

基于元学习和密集残差注意力的遥感图像任意尺度超分辨率重建

魏小源¹, 孟钢², 张浩鹏¹, 姜志国¹

1. 北京航空航天大学 宇航学院, 北京 102206;

2. 北京市遥感信息研究所, 北京 100192

摘要: 超分辨率重建技术在卫星遥感图像信息智能处理领域中有重要的应用。现有面向遥感图像超分辨率重建的深度学习大多只能处理一种比例因子的超分辨率重建任务, 在多尺度层面上缺少泛化性, 难以满足真实遥感图像多倍率连续放大的超分辨率重建任务需求。为解决遥感图像超分辨率重建过程中的多尺度放大问题, 本文采用元学习的方法, 在构建单一自适应模型的基础上实现对遥感图像的任意尺度超分辨率重建, 提升遥感图像的空间分辨率, 利用密集残差网络和通道注意力机制重建遥感图像中地物纹理、目标边缘等丰富细节信息。在真实遥感图像上的定量实验表明, 本文所提方法重建结果的峰值信噪比能达到40 dB以上, 同时在多种数据上的定量和定性实验结果证明了本文方法的有效性。

关键词: 超分辨率重建, 遥感图像, 任意尺度, 元学习, 密集残差网络, 通道注意力机制

中图分类号: TP751.1/P2

引用格式: 魏小源, 孟钢, 张浩鹏, 姜志国. 2024. 基于元学习和密集残差注意力的遥感图像任意尺度超分辨率重建. 遥感学报, 28(7): 1735-1745

Wei X Y, Meng G, Zhang H P and Jiang Z G. 2024. Arbitrary-scale super-resolution reconstruction of remote sensing images based on meta-learning and residual dense attention mechanism. National Remote Sensing Bulletin, 28(7): 1735-1745 [DOI: 10.11834/jrs.20233267]

1 引言

超分辨率重建技术在卫星遥感图像信息智能处理领域中有重要的应用, 经过超分辨率技术重建得到的高空间分辨率遥感图像在目标检测与跟踪、地理资源识别等高层智能任务中有更高的应用价值。传统遥感图像超分辨率重建技术包括双线性插值、双三次插值、边缘保持、基于偏微分方程的全变分 (Dosovitskiy 和 Brox, 2016) 方法和各向异性扩散 (Bavirisetti 和 Dhuli, 2016) 方法等。这些方法虽然计算简单, 但对具有复杂纹理的遥感图像而言存在无法恢复高频细节信息的问题。随着深度学习技术的发展, 基于神经网络的超分辨率重建技术被引入遥感图像处理领域, 其中基于卷积神经网络 (CNN) 的方法是目前最常

用的方法之一, 另外还有基于近年来被广泛应用于图像生成和处理的生成对抗网络 (GAN) 的超分辨率重建技术, 在遥感图像超分辨率重建方面也具有良好的效果, 能够实现更高感知质量的遥感图像超分辨率重建。目前, 用于遥感图像超分辨率重建的前沿深度学习算法包括采用纯数据驱动的流程模型 (任术波等, 2022)、应用混合稀疏表示模型的MSR-SRR (杨雪等, 2022)、针对高分四号卫星中波红外影像的卷积网络 (贺智和贺丹, 2020) 以及针对 Sentinel-2 卫星遥感图像的无监督方法 KN-SRGAN (赵慧岩和李云鹤, 2022) 等。数据流算法采用纯数据驱动的流程模型和优化后的密集残差网络对低分辨率图像进行特征提取, 提升了超分辨率图像的感知质量。KN-SRGAN作为基于生成对抗网络的方法, 重建出的遥感图像具

收稿日期: 2023-07-11; 预印本: 2023-12-01

基金项目: 国家自然科学基金 (编号: 62271017)

第一作者简介: 魏小源, 研究方向为遥感图像超分辨率重建。E-mail: 19375140@buaa.edu.cn

通信作者简介: 张浩鹏, 研究方向为遥感图像处理与解译、空间目标信息处理。E-mail: zhanghaopeng@buaa.edu.cn

有更适合人眼直观视觉感受和具备更佳感知效果等特点。

以上深度学习模型只能处理一种比例因子的超分辨率重建任务，在多尺度层面上缺少泛化性，属于基础学习器的范畴。对于真实遥感图像的超分辨率重建任务，往往要对图像按不同的整数或非整数倍连续放大，若是对每种可能的比例因子都训练一个模型并部署在平台中，会造成极大的算力和空间资源的浪费。因此有必要研究能用单一模型处理任意尺度遥感图像超分辨率重建任务的方法，这种超分辨率重建模型实际上要求学习算法有在任务层面进行学习的能力，而元学习就是一种针对不同任务自适应改变模型内部机制的策略。将元学习和基础学习做对比，基础学习由先验偏置确定假设空间，学习算法在确定好的假设空间内学习单一任务的最优解。元学习旨在学习导致一种算法适应一种任务的原因，以及如何将模型泛化到更多类型的任务中。

针对上述问题，本文提出遥感图像任意尺度超分辨率重建方法，该方法采用元学习 (Vilalta and Drissi, 2002; Pratt 和 Thrun, 1997; Thrun 和 Pratt, 1998) 的思想，根据不同比例因子自适应地调整模型内部参数，完成任意尺度超分辨率重建任务，同时采用带有注意力机制的密集残差网络作为特征提取器，使重建结果具备更清晰、区分度更高的细节。在公开遥感数据和真实卫星遥感图像上的定量和定性实验结果表明，本文所提方法具有良好的任意尺度超分辨率重建能力。

2 研究方法

本文从基于元学习的超分辨率重建方法 Meta-SR (Hu 等, 2019) 出发，采用密集残差网络 RDN (Residual Dense Network) (Zhang 等, 2018b) 和元上采样模块 (Meta Upscale Module) 的组合，实现遥感图像任意尺度超分辨率模型 Meta-RDN 作为基准模型，同时考虑到遥感图像包含丰富的局部地物目标信息，为了使重建结果具备更清晰、区分度更高的细节，将通道注意力 CA (Channel Attention) (Zhang 等, 2018a) 机制引入 Meta-RDN，建立使用密集残差注意力网络 RDCAN (Residual Dense Channel Attention Network) 提取特征的改进模型 Meta-RDCAN。

2.1 Meta-RDCAN 网络结构

对一个给定的从高分辨率图像 I_{HR} 中下采样得到的低分辨率图像 I_{LR} ，超分辨率重建算法的任务是从 I_{LR} 中重建出超分辨率图像 I_{SR} 。其中 I_{SR} 的恢复需要根据特征提取网络从 I_{LR} 中提取的特征图 F_{LR} 以及对应比例因子的上采样滤波器 W 计算得到，因此一个超分辨率重建模型的基本结构是特征提取模块加上采样模块的组合。带有元上采样模块的超分辨率重建算法 Meta-SR 在提取低分辨率图像特征后，将比例因子也作为输入用以计算上采样滤波器权重，最终通过卷积滤波器和特征图运算得到超分辨率图像。该算法中的元上采样模块能自适应地根据输入比例因子来调整模型内部参数，从而实现任意尺度超分辨率重建的功能。本文以密集残差网络作为特征提取器，结合元上采样模块构建遥感图像任意尺度超分辨率重建模型 Meta-RDN。

