光电技术及应用

DOI：10.16818/j.issn1001－5868.2024040104

融合跨阶段局部网络与注意力的遥感图像锐化方法

刘蕙绮，陈应霞*

（长江大学 计算机科学学院，湖北 荆州 434023）

摘　要： 针对遥感图像全色锐化中存在光谱分布不均、空间细节缺失的问题，提出一种融合跨阶段局部网络（CSPNet）与无参数注意力（SimAM）的遥感图像锐化方法。首先，在主干结构中引入 CSPNet，利用普通卷积加跳连的方式替代特征提取中的残差块，以缓解梯度冗余，提升模型学习力。其次，添加 SimAM 块，直接从特征中推导出三维权重，而后反向优化提取到的特征，使得模型能提取到更深层次的特征信息。最后，设计一种可学习作差参数来控制相减权重，以便突出融合图像的边缘信息。实验结果表明，所提方法不仅能改善模型的梯度冗余，还能进一步提升融合图像的空谱分辨率。

关键词： 全色锐化；跨阶段局部网络；无参数注意力；可学习作差参数

中图分类号： TP751　**文章编号：** 1001－5868(2024)05－0837－10

Remote Sensing Image Pansharpening Algorithm Fusing Cross Stage Partial Network and Attention

LIU Huiqi，CHEN Yingxia

（School of Computer Science，Yangtze University，Jingzhou 434023，CHN）

Abstract： A remote sensing image sharpening method is proposed that combines cross stage local network (CSPNet) and parameter free attention (SimAM) to address the issues of uneven spectral distribution and missing spatial details in pansharpening of remote sensing images. Firstly, CSPNet is introduced into the backbone structure to replace the residual blocks in feature extraction with ordinary convolution and skip connections, in order to alleviate gradient redundancy and improve model learning power. Secondly, add SimAM blocks to directly derive 3D weights from the features, and then reverse optimize the extracted features to enable the model to extract deeper level feature information. Finally, design a learnable subtraction parameter to control the subtraction weights, in order to highlight the edge information of the fused image. The experimental results show that the proposed method can not only improve the gradient redundancy of the model, but also further enhance the spatial spectral resolution of the fused image.

Key words： pansharpening; cross stage local network; parameter free attention; parameter for learnable differencing

0　引言

近年来，全色锐化被广泛应用于遥感图像处理中，其目标是将全色图像（Panchromatic，PAN）的高分辨率空间信息与多光谱图像（Multispectral，MS）的高分辨率光谱信息进行融合，以便得到更高空间

收稿日期：2024－04－01.

基金项目：2022 年中国高校产学研创新基金项目（2021RYC06002）；2022 年度湖北省自然资源厅自然资源科技项目（MB1644959/2021-70118）；2022 年湖北省教育厅科研计划项目（B2022040）；上海市多维度信息处理重点实验室开放课题基金资助项目（MIP20222）；中央高校基本科研业务费专项资金资助（MIP20222）.

* **通信作者：**陈应霞

与光谱分辨率的融合图像[1]。全色锐化发展到现在,已提出了众多方法。其中,最为典型的是细节注入法,它大体分为以下两种:

(1)成分替换(Component Substitution,CS)法[2],其原理是在给定域中用PAN图像替换MS图像的分量,代表方法有主成分分析(Principal Component Analysis,PCA)法[3]、亮度-色度-饱和度(Intensity-Hue-Saturation,IHS)变换[4]、施密特正交(Gram-Schmidt,GS)法[5]等。此类方法容易实现,且获得的融合图像空间结构清晰,但存在较为严重的光谱失真。

(2)多分辨率分析(Multi-Resolution Analysis,MRA)法[6],该类方法的思想是将低分辨率多光谱(Low-Spatial Resolution Multi-Spectral,LRMS)图像进行分解,然后利用PAN图像来补充其空间信息,通常会涉及到多尺度特征的细节提取与整合。常用方法有非下采样轮廓波(Nonsubsampled Contourlet,NSCT)变换[7]、拉普拉斯金字塔(Laplacian Pyramid,LP)[8]等。此类方法能较大程度保留图像的光谱信息,但对PAN图和上采样后的MS图间的配准要求很高,这通常很难做到的,易出现空间失真、计算量过大的问题。

总体而言,传统细节注入法通常以假设线性注入模型为前提,但这并不能很好适应传感器的光谱响应特点。尽管此类方法可以归纳为一个细节注入框架,但融合结果在空间和光谱上往往存在不同程度的差异。另外,传统细节注入法虽易实现,但无法对更为复杂的图像进行有效处理。随着深度学习的发展,卷积神经网络(Convolutional Neural Networks,CNN)因其良好的非线性映射性能改进了这一缺点。Masi等[9]提出了一种名为PNN的锐化方法,通过将预插值LRMS图像与PAN图像在输入层堆叠,然后,利用CNN去学习输入图像和锐化后的HRMS图像之间的关系。尽管该算法能在真实遥感数据上获得较好性能,但其未采用可加速收敛的跳跃连接技术,以至在融合效果及收敛性上表现欠佳[10]。针对这一问题,He等[11]提出两种基于图像细节注入的全色锐化方法(DiCNN1和DiCNN2),该方法能有效学习MS图像的细节信息并将其注入到LRMS图像中。而与PNN的不同之处在于它加入了跳跃连接,能够有效缓解梯度爆炸并加速收敛,但它的网络结构过于简单且参数量少,导致不能很好地提取图像的特征,所以融合效果仍未达到理想状态。于是,Deng等[12]将传统锐化框架与CNN相结合,提出一种新的全色锐化网络FusionNet。该网络采用深层次结构预估PAN图和上采样MS图细节的非线性注入模型,将PAN图与上采样MS图差分,直接为网络提供细节。虽然该方法在融合效果上得到了大幅度的提升,但其结构中残差连接较多,在一定程度上存在梯度冗余的问题。

