液晶与显示
Chinese Journal of Liquid Crystals and Displays

文章编号:1007-2780(XXXX)XX-0001-11

FSSENet:频率-空间协同增强的跨传感器超分辨率重建网络

许浪杰[1,2],董　婷[1,2*],邵　攀[1,2],邵　文[1,2]
(1.三峡大学 水电工程智能视觉检测湖北省重点实验室,湖北 宜昌 443002;
2.三峡大学 计算机与信息学院,湖北 宜昌 443002)

摘要:高分辨率遥感影像是实现遥感应用的核心数据源之一。然而,受传感器分辨率、重访周期和天气等因素限制,使得高分辨率遥感影像有时很难覆盖研究区域。针对上述问题,本文提出一种频率-空间协同增强(Frequency-spatial synergistic enhancement,FSSE)的跨传感器超分辨率重建网络FSSENet。FSSENet包含三个主要创新模块:1)构建一种频率空间协同高频低频分离模块,首先利用傅里叶变换在频率域显式分离高低频成分,同时在空间域利用不同感受野的卷积核提取局部(高频)与全局(低频)信息,然后利用自注意力分别实现空间-频率域内高频与低频的独立协同增强,得到质量更优的高频和低频特征,从而提升模型对主体结构和边缘纹理的联合表征能力。2)设计一种低频门控增强的融合模块,首先利用低频特征生成通道和空间门控权重,实现对高频特征的自适应筛选与响应,然后与低频特征进行残差融合,实现高频和低频特征的有效融合,同时抑制噪声并强化边缘纹理细节。3)提出一种感受野渐进扩大的多尺度池化模块,实现模型增强对多尺度地物目标的上下文建模与结构保持能力。为验证FSSENet的有效性,制作两组数据集RSRData-FHJ和RSRData-GH进行实验,实验结果表明,与七种基准方法相比,FSSENet在RSRData-FHJ数据集上PSNR提高0.41至5.29 dB,SSIM实现0.011 5至0.176 7的改进,LPIPS、ERGAS均得到显著效果,为跨传感器遥感图像超分辨率提供了一种有效方案。

关　键　词:跨传感器遥感影像;超分辨率;频域特征;特征交互
中图分类号:TP394.1;TH691.9　　**文献标识码**:A　　**doi**:10.3724/CJLCD.2026-0139　　**CSTR**:32172.14.CJLCD.2026-0139

收稿日期:2026-07-24;**修订日期**:2026-11-02.
基金项目:自然资源部地理国情监测重点实验室开放基金(No.2025NGCM03);国家自然科学基金(No.41901341);湖北省自然科学基金(No.2024AFB867)
Supported by Open Fund for Key Laboratory of Geographic State Monitoring of Ministry of Natural Resources (No.2025NGCM03);Supported by National Natural Science Foundation of China (No.41901341);Natural Science Foundation of Hubei Province (No.024AFB867)
*通信联系人,E-mail:dongt@ctgu.edu.cn

FSSENet:a frequency-spatial synergistic enhancement network for cross-sensor remote sensing image super-resolution

XU Langjie[1,2], DONG Ting[1,2*], SHAO Pan[1,2], SHAO Wen[1,2]

(1. *Hubei Key Laboratory of Intelligent Vision Based Monitoring for Hydroelectric Engineering, China Three Gorges University, Yichang* 443002, *China*;
2. *College of Computer and Information Technology, China Three Gorges University, Yichang* 443002, *China*)

Abstract: High-resolution remote sensing images are one of the core data sources for supporting remote sensing applications such as land-cover monitoring, ecological assessment, disaster identification, and refined interpretation of ground objects. However, owing to the limitations of sensor spatial resolution, revisit period, imaging cost, and weather conditions, high-resolution images are sometimes difficult to obtain or cannot continuously cover the target study areas. To address this problem, this paper proposes FSSENet, a cross-sensor super-resolution reconstruction network based on frequency-spatial synergistic enhancement(FSSE). The proposed network aims to recover high-quality high-resolution remote sensing images from low-resolution observations by jointly enhancing frequency-domain information and spatial-domain structural representation. FSSENet mainly consists of three innovative modules: 1) A frequency-spatial synergistic high- and low-frequency separation module is constructed. In this module, Fourier transform is first employed to explicitly separate high- and low-frequency components in the frequency domain. Meanwhile, convolution kernels with different receptive fields are introduced in the spatial domain to extract local high-frequency information and global low-frequency information. Then, self-attention is used to independently perform synergistic enhancement of high- and low-frequency features between the spatial and frequency domains, thereby obtaining higher-quality high-frequency and low-frequency representations and improving the model's joint representation capability for main structures, edge contours, and texture details. 2) A low-frequency-gated enhanced fusion module is designed. This module first uses low-frequency features to generate channel and spatial gating weights, which enables adaptive selection and response enhancement of high-frequency features. The selected high-frequency details are then fused with low-frequency structural features through residual connections. In this way, the effective fusion of high- and low-frequency features is achieved, while noise interference is suppressed and edge and texture details are further strengthened. 3) A multi-scale pooling module with progressively enlarged receptive fields is proposed. By gradually expanding the receptive field, this module enhances the contextual modeling ability of the network for multi-scale ground objects and improves the preservation of spatial structures in complex remote sensing scenes. To verify the effectiveness of FSSENet, two cross-sensor datasets, RSRData-FHJ and RSRData-GH, are constructed for experiments, providing paired data support for model training and evaluation. Experimental results show that, compared with seven benchmark methods, FSSENet improves PSNR by 0.41 to 5.29 dB and SSIM by 0.115 0 to 0.176 7 on the RSRData-FHJ dataset. In addition, LPIPS and ERGAS also show significant improvements, demonstrating that the proposed method provides an effective solution for cross-sensor remote sensing image super-resolution reconstruction and practical remote sensing image enhancement tasks.

