

文章编号 1004-924X(2025)08-1238-21

面向单幅遥感图像超分辨率的空间自适应及 频率融合网络

杨沂川¹, 马中祺², 周新尧¹, 郑福建¹, 黄 鸿^{1*}

(1. 重庆大学 光电技术与系统教育部重点实验室, 重庆 401331;

2. 北京空间机电研究所, 北京 100094)

摘要: 现有大多遥感图像超分辨率方法, 无法充分挖掘图像中混合尺度的自相似性信息和跨尺度区域间的关联信息, 且忽略了频率域对感知图像高频信息的能力。针对这一问题, 本文提出了一种空间自适应及频率融合网络 (Spatial Adaptation and Frequency Fusion Network, SAF²Net)。SAF²Net 首先引入一种混合尺度空间自适应特征调制模块, 采用类似于特征金字塔的方式获取不同尺度下的判别特征, 丰富多尺度特征的表达能力。随后, 设计了一个全局多尺度感受野选择块, 挖掘跨尺度区域间的关联特征。在此基础上, 引入空间自适应选择块和频率分离选择块, 融合空间-频率互补信息以增强局部特征, 提高模型对图像高频内容的建模能力。在两个公开遥感图像数据集上进行多组实验, SAF²Net 获得的定量评价指标结果均优于其他对比方法。以 UC Merced 数据集 3 倍超分辨率为例, 本文方法相较于次优方法 HAUNet, PSNR 和 SSIM 分别提升了 0.11 dB 和 0.003 3; 在主观视觉质量方面, SAF²Net 能够恢复出更多清晰的纹理细节。实验结果表明, 本文所提出的 SAF²Net 能够从两个不同的角度挖掘混合尺度全局信息, 并有效融合空间-频率互补特征, 在遥感图像超分辨率任务中表现出具有竞争力的重建性能。

关键词: 遥感图像; 超分辨率; 混合尺度特征; 空频互补信息

中图分类号: TP391 文献标识码: A

doi: 10.37188/OPE.20253308.1238

CSTR: 32169.14.OPE.20253308.1238

Spatial adaptation and frequency fusion network for single remote sensing image super-resolution

YANG Yichuan¹, MA Zhongqi², ZHOU Xinyao¹, ZHENG Fujian¹, HUANG Hong^{1*}

(1. Key Laboratory of Optoelectronic Technology and System, Ministry of Education,
Chongqing University, Chongqing 401331, China;

2. Beijing Institute of Space Machinery and Electronics, Beijing 100094, China)

* Corresponding author, *E-mail: hhuang@cqu.edu.cn;

Abstract: Most of the existing methods of remote sensing image super-resolution are unable to fully explore the self-similarity information at hybrid scales and the correlation between cross-scale regions. Moreover, these methods ignore the ability of the frequency domain to perceive the high-frequency information of the images. To address this problem, a Spatial Adaptation and Frequency Fusion Network (SAF²Net)

收稿日期: 2024-12-16; 修订日期: 2025-02-11.

基金项目: 北京市航空智能遥感装备工程技术研究中心开放基金 (No. AIRSE202412); 国家自然科学基金 (No. 42071302)

was proposed. Firstly, SAF²Net introduced a hybrid-scale spatially-adaptive feature modulation, which adopted a feature pyramid-like approach to obtain discriminative features at different scales and enriched the expression ability of multi-scale features. Subsequently, a global multi-scale field selection block was designed to extract the correlation features of cross-scale regions. On this basis, a spatial adaptively selection block and a frequency separation selection block were introduced to fuse space-frequency complementary information to enhance local features, improving the model's ability to model the high-frequency content of images. Multiple sets of experiments are conducted on two remote sensing image datasets, which indicates that the quantitative evaluation metrics obtained by SAF²Net outperform those of other comparative methods. Taking the UCMerced dataset with 3 times super-resolution as an example, the proposed method improves PSNR and SSIM by 0.11 dB and 0.003 3, respectively, in compared with the next best method HAUNet. In terms of the subjective visual quality, SAF²Net is able to recover more clear texture details. The experimental results demonstrate that the SAF²Net proposed is capable of mining the hybrid-scale global information from two different perspectives as well as fusing the space-frequency complementary features effectively, which exhibits competitive performance in the remote sensing image super-resolution task.

Key words: remote sensing image; super resolution; hybrid-scale features; space-frequency complementary information

1 引言

随着遥感应用的日益发展,遥感图像提供了丰富的空间细节和地物特征,在土地覆盖分割^[1]、分类^[2]、目标精确识别^[3]以及变化监测^[4]等诸多研究领域中,扮演着至关重要的角色。然而,由于成像设备的局限性和空间成像条件恶劣,遥感图像质量经常受到运动模糊、大气干扰、超远程成像等因素的影响,导致图像受到不同程度的退化,从而丢失大量的空间细节和地物特征^[5]。因此,提高遥感图像的空间分辨率对于诸多遥感解译任务至关重要。在实际应用中,与直接升级卫星成像设备提高遥感图像空间分辨率相比,超分辨率技术更加灵活和经济^[6]。该技术旨在通过算法从低分辨率(Low Resolution, LR)图像重建出相应高分辨率(High Resolution, HR)图像,并在一定程度上恢复其丢失的空间细节信息。

根据输入图像数量的不同可以分为单幅图像超分辨率(Single Image Super Resolution, SISR)和多幅图像超分辨率(Multiple Image Super-Resolution, MISR)方法。而由于在实际中遥感图像的获取容易受到云雾等掩体的遮挡和物体运动的影响^[7],难以在同一空间位置获取到多幅可用的遥感图像。由此可见,SISR相较于

MISR在遥感图像超分辨率重建任务中具有更广泛的适用性,已成为一个重要的研究课题。

SISR是一个典型的不适定问题(ill-Posed Problem),即重建HR图像需要从输入LR图像中有限的像素信息出发,推断并生成更多的高分辨率像素,而这一过程并非唯一确定,一幅LR图像可能存在多种对应的HR图像,即存在多种可能的解^[8]。早期的相关工作通常基于各种图像先验,来约束无限的HR解空间。双三次插值及其改进算法^[9-10]等基于插值的算法没有较高的复杂度,实时性较好,但由于其没有任何数据外部的先验信息,仅仅依靠局部图像先验而不能结合图像的语义信息进行重建,因此重建结果往往存在边缘、轮廓和高频细节模糊的问题。而早期基于学习的图像超分辨率算法,如基于领域嵌入的算法^[11],基于稀疏表示的算法^[12]以及基于局部线性回归的算法^[13],大多假设LR图像块和对应的HR图像块分布在不同的子空间上且具有相似的局部流形结构,通过学习字典进行图像重建。然而,上述方法往往受限于手工特征的不足和复杂的参数调整策略,无法高效利用图像本身的先验信息,导致重建结果产生过拟合和欠拟合的问题,且应用于地物分布复杂多变的遥感场景时上述算法的性能会急剧下降。

近年来,深度学习在自然图像超分辨率重建任务中表现出了卓越的性能。2016年,第一个基于卷积神经网络(Convolutional Neural Network, CNN)的超分辨率网络^[14]通过三层简单的网络学习插值放大后的低分辨率图像到高分辨率图像之间的映射关系,得到了相较于传统方法更好的重建结果。此后,大量基于CNN的模型被应用于遥感图像超分辨率重建领域,并体现出显著优势^[15]。Lei等^[16]提出一种局部-全局结合网络,旨在融合网络不同层的特征表示来提高重建效果。Xu等^[17]针对多个尺度特征提取的问题,提出一种具有对称结构的深度记忆连接网络,进一步提升了重建效果。Jiang等^[18]提出了一个逐步增强遥感图像的空间分辨率超分网络,该网络利用预训练CNN网络提取低级特征,利用增强型密集连接网络提取结构信息,从而有效增强特征的提取和表达能力。Haut等^[19]通过在局部和全局图像区域提取更丰富的特征,来优化网络设计,减少网络深度的同时提升了超分结果的视觉效果。然而,上述方法并没有充分考虑遥感图像本身的特点,即大多包含尺度不一、纹理多变的地物信息,且同一场景中不同尺度的纹理和结构信息具有一定的相似性。受限于CNN的局部感受野,其仅针对单一尺度的特征进行挖掘,且对精细纹理的小型地物和粗糙纹理的大型地物进行相同的卷积操作,限制了模型对特征的建模能力。

为了克服上述限制,Dong等^[20]利用注意力机制来增强网络的特征表达能力,对多个中间特征进行上采样并相加,从而实现同时关注浅层特征和深层特征。Zhang等^[21]引入连接注意力模块结构,将先前提取的注意力特征与当前特征进行融合,以确保信息在所有注意力模块之间的前向传播,有效利用不同层次特征以提升注意力模块的学习能力。Huan等^[22]通过融合层次特征并结合多尺度膨胀残差块来增强网络的表征能力,提出了一种新的金字塔多尺度残差网络,有效提取遥感图像中包含的多尺度特征。Lei等^[23]提出了一种利用多尺度自相似的网络,通过单尺度和多尺度自相似性,使网络能够在不同尺度间进行特征信息的递归传递和有效集成,从而增强了特征

表达能力,提高了网络对复杂场景的建模能力。随后,Lei等^[24]又提出了一种基于Transformer架构的遥感图像超分辨率网络,通过融合多尺度的高维和低维特征并捕获高维和低维特征之间的依赖关系,提高了遥感图像超分辨率的重建效果。Wang等^[25]提出了一种U形结构的遥感图像超分辨率网络,有效地探索多尺度特征,并通过混合卷积非局部注意力增强全局特征表示能力。Xiao等^[26]通过引入自上而下的选择机制,优化Transformer自注意力机制中的冗余信息,并通过多尺度特征提取增强模型的性能。尽管上述这些遥感图像超分辨率方法取得了较好的性能,但通过简单的插值或卷积无法有效提取不同尺度更有益于超分辨率任务的特征;且仅使用固定大小的卷积核无法进一步学习不同大小区域之间的关联信息,限制了模型的重建性能。此外,大多遥感图像超分辨率方法仅仅针对空间域进行像素级的学习和约束,可能会导致不同像素之间的细微差异混淆,缺乏进一步感知图像高频内容信息的能力。

针对上述问题,本文提出了一种面向单幅遥感图像超分辨率的空间自适应及频率融合网络(Spatial Adaptation and Frequency Fusion Network, SAF²Net),主要贡献如下:(1)设计了一个混合尺度空间自适应特征调制模块以充分利用混合尺度上的自相似信息,挖掘各个尺度中互补的特征信息,丰富模型对多尺度特征的表达能力。(2)设计了一个全局多尺度感受野选择块,通过融合不同大小卷积核提取的特征,一定程度解决局部CNN单一感受野的局限性,提取不同大小区域之间的关联信息,增强模型对不同大小区域特征的建模能力。(3)对于局部特征提取部分,设计了一个结合空间自适应选择和频率分离选择的局部增强特征提取模块,分别从空间域和频率域两个维度,使网络重点关注难以重建的区域,进一步激活有代表性的高频特征信息,细化重建细节。与其他图像超分辨率方法相比,本文方法在两个公开遥感图像数据集和实际遥感数据上的实验结果在定量评价指标及重建图像可视化上均表现出更好的性能,验证了本文方法的有效性。

2 本文方法

2.1 整体网络结构

针对超分辨率过程中不同区域重建难度不一的问题,本文通过充分挖掘和融合遥感图像中

丰富的混合多尺度先验信息,以及融合空间-频率互补信息,提出了一种遥感图像超分辨率网络 SAF²Net,网络结构如图1所示。该模型采用后上采样策略,主要包含三个部分:浅层特征提取块,深层特征提取块和图像上采样重建。

