

# 基于反投影注意力网络的遥感影像超分辨率重建<sup>①</sup>



胡家玥, 田鹏辉

(西安工业大学 计算机科学与工程学院, 西安 710021)

通信作者: 田鹏辉, E-mail: [tphxyz@163.com](mailto:tphxyz@163.com)

**摘要:** 近年来, 在诸如环境监测等一系列工作中, 遥感影像得到了广泛应用. 然而, 目前卫星传感器观测到的影像往往分辨率较低, 很难满足深入研究的需要. 超分辨率 (SR) 目的是提高图像分辨率, 同时提供更精细的空间细节, 完美地弥补了卫星图像的弱点. 因此, 本文提出了一种反投影注意力网络 (back-projection attention network, BPAN) 用于遥感图像的超分辨率重建, 该网络由反投影网络和初始残差注意力块两部分组成. 在反投影网络中, 通过迭代误差反馈机制计算上下投影误差指导图像重建; 在初始残差注意力块中, 引入初始模块融合局部多级特征为重建详细的纹理提供更丰富的信息, 以注意模块自适应地学习不同空间区域的重要性, 促进高频信息的恢复. 为了评价该方法的有效性, 在 AID 数据集上进行了大量的实验, 结果表明, 本文的网络模型提升了传统深度网络的重建性能, 且在视觉效果和客观指标方面有明显提升.

**关键词:** 遥感影响; 超分辨率重建; 注意力; 反投影网络

引用格式: 胡家玥, 田鹏辉. 基于反投影注意力网络的遥感影像超分辨率重建. 计算机系统应用, 2023, 32(12): 211-217. <http://www.c-s-a.org.cn/1003-3254/9317.html>

## Super Resolution Reconstruction of Remote Sensing Images Based on Back-projection Attention Network

HU Jia-Yue, TIAN Peng-Hui

(School of Computer Science and Engineering, Xi'an Technological University, Xi'an 710021, China)

**Abstract:** In recent years, remote sensing images have been widely employed in a series of work such as environmental monitoring. However, the images observed by satellite sensors often have low resolution, which is difficult to meet in-depth research needs. Super resolution (SR) aims to improve image resolution and provides finer spatial details, perfectly compensating for the weaknesses of satellite imagery. Therefore, a back-projection attention network (BPAN) is proposed for SR reconstruction of remote sensing images. The BPAN is composed of the back-projection network and the initial residual attention block. In the back projection network, the iterative error feedback mechanism is adopted to calculate the upper and lower projection errors to guide image reconstruction. In the initial residual attention block, the initial module is introduced to integrate local multilevel features to provide more information for reconstructing detailed textures to focus on the importance of the module to learn different spatial regions adaptively and promote high-frequency information recovery. To evaluate the effectiveness of this method, this study conducts a large number of experiments on AID datasets. The results show that the proposed network model improves the reconstruction performance of traditional deep networks and has significant improvements in visual effects and objective indicators.

**Key words:** remote sensing effect; super resolution (SR) reconstruction; attention; back-projection network

① 收稿时间: 2023-05-17; 修改时间: 2023-06-26; 采用时间: 2023-07-07; csa 在线出版时间: 2023-10-19  
CNKI 网络首发时间: 2023-10-20

图像超分辨率 (SR) 是计算机视觉中的一个热点问题, 其目的是从相应的低分辨率 (LR) 图像中恢复具有丰富高频细节的高分辨率 (HR) 图像。在遥感领域, 从卫星上观测到的图像可以在许多领域发挥重要作用并具有实用价值, 例如环境监测、资源勘探和监测。然而, 由于硬件和低成本成像传感器的限制, 很难直接获得高分辨率 (HR) 图像。因此, 低成本和高效的超分辨率重建技术 (SR) 应运而生, 充分利用低分辨率图像的信息, 重建出质量更好的高分辨率图像。

