

## 基于多分支 CNN 的高光谱与全色影像融合处理

王洪斌, 肖嵩\*\*, 曲家慧\*, 董文倩, 张同振

西安电子科技大学综合业务网理论及关键技术国家重点实验室, 陕西 西安 710071

**摘要** 高光谱与全色影像融合旨在通过融合高空间分辨率的全色影像与低空间分辨率的高光谱影像来获得高空间分辨率的高光谱影像。基于深度卷积神经网络(CNN),提出了一种遥感影像融合方法,利用两个独立的分支网络逐级从高光谱和全色影像中提取光谱和空间特征。该融合网络由两个分支网络和一个主线网络组成,利用两个分支网络分别从高光谱与全色影像中提取空谱特征,主线网络基于分支网络提取的特征,重建得到最终融合的高空间分辨率的高光谱影像。在 CAVE 和 Pavia Center 数据集上分别进行了实验验证,通过对比可以发现,所提出的融合算法在空间细节和光谱保真度上较当前主流算法均表现出更优异的性能。

**关键词** 图像处理; 高光谱影像; 融合; 卷积神经网络; 空谱特征

中图分类号 TP751.1

文献标志码 A

doi: 10.3788/AOS202141.0710001

## Pansharpening Based on Multi-Branch CNN

Wang Hongbin, Xiao Song\*\*, Qu Jiahui\*, Dong Wenqian, Zhang Tongzhen

State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an, Shaanxi 710071, China

**Abstract** Pansharpening aims to obtain hyperspectral images with high spatial resolutions by fusing hyperspectral images with low spatial resolutions and panchromatic images with high spatial resolutions together. This paper introduces a remote sensing image fusion method based on a deep convolutional neural network (CNN), which extracts spectral and spatial features step by step from hyperspectral and panchromatic images using two independent branch networks. The proposed fusion network is composed of two branches and a main network. The two independent branch networks are used for extracting the spatial-spectral features from hyperspectral and panchromatic images, while based on the features extracted from the branch network, the main network is used to reconstruct and the final fused hyperspectral images with high spatial resolutions are obtained. The experimental verifications were conducted on both CAVE and Pavia Center datasets. Through comparison, one can see that the proposed fusion algorithm outperforms the prevailing algorithms in terms of spatial detail and spectral fidelity.

**Key words** image processing; hyperspectral image; fusion; convolutional neural network; spatial-spectral feature

**OCIS codes** 100.4996; 100.4145; 110.4234

## 1 引言

高光谱与全色融合处理是将低空间分辨率高光谱(HS)影像与高空间分辨率的全色(PAN)影像进行信息提取与融合,其目的在于将全色影像的空间分辨率特性与高光谱影像的光谱分辨率特性进行互补,从而生成高空间分辨率的高光谱影像。高光谱

与全色融合被广泛应用于遥感任务(如异常检测<sup>[1]</sup>、空间特征提取<sup>[2]</sup>、视觉影像分析和场景解释<sup>[3]</sup>)的影像增强过程中。

高光谱与全色融合主要有四大类经典方法:成分替换(CS)、多分辨率分析(MRA)、贝叶斯和矩阵分解。其中 CS 方法通常是先将高光谱影像投影到新的空间域中以分离空间和光谱信息,然后将表征

收稿日期: 2020-10-14; 修回日期: 2020-11-02; 录用日期: 2020-11-11

基金项目: 国防预研基金(JY0600090102)、中国长江学者奖励计划(CJT160102)

\* E-mail: jhqu@xidian.edu.cn; \*\* E-mail: xiaosong@mail.xidian.edu.cn

空间信息的分量替换为高空间分辨率的 PAN 影像,最后执行逆变换以获得最终的融合影像。典型的 CS 方法包括:强度-色相饱和度(IHS<sup>[4]</sup>)、主成分分析(PCA<sup>[5]</sup>)、Gram-Schmidt Adaptive(GSA<sup>[6]</sup>)。MRA 类算法通常基于小波变化或拉普拉斯金字塔等,将原影像分解为一系列带通通道,然后在相应的光谱频段插入全色影像分解的高频通道,最后重建得到最终的融合影像。常见的 MRA 策略有:基于平滑滤波的强度调制(SFIM)<sup>[7]</sup>和 MTF-GLP-HPM (Modulation Transfer Function-Generalized Laplacian Pyramid-High Pass Modulation)<sup>[8]</sup>。基于矩阵分解的基本思想是将原始矩阵分解为两个矩阵的乘积,从而对原始高维矩阵进行降维。利用矩阵分解对原始影像进行降维,分解结果有较为明确的物理意义,典型算法包含非负矩阵分解(NMF<sup>[9]</sup>)和耦合非负矩阵分解(CNMF<sup>[10]</sup>)等。贝叶斯类方法一般利用非参数贝叶斯稀疏来表征高光影像与全色影像,利用全色影像的空间结构信息以及高光影像的光谱域信息来重建影像,典型的算法有贝叶斯稀疏算法(Bayesian Sparse)<sup>[11]</sup>等。

近年来,深度学习尤其是卷积神经网络表现出出色的应用前景,已被广泛应用于高光影像与全色融合的遥感影像处理中。深度学习的最大优势在于,可以在训练样本的监督下更新模型中的所有参数,从而减少对先验知识的需求,而且可以期望更高的拟合精度。近年来,基于高光影像与全色融合神经网络(PNN<sup>[12]</sup>)的融合效果已经有了显著的提升,但是由于 PNN 的设计完全是基于超分辨率卷积神经网络(SRCNN<sup>[13]</sup>)提出的,其性能仍有很大的改进空间。Zheng 等<sup>[14]</sup>使用深高光先验和双注意力残差网络,改进了采样过程中的信息丢失问题,并提高了 CNN 的判别能力。Yang 等<sup>[15]</sup>提出了一种新颖的渐进级联深度残差网络,较为出色地解决了融合图像的高频细节丢失问题。Hu 等<sup>[16]</sup>提出了一种新的光谱损失函数以改善融合过程中光谱信息丢失问题。为了充分利用低空间分辨率的高光影像丰富的光谱信息和空间信息以及高空间分辨率的全色影像充足的空间细节,本文提出了一种基于多分支 CNN 的高光影像和全色影像融合方法。该融合方法采用两个基于卷积自编码的分支网络,有效提取了高光影像的空谱特征和全色影像的空间特征,并构建了基于空谱特征的重建网络,利用两个分支网络提取的特征重建了具有高空间分辨率和高光谱分