Key words: cross-sensor remote sensing image; super-resolution; frequency-domain features; feature interaction

1 引 言

遥感影像是生态评估和灾害识别等任务的重要数据来源。高分辨率遥感影像能够提供更清晰的道路、建筑物和农田纹理等信息[1-2]，但高分辨率数据的获取通常受传感器分辨率和重访周期等因素限制，使得高分辨率遥感影像有时很难覆盖研究区域[3-5]。因此，如何在已有低分辨率遥感数据基础上恢复更高分辨率影像，成为遥感影像处理中的重要问题。

超分辨率（Super-Resolution，SR）为突破成像硬件限制提供了一种有效途径[6]。该任务旨在由低分辨率影像恢复高分辨率影像中缺失的空间结构和高频细节信息[7-8]。遥感影像超分辨率重建（Remote Sensing Image Super-Resolution，RSISR）是超分辨率技术在遥感领域的重要应用，与自然图像相比，遥感影像具有地物类型复杂和尺度变化显著等特点，因此在重建过程中容易受到纹理缺失和噪声干扰等问题影响[9]。特别在跨传感器场景下，不同卫星平台在成像机制和空间分辨率等方面存在差异，进而削弱低分辨率与高分辨率影像之间的映射关系[10]。

受传感器大气条件、配准误差及复杂地物分布等因素影响，遥感影像常表现出地物类型多样、尺度差异显著、纹理分布密集和光谱响应复杂等特点[11]。因此，遥感超分辨率不仅需要恢复视觉细节，还需要尽可能保持地物结构和光谱响应的一致性。早期遥感超分辨率方法主要借鉴自然影像超分辨率模型[12]。Ducournau等[13]将卷积神经网络应用于海表温度数据降尺度，验证了深度卷积网络在遥感超分辨率中的可行性；Liebel等[14]利用CNN对Sentinel-2影像进行重建，提升了影像空间细节表达能力；Haut等[15]进一步结合残差结构和通道注意力机制，增强了网络对遥感影像空间特征的提取能力。

近年来，研究者进一步将注意力机制、生成对抗网络和Transformer引入遥感影像超分辨率任务[16-17]。Xu等[18]提出TE-SAGAN，通过自注意力机制提取全局信息，并结合感知损失和对抗损失增强重建影像的细节与纹理；Jia等[19]提出MA-GAN，利用多注意力融合策略提升复杂地物的重建效果。面向跨传感器场景，Li等[20]提出ConvFormerSR，将CNN的局部细节提取能力与Transformer的全局建模能力相结合，并通过特征融合模块整合多层次全局与局部信息。

现有超分辨率方法在跨传感器遥感影像重建场景中仍存在三方面不足。第一，现有方法多数侧重于空间域特征提取，对频率域中高频纹理信息和低频结构信息的显式建模不足，难以充分利用频率域特征对边缘细节与主体结构的互补表达作用[21]；第二，遥感影像超分辨率重建过程中，高频特征通常包含边缘和小尺度目标等细节信息，但也容易混入噪声和无效高频响应，可能导致重建影像出现边缘不稳定等问题；第三，基于卷积神经网络的方法虽然能够有效提取局部纹理、边缘轮廓等细节信息，但受卷积局部感受野限制，对大范围地物结构、全局上下文关系以及长距离依赖的建模能力不足，容易导致重建结果结构表达不完整[22]。

针对上述问题，本文面向Landsat-8与Sentinel-2的遥感影像超分辨率任务，提出一种频率-空间协同增强的跨传感器超分辨率重建网络（FSSENet，Frequency-Spatial Synergistic Enhancement Network）。该网络从频率域和空间域两个角度对遥感影像特征进行协同建模。首先，设计频率空间协同高频低频分离模块（Frequency-Spatial High-Low Separation Module，FSHLSM）对频率域的高低频进行分解，然后通过频率和空间交互的自注意力，增强频域信息与空间信息之间的互补表达；提出低频门控增强的融合模块（Low-Frequency-Gated Enhanced Fusion Module，LFGEFM），通过低频特征约束高频特征中的噪声，增强高频特征；最后，构建了感受野渐进扩大的多尺度池化模块（Progressive Receptive-Field Expansion Multi-Scale Pooling Module，PREMPM），提高模型对不同尺度地物目标的表达能力。通过上述设计，FSSENet能够在跨传感器条件下兼顾结构一致性、边缘清晰度和纹理细节恢复。

