

中图法分类号: TP751 文献标识码: A 文章编号: 1006-8961(2025)04-0989-14

论文引用格式: Yang Y, Wang X Z, Liu X, Huang S Y, Liu Z Y and Wang S Z. 2025. Dual-branch U-Net with hybrid attention for hyperspectral pansharpening. Journal of Image and Graphics, 30(4):0989-1002(杨勇, 王晓争, 刘轩, 黄淑英, 刘紫阳, 王书昭. 2025. 面向高光谱全色锐化的混合注意力双分支U型网络. 中国图象图形学报, 30(4):0989-1002)[DOI:10.11834/jig.240410]

面向高光谱全色锐化的混合注意力双分支U型网络

杨勇¹, 王晓争², 刘轩², 黄淑英^{3*}, 刘紫阳², 王书昭⁴

1. 天津工业大学计算机科学与技术学院, 天津 300387; 2. 天津工业大学控制科学与工程学院, 天津 300387;
3. 天津工业大学软件学院, 天津 300387; 4. 天津工业大学电子与信息工程学院, 天津 300387

摘要: 目的 高光谱(hyperspectral, HS)全色锐化旨在融合高空间分辨率全色(panchromatic, PAN)图像和低空间分辨率高光谱(low resolution hyperspectral, LRHS)图像,生成高空间分辨率高光谱(high resolution hyperspectral, HRHS)图像。现有全色锐化算法往往忽略PAN和HS图像之间的模态差异,从而造成特征提取不精确,导致融合结果中存在光谱畸变和空间失真。针对这一问题,提出一种基于混合注意力机制的双分支U-Net(dual-branch U-Net based on hybrid attention, DUNet-HA),实现PAN与HS图像的多尺度空间—光谱特征的提取和融合。**方法** 设计混合注意力模块(hybrid attention module, HAM)对网络中的每个尺度特征进行编码。在HAM中,利用通道和空间自注意力模块来增强光谱和空间特征,构建一个双交叉注意力模块(double cross attention module, DCAM),通过学习PAN与HS图像跨模态特征的空间—光谱依赖关系来引导两种特征的重建。与经典的混合Transformer结构相比,设计的DCAM可以通过计算与查询位置无关的交叉注意力权重来实现两种图像特征的校正,在降低模型计算量的同时,提升网络的性能。**结果** 在3个广泛使用的HS图像数据集上与11种方法进行对比,在Pavia center数据集中,相比性能第2的方法hyperRefiner,峰值信噪比(peak signal-to-noise ratio, PSNR)提升了1.10 dB,光谱角制图(spectral angle mapper, SAM)降低了0.40;在Botswana数据集中,PSNR提升了1.29 dB,SAM降低了0.14;在Chikusei数据集中,PSNR提升了0.39 dB,SAM降低了0.12。**结论** 实验结果表明,所提出的DUNet-HA结构能更好地融合空间—光谱信息,显著提升高光谱全色锐化结果图像的质量。

关键词: 高光谱全色锐化;模态差异;混合注意力模块(HAM);双交叉注意力模块(DCAM);Transformer;空间—光谱依赖关系

Dual-branch U-Net with hybrid attention for hyperspectral pansharpening

Yang Yong¹, Wang Xiaozheng², Liu Xuan², Huang Shuying^{3*}, Liu Ziyang², Wang Shuzhao⁴

1. School of Computer Science and Technology, Tiangong University, Tianjin 300387, China; 2. School of Control Science and Engineering, Tiangong University, Tianjin 300387, China; 3. School of Software, Tiangong University, Tianjin 300387, China;
4. School of Electronics and Information Engineering, Tiangong University, Tianjin 300387, China

Abstract: **Objective** Hyperspectral (HS) images have rich spectral information and are obtained by sampling hundreds of contiguous narrow spectral bands using spectral imaging systems; however, they typically have low spatial resolution due to the low energy of narrow spectral bands. Single-band panchromatic (PAN) images from PAN imaging systems provide rich

收稿日期: 2024-07-19; 修回日期: 2024-09-10; 预印本日期: 2024-09-17

* 通信作者: 黄淑英 huangshuying@tiangong.edu.cn

基金项目: 国家自然科学基金项目(62072218, 61862030); 天津市自然科学基金项目(24JCZDJC00130)

Supported by: National Natural Science Foundation of China (62072218, 61862030); Tianjin Municipal Natural Science Foundation (24JCZDJC00130)

spatial information but have low spectral resolution. In some remote sensing applications that require high-resolution hyper-spectral (HRHS) images with high spectral and spatial resolution, neither PAN nor HS images can meet the requirements. HS pansharpening aims to fuse the spatial information from PAN images with the spectral information from HS images to obtain HRHS images. This technology has received considerable attention in the field of remote sensing and is of great significance in various remote sensing tasks, such as military surveillance, environmental monitoring, object identification, and classification. HS pansharpening methods are mainly divided into two categories: traditional methods and deep learning (DL)-based methods. The former can be classified into component substitution-based methods, multiresolution analysis-based methods, Bayesian-based methods, and model-based methods. Although these traditional methods are easy to implement and physically interpretable, they often suffer from spatial and spectral distortion issues due to inappropriate prior assumptions and imprecise manual feature extraction. Owing to their powerful feature learning capability, DL-based methods have been widely applied to HS pansharpening tasks. Although these methods have better performance than traditional methods, spectral and spatial distortions still exist in the fused images because of failure to handle spectral and spatial features differently and the complex mapping relationships between multichannel images. With the introduction of the transformer architecture that can learn the global correlation features of images, some researchers have attempted to improve the performance of HS pansharpening methods by establishing relationships between two modal features using this architecture. However, the application of the Transformer architecture is limited by its high computational cost and low parameter efficiency. To effectively fuse PAN and HS images, this work proposes a dual-branch U-Net network based on hybrid attention (DUNet-HA) for HS pansharpening to achieve multiscale feature fusion. Spatial attention branches, spectral attention branches, and dual-cross attention module branches are constructed at each scale. These branches are used to enhance the spatial and spectral features of PAN and HS images and to achieve complementary cross-modal features. The dual-cross attention module is designed to avoid the complex query matrix computation in the Transformer. **Method** The proposed DUNet-HA includes two U-Net branches, one for PAN images and the other for upsampled HS images, to extract and complement texture and spectral features. At each scale, a hybrid attention module (HAM) is constructed to encode features from both types of images. Each HAM comprises a spatial attention module (Spa-A), a spectral attention module (Spe-A), and a dual-cross attention module (DCAM). The Spa-A and Spe-A enhance the texture and spectral features of PAN and HS images, respectively, and the DCAM corrects and complements these features. The enhanced and corrected features are integrated to obtain the encoded features at each scale. The decoder primarily facilitates feature fusion and reconstruction. In addition, we use a DCAM to capture global contextual information and directly integrate encoded features, decoded features, and corrected complementary features at the same scale to handle high-level spatial and spectral features. The proposed DCAM is a novel cross-attention structure that uses query-independent matrix computation instead of the attention computation in Transformer architecture, reducing computational cost. DCAM maps the cross-feature space of PAN and HS images to guide feature interaction for correction and supplementation. **Result** Extensive experiments are conducted on the following three widely used HS datasets to validate the effectiveness of the proposed DUNet-HA: Pavia center, Botswana, and Chikusei. We compare DUNet-HA with several state-of-the-art (SOTA) methods, including five traditional methods (CNMF, GFPCA, SFIM, GSA, and MTF_GLP_HPM) and six DL-based methods (SSFCNN, HyperPNN, DHP-DARN, DIP-Hyperkite, Hyper-DSNet, and HyperRefiner). To evaluate the performance of all methods, we use five objective indicators, namely, spectral cross correlation (SCC), spectral angle mapper (SAM), root mean square error (RMSE), relative dimensionless global error of synthesis (ERGAS), and peak signal-to-noise ratio (PSNR). Experimental results with a scale factor of 4 demonstrate that the proposed method outperforms other SOTA methods in objective results and visual effects. In particular, the PSNR, SAM, and ERGAS of the proposed method on the Pavia center dataset are improved by 1.10 dB, 0.40, and 0.28, respectively, compared with those of the second-best method, HyperRefiner. Its objective results on the other two datasets also surpass those of HyperRefiner. Visual results indicate that our proposed method is superior in recovering fine-grained spatial textures and spectral details. Ablation studies further demonstrate that the DCAM structure significantly improves the fusion process. **Conclusion** This work proposes a dual-branch interactive U-Net network named DUNet-HA for HS pansharpening. This network extracts and reconstructs spatial and spectral information from PAN and HS images through a parallel dual U-Net structure to achieve accurate fusion results. At each scale of

the network's encoder, a HAM is constructed to enhance the spatial features of PAN images and the spectral features of HS images using spatial attention and spectral attention, respectively. A DCAM is utilized to complement these features, reducing the modality differences between PAN and HS image features and enabling their mutual supplementation for feature interaction guidance. Compared with the classic hybrid Transformer attention structure, the DCAM improves network performance while reducing the number of parameters and computational cost. Extensive experimental results on three widely used HS datasets demonstrate that the proposed DUNet-HA outperforms several SOTA methods in quantitative and qualitative evaluations.