考虑到遥感图像中包含丰富的地物信息和复杂的纹理细节，为了在超分辨率重建过程中充分强调遥感图像的局部特征，本文向密集残差网络中引入了通道注意力机制，实现密集残差注意力特征提取网络，从低分辨率输入中提取得到的特征图经过元上采样后输出最终的超分辨率重建结果。进一步改进后的 Meta-RDCAN 网络结构如图 1 所示。Meta-RDCAN 以从 I_{HR} 中经双三次下采样得到的 I_{LR} 和比例因子作为输入， I_{LR} 经过 RDCAN 提取特征后得到 F_{LR} 。随后 F_{LR} 输入元上采样网络，按照由当前比例因子计算得到的上采样滤波器权重将特征值映射为 I_{SR} ，完成一次前向传播过程。然后 I_{SR} 和 I_{HR} 进行比较并求取损失，通过反向传播更新特征提取网络和元上采样网络的参数。

2.2 元上采样网络结构

Meta-SR 方法的核心是使用元上采样模块代替单一尺度超分辨率重建模型中的上采样层，从而实现任意尺度超分辨率重建功能。图 1 中使用的元上采样模块输入为比例因子 r 或一组比例因子构成的向量，以及来自特征提取网络的 F_{LR} ，输出为在当前比例因子下的 I_{SR} 。元上采样模块的功能可用式 (1) 描述

$$I_{SR}(i, j) = \Phi(F_{LR}(i', j'), W(i, j)) \quad (1)$$

式中， $I_{SR}(i, j)$ 表示 I_{SR} 中位置为 (i, j) 的像素值，

$F_{LR}(i', j')$ 表示 I_{LR} 中位置为 (i', j') 的像素特征值, $W(i, j)$ 表示在当前比例因子下的上采样滤波器 w 对像素 (i, j) 的权重, 函数 $\Phi(\cdot)$ 表示计算 I_{SR} 像素值的特征映射函数。该式说明上采样模块具备 3 个功能, 即根据不同比例因子预测不同上采样滤波器 w 的权重预测功能, 将 I_{SR} 中的像素 (i, j) 与 I_{LR} 的像素 (i', j') 相对应的位置投影功能, 以及根据 F_{LR} 和 w 计算 I_{SR} 像素值的特征映射功能。其中, 权重预测功能需要接受来自比例因子的输入。和典型的超分辨率重建网络中的上采样模块不同, 图 1 所示的元上采样模块采用前馈网络来自适应地预测上采样滤波器的权重, 该过程的表达式为

$$W(i, j) = \varphi(v_{ij}; \theta) \quad (2)$$

式中, $W(i, j)$ 表示上采样滤波器 w 对 I_{SR} 中像素 (i, j) 的权重, $\varphi(\cdot)$ 表示以 v_{ij} 为输入, 以 θ 为参数的前馈网络。前馈网络包含两层全连接层和一层 ReLU 激活函数。 v_{ij} 是根据比例因子 r 和位置 (i, j) 计算得到的相对 I_{LR} 中像素 (i', j') 的偏移向量, 计算公式为

$$v_{ij} = \left(\frac{i}{r} - \left\lfloor \frac{i}{r} \right\rfloor, \frac{j}{r} - \left\lfloor \frac{j}{r} \right\rfloor, \frac{1}{r} \right) \quad (3)$$

式中, $\frac{1}{r}$ 项是为了区分成倍数关系的比例因子之间的上采样滤波器权重。例如在对低分辨率图像分别进行 3 倍和 6 倍超分辨率重建时, 对于没有 $\frac{1}{r}$ 参与计算的上采样滤波器, 在 3 倍超分辨率重建图像中的像素 (i, j) 和 6 倍超分辨率重建图像像素 $(2i, 2j)$ 就会具有相同权重, 从而限制超分辨率重建模型的泛化能力。对 I_{SR} 中的每个像素 (i, j) 而言, 其像素值是根据 I_{LR} 在 (i', j') 上的特征值决定的, 但在不同的尤其是非整数的比例因子下, 不同位置上的像素 (i', j') 可能对不同数量的像素 (i, j) 产生影响。例如在 $r = 1.5$ 的超分辨率重建任务中, 有的像素 (i', j') 决定 I_{SR} 的两个像素值, 而有的像素只能决定一个。因此元上采样模块通过向下取整数来实现位置投影功能

$$(i', j') = \left(\left\lfloor \frac{i}{r} \right\rfloor, \left\lfloor \frac{j}{r} \right\rfloor \right) \quad (4)$$

特征映射是根据上采样滤波器的权重 $W(i, j)$ 和 $F_{LR}(i', j')$ 计算得到超分辨率重建图像的像素值。图 1 中的元上采样模块的特征映射功能可用下式描述:

$$\Phi(F_{LR}(i', j'), W(i, j)) = F_{LR}(i', j') W(i, j) \quad (5)$$

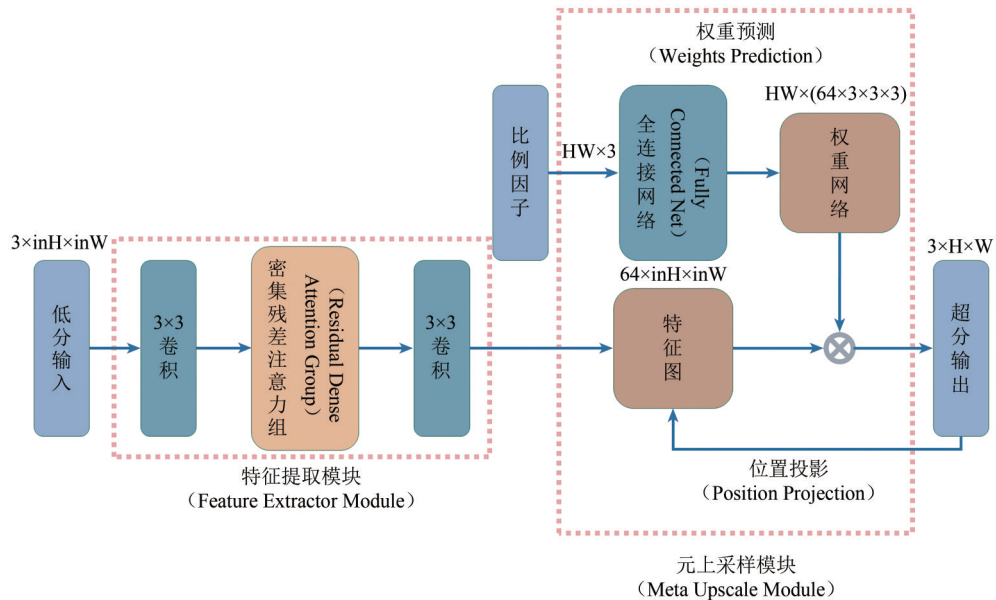


图 1 Meta-RDCAN 模型结构

Fig. 1 Structure of Meta-RDCAN

2.3 密集残差注意力网络结构

为了在超分辨率重建过程中强调遥感图像的局部特征, 提高超分辨率图像中局部细节的重建

效果, 在密集残差网络的基础上引入了通道注意力机制。本文采用图 2 所示轻量化通道注意力机制, 通道注意力层接受上一级特征图输入后, 首

先进行自适应平均池化提取特征图的全局信息，然后通过两层非线性层采样后得到通道权值向量。权值向量和上级特征图相乘可实现对特征图通道的加权，从而完成一次施加通道注意力的过程。该通道注意力层不能随意添加到原网络中，因为一方面加入注意力机制会增加模型参数量，另一方面注意力可能会错误地放大特征图中的噪声，反而造成性能下降。考虑到在密集残差块的前端施加注意力机制相当于舍弃该残差块提取的特征信息，无法充分发挥密集残差和注意力结合的优势，因此本文将通道注意力层插入到图3所示的密集残差块末端，将这种改进后的结构称为密集残差注意力块。同时基于残差连接的思想，在每层注意力的输入端引出一条额外的残差连接和当前注意力层的输出融合，保证对通道的加权不会引发层次信息的丢失。本文在密集残差注意力网络中应用了16组密集残差注意力块，每个密集残差注意力块中有8层卷积层，最终提取得到64通道的特征图。

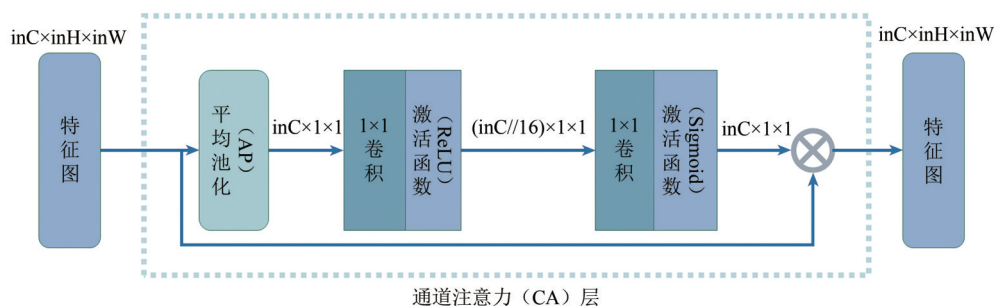


图2 通道注意力结构

Fig. 2 Structure of channel attention

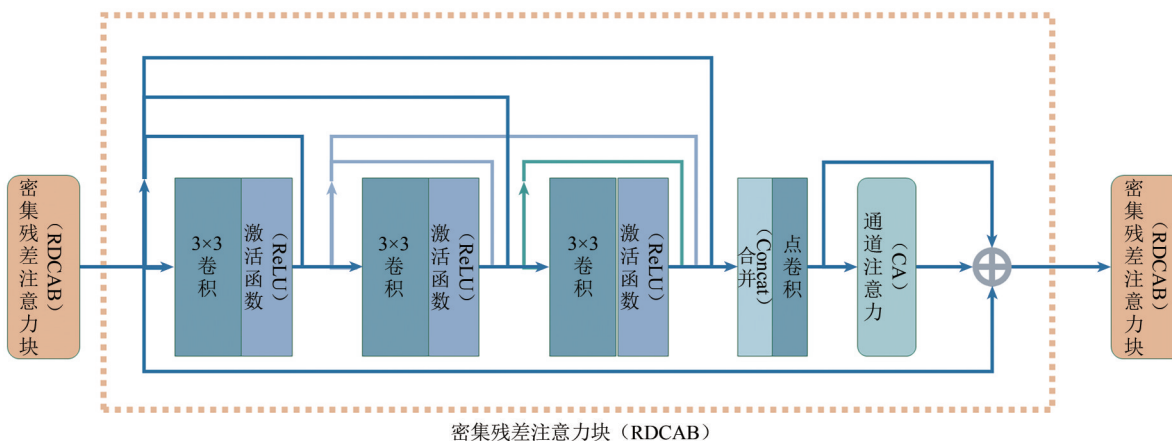


图3 添加通道注意力后的密集残差块结构

Fig. 3 Structure of residual dense block with channel attention

3 数据结果处理与分析

本文对基准模型 Meta-RDN 和改进模型 Meta-RDCAN 的实验结果进行了定性和定量的分析对比。实验过程中，先在 DIV2K 数据集 (Agustsson 和 Timofte, 2017) 上预训练 Meta-RDN 和 Meta-RDCAN，然后分别在遥感数据集 AID (Xia 等, 2017), UCMerced (Yang 和 Newsam, 2010) 和 WIDS (Liu 等, 2019) 上微调。为了研究对遥感图像按不同比例因子进行超分辨率重建时空分辨率变

化带来的影响，在较广的比例因子范围内训练了多组模型，并在澳门科普卫星的真实遥感数据上进行了充分测试。此外，Set5 和作为测试集分割出来的部分 UCMerced 数据也被用于模型测试。评价指标方面，采用有参考指标峰值信噪比 PSNR (Peak Signal-to-Noise Ratio) 和结构相似度 SSIM (Structural Similarity Index Measure) 来衡量超分辨率重建图像和真值图像的差异，无参考指标 NIQE (Mittal 等, 2013) 来衡量超分辨率重建图像的感知质量，以及算法运行时间来衡量图像处理速率。