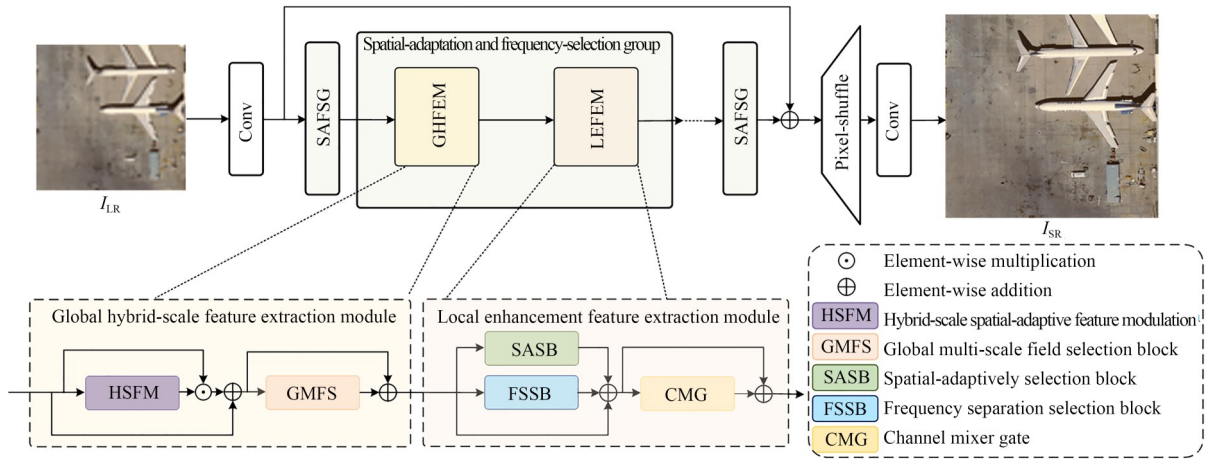


图1 SAF²Net网络整体结构

Fig. 1 Overall structure of the SAF²Net

该网络首先通过一个卷积核大小为 3×3 卷积层 H_{SFE} 将输入的低分辨率图像 I_{LR} 从图像空间转换到特征空间,获得通道数为64的浅层特征 F_0 ,其过程可以表示为:

$$F_0 = H_{SFE}(I_{LR}), \quad (1)$$

其中, $H_{SFE}(\cdot)$ 表示网络的第一层卷积层,即浅层特征提取块。

深层特征提取模块由 m 个空间自适应及频率选择组件 (Spatial-Adaption and Frequency-Selection Group, SAFSG) 构成。浅层特征 F_0 作为第一个深层特征提取模块 SAFSG 的输入,随后通过级联的多个 SAFSG,进一步挖掘和自适应融合图像中的全局混合尺度特征并细化增强局部特征,从而得到增强后的深层特征 F_i ,可以表示为:

$$F_i = H_{SAFSG_i}(F_{i-1}) = H_{SAFSG_i}(H_{SAFSG_{i-1}}(\cdots H_{SAFSG_0}(F_0)\cdots)), \quad (2)$$

其中: $H_{SAFSG_i}(\cdot)$ 表示第 i 个空间自适应及频率选择组件, F_{i-1} 和 F_i 分别是 $H_{SAFSG_i}(\cdot)$ 对应的输入和输出。SAFSG 由全局混合尺度特征提取模块和局部增强特征提取模块组成,接下来将分别在 2.2 节和 2.3 节详细介绍。

为了恢复 HR 目标图像,采用可学习的轻量级上采样层进行快速重建。其中最后一个深层特征提取块 H_{SAFSG_m} 的输出 F_m 与浅层特征 F_0 通过长距离跳跃连接进行融合,作为图像上采样重建部分的输入,防止梯度爆炸的同时加速网络的收敛,如式(3)所示:

$$I_{SR} = \text{Conv}(H_{\text{Rec}}(F_m + F_0)), \quad (3)$$

其中: $H_{\text{Rec}}(\cdot)$ 采用亚像素卷积^[27]操作对特征进行上采样,并通过卷积操作将深层特征映射回与输入 I_{LR} 相同的图像空间,得到最终的重建图像 I_{SR} 。

2.2 全局混合尺度特征提取模块

全局混合尺度特征提取模块 (Global Hybrid-scale Feature Extraction Module, GHFEM) 从两个不同的角度挖掘图像中多尺度特征,有益于模型充分利用全局先验信息,其主要由两部分组成:(1)混合尺度空间自适应特征调制模块 (Hybrid-scale Spatially-adaptive Feature Modulation, HSBM) 从 LR 图像的不同尺度特征表示中挖掘图像内部的特征并进行特征筛选,对深层特征进行调制,引导模型利用不同尺度重复且关键的先验信息进行超分辨率重建;(2)全局多尺度感受野选择块 (Global Multi-scale Field Selec-

tion Block, GMFS)通过不同大小感受野的卷积操作,从不同大小区域中产生多个全局上下文特征,并利用自适应选择机制滤除冗余特征同时融合不同区域产生的特征之间语义信息的关联性,充分利用图像中跨尺度区域互补的先验信息。

2.2.1 混合尺度空间自适应特征调制模块

遥感图像中同一尺度和不同尺度的某些纹理和结构信息总是重复出现,而这些混合尺度的重复先验信息有助于图像超分^[28]。基于此本文提出了一个混合尺度空间自适应特征调制模块 HSFM,其结构如图 2(a)所示。

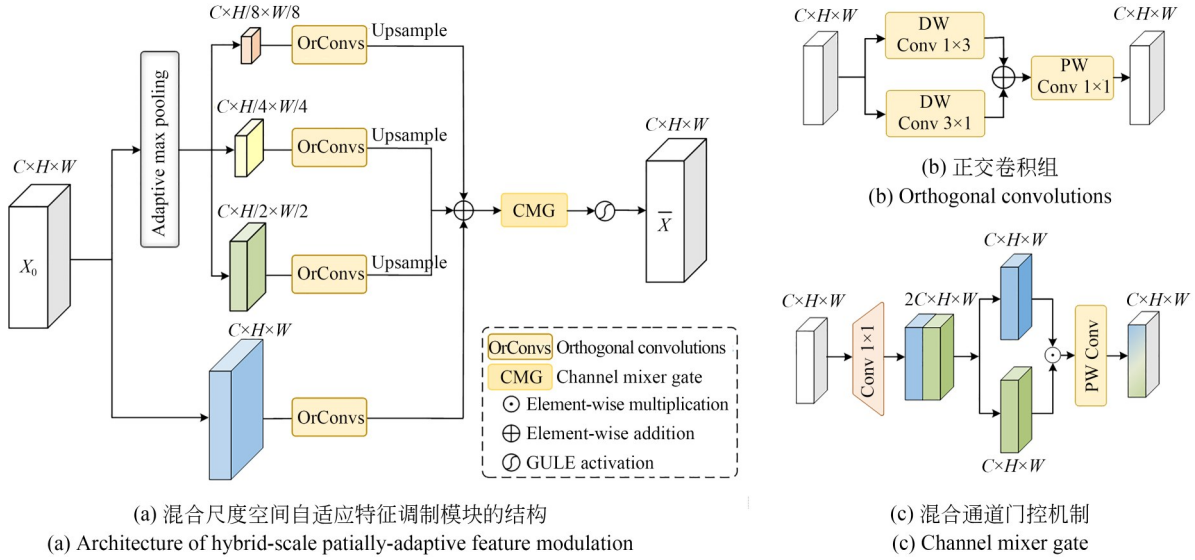


图 2 混合尺度空间自适应特征调制模块

Fig. 2 Hybrid-scale partially-adaptive feature modulation

HSFM 利用一种类似于特征金字塔的方式提取不同尺度的特征,并进行自适应筛选,用以调制深层特征。对于给定的输入特征 $X_0 \in \mathbb{R}^{C \times H \times W}$, 首先将金字塔式自适应最大池化作用于 X_0 以生成不同尺度下的判别特征 $X_i (i = 1, 2, 3)$, 随后通过正交卷积(Orthogonal Convolutions, OrConvs)分别对输入特征 X_0 和生成的判别特征 X_i 进行垂直和水平方向的深度卷积, 细化特征的同时捕捉不同方向的空间信息变化, 并使用一个点卷积融合各个通道之间的上下文关系。最后, 通过一个混合通道门控机制(Channel Mixer Gate, CMG)进一步去除特征相加时带来的冗余信息并更有效地融合不同尺度特征间的非线性互补关系, 上述过程可以表示为:

$$\begin{cases} \bar{X}_0 = \text{OrConvs}(X_0) \\ \bar{X}_i = \uparrow_{(H, W)} \left(\text{OrConvs} \left(\downarrow_{(H/2^i, W/2^i)} (X_0) \right) \right), \\ i = 1, 2, 3, \end{cases} \quad (4)$$

$$\bar{X} = \text{GELU} \left(\text{CMG} \left(\bar{X}_0 + \bar{X}_1 + \bar{X}_2 + \bar{X}_3 \right) \right), \quad (5)$$

其中: $\text{OrConvs}(\cdot)$ 表示正交卷积, 包括垂直和水平方向的深度卷积以及点卷积操作, $\downarrow_{(H/2^i, W/2^i)}$ 表示通过自适应最大池化(Adaptive Max Pooling)将输入特征 X_0 进行下采样, 得到不同尺度的最具判别性的特征 $\bar{X}_i \in \mathbb{R}^{C \times (H/2^i) \times (W/2^i)}$, $\uparrow_{(H, W)}$ 表示通过最近邻插值将细化后的各尺度特征上采样到原始尺寸大小 (H, W) , $\text{CMG}(\cdot)$ 表示通道混合门控机制, 通过 1×1 卷积升维以丰富特征表示, 并沿通道维度将特征分为两部分, 按元素相乘以去除混合尺度中的冗余信息。最后通过 GELU 非线性激活函数以丰富特征表达, 得到融合混合尺度信息的调制系数 $\bar{X}$ 。

进而, 利用得到的调制系数 $\bar{X}$ 与输入特征逐元素相乘实现对深层特征的自适应调制, 并通过残差连接结构提高网络的学习能力, 以更有效地捕获混合尺度全局先验信息。

2.2.2 全局多尺度感受野选择块

传统卷积操作只能探索局部像素之间的关系,其有限的感受野限制了模型的性能,且为了增大感受野一般需要增加网络的深度,这导致了参数数量的急剧增加以及模型收敛的难度。非局部注意力^[29]可以认为是一种自注意力机制,其通过卷积的方式计算整个输入像素与对应区域中像素之

间的关系,从而获得全局上下文特征,一定程度上缓解了上述问题。然而,遥感图像中不同大小区域包含的信息是相对不同的,如较大感受野包含全局先验信息,中等感受野关注图像中主体的结构信息,较小感受野提取图像中的边缘和纹理特征。基于上述考虑,本文设计了一个全局多尺度感受野选择模块 GMFS,如图 3(a)所示。

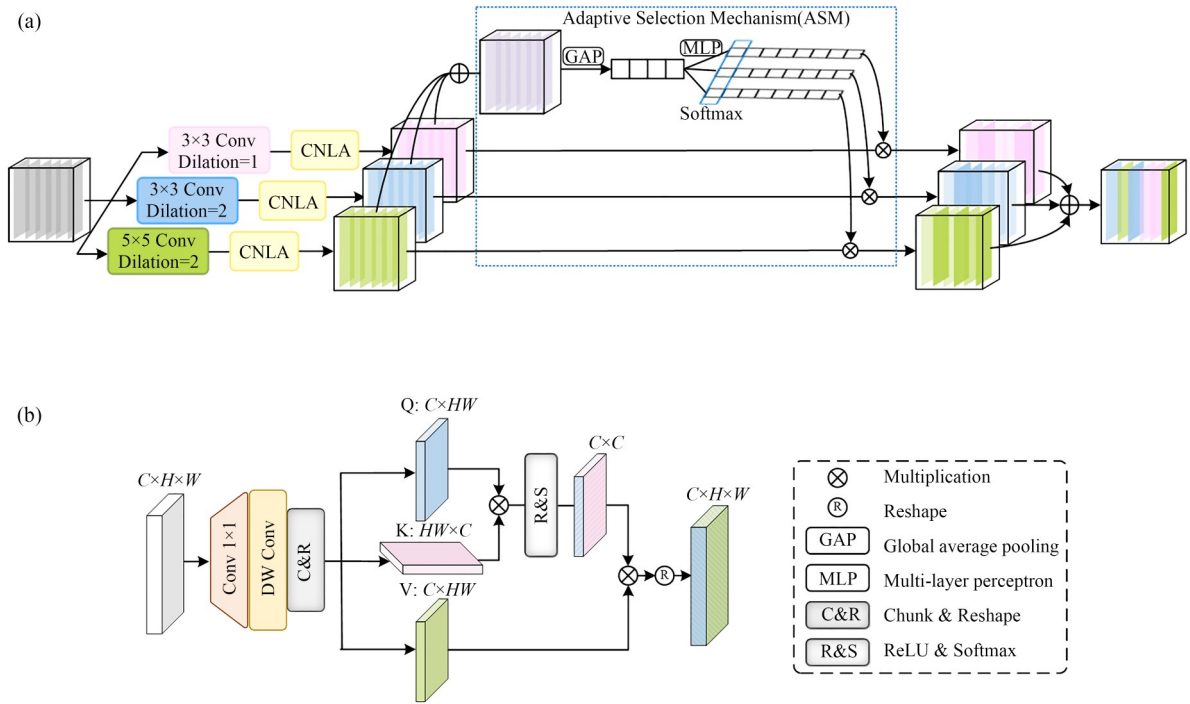


图 3 (a) 全局多尺度感受野选择块; (b) 通道非局部注意力
Fig. 3 (a) Global Multi-scale Field Selection Block; (b) Channel Non-local Attention

