

# 跨尺度混合注意力的遥感图像超分辨率重建<sup>①</sup>



肖振久, 苏 婷, 曲海成, 翟宇琦

(辽宁工程技术大学 软件学院, 葫芦岛 125105)

通信作者: 苏 婷, E-mail: 775583699@qq.com

**摘 要:** 为了解决现有遥感图像超分辨率重建模型对长期特征相似性和多尺度特征相关性关注不足的问题, 提出了一种基于跨尺度混合注意力机制的遥感图像超分辨率重建算法. 首先提出了一个全局层注意力机制 (global layer attention, GLA), 利用层注意力机制加权融合不同层级的全局特征, 建模低分辨率与高分辨率图像特征间的长期依赖关系. 同时, 设计了跨尺度局部注意力机制 (cross-scale local attention, CSLA), 在多尺度的低分辨率特征图中寻找与高分辨率图像匹配的局部信息补丁, 并融合不同尺度的补丁特征, 以优化模型对图像细节信息的恢复能力. 最后, 提出一种局部信息感知损失函数来指导图像的重建过程, 进一步提高了重建图像的视觉质量和细节保留能力. 在 UC-Merced 数据集上的实验结果表明, 本文方法在 3 种放大倍数下的平均 PSNR/SSIM 优于大多数主流方法, 并在视觉效果方面展现出更高的质量和更好的细节保留能力.

**关键词:** 遥感图像; 超分辨率重建; 多尺度特征融合; 注意力机制; 特征相似性

引用格式: 肖振久, 苏婷, 曲海成, 翟宇琦. 跨尺度混合注意力的遥感图像超分辨率重建. 计算机系统应用, 2024, 33(6): 153-160. <http://www.c-s-a.org.cn/1003-3254/9525.html>

## Super-resolution Reconstruction of Remote Sensing Images with Cross-scale Hybrid Attention

XIAO Zhen-Jiu, SU Ting, QU Hai-Cheng, ZHAI Yu-Qi

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

**Abstract:** To address the inadequacy of existing remote sensing image super-resolution reconstruction models in long-term feature similarity and multi-scale feature relevance, this study proposes a novel remote sensing image super-resolution reconstruction algorithm based on a cross-scale hybrid attention mechanism. Initially, the study introduces a global layer attention (GLA) mechanism and employs layer-wise attention to weight and merge global features across different levels, thereby modeling the extended dependency between low-resolution and high-resolution image features. Concurrently, it designs a cross-scale local attention (CSLA) mechanism to identify and integrate local information patches in multi-scale low-resolution feature maps that correspond with high-resolution images, enhancing the model's ability to restore image details. Finally, the study proposes a local information-aware loss function to guide the image reconstruction process, further improving the visual quality and detail preservation of the reconstructed images. Experiments on UC-Merced datasets demonstrate that the proposed method outperforms most mainstream methods in terms of average PSNR/SSIM across three magnification factors and exhibits superior quality and detail preservation in visual results.

**Key words:** remote sensing image; super-resolution reconstruction; multi-scale feature fusion; attention mechanism; feature similarity

① 基金项目: 辽宁省高等学校基本科研项目 (LJKMZ20220699); 辽宁工程技术大学学科创新团队 (LNTU20TD-23)

收稿时间: 2023-12-20; 修改时间: 2024-01-23; 采用时间: 2024-01-29; csa 在线出版时间: 2024-04-19

CNKI 网络首发时间: 2024-04-23

遥感图像在气象观测、农林检测、灾害监测、海洋资源观测、环境监测以及军事侦察等领域中发挥着关键作用。然而,受到成像设备和数据传输信道等外部因素的限制,遥感图像的分辨率通常较低,这对基于这些图像的数据分析任务构成了挑战。鉴于从硬件方面提升遥感图像质量的成本较高,因此,采用图像超分辨率重建技术来提高遥感图像的分辨率和质量成为一个经济有效的解决方案。

## 1 相关研究

图像超分辨率重建的目标是通过特定算法将低分辨率 (low resolution, LR) 图像转换为包含更多高频信息和纹理细节的高分辨率 (high resolution, HR) 图像。传统的超分辨率重建方法<sup>[1-3]</sup>为解决遥感图像的超分辨率重建问题提供了有力的实践支持。随着技术的进步,研究重点逐渐从传统方法转向了基于深度学习的方法。例如, Dong 等人<sup>[4]</sup>在图像超分辨率任务中首次引入卷积神经网络 (CNN), 提出了 SRCNN (super-resolution convolutional neural network), 该网络通过 3 层 CNN 实现了低分辨率图像与高分辨率图像之间的端到端特征映射。为了扩大网络的感受野, Kim 等人<sup>[5]</sup>提出了 DRCN (deeply recursive convolutional network) 通过引入更多的卷积层来增加感受野的大小。然而, DRCN 网络在细节信息的感知质量方面还有提升的空间。Qin 等人<sup>[6]</sup>提出了 DGANet-ISE 网络, 模拟低分辨率和高分辨率图像之间的复杂关系, 从而提高了模型的细节恢复能力。