## 1 相关研究

对于任何给定的低分辨率 (LR) 图像都存在多重解。为了解决这一问题, 近年来提出了许多 SR 技术, 这些技术主要分为基于插值的方法<sup>[1]</sup>、基于重建的方法<sup>[2]</sup>以及基于学习的方法<sup>[3,4]</sup>。基于学习的方法可进一步分为邻域嵌入方法<sup>[5]</sup>、稀疏编码方法<sup>[6]</sup>和基于卷积神经网络 (CNN) 的方法。

Dong 等人<sup>[7]</sup>首先提出了一种三层卷积神经网络 SRCNN, 其性能优于许多传统的 SR 方法。Liebel 等人<sup>[8]</sup>首次提出将 SRCNN 用于遥感图像超分辨率重建。Lei 等人<sup>[9]</sup>设计了一种“分支”结构的网络 (local-global combined network, LGCnet), 来学习遥感图像的多尺度表示, 利用 CNN 随着网络深度加深感受野也随之扩大的特性, 将浅层和深层的特征映射进行级联从而实现局部与全局信息的结合。Haut 等人<sup>[10]</sup>将视觉注意力机制融入到基于残差的网络设计中 (remote sensing residual channel attention network, RSRGAN), 该机制能够引导网络训练过程朝向信息量最大的特征。Ahn 等人<sup>[11]</sup>提出了级联残差网络 (CARN), 该网络使用级联连接来融合局部和全局表示。

尽管这些基于 CNN 的方法在图像 SR 方面取得了重大进展, 但它们平等地对待不同的通道, 这在处理不同通道信息缺乏灵活性。为了解决这个问题, Hu 等人首先提出了 SENet (squeeze-and-excitation network)<sup>[12]</sup>, 该网络设计了一个通道注意模块来提高分类精度。受 SENet 的启发, Zhang 等人提出了一种用于自然图像 SR 的深度残差通道注意网络 (RCAN)<sup>[13]</sup>, 其中设计了一个通道注意 (CA) 模块, 自适应地学习不同通道的重要性。目前基于注意力的方法通常仅从单个级别提取特征。尽管一些文献也考虑了融合多级特征<sup>[9]</sup>, 但它们忽略了不同通道和不同空间区域的不同重要性。

在本文中, 提出反投影注意力网络来克服这些问题。每个初始残差注意块包含初始模块和注意模块。初始模块用于通过使用 3 个卷积层提取局部多级特征, 随后通过使用另一个特征融合层融合提取的特征。注意力模块结合通道注意力和空间注意力使网络更加关注高频信息的区域。

## 2 超分辨率重建方法

考虑到遥感影像丰富的纹理信息和复杂的空间位置信息, 本文提出了一种基于反投影注意力网络的超分辨率模型, 其中使用了密集连接的上投影单元和下投影单元<sup>[14]</sup>, 上投影单元的输出结果不仅作为下投影单元的输入还作为初始残差注意块 (inception residual attention blocks, IRAB)<sup>[15]</sup>的输入, 通过 IRAB 使网络更加关注高频信息的区域。

### 2.1 网络结构

考虑到遥感影像的特点, 本文提出了反投影注意力网络, 该网络采用迭代上采样策略, 其整体结构如图 1 所示, 该网络模型分为插值上采样分支和主网络分支。主网络分支又分为 4 个部分: 特征提取部分、迭代投影部分、重构部分、IRAB。

#### 2.1.1 插值上采样分支

使用上采样分支使我们的网络学习全局残差。该算法如图 1 所示, 可以用式 (1) 表示:

$$I_{\text{bicubic}} = H_{\text{bicubic}}(I_{\text{LR}}) \quad (1)$$

其中,  $H_{\text{bicubic}}(\cdot)$  表示使用双三次插值的上采样算法,  $I_{\text{bicubic}}$  表示插值后的图像。

#### 2.1.2 主网络分支

(1) 特征提取部分: 使用 Conv(3,  $n_0$ ) 从输入构建初始 LR 特征映射  $L_0$ 。然后使用 Conv(1,  $nR$ ) 将维数从  $n_0$  降至  $nR$ , 然后进入投影步骤, 其中  $n_0$  为初始 LR 特征提取中使用的卷积核个数,  $nR$  为每个投影单元中使用的卷积核个数。