GMFS 从不同范围的感受野中产生多个全局上下文特征,并根据跨尺度区域关联信息进行自适应选择以探索最有用的上下文特征。具体而言,令输入特征为 $F \in \mathbb{R}^{C \times H \times W}$,分别用卷积核大小为 $3 \times 3, 3 \times 3, 5 \times 5$ 、空洞率为 1, 2, 2 的三个深度卷积,在不过度增加网络参数数量的前提下从不同感受野中产生多个空间维度的特征表示,可以表示为:

$$F_{s1} = f_{3 \times 3}^1(F), F_{s2} = f_{3 \times 3}^2(F), F_{s3} = f_{5 \times 5}^2(F). \quad (6)$$

随后通过通道非局部注意力 (Channel Non-local Attention, CNLA) 聚合跨像素通道维度的全局上下文信息,捕获不同通道特征之间的变化关系,使模型选择性关注重要的通道级上下文特征,增强模型对全局特征的理解能力,具体结构

如图 3(b)所示。其中,首先通过 1×1 卷积操作对输入特征 F_{si} 进行升维后,利用深度卷积和通道分离及 Reshape 操作生成 Q, K, V ,分别对应于查询矩阵 $Q_{ci} \in \mathbb{R}^{C \times (H \times W)}$,键值矩阵 $K_{ci} \in \mathbb{R}^{(H \times W) \times C}$,数值矩阵 $V_{ci} \in \mathbb{R}^{C \times (H \times W)}$ 。将查询矩阵 Q_{ci} 与键值矩阵 K_{ci} 相乘后归一化得到通道非局部注意力图 $W_{ci} \in \mathbb{R}^{C \times C}$,其过程可以表示为:

$$[Q_{ci}, K_{ci}, V_{ci}] = C\&R(DWConv(Conv_{1 \times 1}(F_{si}))), \quad (7)$$

$$W_{ci} = \text{Softmax}(ReLU(Q_{ci} \otimes K_{ci})), \quad (8)$$

其中: $Conv_{1 \times 1}(\cdot)$ 表示 1×1 卷积的升维操作, $C\&R(\cdot)$ 表示通道分离及 Reshape 操作,

$DWConv(\cdot)$ 表示深度卷积。此外,在 Softmax 操作之前添加了一个 ReLU 激活函数,选择性地控制特征向前传播,以更好地聚合并突出更具代表性的通道上下文信息。随后,将通道注意力图 W_{ci} 与表示空间维度特征的数值矩阵 V_{ci} 相乘,在跨像素通道上下文的指导下聚焦于全局特征中有助于重建任务的更精细的图像内容,可以表示为:

$$U_i = R(W_{ci} \otimes V_{ci}), \quad (9)$$

其中: $U_i (i=1, 2, 3)$ 表示不同大小区域的全局特征表示, $R(\cdot)$ 表示 Reshape 操作,将 U_i 的维度恢复为 $U_i \in \mathbb{R}^{C \times H \times W}$ 。

由于不同区域对于重建的贡献是不同的,为了实现对不同感受野区域特征的自适应选择,受通道选择性注意机制^[30]启发,引入了一种自适应选择机制 (Adaptive Selection Mechanism, ASM),其核心思想是通过动态权重分配实现多尺度感受野特征的自适应融合,增强模型对跨尺度区域关联信息的建模能力。首先将各区域全局特征 U_i 进行相加得到混合特征 U ,通过全局平均池化 (Global Average Pooling, GAP) 进行空间压缩,经过 MLP 进行特征微调后生成对应三个线性特征 $Z = \{Z_1, Z_2, Z_3\}$, $Z \in \mathbb{R}^{3 \times C \times 1 \times 1}$ 。随后通过 Softmax 函数根据线性特征 Z 的通道维度生成对应位置的动态权重以实现对不同感受野区域全局特征贡献的动态调整。具体而言,对于一个线性特征 $Z_i (i=1, 2, 3)$ 的第 k 个通道位置,其权重 $W_{Z_i}^k$ 可以由式(10)得到:

$$W_{Z_i}^k = \frac{e^{Z_i^k}}{e^{Z_1^k} + e^{Z_2^k} + e^{Z_3^k}}, \quad (10)$$

其中, $W_{Z_i}^k$ 满足 $W_{Z_1}^k + W_{Z_2}^k + W_{Z_3}^k = 1$ 。GMFS 最

后输出的特征 $\bar{Y}$ 可以表示为:

$$\bar{Y} = U_1 \times W_{Z_1} + U_2 \times W_{Z_2} + U_3 \times W_{Z_3}. \quad (11)$$

2.3 局部增强特征提取模块

全局混合尺度特征提取模块 GHFEM 从不同的多尺度角度探索全局信息,此外,图像的局部上下文特征也有助于超分辨率网络性能的提。因此本文针对图像中不同区域的重建难度是相对不同的,即平滑的区域易于超分,而纹理密集的区域容易丢失细节这一特点,设计了一个局部增强特征提取模块 (Local Enhancement Feature Extraction Module, LEFEM) 来对图像中的局部信息进行建模,其主要由两个部分组成:(1) 空间自适应选择块 (Spatial Adaptively Selection Block, SASB) 引入空间非局部注意力,通过空间结构关系来自适应缩放局部特征,使网络重点关注图像中难以重建的区域;(2) 频率分离选择块 (Frequency Separation Selection Block, FSSB) 在频域分离图像特征的实部和虚部,进行特征选择,激活更有代表性的高频特征信息,增强模型对细节纹理部分的学习能力。

2.3.1 空间自适应选择块

空间自适应选择块 SASB 由一个主分支和一个空间自适应选择分支组成,如图 4(a) 所示。其中,主分支用于对局部信息进行建模,与常用的级联卷积层来提取特征不同的是,主分支首先通过一个 3×3 的卷积层提取深层空间局部特征,并使通道数加倍,结合 GELU 非线性激活函数以提高特征的表达能力,随后通过一个 1×1 的卷积层将通道数减少到原始输入的维度。而空间自适应分支通过空间非局部注意力 (Spatial Non-local Attention, SNLA) 用于对图像空间

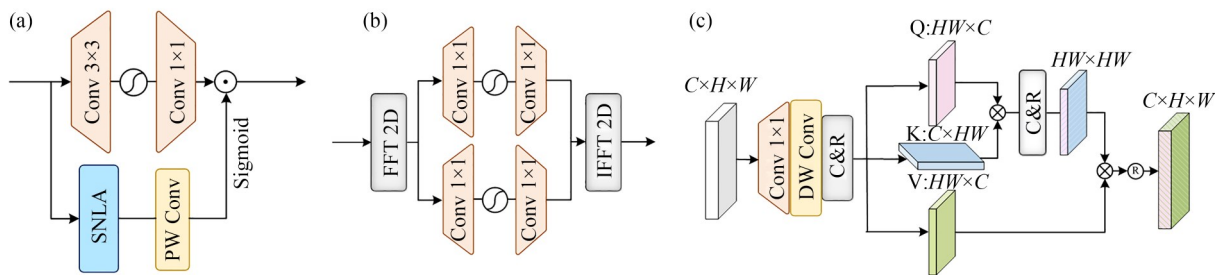


图4 (a) 空间自适应选择块; (b) 频率分离选择块; (c) 空间非局部注意力。

Fig. 4 (a) Spatial adaptively selection block; (b) Frequency separation selection block; (c) Spatial non-local attention

结构信息进行学习,从而使网络更加关注图像中难以重建的高频区域,细化重建细节。与通道非局部注意力CNLA不同的是,空间非局部注意力权重为 $W_{si} \in \mathbb{R}^{HW \times HW}$,与数值矩阵 $V_{si} \in \mathbb{R}^{(H \times W) \times C}$ 相乘后得到空间结构信息,进而得到空间自适应选择权重。SASB工作过程表示如下:

$$W_{si} = \text{Softmax}(\text{ReLU}(Q_{si} \otimes K_{si})), \quad (12)$$

$$\begin{cases} X_{\text{main}} = f_{1 \times 1}(\text{GELU}(f_{3 \times 3}(X_{\text{in}}))) \\ X_{\text{select}} = \text{Sigmoid}(PWConv(R(W_{si} \otimes V_{si}))) \end{cases}, \quad (13)$$

$$X_{\text{Sout}} = X_{\text{main}} \odot X_{\text{select}}, \quad (14)$$

其中: X_{in} 表示输入特征, X_{main} 和 X_{select} 分别是主分支的输出特征以及空间自适应选择权重。通过该权重引导模型对重建难度较大的区域赋予更高的关注度,得到调制后的局部空间特征 X_{Sout} 。

2.3.2 频率分离选择块

在大多数图像超分辨率工作中,仅考虑空间中的特征学习和表达^[31-32],并不能有效帮助高频细节的恢复。而在图像的频域中,能够轻易地分离高频和低频信息。基于此,本文从频域考虑对高频信息进行提取和选择,设计了一个频率分离选择块FSSB,其结构如图4(b)所示。具体而言,首先利用快速傅里叶变换得到特征图频域中的实部 X_{re} 和虚部 X_{im} ,随后通过两个 1×1 卷积操作分别对实部和虚部的特征进行升维和降维以实现通道上下文信息混合,并在过程中使用GELU函数激活有代表性的高频特征,提高模型对高频特征的学习能力,对空间局部特征中高频信息进行补充,可以表示为:

$$[X_{\text{re}}, X_{\text{im}}] = F(X_{\text{in}}), \quad (15)$$

$$X_{\text{Pout}} = F^{-1} \left(\begin{bmatrix} f_{1 \times 1}^C(\text{GELU}(f_{1 \times 1}^{2C}(X_{\text{re}}))) \\ f_{1 \times 1}^C(\text{GELU}(f_{1 \times 1}^{2C}(X_{\text{im}}))) \end{bmatrix} \right), \quad (16)$$

其中: $F(\cdot)$ 和 $F^{-1}(\cdot)$ 分别表示快速傅里叶变换及其逆变换, $f_{1 \times 1}^C(\cdot)$ 和 $f_{1 \times 1}^{2C}(\cdot)$ 分别表示输出通道为 C 与 $2C$ 的 1×1 卷积操作。

最后,通过两次级联局部残差连接,提高网络训练的稳定性及高频特征学习能力,并通过

通道混合门控机制CMG进一步融合空间局部特征 X_{Sout} 与频域局部特征 X_{Fout} 并去除冗余信息,局部增强特征提取模块LEFEM最终的输出 X_{out} 为:

$$\begin{cases} X_{\text{S&F}} = X_{\text{Sout}} + X_{\text{Fout}} + X_{\text{in}} \\ X_{\text{out}} = \text{CMG}(X_{\text{S&F}}) + X_{\text{S&F}} \end{cases}. \quad (17)$$

