文章编号:2096-1472(2023)11-0005-06

DOI:10.19644/j.cnki.issn2096-1472.2023.011.002

# 融合多尺度信息和混合注意力的遥感图像重建

曹春萍,李 昂



摘 要:针对使用单一尺寸的卷积核重建遥感图像效果较差的问题,提出了融合多尺度信息和混合注意力的遥 感图像重建模型。模型使用两种不同的多尺度特征提取块,能有效提取不同感受野下的特征图中的高频和低频特 征,并通过混合注意力机制自适应地调整多尺度特征的权重,利用重建模块重建高清遥感图像。在放大因子为2 时,在 NWPU-RESISC45 和 UCMerced-LandUse 两个测试集上得到的 PSNR(峰值信噪比)和 SSIM(结构相似性)分 别为 37.720 4 dB、37.999 6 dB 和 0.962 1、0.965 4,均优于 DSSR、IRN 和 MPSR 等先进的遥感图像超分辨率重建的 模型,证明了所设计模型的有效性和鲁棒性。

关键词:超分辨率重建;深度学习;多尺度特征;混合注意力机制 中图分类号:TP391 文献标志码:A



# Remote Sensing Image Reconstruction with Multi-scale Information and Hybrid Attention



(School of Optical-Electrical & Computer Engineering, University of Shanghai for Science & Technology, Shanghai 200093, China) ⊠ ccpgcd@163.com; leon\_usst\_365@163.com

**Abstract:** This paper proposes a remote sensing image reconstruction model that integrates multi-scale information and hybrid attention, to address the problem of poor performance in reconstructing remote sensing images using a single size convolutional kernel. Two different multi-scale feature extraction blocks are used in the model, which can effectively extract high-frequency and low-frequency features in feature maps under different receptive fields. The weight of multi-scale features is adaptively adjusted through a hybrid attention mechanism, and the reconstruction module is used to reconstruct high-definition remote sensing images. When the amplification factor is 2, the PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity) obtained on the NWPU-RESISC45 and UCMerced-LandUse test sets are 37.720 4 dB, 37.999 6 dB and 0.962 1, 0.965 4, respectively, which are superior to the advanced models for super-resolution reconstruction of remote sensing images such as DSSR, IRN, and MPSR, demonstrating the effectiveness and robustness of the proposed model.

Key words: super-resolution reconstruction; deep learning; multi-scale feature; hybrid attention mechanism

# 0 引言(Introduction)

遥感图像是通过使用遥感技术(包括飞机、火箭、卫星或其 他空间平台)从空中远距离对地面目标物进行观测而获取到的 图像。目前,遥感图像数据在许多领域都得到有效应用,例如 农林业资源调查、环境监测、灾害监测和评估、城市规划等。然 而,在遥感图像获取的过程中会不可避免地受到外在因素的影 响,包括大气湍流、通道传输能力限制、远距离成像困难等,导 致获取的遥感图像存在分辨率比较低的问题。综合考虑硬件 成本,可以使用遥感图像超分辨率技术获取到更高质量和分辨 率的遥感图像。

# 1 相关技术概述(Overview of related technologies)

传统的遥感图像超分辨率技术有基于插值的方法<sup>[1]</sup>、基于 重构的方法<sup>[2]</sup>和基于深度学习的方法<sup>[3]</sup>。基于插值的方法虽 然计算复杂度较低,但是在进行遥感图像重建时,只能恢复图 像的低频信息,图像效果较差。基于重构的方法通常需要足够 的先验知识,而且重建速度较慢,在先验知识不足的情况下重 建效果不理想。基于深度学习的方法从低分辨率图像提取特 征图,再将特征图与高分辨率图像之间建立映射,重建高分辨 率图像,通过这种方法重建的图像在客观评价指标和视觉效果 方面均明显优于基于插值的方法和基于重构的方法。