2 本文方法

图1所示为本文提出的FSSENet总体网络结构。首先，将输入的低分辨率遥感影像送入卷积层进行浅层特征提取。随后，浅层特征被送入

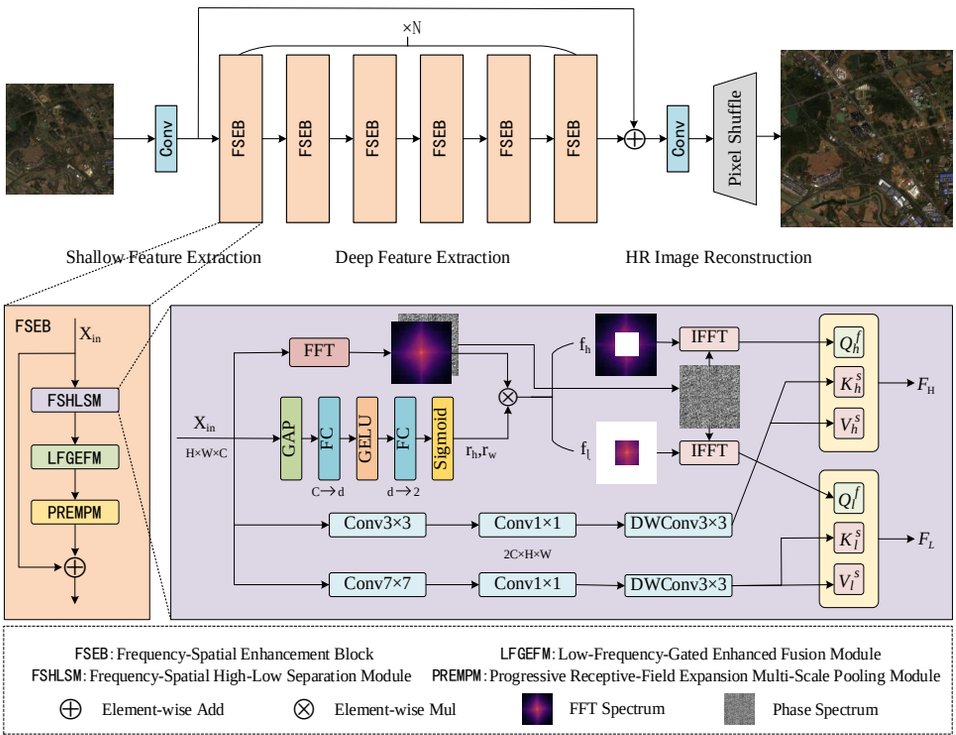


图1 FSSENet总体结构

Fig. 1 Overall architecture of FSSENet

由多个频率-空间增强块(Frequency-Spatial Enhancement Block,FSEB)堆叠构成的深层特征提取模块。FSEB是本文网络的核心组成单元,主要包括频率空间协同高频低频分离模块(FSHLSM)、低频门控增强的融合模块(LFGEFM)以及感受野渐进扩大的多尺度池化模块(PREMPM)。最后,本文在网络中引入全局残差连接,将浅层特征与深层增强特征进行融合。融合后的特征通过卷积映射和像素重排操作完成上采样重建,得到最终的超分辨率遥感影像。

2.1 频率空间协同高频低频分离模块

在跨传感器遥感影像超分辨率重建中,低分辨率影像不仅缺失大量边缘纹理等高频细节,同时也容易受到不同传感器成像差异的影响。现有方法多从空间域直接学习低分辨率影像到高分辨率影像的映射关系,难以显式区分高频信息与低频信息,容易造成重建结果中边缘模糊和纹理不连续等问题。因此,本文提出频率空间协同高频低频分离模块(FSHLSM)。该模块首先利用傅里叶变换在频率域对输入特征进行高低频分解,以分别表征边缘纹理和主体结构信息;同时在空间域采用不同感受野的卷积提取局部细节特征和上下文结构特征,形成空间层面的高低频表示。最终两组高频低频特征通过跨域交叉注意力机制增强频率域与空间域特征之间的互补表达,从而提升模型对遥感影像结构信息和纹理细节的联合重建能力。

如图1所示,给定输入特征 $X_{\text{in}}\in R^{H\times W\times C}$,FSHLSM首先通过快速傅里叶变换将其映射到频率域 F,并根据输入特征的全局统计信息自适应生成高、低频分离掩码。通过频谱分解及逆傅里叶变换,获得频率域引导的高频细节特征和低频结构特征。本文根据输入特征自适应生成低频窗口范围,首先利用全局平均池化获取输入特征的全局语义描述,然后通过两个全连接层预测频率窗口在高度和宽度方向上的比例,最后通过Sigmoid激活函数把比例参数限制在0到1之间:

$$[r_h, r_w]=\delta\left(\mathrm{FC}_2\left(\sigma\left(\mathrm{FC}_1\left(\mathrm{GAP}(X_{\text{in}})\right)\right)\right)\right), \quad (1)$$

其中:$GAP(\cdot)$ 表示全局平均池化,$\sigma(\cdot)$ 表示GELU激活函数,$\delta(\cdot)$ 表示Sigmoid激活函数。r_h 和 r_w 分别控制低频窗口在高度和宽度方向上的范围。

基于预测得到的比例参数,低频窗口掩码定

义为：

$$M_l = I\left(\frac{H}{2}-h,\frac{H}{2}+h\right)\otimes I\left(\frac{W}{2}-w,\frac{W}{2}+w\right),$$
$$h=\left\lfloor\frac{H}{2}\cdot r_h\right\rfloor,\ w=\left\lfloor\frac{W}{2}\cdot r_w\right\rfloor, \tag{2}$$

其中：$I(a,b)$ 表示坐标 $[a,b]$ 的指示函数，$\otimes$ 表示外积操作，M_l 表示低频掩码。之后，通过全1矩阵减去低频掩码来得到高频掩码 M_h。这两个掩码通过元素相乘作用于频谱特征，得到分离后的高频和低频频谱特征图，并通过对应的相位图逆傅里叶变换恢复到空间表示：

$$F_l=\mathrm{IFFT}(F\odot M_l),F_h=\mathrm{IFFT}(F\odot M_h), \tag{3}$$

其中：F_l 和 F_h 分别表示频率域引导下的低频特征和高频特征，$\odot$ 表示逐元素乘法。

在空间域分支中，采用不同感受野的卷积操作提取具有不同空间尺度特征的表示。其中，3×3卷积侧重于提取局部边缘及细粒度纹理信息，7×7卷积用于捕获较大范围内的空间结构与上下文关系，从而形成与频率域高、低频特征相对应的空间特征表示：

$$F_h^{\mathrm{spa}}=\mathrm{Conv}_{3\times3}(X_{\mathrm{in}}),F_l^{\mathrm{spa}}=\mathrm{Conv}_{7\times7}(X_{\mathrm{in}}), \tag{4}$$