2.4 损失函数

在上述相关工作中,大多数采用的 L_1 损失函数对网络模型进行优化。相比于其他损失函数, L_1 损失函数在图像超分辨率任务中,能够提供更有意义的像素级的约束以及更加稳定的梯度,在提高重建结果保真度的同时加快网络模型的收敛速度。对于高分辨率图像 I_{HR} 和 SAF²Net 的重建结果 I_{SR} , L_1 损失函数可以表示为:

$$L_1 = \|I_{\text{SR}} - I_{\text{HR}}\|_1. \quad (18)$$

然而,仅依赖于空间域的像素级损失进行学习和约束,可能会导致不同像素之间的细微差异混淆,缺乏感知高频图像内容信息的能力。此外,由于本文在局部增强特征提取模块部分设计了一个频率分离选择块,空间域的 L_1 损失函数无法有效帮助该模块进行高频特征的估计。因此,本文通过快速傅里叶变换(Fast Fourier Transform, FFT)将 I_{HR} 和 I_{SR} 转换到频率域,从而使频率域的约束参与网络模型的训练过程,该损失函数定义为:

$$L_{\text{Freq}} = \|F(I_{\text{SR}}) - F(I_{\text{HR}})\|_1, \quad (19)$$

其中: $\|\cdot\|_1$ 表示 L_1 范数, $F(\cdot)$ 表示快速傅里叶变换。最终,联合损失函数表示为:

$$L_{\text{all}} = L_1 + \lambda L_{\text{Freq}}, \quad (20)$$

其中, λ 表示频率域损失 L_{Freq} 的权重系数,在本文中设置为 0.01。

3 实验及结果分析

在整个实验中,硬件环境是基于 PANYAN 7048GR 服务器, CPU 为正式版 E5-2096 2.3G-3.6 G 45M Cache 22NM, 内存为 256 G, 两张 NVIDIA GeForce TITAN RTX 显卡。软件环境是基于 Ubuntu 系统的 Pytorch 1.10 深度学习框架, 编程语言为 Python 3.8。

3.1 实验数据集

(1) UCMerced 遥感数据集^[33]

UCMerced 数据集包含机场、住宅区、停车场等 21 类遥感场景,每个类别包含 100 张尺寸为 256×256 的图像,共 2 100 张。为了更好地与其他遥感超分辨率的方法进行对比,本文与其他方法一样,将数据集平均划分为训练集和测试集,各 1 050 张图像,其中训练集中的 20% 作为验证集。

(2) AID 遥感数据集^[34]

AID 数据集包含 30 类遥感场景,共 10 000 张尺寸为 600×600 的图像,涉及 30 类遥感场景,其空间分辨率为 0.5~8 m/pixel。对于该数据集,本文同样遵循之前的其他相关工作的设置,随机选取整个数据集的 80% 作为训练集,并将训练集中每个遥感场景的 5 张图像合并作为验证集,剩下的 20% 作为测试集。

3.2 实验设置

在本文实验中,各数据集中的原始图像被视为 Ground Truth 作为 HR 参考,通过双三次插值(Bicubic Interpolation)得到相应的 LR 图像,从而构建 HR/LR 图像对进行训练和评估。本文主要针对 2 倍、3 倍和 4 倍遥感图像超分辨率任务,从 LR 图像中随机裁剪尺寸大小为 48×48 的图像

块以及与其对应放大倍率的 HR 图像块。此外,训练图像通过随机旋转 90° , 180° , 270° 和水平翻转来增强数据的多样性。

本文网络模型的深层特征提取模块 SAFSG 的级联个数 m 设置为 10,其输入和输出通道数均为 64。模型训练时的批次大小统一设置为 4 (batch size=4),采用 Adam 作为模型的优化器,其中 $\beta_1=0.9$, $\beta_2=0.999$, $\epsilon=10^{-8}$ 。对于 UCMerced 数据集,共训练 500 个 epoch,初始学习率设置为 4×10^{-4} ,通过余弦退火策略调整学习率,最后衰减至 1×10^{-5} ;对于 AID 数据集,共训练 800 个 epoch,初始学习率设置为 4×10^{-4} ,通过余弦退火策略调整学习率,最后衰减至 1×10^{-6} 。实验采用 2.4 节描述的空间域与频率域联合损失函数。

3.3 SAF²Net 算法消融实验

为了验证和评估 SAF²Net 中全局混合尺度特征提取模块 GHFEM 和局部增强特征提取模块 LEFEM 中各个组件的有效性,本文在 UCMerced 数据集 4 倍超分辨率任务中设计了多个消融实验。消融实验的设置与 3.2 节一致,实验结果见表 1。

表 1 SAF²Net 在 UCMerced 数据集 4 倍超分辨率的消融实验

Tab. 1 Ablation experimental results of SAF²Net on UCMerced dataset with $\times 4$ scale factor

Ablation	Variant	PSNR	SSIM
SAF ² Net	—	28.13	0.7748
GHFEM	(1) HSFM: w/o scale8	27.98	0.7716
	(2) HSFM: w/o scale8+scale4	27.97	0.7709
	(3) HSFM: w/o scale8 + scale4+scale2	27.98	0.7707
	(4) HSFM: w/o CMG	27.89	0.7678
	(5) GMFS: w/o ASM	28.01	0.7702
LEFEM	(6) w/o SASB	27.91	0.7690
	(7) w/o FSSB	27.99	0.7710
	(8) w/o SASB + FSSB	27.86	0.7669
	(9) w/o CMG	28.04	0.7730
Loss	(10) w/o L_{Freq}	27.96	0.7709

在本文提出的 SAF²Net 中 GHFEM 主要包含混合尺度空间自适应特征调制模块 HSFM 和全局多尺度感受野选择模块 GMFS。针对 HSFM,本文设置了 4 组实验,其中实验(1)~实验

(3)分别探讨不同尺度信息对于深层特征调制的有效性;实验(4)验证混合通道门控机制 CMG 聚合混合尺度特征的有效性。从表 1 中实验(1)~实验(3)的结果可以看出,缺少某一尺度或多个

尺度的特征信息均会导致 PSNR/SSIM 指标下降,而 SAF²Net 与之相比,当聚合了 scale2, scale4 和 scale8 三种尺度的特征后,模型能够捕获到遥感图像中不同尺度之间的相似性并有效利用它们之间的互补关系,获得更优的 PSNR/SSIM。实验(4)表明,若单纯地将混合尺度的特征通过简单相加进行聚合,会造成特征信息的冗余导致重建性能下降,PSNR/SSIM 相比于 SAF²Net 下降了 0.24 dB/0.0070,说明 CMG 能够在聚合全局混合尺度特征时有效减少特征之间的冗余信息,控制更利于超分任务的精细特征前向流动。此外,本文获取了不同尺度以及聚合混合尺度信息后对应的特征图,如图 5 所示。可以看出,更大尺度包含特征的细节信息,而更小尺度侧重特征的结构、位置等信息,即学习到的特征在不同尺度之间是独立和互补的,聚合这些不同尺度的信息有助于分别从局部和整体挖掘更丰富的特征信息。

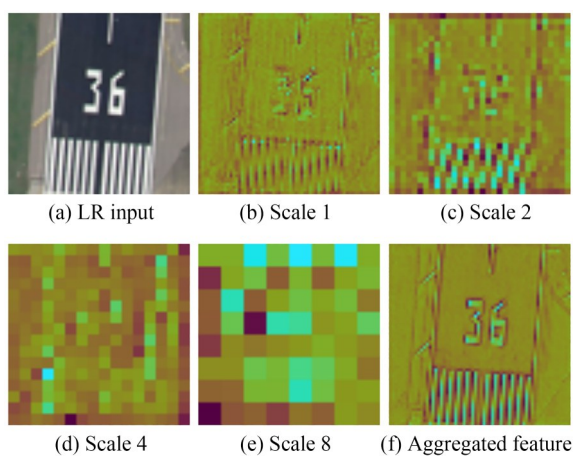


图 5 HSFM 在不同尺度下学习到的特征图

Fig. 5 Visualization of learned feature maps from the HS-FM at different scales

针对 GMFS,本文通过实验(5)来验证自适应选择机制(ASM)能否可以对不同感受野大小区域的特征进行选择从而增强图像的超分重建效果。从实验(5)的结果可以看出,PSNR/SSIM 相比于 SAF²Net 下降了 0.12 dB/0.0046,这是因为直接将不同感受野大小区域的特征相加,同样会产生冗余特征,模型无法有效利用全局结构信息、主体信息和细节纹理信息之间的互补关系;而包含 ASM 完整的 GMFS 能够获得更优的重建

性能,从而证明 ASM 能够有效滤除冗余特征,实现对不同感受野区域全局特征贡献的动态调整。

对于 SAF²Net 中 LEFEM 空间自适应特征选择块 SASB 和频率分离选择块 FSSB,本文设置了 4 组实验来验证其在遥感超分任务中的有效性。实验(6)和实验(7)分别在 SAF²Net 的基础上减去了 SASB 和 FSSB,从而验证这两个组件各自对模型关于局部特征信息的建模能力的影响。实验(6)的结果说明,SASB 能够使模型根据空间自适应选择权重为图像中重建难度较大的边缘纹理部分赋予更多的关注,从而使 PSNR/SSIM 提升 0.22 dB/0.005 8。在实验(7)中,对应的特征图可视化如图 6 所示,可以看出,若没有 FSSB 的参与,将导致特征图在分界处会产生不清晰的边缘,而在 FSSB 的作用下,跑道线等高频部分被增强,即能够激活更多的高频信息参与重建,从而提高模型的超分性能。

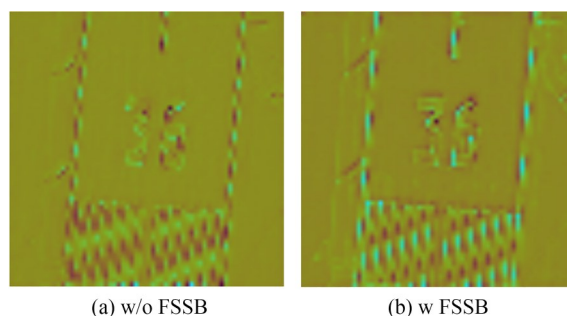


图 6 LEFEM 中 21th 特征图可视化

Fig. 6 Visualization of 21th feature map in LEFEM

实验(8)在 SAF²Net 的基础上减去了 SASB 和 FSSB 两个组件,仅用 SASB 中的主分支进行局部特征建模,PSNR/SSIM 分别降低了 0.27 dB/0.007 9,说明从空间域和频率域两个角度对局部特征进行建模能够建立空间和频率之间的互补关系,进一步细化高频部分的重建细节。同时,实验(9)减去了 CMG 来验证该模块能否有效融合空间域和频率域两个维度的特征。从结果可以看出,相比于 SAF²Net,直接通过相加进行融合导致 PSNR/SSIM 降低了 0.09 dB/0.001 8,说明 CMG 可以将空间域和频率域两个维度的特征转换到同一个维度并去除冗余信息,有助于空频两个维度特征的有效融合。

此外,为了验证本文方法使用的联合损失函

数中频率级约束的有效性,在实验(10)中仅使用空间域的像素级损失函数进行训练。从结果可以看出,使用联合损失函数比仅使用像素级损失函数的 PSNR/SSIM 指标高 0.17 dB/0.003 9,说明该联合损失函数能够从空间域和频率域两个角度提高模型对遥感图像中高频内容信息的感知能力,同时频率域损失函数也能够与模型中关于频率域特征提取相关的组件相互协作,从而进一步加强模型的建模能力。

3.4 客观实验结果分析

为了验证 SAF²Net 网络在遥感图像超分辨率任务中的性能,本文将其与其他优秀的方法在 2 倍、3 倍和 4 倍 3 个不同放大倍率上根据峰值信噪比(Peak Signal to Noise Ratio, PSNR)和结构相似度(Structural Similarity, SSIM)两个指标进行对比。其中 PSNR 基于信噪比的概念,根据 SR 图像与 HR 图像之间的均方误差(Mean Squared Error, MSE)衡量像素级差异,其值越高表明 SR 图像与 HR 图像之间的偏差越小,其具体计算公式如式(21)所示:

$$\begin{cases} MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (I_{SR\ i,j} - I_{HR\ i,j})^2 \\ PSNR = 10 \lg \left(\frac{MAX^2}{MSE} \right) \end{cases}, \quad (21)$$