DONG 等<sup>[4]</sup>率先提出了经典的超分辨率卷积神经网络 (Super-Resolutional Convolution Neural Networks, SRCNN), 利用卷积神经网络(Convolutional Neural Networks, CNN)学 习低分辨率(Low-Resolution, LR)图像与高分辨率(High-Resolution, HR)图像进行特征映射。随后, DONG等<sup>[5]</sup>在 SRCNN 的基础上提出了 FSRCNN(Fast Super-Resolutional Convolution Neural Networks)模型,增加了模型的网络层数, 有效重建出更多的高频细节,但模型的训练难度也有所增加。 KIM 等<sup>[6]</sup>提出深度卷积网络超分模型(Very Deep Super-Resolutional Neural Networks, VDSR),利用多层 CNN 连接实 现特征级联,大大提升了学习率,加快了网络收敛速度,证明了 超分模型的网络深度对超分辨率重建的重要性。随着网络深 度的增加,网络在训练过程中会出现梯度消失和梯度爆炸等问 题,研究者们为了避免此问题的发生,开始钻研新的结构。HE 等<sup>[7]</sup>提出残差网络(Residual Network, ResNet),通过在网络 构中引入残差单元,避免了梯度问题,保证了模型的顺利收敛, 为后续的研究者提供了搭建深层卷积网络的思路 ALIM 等 提出 EDSR(Enhanced Deep Super-Resolution Network),通过 堆叠残差块构建了更深的网络结构,重建出质量更高的图像。 随后,在遥感图像超分辨率重建中,人们不断构建出越来越深 的卷积神经网络用于捕捉更多图像的深层惊息。然而随着网 络层数的不断加深,超分重建再次进入 、瓶颈,研究者们发现 网络到达一定的深度后,想通过增加网络层数提升网络性能, 效果微乎其微。为了重建出更好的图像,研究者继续探索新的 方法。当前,大多数模型对待通道特征是平等的,然而实际上 各通道特征对图像重建的重要性并不相同,因此充分利用通道 特征成为提升重建图像质量的新的突破点。HU 等<sup>[9]</sup>通过在 网络中加入压缩-激励(Squeeze-and-Excitation,SE)模块构建压 缩-激励网络(Squeeze-and-Excitation Network,SENet),计算不 同通道之间的权重,将通道注意力用于图像处理中,重建出更 优质的图像。受通道注意力的启发,WOO等<sup>[10]</sup>在网络中加入 通道注意力和空间注意力,并将两者进行融合,提出了瓶颈注 意力模块(Bottleneck Attention Module, BAM)和卷积注意力 模块(Convolutional Block Attention Module, CBAM), ZHANG 等<sup>[11]</sup>在残差块中融入通道注意力,提出残差通道注意力网络 (Residual Channel Attention Network, RCAN), WOO 等和 ZHANG 等提出的模块和网络模型都取得了更好的重建效果。

当前的遥感图像重建模型大多使用单一尺寸的卷积核,在 应用于遥感图像的重建提取特征时,由于遥感图像目标物体尺 寸差异较大,学习能力有限,捕捉到的特征不足,因此无法重建 出高质量的图像。

为了解决上述问题,本文提出融合多尺度信息和混合注意 力网络(Fusion of Multi-Scale Information and Hybrid Attention Networks,FMSIHAN),设计了两种多尺度特征提取块,通过 多尺度信息融合和混合注意力块(Multi-Scale Information Fusion and Hybrid Attention Block,MSIFHAB)和多尺度类金 字塔特征提取块(Multi-Scale Pyramid-Like Feature Extraction Block,MSPLFEB)提取多尺度特征,并通过 MSIFHAB 中的混 合注意力块对提取的多尺度特征自适应地分配权重,有助于捕 捉图像的高频和低频信息,进而获得更好的重建效果。

#### 2 网络结构(Network structure)

融合多尺度信息和混合注意力网络由浅层特征提取块、多 尺度特征提取块(Multi-Scale Feature Extraction Block, MSFEB)、全局特征融合块和重建模块组成,网络总体结构如 图1所示。