为了提高图像的感知质量, 李强等人<sup>[7]</sup>将生成对抗网络 (generative adversarial network, GAN) 应用到超分辨率重建任务中, 利用生成器和判别器相互协同对抗训练, 生成逼真的纹理。但是该网络重建的图像放大后的细节会产生伪影。Li 等人<sup>[8]</sup>提出了基于循环架构模型的 SR 反馈网络, 递归地对 LR 图像输入进行逐步细化, 以消除重建图像的伪影。

在遥感图像超分辨率重建的上述网络设计中, 细节信息的恢复对于不同实例而言具有统一的重要性。与注意力机制相关的技术<sup>[9,10]</sup>在此方面显示出优势。Dai 等人<sup>[11]</sup>开发了一种新的可训练二阶信道注意模块, 使用二阶特征统计来自适应地调整信道方向特征, 以获取更具鉴别性的局部表示。王庆庆等人<sup>[12]</sup>提出了一种高效的全局注意力机制, 设计了一个交叉自适应特征块来改善图像高频细节信息的缺失, 尽管这种方法

缺乏对中间特征的针对性处理。为了增强图像多尺度特征的表示能力, Lan 等人<sup>[13]</sup>设计了一个带有注意力机制的残差多尺度模块, 该方法利用多层特性并结合信道注意力, 但未充分考虑不同层特征之间的长期依赖关系和多尺度特征之间的相关性。

针对这些挑战, 本文提出了一种基于跨尺度混合注意力机制的超分辨率重建方法。具体来说, 首先通过一个残差组网络处理输入的低分辨率遥感图像, 以获得多个不同尺度的特征信息。基于此, 我们提出了一个全局层注意力机制 (GLA), 通过层注意力<sup>[14]</sup>对不同层级的全局特征进行加权融合, 捕捉图像长期特征的相似性和全局特征的关联性。随后, 我们引入了一个跨尺度局部注意力模块 (CSLA), 用于寻找与高分辨率图像局部信息最匹配的补丁信息, 通过加权融合得到的多尺度补丁特征来构建一个综合的高分辨率特征表示。在图像重建过程中, 引入了一种局部信息感知损失函数, 以约束模型在重建过程中尽可能地恢复细节信息, 同时对低频信息的保留不做过多限制。

## 2 本文方法

本文提出了一种基于跨尺度混合注意力机制的遥感图像超分辨率重建网络, 其整体结构如图 1 所示, 主要包含 3 个部分: 1) 全局层注意力模块, 提取图像的多尺度特征, 通过全局层注意力建模多层次特征的相互依赖性; 2) 跨尺度局部注意力模块, 从多个尺度特征图中提取高分辨率图像的补丁特征; 3) 重建模块, 在局部感知损失和全局对抗损失的引导下引导所提取补丁特征的加权融合。

### 2.1 全局层注意力

为了从低分辨率输入图像中提取丰富的特征表示, 本文设计了一个包含浅层特征提取网络和深层特征提取网络的特征提取架构。给定一个低分辨率输入图像  $I_{LR}$ , 首先经过一个  $3 \times 3$  的卷积层提取图像浅层特征  $F_0$ , 如式 (1) 所示:

$$F_0 = \text{Conv}(I_{LR}) \quad (1)$$

接着, 深度特征提取主要由多个残差组构成这些残差组负责提取低分辨率图像的中间特征  $F_i$ , 其公式如下所示:

$$F_i = H_{RB_i}(F_{i-1}), \quad i = 1, 2, \dots, N \quad (2)$$

其中,  $H_{RB_i}$  表示第  $i$  个残差组,  $N$  表示残差组的数量。

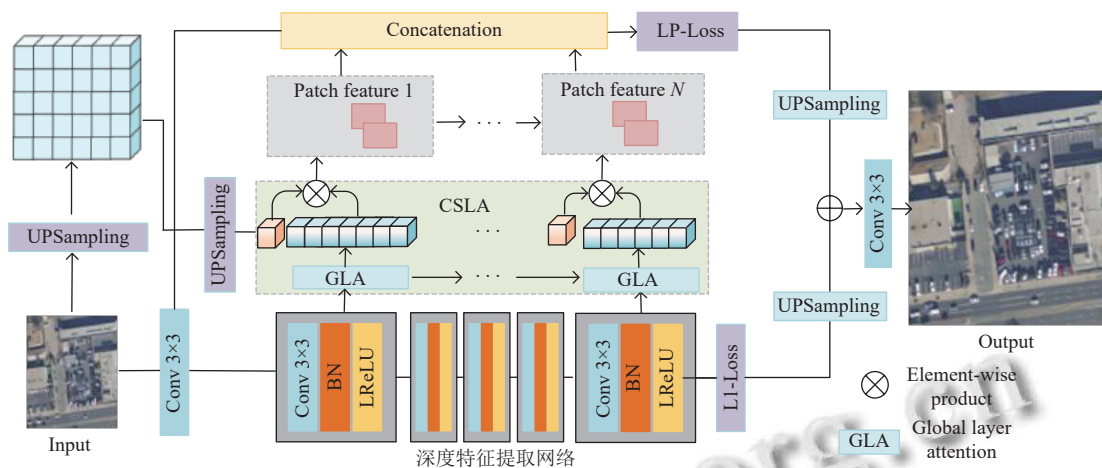


图1 模型结构图

残差组主要通过跳跃连接和密集连接的方式将低层级特征与高层级特征直接相连, 形成绕过中间层的捷径. 这种连接方式有助于从低层级的特征中获取更多细节信息. 然而, 这种直接连接方式无法充分捕获不同深度特征之间的相互依赖关系, 这可能导致模型对图像细节信息的提取不够充分. 为了解决这一局限性, 本研究在残差组结构后引入了全局层注意力模块 (GLA), 其具体结构如图2所示. GLA 模块首先将  $N$  个残差组模块提取的中间特征重构为  $N \times HWC$  维的 2D 矩阵, 并利用矩阵乘法与相应的转置来计算不同层之间的特征

相关性, 其公式如下所示:

$$w_{ji} = \delta(\beta(FG)_i \cdot (\beta(FG))_j^T), i, j = 1, 2, \dots, N \quad (3)$$

其中,  $\delta(\cdot)$  表示 Softmax;  $\beta(\cdot)$  表示重构操作;  $x_{i,j}$  表示第  $i$  个特征组和第  $j$  个特征组之间的相关性系数;  $FG$  表示特征组. 最后, 将重构的特征组与预测相关矩阵相乘, 并将输入特征相加, 其过程如下所示:

$$F_{L_j} = \alpha \sum_{i=1}^N w_{i,j} FG_i + FG_j \quad (4)$$

其中,  $\alpha$  初始值为 0, 之后由网络自动分配.

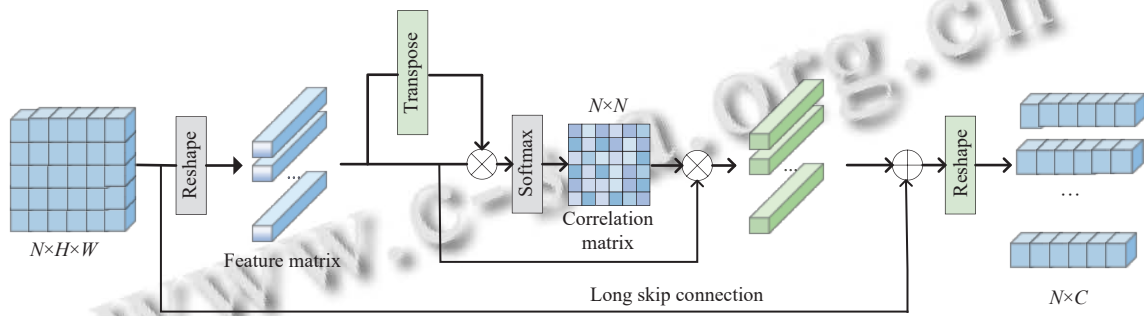


图2 GLA 模块

GLA 模块通过引入层注意力权重, 对不同层级的全局特征进行加权融合, 使得网络能够在不同层级之间灵活地选择和组合特征, 从而建模不同层次特征间的长期依赖性.

### 2.2 跨尺度局部注意力

在解决遥感图像超分辨率重建问题时, 以往的方法通常忽视了多尺度特征相关性的重要性. 为了克服这一问题, 本文提出了一种跨尺度局部注意力模块, 旨

在有效捕获多尺度特征的局部相关性, 以提高重建图像的细节恢复能力. 其具体结构如图3所示.