其中：F_h^{spa} 表示空间域局部细节特征，F_l^{spa} 表示空间域上下文结构特征。

在获得频率域和空间域的高、低频特征后，本文构建跨域交互注意力机制。以高频分支为例，首先将频率域高频特征 F_h 映射为查询向量，Q_h 同时将空间域局部细节特征 F_h^{spa} 映射为键向量 K_h^s 和值向量 V_h^s，并通过缩放点积注意力建模二者之间的关联：

$$Q_h=DW\mathrm{Conv}_{3\times3}\left(\mathrm{Conv}_{1\times1}(F_h)\right), \tag{5}$$

$$K_h^s,V_h^s=DW\mathrm{Conv}_{3\times3}\left(\mathrm{Conv}_{1\times1}(F_h^{\mathrm{spa}})\right), \tag{6}$$

$$F_H=\mathrm{Softmax}\left(\frac{Q_h(K_h^s)^{\mathrm{T}}}{\sqrt{D}}\right)V_h^s, \tag{7}$$

其中：D 表示单个注意力头的特征维度，F_H 表示经过频率域信息增强后的高频交互特征。对于低频分支，采用相同的交互策略，将频率域低频特征与空间域上下文结构特征进行关联建模，得到低频增强特征 F_L，增强网络对主体结构及大范围空间依赖关系的表达能力。通过该机制，FSHLSM能够建立频率与空间之间的特征对应关系，从而增强模型对边缘纹理和主体结构的联合表达能力。

2.2 低频门控增强的融合模块

高频特征通常包含边缘以及小尺度目标等细节信息，但其中可能混入噪声；低频特征包含更稳定的区域结构，能够为细节筛选提供可靠的结构先验。本文设计了低频门控增强的融合模块（LFGEFM），以低频结构信息生成通道和空间门控权重，对增强后的高频特征进行自适应筛选，并通过残差连接完成细节特征与结构特征的有效融合，从而降低无效高频响应对重建结果的干扰。

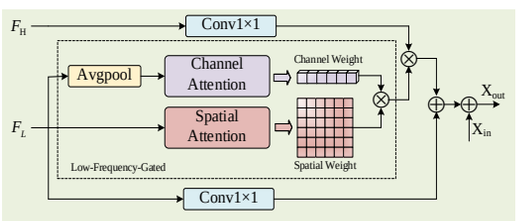


图2 LFGEFM模块结构

Fig. 2 Architecture of the LFGEFM module

如图2所示，给定高频特征 F_H、低频特征 F_L 以及待增强的主分支特征 X_{in}，LFGEFM首先利用低频结构特征生成通道门控权重和空间门控权重：

$$A_c=W_c\left(\mathrm{GAP}(F_L)\right),A_s=W_s(F_L), \tag{8}$$

其中：W_c 和 W_s 表示对应的通道和空间注意力操作，由卷积和ReLU激活函数组成，A_c 表示通道门控权重，A_s 表示空间门控权重，$\delta(\cdot)$ 表示Sigmoid激活函数。二者通过逐元素相乘得到最终的低频门控权重 G，随后，对高频细节特征进行轻量细化，并利用低频门控权重筛选有效高频响应：

$$G=A_c\odot A_s,\hat{F}_h=\mathrm{Conv}_{1\times1}(F_h)\odot G, \tag{9}$$

其中：$\hat{F}_h$ 表示经低频门控保留下来的有效高频细节。

在融合阶段，低频结构特征首先经过卷积投影，然后与筛选后的高频细节进行融合：

$$F_s=\mathrm{Conv}_{1\times1}(F_l)+\hat{F}_h, \tag{10}$$

其中：F_s 表示融合后的特征。

本文并没有对高低频融合后的中间特征进行输出，而是将融合特征 F_s 与初始输入特征 X_{in} 进行残差连接。该方式能够有效保留原始输入信息，降低高低频分解过程中可能带来的信息损失。

2.3 感受野渐进扩大的多尺度池化模块

遥感影像中同时包含建筑物、水体和农田等不同尺度地物，单一感受野难以兼顾局部纹理细节和大范围空间结构。如图3所示，为增强模型对多尺度地物的表征能力，本文设计感受野渐进扩大的多尺度池化模块(PREMPM)。与并行式选择核机制不同，PREMPM采用3×3、5×5和7×7卷积逐级扩大感受野，使特征提取过程由局部纹理逐步过渡到中尺度结构和大范围上下文建模。

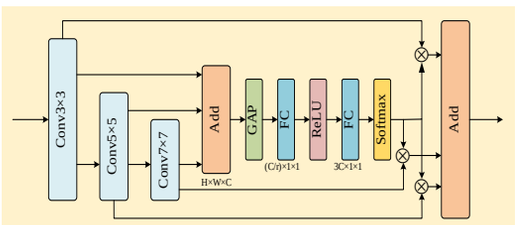


图3 PREMPM模块结构

Fig. 3 Architecture of the PREMPM module

给定输入特征 X，PREMPM首先利用3×3卷积提取局部细节特征 F_3，随后将 F_3 输入5×5卷积以获得中尺度结构特征 F_5，最后通过7×7卷积进一步提取大范围上下文特征 F_7。该过程可表示为：

$$F_3 = \mathrm{Conv}_{3\times3}(X), \tag{11}$$

$$F_5 = \mathrm{Conv}_{5\times5}(F_3), \tag{12}$$

$$F_7 = \mathrm{Conv}_{7\times7}(F_5). \tag{13}$$

在获得三个输出特征后，PREMPM首先对多尺度特征进行聚合，并通过全局平均池化和轻量映射生成不同尺度分支的自适应权重。最终，网络根据输入内容自适应地对三种尺度特征进行加权融合：

$$F_A = F_3 + F_5 + F_7, \tag{14}$$

$$[\alpha_{3,5,7}] = \delta\left(\mathrm{FC}_2\left(\sigma\left(\mathrm{FC}_1\left(\mathrm{GAP}(F_A)\right)\right)\right)\right), \tag{15}$$

$$F_{\mathrm{out}} = \alpha_3 F_3 + \alpha_5 F_5 + \alpha_7 F_7. \tag{16}$$

其中：F_A 表示不同尺度的聚合特征，$\delta(\cdot)$ 表示Sigmoid激活函数，$\sigma(\cdot)$ 表示ReLU激活函数，$\alpha_{3,5,7}$ 表示不同感受野分支的融合权重。