辨率的高光谱影像。为了验证算法的有效性,在 CAVE 数据集和 Pavia Center 数据集上进行了实验验证,结果表明,本文提出的高光影像和全色影像融合算法相比于其他融合算法更好地保持了光谱信息,且更有效地提升了空间性能。

## 2 本文方法

本文算法的网络结构如图 1 所示,主要包括两部分:高光影像和全色影像的特征提取,高光影像空谱的特征重建。在本文算法中,使用  $\mathbf{X}$  表示融合的高空间分辨率高光影像, $\mathbf{X} \in \mathbf{R}^{W \times H \times B}$ , $\mathbf{Y}$  表示可获得的低空间分辨率高光影像, $\mathbf{Y} \in \mathbf{R}^{w \times h \times B}$ , $\mathbf{Z}$  表示可获得的高空间分辨率全色影像, $\mathbf{Z} \in \mathbf{R}^{W \times H}$ ,其中, $\mathbf{R}$  为影像实体, $H$  和  $W$  分别表示全色影像每个波段的长和宽, $h$  和  $w$  分别表示低空间分辨率高光影像每个波段的长和宽, $W \times H = (w \times h) \times r^2$ , $r$  表示全色影像和高光谱影像的空间分辨率比例, $B$  表示高光影像的波段数。

### 2.1 高光影像和全色影像的特征提取

为了有效获取高光影像的空谱特征和全色影像的空间特征,如图 1 所示,本文采用了两个独立且相似结构的卷积自编码网络,通过编码解码网络尽可能准确地对输入的高光谱影像和全色影像进行重构,无监督地学习高光影像和全色影像的隐层特征。由于卷积操作考虑了结构信息,卷积自编码比使用全连接层的自编码更适合处理影像,更利于自动提取输入影像的特征。

#### 2.1.1 高光影像空谱特征提取分支

由于输入的高光谱影像具有与全色影像不同的空间分辨率,本文算法使用最近邻插值方法对高光影像  $\mathbf{Y}$  进行上采样处理,获得具有与全色影像相同空间分辨率的插值高光影像  $\mathbf{Y}_U$ 。将插值高光影像  $\mathbf{Y}_U$  作为网络输入,经过三个卷积层,每个卷积层核大小为  $3 \times 3$ ,步长为 2,获得三个尺度的输出特征,可以表示为

$$\mathbf{F}_H^i = f_{\text{conv}}^i \{ f_{\text{conv}}^{i-1} \cdots [f_{\text{conv}}^1(\mathbf{Y}_U)] \}, i = 1, 2, 3, (1)$$

式中: $\mathbf{F}_H^1$ , $\mathbf{F}_H^2$  和  $\mathbf{F}_H^3$  分别为三个尺度的输出特征; $f_{\text{conv}}^i$  为第  $i$  层卷积操作。解码部分网络与前面译码部分网络对称,使用三个反卷积层重建插值高光影像。解码部分网络还采用了特征融合,将译码部分获得的特征  $\mathbf{F}_H^1$ , $\mathbf{F}_H^2$  与译码部分的特征进行了融合以获得更准确的上下文信息,达到更好的重建效果。这个过程可以表示为

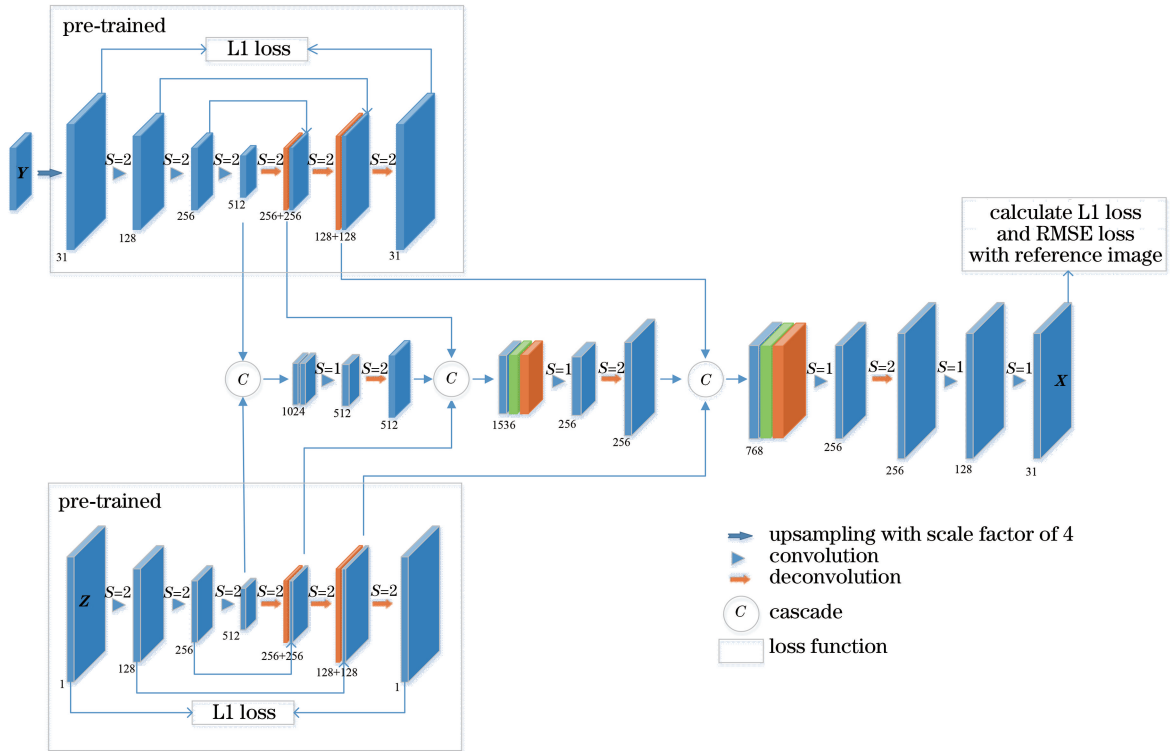


图 1 本文算法的网络结构图

Fig. 1 Network structure in proposed algorithm

$$Y'_U = f_{\text{deconv}} \{ C \{ F_H^1, f_{\text{deconv}} \{ C \{ F_H^2, f_{\text{deconv}} \{ F_H^3 \} \} \} \} \}, \quad (2)$$