首先, 定义输入的多尺度特征  $F_i$  为图像在不同尺度下的信息. 随后, 局部信息搜索器的目标是在每个尺度的低分辨率特征图中搜索与高分辨率图像局部信息最匹配的补丁. 为了实现这一目标, 引入了局部注意力机制来获取每个像素点的局部信息表示. 具体而言, 对于  $H \times W \times C$  的输入特征, 基于局部注意力机制将水平

(X) 方向上每个通道的特征图进行平均池化, 垂直 (Y) 方向同样做平均池化操作, 并且 X 方向和 Y 方向需要并行进行池化操作. 具体计算如式 (5) 所示:

$$\begin{cases} z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} x_c(h, i) \\ z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w) \end{cases} \quad (5)$$

其中,  $x_c$  表示第  $c$  个信道上特征图;  $x_c(h, i)$  表示  $x_c$  在高度为  $h$  处的水平方向特征;  $x_c(j, w)$  表示  $x_c$  在宽度为  $w$  处的垂直方向特征;  $z_c^h(h)$  表示  $x_c$  经过一维平均池化后水平方向的张量;  $z_c^w(w)$  表示  $x_c$  经过一维平均池化后垂直方向的张量. 得到的结果与采样的局部特征进行像素级相似度计算, 此处通过减小采样特征与多尺度特征之间的  $L_1$  距离来减小它们之间的像素差异.  $L_1$  函

数可表示为:

$$L_1 = E_{x_i} \|x_f - x_i\|_1 \quad (6)$$

通过在不同尺度的低分辨率特征图上对局部区域进行局部注意力搜索操作, 得到一系列的局部补丁. 在补丁特征生成阶段, 对从局部搜索中输出的特征图进行拼接 (concat) 操作, 再经过 Softmax 层得到最终的输出. 其过程如式 (7) 所示:

$$f = \varepsilon(\text{Conv}_{1 \times 1}([\![Z^h, Z^w]\!]]) \quad (7)$$

其中,  $Z^h$  表示聚合的水平方向特征图;  $Z^w$  表示聚合的垂直方向特征图;  $[\![\cdot]\!]$  表示拼接 (concat) 操作;  $\varepsilon$  表示非线性函数;  $\text{Conv}_{1 \times 1}$  表示  $1 \times 1$  的卷积操作;  $f \in R^{C/r \times (H+W)}$  表示水平和垂直方向生成的中间特征,  $r$  表示缩减因子.

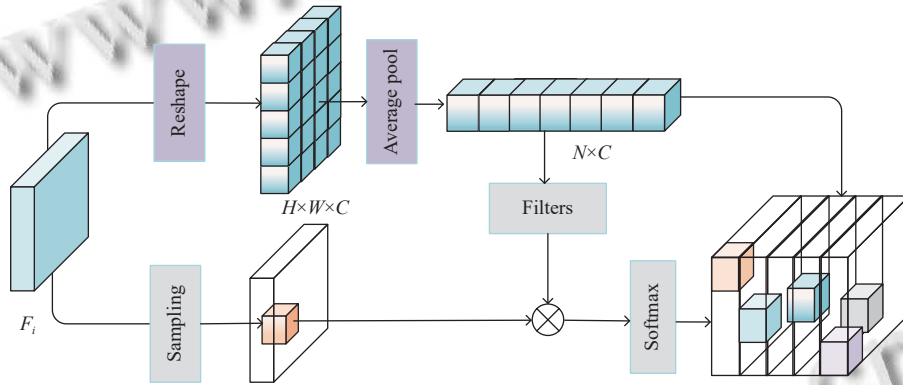


图3 CSLA 模块

沿着空间维数将  $f$  分解成两个单独的特征张量  $f \in R^{C/r \times H}$  和  $f \in R^{C/r \times W}$ . 接着对  $f \in R^{C/r \times H}$  和  $f \in R^{C/r \times W}$  的维度进行调整, 使通道数目  $g^h$  由  $C/r$  个变成  $C$  个. 再通过 Sigmoid 激活函数的作用, 得到水平方向的注意力权重和垂直方向的注意力权重  $g^w$ . 以上过程如式 (8) 所示:

$$\begin{cases} g^h = \rho(\text{Conv}_{1 \times 1}(f^h)) \\ g^w = \rho(\text{Conv}_{1 \times 1}(f^w)) \end{cases} \quad (8)$$

其中,  $\rho$  表示 Softmax 函数.

