4	0.013 5	0.012 2
$MPSNR \uparrow$	38.298 5	38.832 6	39.869 7
$ERGAS \downarrow$	6.290 9	5.897 2	5.191 5
$MSSIM \uparrow$	0.911 2	0.921 9	0.937 6
$CC \uparrow$	0.932 0	0.939 5	0.951 5

由表 1 可知, MFB 模块可以有效地提取低频信息特征和低频信息特征, 与没有 MFB 相比, 采用 MFB 模块后, 本文算法获得了一定的性能提升。加入 MFB 模块后, 虽然各客观指标的改进幅度相对较小, 但 $MPSNR$ 提高了 0.534 1, SAM 降低了 0.252 9, 这证明了 MFB 的引入更有利于网络性能的提升。

AM 模块有助于有效提高网络的表示能力。由表 1 可知, 与没有 AM 相比, 采用 AM 机制后, 本文算法获得了一定的性能提升。加入 AM 模块后, 虽然各客观指标的改进幅度相对较小, $MPSNR$ 提高了 1.571 2, SAM 降低了 0.790 2。由此可见, AM 的引入更有利于光谱特征的表示。

为了从直观上验证本文添加的 MFB 和 AM 模块的有效性, 以 2 幅 Chikusei 数据集的区域为例, 图 6 给出了对比效果。

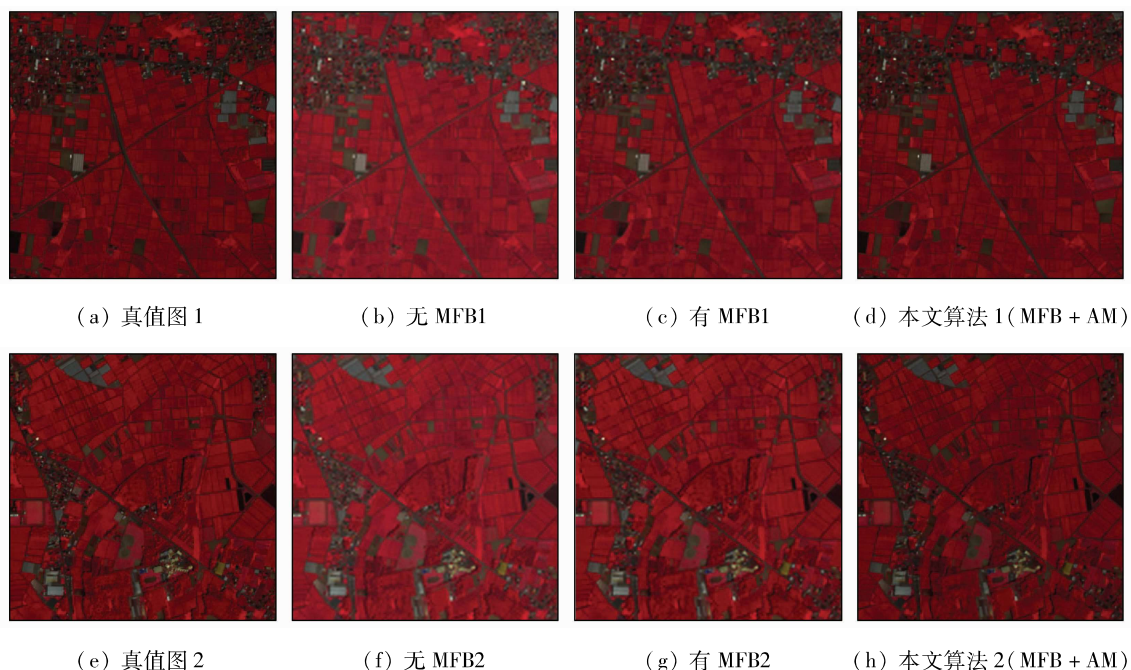


图 6 在 Chikusei 数据集上模块对比

Fig. 6 Module comparison of the Chikusei dataset

从图 6 可以看出,Chikusei 数据集是对农田的遥感影像图,当没有采用 MFB 模块时,田地间的分界线模糊,颜色间亮暗分布不均匀,加入 MFB 模块之后,图像的色差有所改善,且建筑物与农田间的边缘信息更加清晰,说明 MFB 模块可以有效地提取图像的细节信息及轮廓信息。本文算法结合了 MFB 模块与 AM 模块,由图 6(d)和(h)可知,继续加入的 AM 模块,可利用高光谱图像丰富的光谱信息辅助