为解决上述问题,进一步提升锐化效果,本文结合FusionNet模型,提出一种融合跨阶段局部网络与无参数注意力的遥感图像锐化方法,主要研究内容如下:(1)主干网络改进;(2)注意力模块;(3)可学习作差。首先,引入CSPNet网络,利用普通卷积加跳连的方式改进主干结构,缓解模型中梯度信息重复计算的问题,降低不必要的能耗;其次,在网络中设计Attention BasicBlock块,使得网络能够提取到更深层次的特征信息,提升融合图像的空谱分辨率;最后,优化作差方式,设计一种可学习作差参数控制相减权重,以便于更好地凸显融合图像的边缘信息。

1 提出方法

1.1 技术路线

图1所示为本文方法的技术路线图,具体实现路线如下:

(1)首先将获取的原始PAN图进行通道复制,接着对MS图进行上采样,使得大小与PAN图保持一致,然后将上采样后的MS图与PAN图利用可学习作差的方式提取图像的细节信息,并将其反馈给网络,从而对网络中图像的空间信息和潜在光谱信息进行最大化的保留,以达到更好突显融合图像边缘信息的目的。

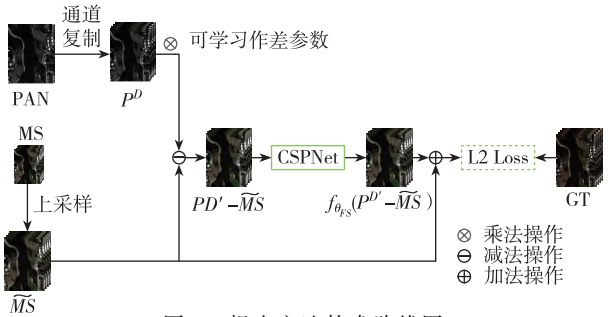


图1 提出方法技术路线图

(2)将步骤(1)提取到的细节信息输入到

CSPNet 网络中，并在该结构中添加 SimAM 模块。

(3)最后将网络的输出与上采样后的 MS 图像相加，采用 L2 损失函数进行网络的训练后获得最终的融合图像。

1.2　主干网络改进

FusionNet 结构相对简单直观，但残差连接过多，存在一定的梯度信息冗余。基于此，本节借鉴 Wang 等[13]提出的 CSPNet 模型的思想，在网络中设计一个含有注意力的模块，即 Attention BasicBlock 块，以解决融合过程中存在的梯度冗余问题。

网络结构如图 2 所示，该网络首先将通道复制得到的 P^D 与上采样得到的$\widehat{MS}$ 经可学习作差参数进行特征提取，接着将提取到的特征图$(P^{D'}-\widehat{MS})$设为$(P^{D'}-\widehat{MS})_0$，再切分为 Part 1：$(P^{D'}-\widehat{MS})'_0$和 Part 2：$(P^{D'}-\widehat{MS})''_0$ 两部分，其中 Part 1 输入$(P^{D'}-\widehat{MS})'_0$ 和$(P^{D'}-\widehat{MS})_T$ 的拼接值，连接到过渡层进行卷积操作后输出$(P^{D'}-\widehat{MS})_Z$ 的值；Part 2 输入$[(P^{D'}-\widehat{MS})''_0,(P^{D'}-\widehat{MS})_1,(P^{D'}-\widehat{MS})_2,\cdots,(P^{D'}-\widehat{MS})_{K-1}]$，经 Attention BasicBlock 块计算后输出$(P^{D'}-\widehat{MS})_T$ 的值，最后再将这两部分连接后输出。利用这种方式，有效增强了网络中信息的流动性，使得模型的泛化力大大提升。

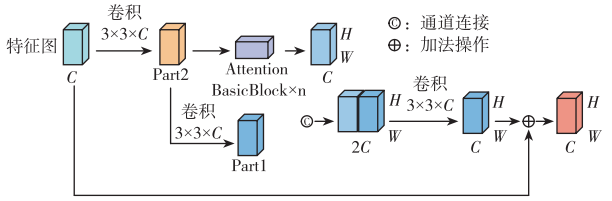


图 2　具有多个 Attention BasicBlock 块的 CSPNet 网络结构

由以上内容，可得到前向传播公式为：

$$(P^{D'}-\widehat{MS})_K=\omega_K*[(P^{D'}-\widehat{MS})''_0,(P^{D'}-\widehat{MS})_1,(P^{D'}-\widehat{MS})_2,\cdots,(P^{D'}-\widehat{MS})_{K-1}] \tag{1}$$

$$(P^{D'}-\widehat{MS})_T=\omega_T*[(P^{D'}-\widehat{MS})''_0,(P^{D'}-\widehat{MS})_1,(P^{D'}-\widehat{MS})_2,\cdots,(P^{D'}-\widehat{MS})_K] \tag{2}$$

$$(P^{D'}-\widehat{MS})_Z=\omega_Z*[(P^{D'}-\widehat{MS})'_0,(P^{D'}-\widehat{MS})_T] \tag{3}$$

式中，$*$ 表示卷积，ω_K 为第 K 个 Attention BasicBlock 层的权重，$(P^{D'}-\widehat{MS})_K$ 为第 K 个 Attention BasicBlock 块的输出，转化后表示为$(P^{D'}-\widehat{MS})_T$ 和$(P^{D'}-\widehat{MS})_Z$。