(2) 投影部分: 初始特征提取之后是一系列投影单元, 交替构建 LR 和 HR 特征映射, 通过密集连接使每个投影单元都可以访问所有先前投影单元的输出。投影单元的结构如图 2 所示。

(3) 重构部分: 使用  $3 \times 3$  的卷积将拼接后的 HR 图像重构为 SR 图像, 其中 fRec 使用 conv(3, 3) 作为重构,  $[IRAB^1, IRAB^2, \dots, IRAB^j]$  是指在每个 IRAB 中产

生的特征图的拼接,称为深度拼接。

### 2.1.3 初始残差注意块 (IRAB)

IRAB 共分为两个模块,分别是 Inception 模块和

Attention 模块,使得 IRAB 通过 Inception 模块捕获多层信息,通过 Attention 模块是网络更加关注携带高频信息的区域,其结构如图 3 所示。

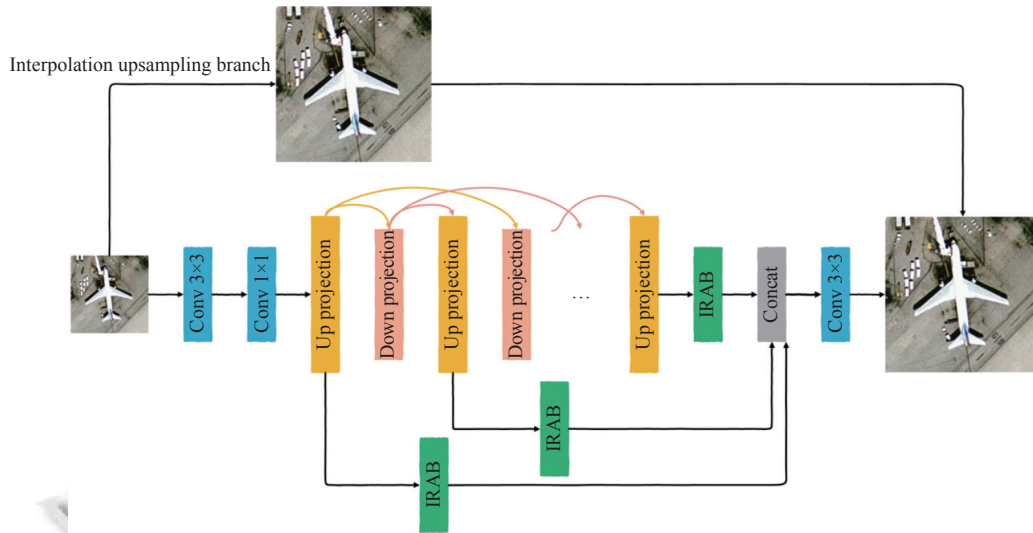


图 1 网络结构

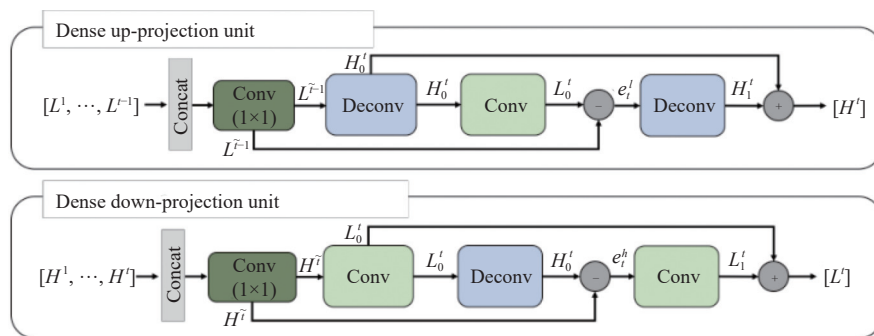


图 2 投影单元

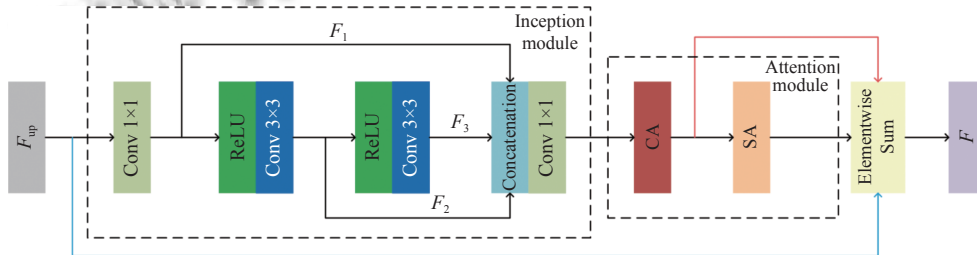


图 3 初始残差注意块

#### (1) Inception 模块

在 Inception 模块中,首先使用了 3 个卷积层(即 Conv(1, C), Conv(3, C), Conv(3, C)) 提取具有不同感知

尺度的特征,随后,利用另一个卷积层对提取的特征进行融合.假设第  $i$  个 IRAB 的输入是  $F_{up}$ , 4 个卷积层可以表示为:

$$\begin{cases} F_1 = f_{1,1}(F_{\text{up}}) \\ F_2 = f_{2,2}(\text{ReLU}(F_1)) \\ F_3 = f_{3,3}(\text{ReLU}(F_2)) \\ F_{\text{Inception}} = f_{4,1}([F_1, F_2, F_3]) \end{cases} \quad (2)$$

其中,  $f(\cdot)$  表示卷积操作,  $k$  是卷积核的大小,  $[\cdot]$  表示拼接操作,  $F_1$ 、 $F_2$  和  $F_3$  表示不同感知尺度的特征图,  $F_{\text{Inception}}$  是 Inception 模块的最终输出, 它将被用作注意力模块的输入。