Fig. 1 Overall network architecture

# 2.1、浅层特征提取块

浅层特征提取块为一个 3×3 的卷积层,用来提取输入 LR 遥感图像的浅层特征,并将特征图的通道数由 3 变为 64,这个 过程可以表示如下:

$$F_0 = f_{\text{Conv3} \times 3}(I_{\text{LR}}) \tag{1}$$

其中: $I_{LR}$ 代表模型输入的低分辨率遥感图像; $f_{Conv3\times3}(\cdot)$ 代表卷积操作,用来提取 $I_{LR}$ 的浅层特征; $F_0$ 代表 $I_{LR}$ 经过浅层特征提取块提取到的浅层特征图。

## 2.2 多尺度特征提取块

每个 MSFEB 由若干个多尺度信息融合和混合注意力块、 若干个多尺度类金字塔特征提取块组成,多尺度特征提取块结 构图如图 2 所示。输入特征图经过多尺度特征提取块的过程 可以表示如下:

 $F_{i} = f_{\text{MSPLFEB}} \left( f_{\text{MSIFHAB}} \left( f_{\text{MSIFHAB}} \left( f_{\text{MSIFHAB}} \left( f_{\text{MSIFHAB}} \left( f_{\text{MSIFHAB}} \left( f_{i-1} \right) \right) \right) \right) \right)$  (2)

其中: $F_{i-1}$ 代表输入多尺度特征提取块的特征图; $f_{MSPLFEBI}(\cdot)$ 、  $f_{MSPLFED2}(\cdot)$ 、 $f_{MSPLFED3}(\cdot)$ 代表多尺度类金字塔特征提取块 的多尺度特征提取操作; $f_{MSIFHABI}(\cdot)$ 、 $f_{MSIFHAB2}(\cdot)$ 代表多 尺度信息融合和混合注意力块的多尺度特征提取操作; $F_i$ 代 表经过多尺度特征提取块后输出的特征图。





#### 2.2.1 多尺度信息融合和混合注意力块

针对大多数图像,传统的卷积神经网络通常采用单一的卷 积层提取图像特征,重建得到的高分辨率图像质量可以达到使 用者的要求。然而,通过卫星等遥感技术获取的遥感图像,其 内部包含的物体尺寸差异较大,在对这类图像进行超分辨率重 建时,若使用较小尺寸的卷积核提取特征,由于感受野较小,所 以对高频信息的提取能力不足;若使用较大尺寸的卷积核提取 特征,将丢失一部分图像像素信息。也就是说,使用单一尺寸 的卷积核进行特征提取时,学习能力有限,无法重建出高质量 的图像。因此,本文设计了多尺度信息融合和混合注意力块, 通过使用卷积层、下采样操作和平均池化层提取特征图中的高 频和低频特征,再通过混合注意力机制,自适应地调整多尺度 特征的权重。多尺度信息融合和混合注意力块结构图如图 3 所示,图 3 中的 W、H、C 分别代表特征图的宽度、高度、通 道数。



图 3 多尺度信息融合和混合注意力块结构图 Fig. 3 Architecture of MSIFHAB

构建多尺度信息融合和混合注意力块的步骤所示。 首先,将输入特征图分别送入三个不同的卷积层,其中第 一个卷积层从下采样操作后的特征图中提取低频特征;第二个 卷积层从原始尺寸的特征图中经过平均池化层提取低频特征; 第三个卷积层从原始尺寸的特征图中提取高频特征。经过下 采样操作得到的特征图经过反卷积恢复到下采样操作前的原 始尺寸。经过三个不同的卷积操作后,得到三种尺度不同的特 征信息,这个过程可以表示如下;

 $F_{C} = \operatorname{ReLU}(T_{\operatorname{Conv4}\times4}(\operatorname{Downs}(F_{\operatorname{MSFin}}))) + \operatorname{Avg}_{\operatorname{Conv3}\times3}(F_{\operatorname{MSFin}}) + \operatorname{ReLU}(f_{\operatorname{Conv3}\times3}(F_{\operatorname{MSFin}}))$ (3)

其中: $F_{MSLin}$ 代表输入特征图; Downs(•)代表下采样操作;  $T_{Conv4\times4}(•)$ 代表反卷积操作; Avg<sub>Conv3×3</sub>(•)代表平均池化操作;  $F_c$  为输出的多尺度特征图。