接着，得到反向权重更新公式为：

$$\omega'_K=f(\omega_K,g''_0,g_1,g_2,\cdots,g_{K-1}) \tag{4}$$

$$\omega'_T=f(\omega_T,g''_0,g_1,g_2,\cdots,g_K) \tag{5}$$

$$\omega'_Z=f(\omega_Z,g'_0,g_T) \tag{6}$$

式中，f 表示权重更新函数；g_K 为第 K 个 Attention BasicBlock 层的梯度。权重 ω'_T 和 ω'_Z 经不同梯度分流，既保持了原网络特征的复用特性，又利用梯度流的截断来避免冗余。

总体来说，借鉴残差的思想，在具体操作中进行梯度分流，并设计 Attention BasicBlock 块。但与残差不同的是，模块中仅以普通卷积加跳连的方式来进行融合。这样设计既缓解了冗余问题，又减少了内存占用率，达到进一步提升融合图像空间与光谱分辨率的目的。

1.3　注意力模块

图 3 为添加 Attention Map 块的 Attention BasicBlock 结构图。在具体的融合过程中，由于输入 LMS 是光谱信息各不相同的多通道图，PAN 是空间信息丰富的单通道图。为了学习到更具区分性的 PAN 图像的空间特征和 LMS 图像的光谱特征，本节使用 Yang 等[14]提出的 SimAM 注意力。具体来说，以神经科学理论为灵感，提出了一个能量函数，对其进行优化，以求解析解的方式使计算权重的速度更快。

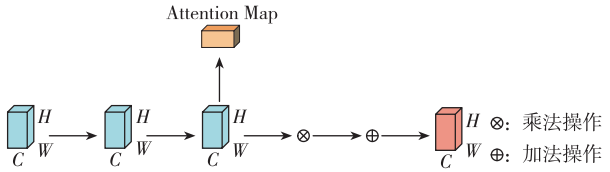


图 3　添加 Attention Map 块的 Attention BasicBlock 网络结构

首先设经可学习作差后提取到的 Part 2 分支特征图$(P^{D'}-\widehat{MS})''_0$ 为 X，且 $X\in\mathbb{R}^{C\times H\times W}$，则具体的能量函数表示为：

$$e_t(\omega_t,b_t,y,x_i)=(y_t-\hat{t})^2+\frac{1}{M-1}\sum_{i=1}^{M-1}(y_o-\hat{x}_i)^2 \tag{7}$$

式中，$\hat{t}=\omega_t t+b_t$，$\hat{x}_i=\omega_t x_i+b_t$ 为 t 与 x_i 的线性变换，而 t 和 x_i 分别表示输入特征图 X 通道中目标神经元数和其他神经元数，ω_t，b_t 是转换的权值和偏差，i 为空间维度上的指数，M 为单个通道上神经元的数目，表示为 $M=H\times W$。在式(7)中，包含的

量都为标量，且 $y_t=\hat{t}$，当 $x_i=y_0$ 的时候取最小值。接着，为计算简便，使用二值化并加入正则项，可到最终的能量函数为：

$$e_t(\omega_t,b_t,y,x_i)=\frac{1}{M-1}\sum_{i=1}^{M-1}(-1-(\omega_t x_i+b_t))^2+(1-(\omega_t t+b_t))^2+\lambda\omega_t^2 \tag{8}$$

由此，得到式(8)的解析解为：

$$\omega_t=-\frac{2(t-\mu_t)}{(t-\mu_t)^2+2\sigma_t^2+2\lambda} \tag{9}$$

$$b_t=-\frac{1}{2}(t+\mu_t)\omega_t \tag{10}$$

式中，$\mu_t=\frac{1}{M-1}\sum_{i=1}^{M-1}x_i$，$\sigma_t^2\ \frac{1}{M-1}\sum_{i}^{M-1}(x_i-\mu_t)^2$ 表示通道中不包含目标神经元的均值和方差。

最后，得到最小能量函数求得融合图中神经元的权值，以此验证目标神经元与其他神经元的线性可分性，公式为：

$$e_t^*=\frac{4(\hat{\sigma}^2+\lambda)}{(t-\hat{\mu})^2+2\hat{\sigma}^2+2\lambda} \tag{11}$$

式中，e_t^* 越小，代表融合图中目标神经元与其他神经元相差越多，重要性越高。$\frac{1}{e_t^*}$表示每个神经元的重要性。接着再对特征进行增强，如下式：

$$\widetilde{X}=\mathrm{sigmoid}\left(\frac{1}{E}\right)\odot X \tag{12}$$

式中，E 表示融合图中神经元 e_t^* 值的集合，使用 sigmoid 函数是为了防止 E 值过大。

总体来说，在网络中添加 SimAM 块，利用能量函数解析解验证特征提取的重要程度，并增强包含更多有用信息的神经元输出，使得融合中模型具有更强的特征表示力，大大提升融合图像的准确性。

1.4 可学习作差

FusionNet 模型使用 $P^D-\widehat{MS}$ 作差来提取图像的细节信息，而直接相减的作差方法有时不能很好突出图像的边缘信息。基于此，本节设计一个可学习作差参数来控制相减的权重。具体来说，定义一个用来控制权重倍率的可学习参数 R，再将 R 乘以通道复制后的 P^D，得到 $P^{D'}$。公式为：

$$P^{D'}=R\times P^D \tag{13}$$

由此，进一步表示为：

$$\widehat{MS}=\widehat{MS}+f_{\Theta_{\mathrm{FS}}}(P^{D'}-\widehat{MS}) \tag{14}$$

