



## 基于混合双分支卷积神经网络和图卷积神经网络的全色锐化方法

王文卿, 张小乔, 何霁, 刘涵, 刘丁

引用本文:

王文卿, 张小乔, 何霁, 等. 基于混合双分支卷积神经网络和图卷积神经网络的全色锐化方法[J]. *智能系统学报*, 2025, 20(3): 649–657.

WANG Wenqing, ZHANG Xiaoqiao, HE Ji, et al. Pansharpening based on hybrid dual-branch convolutional and graph convolutional neural networks[J]. *CAA Transactions on Intelligent Systems*, 2025, 20(3): 649–657.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202401003>

## 您可能感兴趣的其他文章

### 4D卷积神经网络的自闭症功能磁共振图像分类

Classification of the functional magnetic resonance image of autism based on 4D convolutional neural network  
*智能系统学报*. 2021, 16(6): 1021–1029 <https://dx.doi.org/10.11992/tis.202009022>

### 样本仿真结合迁移学习的声呐图像水雷检测

Detection of underwater mine target in sidescan sonar image based on sample simulation and transfer learning  
*智能系统学报*. 2021, 16(2): 385–392 <https://dx.doi.org/10.11992/tis.202101030>

### 改进MobileNet的图像分类方法研究

Research on the improved image classification method of MobileNet  
*智能系统学报*. 2021, 16(1): 11–20 <https://dx.doi.org/10.11992/tis.202012034>

### 融合整体与局部信息的武夷岩茶叶片分类方法

Classification of Wuyi rock tealeaves by integrating global and local information  
*智能系统学报*. 2020, 15(5): 919–924 <https://dx.doi.org/10.11992/tis.202003018>

### 基于小样本学习的LCD产品缺陷自动检测方法

An automatic small sample learning-based detection method for LCD product defects  
*智能系统学报*. 2020, 15(3): 560–567 <https://dx.doi.org/10.11992/tis.201904020>

### 基于卷积神经网络的遥感图像分类研究

Classification of remote-sensing image based on convolutional neural network  
*智能系统学报*. 2018, 13(4): 550–556 <https://dx.doi.org/10.11992/tis.201706078>

DOI: 10.11992/tis.202401003

网络出版地址: <https://link.cnki.net/urlid/23.1538.tp.20250411.1127.002>

# 基于混合双分支卷积神经网络和图卷积神经网络的全色锐化方法

王文卿<sup>1,2</sup>, 张小乔<sup>1</sup>, 何霁<sup>1</sup>, 刘涵<sup>1,2</sup>, 刘丁<sup>1,2</sup>

(1. 西安理工大学自动化与信息工程学院, 陕西 西安 710048; 2. 西安理工大学陕西省复杂系统控制与智能信息处理重点实验室, 陕西 西安 710048)

**摘要:** 多光谱图像全色锐化是遥感影像处理与解译领域的热点问题。相较于传统全色锐化方法, 基于深度学习的全色锐化方法聚焦于图像深层次特征的提取, 大幅提升了融合图像的质量。本文提出一种基于混合双分支卷积神经网络和图卷积神经网络的全色锐化方法, 旨在同时挖掘图像的光谱、空间与非几何结构信息, 以提升融合图像空间分辨率和光谱分辨率。本方法建立在多分辨率分析融合框架的基础上, 利用深度神经网络构建了特征提取、特征融合和图像重构模块。混合双分支网络模块是由 2D 和 3D 卷积神经网络构建, 其中, 2D 卷积神经网络负责挖掘多光谱图像与全色图像的空间特征, 3D 卷积神经网络负责挖掘图像的光谱特征。引入了图卷积神经网络以捕捉图像图结构中节点的空间关系, 从而整合非局部信息。将多光谱图像与全色图像的空间、光谱和非几何特征通过特征融合模块进行融合。将融合特征输入图像重构网络重建高质量多光谱图像。本文算法在 GeoEye-1 和 IKONOS 遥感数据上进行了实验验证, 实验结果表明: 与其他方法相比, 本文算法在主观视觉和客观评价指标上均表现出优秀性能。

**关键词:** 图像融合; 遥感; 图像处理; 深度学习; 卷积神经网络; 机器学习; 特征提取; 图像重构**中图分类号:** TP751 **文献标志码:** A **文章编号:** 1673-4785(2025)03-0649-09

中文引用格式: 王文卿, 张小乔, 何霁, 等. 基于混合双分支卷积神经网络和图卷积神经网络的全色锐化方法[J]. 智能系统学报, 2025, 20(3): 649-657.