Key words: hyperspectral (HS) pansharpening; modality differences; hybrid attention module (HAM); double cross attention module (DCAM); Transformer; spatial-spectral dependency relationship

0 引言

高光谱(hyperspectral, HS)图像通过使用光谱成像系统密集采样数百个连续窄谱段而获得,可提供丰富的光谱信息。然而,由于窄谱段的能量较低,因此HS图像具有较低的空间分辨率。相比之下,来自PAN(panchromatic)成像系统的单波段全色图像可提供丰富的空间信息,但具有较低的光谱分辨率。在一些需要高空间分辨率高光谱(high resolution hyperspectral, HRHS)图像的遥感任务中PAN和HS图像都无法满足要求。因此,高光谱全色锐化(HS pansharpening)任务旨在将高空间分辨率PAN图像中的空间纹理细节与低分辨率HS(low resolution hyperspectral, LRHS)图像的光谱信息进行融合,获得HRHS图像。该任务在遥感领域得到广泛关注,并在军事监视、环境监测和目标识别分类等众多遥感任务中具有重要意义(聂江涛等,2023;Aburaed等,2023)。

传统的HS pansharpening方法主要分为4个类别(Zhang等,2022):成分替代方法(Aiazzi等,2007;Liao等,2015)、多分辨率分析方法(Aiazzi等,2006;Liu,2000)、基于贝叶斯的方法(Wei等,2015)和基于模型的方法(Yokoya等,2012;Lin和Guo,2023)。尽管传统方法易于实施且在物理上可解释,但由于不合适的先验假设和手动特征提取的不精确性,传统方法容易导致空间和光谱的严重失真(Zhang等,2020)。近年来,卷积神经网络(convolutional neural network, CNN)由于其强大的特征学习能力,已广泛应用于HS pansharpening任务(丁晨等,2024)。其中一些方法(Han等,2018;Bandara等,2022)直接通过拼接PAN和上采样LRHS图像来学习融合图像的

空间和光谱特征。然而,由于未考虑两种图像特征之间的差异,直接从混合图像中提取空间和光谱特征的方法易导致融合得到的HRHS图像产生光谱畸变和空间失真。为减少特征混叠所造成的干扰,具有两个独立分支的HS pansharpening任务网络(Peng等,2023;He等,2019)被提出,分别从PAN和LRHS图像中提取空间细节特征和光谱特征。为进一步增强CNN的表示能力,Li等人(2020)将注意力机制引入到模型中,并表现出良好的特征提取能力。尽管这些方法通过注意力机制学习关注图像的固有属性和先验知识,但缺乏对图像长距离特征的学习,未有效利用LRHS和PAN图像特征间的跨模态相关性,从而导致空间—光谱特征提取不足或不准确的问题(叶珍等,2021)。

近年来,具有学习图像特征全局依赖性能力的Transformer架构被提出,一些研究者尝试使用该架构建立两种模态特征之间的关系来提高HS pansharpening任务的性能。Bandara等人(2022)构建一种HyperTransformer结构用于HS pansharpening任务,其构建的HyperTransformer结构被后来众多研究者借鉴(游雪儿等,2024;Peng等,2023;Wang等,2023;Zhou等,2023),其结构如图1所示。该结构通过学习PAN和上采样LRHS图像特征之间的跨模态相关性来融合两种图像的特征。Peng等人(2023)

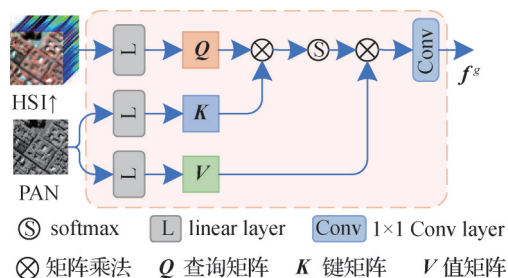


图1 HyperTransformer结构

Fig. 1 HyperTransformer structure

提出一种基于 Transformer 思想的混合注意力模块 S2Block, 通过计算交叉注意力权重矩阵实现 PAN 图像纹理空间细节的有效注入。Zhou 等人(2023)基于多重注意力机制和编码—解码结构实现光谱保持和空间信息融合。

Transformer 结构是基于非局部网络思想而建立的自注意力机制(Vaswani 等, 2017), 非局部相关性计算流程为: 1) 对特征图中每个查询位置建立全局注意力图; 2) 由全局注意力图定义的权重对所有位置特征进行加权求和, 建立查询位置之间的非局部相关性关系。由上述计算流程可以看出, 注意力权重的生成与每个查询位置有关, 每个查询位置都需要建立自己的注意力权重图, 而这也导致 Transformer 结构往往具有较高的计算量。Cao 等人(2019)研究发现, Transformer 结构在视觉任务中建模的全局上下文注意力权重图在图像中的不同查询位置几乎相同, 这也表明该结构对所有位置进行权重图计算存在冗余。

针对上述问题, 本文提出一种基于混合多注意力机制的双分支 U-Net (dual-branch U-Net network based on hybrid attention, DUNet-HA) 用于高光谱全色锐化, 实现不同尺度的特征融合。为更好地提取两种图像的自身特征和交叉特征, 混合注意力模块 (hybrid attention module, HAM) 被构建以实现每个尺度下特征的提取和融合。在 HAM 中, 为缩小两种特征间的模态差异性, 设计一种双交叉注意力模块 (double cross attention module, DCAM), 通过建立 PAN 与 HS 图像上下文特征间的关系, 实现两种特征的校正和补充。

本文的主要贡献总结如下: 1) 提出一种基于 DUNet-HA 的 HS pansharpening 网络, 通过构建并行交互的双分支 U-Net 结构来实现两种图像的多尺度特征提取和融合, 从而获得更加准确的融合结果。2) 在网络的每个尺度编解码层, 构建一个 HAM, 其包含空间注意力分支、光谱注意力分支和 DCAM 分支, 分别用于增强 PAN 与 LRHS 图像的空间和光谱特征, 以及实现跨模态特征的相互补充。3) 基于 Transformer 结构, 提出一种基于查询无关的交叉注意力块 (cross attention block, CAB), 通过建立学习 PAN 与 HS 图像之间跨模态特征的空间—光谱依赖关系来提取有效的空间—光谱特征。基于该结构, 提出具有双分支 CAB 的 DCAM 以实现 PAN 与 HS 图

像之间特征的校正和补充。

1 提出的方法

1.1 双分支 U-Net 网络

考虑到 PAN 与 LRHS 图像之间的模态特征差异, 本文提出 DUNet-HA, 如图 2 所示, 其包含上下两个 U-Net 分支, 分别用于 PAN 和上采样 LRHS 图像中的纹理和光谱特征的提取及校正补充, 并且每个尺度下构建一个 HAM 来实现两种图像特征的编码。HAM 由空间注意力模块 (spatial attention, Spa-A)、光谱注意力模块 (spectral attention, Spe-A) 以及 DCAM 组成。其中, Spa-A 和 Spe-A 用于增强 PAN 与 HS 图像的纹理特征和光谱特征, DCAM 用于实现两种特征的校正和补充, 增强的特征与校正特征被整合, 进而得到每个尺度的编码特征。而解码器主要用于实现特征的融合和重建。因此, 本文只使用 DCAM 捕获全局上下文信息, 直接将同一尺度的编码特征、解码特征以及校正补充特征进行整合, 以更好地处理高层次空间和光谱特征。本文提出的 DCAM 为与查询计算无关的交叉注意力结构, 用于 PAN 与 HS 图像的交叉特征空间映射, 旨在进行特征的交互引导, 实现两种特征的校正补充, 并以残差方式将获得的校正补充特征传递到两个 U-Net 分支的编解码层, 以实现特征的补充。网络具体执行过程如下:

首先, $PAN \in \mathbf{R}^{1 \times H \times W}$ 和上采样的 LRHS 图像 ($LRHS \uparrow \in \mathbf{R}^{C \times H \times W}$, 其中, C 表示 LRHS 图像的光谱通道数, H 和 W 表示图像高度和宽度) 经过一组残差网络进行浅层特征提取, 其操作定义为

$$f_{en_p}^0 = RBs(PAN), f_{en_h}^0 = RBs(LRHS \uparrow) \quad (1)$$

式中, $f_{en_p}^0, f_{en_h}^0 \in \mathbf{R}^{C' \times H \times W}$ 表示 PAN 和上采样 LRHS 图像的浅层特征, C' 表示特征图的通道数, $RBs(\cdot)$ 表示残差网络特征提取操作。提取的浅层特征被分别送入编码器的 HAM 中, 进行特征增强和校正。每个尺度中 HAM 的执行过程为

$$f_{o_p}^i = Spa-A(f_{en_p}^{i-1}) + f_{DCAM}^{pan}(f_{en_p}^{i-1}, f_{en_h}^{i-1}) \quad (2)$$

$$f_{o_h}^i = Spe-A(f_{en_h}^{i-1}) + f_{DCAM}^{hsi}(f_{en_p}^{i-1}, f_{en_h}^{i-1}) \quad (3)$$

$$f_{en_p}^i = D(f_{o_p}^i), f_{en_h}^i = D(f_{o_h}^i) \quad (4)$$

式中, $f_{o_p}^i$ 和 $f_{o_h}^i$ 表示第 i ($i = 1, 2, 3$) 个尺度的编码特征, $Spa-A(\cdot)$ 和 $Spe-A(\cdot)$ 分别表示空间注意力和通道注

注意力计算操作,其具体实现采用Woo等人(2018)提出的CBAM网络中的结构。 $f_{\text{DCAM}}^{\text{pan}}(\cdot)$ 和 $f_{\text{DCAM}}^{\text{hsi}}(\cdot)$ 分别表示DCAM中对PAN和HS图像的特征校正操作。 $D(\cdot)$ 表示2倍下采样操作,即使用 2×2 的卷积核、步幅为2的卷积层进行下采样,并应用深度可分离卷积层来增加特征图的通道数。降采样特征 $f_{\text{en}_p}^i$ 和 $f_{\text{en}_h}^i$ 用于下一尺度HAM的输入。

解码器中,每个尺度利用上采样操作恢复特征图的尺寸,并直接将该尺度下的编码特征、解码特征及来自DCAM的校正特征进行整合,以获得当前尺度的解码特征。其执行过程为

$$f_{\text{de}_p}^i = U(f_{\text{en}_p}^{i+1}) + f_{\text{en}_p}^i + f_{\text{DCAM}}^{\text{pan}}(U(f_{\text{en}_p}^{i+1}), U(f_{\text{en}_h}^{i+1})) \quad (5)$$

$$f_{\text{de}_h}^i = U(f_{\text{en}_h}^{i+1}) + f_{\text{en}_h}^i + f_{\text{DCAM}}^{\text{hsi}}(U(f_{\text{en}_p}^{i+1}), U(f_{\text{en}_h}^{i+1})) \quad (6)$$

式中, $U(\cdot)$ 为2倍上采样操作,得到的融合特征 $f_{\text{de}_p}^i$ 和 $f_{\text{de}_h}^i$ 表示第 i ($i = 1, 2$)个尺度的解码特征。将最后一层解码特征 $f_{\text{de}_p}^1$ 和 $f_{\text{de}_h}^1$ 进行连接,然后通过一组残差块进行特征融合和通道调整,得到最终融合结果HRHS,即

$$\text{HRHS} = \text{RBs}(\text{Cat}(f_{\text{de}_h}^1, f_{\text{de}_p}^1)) \quad (7)$$

式中, $\text{Cat}(\cdot)$ 表示特征连接操作, $\text{HRHS} \in \mathbf{R}^{C \times H \times W}$ 表示最终融合结果。

1.2 HAM结构

在编码器阶段为了在每个尺度下更好地获得融合特征,提出混合注意力模块HAM用于PAN与HS图像的自注意力特征和交叉注意力补充特征的提取,其中Spa-A和Spe-A模块具体操作为

$$f_{\text{en}_p}^{i\text{-sa}} = \text{Sig}(\text{Conv}(\text{cat}(\text{AP}(f_{\text{en}_p}^i), \text{MP}(f_{\text{en}_p}^i)))) \times f_{\text{en}_p}^i \quad (8)$$

$$f_{\text{en}_h}^{i\text{-ca}} = \text{Sig}(\text{Squ}((\text{AP}(f_{\text{en}_h}^i) + \text{MP}(f_{\text{en}_p}^i)))) \times f_{\text{en}_h}^i \quad (9)$$

式中, $f_{\text{en}_p}^{i\text{-sa}}$ 和 $f_{\text{en}_h}^{i\text{-ca}}$ 分别表示经过Spa-A和Spe-A模块得到的增强特征, $\text{Sig}(\cdot)$ 表示sigmoid函数, $\text{AP}(\cdot)$ 和 $\text{MP}(\cdot)$ 分别表示平均池化和最大池化操作, $\text{Conv}(\cdot)$ 表示卷积核为 7×7 大小的卷积操作, $\text{Squ}(\cdot)$ 表示通道的压缩膨胀操作(Woo等,2018),其主要包括两层卷积核大小为 1×1 的卷积操作和一个ReLU(rectified linear unit)激活函数。

虽然Transformer架构具有学习图像特征全局依赖性的能力,但是其具有较大的计算复杂度,并且对所有位置进行权重图的计算存在冗余。因此,为降低网络的计算复杂性,本文提出一种全新的交叉注意力块(cross attention block, CAB),如图2中右下所示,利用基于查询无关的矩阵计算代替Transformer架构中的注意力计算,通过学习PAN和上采样

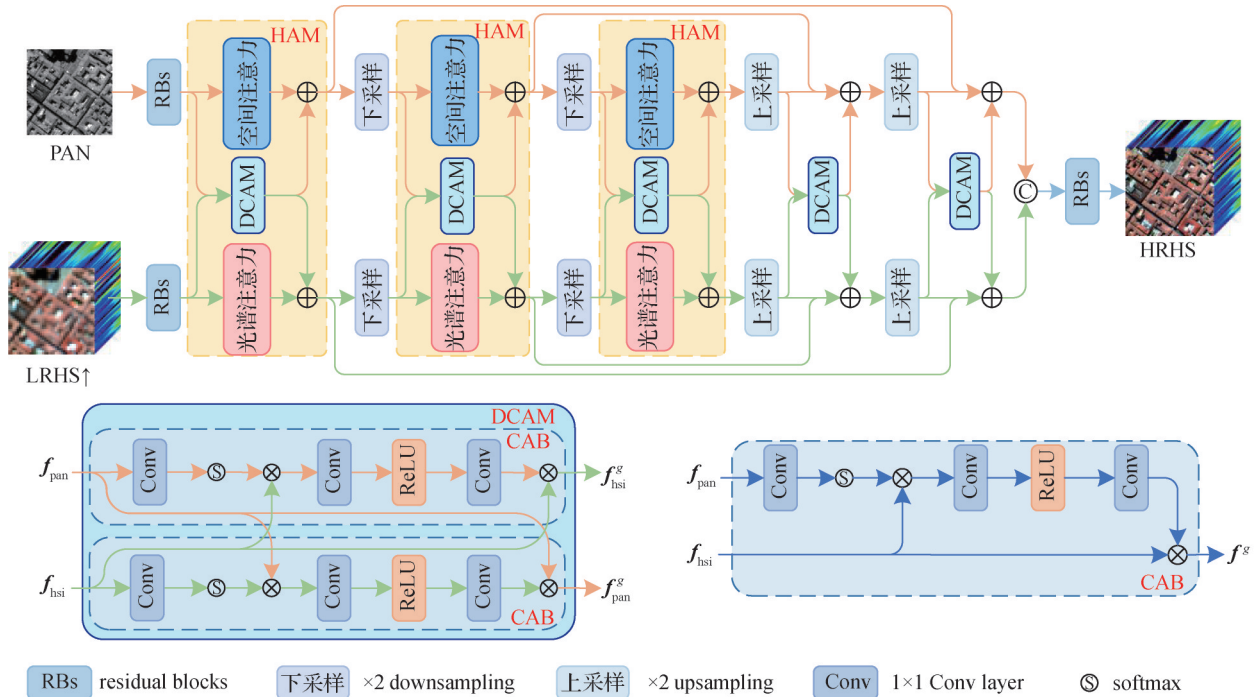


图2 DUNet-HA整体框架

Fig. 2 Overall framework of the DUNet-HA

LRHS图像模态特征的相关性来对图像特征进行校正和补充。基于该结构,设计包含两个CAB分支的DCAM,如图2中左下所示,输出PAN和HS图像的校正补充特征。下面将详细介绍DCAM的构建过程。