式中，$\widehat{MS}$ 和 $\widehat{MS}$ 是由 $\widehat{MS}_i$，$i=1,2,\cdots,B$ 和 $\widehat{MS}_i$，$i=1,2,\cdots,B$ 波段叠加得到的，$\widehat{MS}_i$ 为高空间分辨率 MS 的第 i 个波段，$\widehat{MS}_i$ 是低空间分辨率 MS 上采样后的第 i 个波段，$f_{\Theta_{\mathrm{FS}}}$ 为 FS 的非线性映射函数。

总体来说，实现较简单，先设计一个可学习作差参数控制权重，该参数初始设置为 1，后面会随着学习的变化而不断发生变化。以这种方式提取图像的细节信息，再反馈给网络，使得融合图像在保持边界信息更清晰的同时，最大程度的保留图像的空间信息和光谱信息。

1.5 损失函数

在模型训练的时候，会对每一个训练样本进行损失函数的计算。先是利用正向传输使得模型产生预测值，接着用损失函数来计算预测值与真实值间的误差，然后通过数据的反向传输去调整参数，最后使得预测值更加接近真实值，模型预测更准确。

在目前的锐化算法中，大部分都使用 L2 作为损失函数来优化整个网络。因为 L2 损失函数可以惩罚较大的异常项，所以本文使用 L2 损失函数$\widehat{MS}$对图像与 GT 的相似性进行约束，使其尽可能接近真实值，具体表示如下：

$$Loss(\Theta_{\mathrm{FS}})=\frac{1}{n}\sum_{k=1}^{n}\left\|\widehat{MS}_{\{k\}}+f_{\Theta_{\mathrm{FS}}}(P_{\{k\}}^{D'}-\widehat{MS}_{\{k\}}-GT_{\{k\}})\right\|_F^2 \tag{15}$$

式中，n 表示训练样本数，$\|\cdot\|_F$ 为 Frobenius 范数，$GT_{\{k\}}$ 表示从 GT 参考图中提取的第 k 个样本。通过最小化损失函数，网络将强制自动学习参数 Θ_{FS} 的最优映射。

1.6 提出算法步骤

具体融合步骤表述如下：

(1)首先对获取的原始 PAN 图像进行通道复制操作，接着将 MS 图像上采样到 PAN 图大小，得到 LMS；

(2)将步骤(1)得到的 PAN 和 LMS 作为输入，经可学习作差参数 R 提取图像的细节信息；

(3)将提取到的信息输送到两支路 CSPNet 网络中：支路 1 仅一层卷积提取特征信息；支路 2 利用式(11)求得融合图中目标神经元的权值，再经 n 个 Attention BasicBlock 块计算，划分不同的梯度流；

(4)再将两支路输送到新卷积中进行特征融合；

(5)接着将融合结果与输入特征图经跳跃连接后输出；

(6)最后输出结果经非线性映射后，再利用跳跃连接与 LMS 融合，以得到最终的融合图像。

2　实验结果与分析

2.1　实验准备与数据集

为证明所提算法的有效性，将实验结果与一些具有代表性的方法进行对比，如传统方法：GSA[15]，BDSD-PC[16]和 MTF-GLP-FS[17]，基于深度学习的方法：PNN[9]，BDPN[10]，MSDCNN[10]，DRPNN[10]，DiCNN1[11]和 FusionNet[12]。文中用到的客观评价指标有 SAM[18]，ERGAS[18]，SF[18]，RMSE[18]，SID[19]，Q4[20]，RASE[18]和 SCC[18]。其中，若 SF，Q4 和 SCC 的指标数值高，则证明得到的融合图像表现较好，SAM，ERGAS，RASE，RMSE 和 SID 的指标越低，表示融合图像质量越好。文中在同样设备下对所用方法展开研究，深度学习模型均采用 Pytorch 架构，并在 GeForce RTX 4060 上进行训练，测试使用 Matlab (2019 a)。

实验分别在 QuickBird 和 N_NWPU 数据集上进行，其中 N_NWPU 数据集是根据 Wald 协议[21]，以 NWPU-RESISC45 数据集为基础设计的。设计过程为：先对 NWPU-RESISC45 数据集中的图像进行切割，得到大小为 256×256，64×64×4 的原始 PAN 和 MS 图像；接着对 PAN 和 MS 进行滤波操作，然后再按照采样率 4 对其进行下采样，以得到低分辨率的 PAN 和 MS 图像；最后利用 23-tap polynomial[22]插值上采样得到 LMS 图像。构建的训练数据集包括参考图(原始 MS)，下采样后的 PAN 和 MS，以及上采样得到的 LMS 图像，其中模拟每幅尺寸大小为 64×64×4，64×64，16×16×4 和 64×64×4 的 10 400 个图像对，随机按照 8∶2 划分训练集和验证集。(下载地址：(1) PanCollection for Survey Paper (liangjiandeng.github.io)；(2) Gong CHENG (gcheng-nwpu.github.io))。