式中:  $Y'_U$  为重建的插值高光谱影像;  $f_{\text{deconv}}$  为反卷积操作;  $C$  表示级联操作。令  $F_H^4$  和  $F_H^5$  分别为第一个反卷积操作和第二个反卷积操作的输出, 则  $F_H^4 = f_{\text{deconv}}(F_H^3)$ ,  $F_H^5 = f_{\text{deconv}}[C(F_H^4, F_H^2)]$ 。获得的隐层特征  $\{F_H^i\}_{i=1,2,3,4,5}$  是提取的高光谱影像的多尺度空谱特征。

### 2.1.2 全色影像空间特征提取分支

提取全色影像空间特征的网络与高光谱影像空谱特征的卷积自编码网络基本相似, 将获得的全色影像  $Z$  作为网络输入, 经过三次卷积操作, 获得三个尺度的输出特征  $F_P^1$ 、 $F_P^2$  和  $F_P^3$ 。将  $F_P^3$  输入到编码网络, 经过反卷积操作和特征融合, 获得两次反卷积操作的输出  $F_P^4$  和  $F_P^5$ 。隐层特征  $\{F_P^i\}_{i=1,2,3,4,5}$  是提取的高光谱影像的多尺度空间特征。

### 2.1.3 损失函数

由于损失函数使用 L2 范数平滑影像边缘地区, 因此本文高光谱影像空谱特征提取卷积自编码网络和全色影像空间特征提取卷积自编码网络均使用 L1 范数, 以最小化输入影像和重建影像之间的误差, 进而优化网络。高光谱影像特征提取网络和全色影像特征提取网络的损失函数分别为

$$L_H = \frac{1}{S} \sum_{s=1}^S \| Y'_{U,s} - Y_{U,s} \|_1, \quad (3)$$

$$L_P = \frac{1}{S} \sum_{s=1}^S \| Z'_s - Z_s \|_1, \quad (4)$$

式中:  $L_H$  为高光谱影像特征提取网络的损失函数;  $L_P$  为全色影像特征提取网络的损失函数;  $S$  为训练样本数;  $Y_{U,s}$  和  $Y'_{U,s}$  分别为第  $s$  对输入的参考高光谱影像和重建的高光谱影像;  $Z_s$  和  $Z'_s$  分别为第  $s$  对输入的参考全色影像和重建的全色影像。

## 2.2 高光谱影像重建

### 2.2.1 基于空谱特征的高光谱影像重建网络

利用卷积自编码网络提取了充足的高光谱影像空谱特征  $\{F_H^i\}_{i=1,2,3,4,5}$  和全色影像空间特征  $\{F_P^i\}_{i=1,2,3,4,5}$  后, 本文提出了基于空谱特征的高光谱影像重建网络, 依次将提取的全色影像和高光谱影像对应的多尺度特征输入重建网络, 有效提高了重建的高光谱影像  $X$  的空间分辨率和光谱分辨率。小尺度高光谱空谱特征  $F_H^3$  和小尺度全色空间特征  $F_P^3$  级联作为网络的输入, 经过反卷积操作后, 上采样到较高尺度, 并与提取的高光谱影像相应尺度的特征  $C(F_H^2, F_H^4)$  和  $C(F_P^2, F_P^4)$  级联, 以作为下一层反卷积层的输入。反卷积后的输出同样与提取的高光谱影像相应尺度的特征  $C(F_H^1, F_H^5)$  和



$C(\mathbf{F}_H^1, \mathbf{F}_H^5)$  级联, 经过一层卷积操作, 实现特征融合, 并经过一层反卷积操作, 获得与全色影像相同尺寸的输出图。该过程可以表示为

$$\mathbf{F}_{HP}^1 = f_{\text{deconv}} \{ f_{\text{conv}} [C(\mathbf{F}_H^3, \mathbf{F}_P^3)] \}, \quad (5)$$

$$\mathbf{F}_{HP}^2 = f_{\text{deconv}} \{ f_{\text{conv}} \{ C [C(\mathbf{F}_H^2, \mathbf{F}_H^4), C(\mathbf{F}_P^2, \mathbf{F}_P^4), \mathbf{F}_{HP}^1] \} \}, \quad (6)$$

$$\mathbf{F}_{HP}^3 = f_{\text{deconv}} \{ f_{\text{conv}} \{ C [C(\mathbf{F}_H^1, \mathbf{F}_H^5), C(\mathbf{F}_H^1, \mathbf{F}_H^5), \mathbf{F}_{HP}^2] \} \}, \quad (7)$$

式中:  $\mathbf{F}_{HP}^1$ 、 $\mathbf{F}_{HP}^2$  和  $\mathbf{F}_{HP}^3$  分别为三次反卷积操作的输出结果。在学习了高光谱和全色影像充足的各尺度特征后, 对获得的特征  $\mathbf{F}_{HP}^3$  进行两次卷积操作, 可获得最终的融合的高空间分辨率高光谱影像  $\mathbf{X}$ 。

### 2.2.2 损失函数

为了以更高的空谱保真度重建高空间分辨率高光谱影像, 本文基于空谱特征的高光谱影像重建网络不仅使用 L1 范数约束重建的高光谱影像与参考高光谱影像之间的误差, 还采用用于测量融合影像空谱失真程度的全局指标均方根误差 (Root mean squared error, RMSE) 同时约束高光谱影像重建网络的误差。损失函数为

$$L_R = \frac{1}{T} \sum_{t=1}^T \left[ \|\mathbf{X}_t - \mathbf{R}_t\|_1 + \frac{\|\mathbf{X}_t - \mathbf{R}_t\|_F}{\sqrt{W \times H \times B}} \right], \quad (8)$$