其中: MSE 为 SR 图像和 HR 图像之间的均方误差, M 和 N 分别为图像的高度和宽度, $I_{SR\ i,j}$ 和 $I_{HR\ i,j}$ 分别为 SR 图像和 HR 图像中对对应像素的灰度值,

MAX 为像素值的最大值,本文中 $MAX = 255$ 。

SSIM 在设计上考虑了人眼的主观感受,结合亮度、对比度和结构三个角度进行计算,评估 SR 图像与 HR 图像之间的相似性,其值越接近 1 表示超分结果与 HR 图像更为相似,即重建图像质量越好,具体计算公式如式(22)所示:

$$SSIM = \frac{(2\mu_{SR}\mu_{HR} + c_1)(2\sigma_{SR,HR} + c_2)}{(\mu_{SR}^2 + \mu_{HR}^2 + c_1)(\sigma_{SR}^2 + \sigma_{HR}^2 + c_2)}, \quad (22)$$

其中: $\mu_{SR}, \mu_{HR}, \sigma_{SR}, \sigma_{HR}$ 分别表示 SR 图像和 HR 图像的均值、方差, $\sigma_{SR,HR}$ 表示二者之间的协方差, c_1 和 c_2 为两个常量。此外,为了进一步评估 SAF²Net 的重建性能,在本文中针对遥感图像的结构和语义一致性,额外引入 SCC 和 SAM 两个指标,以更全面地对重建质量进行定量分析,其中 SCC 值越高, SAM 值越低表示重建图像质量越好。两个遥感数据集的定量评价结果分别展示在表 2 和表 3 中。

对比方法包括 Bicubic 插值^[9], SRCNN^[14], VDSR^[35], LGCNet^[16], DCM^[19], TransENet^[24], HSENet^[23], HAUNet^[25], TTST^[26] 和 TSFNet^[36]。其中 Bicubic 插值是传统方法,常作为性能评价的基线, SRCNN 和 VDSR 是自然图像超分辨率的经典方法,其余对比方法是最近几年遥感图像超分辨率的优秀方法。从结果可以看出,对于 UC-Merced 数据集, SAF²Net 在 3 倍超分任务中 PSNR/SSIM 分别提升 0.11~2.99 dB/0.003~−0.087 8;对于 AID 数据集, SAF²Net 在 4 倍超

表 2 不同算法在 UCMerced 测试集中的评价指标结果

Tab. 2 Evaluation metrics of different methods over the UCMerced test dataset

Scale	Metric	Bicubic	SRCNN	VDSR	LGCNet	DCM	TransENet	HSENet	HAUNet	TTST	TSFNet	本文
UCMerced	PSNR	30.76	32.84	33.87	33.48	33.65	34.03	34.22	34.46	34.39	<u>34.48</u>	34.53
	SSIM	0.878 9	0.915 2	0.927 4	0.923 5	0.927 4	0.930 1	0.932 7	0.933 3	0.932 4	<u>0.933 5</u>	0.934 2
	×2											
	SCC	0.507 9	0.595 6	0.619 6	0.605 6	0.631 8	0.636 6	0.634 1	0.643 7	0.645 5	<u>0.646 1</u>	0.647 8
	SAM	0.071 2	0.055 6	0.051 9	0.054 0	0.050 4	0.049 8	0.050 0	0.048 8	0.048 5	<u>0.048 4</u>	0.047 6
	PSNR	27.46	28.66	29.75	29.28	29.52	29.92	30.00	<u>30.34</u>	30.25	30.32	30.45
UCMerced	SSIM	0.763 1	0.803 8	0.834 6	0.823 8	0.839 4	0.840 8	0.842 0	<u>0.847 6</u>	0.845 8	0.847 2	0.850 9
	×3											
	SCC	0.276 6	0.358 1	0.394 1	0.365 4	0.405 5	0.407 6	0.413 1	0.423 6	0.420 1	<u>0.424 8</u>	0.437 2
	SAM	0.103 9	0.088 3	0.082 9	0.086 8	0.082 0	0.080 7	0.080 6	<u>0.077 9</u>	0.078 7	0.078 1	0.076 6
	PSNR	25.65	26.78	27.54	27.02	27.22	27.77	27.73	<u>28.06</u>	27.98	27.92	28.13
	SSIM	0.672 5	0.721 9	0.752 2	0.733 3	0.752 8	0.763 0	0.762 3	<u>0.772 6</u>	0.769 8	0.769 5	0.774 8
UCMerced	×4											
	SCC	0.150 0	0.222 5	0.258 9	0.224 6	0.261 2	0.270 7	0.269 2	<u>0.293 2</u>	0.289 2	0.292 0	0.301 5
	SAM	0.126 9	0.111 8	0.105 5	0.110 6	0.105 3	0.102 7	0.104 3	<u>0.099 7</u>	0.100 6	0.100 7	0.098 1

表 3 不同算法在 AID 测试集中的评价指标结果

Tab. 3 Evaluation metrics of different methods over the AID test dataset

Scale	Metric	Bicubic	SRCNN	VDSR	LGCNet	DCM	TransENet	HSENet	HAUNet	TTST	TSFNet	本文
AID	PSNR	32.39	34.49	35.11	34.80	35.21	35.28	35.50	35.58	<u>35.60</u>	35.54	35.67
	SSIM	0.890 6	0.928 6	0.934 0	0.932 0	0.936 6	0.937 4	0.938 3	0.938 9	<u>0.939 2</u>	0.938 5	0.940 3
	×2 SCC	0.520 2	0.600 6	0.618 1	0.611 6	0.639 0	0.652 9	0.662 6	0.661 6	<u>0.663 8</u>	0.659 8	0.666 3
	SAM	0.074 2	0.058 0	0.054 4	0.055 8	0.053 2	0.052 9	0.052 4	0.051 8	<u>0.051 3</u>	0.052 0	0.050 1
AID	PSNR	29.08	30.55	31.17	30.73	31.31	31.45	31.49	<u>31.64</u>	31.62	31.54	31.76
	SSIM	0.786 3	0.837 2	0.851 1	0.841 7	0.856 1	0.859 5	0.858 8	0.861 9	<u>0.862 6</u>	0.859 5	0.863 5
	×3 SCC	0.274 3	0.353 4	0.380 0	0.356 4	0.400 5	0.409 2	0.405 3	0.414 4	<u>0.416 0</u>	0.409 6	0.418 9
	SAM	0.105 5	0.089 4	0.083 6	0.087 6	0.082 0	0.080 5	0.080 6	0.079 2	<u>0.079 0</u>	0.080 3	0.078 0
AID	PSNR	27.30	28.40	28.99	28.61	29.17	29.38	29.32	29.49	29.40	<u>29.50</u>	29.59
	SSIM	0.703 6	0.756 1	0.775 3	0.762 6	0.782 4	0.790 9	0.786 7	0.792 4	0.789 0	<u>0.792 6</u>	0.794 5
	×4 SCC	0.141 6	0.211 1	0.242 7	0.218 2	0.267 5	0.284 3	0.276 5	0.293 5	0.292 2	<u>0.296 0</u>	0.298 3
	SAM	0.127 5	0.112 4	0.105 5	0.109 9	0.103 2	0.100 5	0.101 7	0.099 7	0.100 3	<u>0.099 6</u>	0.098 7

注:加粗表示最优指标,下划线表示次优指标

分任务中 PSNR/SSIM 分别提升 0.09~2.29 dB/0.001 9~0.090 9。由于遥感图像与自然图像的数据分布存在较大差异,本文所提算法相较于传统方法和自然图像的超分辨率方法在定量指标上取得了大幅提升;对比其他遥感图像超分辨率方法,例如 HAUNet 使用固定尺寸的卷积核提取特征,导致其难以自适应融合不同尺度区域的关联信息;TSFNet 虽然通过提取由粗到细的双分支空间频率信息,但并未进行特征筛选以及融合空间频率互补特征,而本文从两个不同的角度挖掘全局混合尺度信息以及不同大小区域的

关联特征,并结合空间域和频率域进行高频局部特征的建模,在所有放大倍率上都获得了最优的评价指标结果,从定量的角度证明了 SAF²Net 在遥感图像超分辨率任务中可以获得具有竞争力的性能。

为了进一步对比 SAF²Net 与各个方法在不同遥感场景下的超分性能,本文分别对比了 UC-Merced 数据集 3 倍超分辨率任务和 AID 数据集 4 倍超分辨率任务中每个遥感场景类别超分结果的评价指标,具体如表 4 和表 5 所示。

表 4 不同算法在 UCMerced 测试集上 21 类遥感场景 3 倍超分辨率的 PSNR/SSIM

Tab. 4 PSNR/SSIM of different methods on 21 types of RS scenes on UCMerced test dataset for×3 SR

Class	Baseline		自然图像超分方法			
	Bicubic		SRCNN		VDSR	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Agricultural	26.86	0.542 4	27.47	0.591 4	27.62	0.611 1
Airplane	26.71	0.783 7	28.24	0.833 2	29.91	0.845 9
Baseball diamond	33.33	0.835 1	34.33	0.867 1	35.09	0.881 6
Beach	36.14	0.896 7	37.00	0.915 5	37.43	0.918 8
Buildings	25.09	0.774 3	26.84	0.827 6	28.57	0.860 6
Chaparral	25.21	0.700 2	26.11	0.745 3	26.78	0.775 2
D-Residential	25.76	0.773 3	27.41	0.818 4	28.84	0.860 2
Forest	27.53	0.703 5	28.24	0.740 4	28.56	0.766 5
FreeWay	27.36	0.766 7	28.69	0.803 3	30.33	0.838 7
Golf Course	35.21	0.881 5	36.15	0.890 3	36.40	0.901 7

续表 4 不同算法在 UCMerced 测试集上 21 类遥感场景 3 倍超分辨率的 PSNR/SSIM

Tab. 4 PSNR/SSIM of different methods on 21 types of RS scenes on UCMerced test dataset for $\times 3$ SR

Class	Baseline		自然图像超分方法					
	Bicubic		SRCNN				VDSR	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Harbor	21.25	0.754 8	22.82	0.822 8	24.39	0.889 2		
Intersection	26.48	0.760 4	27.67	0.818 4	28.82	0.850 7		
M-Residential	25.68	0.746 6	27.06	0.807 4	28.43	0.845 3		
Mobile Homepark	22.25	0.693 9	23.89	0.745 9	25.25	0.804 4		
Overpass	24.59	0.706 4	25.65	0.735 6	27.78	0.802 0		
Parking Lot	21.75	0.742 5	23.11	0.788 9	24.40	0.829 6		
River	28.12	0.720 4	28.89	0.745 1	29.18	0.791 9		
Runway	29.30	0.764 2	30.61	0.808 9	31.48	0.827 1		
S-Residential	28.34	0.802 8	29.40	0.834 3	31.44	0.861 8		
Storage Tanks	29.97	0.835 2	31.33	0.870 2	32.77	0.894 6		
Tennis Court	29.75	0.839 8	30.98	0.867 2	32.36	0.888 2		
Average	27.46	0.763 1	28.66	0.803 8	29.75	0.834 6		