通过该设计，PREMPM能够根据输入影像内容动态选择合适的尺度信息，在增强局部纹理细节的同时保持较大范围的结构一致性。

3 实验结果

3.1 数据集

为了弥补跨传感器遥感超分辨率数据集的不足和降低不同传感器成像特性差异对超分辨率重建的影响，本文基于Google Earth Engine(GEE)平台构建Landsat-8与Sentinel-2跨传感器配对遥感影像数据集，并从影像获取、空间对应及图像块裁剪等方面对数据进行预处理。在遥感影像获取阶段，选取Landsat-8与Sentinel-2的地表反射率产品，分别作为低分辨率影像和高分辨率参考影像。地表反射率产品能够在一定程度上降低大气条件对影像辐射特性的影响。针对不同传感器在波段响应及观测条件方面存在的差异，在数据构建过程中选取具有对应光谱特征的波段和云量最低的影像，为跨传感器超分辨率提供数据基础。

利用GEE平台提供的地理参考信息，对Landsat-8与Sentinel-2影像进行空间范围匹配，使不同空间分辨率的影像覆盖相同的地理区域。对两类影像进行成对裁剪，将30m分辨率的Landsat-8影像裁剪为160×160的LR影像块，对应的10m分辨率Sentinel-2影像被裁剪为480×480的HR影像块，构建了空间位置一致的HR/LR影像对数据集，即RSRData-FHJ和RSRData-GH。RSRData-FHJ数据集覆盖中国的福建省、湖南省和江苏省，其中包含训练集1044对和测试集999对，而RSRData-GH数据集涵盖了中国的广东省和湖北省，其中包含训练集1038对和测试集872对。

3.2 实验细节

本文采用PyTorch深度学习框架来实现所提出的SR模型，并在配备NVIDIA GeForce RTX 3060 GPU的计算平台上完成所有实验。训练过程中，模型以随机裁剪得到的64×64像素LR影像作为输入，其对应的HR影像大小为192×192像素。为增强模型的泛化能力，我们对HR/LR影像对进行随机旋转等数据增强操作。初始学习率设置为 1×10^{-4}，并采用Adam优化器对模型进行训练，批处理大小(Batch size)设置为2，总训练轮数约为300个epoch。对于模型参数设置，FSEB模块的数量配置为10个。

为了全面评价不同模型在遥感图像超分辨率重建任务中的性能，本文主要采用四个指标对重建结果进行定量评估。峰值信噪比（Peak Signal-to-Noise Ratio，PSNR）：用于衡量重建影像与参考高分辨率影像之间的像素级误差，反映模型的整体重建精度；结构相似性（Structural Similarity Index Measure，SSIM）：主要从亮度、对比度和结构三个方面评价影像相似性，用于衡量重建结果对地物结构和空间纹理的保持能力；学习感知图像块相似度（Learned Perceptual Image Patch Similarity，LPIPS）：通过预训练深度网络提取深层感知特征，衡量重建影像与参考影像在视觉感知层面的差异，能够反映图像纹理真实感和感知质量[23]；相对全局无量纲合成误差（Relative Dimensionless Global Error in Synthesis，ERGAS）：用于衡量重建影像的全局相对误差，常用于遥感图像重建任务中的综合质量评价[14]。一般而言，PSNR和SSIM越高表示重建效果越好，而LPIPS和ERGAS越低表示重建影像与参考影像越接近。

3.3 实验结果

FSSENet进一步与七种基准的超分辨率方法进行比较和分析，即Bicubic、SRCNN[24]、TransENet[25]、HAT[26]、ConvFormerSR[20]、SRConvNet[27]和DMNet[28]。不同方法在两个数据集上的单幅影像重建结果如图4和图5所示。

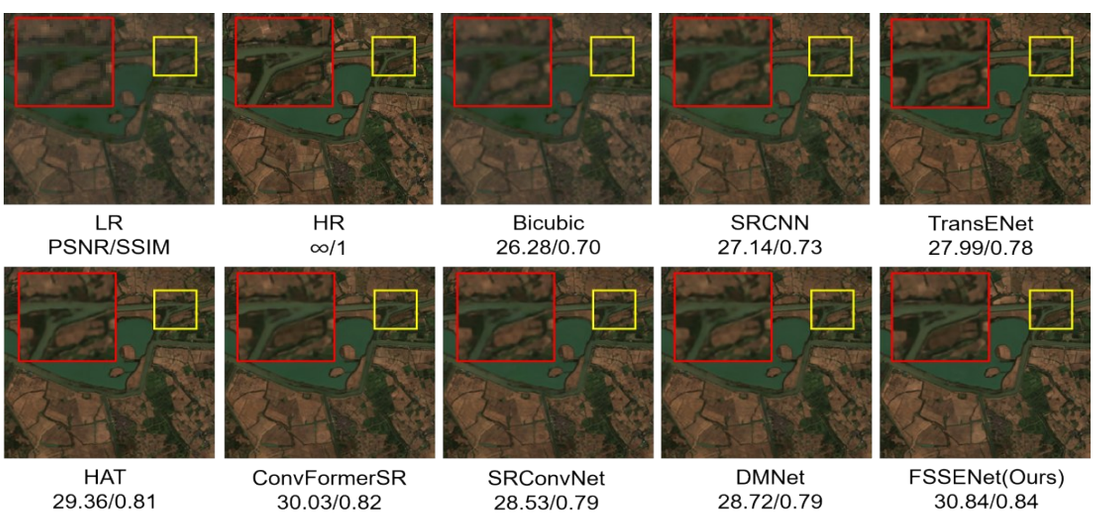


图4 FSSENet在RSRData-FHJ数据集上与其他模型的实验对比图

Fig. 4 Experimental comparison of FSSENet with other models on RSRData-FHJ dataset