最后将第  $c$  通道特征图在对应的坐标位置与  $g^h$  和  $g^w$  的第  $c$  通道的特  $g_c^h$  征张量  $g_c^w$  和相乘, 得到的最终输出补丁特征图  $p_c$ , 如式 (9) 所示:

$$p_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (9)$$

在局部信息搜索器得到各尺度的局部信息表示后,

接下来是特征融合器的工作. 特征融合器的目标是将不同尺度的补丁特征进行加权融合, 得到一个综合的高分辨率特征表示. 具体而言, 采用自适应加权融合策略, 通过学习每个尺度补丁特征的权重, 将它们进行线性加权融合. 假设  $\theta$  表示局部搜索得到的最佳特征补丁,  $\delta$  表示对应上采样得到 HR 特征, 则综合的高分辨率特征表示  $\psi$  可以表示为:

$$\varphi(X_{i,j}, P_{g,h}) = \theta(X_{i,j})^T \delta(P_{g,h}) \quad (10)$$

通过这种方式, 使得模型能够更加关注于细节特征的恢复, 并提高了模型的细节特征表示能力. 通过在通道级别执行特征加权操作, 模型能够更好地适应复杂的数据分布和任务要求, 从而提高模型性能. 通过融合每个尺度的补丁特征, 模型的图像细节信息恢复能力得到了最大化.

### 2.3 图像重建

重建部分首先采用亚像素卷积作为上采样模块,通过对特征图的每个像素块进行插值,并将像素值分配到高分辨率图像的对应位置,从而实现图像的细节恢复.利用 Conv 3×3 层对多通道的特征图进行卷积操作,生成最终的三通道重建图像.其公式如下所示:

$$I_{SR} = U_{\uparrow}(F_0 + F_i + F_L) \quad (11)$$

其中,  $I_{SR}$  表示重建图像;  $U_{\uparrow}$  表示上采样.

重建过程中为了确保重建图像与原始高分辨率图像在局部区域与补丁特征的一致性,设计了一个局部感知损失函数用于计算重建图像特定区域与对应的补丁特征之间的特征差异,具体过程可表示为式(12):

$$L_{LP} = \sum_{i,j} \left[ y_{(i,j)} \times d_{(i,j)}^2 + (1 - y_{(i,j)}) \times \max(0, m - d_{(i,j)})^2 \right] \quad (12)$$

其中,  $y_{(i,j)}$  表示两个特征是否长期依赖,  $d(i, j)$  表示两个特征的余弦距离,  $m$  表示一个边界值.

可以根据图像的特征差异进行调整,内容损失函数使用  $L_1$  损失.最终,总的损失函数可以表达为:

$$L_{total} = \lambda L_{LP} + \eta L_1 \quad (13)$$

其中,  $\lambda$  和  $\eta$  为损失函数的加权系数.

## 3 实验结果分析

### 3.1 实验设置

对于跨尺度局部注意力,设置 patch 大小  $p=3$  和  $s=1$  进行密集采样,使用  $3 \times 3$  作为所有卷积层的滤波器大小(注意力中内核大小为  $1 \times 1$ ).提取到的所有中间特征通道  $C=128$ .训练样本统一采用了随机旋转  $90^\circ$ 、 $180^\circ$  和水平翻转进行数据增强.模型架构基于 PyTorch 框架实现,在一台 Nvidia RTX3090 服务器上进行训练.为了优化模型,使用了 Adam 作为训练优化器,初始学习速率被设置为  $1E-4$ ,每 100 个 epochs 减少  $1/10$ .训练停止在 500 epochs.

### 3.2 数据集和评价指标

UC-Merced 数据集<sup>[15]</sup>: UC-Merced 数据集由美国地质调查局的国家地图城区图像集合中手动提取而成,涵盖了 20 个城市,包含 21 个场景类别(如农业、棒球场、密集住宅等).每个类别包含 100 张  $256 \times 256$  像素的图像,空间分辨率为 0.3 m/像素.将数据集分为训练集和测试集,其中 20% 的训练集作为验证集.

评价指标:实验同时采用两种常用的超分辨率重建效果的评价指标.峰值信噪比(peak signal-to-noise ratio, PSNR)和结构相似性(structural similarity, SSIM).PSNR 的值越高表示图像的质量越好,SSIM 的值越接近 1 说明两幅图像越相似.对于可视化结果,以人类主观视觉感知作为定性评估标准.

### 3.3 消融实验

为了验证本文提出的各个组件在遥感图像超分辨率重建中的有效性,我们在 UC-Merced 数据集上进行了一系列的消融实验.实验目的是探究 GLA 模块、CSLA 模块以及局部感知损失(LP-Loss)对 2、3 和 4 倍超分辨率重建效果的影响,以此评估每个组件的重要性.所有模型均在相同的实验设置下进行训练以确保公平性.实验结果如表 1-表 3 所示.