式中:  $T$  为训练样本数;  $\|\cdot\|_F$  表示 F 范数;  $\mathbf{X}_t$  和  $\mathbf{R}_t$  分别为第  $t$  对重建高光谱影像和参考高光谱影像。

## 3 分析与讨论

### 3.1 数据集

为了评估本文所提出的融合方法的性能, 分别在两个被广泛用于高光谱影像融合的数据集上进行了实验, 这两个数据集分别是 CAVE、Pavia Center。

CAVE 数据集是哥伦比亚大学计算机视觉实验室通过广义分类像素相机获取的, 其中包含了 32 个室内的高空间分辨率 HS 影像。这些 HS 影像在 400~700 nm 波长范围内有 31 个波段, 空间分辨率为 512 pixel×512 pixel。从中随机选取 22 张作为训练集, 剩余 10 张作为测试集。由于训练样本数量有限, 对参考影像进行进一步的划分, 将 22 张尺寸为 512×512×31 的 HS 影像均分成 16 张 128×128×31 的影像。最终我们的训练集共有 22×16 张尺寸为 128×128×31 的 HS 影像, 测试集有 10

张尺寸为 512×512×31 的 HS 影像。

Pavia Center 数据集是通过 ROSIS 传感器在意大利北部 Pavia 上空拍摄的, 其在 430~860 nm 光谱范围内有 115 个波段, 空间分辨率为 1096 pixel×1096 pixel。为了提高效率, 将 13 个数据质量较差波段的数据去除, 最终的参考影像的尺寸为 960×640×102。将参考影像无重叠地均分成 12 份, 得到 160×320×102 张大块影像, 其中 9 张用来训练, 剩余 3 张用作测试。再将每一张大块影像以步长 8 滑动分割成有重叠的 21 张尺寸为 160 pixel×160 pixel 的小样本。最终我们拥有 9×21 张尺寸为 160 pixel×160 pixel 的训练样例以及 3×21 张尺寸为 160 pixel×160 pixel 的测试样例。

对这两个数据集进行上述的预处理后, 根据 Wald 协议, 利用参考影像生成融合影像所需要的 PAN 影像和低空间分辨率的 HS 影像。Wald 协议是基于尺度不变性的假设, 具体地, PAN 影像是通过对参考影像的可见光谱带求平均生成的, 即生成的 PAN 影像的通道数为 1, 尺度与参考影像保持一致; 低空间分辨率的 HS 影像是将参考影像通过尺寸为 9×9 的高斯滤波器, 并采用 4 为比例因子对其进行下采样获得的。最终, CAVE 数据集得到的 HS 尺寸为 32×32×31, Pavia Center 得到的 HS 尺寸为 40×40×102。然后将生成的 PAN 影像和低空间分辨率 HS 影像进行融合, 最终生成高空间分辨率的 HS 影像, 并与原始 HS 影像进行比较。

### 3.2 对比算法

为了评估本文所提出的算法性能, 采用当前 6 种被广泛用于高光谱与全色融合的优秀算法进行比较, 这 6 种算法分别是 PCA、GSA、SFIM、CNMF、Bayesian sparse 和 MTF-GLP-HPM。其中 PCA、GSA 属于 CS 算法, 这两种算法简单而高效, 其融合影像的空间细节有着很高的保真度, 而且 GSA 是 GS 的改进版算法, 该算法已被成功应用于很多商业软件中。SFIM、MTF-GLP-HPM 是 MRA 类典型的优秀算法, 能够良好地保存影像的频谱内容, 具有时间相干性和频谱一致性。CNMF 基于矩阵分解, 该算法在多个数据集上均可以在空间和光谱方面生成高质量的融合影像。Bayesian sparse 是贝叶斯类算法, 可以通过贝叶斯融合模型的后验分布对其融合过程进行直观解释, 而且有着较佳的融合效果。

### 3.3 评价指标

本文采用广泛使用的四个质量指标来客观评估融合的结果。这四个指标分别为相关系数 (CC)、光



谱角度映射(SAM)、均方根误差和相对全局合成误差(ERGAS)。

相关系数表征的是两幅影像之间的几何变化程度:

$$C_{CC}(\hat{\mathbf{D}}, \mathbf{D}) = \frac{1}{m_\lambda} \sum_{i=1}^{m_\lambda} C_{CCS}(\hat{\mathbf{D}}^i, \mathbf{D}^i), \quad (9)$$

$$C_{CCS}(\mathbf{A}, \mathbf{B}) = \frac{\sum_{j=1}^n (A_j - \mu_A)(B_j - \mu_B)}{\sqrt{\sum_{j=1}^n (A_j - \mu_A)^2 \sum_{j=1}^n (B_j - \mu_B)^2}}, \quad (10)$$

式中:  $C_{CC}$  为相关系数;  $C_{CCS}$  为单波段的互相关系数;  $\hat{\mathbf{D}}^i$  为融合影像,  $\hat{\mathbf{D}}^i \in \mathbf{R}^{m_\lambda \times n}$ , 其中  $m_\lambda$  为光谱波段数量,  $n$  为像素个数;  $\mathbf{D}^i$  为参考影像,  $\mathbf{D}^i \in \mathbf{R}^{m_\lambda \times n}$ ;  $\mathbf{A}$ ,  $\mathbf{B}$  为单通道影像,  $\mathbf{A}, \mathbf{B} \in \mathbf{R}^{1 \times n}$ ;  $A_j$  表示  $\mathbf{A}$  的第  $j$  个元素;  $\mu_A$  为  $\mathbf{A}$  的样本均值,  $\mu_A = \frac{1}{n} \sum_{j=1}^n A_j$ ;  $B_j$  为  $\mathbf{B}$  的第  $j$  个元素;  $\mu_B$  为  $\mathbf{B}$  的样本均值。CC 最优值为 1。

光谱角度映射表征的是光谱形状的保持程度:

$$S_{SAM}(\hat{\mathbf{D}}, \mathbf{D}) = \frac{1}{n} \sum_{j=1}^n S_{SAM}(\hat{\mathbf{d}}_j, \mathbf{d}_j), \quad (11)$$

$$S_{SAM}(\mathbf{a}, \mathbf{b}) = \arccos\left(\frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\mathbf{a}\| \|\mathbf{b}\|}\right), \quad (12)$$