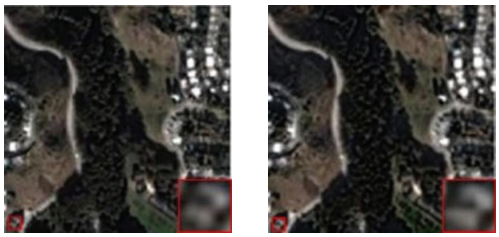
2.2　对比实验

2.2.1　QuickBird 数据集

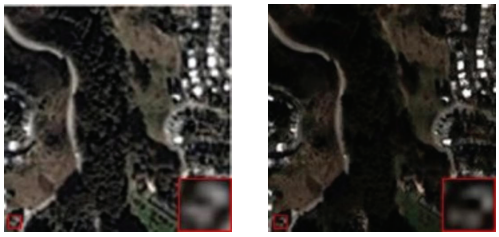
图 4 为不同算法在 QuickBird 数据集下的融合图。从主观效果看，传统 GSA(图 4(c))、BDSD-PC(图 4(d))和 MTF-GLP-FS(图 4(e))算法融合结果模糊，在空间信息方面过于锐化，且和参考图像比较，局部有不自然的颜色。PNN(图 4(f))算法在空间和光谱上相较于传统方法有明显提升，但在河流边缘上，清晰度不如参考图像。BDPN(图 4(g))存在一定程度的光谱失真，融合图像整体颜色比参考图像暗。MSDCNN(图 4(h))算法在光谱上严重失真，在空间质量方面的表现也较差，融合图像平滑且细节信息损失较多。DRPNN(图 4(i))、DiCNN1(图 4(j))和 FusionNet(图 4(k))算法融合图在空间上提升较大，但比较局部放大区域，发现有轻微的光谱失真。本文算法(图 4(l))在空间信息和光谱特征上都比其他方法略有优势，且在细节信息的保持方面和参考图像比较一致。



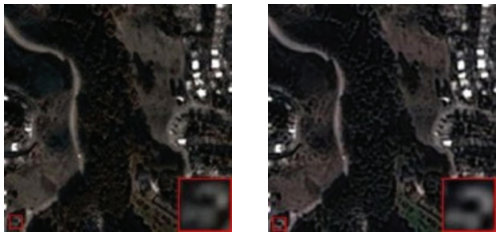
(a) LMS　(b) 参考图



(c) GSA　(d) BDSD-PC



(e) MTF-GLP-FS　(f) PNN



(g) BDPN　(h) MSDCNN



(i) DRPNN　(j) DiCNN1



(k) FusionNet



(l) 本文算法

图 4 测试图 1 在不同算法下的锐化结果

表 1 为测试图 1 的客观评价指标。从表 1 可知，传统 GSA，BDSD-PC 和 MTF-GLP-FS 算法在 SAM，SID 和 RASE 指标上的值较大，在 SF 和 SCC 指标上的值最小，表明这三种算法融合图像在空间细节和光谱保真方面呈现的效果较差，融合效果不理想。BDPN 和 MSDCNN 在各指标上均优于传统方法，但提升不大，说明融合结果仍存在光谱失真和空间损失。PNN，DRPNN，DiCNN1 和 FusionNet 指标表现的都较好，表明融合结果在空间和光谱上都有显著的提升。本文算法在各指标上均表现最佳，说明融合结果能最大程度的保留图像的空间细节和光谱特征，与图 4 的视觉分析结果保持一致。

表 1 测试图 1 融合结果的评价指标

评价指标	SAM ↓	ERGAS ↓	SF ↑	RMSE ↓	SID ↓	Q4 ↑	RASE ↓	SCC ↑
GSA	7.162 1	5.861 9	43.928 6	40.374 2	0.052 9	0.821 1	24.436 7	0.859 3
BDSD-PC	7.804 0	6.144 9	44.455 3	42.092 1	0.056 2	0.776 5	25.476 5	0.840 2
MTF-GLP-FS	7.702 5	5.878 7	45.826 0	41.157 0	0.057 7	0.805 8	24.910 5	0.837 7
PNN	5.185 3	4.303 0	56.628 6	29.240 6	0.028 7	0.880 8	17.698 0	0.891 0
BDPN	6.944 0	5.297 4	54.748 4	35.727 9	0.045 0	0.828 0	21.624 5	0.891 0
MSDCNN	7.421 9	5.725 6	47.099 0	39.421 0	0.054 2	0.801 0	23.859 8	0.849 2
DRPNN	4.405 5	3.330 5	62.896 6	22.345 3	0.019 8	0.909 7	13.708 1	0.956 5
DiCNN1	4.986 8	4.088 5	58.035 3	27.801 5	0.026 4	0.881 5	16.827 0	0.935 1
FusionNet	4.532 7	3.435 2	61.758 1	23.227 5	0.021 0	0.904 4	14.058 6	0.952 2
本文算法	4.418 1	3.385 4	63.277 5	22.648 5	0.019 6	0.912 0	13.524 6	0.956 8

2.2.2 N_NWPU 数据集

图 5～6 是两组融合效果图。从主观角度分析，传统 GSA、BDSD-PC 和 MTF-GLP-FS 算法在空间上表现较好，但存在光谱失真。PNN 中图像细节信息相较于传统方法有明显的增强，但融合图像颜色偏淡。BDPN 算法有少许空间信息丢失和较多光谱信息损失，整体失真严重。MSDCNN 不仅存在严重的光谱扭曲，在空间细节保留方面也最弱，锐化后图像有伪影且与参考图像相差极大。DRPNN、DiCNN1、FusionNet 和本文算法融合结果都比较好，但本文算法无论从整体上看，还是从局部放大区域来看，光谱信息与参考图像最为一致，边缘信息突出且融合图像的质量也最高。