英文引用格式: WANG Wenqing, ZHANG Xiaoqiao, HE Ji, et al. Pansharpening based on hybrid dual-branch convolutional and graph convolutional neural networks[J]. CAAI transactions on intelligent systems, 2025, 20(3): 649-657.

## Pansharpening based on hybrid dual-branch convolutional and graph convolutional neural networks

WANG Wenqing<sup>1,2</sup>, ZHANG Xiaoqiao<sup>1</sup>, HE Ji<sup>1</sup>, LIU Han<sup>1,2</sup>, LIU Ding<sup>1,2</sup>

(1. School of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048, China; 2. Shaanxi Key Laboratory of Complex System Control and Intelligent Information Processing, Xi'an University of Technology, Xi'an 710048, China)

**Abstract:** The pansharpening of multispectral images represents a trending research topic in remote sensing image processing and interpretation. Moreover, compared with traditional pansharpening methods, deep learning-based pansharpening methods mainly extract deep features, thereby greatly improving the quality of fused images. Here, a method based on hybrid dual-branch convolutional neural network (CNN) and graph convolutional neural network (GCNN) is proposed to simultaneously extract spectral information, spatial information, and non-geometric structural information and improve the spatial and spectral resolutions of fused images. This hybrid method comprises the construction of a multi-resolution analysis fusion framework, followed by the construction of a feature extraction module, a feature fusion module, and an image reconstruction module based on deep neural networks. First, the hybrid dual-branch network module was constructed using 2D and 3D CNNs that focus on extracting spatial and spectral features, respectively. Second, GCNN was introduced to capture the spatial relationships of the nodes in the graph structure of the image and integrate non-local information. Afterward, the spatial, spectral, and non-geometric features extracted from multispectral and panchromatic images were fused by the feature fusion module. Finally, the fused features were input into the image reconstruction network to reconstruct the high-quality multispectral images. The proposed method was experimentally validated using GeoEye-1 and IKONOS remote sensing data. Compared with other methods, the experimental results obtained by the proposed method reveal its excellent performance in subjective and objective vision evaluations.

**Keywords:** image fusion; remote sensing; image processing; deep learning; convolutional neural network; machine learning; feature extraction; image reconstruction

收稿日期: 2024-01-02. 网络出版日期: 2025-04-11.

基金项目: 国家自然科学基金项目(62376214, 92270117); 陕西省自然科学基金项目(2023-JC-YB-533).

通信作者: 刘涵, E-mail: [liuhan@xaut.edu.cn](mailto:liuhan@xaut.edu.cn).

随着 IKONOS、GeoEye-1、高分五号等各类遥感卫星的成功发射, 遥感影像数据呈现爆发式增长, 在国民经济和国防建设等领域的作用日益凸

显。卫星传感器自身成像机制的局限,导致其无法使同一幅遥感图像兼具高空间分辨率和高光谱分辨率<sup>[1]</sup>。然而,各类遥感影像处理与解译任务,如目标检测<sup>[2]</sup>、变化监测<sup>[3]</sup>、地物分类<sup>[4]</sup>等,对高空间分辨率多光谱 (high resolution multispectral, HRMS) 图像具有迫切需求。因此,全色锐化技术应运而生。该技术是将具有高空间分辨率的全色 (Panchromatic, PAN) 图像和低空间分辨率的多光谱 (low resolution multispectral, LRMS) 图像融合,生成兼具高空间分辨率和高光谱分辨率的图像,即 HRMS 图像<sup>[5]</sup>。