存在缺陷的空间信息,重建更高空间分辨率的高光谱图像,实验效果不管是纹理还是颜色更接近于真值。本文算法可以获得有效的特征,使得特征信息更加有效、完整,达到了较理想的超分辨率效果。

同时为了验证方法的泛化能力,使用 Pavia Centre scene 数据集进行测试,以 2 幅 Pavia Centre scene 数据集的区域为例,结果如图 7 所示。

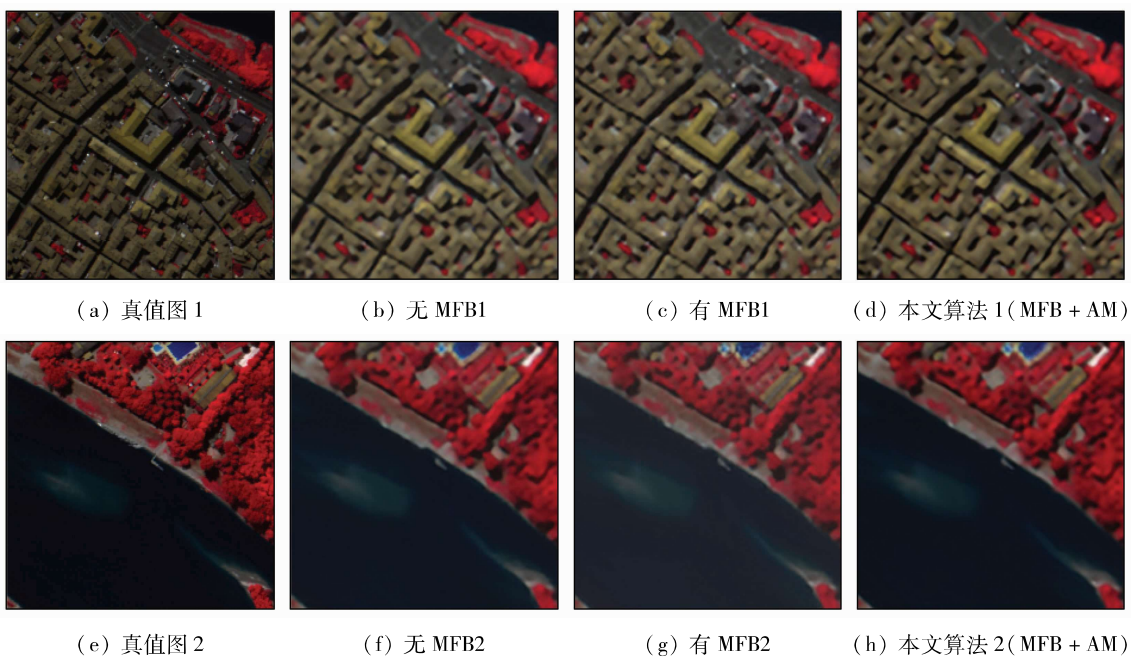


图 7 在 Pavia Centre scene 数据集上模块对比

Fig. 7 Module comparison of the Pavia Centre scene dataset

从图 7 可以看出,Pavia Centre scene 数据集是对水、沥青、砖和阴影等地方的遥感影像图,当没有采用 MFB 模块时,砖的分界线模糊,颜色间亮暗分

布不均匀,阴影与沥青边缘也模糊不清,加入 MFB 模块之后,图像的色差有所改善,且砖、阴影和沥青的边缘信息更加清晰,说明 MFB 模块可以有效地提

取图像的细节信息及轮廓信息。图 7(d) 和(h) 是本文算法的结果,本文算法结合 MFB 和 AM 模块后,图像中房屋间的边缘特征得到了更好的重建。相比于其他的实验结果,本文算法的蓝色屋顶更加清晰,阴影部分更加精细。通过效果图的对比可以看出,本文算法对复杂的地物目标超分辨率后,重建的细节特征更加明显。

3.3.2 分组策略

为了有效地利用高光谱图像相邻光谱波段间的相关关系,减少模型参数,本文设计了一种分组策略,即沿光谱维度将高光谱图像分成若干组,为尽可能地保留组间光谱波段的关联性,每组内设置重叠波段数 P 。为验证分组策略的可行性,在 Pavia Centre scene 数据集上,进行分组验证。首先通过将每组光谱通道数 g 固定为 8,探讨重叠因子 P 对图像超分辨率后 $MPSNR$ 的影响,实验结果如图 8 所示。