式中:  $\hat{\mathbf{d}}_j$  为融合图像第  $j$  个像素的光谱矢量;  $\mathbf{d}_j$  为参考图像第  $j$  个像素的光谱矢量;  $\mathbf{a}, \mathbf{b}$  为融合图像和参考图像第  $j$  个像素的光谱矢量,  $\mathbf{a}, \mathbf{b} \in \mathbf{R}^{m_\lambda}$ ;  $\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}^T \mathbf{b}$  为  $\mathbf{a}$  和  $\mathbf{b}$  的内积;  $\|\cdot\|$  是 L2 范数。SAM 的最优值为 0。该值是通过对所有影像像素的值进行求平均得到的。

均方根误差表征的是参考影像与融合影像之间的标准差:

$$E_{RMSE}(\hat{\mathbf{D}}, \mathbf{D}) = \frac{\|\hat{\mathbf{D}} - \mathbf{D}\|_F}{\sqrt{n \cdot m_\lambda}}, \quad (13)$$

该值是通过计算参考影像和融合影像的 L2 范数得到的。其中,  $\|\mathbf{D}\|_F = \sqrt{\text{trace}(\mathbf{D}^T \mathbf{D})}$  是  $\mathbf{D}$  的矩阵范数, RMSE 的最优值为 0, 其中  $\text{trace}(\cdot)$  为求迹运算。

相对全局合成误差是表征融合影像质量的全局指标:

$$E_{ERGAS}(\hat{\mathbf{D}}, \mathbf{D}) = 100d \sqrt{\frac{1}{m_\lambda} \sum_{k=1}^{m_\lambda} \left(\frac{E_{RMSE,k}}{\mu_k}\right)^2}, \quad (14)$$

$$E_{RMSE,k} = \frac{\|\hat{\mathbf{D}}^k - \mathbf{D}^k\|_F}{\sqrt{n}}, \quad (15)$$

式中:  $E_{RMSE,k}$  为第  $k$  个波段的均方根误差;  $\hat{\mathbf{D}}^k$  为融合影像第  $k$  个波段的影像;  $\mathbf{D}^k$  为参考影像第  $k$  个波段的影像;  $\mu_k$  为第  $k$  个波段的样本均值;  $d$  为 PAN 和 HS 影像的线性分辨率之比。ERGAS 的最优值为 0。

### 3.4 主观分析

分别在 CAVE 和 Pavia Center 数据集上进行了实验, 采用的学习率为 0.001, 训练次数为 1000, 单批次有 16 张影像。从测试集中各选取了两张进行主观分析, 融合得到的影像如图 2~5 所示, 每组影像有两幅图, 分别是对应算法得到的融合影像以及局部放大影像。

如图 2 所示, 该组影像使用的是 Pavia Center 数据集第 6 大块影像的第 19 小块影像, 选取了 [22, 34, 46] 通道。可以观察到, 参考影像局部放大图中虚线框区域内有明显的三辆车, 其中最右侧为一蓝色车辆。通过对比可以发现, SFIM、PCA、MTF-GLP-HPM 算法可以在相同区域内隐约看到紫色的像素点, 这说明存在一定程度的光谱失真; 除了 Bayesian sparse 算法外, 其余算法得到的融合影像的虚线框区域内中有三个物体; 但是相较于本文所提出的算法, 无论是纹理的清晰程度还是光谱的保真度, 其余算法都十分逊色。

如图 3 所示, 该组影像使用的是 Pavia Center 数据集第 6 大块影像的第 8 小块影像, 选取了 [22, 34, 46] 通道。由图 3 可见, 放大图中虚线框区域内有 6 棵树, 左侧上有一些蓝色和白色的车辆。对比各个算法得到的融合影像, 可以看出, PCA 得到的融合影像信息缺失严重, 光谱保真度较差; SFIM、GSA、CNMF 均可以大致展现出树的轮廓, 但是细节不够完善, 并不能清晰地分离出树的影子与树本体, 而且左侧车辆颜色也基本看不清, 说明其光谱失真较为严重; 而本文所提算法对树的纹理和光谱细节均有细致的表达, 而且对于旁边的车辆, 也基本能展现出车辆的光谱信息。

如图 4 所示, 该组影像使用的是 CAVE 数据集, 选取的是第 28 张影像的 [5, 13, 25] 通道。该影像包括色块和字母两部分内容。通过对比可以发现, 在 SFIM、MTF-GLP-HPM 的融合结果中, 过锐化较为严重, 出现明显的空间失真; GSA 和 Bayesian sparse 的融合结果明显模糊, 其空间细节