随着信息处理技术的不断发展,诸多全色锐化方法被提出。传统的全色锐化方法主要分为 3 类:分量替换、多分辨率分析和变分优化<sup>[6-7]</sup>。分量替换方法的核心思想是将 LRMS 图像投影到其他特征空间,并利用直方图匹配后的 PAN 图像替换其第一主分量,然后通过逆变换得到融合图像。其主要代表方法是强度-色调-饱和度<sup>[8]</sup>和主成分分析<sup>[9]</sup>。虽然分量替换方法能保持融合图像的空间分辨率,但 LRMS 和 PAN 图像之间的局部差异会导致融合图像的光谱失真。多分辨率分析方法采用小波变换<sup>[10]</sup>、轮廓波变换等多分辨率技术,将从 PAN 图像提取的空间信息注入 LRMS 图像。与分量替换方法相比,该类方法能更好地保留光谱信息,但会产生局部区域的空间失真。变分优化方法将融合过程视为图像恢复问题,利用先验正则化约束来构造变分融合能量函数,通过最小化能量函数来重建融合图像。常见的变分优化方法有低秩先验<sup>[11]</sup>、稀疏先验<sup>[12]</sup>和拉普拉斯先验<sup>[13]</sup>。

近年来,深度学习方法因其突出的数据处理和计算能力在全色锐化领域广受青睐。基于深度学习的全色锐化方法分为两类:基于卷积神经网络 (convolutional neural network, CNN) 和基于生成对抗网络 (generative adversarial networks, GAN) 的方法。文献 [14] 创造性地将 CNN 应用于全色锐化领域。该方法首先对上采样 LRMS 图像和 PAN 图像进行拼接,再将拼接后的特征输入三层网络进行融合。文献 [15] 提出了一种基于双流分支网络 (two-stream fusion network, TFNet) 的方法,其使用两个子网络分别对 LRMS 和 PAN 图像进行特征提取,随后经过特征融合与图像重建网络得到融合图像。此外,文献 [16] 提出了一种用于全色锐化的深度网络 PanNet。该网络在高通滤波域训练网络参数,将上采样多光谱图像添加到 CNN 的输出,从而将光谱信息直接传递至重建图像。基于 GAN 的全色锐化方法需要将 PAN 和 LRMS 作为生成器的输入,再将其映射到所需的

HRMS 图像,同时利用鉴别器实现对抗训练生成 HRMS 图像。文献 [17] 提出了一种 PSGAN (GAN for pansharpening) 方法,该方法在 GAN 中添加了双流输入,通过堆叠输入和注意力机制,有效生成高质量的 HRMS 图像。

3D CNN 是在 2D CNN 基础上衍生而来,其考虑多张图像的关联信息或视频的时序信息,引入了新的维度信息,能够提取 3 个维度的联合特征。在遥感图像领域,3D CNN 被广泛应用到多光谱/高光谱图像分类<sup>[18-19]</sup>。高光谱图像数据的光谱信息在分类中起到重要作用。3D CNN 通过利用波段间的相关特性,能够挖掘光谱波段的联合空谱特征,从而提升分类性能。受此启发,本文利用 3D CNN 来提取图像的光谱特征信息。

近年来,图卷积神经网络 (graph convolutional neural network, GCNN)<sup>[20]</sup> 在诸多领域也颇受关注。GCNN 的现有应用大多聚焦于图形数据,例如社交网络、物理系统和知识图<sup>[21]</sup>。其主要优点是能够充分表达信息之间的传播和数据的交互。文献 [22] 提出了 GCResNet, 该网络将 GCNN 应用到图像去模糊领域,使性能得到显著提升。文献 [23] 提出了 GCPNet, 该网络将知识蒸馏与 GCNN 相结合,设计了用于 LRMS 和 PAN 融合的光谱图卷积神经网络模块和空谱图卷积神经网络模块,取得了较好的融合结果。这两个模块分别捕捉了全色图像和多光谱图像的光谱特征和空间信息,从而得到较好的融合结果,展现了 GCNN 在提取非局部结构信息方面的优势。

考虑到 2D CNN、3D CNN 和 GCNN 能够挖掘遥感图像不同层次的结构特征,本文提出了一种基于混合双分支卷积神经网络和图卷积神经网络的全色锐化方法,通过利用 2D CNN、3D CNN 和 GCNN 挖掘 LRMS 图像与 PAN 图像的空间、光谱与非局部结构特征。本文主要贡献如下:

- 1) 以深度神经网络为基础,构建了特征提取、特征融合和图像重构 3 个网络模块,增强了融合框架的可解释性。
- 2) 通过结合 2D CNN 和 3D CNN,构建了混合双分支特征提取网络,分别从多光谱图像与全色图像中提取出空间信息和光谱特征。
- 3) 引入了 GCNN,有效地捕捉了图像非局部空间关系,为融合图像重建提供了更丰富的特征。