其次,为了充分利用多尺度特征信息,本文提出混合注意 力机制,混合注意力块(Hybrid Attention Block,HAB)结构图 如图 4 所示,通过对不同的尺度特征自适应地分配不同的权 重,使网络能学习到对图像重建更有价值的特征,这个过程可 以表示如下:

 $F_{s} = [CA(F_{c}) + F_{c}] + SA(CA(F_{c})) + [SA(F_{c}) + CA(F_{c})]$ (4)

其中:CA(•)代表通道注意力机制,与 CBAM 中的通道注意 力保持一致;SA(•)代表空间注意力机制,和 CBAM 中的空间 注意力保持一致;F<sub>s</sub>代表输出的加权多尺度特征图。



Fig. 4 Architecture of HAB

最后,通过一个3×3的卷积层将加权后得到的特征图和 加权前的多尺度特征图融合为多尺度信息融合特征图,并且通 过跳跃连接避免模型在训练过程中出现梯度消失等问题,这个 过程可以表示如下:

$$F_{MSFout} = F_{MSFin} + F_C + f_{Conv3\times3}(F_S)$$
(5)  
其中,  $F_{MSFout}$  为输出的多尺度信息融合特征图。

#### 2.2.2 多尺度类金字塔特征提取块

多尺度类金字塔特征提取块结构图如图 5 所示, K 为卷积 核尺寸, N 为输入通道数, D 为扩张率。多尺度类金字塔特征 提取块由四个部分组成, 第一部分包含四个不同感受野的卷积 层, 第二部分包含三个不同感受野的卷积层, 第三部分是一个 将通道数恢复到 64 的1×1的卷积层, 第四部分是一个跳跃 连接。



图 5 多尺度类金字塔特征提取块结构图 Fig. 5 Architecture of MSPLFEB

构建多尺度类金字塔特征提取块的步骤如下。

首先,将输入特征图分别送入多尺度类金字塔特征提取块 第一部分的四个不同感受野的卷积层,这四个卷积层的输入通 道数为 64,输出通道数为 16,卷积核尺寸分别为 1×1、3×3、 3×3 和 3×3,扩张率分别为 1、1、2 和 3,这个过程可以表示 如下:

 $F_{\text{MSP1}} = \operatorname{ReLU}(f_{\text{Conv1}\times 1, d1}^{64 \rightarrow 16}(F_{\text{MSP-in}})) + \operatorname{ReLU}(f_{\text{Conv3}\times 3, d1}^{64 \rightarrow 16}(F_{\text{MSP-in}})) +$ 

ReLU( $f_{Conv3\times3,d2}^{6i+16}(F_{MSP:in})$ )+ReLU( $f_{Conv3\times3,d2}^{6i+16}(F_{MSP:in})$ )(6) 其中: $F_{MSP:in}$ 代表输入特征图; $f_{Conv3\times1,d1}^{6i+16}$ (•)、 $f_{Conv3\times3,d1}^{6i+16}$ (•)、  $f_{Conv3\times3,d2}^{6i+16}$ (•)和 $f_{Conv3\times3,d3}^{6i+16}$ (•)分别代表输入通道数为 64,输 出通道数为 16,卷积核尺寸分别为 1×1、3×3、3×3 和 3×3, 扩张率分别为 1、1、2 和 3 的卷积操作; $F_{MSP1}$ 代表经过第一部 分输出的多尺度特征图。

其次,将 F<sub>MSP1</sub>分别送入多尺度类金字塔特征提取块第二 部分的三个不同感受野的卷积层,这三个卷积层的输入通道数 为16,输出通道数为32,卷积核尺寸分别为1×1、3×3和3× 3,扩张率分别为1、1和2,这个过程可以表示如下:

$$F_{MSP2} = \text{ReLU}(f_{Conv1\times1,d1}^{10\to32}(F_{MSP1})) + \text{ReLU}(f_{Conv3\times3,d1}^{10\to32}(F_{MSP1})) + \text{ReLU}(f_{Conv3\times3,d2}^{10\to32}(F_{MSP1}))$$
(7)