首先,针对图1中注意力计算冗余的问题,构建CAB,如图2中右下所示。以PAN的特征来校正HS特征为例,CAB主要分3部分计算:1)构建上下文互相关矩阵。利用 1×1 的卷积获得注意力权重并执行池化操作获得全局上下文特征,通过矩阵相乘获得上下文互相关矩阵。2)利用上下文互相关矩阵进行增强变换。通过卷积层对通道进行压缩膨胀以捕获通道间相关依赖关系。3)实现特征校正操作。将经过增强获得的全局上下文权重与HS的特征图相乘,获得最终引导校正特征。然后,针对PAN与HS图之间进行相互特征引导,构建DCAM。对双分支U-Net结构中每个尺度的编码特征和解码特征都采用一个DCAM来实现两种特征的校正,具体为

$$f_{c,p}^i = \text{Squ}(S(\text{Conv}_{1 \times 1}(f_p^i)) \times f_h^i) \times f_p^i \quad (10)$$

$$f_{c,h}^i = \text{Squ}(S(\text{Conv}_{1 \times 1}(f_h^i)) \times f_p^i) \times f_h^i \quad (11)$$

式中, f_p^i 和 f_h^i 分别表示第*i*个尺度PAN和HS的编码特征或解码特征, $f_{c,p}^i$ 和 $f_{c,h}^i$ 分别表示编码或解码阶段中第*i*个尺度PAN和HS的校正特征。 $\text{Conv}(\cdot)$ 表示卷积核大小为 1×1 的卷积操作, $S(\cdot)$ 表示softmax操作。 $\text{Squ}(\cdot)$ 表示压缩膨胀操作(Woo等,2018)。

1.3 损失函数

已经证明在HS pansharpening任务中 L_1 损失表现优于 L_2 (Zhuo等,2022)。因此重建损失定义为

$$L_{\text{ref}} = \frac{1}{CHW} \|H_{\text{GT}} - I_{\text{HRHS}}\|_1 \quad (12)$$

式中, $H_{\text{GT}} \in \mathbf{R}^{C \times H \times W}$ 表示真实图像(ground truth, GT), I_{HRHS} 表示网络预测的HRHS图像。

2 提出的方法

2.1 数据集与实验设置

为验证所提出的DUNet-HA的有效性,在3个广泛使用的HS数据集(Pavia center(Plaza等,2009)、Botswana(Ungar等,2003)和Chikusei(Yokoya和Iwasaki,2016))上进行了大量实验。将DUNet-HA与一些最先进的方法进行比较,包括CNMF(coupled non-negative matrix factorization)(Yokoya等,2012)、CFPCA

(graphical filter principal component analysis)(Liao等,2015)、SFIM(smoothing filter-based intensity modulation)(Liu,2000)、GSA(gram-schmidt adaptive)(Aiazzi等,2007)、MTF_GLP_HPM(modulation transfer function generalized Laplacian pyramid with high-pass modulation)(Aiazzi等,2006)、SSFCNN(spatial and spectral fusion with CNN)(Han等,2018)、Hyper-PNN(hyperspectral pansharpening via spectrally predictive CNN)(He等,2019)、DHP-DARN(deep hyperspectral prior and dual-attention residual network)(Zheng等,2020)、DIP-Hyperkite(deep hyperspectral prior and novel over-complete network for hyperspectral pansharpening)(Bandara等,2022)、Hyper-DSNet(deep-shallow fusion network for hyperspectral pansharpening)(Zhuo等,2022)和HyperRefiner(the refined hyperspectral pansharpening network)(Zhou等,2023)。为客观评估所有对比方法的性能,采用5个客观指标(Zhuo等,2022;Bandara等,2022),包括光谱交叉相关(spectral cross correlation, SCC)、光谱角映射(spectral angle mapping, SAM)、均方根误差(root mean square error, RMSE)、无量纲全局相对误差(relative dimensionless global error of synthesis, ERGAS)和峰值信噪比(peak signal-to-noise ratio, PSNR)。

在网络训练方面,遵循Wald协议(Wald,2000)对数据集进行处理,并参考Bandara等人(2022)的研究进行相关设置。为公平比较,使用Python 3.9.13和PyTorch 1.13.1在Ubuntu 20.04系统上用单个NVIDIA GeForce RTX A6000 GPU对所提方法进行训练。

2.2 实验结果与分析

表1—表3提供了所有对比方法在3个数据集上融合结果的客观评估结果。其中最高的SCC值表明融合的HSI更接近于实际HSI,且几何失真较小。最小的SAM和ERGAS值进一步表明本文方法引入的光谱失真较少。从表中可看出,与传统方法相比,基于深度学习(deep learning, DL)的方法表现更好,而本文方法在所有评估指标上都优于所有对比方法。

图3—图5展示了不同比较方法在Pavia center、Botswana和Chikusei数据集上获得的融合结果的伪RGB图像。为更清楚地观察融合结果之间的差异,本文计算并显示融合结果与真实值之间的平均绝对

表 1 Pavia center数据集上不同方法客观指标
Table 1 The objective metrics for different methods on the Pavia center dataset

方法	PSNR/dB(↑)	SAM(↓)	SCC(↑)	ERGAS(↓)	RMSE×10 ⁻² (↓)
CNMF	32.024 4	7.236 4	0.946 2	4.819 5	2.709 0
GFPCA	27.020 6	9.777 0	0.857 8	8.266 1	4.798 0
GSA	30.022 0	8.457 9	0.926 3	5.950 8	3.471 0
MTF_GLP_HPM	31.367 2	8.885 9	0.926 7	5.937 9	3.110 0
SFIM	28.276 7	8.650 5	0.896 5	7.431 0	4.116 0
SSFCNN	35.589 0	5.778 7	0.976 1	3.336 2	1.810 4
HyperPNN	37.037 0	5.295 4	0.980 6	2.941 9	1.565 6
DHP-DARN	37.159 2	5.268 5	0.982 4	2.887 6	1.529 9
DIP-Hyperkite	36.958 6	5.363 0	0.980 9	2.936 3	1.566 1
Hyper-DSNet	37.172 5	5.154 1	0.982 2	2.892 9	1.524 5
HyperRefiner	<u>37.719 3</u>	<u>5.046 1</u>	<u>0.985 6</u>	<u>2.774 8</u>	<u>1.426 5</u>
DUNet-HA(本文)	38.821 1	4.646 3	0.987 1	2.496 9	1.286 1

注:加粗、下划线字体表示各列最优、次优结果。“↑”表示值越大越好,“↓”表示值越小越好。

表 2 Botswana数据集上不同方法客观指标
Table 2 The objective metrics for different methods on the Botswana dataset

方法	PSNR/dB(↑)	SAM(↓)	SCC(↑)	ERGAS(↓)	RMSE×10 ⁻² (↓)
CNMF	33.600 6	2.301 8	0.954 9	3.815 1	4.508 0
GFPCA	30.006 5	2.731 4	0.916 2	5.827 9	7.006 0
GSA	36.393 7	2.327 2	0.963 1	2.927 5	3.175 0
MTF_GLP_HPM	36.995 8	2.229 2	0.963 7	2.778 7	3.177 0
SFIM	34.170 4	2.425 9	0.955 7	3.839 8	4.263 0
SSFCNN	42.354 2	1.916 4	0.981 8	4.186 4	1.388 0
HyperPNN	42.586 1	1.990 9	0.977 8	1.825 1	1.395 5
DHP-DARN	41.162 9	1.987 7	0.981 8	2.589 2	1.358 9
DIP-Hyperkite	42.477 3	1.772 8	0.982 5	1.872 4	1.289 7
Hyper-DSNet	43.700 4	1.841 8	0.980 9	1.472 4	1.299 8
HyperRefiner	<u>44.245 5</u>	<u>1.675 1</u>	<u>0.986 4</u>	<u>1.399 7</u>	<u>1.165 5</u>
DUNet-HA(本文)	45.533 2	1.530 3	0.988 6	1.268 7	1.034 7

注:加粗、下划线字体表示各列最优、次优结果。“↑”表示值越大越好,“↓”表示值越小越好。

误差(mean absolute error, MAE)图,并将局部区域放大图显示在结果下方。其图中右侧图例表明 MAE 图中网络预测结果与 GT 的差异程度,蓝色越深,表明网络预测结果与 GT 越接近,黄色越深,表明网络预测结果与 GT 差异越大。从这些图中可以观察到,传统方法的结果在空间和光谱上都表现出更严重的失真,而本文的结果和对应的 MAE 图相比其