3.1 数据集预处理

本文使用的数据集包括 DIV2K (Agustsson 和 Timofte, 2017)、AID (Xia 等, 2017)、UCMerced (Yang 和 Newsam, 2010)、WIDS (Liu 等, 2019)、Set5 (Kulkarni 等, 2016) 和来自澳门科普卫星的真实遥感图像数据。DIV2K 数据集是一个广泛用于图像超分辨率研究的数据集, 包含 800 张来自各种场景的不同尺寸的高分辨率图像。AID 数据集是一种广泛用于遥感图像分类任务的公共数据集, 包含来自 Google Earth 等卫星影像在内的 10 种不同场景共 10000 张的遥感图像, 空间分辨率在 0.5—8 m。AID 数据集的图像场景包括: 机场、海滨、农田、森林、工业区、公园、停车场、铁路、市区和河流, 每个场景下的图像数量相等, 且每张图像的分辨率为 600×600 像素。该数据集的特点是图像场景种类多样, 充分考虑了遥感图像的特点, 可以用于各种图像分类、目标检测和识别任务的训练和测试。UCMerced 数据集是一个用于地物分类的公共遥感图像数据集, 包含 21 类总计 2100 张大小为 256×256 像素的地物图像, 空间分辨率为 0.3 m。这些图像拍摄于不同的季节和时间, 是广泛用于遥感图像分类算法评估和测试的数据集之一。WIDS 数据集源于机器学习比赛 WiDSDataathon2019, 该数据集中图像的空间分辨率为 3 m, 共有 11000 张大小为 256×256 像素的卫星影像。Set5 数据集是常用于测试图像超分辨率重建算法性能的基准数据集, 包含 5 张不同类别的低分辨率彩色图片, 因其简单和易于使用而广泛应用于图像超分辨率重建算法的快速测试和评估中。澳门科普卫星数据源于 2022 年搭载在天舟五号上发射的“澳门学生科普卫星一号”的遥感影像。该影像是一张大小为 2048×2048 像素的三通道彩色图像, 空间分辨率为 8 m。实验中对原图进行了大小为 512×512 像素的随机裁剪, 并从中挑选出 7 张具有丰富地物目标信息和纹理细节的图像构成澳门科普卫星数据集用于模型测试。

在进行有监督训练和有参考质量评估时, 需要对训练数据和测试数据按一定比例因子进行下采样以获得低分—高分图像对。本文采用双三次下采样对训练和测试数据进行预处理, 处理结果如表 1 中所示, 其中训练数据以 0.1 为步长, 测试数据以 0.5 为步长进行下采样, UCMerced 数据集按

8 : 2 的比例划分为训练集和测试集。

表 1 数据集预处理

Table 1 Pre-processing of datasets

数据集	类型	格式	空间分辨率/m	下采样范围	用途
DIV2K	通用	png	—	1.1—4.0	训练
AID	遥感	ipg	0.5—0.8	1.1—4.0	训练
UCMerced	遥感	tif	0.3	1.1—4.0	训练
WIDS	遥感	tif	3	1.1—6.0	训练
Set5	通用	png	—	1.5—4.0	测试
UCMerced	遥感	tif	0.3	1.5—4.0	测试
澳门科普卫星数据	遥感	tif	8	1.5—4.0	测试

对遥感图像而言, 经过下采样处理后得到的低分图像相比原图更加模糊, 相当于降低了空间分辨率。在遥感图像超分辨率重建任务中, 输入图像的空间分辨率将对超分辨率重建结果产生影响, 比如高空间分辨率的遥感图像具有更清晰密集的细节, 在超分辨率重建过程中有足够的特征信息用于重建, 但低空间分辨率的遥感图像缺少纹理、边缘和轮廓等信息, 使超分辨率重建效果下降。本文进行的下采样操作会引起实际输入的遥感图像和原图像空间分辨率的不一致。为了研究空间分辨率变化带来的影响, 实验中在表 1 中的多组遥感数据集上训练了 Meta-RDN 和 Meta-RDCAN, 并且对部分遥感数据集的下采样范围进行扩大, 目的是使下采样后的训练数据空间分辨率范围一定程度上能覆盖到测试数据下采样后的空间分辨率范围。

这种覆盖关系如图 4 中所示, 其中不同的遥感数据集用颜色加以区分, 在数轴上的范围代表了数据集经下采样后相当的空间分辨率范围。原图像空间分辨率到下采样后相当空间分辨率之间的关系由式 (6) 给出:

$$R_{\text{LR}} = R_{\text{HR}} S \tag{6}$$

式中, R_{HR} 和 R_{LR} 分别表示原始遥感数据和下采样后数据的空间分辨率, S 表示下采样因子。

从图 4 中可看出, 模型在遥感数据集 AID 上训练时, 可以获得测试时处理的低空间分辨率图像的经验。在 UCMerced 上训练的模型能获得更充足的重建图像细节的经验, 但该训练数据与澳门科普卫星数据的空间分辨率区间没有交集, 因此在测试时, 相当于要求模型在新的空间分辨率范围内泛化。

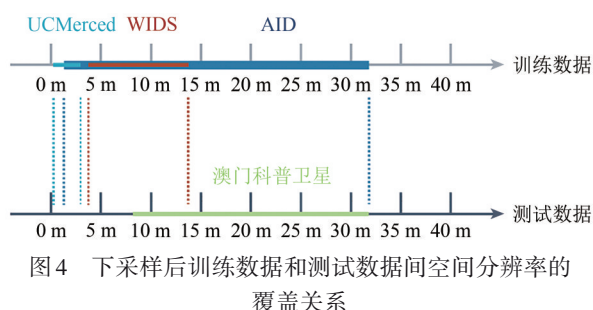


Fig. 4 The coverage relationship of spatial resolution between training data and test data after downsampling

3.2 实验设置

实验过程中采用两块 RTX 2080Ti 进行训练，批量 (Batch) 大小为 2，训练轮次 (Epochs) 为 200。损失函数方面没有采用超分辨率重建模型常用的 L2 损失，而是遵循 Meta-SR 的设置使用 L1 损失函数。L1 损失函数直接衡量模型超分辨率重建结果与原始高分辨率图像之间的差异，计算方法是将超分辨率重建图像与高分辨率图像之间的差取绝对值，然后对所有差值求和并除以样本数量，求得平均绝对误差后反向传播并更新网络参数。优化策略方面采用自适应矩估计 ADAM (Adaptive

Moment Estimation) 优化算法，该算法使用梯度的一阶矩估计和二阶矩估计动态地调整每个参数的学习率。自适应矩估计使用指数移动平均法来估计梯度的一阶矩和二阶矩，并通过偏差校正来纠正估计的偏差。相对于传统的随机梯度下降和其他优化算法，自适应矩估计算法在训练深度神经网络时往往可以取得更好的结果，并且具有一定的鲁棒性。实验中设置参数学习率 (Learning Rate) 为 10^{-4} ，学习率衰减 (Learning Rate Decay) 为 200。

本文在 DIV2K 和 WIDS 上按照不同方案训练了 Meta-RDCAN 并绘制了 L1 损失的变化曲线，结果如图 5 所示。根据损失曲线，模型训练过程中损失接近 5.0 后下降变得十分缓慢，但继续训练确实能提升模型性能。图 5 (a) 和图 5 (b) 中在单一数据集上训练模型时，损失在中止训练时高于 5.0，而图 5 (c) 中在 WIDS 上微调的预训练模型最终损失能降到 5.0 以下。因此为了节省训练成本，同时保证算法性能不会因为提前终止而产生明显下降，本实验采用先在通用数据集 DIV2K 上预训练，再在各组遥感数据集上微调的训练方案。