表 1 不同模块组合下的 UC-Merced 测试集上的客观评价指标(×2)

组合	PSNR (dB)	SSIM
HSENet	34.22	0.933
HSENet+GLA	34.53	0.941
HSENet+CSLA	34.56	0.946
HSENet+LP-Loss	34.38	0.937
HSENet+GLA+CSLA	34.62	0.952
HSENet+GLA+LP-Loss	34.57	0.947
HSENet+CSLA+LP-Loss	34.59	0.950
HSENet+CSLA+GLA+LP-Loss	34.72	0.954

表 2 不同模块组合下的 UC-Merced 测试集上的客观评价指标(×3)

组合	PSNR (dB)	SSIM
HSENet	30.00	0.842
HSENet+GLA	30.24	0.851
HSENet+CSLA	30.30	0.856
HSENet+LP-Loss	30.07	0.847
HSENet+GLA+CSLA	30.37	0.861
HSENet+GLA+LP-Loss	30.29	0.857
HSENet+CSLA+LP-Loss	30.31	0.860
HSENet+CSLA+GLA+LP-Loss	30.42	0.867

表 3 不同模块组合下的 UC-Merced 测试集上的客观评价指标(×4)

组合	PSNR (dB)	SSIM
HSENet	27.73	0.762
HSENet+GLA	27.89	0.771
HSENet+CSLA	27.93	0.779
HSENet+LP-Loss	27.80	0.768
HSENet+GLA+CSLA	27.99	0.787
HSENet+GLA+LP-Loss	27.91	0.775
HSENet+CSLA+LP-Loss	27.94	0.783
HSENet+CSLA+GLA+LP-Loss	28.05	0.792

从表1-表3可以看出,从结果中可以看出,单独引入GLA或CSLA模块时,两项客观指标均有所提升.这表明GLA和CSLA模块对提高模型重建效果是有效的.特别是当同时引入GLA和CSLA模块时,两项指标的值比单独引入GLA或CSLA时更高,说明这两个模块的组合在提取深度特征上具有相互促进的作用.而当只引入LP-Loss时,尽管两项指标有所提升,但提升效果并不明显.当同时引入GLA和LP-Loss或CSLA和LP-Loss时,效果接近于只引入GLA+CSLA模块,没有显著差异.但当3个模块(GLA、CSLA和LP-Loss)一起引入时,两项指标值相较于不添加任何模块时均有显著提升,这进一步验证了改进模型的有效性.

### 3.4 与现有方法的对比

为了进一步验证本文模型的有效性,将本文方法与现有的几种主流超分辨率重建方法进行对比,包括

Bicubic<sup>[1]</sup>、SRCNN<sup>[4]</sup>、FSRCNN<sup>[16]</sup>、LGCNet<sup>[17]</sup>、DCM<sup>[18]</sup>、DGANet-ISE<sup>[6]</sup>和HSENet<sup>[19]</sup>这7种超分辨率方法(放大因子为2、3和4).表4展示了这些方法在UC-Merced数据集上放大因子为2、3和4时的PSNR和SSIM指标.结果显示,本文方法在3种放大因子下的平均PSNR/SSIM值比HSENet提升了0.41 dB/0.025,与其他6种方法相比也均有明显提升.体现了本文方法的有效性.

表5为不同方法在所有21类UC-Merced数据集上放大因子为3时的PSNR指标.由于DGANet-ISE没有报告在放大因子为3时的表现,所以没有列出其重建结果.从表5中结果可以看出,本文方法在15个UC-Merced场景类别中获得了最佳的PSNR结果.HSENet在其他6个类别中获得了最佳的PSNR结果.与HSENet相比,本文方法的重建效果更好.

表4 在UC-Merced数据集上不同算法的客观评价指标

Scale	Index	Bicubic	SRCNN	FSRCNN	LGCNet	DCM	DGANet-ISE	HSENet	Ours
×2	PSNR (dB)	30.76	32.84	33.18	33.48	33.65	33.68	34.22	34.72
	SSIM	0.879	0.915	0.920	0.924	0.927	0.934	0.933	0.954
×3	PSNR (dB)	27.46	28.66	29.09	29.28	29.52	—	30.00	30.42
	SSIM	0.763	0.804	0.817	0.824	0.839	—	0.842	0.867
×4	PSNR (dB)	25.65	26.78	26.93	27.02	27.22	27.31	27.73	28.05
	SSIM	0.673	0.722	0.727	0.733	0.753	0.767	0.762	0.792

表5 UC-Merced数据集上每个场景类别的平均PSNR(×3)(dB)