他基于深度学习的方法结果包含的残差信息最少。这也表明本文的结果在空间和光谱保真度上更好,所提出的 DUNet-HA 方法优于其他基于深度学习的方法。

为进一步评估基于深度学习 HS pansharpening 的光谱恢复能力,图 6 展示了在 3 个模拟数据集上随机选取 3 个像素位置的光谱差异值曲线。通过将网

表 3 Chikusei数据集上不同方法客观指标
Table 3 The objective metrics for different methods on the Chikusei dataset

方法	PSNR/dB(↑)	SAM(↓)	SCC(↑)	ERGAS(↓)	RMSE×10 ⁻² (↓)
CNMF	35.670 1	3.793 3	0.901 3	6.601 8	1.973 0
GFPCA	32.810 5	4.985 2	0.847 0	9.230 7	2.714 0
GSA	35.432 9	4.221 7	0.913 6	8.009 2	1.921 0
MTF_GLP_HPM	30.685 9	7.709 0	0.705 2	32.004 5	4.544 0
SFIM	29.997 7	6.894 2	0.723 3	25.721 0	4.014 0
SSFCNN	40.822 0	2.482 9	0.971 2	4.353 7	1.018 8
HyperPNN	40.647 1	2.566 8	0.969 9	4.588 2	1.031 8
DHP-DARN	41.041 9	2.421 7	0.974 3	4.261 4	0.983 7
DIP-Hyperkite	41.650 3	2.320 7	0.976 1	4.089 6	0.924 6
Hyper-DSNet	41.699 0	2.331 5	0.976 5	4.015 0	0.919 7
HyperRefiner	<u>42.995 8</u>	<u>2.111 2</u>	<u>0.982 9</u>	<u>3.512 6</u>	<u>0.804 6</u>
DUNet-HA(本文)	43.384 3	1.990 8	0.983 6	3.390 5	0.772 5

注:加粗、下划线字体表示各列最优、次优结果。“↑”表示值越大越好,“↓”表示值越小越好。

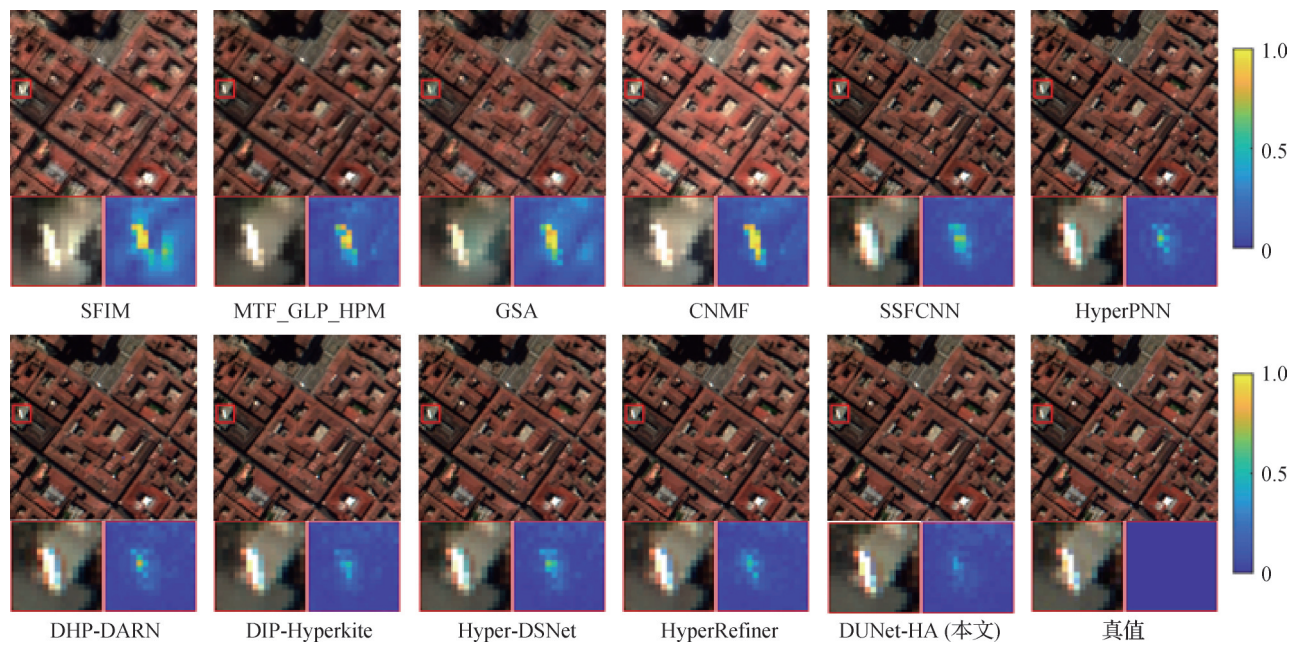


图 3 Pavia center数据集上不同方法融合结果

Fig. 3 Fusion results of different methods on the Pavia center dataset

络融合结果图与参考图对应像素位置的光谱曲线进行相减得到光谱差异值曲线。从图6红色线条可以看出,本文提出方法显示出较低的光谱差异值,得到的融合结果光谱特征更接近参考图,表明所提方法具有良好的光谱保持能力。

此外,为更好地证明所提方法的有效性,本文对融合结果进行了语义分割下游任务处理。在软件ENVI

中使用经典的卫星图像无监督语义分割K-Means算法来评估不同方法的融合结果,其分类类别设置为4,最大迭代次数设置为5,分类结果如图7所示。从图7可以看出,本文提出方法的分类结果最接近参考图GT,由此可进一步证明本文方法的有效性。