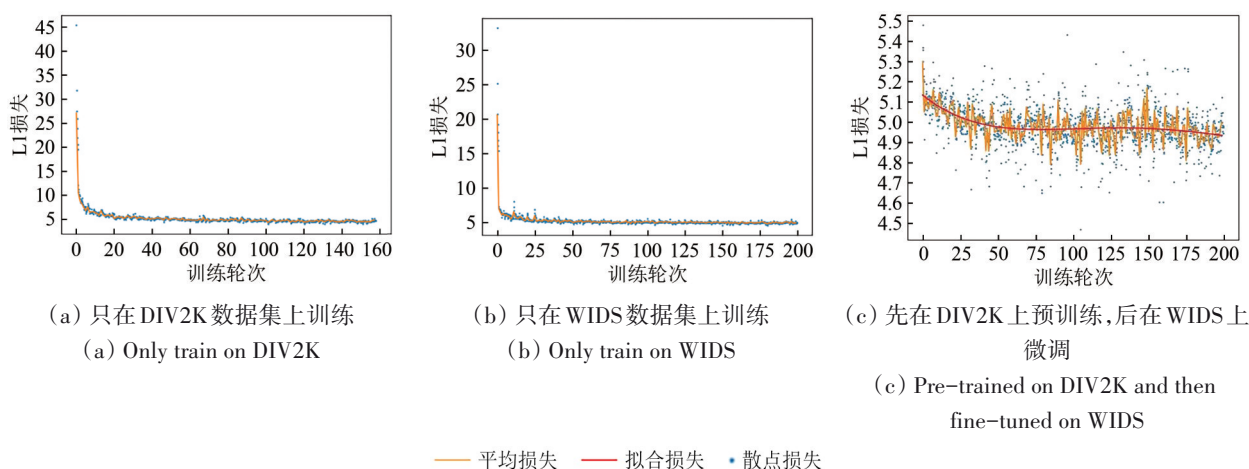


Fig. 5 The loss curve of Meta-RDCAN under different training regimens

3.3 实验结果

实验过程中在遥感数据集 AID, UCMerced 和 WIDS 上分别训练了 Meta-RDN 和 Meta-RDCAN，然后在 Set5, UCMerced 和澳门科普卫星数据上进行测试。测试的比例因子设置在 1.5—4.0，步长为 0.5。因此一共有 6 组模型在 3 组数据上进行共计 18 次测试，每次测试分别包含 6 个比例因子的结

果。在澳门遥感卫星数据上计算有参考指标峰值信噪比 (dB)、结构相似度和算法运行时间 (s) 的结果如表 2 中所列。表 3 和表 4 中列出了部分在 UCMerced 和 Set5 上计算得到的以上指标，各表中横向对比时更高的峰值信噪比已用粗体标出表 3 和表 4 中列出了部分在 UCMerced 和 Set5 上计算得到的以上指标，各表中横向对比时更高的峰值信噪

比已用粗体标出, 算法运行时间取在不同比例因子下重建所耗时间的平均值, 且在不同数据集上训练的模型用模型名称加训练数据集区分。

表 2 在澳门科普卫星数据上的测试结果
Table 2 Test results on Macao science popularization satellite data

模型	因子	PSNR/dB	SSIM	时间/s	模型	因子	PSNR/dB	SSIM	时间/s
Meta-RDN+AID	1.5	40.42	0.9729	0.27	Meta-RDCAN+AID	1.5	42.70	0.9768	0.35
	2.0	40.50	0.9814			2.0	42.76	0.9901	
	2.5	38.75	0.9720			2.5	41.06	0.9854	
	3.0	37.32	0.9606			3.0	39.63	0.9795	
	3.5	35.92	0.9468			3.5	38.14	0.9707	
	4.0	34.63	0.9307			4.0	36.68	0.9588	
Meta-RDN+UCMerced	1.5	42.47	0.9754	0.27	Meta-RDCAN+UCMerced	1.5	42.38	0.9761	0.35
	2.0	42.44	0.9892			2.0	42.39	0.9896	
	2.5	40.76	0.9840			2.5	40.73	0.9846	
	3.0	39.25	0.9774			3.0	39.20	0.9780	
	3.5	37.76	0.9680			3.5	37.68	0.9685	
	4.0	36.29	0.9555			4.0	36.19	0.9559	
Meta-RDN+WIDS	1.5	39.10	0.9627	0.28	Meta-RDCAN+WIDS	1.5	38.65	0.9562	0.37
	2.0	39.78	0.9803			2.0	39.75	0.9801	
	2.5	37.90	0.9701			2.5	37.98	0.9709	
	3.0	36.85	0.9596			3.0	36.98	0.9619	
	3.5	35.47	0.9452			3.5	35.63	0.9489	
	4.0	34.35	0.9298			4.0	34.51	0.9348	

注: 黑体表示横向对比时更高的峰值信噪比。

表 3 在 UCMerced 上的部分测试结果
Table 3 Partial test results on UCMerced dataset

模型	因子	PSNR/dB	SSIM	时间/s	模型	因子	PSNR/dB	SSIM	时间/s
Meta-RDN+AID	2.0	32.58	0.8909	0.14	Meta-RDCAN+AID	2.0	34.78	0.9337	0.17
	3.0	29.23	0.7895			3.0	30.55	0.8386	
	4.0	27.44	0.7139			4.0	28.46	0.7638	
Meta-RDN+UCMerced	2.0	34.58	0.9310	0.18	Meta-RDCAN+UCMerced	2.0	34.57	0.9311	0.17
	3.0	30.48	0.8390			3.0	30.52	0.8433	
	4.0	28.37	0.7606			4.0	28.36	0.7619	
Meta-RDN+WIDS	2.0	31.33	0.8687	0.12	Meta-RDCAN+WIDS	2.0	31.53	0.8749	0.19
	3.0	28.50	0.7651			3.0	28.69	0.7744	
	4.0	26.93	0.6905			4.0	27.09	0.6991	

注: 黑体表示横向对比时更高的峰值信噪比。

表 4 在 Set5 上的部分测试结果
Table 4 Partial test results on Set5 dataset

模型	因子	PSNR/dB	SSIM	时间/s	模型	因子	PSNR/dB	SSIM	时间/s
Meta-RDN+AID	2.0	35.13	0.9298	0.12	Meta-RDCAN+AID	2.0	37.53	0.9590	0.17
	3.0	32.01	0.8852			3.0	33.77	0.9226	
	4.0	29.93	0.8399			4.0	31.44	0.8847	
Meta-RDN+UCMerced	2.0	37.21	0.9750	0.12	Meta-RDCAN+UCMerced	2.0	37.39	0.9583	0.14
	3.0	33.47	0.9180			3.0	33.58	0.9209	
	4.0	31.15	0.8774			4.0	31.24	0.8805	
Meta-RDN+WIDS	2.0	33.77	0.9234	0.13	Meta-RDCAN+WIDS	2.0	34.08	0.9235	0.15
	3.0	31.12	0.8767			3.0	31.39	0.8792	
	4.0	29.36	0.8318			4.0	29.58	0.9357	