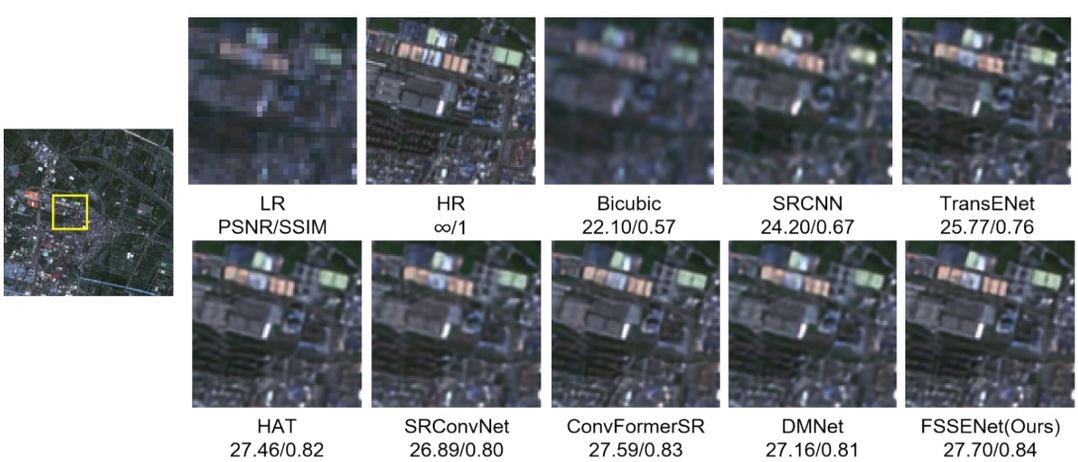


图5 FSSENet在RSRData-GH数据集上与其他模型的实验对比图

Fig. 5 Experimental comparison of FSSENet with other models on RSRData-GH dataset

在图4中，红色框内影像为黄色框区域的放大效果图。可以看出，Bicubic和SRCNN的重建结果整体较为模糊，在建筑物边缘和道路边界处存在明显的细节缺失；TransENet、HAT和ConvFormerSR等深度学习方法能够提升影像清晰度，但在局部复杂地物区域仍存在纹理不连续或局部伪影等现象。相比之下，FSSENet能够更准确地恢复地物边缘和局部纹理细节，重建结果更加完整和清晰，且未出现明显的噪声和伪影。

图5展示了不同方法在RSRData-GH数据集上的视觉对比结果，左侧为原始遥感图像，右侧各图像为不同超分辨率重建模型对黄色框区域的放大重建结果。通过对比可以看出，Bicubic、SRCNN和TransENet等方法对高频纹理的恢复不足，建筑物边缘较为模糊；ConvFormerSR和HAT能够恢复更多空间细节，但在密集建筑区域仍可能出现边缘粘连或纹理失真。FSSENet在建筑物轮廓、道路边界以及局部纹理连续性方面更接近HR参考图像。

此外，在RSRData-FHJ和RSRData-GH两个数据集上对FSSENet进行了全面的定量评估，如表1所示，FSSENet在PSNR、SSIM、LPIPS和ERGAS方面优于其他模型。具体而言，对于RSRData-FHJ数据集，FSSENet与其他模型相比，PSNR实现了0.41至5.29 dB的改进，SSIM实现了0.1150-0.1767的改进。LPIPS值比其他模型低0.0348-0.2455，ERGAS值低0.3248-6.3492。同样，对于RSRData-GH数据集，FSSENet在PSNR中实现了0.19至5.25 dB的增加，在SSIM中实现了0.0090-0.2286的增加，和其他超分模型相比LPIPS值降低了0.0236-0.3810，ERGAS值降低了0.1062-4.9023。与RSRData-FHJ相比，RSRData-GH包含更加复杂的地物分布和跨区域成像差异，FSSENet仍能够保持较高的重建精度。综合两个数据集的实验结果，FSSENet在像素级精度、结构保持以及感知质量等方面均取得稳定提升。

表1 不同方法在RSRData-FHJ和RSRData-GH数据集上的定量比较

Tab. 1 Quantitative comparison of different methods on the RSRData-FHJ and RSRData-GH datasets

Method	RSRData-FHJ					RSRData-GH				
	PSNR(dB) ↑	SSIM ↑	LPIPS ↓	ERGAS ↓	Time ↓	PSNR(dB) ↑	SSIM ↑	LPIPS ↓	ERGAS ↓	Time ↓
Bicubic	25.71	0.673 5	0.478 5	13.632 1	—	23.11	0.614 4	0.554 9	10.677 9	—
SRCNN	27.44	0.734 8	0.369 1	10.896 9	1 623	24.98	0.696 3	0.429 9	8.458 3	1 588
TransENet	29.29	0.801 8	0.323 3	8.808 3	4 572	27.11	0.794 8	0.302 7	6.631 7	4 038
HAT	30.58	0.836 6	0.278 1	7.607 7	5 609	28.17	0.834 0	0.214 7	5.880 9	5 785
ConvFormerSR	30.59	0.838 7	0.267 8	7.622 6	5 473	28.13	0.831 5	0.197 5	5.951 5	3 136
SRConvNet	29.82	0.814 6	0.295 5	8.312 5	1 914	27.62	0.811 9	0.269 4	6.279 7	1 747
DMNet	29.83	0.815 5	0.315 3	8.261 8	2 167	27.83	0.821 1	0.261 1	6.112 8	2 279
FSSENet (Ours)	**31.00**	**0.850 2**	**0.233 0**	**7.282 9**	**3 577**	**28.36**	**0.843 0**	**0.173 9**	**5.774 7**	**3 952**