(a) LMS



(b) 参考图



(c) GSA



(d) BDSD-PC



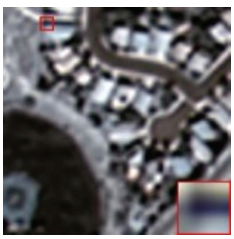
(e) MTF-GLP-FS



(f) PNN



(g) BDPN



(h) MSDCNN


(i) DRPNN

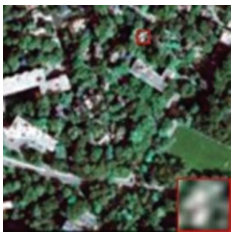

(j) DiCNN1


(k) FusionNet


(l) 本文算法

图 5　测试图 2 在不同算法下的锐化结果


(a) LMS

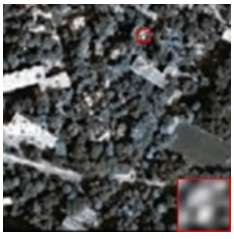

(b) 参考图

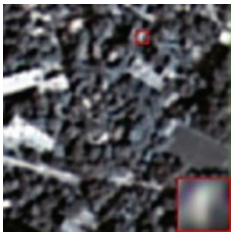

(c) GSA

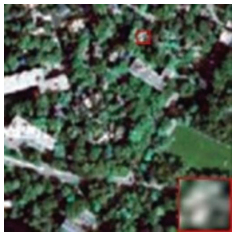

(d) BDSD-PC


(e) MTF-GLP-FS


(f) PNN


(g) BDPN


(h) MSDCNN


(i) DRPNN


(j) DiCNN1


(k) FusionNet

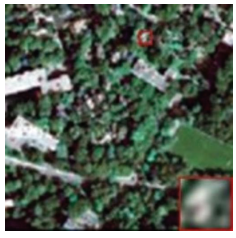

(l) 本文算法

图 6　测试图 3 在不同算法下的锐化结果

由表 2～3 可知，传统 GSA，BDSD-PC 和 MTF-GLP-FS 算法 SF 较大，说明融合结果在边缘上表现较好。PNN 中，部分 RMSE，RASE 不如传统方法，表明融合图像有轻微光谱失真。BDPN，MSDCNN 中 SID，SAM 值较大，SF 值偏低，说明锐化后图像光谱信息损失较多，空间信息保留较少。DRPNN，DiCNN1 和 FusionNet 指标值较好，说明融合结果在空间和光谱上都表现良好。本文算法各指标均最佳，说明在保持源图像空间信息不受损的前提下，最大化保留了图像光谱信息，达到了预期效果。

2.3　残差实验

为进一步验证所提方法，以测试图 3 为例进行残差实验，即将融合图和参考图作差后求均值，若求得结果中含有的细节越少，表明算法质量越高。从图 7 残差图来看，传统 GSA，BDSD-PC 和 MTF-GLP-FS 方法含有细节较多，说明融合效果欠佳，有光谱失真或空间模糊的问题。PNN，BDPN，MSDCNN，DRPNN，DiCNN1，FusionNet 方法比传统算法效果好，但也存在较清晰的纹理，融合效果还需提升。所提方法残差图含有的细节最少，融合质量最高。

2.4　消融实验

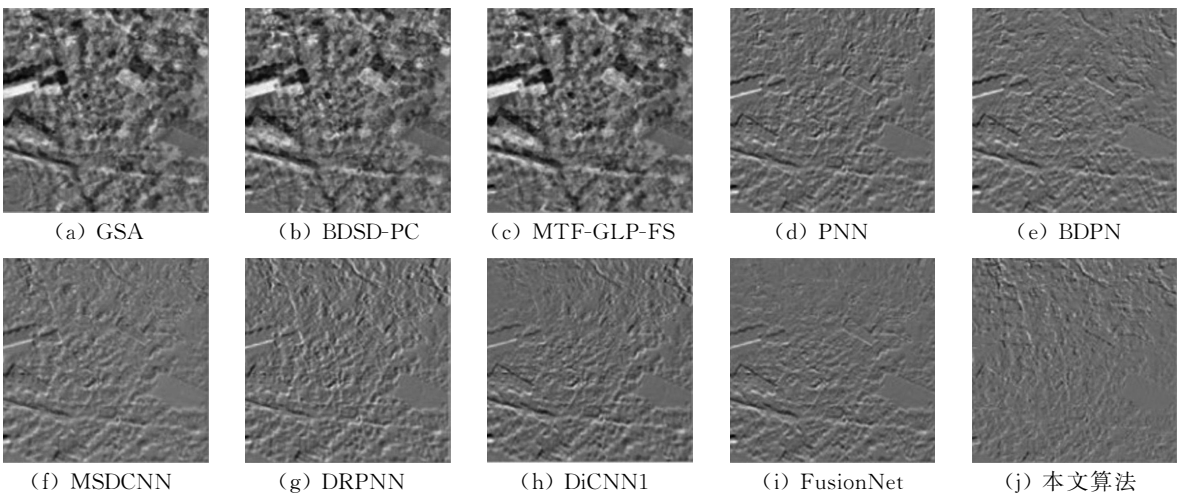
为验证改进的有效性，本节在 N_NWPU 数据集上进行消融实验。为方便，依次简化为 A，B，C，其中 Base 表示 FusionNet。从表 4 指标上能看出，与 Base 相比，A 与 A＋B 都得到了很好的融合结果，验证了主干网络改进和注意力的可行性。接着在其基础上又添加了 C，融合结果提升得更大。