## (2) Attention 模块

LR 遥感场景图像包含丰富的低频和有价值的高频特征。低频特征似乎更平坦, 高频特征通常是区域、充满了边缘、纹理和其他细节。在 SR 领域中, 高频通道特征对 HR 重构更为重要。因此我们使用 Attention 模块, 该模块由通道注意模块和空间注意模块两个子模块组成。

通道注意模块用于提取不同通道的重要性, 整个通道注意模块的过程如下:

$$F_c = A_{\text{channel}} \otimes F_{\text{Inception}} \quad (3)$$

$$F_{ca} = F_c + F_{\text{Inception}} \quad (4)$$

其中,  $F_c$  表示中间特征图,  $A_{\text{channel}}$  表示通道注意向量,  $F_{ca}$  表示输出的特征图,  $\otimes$  表示逐元素相乘。

考虑到不同的池化操作可能会收集到不同的重要特征, 本文使用两种池化操作来压缩输入特征图的信息, 之后得到两个通道向量, 其中每个通道向量的大小为  $R^{1 \times 1 \times C}$ , 随后将每个通道向量输入到具有一个隐藏层的共享全连接层 (SFCL), 使用 elementwise 求和后使用激活函数, 可以得到通道注意向量  $A_{\text{channel}}$ 。因此, 计算通道细化特征图  $F_{ca}$  过程如下:

$$F_c = \sigma(\text{SFCL}(\text{MaxPool}(F_{\text{Inception}})) + \text{SFCL}(\text{AvgPool}(F_{\text{Inception}}))) \otimes F_{\text{Inception}} \quad (5)$$

$$\text{SFCL}(\text{MaxPool}(F_{\text{Inception}})) = \sigma(W_u \delta(W_d \text{MaxPool}(F_{\text{Inception}}))) \quad (6)$$

$$\text{SFCL}(\text{AvgPool}(F_{\text{Inception}})) = \sigma(W_u \delta(W_d \text{AvgPool}(F_{\text{Inception}}))) \quad (7)$$

其中,  $\delta$  表示 ReLU 激活函数,  $\sigma$  表示 Sigmoid 函数,  $W_d$  和  $W_u$  分别是两个 MLP 全连接层的前后权重集。