类别	Bicubic	SRCNN	FSRCNN	LGCNet	DCM	HSENet	Ours
农业	26.86	27.47	27.61	27.66	29.06	27.64	<b>29.58</b>
海滩	26.71	28.24	28.98	29.12	30.77	30.09	<b>32.30</b>
棒球场	33.33	34.33	34.64	34.72	33.76	<b>35.05</b>	32.96
飞机	36.14	37.00	37.21	37.37	36.38	37.69	<b>38.25</b>
建筑群	25.09	26.84	27.50	27.81	28.51	28.95	<b>30.74</b>
丛林	25.21	26.11	26.21	26.39	26.81	26.70	<b>28.52</b>
森林	25.76	27.41	28.02	28.25	28.79	<b>29.24</b>	25.71
高速公路	27.53	28.24	28.35	28.44	28.16	28.59	<b>30.22</b>
高尔夫球场	27.36	28.69	29.27	29.52	30.45	30.63	<b>32.17</b>
海港	35.21	36.15	36.43	36.51	34.43	36.62	<b>37.53</b>
道路	21.25	22.82	23.29	23.63	26.55	24.88	<b>26.39</b>
中型住宅	26.48	27.67	28.06	28.29	29.28	29.21	<b>31.26</b>
移动家庭公园	25.68	27.06	27.58	27.76	27.21	<b>28.55</b>	26.43
停车场	22.25	23.89	24.34	24.59	26.05	25.70	<b>27.95</b>
河流	24.59	25.65	26.53	26.58	27.77	<b>28.22</b>	25.41
跑道	21.75	23.11	23.34	23.69	24.95	24.66	<b>26.57</b>
稀疏住宅	28.12	28.89	29.07	29.12	28.89	<b>29.22</b>	27.60
网球场	29.30	30.61	31.01	31.15	32.53	31.15	<b>33.52</b>
储油罐	28.34	29.40	30.23	30.53	29.81	31.64	<b>32.34</b>
密集住宅	29.97	31.33	31.92	32.17	29.02	<b>32.95</b>	30.51
立交桥	29.75	30.62	31.34	31.58	30.76	32.71	<b>32.95</b>
平均值	27.46	28.23	29.09	29.28	29.52	30.00	<b>30.42</b>

本文方法的重建效果如图4所示,可以观察到,相较于LR图像,本文方法生成的图像能够更清晰地展现出一些细节信息,如纹理和边缘等.甚至在图像的某些区域,本文方法能够恢复出与HR图像相当的纹理和边缘等信息.

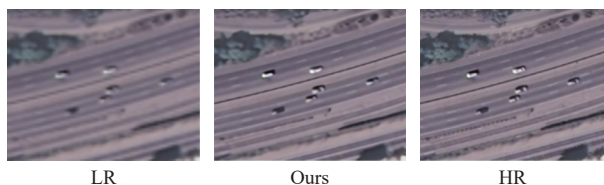


图4 本文方法的重建结果

为了更直观地展示本文方法的有效性,定性地展示了本文模型以及其他几种对比模型的重建结果.由于篇幅原因,这里只列出UC-Merced数据集在放大因子为3和4时的重建结果.从图5和图6中可以看出,Bicubic方法的重建效果模糊,其他几种超分辨率模型的重建效果明显优于Bicubic方法.本文方法重建效果优于HSENet方法,重建后的图像能够保留更多的细节.总的来说,与其他方法相比,本文方法能够获得更清晰的边缘和细节,从而获得更好的重建结果.

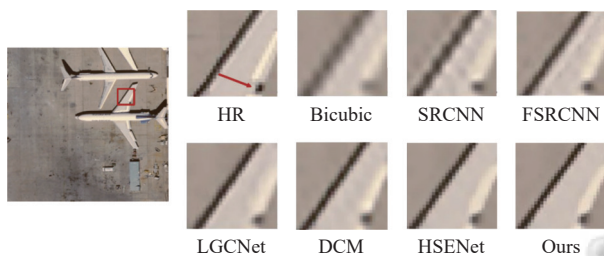


图5 不同方法在UC-Merced数据集上的结果比较( $\times 3$ )



图6 不同方法在UC-Merced数据集上的结果比较( $\times 4$ )

#### 4 结论与展望

本文提出了一种基于跨尺度混合注意力机制的超分辨率重建方法,旨在解决遥感图像重建中存在的长

期特征依赖性和多尺度相关性不足的问题.通过引入全局层注意力机制和跨尺度局部注意力模块,更好地捕捉图像的长期特征相似性和局部细节信息,并将它们有效地融合到最终的高分辨率重建结果中.此外,提出了一种局部信息感知损失约束模型,用于在重建过程中平衡高频细节信息的恢复和低频信息的保留.在实验验证中,基于公共数据集对本文方法进行了评估,并与其他现有方法进行了比较.结果显示,在不同倍数的超分辨率重建任务中,本文方法在PSNR/SSIM值上均优于其他方法,并且在视觉效果上展现出更高的细节信息质量.本文提出的基于跨尺度混合注意力机制的超分辨率重建方法在遥感图像处理领域具有重要的理论和实用价值.未来的研究可以进一步拓展该方法在其他领域的应用,并探索更加高效和精确的超分辨率重建算法,以满足不断增长的遥感数据处理需求.