2.3 消融实验

进行消融实验以证明网络中HAM、Spa-A&Spe-

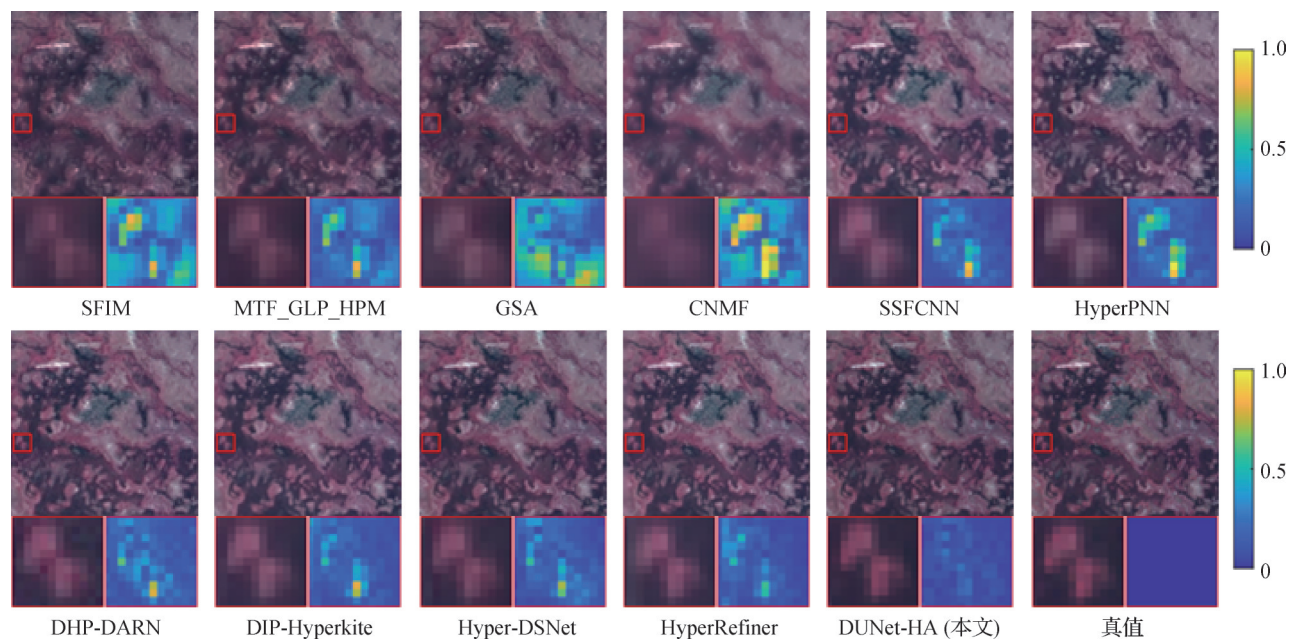


图 4 Botswana 数据集上不同方法融合结果

Fig. 4 Fusion results of different methods on the Botswana dataset

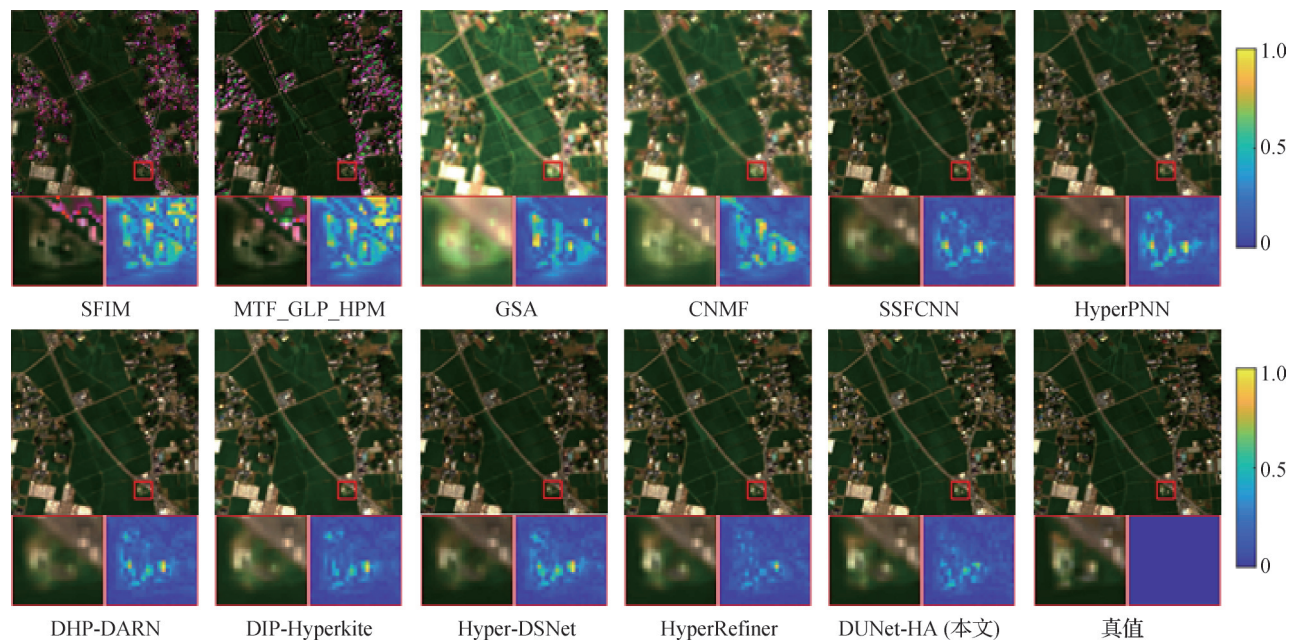


图 5 Chikusei 数据集上不同方法融合结果

Fig. 5 Fusion results of different methods on the Chikusei dataset

A 和 DCAM 模块的有效性。

在第 1 个消融实验中,在 Pavia center 数据集上验证 HAM 的有效性。去除 HAM 模块,使用残差网络替代 Spa-A&Spe-A 注意力结构,同时每尺度下 PAN 与 HS 特征不再经过 DCAM 结构校正补充(即去除 DCAM),而是直接进行交叉相加。其结果如表 4 所示。从结果对比可以看出,去除 HAM 模块时,网

络性能明显下降,表明通过 HAM 结构进行自注意力增强和交叉注意力校正对最终融合结果是有益的,证明所提 HAM 结构的有效性。

为进一步证明 HAM 中 Spa-A&Spe-A 和 DCAM 的有效性,在第 2 个消融实验中,在 Pavia center 数据集上对其进行验证。使用残差网络代替 Spa-A&Spe-A 注意力结构,并通过将每尺度下 PAN 与 HS

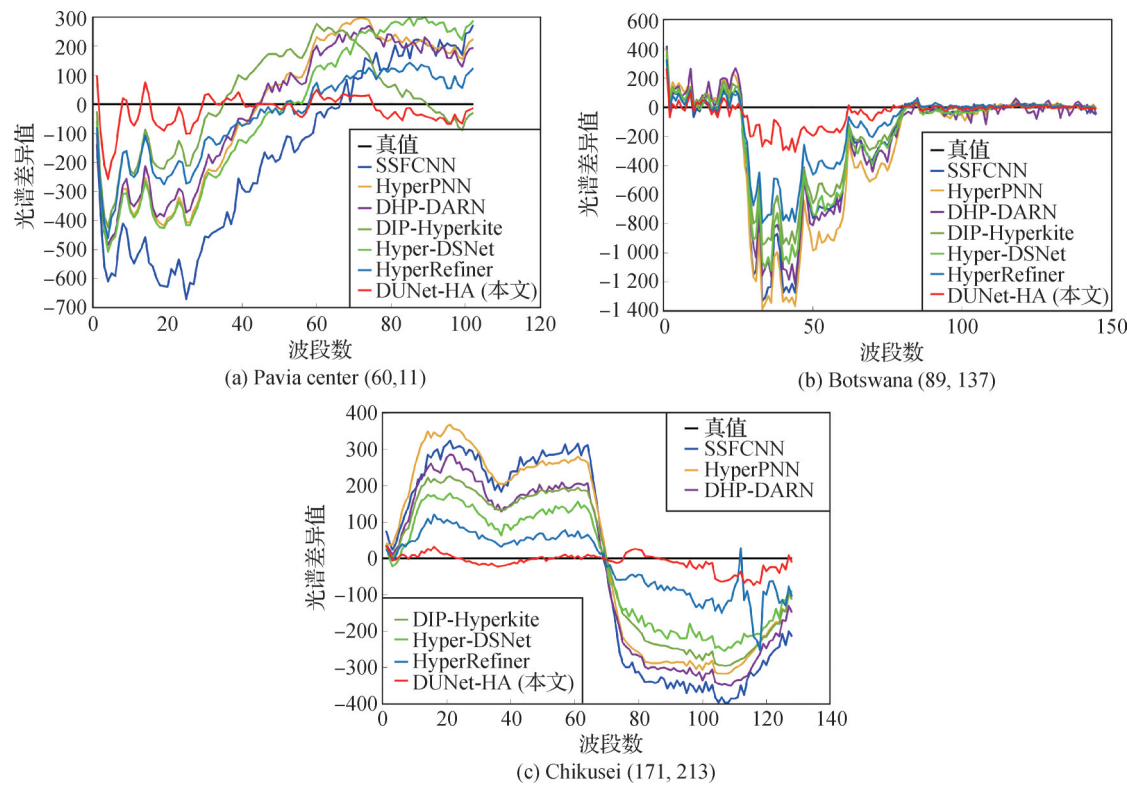


图6 不同方法在Pavia center、Botswana和Chikusei数据集上随机位置光谱差异值曲线
Fig. 6 The spectrum difference value at random pixels for different methods on the Pavia center, Botswana, and Chikusei datasets

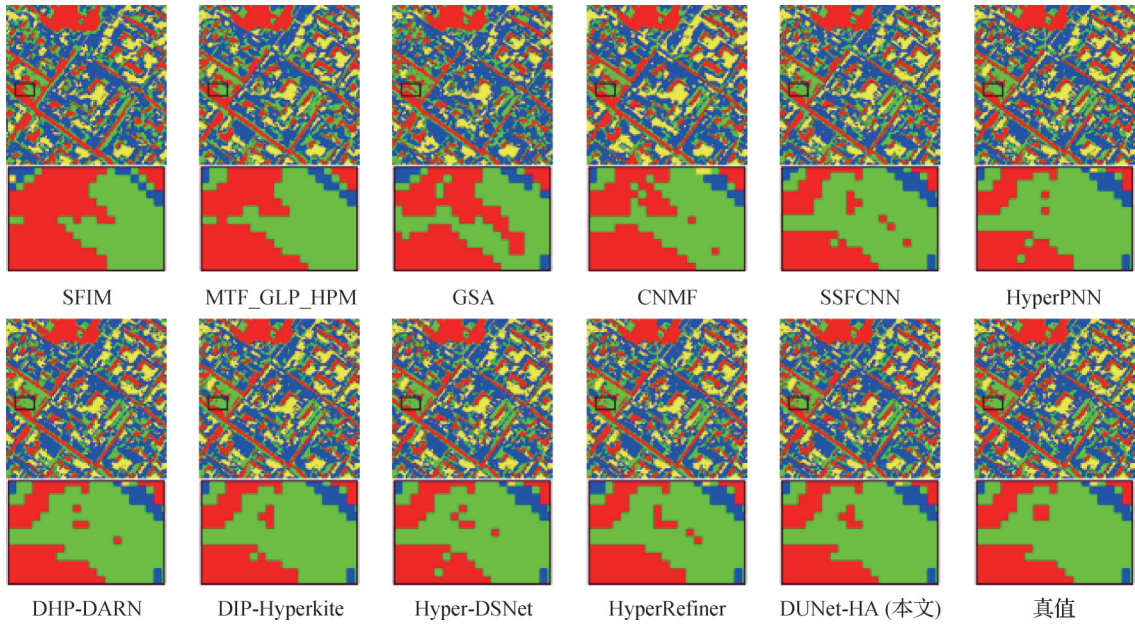


图7 Pavia center数据集上不同融合方法语义分割结果
Fig. 7 Results of semantic segmentation for different methods on the Pavia center dataset