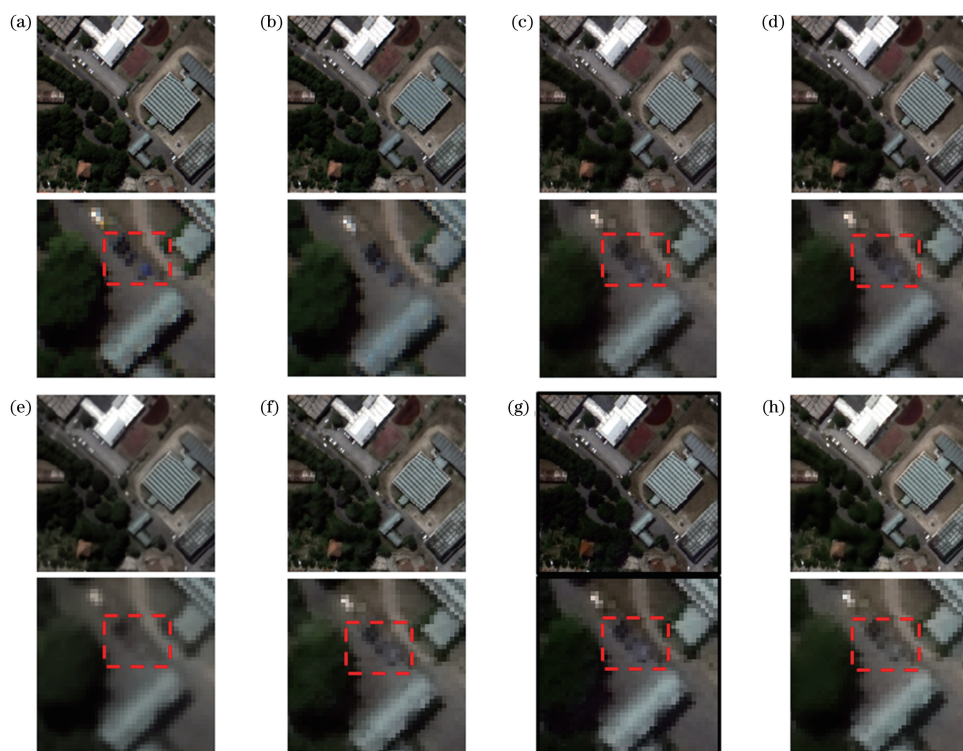


图 2 不同算法对 Pavia Center 数据集第 40 张影像的融合结果。(a)参考影像;(b)所提算法;(c) SFIM;  
(d) MTF-GLP-HPM;(e) Bayesian sparse;(f) GSA;(g) PCA;(h) CNMF

Fig. 2 Fusion results of 40th image in Pavia Center dataset by different algorithms. (a) Reference image;  
(b) proposed algorithm; (c) SFIM; (d) MTF-GLP-HPM; (e) Bayesian sparse; (f) GSA; (g) PCA; (h) CNMF

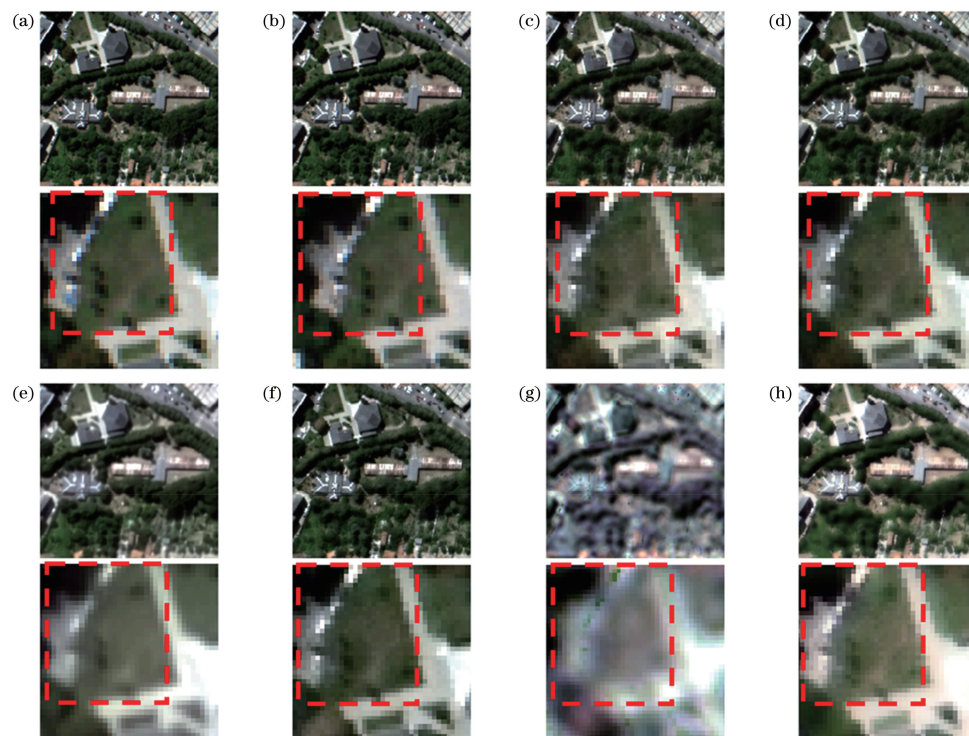


图 3 不同算法对 Pavia Center 数据集第 24 张影像的融合结果。(a)参考影像;(b)所提算法;(c) SFIM;  
(d) MTF-GLP-HPM;(e) Bayesian sparse;(f) GSA;(g) PCA;(h) CNMF

Fig. 3 Fusion results of 24th image in Pavia Center dataset by different algorithms. (a) Reference image;  
(b) proposed algorithm; (c) SFIM; (d) MTF-GLP-HPM; (e) Bayesian sparse; (f) GSA; (g) PCA; (h) CNMF

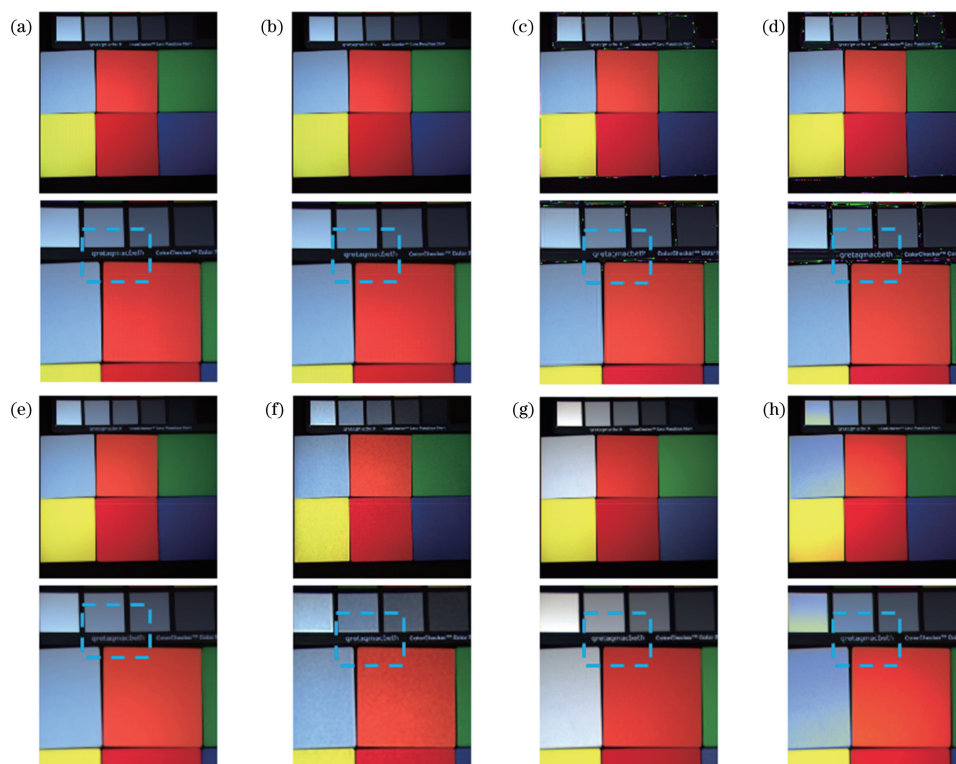


图 4 不同算法对 CAVE 数据集第 24 张影像的融合结果。(a)参考影像;(b)所提算法;(c) SFIM;  
(d) MTF-GLP-HPM;(e) Bayesian sparse;(f) GSA;(g) PCA;(h) CNMF