Class	遥感图像超分方法											
	TransENet		HSENet		HAUNet		TTST		TSFNet		本文	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Agricultural	28.02	0.622 8	27.64	0.613 9	28.07	<u>0.626 0</u>	27.67	0.613 3	27.90	0.615 4	<u>28.04</u>	0.634 6
Airplane	29.94	0.853 4	30.09	0.852 8	30.41	0.856 2	<u>30.45</u>	0.857 8	30.50	<u>0.858 9</u>	30.33	0.859 2
Baseball diamond	35.04	0.882 8	35.05	0.882 8	35.34	0.884 1	35.35	0.885 2	<u>35.37</u>	<u>0.885 5</u>	35.39	0.885 7
Beach	37.53	0.920 4	37.69	0.921 4	37.90	<u>0.923 7</u>	37.69	0.922 1	37.73	0.922 2	<u>37.88</u>	0.924 1
Buildings	28.81	0.869 1	28.95	0.869 5	<u>29.47</u>	<u>0.876 8</u>	29.23	0.871 9	29.45	0.874 9	29.55	0.878 5
Chaparral	26.69	0.778 2	26.70	0.780 0	26.86	0.785 4	26.90	0.784 6	<u>26.93</u>	<u>0.785 9</u>	26.98	0.786 6
D-Residential	29.11	0.869 2	29.24	0.869 2	<u>29.63</u>	<u>0.875 1</u>	29.33	0.870 5	29.55	0.874 1	29.77	0.878 7
Forest	28.59	0.770 2	28.59	0.771 0	28.69	<u>0.775 7</u>	28.65	0.775 2	<u>28.70</u>	0.775 1	28.73	0.776 6
FreeWay	30.38	0.843 6	30.63	0.846 3	30.96	<u>0.855 1</u>	30.90	0.852 0	<u>30.98</u>	0.854 9	31.16	0.859 3
Golf Course	36.68	0.903 8	36.62	0.903 8	36.91	<u>0.905 0</u>	36.75	0.904 5	36.78	0.904 7	<u>36.87</u>	0.905 9
Harbor	24.72	0.900 8	24.88	0.904 8	<u>25.51</u>	<u>0.916 3</u>	25.34	0.910 7	25.46	0.913 5	25.71	0.919 3
Intersection	29.03	0.857 3	29.21	0.859 0	<u>29.52</u>	0.865 0	29.42	0.865 2	29.50	<u>0.865 9</u>	29.71	0.871 0
M-Residential	28.47	0.846 8	28.55	0.849 7	28.91	0.855 9	28.95	0.857 2	<u>29.12</u>	<u>0.857 4</u>	29.01	0.858 6
Mobile Homepark	25.64	0.816 1	25.70	0.818 2	26.14	0.826 0	26.06	0.823 5	<u>26.25</u>	<u>0.827 6</u>	26.31	0.831 7
Overpass	27.83	0.808 2	28.22	0.811 4	<u>28.40</u>	0.816 1	28.21	0.813 8	28.36	<u>0.818 1</u>	28.81	0.824 9
Parking Lot	24.45	0.836 8	24.66	0.840 4	25.25	0.854 3	25.40	0.854 8	<u>25.57</u>	<u>0.859 3</u>	25.53	0.860 5
River	29.25	0.793 1	29.22	0.793 5	29.40	0.796 2	29.34	0.796 6	29.29	0.796 0	<u>29.39</u>	0.796 8
Runway	31.25	0.830 4	31.15	0.831 0	31.56	0.835 2	31.85	<u>0.839 2</u>	31.26	0.832 4	<u>31.76</u>	0.840 8
S-Residential	31.57	0.865 8	31.64	0.868 7	<u>31.87</u>	<u>0.868 9</u>	31.74	0.868 3	31.85	0.868 2	31.91	0.870 5
Storage Tanks	32.71	0.896 5	32.95	0.897 4	<u>33.16</u>	0.899 9	33.03	0.898 8	33.14	<u>0.901 0</u>	33.36	0.902 2
Tennis Court	32.51	0.894 4	32.71	0.896 0	<u>33.12</u>	<u>0.902 3</u>	32.92	0.899 4	33.07	0.899 8	33.24	0.905 3
Average	29.92	0.840 8	30.00	0.842 0	<u>30.34</u>	<u>0.847 6</u>	30.25	0.845 8	30.32	0.847 2	30.45	0.850 9

表 5 不同算法在 AID 测试集上 30 类遥感场景 4 倍超分辨率的 PSNR/SSIM

Tab. 5 PSNR/SSIM of different methods on 30 types of RS scenes on AID test dataset for $\times 4$ SR

Class	Baseline		自然图像超分方法			
	Bicubic		SRCNN		VDSR	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Airport	27.03	0.734 4	28.17	0.781 3	28.82	0.801 2
Bare Land	34.88	0.827 8	35.63	0.853 5	35.98	0.858 3
Baseball Field	29.06	0.787 8	30.51	0.827 5	31.18	0.840 3
Beach	31.07	0.785 0	31.92	0.813 6	32.29	0.819 5
Bridge	28.98	0.791 5	30.41	0.828 2	31.19	0.842 1
Center	25.26	0.668 1	26.59	0.729 8	27.48	0.760 1
Church	22.15	0.558 6	23.41	0.644 8	24.12	0.679 4
Commercial	25.83	0.693 5	27.05	0.755 8	27.62	0.777 2
D—Residential	23.05	0.591 2	24.13	0.669 1	24.70	0.699 3
Desert	38.49	0.900 9	38.84	0.913 0	39.13	0.914 8
Farmland	32.30	0.808 1	33.48	0.839 3	34.20	0.852 9
Forest	27.39	0.619 9	28.15	0.681 7	28.36	0.690 2
Industrial	24.75	0.655 2	26.00	0.721 8	26.72	0.751 5
Meadow	32.06	0.707 0	32.57	0.735 3	32.77	0.739 5
M—Residential	26.09	0.643 6	27.37	0.711 3	28.06	0.735 8
Mountain	28.04	0.689 0	28.90	0.739 8	29.11	0.747 7
Park	26.23	0.681 2	27.25	0.738 3	27.69	0.754 9
Parking	22.33	0.671 5	24.01	0.754 9	25.21	0.795 6
Playground	27.27	0.734 8	28.72	0.784 6	29.62	0.809 8
Pond	28.94	0.771 5	29.85	0.804 9	30.26	0.814 6
Port	24.69	0.742 2	25.82	0.790 7	26.43	0.809 5
Railway Station	26.31	0.668 6	27.55	0.732 4	28.19	0.757 8
Resort	25.98	0.694 2	27.12	0.748 4	27.71	0.768 1
River	29.61	0.725 1	30.48	0.765 2	30.82	0.775 0
School	24.91	0.667 3	26.13	0.731 1	26.78	0.755 5
S—Residential	25.41	0.553 6	26.16	0.616 5	26.46	0.630 0
Square	26.75	0.705 8	28.13	0.764 0	28.91	0.787 5
Stadium	24.81	0.689 3	26.10	0.746 7	26.88	0.774 0
Storage Tanks	24.18	0.636 1	25.27	0.695 9	25.86	0.721 7
Viaduct	25.86	0.649 0	27.03	0.712 3	27.74	0.741 8
Average	27.30	0.703 6	28.40	0.756 1	28.99	0.775 3

续表 5 不同算法在 AID 测试集上 30 类遥感场景 4 倍超分辨率的 PSNR/SSIM

Tab. 5 PSNR/SSIM of different methods on 30 types of RS scenes on AID test dataset for $\times 4$ SR

Class	遥感图像超分方法											
	TransENet		HSENet		HAUNet		TTST		TSFNet		本文	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Airport	29.23	0.818 2	29.22	0.814 7	29.32	0.816 3	29.29	0.815 4	<u>29.38</u>	<u>0.819 5</u>	29.47	0.821 1
Bare Land	36.20	0.862 6	36.21	0.862 3	36.40	<u>0.865 7</u>	36.20	0.864 4	36.18	0.864 2	<u>36.22</u>	0.868 3
Baseball Field	31.59	0.853 8	31.61	0.851 3	31.67	0.855 4	31.70	0.856 4	<u>31.76</u>	<u>0.856 9</u>	31.80	0.857 2
Beach	32.55	0.827 0	32.55	0.825 1	32.69	0.827 3	32.54	0.827 0	32.58	<u>0.829 3</u>	<u>32.62</u>	0.830 1
Bridge	31.63	0.854 4	31.61	0.853 2	31.77	<u>0.855 6</u>	31.68	0.851 6	<u>31.79</u>	0.854 4	31.82	0.856 7
Center	28.03	0.784 8	27.68	0.779 6	<u>28.14</u>	<u>0.786 5</u>	27.87	0.785 9	27.96	0.784 1	28.19	0.788 5
Church	24.51	0.703 8	24.38	0.698 9	24.58	0.705 4	24.47	0.700 5	<u>24.59</u>	<u>0.707 8</u>	24.60	0.708 7
Commercial	27.97	0.794 2	27.96	0.789 4	28.05	0.796 6	27.99	0.788 9	<u>28.09</u>	<u>0.797 8</u>	28.13	0.798 4
D-Residential	25.13	0.729 6	25.31	0.725 7	25.24	0.735 2	25.39	0.720 7	<u>25.48</u>	<u>0.737 6</u>	25.56	0.738 8
Desert	39.31	0.918 2	39.25	0.917 6	39.58	<u>0.918 8</u>	39.38	0.917 0	39.47	0.918 5	<u>39.50</u>	0.921 9
Farmland	34.58	0.864 5	34.46	0.862 2	34.72	<u>0.866 5</u>	34.45	0.859 9	34.63	0.864 0	<u>34.68</u>	0.869 6
Forest	28.56	0.705 9	28.68	0.703 0	28.61	0.711 9	28.67	0.702 9	<u>28.70</u>	0.704 8	28.72	<u>0.708 1</u>
Industrial	27.21	0.779 3	27.22	0.775 6	27.30	0.782 6	27.23	0.769 9	<u>27.41</u>	<u>0.784 5</u>	27.48	0.786 8
Meadow	32.66	0.748 4	32.67	0.745 9	33.02	0.747 4	32.79	0.746 2	32.86	<u>0.748 6</u>	<u>32.92</u>	0.749 6
M-Residential	28.45	0.754 4	28.56	0.750 0	28.55	0.756 4	28.60	0.754 7	<u>28.65</u>	<u>0.756 9</u>	28.75	0.758 2
Mountain	29.28	0.758 2	29.09	0.755 8	29.32	<u>0.759 5</u>	29.11	0.758 0	29.25	0.758 7	<u>29.28</u>	0.762 3
Park	28.01	0.770 2	28.06	0.768 4	28.08	0.771 5	28.06	0.769 6	<u>28.17</u>	<u>0.772 8</u>	28.23	0.773 5
Parking	26.40	0.835 5	26.30	0.823 4	26.63	0.836 8	26.45	0.829 6	<u>26.84</u>	<u>0.837 2</u>	26.92	0.838 8
Playground	30.30	0.832 0	30.12	0.826 2	30.16	0.834 1	30.15	0.831 1	<u>30.40</u>	<u>0.834 4</u>	30.48	0.835 9
Pond	30.53	0.823 1	30.43	0.821 3	30.60	<u>0.824 3</u>	30.48	0.821 8	30.54	0.824 0	<u>30.57</u>	0.824 9
Port	26.91	0.825 6	26.88	0.821 9	27.01	0.827 7	26.99	0.825 9	<u>27.08</u>	<u>0.828 0</u>	27.18	0.830 0
Railway Station	28.61	0.779 1	28.71	0.772 4	28.70	0.782 8	28.78	0.783 1	<u>28.85</u>	<u>0.783 9</u>	28.92	0.785 3
Resort	28.08	0.783 5	27.92	0.779 6	28.18	<u>0.785 1</u>	27.94	0.783 6	28.12	0.785 0	<u>28.16</u>	0.787 6
River	31.00	0.783 0	30.89	0.781 1	31.06	0.783 6	30.90	0.781 3	<u>30.07</u>	<u>0.783 8</u>	31.13	0.784 3
School	27.22	0.778 3	27.17	0.771 2	27.31	0.779 6	27.28	0.778 4	<u>27.34</u>	<u>0.779 8</u>	27.42	0.780 8
S-Residential	26.63	0.645 2	26.74	0.641 3	26.66	<u>0.646 9</u>	26.74	0.645 9	<u>26.80</u>	0.645 6	26.83	0.648 0
Square	29.39	0.804 5	29.16	0.800 7	<u>29.52</u>	<u>0.806 6</u>	29.18	0.803 5	29.45	0.806 0	29.58	0.808 8
Stadium	27.41	0.796 5	27.30	0.789 6	<u>27.53</u>	0.798 9	27.33	0.796 8	27.50	<u>0.799 8</u>	27.62	0.801 0
Storage Tanks	26.20	0.739 4	26.28	0.736 0	26.27	0.741 2	26.22	0.739 8	<u>26.41</u>	<u>0.742 6</u>	26.46	0.744 3
Viaduct	28.21	0.762 9	28.05	0.757 8	<u>28.29</u>	0.764 4	28.07	0.761 2	28.25	<u>0.765 7</u>	28.32	0.768 5
Average	29.38	0.790 9	29.32	0.786 7	29.49	0.792 4	29.40	0.789 0	<u>29.50</u>	<u>0.792 6</u>	29.59	0.794 5