注: 黑体表示横向对比时更高的峰值信噪比。

纵向比较表2中的峰值信噪比,可以看出当比例因子增大至3.0左右时,无论是Meta-RDN还是Meta-RDCAN,计算出的峰值信噪比基本下降到40.0以下。一般认为重建图像的峰值信噪比大于35.0时,图像质量可被接受,而峰值信噪比大于40.0时重建图像具有较高的质量。通过观察图6中按不同比例因子进行重建的澳门科普卫星数据可以发现,当比例因子在2.5以下时,图像中的飞机、油罐等目标能有较清晰的轮廓,而且部分并排的

目标也有明显的能用肉眼区分的边缘。而当比例因子超过4.0时,一些细小的飞机目标在重建结果中已经消失,排列较为紧密的目标也在超分辨率重建过程中被混淆,无法清楚地判断出多个目标实例,这种重建图像是无法用于下游目标检测与识等任务的。综合定性和定量的分析,可以给出所提模型Meta-RDCAN的适用范围,即适用于比例因子在4.0以内的遥感图像任意尺度超分辨率重建任务。

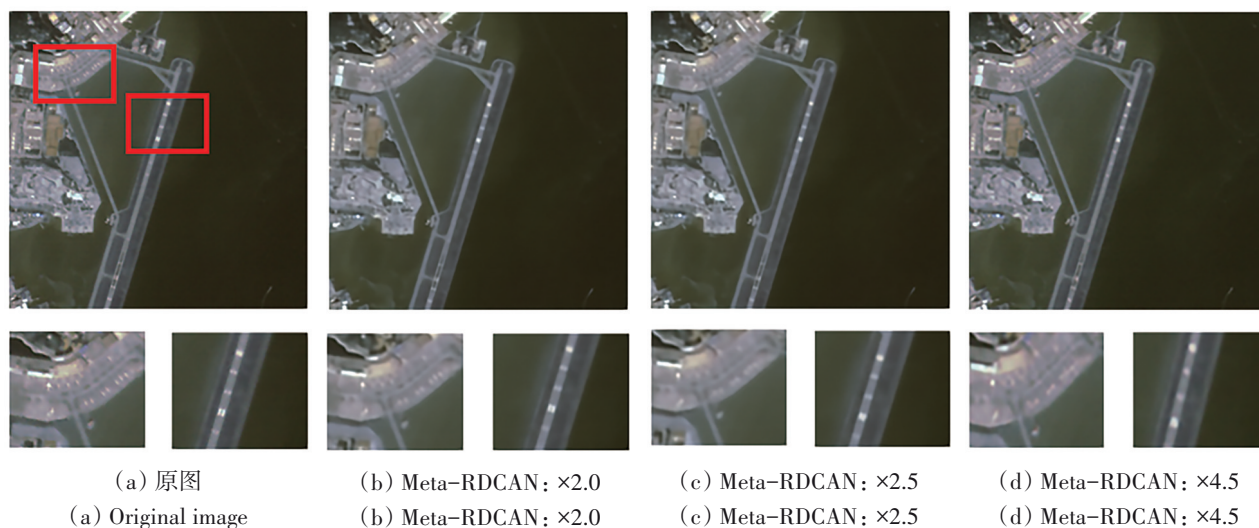


图6 Meta-RDCAN在2.0、2.5及4.5比例因子下对澳门科普卫星数据的重建结果

Fig. 6 Meta-RDCAN reconstruction results on Macao data under 2.0, 2.5 and 4.5 scale factors

另外,纵向对比表2和表3中在不同数据集上训练模型的测试结果可看出,在AID和UCMerced上训练的模型效果几乎总是优于在WIDS上训练的模型。这一结果可借助图4中的训练数据和测试数据的对应关系解释。当在澳门科普卫星数据上测试时,AID包含全部测试数据可能涵盖的空间分辨率,而UCMerced和WIDS数据集都不能完全覆盖测试数据,但UCMerced数据集的空间分辨率更高,训练过程中能提供更多的特征信息和更丰富的重建细节的经验。在UCMerced上测试时,UCMerced训练数据已完全覆盖测试可能出现的空间分辨率,AID数据集也包含足够充足的样本让模型学习到重建高空间分辨率遥感图像的经验,反之WIDS数据集与测试数据完全没有空间分辨率的交集,因此训练得到的模型性能不如其他两者。该实验结果说明了空间分辨率在遥感图像超分辨率重建任务中产生的重要影响,也说明了用于训

练的数据尽可能选择覆盖空间分辨率范围广,包含丰富细节的遥感图像。

最后,横向对比表2、表3和表4中的各项指标,可以说明在密集残差块中添加通道注意力确实能提升网络性能。在图7和图8的定性比较结果中,可看出Meta-RDCAN能重建出草坪中心目标下半部分的完整边缘和蝴蝶翅膀的纹理细节,而Meta-RDN的重建结果则可能缺少边缘或者存在细节模糊的问题。另外,通过对比算法运行时间能说明取得提升的代价是计算速率的降低。值得注意的是在有些横向对比结果中,即使峰值信噪比取得较高的值,结构相似度也可能取得较低的值,以及部分纵向对比中,峰值信噪比降低的同时结构相似度反而升高。这一现象说明峰值信噪比和结构相似度这两种有参考评价指标之间不一定是正相关关系,在分析模型性能优劣时还需根据具体任务具体分析更能准确反映算法的优缺点的指标。

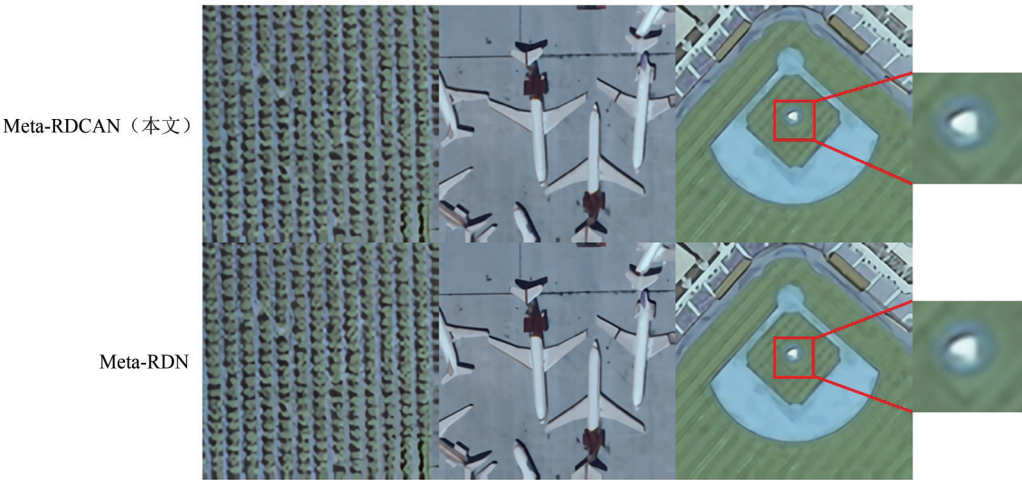


图7 Meta-RDCAN 和 Meta-RDN 在 3.5 比例因子下对 UCMerced 数据的重建结果

Fig. 7 Meta-RDCAN and Meta-RDN reconstruction results on UCMerced data under 3.5 scale factor