## 1 本文算法

### 1.1 模型整体框架

本文方法的网络框架如图 1 所示。

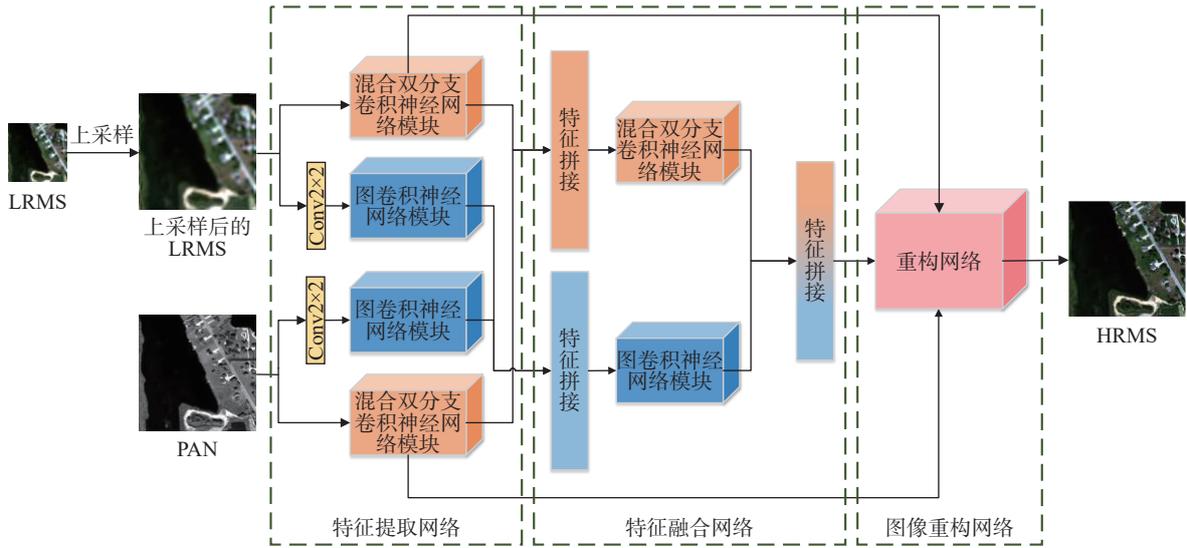


图 1 本文算法框架

Fig. 1 Framework of the proposed method

图 1 该网络主要包括 3 个阶段: 特征提取阶段、特征融合阶段与图像重构阶段。其中, 特征提取和特征融合充当编码器, 从输入 PAN 和 LRMS 图像中提取和编码特征; 图像重构阶段要对融合后的特征进行解码, 重建出所需的 HRMS 图像。具体实现细节描述如下:

首先, 将 LRMS 图像上采样至与 PAN 图像相同尺寸, 并将 PAN 与上采样的 LRMS 图像的每一个波段进行直方图匹配。然后, 将直方图匹配后的 PAN 图像在通道方向进行拼接, 并定义为  $P_H$ 。最后, 以上采样的 LRMS 图像和直方图匹配后的 PAN 图像作为网络输入。

特征提取网络, 采用两个网络模块分别对上采样的 LRMS 图像和直方图匹配后的 PAN 图像进行提取特征。第 1 个网络模块为混合双分支卷积神经网络。在该网络中, 2D CNN 由两个  $3 \times 3$  的卷积核和一个  $2 \times 2$  的卷积核构成, 旨在从两幅图像中提取出空间特征。3D CNN 由两个  $3 \times 3 \times 3$  的卷积核和一个  $1 \times 2 \times 2$  的卷积核构成, 旨在提取两幅图像的光

谱信息。第 2 个网络模块为 GCNN 模块。该模块利用了 GCNN 在捕捉节点中丰富的上下文空间关系和整合非局部信息方面的能力, 以提取多光谱图像和全色图像中的非结构化特征。最后, 将两个不同源图像的特征拼接, 送入特征融合网络。