其中: $f_{Conv1\times1,d1}^{16\to32}(\cdot)$ 、 $f_{Conv3\times3,d1}^{16\to32}(\cdot)$ 和 $f_{Conv3\times3,d2}^{16\to32}(\cdot)$ 分别代表 输入通道数为16,输出通道数为32,卷积核尺寸分别为1×1、 3×3和3×3,扩张率分别为1、1和2的卷积操作; $F_{MSP2}$ 代表 经过多尺度类金字塔特征提取块第二部分输出的多尺度特 征图。

最后,将 F<sub>MSP2</sub> 送入多尺度类金字塔特征提取块第三部分的卷积层,并通过跳跃连接将第三部分的输出结果与多尺度类 金字塔特征提取块的输入特征图进行残差连接,这个过程可以 表示如下:

$$F_{\text{MSP-out}} = F_{\text{MSP-in}} + f_{\text{Conv1}\times1}^{32 \to 64} (F_{\text{MSP2}})$$
(8)

其中: $f_{Convl\times 1}^{32 \to 64}$ (•)代表输入通道数为 32、输出通道数为 64、卷 积核尺寸为 1×1 的卷积操作; $F_{MSP-out}$ 代表最后输出的多尺度 特征图。

#### 2.3 全局特征融合块

经过 n 个多尺度特征提取块得到多尺度特征图后,通过一个 3×3 的卷积层将其整合,这个过程可以表示如下:

$$F_{\rm MF} = f_{\rm Conv3\times3}(F_n) \tag{9}$$

其中:F<sub>n</sub>代表经过n个多尺度特征提取块得到的多尺度特征图;F<sub>MF</sub>代表整合后得到的多尺度特征整合图。

通过一个跳跃连接,将通过浅层特征提取块得到的浅层特 征图和多尺度特征整合图进行融合,得到全局特征融合图,这 个过程可以表示如下:

$$F_{\rm GF} = F_0 + F_{\rm MF} \tag{10}$$

其中,FGF代表全局特征融合图。

## 2.4 重建模块

本文所提模型中重建模块的作用是将全局特征融合图映 射成超分辨率图像,由两个部分组成,包括亚像素卷积层和 个 3×3 的卷积层。全局特征融合图经过亚像素卷积层上采样 后,再通过一个 3×3 的卷积层重建得到高分辨率图像:

$$I_{\rm SR} = f_{\rm Conv3\times3}({\rm Pixel}(F_{\rm GF})) \qquad (11)$$

其中:Pixel(•)代表亚像素卷积操作; Isr 代表低分辨率遥感 图像经过本文模型重建得到的高分辨率递感图像。

## 2.5 损失函数

本文模型使用的损失函数是平均绝对误差(Mean Absolute Error, MAE),它是目前深度学习领域遥感图像进行超分辨率重建广泛使用的损失函数:

$$L_{1}(\theta) = \frac{1}{N} \sum_{i=1}^{N} \| \operatorname{Model}(I_{\mathrm{LR}}^{i}) - I_{\mathrm{HR}}^{i} \|_{1}$$
(12)

其中:θ代表本文模型中的所有参数;N代表训练数据的总数 量;I<sup>i</sup><sub>LR</sub>代表输入的低分辨率图像;I<sup>i</sup><sub>HR</sub>代表原始高清图像; Model(•)为本文所提模型;Model(I<sup>i</sup><sub>LR</sub>)代表通过模型重建得 到的高清图像。

### 3 实验(Experiment)

#### 3.1 数据集

本文实验使用的两个数据集均是真实遥感数据集,分别为 NWPU-RESISC45和UCMerced-LandUse。NWPU-RESISC45 数据集是西北工业大学公布的用于遥感图像场景分类的大规 模公开数据集,数据集共有31500幅图像,其中包含45个类 别,每个类别分别包含700幅图像。UCMerced-LandUse数据 集是由 UC Merced 计算机视觉实验室公布的用于遥感图像场 景分类的公开数据集,共有 2 100 幅图像,其中包含 21 个类别, 每个类别分别包含 100 幅图像。本文从 NWPU-RESISC45 数 据集中 airplane 类别的 700 幅图像中随机选取 500 幅图像作为 训练集,剩余的 200 幅图像随机选取一半设置为验证集,另一 半设置为测试集,命名为 NTest<sup>[12]</sup>。设置 UCMerced-LandUse 数据集中 airplane 类别的 100 幅图像为第二个测试集,命名为 UTest。