特征直接进行交叉相加来代替DCAM结构,结果如表5所示。从表5中可以看出,在完全去除Spa-A&Spe-A或DCAM时,网络性能都会出现一定的下降。结果表明Spa-A&Spe-A和DCAM对网络性能是有提升效果的。

在第3个消融实验中,在Botswana数据集上,分析所提出的CAB与以往基于Transformer结构的混合注意力模块的性能。采用Bandara等人(2022)提出的混合Transformer结构代替DCAM中的双分支结构,如图8所示。

表 4 HAM 结构消融客观指标
Table 4 The objective metrics of the HAM structure ablation study

HAM	PSNR/dB(↑)	SAM(↓)	RMSE×10 ⁻² (↓)	ERGAS(↓)
×	38.433 8	4.755 2	1.331 9	2.586 7
√	38.821 1	4.646 3	1.286 1	2.496 9

注:加粗字体表示各列最优结果。“√”和“×”分别表示使用和未使用对应模块。“↑”表示值越大越好,“↓”表示值越小越好。

表 5 HAM 内部结构消融客观指标
Table 5 Objective metrics for the ablation study of the internal structure of HAM

DCAM	Spa-A&Spe-A	PSNR/dB(↑)	SAM(↓)	RMSE×10 ⁻² (↓)	ERGAS(↓)
×	×	38.433 8	4.755 2	1.331 9	2.586 7
×	√	38.667 0	4.704 2	1.291 5	2.515 6
√	×	38.561 2	4.721 8	1.314 8	2.563 5
√	√	38.821 1	4.646 3	1.286 1	2.496 9

注:加粗字体表示各列最优结果。“√”和“×”分别表示使用和未使用对应模块。“↑”表示值越大越好,“↓”表示值越小越好。

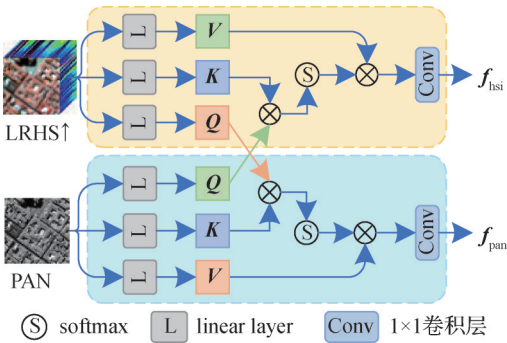


图 8 双分支混合 Transformer 结构

Fig. 8 Dual-branch hybrid Transformer structure

然后,在本文网络中用图 8 结构替换 DCAM,其结果如表 6 所示。结果表明,DCAM 在使用更少的

参数量和计算量的情况下,实现了更好的融合性能,其结果证明该结构能实现对两种特征的有效补充。

为了进一步验证模型的有效性,给出了基于深度学习的对比方法在 Chikusei 数据集上测得的模型参数量和计算复杂度,如表 7 所示。虽然本文模型相比 HyperPNN、DHP-DARN 和 Hyper-DSNet 具有更多的参数量和计算量,但结合表 1 的性能比较,可以看到本文模型获得了远优于这些方法的性能。另外,与较先进的 HyperRefiner 方法相比,本文方法在使用较少参数的情况下取得了更好的结果。综合来看,本文提出的 DUNet-HA 方法具有最优的 HS pan-sharpening 性能。

表 6 DCAM 与双分支混合 Transformer 结构性能比较
Table 6 Performance comparison between DCAM and the dual-branch hybrid Transformer structure

注意力类型	PSNR/dB(↑)	SAM(↓)	ERGAS(↓)	RMSE×10 ⁻² (↓)	Params/M(↓)	FLOPs/G(↓)
双分支混合 Transformer	45.373 9	1.565 9	1.280 2	1.054 6	28.20	142.15
DCAM	45.533 2	1.530 3	1.268 7	1.034 7	16.99	114.79

注:加粗字体表示各列最优结果。“↑”表示值越大越好,“↓”表示值越小越好。Params:parameters;FLOPs:floating point operations per second。

3 结 论

本文提出一种名为 DUNet-HA 的双分支交互的

U-Net 网络用于 HS pansharpening。该网络通过构建并行的双 U-Net 结构来提取和重建 PAN 与 HS 图像的空间和光谱信息,以获得更精确的融合结果。在网络的每个尺度编码层,一个 HAM 被构建来利用空

表7 基于深度学习的方法模型参数量和计算量比较
Table 7 Comparison of parameters and FLOPs of
DL-based methods

方法	Params/M	FLOPs/G
HyperPNN	0.52	17.72
DHP-DARN	0.45	58.12
DIP-Hyperkite	0.97	425.28
Hyper-DSNet	0.53	43.78
HyperRefiner	19.32	170.00
DUNet-HA(本文)	16.99	173.27

注：Params：parameters；FLOPs：floating point operations per second。

间注意力和光谱注意力分别实现PAN与HS图像空间特征和光谱特征的增强,同时利用所构建的DCAM实现两种特征的信息互补。在解码层,仅使用DCAM来对解码特征进行特征校正和补充。DCAM通过构建一个具有双分支CAB的结构以特征交互引导的方式来缩小PAN与HS图像特征的模态差异性,同时实现两种特征的相互补充。相比经典Transformer混合注意力结构,DCAM在降低参数量和计算量的同时,提升了网络性能。在3个广泛使用的HS数据集进行的大量实验结果表明,所提出的DUNet-HA在定量和定性评估中均优于一些SOTA(state-of-the-art)方法。虽然本文所提出的DCAM结构能够学习PAN与HS图像之间跨模态特征的空间—光谱依赖关系,且有效性得到了验证,但针对PAN和HS图像特征的全局自相关性并未进行深入探讨。未来的工作将针对PAN图像的空间特征自相关性和HS图像的光谱特征自相关性问题进行探索,将Transformer结构和CNN相结合并提出更优的模型来实现HS pansharpening任务。

参考文献(References)

Aburaed N, Alkhatib M Q, Marshall S, Zabalza J and Al Ahmad H. 2023. A review of spatial enhancement of hyperspectral remote sensing imaging techniques. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 16: 2275-2300 [DOI: 10.1109/JSTARS.2023.3242048]

Aiazzi B, Alparone L, Baronti S, Garzelli A and Selva M. 2006. MTF-tailored multiscale fusion of high-resolution MS and Pan imagery. *Photogrammetric Engineering & Remote Sensing*, 72(5): 591-596 [DOI: 10.14358/PERS.72.5.591]

Aiazzi B, Baronti S and Selva M. 2007. Improving component substitution pansharpening through multivariate regression of MS + Pan data. *IEEE Transactions on Geoscience and Remote Sensing*, 45(10): 3230-3239 [DOI: 10.1109/TGRS.2007.901007]

Bandara W G C and Patel V M. 2022. HyperTransformer: a textural and spectral feature fusion Transformer for pansharpening//*Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New Orleans, USA: IEEE: 1767-1777 [DOI: 10.1109/CVPR52688.2022.00181]

Bandara W G C, Valanarasu J M J and Patel V M. 2022. Hyperspectral pansharpening based on improved deep image prior and residual reconstruction. *IEEE Transactions on Geoscience and Remote Sensing*, 60: #5520816 [DOI: 10.1109/TGRS.2021.3139292]

Cao Y, Xu J R, Lin S, Wei F Y and Hu H. 2019. GCNet: non-local networks meet squeeze-excitation networks and beyond//*Proceedings of 2019 IEEE/CVF International Conference on Computer Vision*. Seoul, Korea (South): IEEE: 1971-1980 [DOI: 10.1109/ICCVW.2019.00246]

Ding C, Chen J Y, Zheng M M, Zhang L, Wei W and Zhang Y N. 2024. Survey of hyperspectral image change detection method. *Journal of Image and Graphics*, 29(6): 1714-1729 (丁晨, 陈静怡, 郑萌萌, 张磊, 魏巍, 张艳宁. 2024. 高光谱图像变化检测技术研究进展. *中国图象图形学报*, 29(6): 1714-1729) [DOI: 10.11834/jig.240031]