在特征融合网络中, 将 LRMS 和 PAN 两幅图像由特征提取网络的混合双分支卷积神经网络模块提取出的特征进行拼接, 输入特征融合网络的混合双分支卷积神经网络模块。同时, 将两幅图像由特征提取网络的图卷积神经网络模块提取出的特征进行拼接, 输入特征融合网络中的图卷积神经网络模块。然后, 将特征融合网络输出的两个特征进行整合拼接, 以获取最终融合特征图, 作为图像重构网络的输入。图像重构网络如图 2(c) 所示。考虑到随着网络层数的增加, 卷积后获取的高级特征会失去原始图像的纹理细节, 因此在图像重构网络中使用跳跃连接, 将包含图像原始细节纹理的低级特征与高级特征连接, 以弥补特征融合过程中的原始细节丢失。

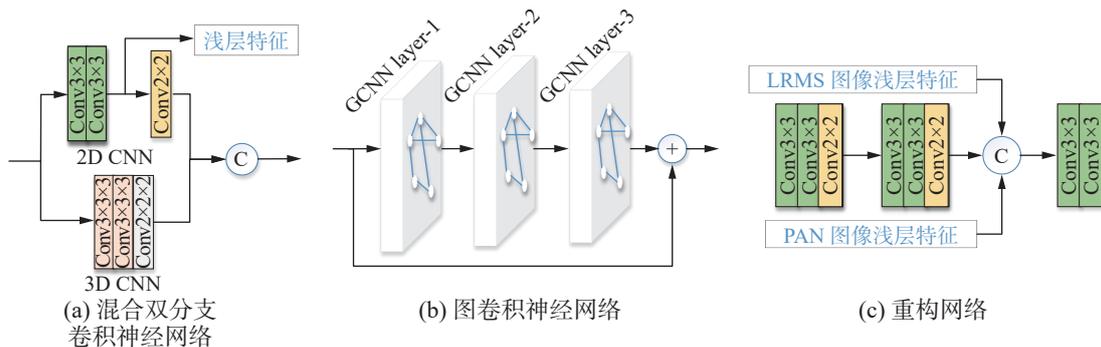


图 2 网络模块结构示意图

Fig. 2 Schematic diagram of network module structure

### 1.2 混合双分支卷积神经网络

本文算法提出的混合双分支卷积神经网络模块具体结构如图 2(a) 所示, 由 2D CNN 和 3D CNN 网络构成。其中, 2D CNN 专注于提取 LRMS 和 PAN 图像中的空间信息, 而 3D CNN 主要用于提取图像中的光谱信息。3D CNN 具有直接处理多通道数据的能力, 表明其可以利用跨光谱通道的卷积操作来捕获不同波段之间的相关性和特征。3D CNN 通过在不同光谱维度上进行卷积操作, 同时考虑来自多个光谱波段的信息, 从而更全面地捕获图像中的光谱特征。该模块应用于特征提取和特征融合两个阶段, 在特征提取阶段, 其主要作用是分别从 LRMS 和 PAN 图像中挖掘光谱特征与空间信息, 并将提取的特征拼接后输入特征融合网络进行进一步融合。模块中的 2D CNN 和 3D CNN 均由 3 个卷积层构成, 其中 2D CNN 第  $l$  层的输出特征图可以表示为

$$\begin{cases} \mathbf{F}_{P_{2D}}^{(l)} = f_{2D}^{(l)}(\mathbf{P}_H) \\ \mathbf{F}_{M_{2D}}^{(l)} = f_{2D}^{(l)}(\tilde{\mathbf{M}}) \end{cases} \quad (1)$$

式中:  $\mathbf{P}_H$  为 PAN 图像,  $\tilde{\mathbf{M}}$  为上采样的 LRMS 图像。  $f_{2D}^{(l)}(\cdot)$  为 2D CNN 第  $l$  层,  $\mathbf{F}_{P_{2D}}^{(l)}$  为 PAN 图像经过  $l$  层 2D CNN 后的特征,  $\mathbf{F}_{M_{2D}}^{(l)}$  为 LRMS 图像经过  $l$  层 2D CNN 后的特征。3D CNN 第  $l$  层的输出特征图可以表示为

$$\begin{cases} \mathbf{F}_{P_{3D}}^{(l)} = f_{3D}^{(l)}(\mathbf{P}_H) \\ \mathbf{F}_{M_{3D}}^{(l)} = f_{3D}^{(l)}(\tilde{\mathbf{M}}) \end{cases} \quad (2)$$