通过将图像进行随机水平翻转以及随机旋转 90°、180°、 270°增强数据集,获得更多的训练数据。

#### 3.2 实验设置

本文实验的硬件环境为 Intel(R) Xeon(R) Platinum 8255C CPU处理器及 NVIDIA RTX 2080TI 11 GB GPU;软件 环境为 PyTorch 1. 8.1 框架、Python 3.8 和 64 位 Ubuntu18.04 操作系统。本文所提模型采用 ADAM 算法进行优化,优化器 参数设置为  $beta_1=0.9, beta_2=0.999, \epsilon=10^{-8}$ ,训练批大小设 置为 16。训练开始时,学习速率设置为  $10^{-4}$ ,在迭代 200 个 epoch 后,学习速率降低一半,总迭代次数为 400 个 epoch。在实验过程中,使用峰值信噤比(Peak Signal to Noise Ratio,PSNR)和结构相似性(Structural Similarity,SSIM)作为评估指标,PSNR 值和 SSIM 值越大,则重建效果越好。

图 6 给出了训练过程中损失函数的变化趋势,横坐标代表 迭代次数,纵坐标代表训练的损失,可以看出当迭代次数达到 300 个 epoch 后,损失趋于稳定。





#### 3.3 消融实验

为验证本文所提模型中使用的多尺度信息融合和混合注 意力块(MSIFHAB)、多尺度类金字塔特征提取块(MSPLFEB) 及混合注意力块(HAB)的有效性,通过组合不同的块设置消 融实验。消融实验中,本文所提模型中的多尺度特征提取块 (MSFEB)的个数设置为1个,放大因子设置为2,总共迭代200 个 *epoch*。不同的块组合及实验结果如表1所示。其中:M1 模型为去除 MSPLFEB 块和 MSIFHAB 块中的 HAB 块;M2 模型为去除 MSPLFEB 块;使用和 CBAM 相同的通道注意力 机制(CA)和空间注意力机制(SA) 替换 MSIFHAB 块中的 HAB块;M3 模型为去除 MSPLFEB 块;M4 模型为去除 MSIFHAB 块;M5 模型为同时包含 MSPLFEB 块和

#### 第26卷第11期

#### MSIFHAB块。

从表1中M1、M2、M3模型在测试集上得到的PSNR值不 难看出,没有注意力机制的M1模型重建得到的图像在两个测 试集上得到的PSNR值均为最低且明显低于其他模型。加入 CA和SA后的M2模型重建得到的图像在测试集上得到的 PSNR值相比M1模型有了明显提升,而加入HAB的M3模 型在两个测试集上得到的PSNR值比M2模型提高了0.07dB 和0.08dB。由此可见,注意力机制在遥感图像超分辨率重建 网络中发挥了重要的作用,而本文提出的HAB相比其他注意 力机制表现出更优的性能。从表1中M3、M4、M5模型在测试 集上得到的PSNR值可以看出,本文模型单独保留MSPLFEB 块或MSIFHAB块后的指标均低于同时包含两个块的模型,证 明了本文模型各块的有效性。

## 表1 不同块组合在测试集上的平均 PSNR 值 Tab.1 Average PSNR value of different block combinations on the test sets

模型	MSPLFEB	MSIFHAB	HAB	CA+SA	PSNR/dB		
					UTest	NTest	
M1	×	$\checkmark$	×	×	35.38	35.76	
M2	$\times$	$\checkmark$	$\times$	$\checkmark$	36.67	37.04	
M3	$\times$	$\checkmark$	$\checkmark$	$\times$	36.74	37.12	
M4	$\checkmark$	$\times$	$\times$	$\times$	35.28	35.68	
M5	$\checkmark$	$\checkmark$	$\checkmark$	$\times$	36.96	37.35	