Fig. 4 Fusion results of 24th image in CAVE dataset by different algorithms. (a) Reference image;

(b) proposed algorithm; (c) SFIM; (d) MTF-GLP-HPM; (e) Bayesian sparse; (f) GSA; (g) PCA; (h) CNMF

注入不足;PCA 和 CNMF 的融合结果有明显的光谱失真问题,其中 PCA 的蓝色色块趋向于灰白色,而且放大区域的虚线框内字母颜色也接近灰白色,CNMF 的蓝色色块混有黄色;本文算法得到的融合影像在空间细节和光谱特性上都非常完善,十分接近参考影像。

如图 5 所示,该组影像使用的是 CAVE 数据集,选取的是第 30 张影像的[5,13,25]通道。该组影像包含色球和色块两部分内容。通过主观对比可以发现,SIFM、MTF-GLP-HPM 放大区域的纹理特征有很好的保真度,但是其余位置有明显的空间失真,注入细节过多;PCA 依然有光谱保真度问题,融合图右侧的蓝色色块变成了灰色;Bayesian spare、GSA 和 CNMF 这几个算法的光谱保真度很好,但是与参考影像相比,其虚线框内的纹理特征都没有表现出来,说明其空间细节注入不够;而本文算法的融合结果不仅在光谱上没有任何失真问题,而且在空间细节上也处理得很好。

### 3.5 客观分析

表 1 详细展示了 3.3 节中提到的四个指标在

Pavia Center 数据集上的评估结果。该数据是对所有测试集求取评估指标后求平均的结果。通过对比不难发现,PCA 算法的突出问题是 CC 指标低,说明其空间细节处理不足,但是其 RMSE 指标较佳,说明该算法与参考影像整体上还是很逼近的;Bayesian sparse 和 CNMF 在该数据集上表现相对平庸;MTF-GLP-HPM 和 SFIM 在 6 个对比算法中表现出了优异的性能,但是与本文提出的算法相比,其光谱保真度较低,而且全局误差也明显较高。

表 2 详细展示了 3.3 节中提到的四个指标在 CAVE 数据集上的评估结果。通过对比可以发现,在 Pavia Center 数据集上表现优异的 MTF-GLP-HPM 算法在 CAVE 数据集上的表现较为中庸;PCA 算法在该数据集上的评估结果明显欠佳;CNMF 算法的 CC 指标还是很不错的,但是其 SAM 指标明显比其他算法差,说明其光谱保真度是有问题的;而 Bayesian sparse 和 GSA 算法在该数据集上表现出良好的性能;但是从各指标上来看,本文提出的算法无论是在空间细节上还是在光谱保真度上依然是最优异的。



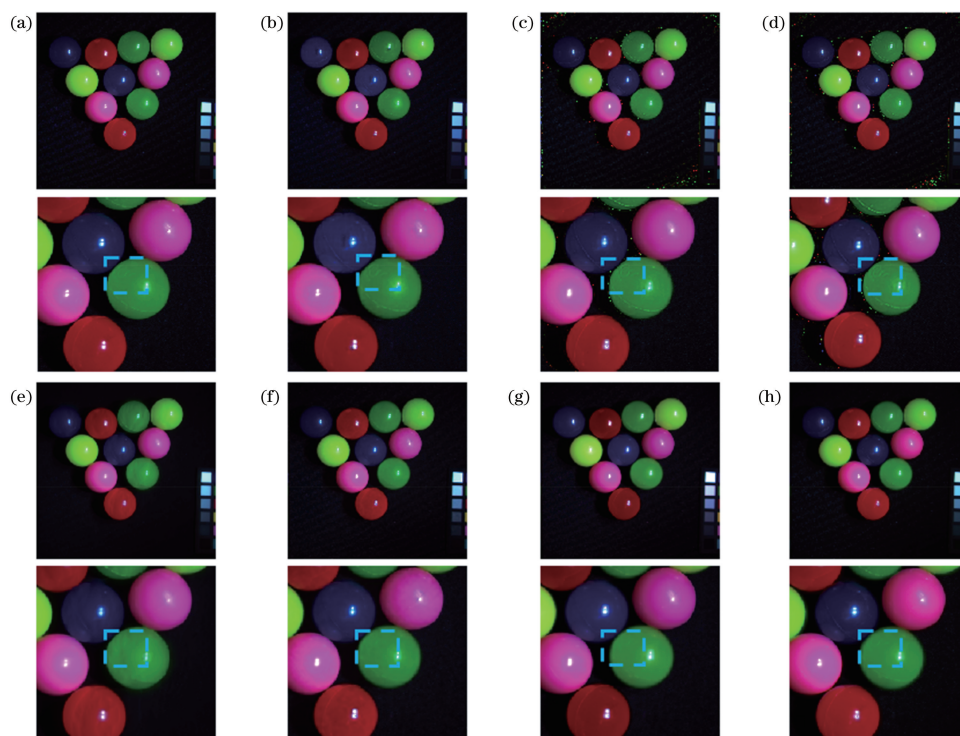


图 5 不同算法对 CAVE 数据集第 5 张影像的融合结果。(a)参考影像;(b)所提算法;(c) SFIM;  
(d) MTF-GLP-HPM;(e) Bayesian sparse;(f) GSA;(g) PCA;(h) CNMF