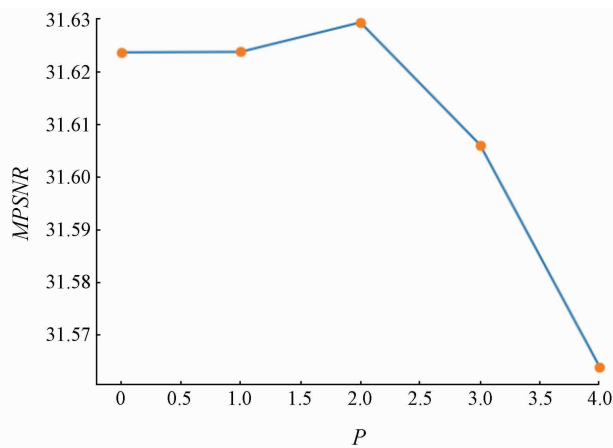


图 8 重叠因子对 $MPSNR$ 的影响

Fig. 8 Effect of overlap factor on $MPSNR$

从图 8 可以看出,重叠因子与 $MPSNR$ 是非线

性关系,并不是越大越好,当它取到临界值 $P = 2$ 时,此时网络中 $MPSNR$ 最大。可见,重叠因子为 2 时,已满足对组间光谱相关性的充分利用,若继续增加重叠波段,可能会对组内光谱信息造成干扰,因此分组时选择重叠因子为 2。为进一步确定组内光谱波段数 g 的选择,固定 $P = 2$ 进行实验验证,结果如图 9 所示。

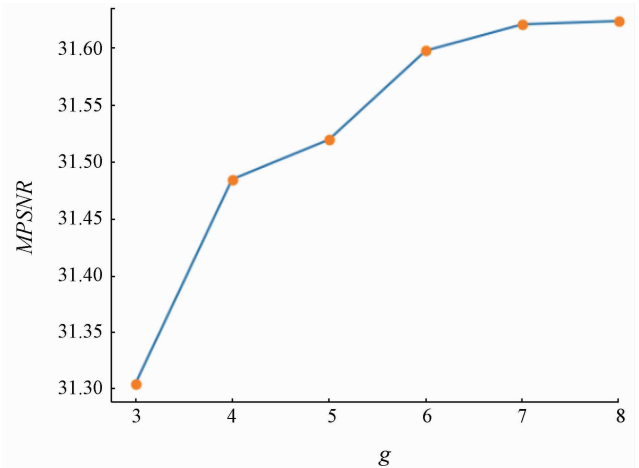


图 9 每组光谱通道数

Fig. 9 Number of spectral channels per group

从图 9 可以看出,当每组光谱波段数为 8 时,网络中 $MPSNR$ 最大,使用分组策略后,光谱重构性能有了很大的提高。经过多次实验表明当 $g = 8, P = 2$ 时,实验效果最佳。

3.4 与现有的算法对比实验

最后,为验证本文算法的有效性,在 Chikusei 数据集上,将本文算法与现有的算法进行比较,对比的算法有 Bicubic^[18], VDSR^[19], TLCNN^[20], EDSR^[21], RCAN^[22], SAN^[23], 3DCNN^[24], GDRRN^[25], DeepPrior^[26]。实验对比结果见表 2。

表 2 在 Chikusei 数据集上不同算法的对比结果

Tab. 2 Comparison results of different algorithms on the Chikusei dataset

算法	$RMSE \downarrow$	$MPSNR \uparrow$	$CC \uparrow$	$SAM \downarrow$	$ERGAS \downarrow$	$MSSIM \uparrow$
Bicubic	0.015 6	37.637 7	0.921 2	3.404 0	6.756 4	0.894 9
VDSR	0.014 8	37.775 5	0.922 7	3.664 2	6.870 8	0.906 5
TLCNN	0.015 0	37.725 1	0.919 6	3.857 3	6.752 2	0.900 8
3DCNN	0.014 0	38.609 1	0.935 5	3.117 4	6.002 6	0.912 7
GDRRN	0.013 7	38.719 8	0.936 9	2.500 0	5.954 0	0.919 3
DeepPrior	0.014 7	38.192 3	0.929 3	3.559 0	6.209 6	0.901 0
本文算法	0.012 2^①	39.869 7	0.951 5	2.444 6	5.191 5	0.937 6