此外，为评估不同方法的计算效率，本文对各模型的单幅影像平均重建时间进行了比较，结果如表1所示。在RSRData-FHJ和RSRData-GH测试集上，FSSENet的平均重建时间分别为3577 ms和3952 ms，均低于TransENet和HAT，但高于部分轻量化方法。综合来看，FSSENet虽未取得最优推理速度，但在PSNR、SSIM、LPIPS和ERGAS指标上均表现最优，表明该方法在一定计算开销下能够获得更高的重建精度，实现重建质量与计算效率之间的有效权衡。

上述结果验证了所提网络在两组跨传感器配对数据集上的重建有效性，但由于尚未单独评估不同程度光谱偏移和几何错位条件下的性能变化，现有实验结果不能直接证明模型具有显式的光谱校正或几何配准能力。后续研究将进一步探索面向跨传感器成像差异的光谱一致性约束和可学习特征对齐策略，以增强模型对复杂成像条件的适应能力。

3.4 消融实验

为了分析FSSENet中各关键模块对跨传感器遥感超分性能的影响，本文在RSRData-FHJ数据集上开展了消融实验，结果如表2所示。

表2 FSSENet关键模块消融实验结果

Tab. 2 Ablation study results of key modules in FSSENet

Method	Cross Attention	LFGEFM	PREMPM	PSNR ↑	SSIM ↑
A	×	√	√	30.63	0.842 7
B	√	×	√	30.90	0.848 8
C	√	√	×	30.38	0.838 4
D	√	√	√	**31.00**	**0.850 2**

其中，A表示去除FSHLSM中的频率-空间交互注意力机制；B表示去除低频门控增强的融合模块LFGEFM，并用Concat操作拼接高频低频特征；C表示去除渐进式感受野扩大多尺度模块PREMPM；D表示完整的FSSENet模型。所有实验均采用相同的数据划分、训练策略和评价指标，以保证结果的可比性。

从表2可以看出，完整模型D的PSNR/SSIM在所有消融对比模型中表现最好。模型A的PSNR和SSIM相比完整模型分别降低0.37 dB和0.0075。这说明频率域与空间域之间的交互建模有助于提高结构信息和细节纹理的联合表达能力。

当去除LFGEFM后，模型B相比完整模型分别下降0.10 dB和0.0014。虽然整体下降幅度相对较小，但低频门控增强的融合模块仍能为模型带来稳定增益。因为低频结构特征能够为高频细节选择提供约束，使模型在补充边缘和纹理信息时减少无效高频响应或伪纹理的干扰，从而提高重建效果。

模型C的PSNR和SSIM相比完整模型分别降低0.62 dB和0.0118，是三组消融设置中下降最明显的一组。遥感图像中同时包含建筑物、道路和水体等不同尺度地物，单一感受野难以兼顾局部纹理和大范围空间结构。PREMPM通过引入不同尺度卷积分支并进行自适应融合，使模型能够根据输入内容动态选择合适的上下文信息，从而提升多尺度地物的重建质量。总体来看，去除任一关键模块都会导致模型性能下降。消融实验结果与表1中的对比实验结果相一致，进一步验证了FSSENet各模块设计的合理性。

4 结　论

高分辨率遥感影像有时很难覆盖研究区域，本文针对Landsat-8到Sentinel-2跨传感器遥感影像3倍超分辨率重建中存在的成像机制差异、细节缺失和边缘模糊等问题，提出了一种频率-空间细节重建网络FSSENet。该网络通过频率分解对深层特征进行高低频建模，以分别表征纹理细节和主体结构；同时构建频率空间协同高频低频分离模块，增强频域信息与空间结构信息之间的互补表达。在此基础上，本文进一步引入低频门控增强的融合模块和感受野渐进扩大的多尺度池化模块，以提高模型对边缘纹理和多尺度地物结构的恢复能力。

基于RSRData-FHJ和RSRData-GH数据集，本文对所提方法进行了实验验证。结果表明，FSSENet在两个数据集上均取得了较好的重建性能，在RSRData-FHJ数据集上获得31.00 dB的PSNR和0.8502的SSIM，在RSRData-GH数据集上获得28.36 dB的PSNR和0.8430的SSIM，同时在LPIPS和ERGAS指标上也优于多种对比方法。总体来看，FSSENet能够有效融合频率域细节信息与空间域结构信息，为跨传感器遥感影像超分辨率重建提供了一种可行方法。未来将进一步扩展更多传感器组合和地理区域数据，并结合模型复杂度分析，提高方法在实际遥感应用中的适用性。

致谢：本文得到三峡大学先进计算中心资助支持

参 考 文 献：

[1] Jia P., Tian D., Wang Y. *et al*. *Airborne optical imaging technology*: a road map in CIOMP. Light Sci Appl 14, 119 (2025). https://doi.org/10.1038/s41377-025-01776-3

[2] KuhnJeff, LodieuNicolas, Rafael Rebolo López, *et al*. Creating ground-based telescopes for imaging interferometry: SELF and ELF[J]. *Light: Advanced Manufacturing* 6, 33(2025). doi: 10.37188/lam.2025.033

[3] 赵晓，杜光磊. 基于条件先验增强和扩散模型的遥感图像超分辨重建算法[J]. 液晶与显示，2025,40(7):1067-1079.
ZHAO X, DU G L. Super-resolution of remote sensing images based on conditional prior enhancement and diffusion models[J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(7): 1067-1079. (in Chinese)

[4] 李智,高连如,郑珂，等. 高分辨率遥感影像场景分类研究进展[J]. 遥感学报，2024,28(11):2739-2760.
LI Z, GAO L R, ZHENG K, *et al*. Research progress on high-resolution remote sensing image scene classification [J]. *Journal of Remote Sensing*, 2024, 28(11): 2739-2760. (in Chinese)

[5] Karwowska K, Wierzbicki D. Using super-resolution algorithms for small satellite imagery: A systematic review[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2022, 15: 3292-3312.