#### 3.4 对比实验

为了验证本文模型的有效性,本文选取了7个具有代表意义的重建模型(包括 SRCNN、IDN<sup>[13]</sup>、LGCNet<sup>[14]</sup>、RCAN、MPSR、IRN<sup>[15]</sup>和DSSR<sup>[16]</sup>)设置对比实验,分别从客观评价指标和主观视觉效果两个方面进行分析。

## 3.4.1 客观评价指标

表 2 和表 3 分别给出了放大因子为 2、3、4 时,名模型在两个测试集上的 PSNR 值和 SSIM 值,表中的黑体数字表示模型的最佳结果。

## 表2 各模型在测试集上的 PSNR 值

Tab.2 PSNR values of each model on the test set

					- 平	<u>1M</u> ∶qR
齿刑		UTest			NTest	
医室	r=2	r=3	r=4	r=2	r=3	r=4
SRCNN	35.860 6	31.267 9	28.607 4	36.4818	32.268 0	29.9628
IDN	37.6464	32.554 6	29.7357	37.763 6	33.299 3	30.874 8
LGCNet	37.412 3	32.106 1	29.189 6	37.5750	32.927 8	30.408 4
RCAN	37.717 2	32.678 6	29.738 3	37.852 5	33.394 6	30.908 2
MPSR	37.613 0	32.575 2	29.729 0	37.899 5	33.457 5	30.977 0
IRN	37.594 6	32.633 3	29.780 9	37.8528	33.415 7	30.9511
DSSR	37.664 5	32.669 8	29.737 0	37.8484	33.400 5	30.909 8
本文模型	37.7204	32.751 3	29.8407	37.9996	33. 473 9	31.0119

表3 各模型在测试集上的 SSIM 值

Tab.3 SSIM values of each model on the test set

書刊	UTest			NTest		
医室	r=2	r=3	r=4	r=2	r=3	r=4
SRCNN	0.9507	0.882 2	0.8147	0.952 6	0.886 4	0.824 6
IDN	0.960 3	0.900 5	0.840 5	0.961 8	0.903 5	0.846 6
LGCNet	0.960 5	0.896 0	0.8309	0.9607	0.898 1	0.8368

					\$	卖表
模型	UTest			NTest		
	r=2	r=3	r=4	r=2	r=3	r=4
RCAN	0.961 0	0.900 6	0.839 2	0.962 5	0.904 6	0.847 1
MPSR	0.9604	0.900 1	0.8394	0.962 3	0.905 8	0.848 8
IRN	0.961 2	0.901 1	0.841 1	0.9627	0.905 3	0.848 3
DSSR	0.960 6	0.900 5	0.839 0	0.962 3	0.904 6	0.847 0
本文模型	0.9621	0.9019	0.8443	0.9654	0.9061	0.8508

从表 2 和表 3 中可以看出,出现最早的 SRCNN 模型在所 有场景下得到的 PSNR 值和 SSIM 值均为最低。其他模型随 着模型层数的加深以及更加有效的块的加入, PSNR 值和 SSIM 值相对于 SRCNN 模型均有明显增长。在所有场景里, 本文提出的模型性能均为最优。

## 3.4.2 主观视觉效果

本文选取了两个测试集中的三幅遥感图像的重建结果进 行展示,图7、图8和图9分别为各模型在放大因子为2、3、4时 的重建效果对比图。在这三幅图片中,左边的大图为高清图 像,在高清图像中用方框标记了关键区域,右边八个小图则是 不同模型重建后的图像的标记区域的放大图。通过仔细对比 发现,本文所提模型重建得到的遥感图像在飞机边缘和纹理细 节等方面,的优于其他模型重建后的遥感图像。在放大因子为 4时可以看到 SRCNN、IDN、LGCNET等模型重建后的图像整 体仍比较模糊,无法看到清晰的边界,而本文所提模型重建后 的图像已经可以清晰地看到飞机的边缘轮廓。通过主动视觉 效果对比实验,证明了本文所提模型中的多尺度特征提取块能 更清晰地恢复遥感图像的高频信息,重建得到的高清遥感图像 拥有更好的主观视觉效果。