空间注意模块使网络更加关注携带高频信息的区域, 整个空间注意模块的过程如下:

$$F_{sa} = A_{\text{spatial}} \otimes F_{ca} \quad (8)$$

其中,  $F_{sa}$  表示输出的特征图,  $A_{\text{spatial}}$  表示空间注意向量,  $\otimes$  表示逐元素相乘。

沿着  $F_{ca}$  的通道, 利用平均池化和最大池化操作, 对不同通道中每个位置的空间信息进行聚合, 分别生成两个特征图, 之后应用 Conv(5, 2) 得到融合后的特征图  $F_{\text{fused}}$ , 使用 Conv(1, 1) 将融合特征图压缩为一个通道, 最后使用 Sigmoid 函数得到空间特征权重  $A_{\text{spatial}}$ , 计算过程如下:

$$A_{\text{spatial}} = \sigma(f(F_{\text{fused}})) \quad (9)$$

$$F_{\text{fused}} = \delta(f([\text{MaxPool}(F_{ca}), \text{AvgPool}(F_{ca})])) \quad (10)$$

其中,  $\delta$  表示 ReLU 激活函数,  $\sigma$  表示 Sigmoid 函数,  $f(\cdot)$  表示卷积操作。

## 2.2 损失函数

损失函数指导着网络优化的方向, 目前已经提出了许多损失函数, 例如 L1 损失和 L2 损失。其中, L1 损失函数已被证明在 SR 任务中更有用, 因此本文使用 L1 损失函数, 给定一个训练集  $\{I_{\text{LR}}^i, I_{\text{HR}}^i\}_{i=1}^N$ , 包含  $N$  个 LR 图像及其对应的 HR 图像, 利用式 (11) 最小化 L1 损失函数, 得到参数  $\theta$ 。

$$L(\theta) = \frac{1}{N} \sum_{i=1}^N \| \hat{I}_{\text{HR}}^i - I_{\text{HR}}^i \| \quad (11)$$

## 3 实验分析

### 3.1 数据处理及环境配置

本文选用 AID<sup>[16]</sup> 的遥感影像作为实验数据集, 该数据集来自 Google Earth, 包含 30 个类别的场景图像, 其中每个类别有约 220–420 张, 整体共计 10000 张, 其中每张像素大小为 600×600。从每个类中随机抽取 10 张图像组成测试集, 从每个类中随机抽取 5 张图像组成验证集, 剩余的 9550 张图像组成训练集, 在实验过程中首先将 RGB 通道转为 YCrCb 模式, 将 Y 通道的数据提取出来作为主要处理对象, 通过对 Y 通道在特定的缩放因子上进行双三次插值下采样得到 LR 图像。最后, 使用数据增强技术扩展训练集, 并随机旋转 90°、180°、270° 和翻转。

为了验证本文网络的性能, 本文的实验环境如表 1 所示。

在训练时, 针对不同的缩放因子, 投影单元中所使用的卷积核的大小是不同的。对于 ×2 的任务, 使用



6×6, stride=2, padding=2的卷积核;对于×3的任务,使用7×7, stride=3, padding=2的卷积核;对于×4的任务,使用8×8, stride=4, padding=2的卷积核.采用Adam优化器对网络进行优化,其中Adam动量参数设置为 $\beta_1=0.9$ ,  $\beta_2=0.999$ ,  $\epsilon=10^{-8}$ , 初始学习率为 $1 \times 10^{-4}$ , 50个epoch后,学习率降低到初始学习率的一半.

表1 实验环境

环境	配置
系统	Ubuntu 20.04
CPU	Intel(R) Xeon(R) Platinum 8255C
内存	40 GB
GPU	RTX 2080 Ti
框架	Python 3.8, PyTorch 1.10.0

### 3.2 评价指标

本文选择峰值信噪比 (peak signal-to-noise ratio,  $PSNR$ )<sup>[17]</sup>和结构相似度 (structure similarity index,  $SSIM$ )<sup>[18]</sup>作为评价指标,其公式如式(12)、式(14):

$$PSNR(I, \hat{I}) = 10 \log \frac{255^2}{MSE(I, \hat{I})} \quad (12)$$

其中,

$$MSE(I, \hat{I}) = \frac{\sum_{w=1}^W \sum_{h=1}^H (I(h, w) - \hat{I}(h, w))^2}{H \times W} \quad (13)$$