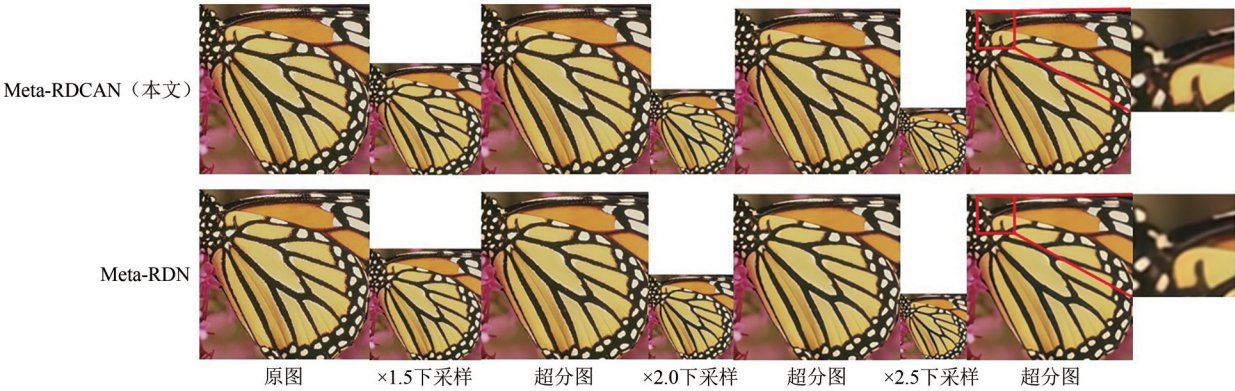


图8 Meta-RDCAN 和 Meta-RDN 在 3 种比例因子下对 Set5 数据的重建结果

Fig. 8 Meta-RDCAN and Meta-RDN reconstruction results on Set5 data under three different scale factors

对于遥感图像任意尺度超分辨率重建任务，仅使用有参考评价指标会带来如必需高分辨率图像作为真值等诸多限制，不符合遥感图像缺少高分辨率图像的实际应用场景。因此本文还引入无参考图像质量评价指标 NIQE 来评价超分辨率重建图像的感知质量。NIQE 是一种将输入图像与预先用自然图像集建立好的模型对比的方法，其值越低代表图像感知质量越好。由于不需要低分—高分图像对，超分辨率重建算法可以直接在原始测试数据上运行，也在一定程度上淡化了图4所示的空间分辨率改变的影响。表5是 Meta-RDN 和 Meta-RDCAN 直接对澳门科普卫星数据按照 1.5 和 2.0 倍率进行超分辨率重建的结果，且已用粗体标出横向对比时更优的结果。可以看出用 NIQE 作为评价指标时，也能得出上述使用有参考指标得到的结论，该结果进一步证明了使用通道注意力机制的有效性。

表5 在澳门科普卫星数据上计算 NIQE 的结果

Table 5 Results of calculated NIQE on Macao satellite data		
模型	因子	NIQE ↓
Meta-RDN+AID	1.5	5.227
	2.0	4.830
Meta-RDN+UCMerced	1.5	4.382
	2.0	5.071
Meta-RDN+WIDS	1.5	4.917
	2.0	4.905
Meta-RDCAN+AID	1.5	4.619
	2.0	4.843
Meta-RDCAN+UCMerced	1.5	4.229
	2.0	4.715
Meta-RDCAN+WIDS	1.5	4.958
	2.0	4.805

注: 黑体表示横向对比时更高的 NIQE。

4 结 论

针对遥感图像任意尺度超分辨率重建的问题，

本文提出采用元学习和密集残差注意力网络的超分辨率重建方法 Meta-RDCAN。本方法应用的元上采样模块通过权重预测、位置投影和特征映射3大功能实现任意尺度超分辨率重建,同时从充分提取遥感图像局部地物目标信息的角度出发设计密集残差注意力特征提取网络,恢复遥感图像的细节信息。

本文在 DIV2K、AID、UCMerced、WIDS、Set5 和来自澳门科普卫星的真实遥感图像数据上进行了充分实验,分析了空间分辨率变化对超分辨率重建结果的影响,并基于损失曲线验证了先在通用数据集上预训练、然后在遥感数据集上微调的训练方案的合理性。

基于不同比例因子的测试结果表明本文所提模型适用于比例因子在4.0以内的遥感图像任意尺度超分辨率重建任务。对比实验结果说明添加通道注意力的改进模型在峰值信噪比和结构相似度上能取得比基准模型更好的表现;基于无参考指标 NIQE 的对比结果也能说明改进模型超分辨率重建结果的感知质量优于基准模型。以上研究证明了基于元学习和密集残差注意力的遥感图像任意尺度超分辨率重建方法的有效性。本文的主要贡献包括两方面:

(1) 针对遥感图像任意尺度超分辨率重建问题,采用元学习的方法自适应调整模型内部参数,实现用单一模型对单张遥感图像进行连续的整数和非整数倍超分辨率重建。

(2) 针对重建结果中细节缺失、地物目标边缘不清晰的问题,采用通道注意力机制改进密集残差网络,提升了超分辨率重建结果的质量。

参考文献(References)