图 7 放大因子为 2 时各模型重建效果对比 Fig. 7 Comparison of reconstruction effects among different models when the amplification factor is 2



图 8 放大因子为 3 时各模型重建效果对比

Fig. 8 Comparison of reconstruction effects among different models when the amplification factor is 3





Fig. 9 Comparison of reconstruction effects among different models when the amplification factor is 4

# 4 结论(Conclusion)

本文提出了一种基于融合多尺度信息和混合注意力的遥 感图像超分辨率重建模型。多尺度信息融合和混合注意力块 能充分提取多尺度特征,并利用混合注意力机制自适应地调整 多尺度特征通道之间和空间区域的权重,增强了网络重建的性 能。多尺度类金字塔特征提取块通过使用不同尺寸、不同扩张 率的卷积核增大了网络的感受野,使网络可以学习到更多的高 频特征。本文设计了丰富的消融实验以及对比实验,并对实验 结果进行了定性、定量分析,证明了本文所提模型的有效性和 鲁棒性。未来的研究将致力于优化网络结构和缩短模型的训 练时间及进一步提高图像重建效果。

# 参考文献(References)

- [1] KEYS R. Cubic convolution interpolation for digital image processing[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1981, 29(6):1153-1160.
- [2] SUN J, XU Z B, SHUM H Y. Image super-resolution using gradient profile prior [C] // CVPR Organizing Committee Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2008:1-8.
- [3] YANG C Y, YANG M H. Fast direct super-resolution by simple functions[C]//Institute of Electrical and Electronics Engineers. Proceedings of the 2013 IEEE International Conference on Computer Vision, Piscataway; IEEE, 2014; 561-568.
- [4] DONG C, LOY C C, HE K M, et al. Image super-resolution using deep convolutional networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(2):295-307.
- [5] DONG C, LOY C C, TANG X O. Accelerating the super-resolution convolutional neural network[C]//European Conference on Computer Vision, Lecture Notes in Computer Science, Cham; Springer, 2016; 391-407.
- [6] KIM J,LEE J K,LEE K M. Accurate image super-resolution using very deep convolutional networks[C]//CVPR Organizing Committee. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE,2016:1646-1654.

- [7] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C]//CVPR Organizing Committee. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016; 770-778.
- [8] LIM B,SON S,KIM H,et al. Enhanced deep residual networks for single image super-resolution[C]//CVPR Organizing Committee. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE, 2017:1132-1140.
- [9] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C] // CVPR Organizing Committee. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018:7132-7141.
- [10] WOO S, PARK J, LEE J Y, et al. CBAM; convolutional block attention module [C]//Springer Science. Proceedings of the European Conference on Computer Vision. Berlin: Springer, 2018:3-19.
- [11] ZHANG Y LILIK P, LI K, et al. Image super-resolution using very deep residual channel attention networks [C] // Springer Science. Proceedings of the European Conference on Computer Vision. Berlin: Springer, 2018: 294-310.
- [12] 刘丛,王亚新,基于双并行残差网络的遥感图像超分辨率 重建[J],模式识别与人工智能,2021,34(8):760-767.
- HUI Z, WANG X M, GAO X B. Fast and accurate single image super-resolution via information distillation network[C]//CVPR Organizing Committee. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018:723-731.
- [14] LEI S, SHI Z W, ZOU Z X. Super-resolution for remote sensing images via local-global combined network[J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14 (8): 1243-1247.
- [15] XIE H, JIANG H, LIU X, et al. Super resolution for remote sensing images via improved residual network[C//ICMCCE Organizing Committee, Proceedings of the 2020 5th International Conference on Mechanical, Control and Computer Engineering, Piscataway; IEEE, 2021; 2295-2298.
- [16] DONG X Y, SUN X, JIA X P, et al. Remote sensing image super-resolution using novel dense-sampling networks [J].
   IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(2):1618-1633.

# 作者简介:

- 曹春萍(1968-),女,硕士,副教授。研究领域:数据挖掘,图像 处理。
- 李 昂(1995-),男,硕士生。研究领域:人工智能与深度学习, 图像处理。