式中:  $f_{3D}^{(l)}(\cdot)$  为 3D CNN 的第  $l$  层,  $\mathbf{F}_{P_{3D}}^{(l)}$  为 PAN 图像经过  $l$  层 3D CNN 后的特征,  $\mathbf{F}_{M_{3D}}^{(l)}$  表示 PAN 图像经过  $l$  层 3D CNN 后的特征。LRMS 和 PAN 图像经过第 1 个混合双分支卷积神经网络模块后的特征图记为  $\mathbf{F}_P$  和  $\mathbf{F}_M$ , 经过三层卷积后, 则有

$$\begin{cases} \mathbf{F}_P = \text{ConCat}(f_{2D}^{(3)}(\mathbf{P}_H), f_{3D}^{(3)}(\mathbf{P}_H)) \\ \mathbf{F}_M = \text{ConCat}(f_{2D}^{(3)}(\tilde{\mathbf{M}}), f_{3D}^{(3)}(\tilde{\mathbf{M}})) \end{cases} \quad (3)$$

式中 ConCat 为将特征在通道方向上拼接。

### 1.3 图卷积神经网络

图卷积神经网络作为一种独特的神经网络结构, 与常见的卷积神经网络存在显著差异。它通过图卷积的方式, 有效地从图数据中提取关键特征, 进而实现节点分类、图分类以及边预测等任务。图卷积可以定义为简单形式<sup>[20]</sup>:

$$\mathbf{Z} = \hat{\mathbf{A}}\mathbf{F}\boldsymbol{\Theta} \quad (4)$$

式中:  $\mathbf{Z}$  是卷积信号矩阵,  $\hat{\mathbf{A}}$  是邻接矩阵,  $\boldsymbol{\Theta}$  是滤波器参数矩阵。受到文献 [23-24] 的启发, 全局网络也可以近似地看作图卷积神经网络。全局网络的结构如图 3 所示: 输入特征先分别经过  $\alpha(\cdot)$ 、 $\beta(\cdot)$  和  $\gamma(\cdot)$  3 个  $1 \times 1$  的卷积核, 通过对图像中所有位置的像素进行加权平均, 从而捕捉图像中远距离的关系; 再通过重塑操作对输入张量的维度重新排列或重新组合, 以计算全局范围内像素间的关联性。

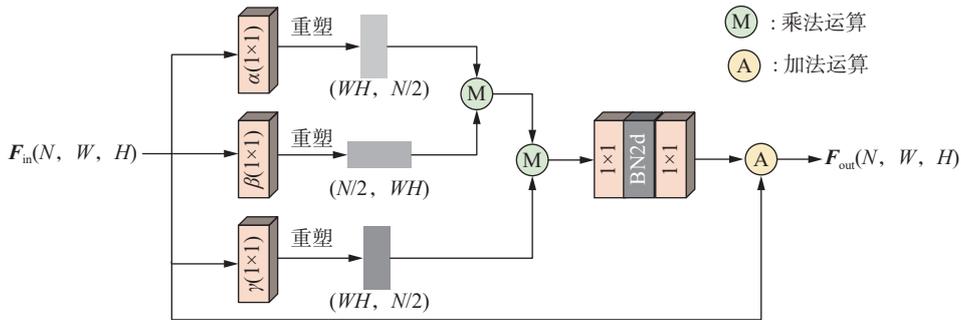


图 3 图卷积神经网络结构

Fig. 3 Graph convolutional neural network structure

重塑后的特征由输入特征  $\mathbf{F}_{in} \in \mathbf{R}^{N \times W \times H}$  变成  $\mathbf{F}_\alpha \in \mathbf{R}^{WH \times (N/2)}$ 、 $\mathbf{F}_\beta \in \mathbf{R}^{(N/2) \times WH}$  和  $\mathbf{F}_\gamma \in \mathbf{R}^{WH \times (N/2)}$ , 其中  $\mathbf{F}_\alpha$ 、 $\mathbf{F}_\beta$  和  $\mathbf{F}_\gamma$  是输入特征经过卷积层后再重塑的特征。 $\mathbf{F}_\alpha$  与  $\mathbf{F}_\beta$  做乘法运算, 实现通道间的交互和融合。通过元素级的乘法, 可以实现两个特征图之间的像素级别交互, 从而加强特征的表达。随后, 经过 Softmax 激活后, 再与  $\mathbf{F}_\gamma$  相乘, 进行进一步的通道交互。交互后的特征图像再经过两个  $1 \times 1$  的卷积层和一个 BN(batch normalization) 层之后, 通过重塑将特征输出, 记为  $\mathbf{F}_{out}$ 。此时