#### 参考文献

- Keys R. Cubic convolution interpolation for digital image processing. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1981, 29(6): 1153–1160. [doi: 10.1109/TASSP.1981.1163711]
- Sun J, Zhu JJ, Tappen MF. Context-constrained hallucination for image super-resolution. *Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Francisco: IEEE, 2010. 231–238.
- Yang SY, Wang ZY, Zhang L, *et al.* Dual-geometric neighbor embedding for image super resolution with sparse tensor. *IEEE Transactions on Image Processing*, 2014, 23(7): 2793–2803. [doi: 10.1109/TIP.2014.2319742]
- Dong C, Loy CC, He KM, *et al.* Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(2): 295–307. [doi: 10.1109/TPAMI.2015.2439281]
- Kim J, Lee JK, Lee KM. Deeply-recursive convolutional network for image super-resolution. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016. 1637–1645.
- Qin MJ, Mavromatis S, Hu LS, *et al.* Remote sensing single-image resolution improvement using a deep gradient-aware network with image-specific enhancement. *Remote Sensing*, 2020, 12(5): 758. [doi: 10.3390/rs12050758]
- 李强, 汪西原, 何佳玮. 基于生成对抗网络的遥感图像超分辨率重建改进算法. *激光与光电子学进展*, 2023, 60(10): 1028010.

- 8 Li Z, Yang JL, Liu Z, *et al.* Feedback network for image super-resolution. Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 3867–3876.
- 9 Zhang YL, Li KP, Li K, *et al.* Image super-resolution using very deep residual channel attention networks. Proceedings of the 15th European Conference on Computer Vision. Munich: Springer, 2018. 286–301.
- 10 Yang X, Fan JF, Wu CH, *et al.* NasmamSR: A fast image super-resolution network based on neural architecture search and multiple attention mechanism. Multimedia Systems, 2022, 28(1): 321–334. [doi: [10.1007/s00530-021-00841-2](https://doi.org/10.1007/s00530-021-00841-2)]
- 11 Dai T, Cai JR, Zhang YB, *et al.* Second-order attention network for single image super-resolution. Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 11065–11074.
- 12 王庆庆, 辛月兰, 赵佳, 等. 高效全局注意网络的图像超分辨率重建. 激光与光电子学进展, 2024, 61(10): 1010006.
- 13 Lan RS, Sun L, Liu ZB, *et al.* MADNet: A fast and lightweight network for single-image super resolution. IEEE Transactions on Cybernetics, 2021, 51(3): 1443–1453. [doi: [10.1109/TCYB.2020.2970104](https://doi.org/10.1109/TCYB.2020.2970104)]
- 14 Niu B, Wen WL, Ren WQ, *et al.* Single image super-resolution via a holistic attention network. Proceedings of the 16th European Conference on Computer Vision. Glasgow: Springer, 2020. 191–207.
- 15 Yang Y, Newsam S. Bag-of-visual-words and spatial extensions for land-use classification. Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems. San Jose: ACM, 2010. 270–279.
- 16 Dong C, Loy CC, Tang XO. Accelerating the super-resolution convolutional neural network. Proceedings of the 14th European Conference on Computer Vision. Amsterdam: Springer, 2016. 391–407.
- 17 Lei S, Shi ZW, Zou ZX. Super-resolution for remote sensing images via local-global combined network. IEEE Geoscience and Remote Sensing Letters, 2017, 14(8): 1243–1247. [doi: [10.1109/LGRS.2017.2704122](https://doi.org/10.1109/LGRS.2017.2704122)]
- 18 Haut JM, Paoletti ME, Fernandez-Beltran R, *et al.* Remote sensing single-image superresolution based on a deep compendium model. IEEE Geoscience and Remote Sensing Letters, 2019, 16(9): 1432–1436. [doi: [10.1109/LGRS.2019.2899576](https://doi.org/10.1109/LGRS.2019.2899576)]
- 19 Lei S, Shi ZW. Hybrid-scale self-similarity exploitation for remote sensing image super-resolution. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5401410.

(校对责编: 孙君艳)