注:加粗表示最优指标,下划线表示次优指标

对于 UCMerced 数据集中的 21 类遥感场景,本文提出的 SAF²Net 在 15 个和 21 个遥感场景中分别获得了最优的 PSNR/SSIM,在 5 个遥感场景中获得了次优的 PSNR;对于 AID 数据集中的 30 类遥感场景,SAF²Net 在 22 个和 29 个遥感场景中分别获得了最优的 PSNR/SSIM,在 8 个和 1 个遥感场景中分别获得了次优的 PSNR/SSIM。

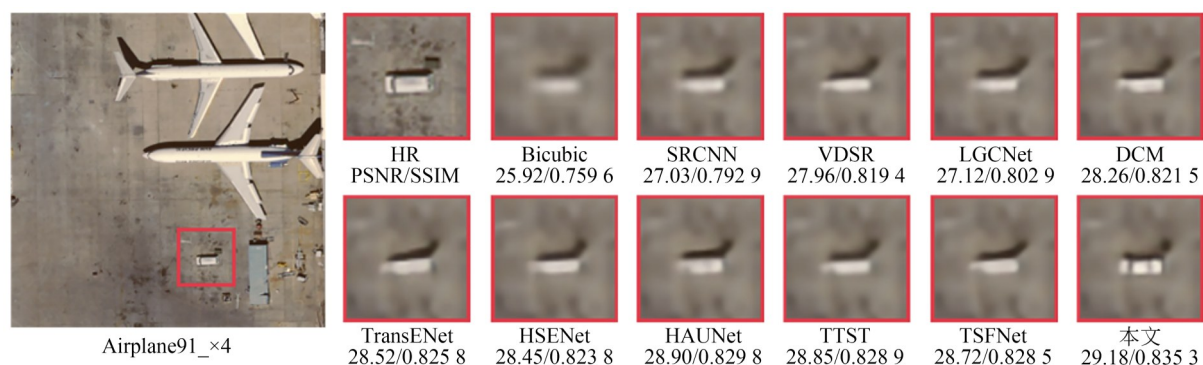
从结果可以看出,由于本文方法分别从两个不同的角度捕获全局混合尺度结构特征,利用了更多有效的全局先验信息,在 SSIM 指标上具有突出优势。对于沙滩(Beach)、沙漠(Desert)等高频细节较少的遥感场景,无需对高频信息进行建模,而本文方法由于使用像素级和频率级联合损失函数,侧重对超分过程中的高频信息进行学习建模,导致像素级指标 PSNR 在这类场景中存在一定程度的下降,取得了次优的结果。然而,得益于充分挖掘混合尺度的先验信息,并通过空间

域和频率域共同选择更具代表性的高频特征信息,SAF²Net 对于机场(Airport)、住宅区(Residential)、火车站(Railway Station)等具有更多重复性结构或复杂高频纹理细节的场景能够表现出更好的重建性能。

3.5 主观实验结果分析

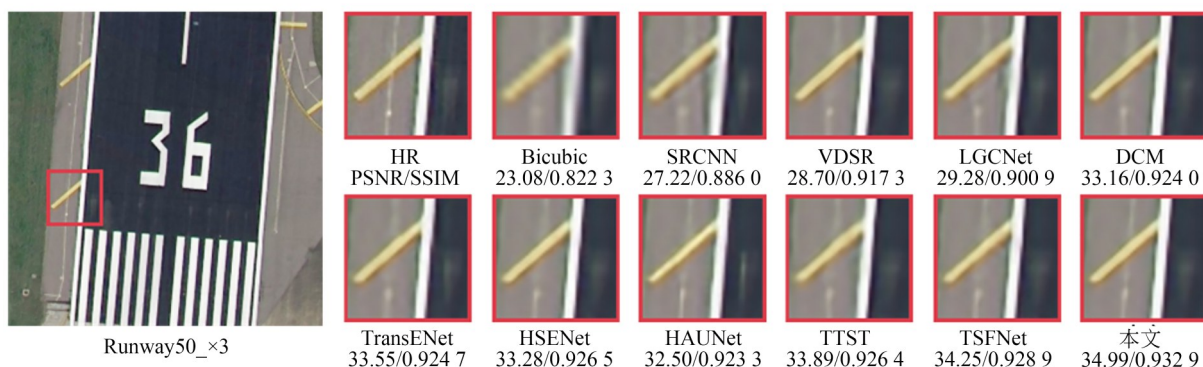
为了直观对比不同方法超分结果的视觉质量,图 7 和图 8 分别展示了两个数据集在放大倍率为 4 倍和 3 倍时的超分结果。

在 UCMerced 数据集 Airplane91 图像的 4 倍超分结果中,可以看出随着对比方法性能的不断提升小车的形状逐渐清晰可见,但对于更为细小的车窗却无法恢复,而本文算法除了能够得到与 HR 更相似的小车形状,也能够一定程度上恢复小车前后车窗更细小的纹理细节;在 Runway50 图像的 3 倍超分结果中,可以看出 Bicubic, SRCNN, LGCNet 等方法恢复的跑道线边缘存在明显的伪影和模糊,而 HAUNet 虽然能够得到较



(a) Airplane91 图像的 4 倍超分结果

(a) Airplane91 with $\times 4$ scale factor



(b) Runway50 图像的 3 倍超分结果

(b) Runway50 with $\times 3$ scale factor

图 7 UCMerced 数据集中不同方法的可视化结果

Fig. 7 Super-resolved outputs of different methods on UCMerced dataset

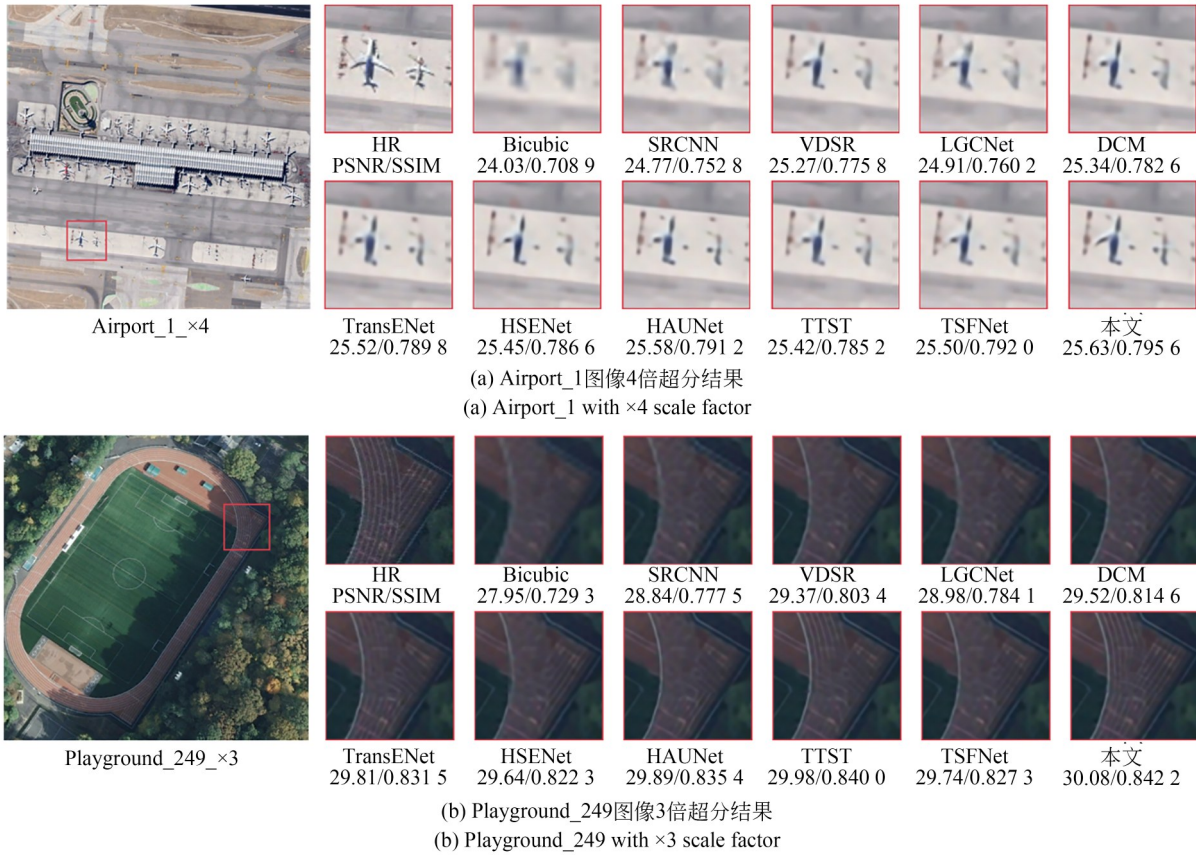


图 8 AID 数据集中不同方法的可视化结果

Fig. 8 Super-resolved outputs of different methods on AID dataset

为清晰的跑道线,但与 HR 相比缺失了一些轮廓细节信息,相比之下本文方法能够恢复出最清晰的跑道线和边缘纹理。

AID 数据集中 Airport_1 图像 4 倍超分结果和 Playground_249 图像 3 倍超分如图 8(a)和(b)所示,对比各算法的可视化结果,SAF²Net 能够得到更加高质量的结果,如轮廓更加清晰和完整的飞机机翼以及更少伪影的边缘和更加清晰的跑道线。实验结果表明,相比于其他对比方法,本文方法 SAF²Net 能够更有效地利用全局混合尺度的先验信息,并有效利用空间域和频率域之间的互补性,针对图像中高频细节区域进一步增强局部特征的建模,从而在主观视觉质量上也能够取得更有竞争力的超分结果。

3.6 实际遥感数据的主观实验结果

为了验证本文方法在遥感图像超分辨率任务中的泛化能力,本文分别基于哨兵 2 号(Sentinel-2)卫星数据和 HSRS-SC 黑河生态水文遥感

试验航空数据^[37],选取可见光三个波段生成对应的 RGB 遥感图像,其空间分辨率分别为 10 m/pixel 和 1 m/pixel。将生成的 RGB 遥感图像作为模型的输入,进行放大倍率为 4 的超分辨率重建实验,并展示部分方法的主观实验结果,如图 9 和图 10 所示。

从结果可以看出,本文方法面对不同传感器获取的实际遥感数据,相较于本文其他对比方法都能够获得具有更优的主观视觉质量的超分结果。具体来说,由于 SAF²Net 从遥感图像特点出发,充分挖掘图像中混合尺度的自相似性以及不同大小区域之间的关联特征,并从空间域和频率域两个维度共同进行局部特征的细化,有效提高了模型对高频部分的建模能力,成功地恢复出了更多逼真的边缘轮廓和清晰的纹理细节。上述实验结果进一步验证了 SAF²Net 模型对于实际遥感图像数据具有一定的泛化性能。