[6] Wei Ren, LiangQianxi, Peng Xi. Multifocal metalenses drive super-resolution imaging of brain organoids[J]. *Light: Advanced Manufacturing* 6, 71(2025). doi: 10.37188/lam.2025.071

[7] 李彬，喻夏琼，王平，等. 基于深度学习的单幅影像超分辨率重建综述[J]. 计算机工程与科学，2021，43(01)：112-124.
LI B, YU X Q, WANG P, *et al*. A review of single image super-resolution reconstruction based on deep learning [J]. *Computer Engineering & Science*, 2021, 43(1):112-124. (in Chinese)

[8] 李明锴,徐其志. 分级监督范式指导下的遥感影像超分辨率方法[J]. 遥感学报，2024，28(7)：1746-1759.
LI M K, XU Q Z. Remote sensing image super-resolution method guided by hierarchical supervision paradigm[J]. *Journal of Remote Sensing*, 2024, 28(7):1746-1759. (in Chinese)

[9] Wang P, Bayram B, Sertel E. A comprehensive review on deep learning based remote sensing image super-resolution methods[J]. *Earth—Science Reviews*, 2022, 232: 104110.

[10] Wang J, Ma A, Zhong Y, *et al*. Cross-sensor domain adaptation for high spatial resolution urban land-cover mapping: From airborne to spaceborne imagery[J]. *Remote Sensing of Environment*, 2022, 277: 113058.

[11] 付奕博，谢东海，王志博，等. 基于条件控制扩散模型的遥感影像超分辨率增强算法[J]. 地球信息科学学报，2024，26(10)：2384-2393.
FU Y B, XIE D H, WANG Z B, *et al*. Remote sensing image super-resolution enhancement algorithm based on conditional control diffusion model [J]. *Journal of Geo—information Science*, 2024, 26 (10) : 2384-2393. (in Chinese)

[12] 贺兴，王磊，张鹏超，等. 基于多维注意力网络的图像超分辨率重建[J]. 液晶与显示，2025,40(7):1056-1066.
HE Xing, WANG Lei, ZHANG Pengchao, *et al*. Image super-resolution reconstruction based on multidimensional attention network[J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(7): 1056-1066.

[13] AurelienDucournau, RonanFablet. Deep learning for ocean remote sensing: an application of convolutional neural networks for super-resolution on satellite-derived SST data[J]. *2016 9TH IAPR WORKSHOP ON PATTERN RECOGNITION IN REMOTE SENSING (PRRS)*, 2016.

[14] Liebel L, Körner M. Single-image super resolution for multispectral remote sensing data using convolutional neural networks[J]. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences — ISPRS Archives*, 2016, 41: 883-890.

[15] Juan Mario Haut, Fernández-BeltranRubén, Mercedes Eugenia Paoletti, *et al*. Remote Sensing Image Superresolution Using Deep Residual Channel Attention [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(11): 9277-9289.

[16] Hou Y., Wang W., Fu Y. *et al*. *Multi—resolution analysis enables fidelity—ensured deconvolution for fluorescence microscopy. eLight* 4, 14 (2024). https://doi.org/10.1186/s43593-024-00073-7

[17] Zhang Y., Yang E., Yoon H. H. *et al*. *Reconstructive spectrometers*: hardware miniaturization and computational reconstruction. eLight 5, 23 (2025). https://doi.org/10.1186/s43593-025-00101-0

[18] Xu Yongyang, Luo Wei, Hu Anna, *et al*. TE-SAGAN: An Improved Generative Adversarial Network for Remote Sensing Super-Resolution Images[J]. *Remote Sensing*, 2022, 14(10): 2425.

[19] Jia Sen, Wang Zhihao, Li Qingquan, *et al*. Multiattention Generative Adversarial Network for Remote Sensing Image Super-Resolution[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60.

[20] Li Junjie, Meng Yizhuo, Tao Chongxin, *et al*. ConvFormerSR: Fusing Transformers and Convolutional Neural Networks for Cross-Sensor Remote Sensing Imagery Super-Resolution[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62.

[21] Fang J, Yu Y, Wang Z, *et al*. An image arbitrary-scale super-resolution network using frequency-domain information[J]. *ACM Transactions on Multimedia Computing, Communications and Applications*, 2023, 20(3): 1-23.

[22] Xiao Y, Yuan Q, Jiang K, *et al*. Frequency-assisted mamba for remote sensing image super-resolution[J]. *IEEE Transactions on Multimedia*, 2024, 27: 1783-1796.

[23] Chung M, Jung M, Kim Y. Enhancing remote sensing image super-resolution guided by bicubic-downsampled low-resolution image[J]. *Remote Sensing*, 2023, 15(13): 3309.

[24] Dong C, Loy C C, He K, *et al*. Learning a deep convolutional network for image super-resolution[C]//European conference on computer vision. Cham: Springer International Publishing, 2014: 184-199.

[25] Lei Sen, Shi Zhenwei, Mo Wenjing. Transformer-Based Multistage Enhancement for Remote Sensing Image Super-Resolution[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60.

[26] Chen X, Wang X, Zhou J, *et al*. Activating more pixels in image super-resolution transformer[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2023: 22367-22377.

[27] Li Feng, Cong Runmin, Wu Jingjing, *et al*. SRConvNet: A Transformer-Style ConvNet for Lightweight Image Super-Resolution[J]. *International Journal of Computer Vision*, 2025, 133(1): 173-189.

[28] Li W, Guo H, Hou Y, *et al*. Dual-domain modulation network for lightweight image super-resolution[J]. *IEEE Transactions on Multimedia*, 2026.

作者简介:



许浪杰,男,硕士研究生,2024年于湖北汽车工业学院获得学士学位,主要从事深度学习、遥感影像超分处理的研究。E-mail:2195470052@qq.com



董　婷,女,博士,副教授,2016年于武汉大学获工学博士学位,主要从事模式识别,遥感图像处理,3S技术等相关教学研究工作。E-mail:dongt@ctgu.edu.cn