Fig. 5 Fusion results of 5th image in CAVE dataset by different algorithms. (a) Reference image; (b) proposed algorithm;  
(c) SFIM; (d) MTF-GLP-HPM; (e) Bayesian sparse; (f) GSA; (g) PCA; (h) CNMF

表 1 Pavia Center 数据集上的评估结果

Table 1 Evaluation results on Pavia Center dataset

Parameter	Bayesian sparse	CNMF	GSA	MTF-GLP-HPM	PCA	SFIM	Proposed algorithm
CC	0.9337	0.9251	0.9448	0.9675	0.9145	0.9631	0.9791
SAM	8.3296	7.1236	7.4254	6.2297	7.6162	6.2508	4.9153
RMSE	0.0363	0.0385	0.0317	0.0250	0.0370	0.0267	0.0185
ERGAS	5.7224	5.7865	4.9169	3.6290	5.7443	4.0159	2.8287

表 2 CAVE 数据集上的评估结果

Table 2 Evaluation results on CAVE dataset

Parameter	Bayesian sparse	CNMF	GSA	MTF-GLP-HPM	PCA	SFIM	Proposed algorithm
CC	0.9942	0.9925	0.9936	0.9884	0.9893	0.9895	0.9968
SAM	5.5619	6.6133	5.8892	6.5877	6.5012	6.1081	4.4402
RMSE	0.0140	0.0168	0.0163	0.0199	0.0230	0.0193	0.0123
ERGAS	3.6471	3.9827	3.8574	5.3817	4.1675	5.0852	2.6174

## 4 结 论

为了充分利用高光谱影像的空谱特征和全色影像丰富的空间特征,提出了一种基于多分支 CNN 的高光谱和全色影像融合方法。该方法构建了两个独立的卷积自编码分支网络,并逐级提取高光谱影像的空谱特征和全色影像的空间特征;设计了基于

空谱特征的高光谱影像重建网络,利用提取的全色影像和高光谱影像的多尺度空间和光谱,重建融合得到高空谱保真度的高光谱影像。对所提算法在 CAVE 数据集和 Pavia Center 数据集上进行了模拟数据实验,并对比和分析了所提算法与几种主流融合算法的融合性能。主观和客观的实验结果表明,所提算法与其他算法相比,能够在保持原始高光谱

影像丰富光谱信息的同时,更有效地提升空间性能,获得高空谱保真度的融合高光谱影像。下一步将研究所提出的融合算法在真实高光谱影像和全色影像数据上的有效性。

### 参 考 文 献

- [1] Zhang X H, Hao R F, Li T Y. Hyperspectral abnormal target detection based on low rank and sparse matrix decomposition-sparse representation [J]. *Laser & Optoelectronics Progress*, 2019, 56(4): 042801.  
张晓慧, 郝润芳, 李廷鱼. 基于低秩稀疏矩阵分解和稀疏字典表达的高光谱异常目标检测[J]. *激光与光电子学进展*, 2019, 56(4): 042801.
- [2] Liu X G, Zhao H J, Li N. Feature extraction based on multifractal spectrum for hyperspectral data [J]. *Acta Optica Sinica*, 2009, 29(3): 844-847.  
刘小刚, 赵慧洁, 李娜. 基于多重分形谱的高光谱数据特征提取[J]. *光学学报*, 2009, 29(3): 844-847.
- [3] Bi X J, Zhou Z Y. Hyperspectral image classification algorithm based on two-channel generative adversarial network [J]. *Acta Optica Sinica*, 2019, 39(10): 1028002.  
毕晓君, 周泽宇. 基于双通道 GAN 的高光谱图像分类算法[J]. *光学学报*, 2019, 39(10): 1028002.
- [4] Tu T M, Su S C, Shyu H C, et al. A new look at IHS-like image fusion methods [J]. *Information Fusion*, 2001, 2(3): 177-186.
- [5] Chavez P S, Jr, Kwarteng A Y. Extracting spectral contrast in Landsat Thematic Mapper image data using selective principal component analysis [J]. *Photogrammetric Engineering and Remote Sensing*, 1989, 55(3): 339-348.
- [6] Aiazzi B, Baronti S, Selva M. Improving component substitution pansharpening through multivariate regression of MS+Pan data [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2007, 45(10): 3230-3239.
- [7] Liu J G. Smoothing filter-based intensity modulation: a spectral preserve image fusion technique for improving spatial details [J]. *International Journal of Remote Sensing*, 2000, 21(18): 3461-3472.
- [8] Aiazzi B, Baronti S, Lotti F, et al. A comparison between global and context-adaptive pansharpening of multispectral images [J]. *IEEE Geoscience and Remote Sensing Letters*, 2009, 6(2): 302-306.
- [9] Liu X S, Xia W, Wang B, et al. An approach based on constrained nonnegative matrix factorization to unmix hyperspectral data [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2011, 49(2): 757-772.
- [10] Yokoya N, Yairi T, Iwasaki A. Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2012, 50(2): 528-537.
- [11] Akhtar N, Shafait F, Mian A. Bayesian sparse representation for hyperspectral image super resolution [C] // 2015 IEEE Conference on Computer Vision and Pattern Recognition, June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015: 15524066.
- [12] Masi G, Cozzolino D, Verdoliva L, et al. Pansharpening by convolutional neural networks [J]. *Remote Sensing*, 2016, 8(7): 594.
- [13] Dong C, Loy C C, He K M, et al. Learning a deep convolutional network for image super-resolution [M] // Fleet D, Pajdla T, Schiele B, et al. *Computer Vision-ECCV 2014. Lecture Notes in Computer Science*. Cham: Springer, 2014, 8692: 184-199.
- [14] Zheng Y X, Li J J, Li Y S, et al. Hyperspectral pansharpening using deep prior and dual attention residual network [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 58(11): 8059-8076.
- [15] Yang Y, Tu W, Huang S Y, et al. PCDRN: progressive cascade deep residual network for pansharpening [J]. *Remote Sensing*, 2020, 12(4): 676.
- [16] Hu J W, Du C G, Fan S S. Two-stage pansharpening based on multi-level detail injection network [J]. *IEEE Access*, 2020, 8: 156442-156455.