其中,  $W$ 和 $H$ 分别是图像的宽度和高度.  $PSNR$ 值越大, 重构效果越好.

$$SSIM(I, \hat{I}) = \frac{(2\mu_I \mu_{\hat{I}} + C_1)(\sigma_{I\hat{I}} + C_2)}{(\mu_I^2 + \mu_{\hat{I}}^2 + C_1)(\sigma_I^2 + \sigma_{\hat{I}}^2 + C_2)} \quad (14)$$

$SSIM(\cdot)$ 表示HR图像 $I$ 与重建后的SR图像 $\hat{I}$ 在亮度、对比度和结构上的相似性.  $SSIM$ 值越大, 表示图像质量越高.

### 3.3 实验结果及分析

为了验证本文方法的重建性能, 选择SRCNN、SAN、DSSR和DASR这4种方法进行2倍、3倍和4倍的重建实验对比, 实验结果见表2.

表2给出了本文方法及对比方法在不同数据集上的实验结果, 并以加粗字体表示最优结果, 从表格可以看出, 本文提出的方法在两个指标上基本优于比较方法, 只在个别类中略低于比较方法, 但差值并不显著. 通过观察所有测试图像的平均 $PSNR$ 和平均 $SSIM$ , 证明了本文模型在超分辨率重建任务上的有效性.

表2 不同方法在部分数据集上的 $PSNR$ (dB)和 $SSIM$ 结果

Image	Scale	SRCNN	SAN	DSSR	DASR	Ours
bareland	×2	37.62/0.9325	38.97/0.9645	38.99/0.9649	39.04/0.9658	<b>39.10/0.9665</b>
	×3	33.89/0.7873	36.64/0.8917	36.68/0.8923	36.70/0.8928	<b>36.79/0.8933</b>
	×4	32.25/0.7967	34.51/0.8530	34.52/0.8533	34.57/0.8537	<b>34.59/0.8542</b>
desert	×2	33.64/0.8451	35.33/0.9143	35.32/0.9146	35.36/0.9148	<b>35.49/0.9155</b>
	×3	30.13/0.7554	32.27/0.8492	32.30/0.8499	32.31/0.8497	<b>32.37/0.8506</b>
	×4	29.56/0.7064	30.19/0.7660	30.21/0.7660	30.23/0.7663	<b>30.30/0.7668</b>
farmland	×2	32.60/0.8651	34.75/0.8923	34.76/0.8925	34.79/0.8931	<b>34.81/0.8926</b>
	×3	29.80/0.7734	30.99/0.8072	30.97/0.8071	31.01/0.8077	<b>31.05/0.8075</b>
	×4	27.94/0.7325	29.30/0.7661	29.30/0.7663	29.33/0.7665	<b>29.33/0.7668</b>
playground	×2	31.96/0.8657	34.46/0.9020	34.50/0.9026	34.52/0.9030	<b>34.58/0.9039</b>
	×3	29.30/0.7845	30.39/0.8328	30.42/0.8332	30.47/0.8336	<b>30.47/0.8338</b>
	×4	27.85/0.7425	28.60/0.7783	28.61/0.7789	28.66/0.7794	<b>28.69/0.7798</b>
bridge	×2	29.56/0.7959	32.24/0.8845	32.27/0.8846	32.32/0.8851	<b>32.35/0.8854</b>
	×3	26.20/0.7869	27.82/0.8026	27.84/0.8029	27.89/0.8032	<b>27.86/0.8032</b>
	×4	23.67/0.6578	25.22/0.7642	25.23/0.7646	25.25/0.7649	<b>25.30/0.7655</b>

图4-图6给出了5种方法在×2、×3和×4不同放大倍数下对遥感影像测试集中部分影响的重建结果, 并与真实高分影像(GT)进行对比. 从图中可以看出本文所提方法重建得到的图像与原始图像更加接近, 影像细节更加丰富, 说明本文方法能够更好地重建出细节特征.