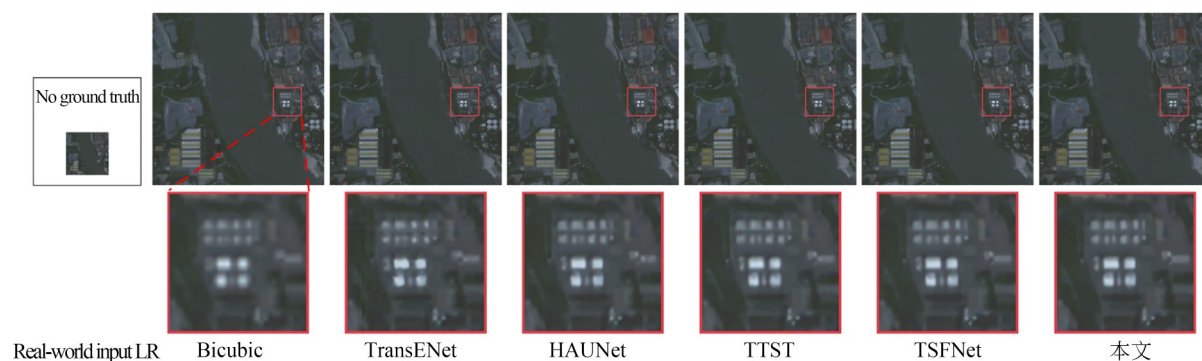


图9 哨兵2号卫星数据中不同方法的4倍超分可视化结果

Fig. 9 Super-resolved outputs of different methods on Sentinel-2 data

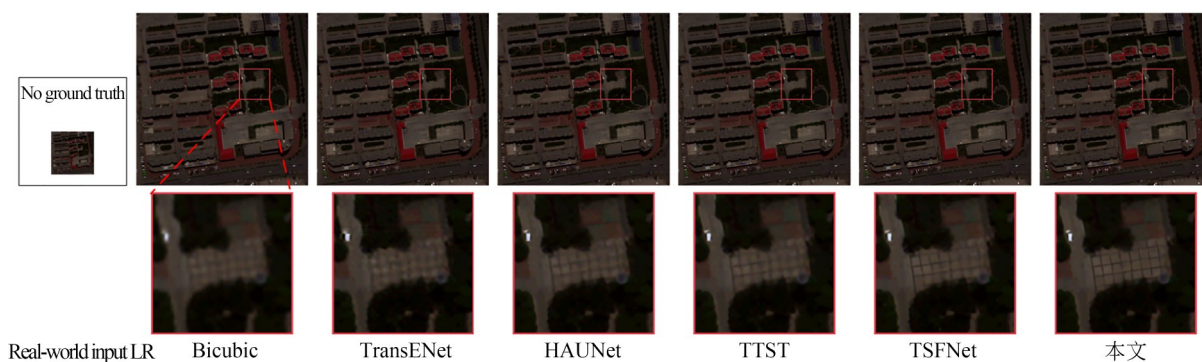


图10 HSRS-SC数据集中不同方法的4倍超分可视化结果

Fig. 10 Super-resolved outputs of different methods on HSRS-SC dataset

4 结 论

本文提出了一种面向单幅遥感图像超分辨率的空间自适应及频率融合网络 SAF²Net。该网络主要包括全局混合尺度特征提取模块和局部特征增强模块。其中,首先利用混合尺度空间自适应特征调制模块,挖掘遥感图像中的自相似性信息,实现不同尺度的全局特征表征;随后,通过全局多尺度感受野选择块自适应融合不同感受野的特征,捕获跨尺度区域关联特征以提升模型对全局上下文信息的建模能力。局部增强特征提取模块则联合空间自适应选择块与频率分离选择块,从空间域和频率域两个维度提取局部特征,有效融合空频互补信息,着重提升模型对图像高频内容的建模能力。

在 UCMerced 和 AID 两个遥感数据集上的对比实验表明,本文方法获得的 PSNR 和 SSIM

指标和主观视觉质量均优于其他对比方法。此外,针对广泛存在的包含大量重复性结构或复杂高频纹理的遥感场景类别,本文所提出的方法在超分辨率任务中展现出了具有竞争力的性能表现。然而,本文重点关注的是提升模型在预定义退化类型下的重建性能。在下一步的研究中,将考虑探索在未知且复杂退化类型条件下的超分辨率重建方法,进一步扩展遥感图像超分辨率方法在实际应用中的有效性。

作者贡献声明:

杨沂川:调查研究,论文构思,软件,验证,实验结果分析,初稿写作;

马中祺:提供资源,调查研究,验证;

周新尧:实验数据整理,编辑写作;

郑福建:实验数据整理,调查研究;

黄鸿:获取资助,项目管理,提供资源,指导,审核与编辑写作。

参考文献:

- [1] WU L S, FANG L Y, YUE J, *et al.* Deep bilateral filtering network for point-supervised semantic segmentation in remote sensing images[J]. *IEEE Transactions on Image Processing*, 2022, 31: 7419-7434.
- [2] ZHENG F J, LIN S, ZHOU W, *et al.* A lightweight dual-branch swin transformer for remote sensing scene classification[J]. *Remote Sensing*, 2023, 15(11): 2865.
- [3] LIU B H, XU C Y, CUI Z, *et al.* Progressive context-dependent inference for object detection in remote sensing imagery[J]. *IEEE Transactions on Image Processing*, 2022, 32: 580-590.
- [4] XIONG J W, LIU F, WANG X Y, *et al.* Siamese transformer-based building change detection in remote sensing images[J]. *Sensors*, 2024, 24(4): 1268.
- [5] JIA S, WANG Z H, LI Q Q, *et al.* Multiattention generative adversarial network for remote sensing image super-resolution[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5624715.
- [6] 易见兵, 陈俊宽, 曹锋, 等. 轻量级重参数化的遥感图像超分辨率重建网络设计[J]. *光学精密工程*, 2024, 32(2): 268-285.
YI J B, CHEN J K, CAO F, *et al.* Design of Lightweight re-parameterized remote sensing image super-resolution network[J]. *Opt. Precision Eng.*, 2024, 32(2): 268-285. (in Chinese)
- [7] HUANG N B, YANG Y, LIU J J, *et al.* Single-image Super-Resolution for Remote Sensing Data Using Deep Residual-Learning Neural Network[C]. *Neural Information Processing*. Cham: Springer International Publishing, 2017: 622-630.
- [8] HAUT J M, FERNANDEZ-BELTRAN R, PAOLETTI M E, *et al.* A new deep generative network for unsupervised remote sensing single-image super-resolution[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(11): 6792-6810.
- [9] KEYS R. Cubic convolution interpolation for digital image processing[J]. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1981, 29(6): 1153-1160.
- [10] THURNHOFER S. Edge-enhanced image zooming[J]. *Optical Engineering*, 1996, 35(7): 1862.
- [11] CHANG H, YEUNG D Y, XIONG Y M. Super-resolution through neighbor embedding[C]. *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2004. *CVPR 2004. June 27 - July 2, 2004, Washington, DC, USA*. IEEE, 2004: 1.
- [12] 朱福珍, 刘越, 黄鑫, 等. 改进的稀疏表示遥感图像超分辨率重建[J]. *光学精密工程*, 2019, 27(3): 718.
ZHU F Z, LIU Y, HUANG X, *et al.* Remote sensing image super-resolution based on improved sparse representation[J]. *Opt. Precision Eng.*, 2019, 27(3): 718. (in Chinese)
- [13] TIMOFTE R, DE V, GOOL L V. Anchored neighborhood regression for fast example-based super-resolution[C]. *2013 IEEE International Conference on Computer Vision. December 1-8, 2013, Sydney, NSW, Australia*. IEEE, 2013: 1920-1927.
- [14] 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.
- [15] LI Q, GONG M G, YUAN Y, *et al.* Symmetrical feature propagation network for hyperspectral image super-resolution[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5536912.
- [16] 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.
- [17] XU W J, XU G L, WANG Y, *et al.* High quality remote sensing image super-resolution using deep memory connected network[C]. *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium. July 22-27, 2018, Valencia, Spain*. IEEE, 2018: 8889-8892.
- [18] JIANG K, WANG Z Y, YI P, *et al.* A progressively enhanced network for video satellite imagery superresolution[J]. *IEEE Signal Processing Letters*, 2018, 25(11): 1630-1634.
- [19] HAUT J M, PAOLETTI M E, FERNANDEZ-BELTRAN R, *et al.* Remote sensing single-image superresolution based on a deep compendium model[J]. *IEEE Geoscience and Remote Sensing*

- Letters*, 2019, 16(9): 1432-1436.
- [20] 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.
- [21] ZHANG X R, LI Z Y, ZHANG T Y, *et al.* Remote sensing image super-resolution via dual-resolution network based on connected attention mechanism [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 60: 5611013.
- [22] HUAN H, LI P C, ZOU N, *et al.* End-to-end super-resolution for remote-sensing images using an improved multi-scale residual network[J]. *Remote Sensing*, 2021, 13(4): 666.
- [23] LEI S, SHI Z W. Hybrid-scale self-similarity exploitation for remote sensing image super-resolution [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 60: 5401410.
- [24] LEI S, SHI Z W, MO W J. Transformer-based multistage enhancement for remote sensing image super-resolution[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 60: 5615611.
- [25] WANG J R, WANG B L, WANG X X, *et al.* Hybrid attention-based U-shaped network for remote sensing image super-resolution [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 5612515.
- [26] XIAO Y, YUAN Q Q, JIANG K, *et al.* TTST: a top-k token selective transformer for remote sensing image super-resolution[J]. *IEEE Transactions on Image Processing*, 2024, 33: 738-752.
- [27] SHI W Z, CABALLERO J, HUSZÁR F, *et al.* Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 1874-1883.
- [28] XIAO Y, YUAN Q Q, HE J, *et al.* Remote sensing image super-resolution via cross-scale hierarchical trans-former[J]. *Geo-Spatial Information Science*, 2023, 27(6):1914-1930.
- [29] MEI Y Q, FAN Y C, ZHOU Y Q. Image super-resolution with non-local sparse attention [C]. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 20-25, 2021. Nashville, TN, USA. IEEE, 2021: 3516-3525.
- [30] LI X, WANG W H, HU X L, *et al.* Selective kernel networks[C]. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 15-20, 2019. Long Beach, CA, USA. IEEE, 2019: 510-519.
- [31] 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[J/OL]. 激光与光电子学进展, 2023: 1-20. (2023-11-09). [http://kns.cnki.net/KCMS/detail/detail.aspx? filename=JG-DJ2023110701C&dbname=CJFD&dbcode=CJFQ](http://kns.cnki.net/KCMS/detail/detail.aspx?filename=JG-DJ2023110701C&dbname=CJFD&dbcode=CJFQ).
- WANG Q Q, XIN Y L, ZHAO J, *et al.* Image Super-Resolution Reconstruction based on Efficient Global Attention Network[J/OL]. *China Industrial Economics*, 2023: 1-20. (2023-11-09). [http://kns.cnki.net/KCMS/detail/detail.aspx? filename=JGDJ2023110701C&dbname=CJFD&dbcode=CJFQ](http://kns.cnki.net/KCMS/detail/detail.aspx?filename=JGDJ2023110701C&dbname=CJFD&dbcode=CJFQ). (in Chinese)
- [32] 周颖, 裴盛虎, 陈海永, 等. 基于多尺度自适应注意力的图像超分辨率网络[J]. 光学精密工程, 2024, 32(6): 843-856.
- ZHOU Y, PEI S H, CHEN H Y, *et al.* Image super-resolution network based on multi-scale adaptive attention[J]. *Opt. Precision Eng.*, 2024, 32(6): 843-856. (in Chinese)
- [33] YANG Y, NEWSAM S. Bag-of-visual-words and spatial extensions for land-use classification [C]. *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*. San Jose California. ACM, 2010: 270-279.
- [34] XIA G S, HU J W, HU F, *et al.* AID: a benchmark data set for performance evaluation of aerial scene classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 55(7): 3965-3981.
- [35] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks [C]. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 1646-1654.
- [36] WANG J R, LU Y T, WANG S Z, *et al.* Two-stage spatial-frequency joint learning for large-fac-

tor remote sensing image super-resolution [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 5606813.

- [37] 徐科杰, 邓培芳, 黄鸿. HSRS-SC: 面向遥感场景分类的高光谱图像数据集[J]. 中国图象图形学

报, 2021, 26(8): 1809-1822.

XU K J, DENG P F, HUANG H. HSRS-SC: a hyperspectral image dataset for remote sensing scene classification [J]. *Journal of Image and Graphics*, 2021, 26(8): 1809-1822. (in Chinese)

作者简介:



杨沂川(2000—),男,重庆人,硕士研究生,2022年于重庆大学获得学士学位,主要从事图像增强和遥感图像超分辨率重建方面的研究。E-mail: yangyichuan@stu.cqu.edu.cn

通讯作者:



黄 鸿(1980—),男,湖南新宁人,博士,教授,博士生导师,2005年于重庆大学获得硕士学位,2008年于重庆大学获得博士学位,主要从事图像处理、机器视觉和图像测量方面的研究。E-mail: hhuang@cqu.edu.cn