①加粗字体为最优值。

表 2 中, VDSR 算法在引入了深度残差网络后进行图像超分辨,取得了一定的效果,但其残差结构不如本文设计得巧妙,并未有较好的超分辨率效果; 3DCNN 通过给定大量的原始图像及其损坏图像的实例对,可以很好地利用光谱相关性,将退化的图像恢复到原始图像,但其计算复杂度非常大,检测性能受到一定的制约; GDRRN 通过设计一组递归模块

并将其嵌入到全局残差结构中,提出了一种分组深度递归残差网络,这种群卷积和递归结构保证了它能够产生很好的性能,可见分组卷积模式有利于高光谱图像信息充分提取; DeepPrior 开发了一种有效的单幅高光谱图像恢复算法,由于高光谱训练样本有限,光谱波段维度高,很难充分挖掘高光谱图像空间信息和光谱之间的相关性。综上所述,本文提出

的 MASR 模型沿光谱波段进行分组卷积,使用不同卷积核精细化地提取原始图像信息,并且充分挖掘了空谱之间的相关性,超分辨率效果优于 TLCNN, 3DCNN, GDRRN 和 DeepPrior 等现有算法。

为了验证本文算法的泛化能力,使用 Pavia Centre scene 数据集与其他算法进行再次比较,对比结果见表 3。

表 3 在 Pavia Centre scene 数据集上不同算法的对比结果
Tab. 3 Comparison results of different algorithms on Pavia Centre scene dataset

算法	$RMSE \downarrow$	$MPSNR \uparrow$	$CC \uparrow$	$SAM \downarrow$	$ERGAS \downarrow$	$MSSIM \uparrow$
Bicubic	0.043 7	27.587 4	0.859 4	6.139 9	6.881 4	0.696 1
VDSR	0.041 9	27.882 1	0.865 9	6.700 4	6.699 1	0.724 2
RCAN	0.037 6	28.816 5	0.891 7	5.978 5	6.048 5^①	0.771 9
TLCNN	0.043 1	27.668 2	0.856 3	6.901 3	6.913 9	0.714 1
3DCNN	0.039 6	28.411 4	0.881 3	5.866 9	6.266 5	0.750 1
DeepPrior	0.041 0	28.106 1	0.872 3	6.266 5	6.484 5	0.736 5
本文算法	0.028 1	31.942 2	0.940 6	6.557 7	6.742 3	0.878 6

①加粗字体为最优值。

由表 3 可知,在不同的高光谱数据集 Pavia Centre scene 上,本文算法仍然优于其他先进算法。其中 $MPSNR$ 为 31.942 2, $RMSE$ 为 0.028 1。通过以上 2 个对比实验可以看出,本文算法使用的 MFB 有效地提取了图像的低频信息和高频信息,AM 能够

有效地增强特征表达能力。

以上 2 个数据集均为遥感高光谱数据集,为验证本文算法的有效性,在普通高光谱数据集 CAVE 上实现进一步的实验验证,对比结果见表 4。

表 4 在 CAVE 数据集上不同算法的对比结果
Tab. 4 Comparison results of different algorithms on CAVE dataset

算法	$RMSE \downarrow$	$MPSNR \uparrow$	$CC \uparrow$	$SAM \downarrow$	$ERGAS \downarrow$	$MSSIM \uparrow$
Bicubic	0.021 2	34.721 4	0.986 8	4.175 9	5.271 9	0.927 7
EDSR	0.014 9	38.157 5	0.993 1	3.549 9	3.592 1	0.952 2
RCAN	0.014 2	38.758 5	0.993 5^①	3.605 0	3.417 8	0.953 0
SAN	0.014 3	38.718 8	0.993 5	3.595 1	3.420 0	0.953 1
3DCNN	0.021 2	34.985 3	0.986 2	4.229 7	7.318 2	0.954 9
GDRRN	0.014 5	38.450 7	0.993 4	3.414 3	3.508 6	0.953 8
本文算法	0.013 9	39.172 9	0.986 5	3.384 2	4.328 4	0.957 2