## 4 结论与展望

本文提出了一种基于反投影注意力网络的超分辨

率重建方法, 在该网络中, 联合了深度反投影网络和初始残差注意块. 利用上、下投影单元的误差反馈来引导网络达到更好的效果, 同时将上投影单元的输出结果作为初始残差注意块的输入, 通过初始模块融合不同网络深度的多层次特征, 为重构提供更详细的信息, 同时不引入过多的参数; 通过注意模块使网络更加关注高频信息的区域, 有助于网络恢复更完整的边缘和更详细的纹理. 在测试图像上的实验结果表明, 与其他方法相比, 该方法可以获得较好的效果.

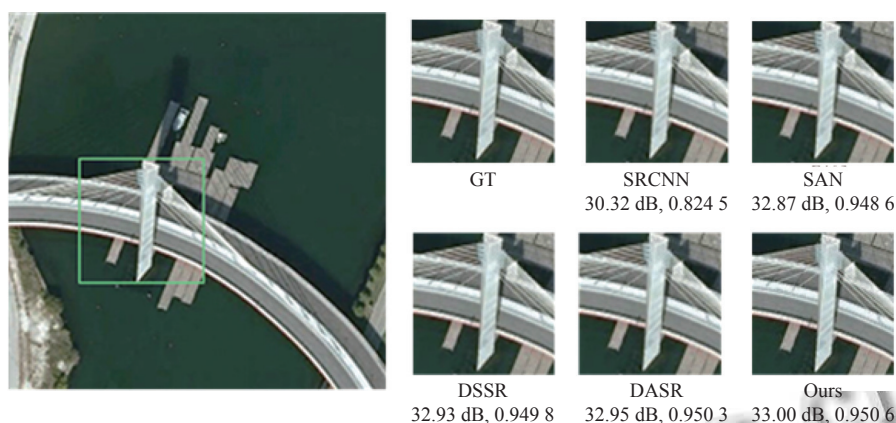


图4 bridge 测试集 2 倍超分辨率重建结果对比

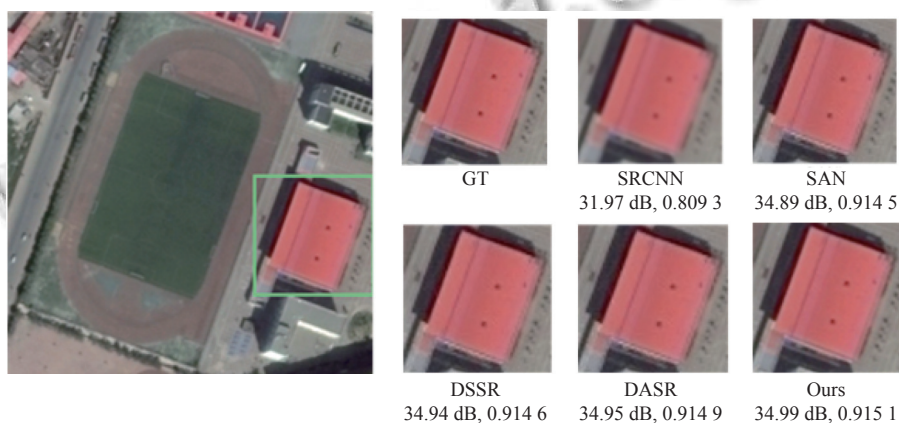


图5 playground 测试集 3 倍超分辨率重建结果对比

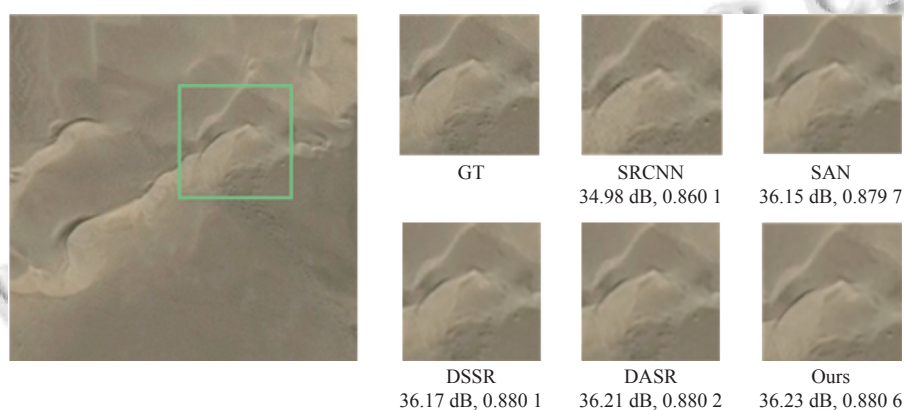


图6 desert 测试集 4 倍超分辨率重建结果对比

参考文献

1 Zhang L, Wu XL. An edge-guided image interpolation algorithm via directional filtering and data fusion. IEEE Transactions on Image Processing, 2006, 15(8): 2226–2238. [doi: 10.1109/TIP.2006.877407]

2 Kim KI, Kwon Y. Single-image super-resolution using

sparse regression and natural image prior. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(6): 1127–1133. [doi: 10.1109/TPAMI.2010.25]

3 Huang JB, Singh A, Ahuja N. Single image super-resolution from transformed self-exemplars. Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern

- Recognition. Boston: IEEE, 2015. 5197–5206.
- 4 Schuler S, Leistner C, Bischof H. Fast and accurate image upscaling with super-resolution forests. Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015. 3791–3799.
  - 5 Chang H, Yeung DY, Xiong YM. Super-resolution through neighbor embedding. Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. Washington: IEEE, 2004. 1.
  - 6 Yang JC, Wright J, Huang TS, *et al.* Image super-resolution via sparse representation. IEEE Transactions on Image Processing, 2010, 19(11): 2861–2873. [doi: [10.1109/TIP.2010.2050625](https://doi.org/10.1109/TIP.2010.2050625)]
  - 7 Dong C, Chen CL, He KM, *et al.* Learning a deep convolutional network for image super-resolution. Proceedings of the 13th European Conference on Computer Vision. Zurich: Springer, 2014. 184–199.