- Agustsson E and Timofte R. 2017. NTIRE 2017 challenge on single image super-resolution: dataset and study//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Honolulu: IEEE: 1122-1131 [DOI: 10.1109/cvprw.2017.150]
- Bavirisetti D P and Dhuli R. 2016. Fusion of infrared and visible sensor images based on anisotropic diffusion and Karhunen-Loeve transform. IEEE Sensors Journal, 16(1): 203-209 [DOI: 10.1109/jsen.2015.2478655]
- Dosovitskiy A and Brox T. 2016. Inverting visual representations with convolutional networks//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE: 4829-4837 [DOI: 10.1109/cvpr.2016.522]
- He Z and He D. 2020. Deep learning-based super-resolution for GF-4 satellite imagery. Journal of Remote Sensing (in Chinese), 24(12): 1500-1510 (贺智, 贺丹. 2020. 基于深度学习的高分四号卫星图像超分辨率重建. 遥感学报, 24(12): 1500-1510) [DOI: 10.11834/jrs.20208397]
- Hu X C, Mu H Y, Zhang X Y, Wang Z L, Tan T N and Sun J. 2019. Meta-SR: a magnification-arbitrary network for super-resolution//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE: 1575-1584 [DOI: 10.1109/CVPR.2019.00167]
- Kulkarni K, Lohit S, Turaga P, Kerviche R and Ashok A. 2016. ReconNet: non-iterative reconstruction of images from compressively sensed measurements//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE: 449-458 [DOI: 10.1109/cvpr.2016.55]
- Liu E, Logan J, Lee M, Roy S, Wang S, Valerie V M and Venkatesh V. 2019. WiDS Datathon 2019. Kaggle. <https://kaggle.com/competitions/widsdatathon2019>
- Mittal A, Soundararajan R and Bovik A C. 2013. Making a "completely blind" image quality analyzer. IEEE Signal Processing Letters, 20(3): 209-212 [DOI: 10.1109/lsp.2012.2227726]
- Pratt L and Thrun S. 1997. Guest editors' introduction. Machine Learning, 28(1): 5 [DOI: 10.1023/A:1007322005825]
- Ren S B, Meng Q and Wu Z. 2022. Flow-based super-resolution reconstruction of remote sensing images. Chinese Space Science and Technology, 42(6): 99-106 (任术波, 孟倩, 吴钻. 2022. 基于流的遥感图像超分辨率重建. 中国空间科学技术, 42(6): 99-106) [DOI: 10.16708/j.cnki.1000-758X.2022.0088]
- Thrun S and Pratt L. 1998. Learning to learn: introduction and overview//Learning to Learn. New York: Springer: 3-17 [DOI: 10.1007/978-1-4615-5529-2_1]
- Vilalta R and Drissi Y. 2002. A perspective view and survey of meta-learning. Artificial Intelligence Review, 18(2): 77-95 [DOI: 10.1023/A:1019956318069]
- Xia G S, Hu J W, Hu F, Shi B G, Bai X, Zhong Y F, Zhang L P and Lu X Q. 2017. AID: a benchmark data set for performance evaluation of aerial scene classification. IEEE Transactions on Geoscience and Remote Sensing, 55(7): 3965-3981 [DOI: 10.1109/tgrs.2017.2685945]
- Yang X, Li F, Lu M, Xin L, Lu X T and Zhang N. 2022. New super-resolution reconstruction method based on mixed sparse representations. National Remote Sensing Bulletin, 26(8): 1685-1697 (杨雪, 李峰, 鹿明, 辛蕾, 鲁啸天, 张南. 2022. 混合稀疏表示模型的超分辨率重建. 遥感学报, 26(8): 1685-1697) [DOI: 10.11834/jrs.20219409]
- Yang Y and Newsam S. 2010. 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: Association for Computing Machinery: 270-279 [DOI: 10.1145/1869790.1869829]
- Zhang Y L, Li K P, Li K, Wang L C, Zhong B N and Fu Y. 2018a. Image super-resolution using very deep residual channel attention networks//15th European Conference on Computer Vision. Mu-

nich: Springer: 294-310 [DOI: 10.1007/978-3-030-01234-2_18]
 Zhang Y L, Tian Y P, Kong Y, Zhong B N and Fu Y. 2018b. Residual dense network for image super-resolution//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE: 2472-2481 [DOI: 10.1109/CVPR.2018.00262]

Zhao H Y and Li Y H. 2022. Super-resolution analysis of Sentinel-2 images using generative adversarial networks. *Journal of Computer Applications*, 42(S1): 298-304 (赵慧岩, 李云鹤. 2022. 基于生成对抗网络的 Sentinel-2 遥感图像超分辨率分析. *计算机应用*, 42(S1): 298-304) [DOI: 10.11772/j.issn.1001-9081.2021060953]

Arbitrary-scale super-resolution reconstruction of remote sensing images based on meta-learning and residual dense attention mechanism

WEI Xiaoyuan¹, MENG Gang², ZHANG Haopeng¹, JIANG Zhiguo¹

1. School of Astronautics, Beihang University, Beijing 102206, China;

2. Beijing Institute of Remote Sensing Information, Beijing 100192, China

Abstract: Super-resolution reconstruction technology plays an important role in the intelligent processing of satellite remote sensing images. Existing deep learning methods for super-resolution remote sensing image reconstruction can only handle super-resolution tasks with a single scale factor, lacking generalization at the multiscale level and failing to meet the requirements of real super-resolution remote sensing image reconstruction for continuous zooming at multiple magnification levels.

To address the problem of arbitrary-scale super-resolution reconstruction and enhance the quality of reconstructed real remote sensing images, this paper proposes a super-resolution reconstruction method called the Meta-RDCAN, which utilizes meta-learning and residual dense channel attention network. The proposed method employs a meta-upscale module that incorporates three functions: weight prediction, location projection, and feature mapping. The module adaptively adjusts the internal parameters of a model according to different scale factors for arbitrary-scale super-resolution reconstruction. From the perspective of extracting detailed information of local land objects in a remote sensing image, a dense residual network with an attention mechanism is used as a feature extractor, enabling the reconstructed results to possess clear and distinguishable details. Extensive experiments are conducted on various datasets, including DIV2K, AID, UCMerced, WIDS, Set5, and real remote sensing image from Macao Science Popularization Satellite. The influence of variations in spatial resolution on the super-resolution reconstruction results is analyzed, and the effectiveness of the training scheme, which involves pretraining on a general dataset followed by fine-tuning on a remote sensing dataset, is validated using the loss curve.

The test results with different scale factors demonstrate that the proposed model is suitable for arbitrary-scale super-resolution reconstruction tasks in remote sensing images, having scale factors of up to 4.0. Comparative experimental results show that the improved model with added channel attention achieves enhanced performance in terms of Peak Signal-to-Noise Ratio (PSNR) and Structural SIMilarity (SSIM), compared with the baseline model. On real remote sensing data, the reconstruction results of the proposed model achieve a PSNR of over 40 dB and an SSIM of over 0.95. The comparison based on the no-reference quality indicator NIQE confirms that the perceived quality of the super-resolution reconstruction results of the improved model surpasses that of the baseline model.

The proposed method for arbitrary-scale super-resolution reconstruction is effective for remote sensing images by utilizing meta-learning and dense residual channel attention. The main contributions of this paper include two aspects. First, for the arbitrary-scale super-resolution reconstruction of remote sensing image, the meta-learning approach is employed to adaptively adjust the internal parameters of a model. This approach enables continuous integer- and non-integer-scale super-resolution reconstruction of a single remote sensing image with a single model. Second, to address issues, such as missing details and unclear edges of geographic features, in reconstruction results, a channel-attention mechanism is applied to enhance the dense residual network and improve the quality of super-resolution reconstruction results.

Key words: super-resolution reconstruction, remote sensing image, arbitrary-scale, meta-learning, residual dense network, channel attention mechanism

Supported by National Natural Science Foundation of China (No. 62271017)