①加粗字体为最优值。

由表 4 可知,在普通的高光谱图像 CAVE 数据集中,本文算法依然实现了较好的超分辨率效果,其中 $RMSE$ 为 0.013 9, $MPSNR$ 为 39.172 9。通过上述对比实验结果可以看出,本文算法针对普通高光谱数据集也适用,本文的 MFB 有效地提取了特征的高低频信息,AM 增强了图像的空间和光谱信息,捕获了更加敏感的特征信息,使得最终得到的特征图的信息更加完整且有效。

据的空间和光谱相关性特征,增强网络对图像细节信息的敏感性,更有效地重建超分辨率图像特征。

3)通过模块分析实验可以看出,本文使用的 2 个模块可以有效地提升重建结果的质量。该网络中也采用了像素反卷积层,缓解了特征图因上采样产生的棋盘格效应。在高光谱数据集上的实验结果表明,本文算法优于其他现有算法,在超分辨率上表现优异。

4 结论

本文提出的多感受野特征与空谱注意力结合的高光谱图像超分辨率算法,在高光谱图像超分辨率上取得了较好的效果,具体结论如下:

1)本文算法在单个卷积层中设计多尺度卷积核,充分提取多感受野特征,保证图像中高频信息和低频信息完整性。

2)空谱结合注意力模块充分利用了高光谱数

参考文献 (References):

- [1] Muhammad U, Arif M, Ajmal M. Hyperspectral face recognition using 3D-DCT and partial least squares[C]//Burghardt; BMVA Press. UK; University of Bristol, 2013: 57.1-57.10.
- [2] Lowe A, Harrison N, French A P. Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress[J]. Plant Methods, 2017, 13(1): 80-92.
- [3] Lin J, Clancy N T, Qi J, et al. Dual-modality endoscopic probe for tissue surface shape reconstruction and hyperspectral imaging enabled by deep neural networks[J]. Medical Image Analysis, 2018,