  - 8 Liebel L, Körner M. Single-image super resolution for multispectral remote sensing data using convolutional neural networks. ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Copernicus Publications, 2016. 883–890.
  - 9 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)]
  - 10 Haut JM, Fernandez-Beltran R, Paoletti ME, *et al.* Remote sensing image superresolution using deep residual channel attention. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(11): 9277–9289. [doi: [10.1109/TGRS.2019.2924818](https://doi.org/10.1109/TGRS.2019.2924818)]
  - 11 Ahn N, Kang B, Sohn KA. Fast, accurate, and lightweight super-resolution with cascading residual network. Proceedings of the 15th European Conference on Computer Vision. Munich: Springer, 2018. 256–272.
  - 12 Hu J, Shen L, Albanie S, *et al.* Squeeze-and-excitation networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011–2023. [doi: [10.1109/TPAMI.2019.2913372](https://doi.org/10.1109/TPAMI.2019.2913372)]
  - 13 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. 294–310.
  - 14 Haris M, Shakhnarovich G, Ukita N. Deep back-projection networks for super-resolution. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 1664–1673.
  - 15 Lei PC, Liu C. Inception residual attention network for remote sensing image super-resolution. International Journal of Remote Sensing, 2020, 41(24): 9565–9587. [doi: [10.1080/01431161.2020.1800129](https://doi.org/10.1080/01431161.2020.1800129)]
  - 16 Xia GS, Hu JW, Hu F, *et al.* AID: A benchmark data set for performance evaluation of aerial scene classification. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(7): 3965–3981. [doi: [10.1109/TGRS.2017.2685945](https://doi.org/10.1109/TGRS.2017.2685945)]
  - 17 褚江, 陈强, 杨曦晨. 全参考图像质量评价综述. 计算机应用研究, 2014, 31(1): 13–22.
  - 18 Kim J, Lee JK, Lee KM. Accurate image super-resolution using very deep convolutional networks. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016. 1646–1654.

(校对责编: 牛欣悦)