$\mathbf{F}_{out} \in \mathbf{R}^{N \times W \times H}$  与原输入  $\mathbf{F}_{in}$  大小一致。上述过程可表示为

$$\mathbf{F}_{out} = \alpha(\mathbf{F}_{in})\beta^T(\mathbf{F}_{in})\gamma(\mathbf{F}_{in})\boldsymbol{\Theta} \quad (5)$$

将  $\alpha(\mathbf{F}_{in})\beta^T(\mathbf{F}_{in})$  视作邻接矩阵  $\hat{\mathbf{A}}$ , 则上文描述的运算机理可近似地看作图卷积过程。

### 1.4 损失函数

根据文献 [25-26] 中提出的合成协议和一致性协议, 本文认为融合后的 HRMS 图像的光谱应尽可能与 LRMS 图像保持一致。本文将 HRMS 和输入 LRMS 之差的  $L_1$  范数作为光谱损失函数  $L_{\text{Spectral}}$ ,

其定义为

$$L_{\text{Spectral}} = \left\| \hat{M} - \tilde{M} \right\|_1 \quad (6)$$

其中  $\hat{M}$  表示 HRMS 图像。

为保持空间细节尽可能与原始 PAN 图像  $P$  一致, 本文用 HRMS 图像和 PAN 图像之差的  $L_1$  范数作为空谱损失函数  $L_{\text{Spatial}}$ , 其定义为

$$L_{\text{Spatial}} = \left\| \hat{M} - P \right\|_1 \quad (7)$$

除了光谱约束和空谱约束, 本文还考虑到了图像的边缘细节纹理, 用 HRMS 和 PAN 梯度之差的  $L_1$  范数来约束, 其定义为

$$L_{\text{grad}} = \left\| \nabla \hat{M} - \nabla P \right\|_1 \quad (8)$$

式中  $\nabla \hat{M}$  和  $\nabla P$  分别表示 HRMS 和 PAN 图像的梯度。本文设计的损失函数:

$$L = \mu L_{\text{Spectral}} + \nu L_{\text{Spatial}} + L_{\text{grad}} \quad (9)$$

式中  $\mu$  和  $\nu$  分别是光谱损失和空谱损失的权重。

第 2 章的消融实验结果显示, 当  $\mu = 0.1, \nu = 1$  时, 模型效果达到最好。

## 2 实验结果与分析

### 2.1 数据集

为评估所提模型的性能, 本文选择来源于 IKONOS 和 GeoEye-1 两个卫星传感器的数据集对模型进行测试。其中 LRMS 图像包括红、绿、蓝和近红外 4 个波段, PAN 图像仅具有单个波段。两类卫星图像的辐射分辨率均为 11 bits。两类卫星的主要特性如表 1 所示。LRMS 图像大小为  $200 \times 200$ 。PAN 图像的大小为  $800 \times 800$ 。图像数据被随机分配为训练集和测试集, 分配比例为 9:1。在训练过程中, 训练集裁成尺寸为  $8 \times 8$  和  $32 \times 32$  大小的块, 进而增加训练样本的数量。具体划分情况见表 2。

表 1 两类遥感卫星特性

Table 1 Characteristics of two remote sensing satellites

卫星	光谱范围/nm					空间分辨率/m	
	蓝	绿	红	近红外	PAN	PAN	LRMS
IKONOS	450 ~ 530	520 ~ 610	640 ~ 720	760 ~ 860	450 ~ 900	1.0	4
GeoEye-1	450 ~ 510	510 ~ 580	655 ~ 690	780 ~ 920	450 ~ 900	0.5	2

表 2 数据集划分

Table 2 Division of the dataset

卫星	源图像数量/组	数据类型	组数	裁剪后的组数
IKONOS	240×(200×200,800×800)	训练集	216	13 824×(8×8,32×32)
		测试集	24	24×(50×50,200×200)
GeoEye1	250×(200×200,800×800)	训练集	225	14 400×(8×8,32×32)
		测试集	25	25×(50×50,200×200)