- 48:162 – 176.
- [4] Akhtar N, Shafait F, Mian A S. Sparse spatio – spectral representation for hyperspectral image super – resolution [C]//European Conference on Computer Vision, 2014:63 – 78.
- [5] Dian R, Li S, Guo A, et al. Deep hyperspectral image sharpening [J]. IEEE Transactions on Neural Network Learning Systems, 2018, 29(11):5345 – 5355.
- [6] Huang H, Yu J, Sun W. Super – resolution mapping via multi – dictionary based sparse representation [C]//IEEE International Conference on Acoustic, Speech and Signal Processing, 2014:3523 – 3527.
- [7] 练秋生, 张钧芹, 陈书贞. 基于两级字典与分频带字典的图像超分辨率算法 [J]. 自动化学报, 2013, 39(8):1310 – 1320.
Lian Q S, Zhang J Q, Chen S Z. Image superresolution algorithm based on two – level dictionary and subband dictionary [J]. Acta Automatica Sinica, 2013, 39(8):1310 – 1320.
- [8] Yang J, Wright J, Huang T S, et al. Image super – resolution via sparse representation [J]. IEEE Transactions on Image Processing, 2010, 19(11):2861 – 2873.
- [9] Li X, Wei H W, Zhang H Q. Super – resolution reconstruction of single remote sensing image combined with deep learning [J]. Journal of Image and Graphics, 2018, 23(2):209 – 218.
- [10] Qiao J, Song H, Zhang K, et al. Image super – resolution using conditional generative adversarial network [J]. IET Image Processing, 2019, 13(14):2673 – 2679.
- [11] Pan J T, Kevin M G, Elisa S, et al. Shallow and deep convolutional networks for saliency prediction [C]//IEEE Conference on Computer Vision and Pattern Recognition, 2016:598 – 606.
- [12] Li K, Dai D X, Ender K, et al. Hyperspectral image super – resolution with spectral mixup and heterogeneous datasets [EB/OL]. (2021 – 1 – 19) [2021 – 3 – 8]. <https://arxiv.org/pdf/2101.07589.htm>.
- [13] Aggarwal H K, Majumdar A. Hyperspectral image denoising using spatio – spectral total variation [J]. IEEE Geoscience and Remote Sensing Letters, 2016, 13(3):442 – 446.
- [14] Loncan L, De Almeida L B, Briottet X, et al. Hyperspectral pan-sharpening [J]. IEEE Geoscience and Remote Sensing Magazine, 2015, 3(3):27 – 46.
- [15] Wald L. Data fusion: Definitions and architectures: Fusion of images of different spatial resolutions [M]. Presses des MINES, Computer Science, 2002.
- [16] Peng J Y, Shi C Y, Laugeman E, et al. Implementation of the structural SIMilarity (SSIM) index as a quantitative evaluation tool for dose distribution error detection [J]. Medical Physics, 2020, 47(4):1907 – 1919.
- [17] Mei S, Yuan X, Ji J, et al. Hyperspectral image spatial super – resolution via 3D full convolutional neural network [J]. Remote Sensing, 2017, 9(11):1139.
- [18] 李成轶, 田淑芳. 基于字典学习的遥感影像超分辨率融合方法 [J]. 国土资源遥感, 2017, 29(1):50 – 56. doi:10.6046/gtzyyg.2017.01.08.
Li C Y, Tian S F. Super – resolution fusion method for remote sensing image based on dictionary learning [J]. Remote Sensing for Land and Resource, 2017, 29(1):50 – 56. doi:10.6046/gtzyyg.2017.01.08.
- [19] Kim J, Lee J K, Lee K M. Accurate image super – resolution using very deep convolutional networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016:1646 – 1654.
- [20] Yuan Y, Zheng X, Lu X. Hyperspectral image superresolution by transfer learning [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2017, 10(5):1963 – 1974.
- [21] Lim B, Son S, Kim H, et al. Enhanced deep residual networks for single image super – resolution [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017:136 – 144.
- [22] Zhang Y, Li K P, Li K, et al. Image super – resolution using very deep residual channel attention networks [C]//Proceedings of the European Conference on Computer Vision (ECCV), 2018:286 – 301.
- [23] Dai T, Cai J, Zhang Y, et al. Second – order attention network for single image super – resolution [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019:11065 – 11074.
- [24] Mei S, Yuan X, Ji J, et al. Hyperspectral image spatial super – resolution via 3D full convolutional neural network [J]. Remote Sensing, 2017, 9(11):1139.
- [25] Li Y, Zhang L, Ding L C, et al. Single hyperspectral image super – resolution with grouped deep recursive residual network [C]//Proceedings of the IEEE International Conference on Multimedia Big Data, 2018:1 – 4.
- [26] Sidorov O, Hardeberg J Y. Deep hyperspectral prior: Single – image denoising, inpainting, super – resolution [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision Workshop, 2019:3844 – 3851.

Hyperspectral super – resolution combining multi – receptive field features with spectral – spatial attention

QU Haicheng, WANG Yaxuan, SHEN Lei

(College of Software, Liaoning Technical University, Huludao 125105, China)

Abstract: To address the problem that image details are liable to be lost in the process of hyperspectral super – resolution, this study proposed a hyperspectral super – resolution algorithm that combines multi – receptive field features and spectral – spatial attention. By fully using the high – and low – frequency information in hyperspectral

images, this algorithm reduces the loss of image details and improves the hyperspectral super-resolution effects. First, in the feature extraction stage, convolution with different sizes of convolutional kernels is used to obtain multi-scale receptive field features. This assists in extracting more high- and low-frequency information from low-resolution images, thus retaining the features of original images. Then, the acquired image features are enhanced by the spatial-spectral attention mechanism, and the reconstruction of spatial-dimension features is conducted using spectral-dimension information. Finally, the features of various groups are fused, and the checkerboard pattern is relieved by applying the pixel deconvolution layer. As a result, clear and high-resolution images can be produced. The proposed super-resolution algorithm that combines multi-receptive field features with spectral-spatial attention was applied to two public datasets Chikusei and Pavia Center Scene, achieving peak signal-to-noise ratios of 39.869 7 and 31.942 2, respectively and structural similarity of 0.937 6 and 0.878 6, respectively. Therefore, the super-resolution algorithm enjoys obvious performance advantages compared to the latest super-resolution algorithms. Overall, the algorithm proposed in this study integrates the advantages of the multi-receptive field feature extraction module and the spatial-spectral attention module and can significantly improve image details.

Keywords: hyperspectral image; image super resolution; multi-receptive field feature extraction; attention mechanism

(责任编辑: 陈 理)