表 2 测试图 2 融合结果的评价指标

评价指标	SAM ↓	ERGAS ↓	SF ↑	RMSE ↓	SID ↓	Q4 ↑	RASE ↓	SCC ↑
GSA	3.530 1	1.737 7	26.335 1	5.447 0	0.015 7	0.967 1	6.541 6	0.985 8
BDSD-PC	3.306 6	1.701 3	25.656 2	5.255 7	0.015 9	0.966 4	6.312 0	0.987 0
MTF-GLP-FS	3.635 1	2.154 8	26.142 1	6.936 6	0.017 9	0.961 5	8.330 6	0.983 3
PNN	2.565 5	2.332 6	20.171 6	7.641 9	0.015 1	0.949 1	9.177 7	0.977 1
BDPN	3.968 4	1.717 6	23.215 0	5.531 5	0.017 5	0.953 7	6.643 2	0.991 1
MSDCNN	4.772 3	4.857 3	10.011 6	16.552 2	0.021 8	0.817 3	19.878 7	0.757 7
DRPNN	0.899 5	0.501 2	26.206	1.610 8	0.005 0	0.996 5	1.934 5	0.999 3
DiCNN1	0.690 1	0.274 5	26.423 7	0.888 1	0.003 3	0.998 2	1.066 6	0.999 5
FusionNet	0.679 2	0.296 5	26.426 4	0.916 3	0.002 9	0.997 4	1.100 5	0.999 4
本文算法	0.530 5	0.226 7	26.467 9	0.733 2	0.002 2	0.998 8	0.880 6	0.999 6

表 3 测试图 3 融合结果的评价指标

评价指标	SAM ↓	ERGAS ↓	SF ↑	RMSE ↓	SID ↓	Q4 ↑	RASE ↓	SCC ↑
GSA	2.789 6	1.539 8	19.565 2	5.346 6	0.004 8	0.957 2	6.099 2	0.985 4
BDSD-PC	2.648 6	1.415 7	19.499 3	4.729 4	0.002 6	0.957 0	5.395 1	0.983 9
MTF-GLP-FS	2.760 4	1.840 8	19.090 4	6.324 5	0.004 1	0.953 0	7.214 8	0.980 8
PNN	2.475 2	1.504 9	17.719 9	5.081 0	0.005 1	0.952 3	5.796 2	0.981 3
BDPN	3.927 7	1.785 4	18.622 2	6.451 9	0.008 6	0.957 2	7.360 0	0.987 3
MSDCNN	5.069 9	4.025 7	9.069 7	14.370 3	0.020 1	0.811 3	16.393 0	0.773 9
DRPNN	0.813 5	0.379 5	19.520 1	1.355 4	0.000 3	0.988 0	1.546 1	0.998 6
DiCNN1	0.744 3	0.345 8	19.527 5	1.229 5	0.000 3	0.986 8	1.402 6	0.998 8
FusionNet	0.744 0	0.348 1	19.536 3	1.219 3	0.000 2	0.985 1	1.390 9	0.998 7
本文算法	0.646 3	0.301 8	19.754 0	1.078 4	0.000 1	0.988 2	1.230 2	0.998 9



(a) GSA (b) BDSD-PC (c) MTF-GLP-FS (d) PNN (e) BDPN

(f) MSDCNN (g) DRPNN (h) DiCNN1 (i) FusionNet (j) 本文算法

图 7 测试图 3 残差图

表 4 N_NWPU 数据集上消融实验的评价指标

方法	A	B	C	SAM ↓	ERGAS ↓	SF ↑	RMSE ↓	SID ↓	Q4 ↑	RASE ↓	SCC ↑
Base				1.023 8	0.483 6	46.466 6	1.322 6	0.000 9	0.998 7	1.764 7	0.999 5
Base+A	√			0.638 5	0.281 3	46.603 8	0.807 9	0.000 3	0.999 5	1.078 0	0.999 7
A+B	√	√		0.593 0	0.275 1	46.691 1	0.791 1	0.000 2	0.999 6	1.055 5	0.999 7
A+B+C	√	√	√	0.569 7	0.265 9	46.955 6	0.766 4	0.000 1	0.999 7	1.022 6	0.999 8

2.5 运行效率

运算速度也是评价算法性能的一个关键要素。本节在 QuickBird 和 N_NWPU 数据集上进行验证。从表 5 可以看出，虽然 PNN，BDPN，

MSDCNN，DRPNN，DiCNN1 和 FusionNet 方法所耗费的时间较短，但所提方法运行时间最少，验证了改进后网络不仅能缓解梯度冗余，还能提升计算效率，增强 CNN 的学习力。

表 5　QB 和 N_NWPU 数据集上平均测试时间

数据集	方法	测试时间/s
QuickBird	PNN	0.365
	BDPN	0.427
	MSDCNN	0.373
	DRPNN	0.393
	DiCNN1	0.365
	FusionNet	0.368
	本文算法	0.319
N_NWPU	PNN	0.094
	BDPN	0.118
	MSDCNN	0.106
	DRPNN	0.120
	DiCNN1	0.090
	FusionNet	0.112
	本文算法	0.082

3　结论

为解决遥感图像锐化中存在的空间模糊和光谱失真的问题，基于 FusionNet 模型，提出了一种新的全色锐化方法。该方法借鉴 CSPNet 和 SimAM 的思想，设计一个 Attention BasicBlock 块，仅使用卷积跳连的方式替换残差结构，克服了 FusionNet 因残差连接较多而造成的梯度冗余问题。另外，还设计了一种可学习作差方法，进一步提取原始图像中深层次的特征信息。实验结果表明，所提方法在 QuickBird 和 N_NWPU 数据集上都获得了最佳融合效果，且客观指标也表现出了最佳性能。此外，残差实验、消融实验和运行效率也证明了该方法对性能提升起着积极作用。下一步将在损失函数上展开研究，以求实现更好效果。

参考文献：

[1] 陈毛毛，郭　擎，刘明亮，等. 密集卷积残差网络的遥感图像融合[J]. 遥感学报，2021，25(6)：1270-1283.
Chen Maomao，Guo Qing，Liu Mingliang，et al. Pansharpening by residual network with dense convolution for remote sensing images[J]. National Remote Sensing Bulletin，2021，25(6)：1270-1283.