Han X H, Shi B X, Zheng Y Q. 2018. SSF-CNN: spatial and spectral fusion with CNN for hyperspectral image super-resolution//*Proceedings of the 25th IEEE International Conference on Image Processing*. Athens, Greece: IEEE: 2506-2510 [DOI: 10.1109/ICIP.2018.8451142]

He L, Zhu J W, Li J, Plaza A, Chanussot J and Li B. 2019. HyperPNN: hyperspectral pansharpening via spectrally predictive convolutional neural networks. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(8): 3092-3100 [DOI: 10.1109/JSTARS.2019.2917584]

Li J J, Cui R X, Li B, Song R, Li Y S, Dai Y C and Du Q. 2020. Hyperspectral image super-resolution by band attention through adversarial learning. *IEEE Transactions on Geoscience and Remote Sensing*, 58(6): 4304-4318 [DOI: 10.1109/TGRS.2019.2962713]

Liao W Z, H X, Van C F, Gautama S, Philips A P, W, Liu H, Zhu T T, Shimoni M, Moser G. 2015. Processing of multiresolution thermal hyperspectral and digital color data: Outcome of 2014 IEEE GRSS data fusion contest. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(6): 2984-2996 [DOI: 10.1109/JSTARS.2015.2420582]

Lin B H and Guo Y L. 2023. Deep hyperspectral and multispectral image fusion via probabilistic matrix factorization. *IEEE Transac-*

- tions on Geoscience and Remote Sensing, 61: #5503914 [DOI: 10.1109/TGRS.2023.3244992]
- Liu J G. 2000. Smoothing filter-based intensity modulation: a spectral preserve image fusion technique for improving spatial details. *International Journal of Remote Sensing*, 21(18): 3461-3472 [DOI: 10.1080/014311600750037499]
- Nie J T, Zhang L, Wei W, Yan Q S, Ding C, Chen G C and Zhang Y N. 2023. A survey of hyperspectral image super-resolution method. *Journal of Image and Graphics*, 28(6): 1685-1697 (聂江涛, 张磊, 魏巍, 闫庆森, 丁晨, 陈国超, 张艳宁. 2023. 高光谱图像超分辨率重建技术研究进展. *中国图象图形学报*, 28(6): 1685-1697) [DOI: 10.11834/jig.230038]
- Peng S R, Guo C H, Wu X and Deng L J. 2023. U2Net: a general framework with spatial-spectral-integrated double U-Net for image fusion//*Proceedings of the 31st ACM International Conference on Multimedia*. New York, USA: Association for Computing Machinery: 3219-3227 [DOI: 10.1145/3581783.3612084]
- Plaza A, Benediktsson J A, Boardman J W, Brazile J, Bruzzone L, Camps-Valls G, Chanussot J, Fauvel M, Gamba P, Gualtieri A, Marconcini M, Tilton J C and Trianni G. 2009. Recent advances in techniques for hyperspectral image processing. *Remote Sensing of Environment*, 113(S1): S110-S122 [DOI: 10.1016/j.rse.2007.07.028]
- Ungar S G, Pearlman J S, Mendenhall J A and Reuter D. 2003. Overview of the earth observing one (EO-1) mission. *IEEE Transactions on Geoscience and Remote Sensing*, 41(6): 1149-1159 [DOI: 10.1109/TGRS.2003.815999]
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, Kaiser L and Polosukhin I. 2017. Attention is all you need//*Proceedings of the 31st International Conference on Neural Information Processing Systems*. Red Hook, USA: Curran Associates Inc.: 5998-6008 [DOI: 10.5555/3295222.3295349]
- Wald L. 2000. Quality of high resolution synthesised images: is there a simple criterion?//*Proceedings of the 3rd Conference "Fusion of Earth Data: Merging Point Measurements, Raster Maps and Remotely Sensed Images"*. Sophia Antipolis, France: SEE/URISCA: 99-103 [DOI: 10.1109/IGARSS.2000.859923]
- Wang X H, Wang X Y, Song R X, Zhao X Y and Zhao K Y. 2023. MCT-Net: multi-hierarchical cross Transformer for hyperspectral and multispectral image fusion. *Knowledge-Based Systems*, 264: #110362 [DOI: 10.1016/j.knsys.2023.110362]
- Wei Q, Bioucas-Dias J, Dobigeon N and Tournet J Y. 2015. Hyperspectral and multispectral image fusion based on a sparse representation. *IEEE Transactions on Geoscience and Remote Sensing*, 53(7): 3658-3668 [DOI: 10.1109/TGRS.2014.2381272]
- Woo S, Park J, Lee J Y and Kweon I S. 2018. CBAM: convolutional block attention module//*Proceedings of the 15th European Conference on Computer Vision*. Munich, Germany: Springer: 3-19 [DOI: 10.1007/978-3-030-01234-2_1]
- Ye Z, Bai L and He M Y. 2021. Review of spatial-spectral feature extraction for hyperspectral image. *Journal of Image and Graphics*, 26(8): 1737-1763 (叶珍, 白璘, 何明一. 2021. 高光谱图像空谱特征提取综述. *中国图象图形学报*, 26(8): 1737-1763) [DOI: 10.11834/jig.210198]
- Yokoya N and Iwasaki A. 2016. Airborne hyperspectral data over Chikusei [EB/OL]. [2024-07-04]. https://www.researchgate.net/publication/304013716_Airborne_hyperspectral_data_over_Chikusei
- Yokoya N, Yairi T and Iwasaki A. 2012. Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 50(2): 528-537 [DOI: 10.1109/TGRS.2011.2161320]
- You X E, Su Y C, Jiang M Y, Li P F, Liu D S and Bai J Y. 2024. Deep embedded Transformer network with spatial-spectral information for unmixing of hyperspectral remote sensing images. *Journal of Image and Graphics*, 29(8): 2220-2235 (游雪儿, 苏远超, 蒋梦莹, 李鹏飞, 刘东升, 白晋颖. 2024. 深度嵌套式Transformer网络的高光谱图像空谱解混方法. *中国图象图形学报*, 29(8): 2220-2235) [DOI: 10.11834/jig.230393]
- Zhang B, Wu Y F, Zhao B Y, Chanussot J, Hong D F, Yao J and Gao L R. 2022. Progress and challenges in intelligent remote sensing satellite systems. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15: 1814-1822 [DOI: 10.1109/JSTARS.2022.3148139]
- Zhang L, Nie J T, Wei W, Zhang Y N, Liao S C and Shao L. 2020. Unsupervised adaptation learning for hyperspectral imagery super-resolution//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, USA: IEEE: 3073-3082 [DOI: 10.1109/CVPR42600.2020.00314]
- Zheng Y X, Li J J, Li Y S, Guo J, Wu X Y and Chanussot J. 2020. Hyperspectral pansharpening using deep prior and dual attention residual network. *IEEE Transactions on Geoscience and Remote Sensing*, 58(11): 8059-8076 [DOI: 10.1109/TGRS.2020.2986313]
- Zhou B, Zhang X F, Chen X, Ren M and Feng Z Y. 2023. HyperRefiner: a refined hyperspectral pansharpening network based on the autoencoder and self-attention. *International Journal of Digital Earth*, 16(1): 3268-3294 [DOI: 10.1080/17538947.2023.2246944]
- Zhuo Y W, Zhang T J, Hu J F, Dou H X, Huang T Z and Deng L J. 2022. A deep-shallow fusion network with multidetail extractor and spectral attention for hyperspectral pansharpening. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15: 7539-7555 [DOI: 10.1109/JSTARS.2022.3202866]

作者简介

杨勇,男,教授,博士生导师,主要研究方向为遥感图像处理、计算机视觉与机器学习。E-mail: yangyong@tiangong.edu.cn
黄淑英,通信作者,女,教授,主要研究方向为遥感图像处理、计算机视觉与机器学习。
E-mail: huangshuying@tiangong.edu.cn
王晓争,男,博士研究生,主要研究方向为遥感图像处理与机

器学习。E-mail: xiaozhengwang95@163.com
刘轩,男,硕士研究生,主要研究方向为遥感图像处理与机器学习。E-mail: lx1403963990@outlook.com
刘紫阳,男,博士研究生,主要研究方向为遥感图像处理与机器学习。E-mail: qduliuzy@163.com
王书昭,男,博士研究生,主要研究方向为图像处理与机器学习。E-mail: 794446715@qq.com