[2] 李树涛，李聪好，康旭东. 多源遥感图像融合发展现状与未来展望[J]. 遥感学报，2021，25(1)：148-166.
Li Shutao，Li Congyu，Kang Xudong. Development status and future prospects of multi-source remote sensing image fusion [J]. National Remote Sensing Bulletin，2021，25(1)：148-166.

[3] Wu L，Yin Y J，Jiang X Y，et al. Pansharpening based on multi-objective decision for multi-band remote sensing images [J]. Pattern Recognition，2021，118：108022.

[4] Singh S，Anand R S. Multimodal medical image fusion using hybrid layer decomposition with CNN-based feature mapping and structural clustering[J]. IEEE Trans. Instrum. Meas.，2020，69(6)：3855-3865.

[5] Zhou H Y，Liu Q J，Weng D W，et al. Unsupervised cycle-consistent generative adversarial networks for pan-sharpening [J]. IEEE Trans. Geosci. Remote Sens.，2021，60：1-14.

[6] 胡建文，汪泽平，胡　佩. 基于深度学习的空谱遥感图像融合综述[J]. 自然资源遥感，2023，35(1)：1-14.
Hu Jianwen，Wang Zeping，Hu Pei. A review of pan-sharpening methods based on deep learning[J]. Remote Sensing for Natural Resources，2023，35(1)：1-14.

[7] Ma D，Lai H C. Remote sensing image matching based improved ORB in NSCT domain[J]. J. Indian Soc. Remote Sens.，2019，47(5)：801-807.

[8] Liu Y，Wang L，Cheng J，et al. Multi-focus image fusion：a survey of the state of the art[J]. Information Fusion，2020，64：71-91.

[9] Masi G，Cozzolino D，Verdoliva L，et al. Pansharpening by convolutional neural networks[J]. Remote Sens.，2016，8(7)：594.

[10] 邓良剑，冉　燃，吴　潇，等. 遥感图像全色锐化的卷积神经网络方法研究进展[J]. 中国图象图形学报，2023，28(1)：57-79.
Deng Liangjian，Xiao Ran，Wu Xiao，et al. CNN-based remote sensing pan-sharpening：A critical review[J]. J. of Image and Graphics，2023，28(1)：57-79.

[11] He L，Rao Y Z，Li J，et al. Pansharpening via detail injection based convolutional neural networks[J]. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.，2019，12(4)：1188-1204.

[12] Deng L J，Vivone G，Jin C，et al. Detail injection-based deep convolutional neural networks for pansharpening[J]. IEEE Trans. Geosci. Remote Sens.，2021，59(8)：6995-7010.

[13] Wang C Y，Liao H Y M，Wu Y H，et al. CSPNet：A new backbone that can enhance learning capability of CNN[C]// 2020 IEEE/CVF Conf. on Computer Vision and Pattern Recognition Workshops(CVPRW). IEEE，2020：1571-1580.

[14] Yang L X，Zhang R Y，Li L D，et al. SimAM：A simple，parameter-free attention module for convolutional neural networks [C]// International Conference on Machine Learning. PMLR，2021，139：11863-11874.

[15] 张立福，赵晓阳，孙雪剑，等. 高分五号高光谱数据融合方法比较[J]. 遥感学报，2022，26(4)：632-645.
Zhang Lifu，Zhao Xiaoyang，Sun Xuejian，et al. Comparison of fusion methods on GF-5 hyperspectral data [J]. National

Remote Sensing Bulletin, 2022, 26(4): 632-645.

[16] Vivone G. Robust band-dependent spatial-detail approaches for panchromatic sharpening [J]. IEEE Trans. Geosci. Remote Sens., 2019, 57(9): 6421-6433.

[17] Vivone G, Restaino R, Chanussot J. Full scale regression based injection coefficients for panchromatic sharpening[J]. IEEE Trans. Image Process., 2018, 27: 3418-3431.

[18] 唐霖峰, 张 浩, 徐 涵, 等. 基于深度学习的图像融合方法综述[J]. 中国图象图形学报, 2023, 28(1): 3-36.
Tang Linfeng, Zhang Hao, Xu Han, et al. Deep learning based image fusion: A survey[J]. Journal of Image and Graphics, 2023, 28(1): 3-36.

[19] 方 帅, 潮 蕾, 曹风云. 自适应权重注入机制遥感图像融合[J]. 中国图象图形学报, 2020, 25(3): 546-557.
Fang Shuai, Chao Lei, Cao Fengyun. New pan-sharpening method based on adaptive weight mechanism[J]. Journal of Image and Graphics, 2020, 25(3): 546-557.

[20] Liu X, Liu Q, Wang Y. Remote sensing image fusion based on two-stream fusion network [J]. Information Fusion, 2020, 55: 1-15.

[21] Zhang Y L, Tian Y P, Kong Y, et al. Residual dense network for image super-resolution[C]// 2018 IEEE/CVF Conf. on Computer Vision and Pattern Recognition. IEEE, 2018: 2472-2481.

[22] Aiazzi B, Alparone L, Baronti S, et al. Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis [J]. IEEE Trans. Geosci. Remote Sens., 2002, 40(10): 2300-2312.

作者简介:

刘蕙绮(1999—),女,黑龙江省哈尔滨市人,硕士研究生,主要研究方向为机器学习、图像处理;

陈应霞(1978—),男,湖北省黄冈市人,博士,讲师,硕士生导师,研究方向为机器学习、图像处理。