### 2.2 实验环境

本文实验在 Python 3.8 环境下使用 PyTorch 框架进行训练和测试。硬件实验平台 GPU 为 24 GB RTX 3090。最大迭代次数设置为 6 000 次, 批大小设置为 256。初始学习率设置为 0.001, 并每隔 2 000 次迭代衰减 50%。空间损失项和光谱损失项的权重系数分别设置为 0.1 和 1。在 IKONOS 和 GeoEye-1 数据集上训练时长约为 2 h。

### 2.3 评价指标

为评估所提方法的优越性, 本文选取了 6 种具有代表性的评价指标, 分别为: 光谱角映射 (spectral angle mapper, SAM)、合成无量纲全局相对误差 (erreur relative globale adimensionnelle synthèse, ERGAS)、相对平均光谱误差 (relative average spectral error, RASE)、空间相关系数 (spatial correlation coefficient, SCC)、通用图像质量指标

$Q$ (universal image quality)、结构相似性 (structural similarity, SSIM)<sup>[23]</sup> 指标。

### 2.4 实验结果与分析

本文选取 6 种算法作为对比, 分别是 4 种传统方法和 2 种深度学习方法。传统的对比算法包括 MTF-GLP(generalized Laplacian pyramid with an MTF-matched filter)<sup>[27]</sup>、GS2\_GLP(Gram-Schmidt mode 2 algorithm with generalized Laplacian pyramid)<sup>[28]</sup>、BDS-PC(robust band dependent spatial detail with physical constraints)<sup>[29]</sup>、AWLP\_H(additive wavelet luminance proportional with haze correction)<sup>[30]</sup>。基于 CNN 的方法包括 TFNet<sup>[15]</sup> 和 PanNet<sup>[16]</sup>。

图 4 给出了 GeoEye-1 数据测试图像在不同算法下的融合结果。从图中可以看出, GS2\_GLP 产生了一定的光谱失真, 主要体现在图像下方的植被区域。BDS-PC 方法的融合图像产生了严重的

光谱失真, 尤其在水域和植被区域, 且植被区域出现了过度饱和。AWLP-H 及 MTF-GLP 方法的融合图像相比于前两种方法有更好的光谱质量, 但存在一定程度的伪影, 如绿框所标区域。相比于传统算法, TFNet 和 PanNet 方法的融合图像在光谱和空间上都更接近于参考图像, 但依然存在一些光谱和空间失真。图 4(h) 和 (i) 正中间上方水域和陆地的交接

处存在少许空间细节失真。表 3 给出了 24 幅测试图像在不同融合方法下的平均客观评价指标。从表中可以看出, 传统方法的评价指标数值均低于深度学习算法。本文算法在 SAM、RASE、SCC 和 SSIM 4 个指标上具有最佳数值结果, 而在其他 2 个指标具有次最佳数值结果。其中, RASE 指标较其他方法降低了 0.44。

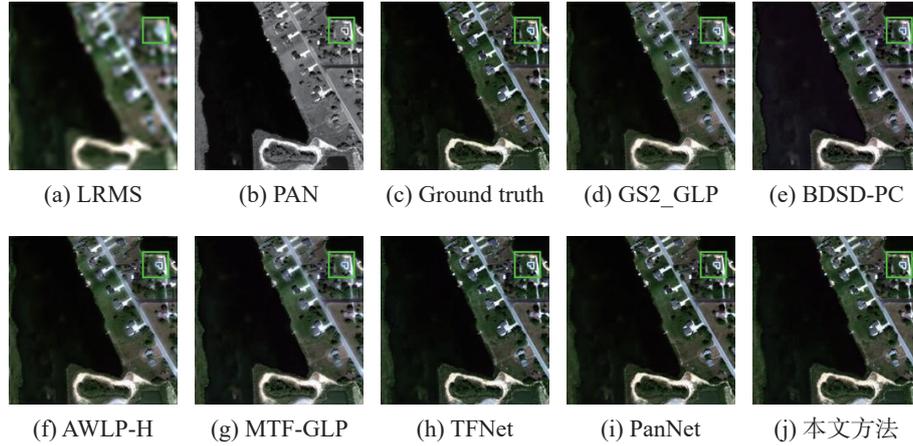


图 4 GeoEye-1 数据集上 7 种融合方法的融合结果

Fig. 4 Fusion results of seven fusion methods on the GeoEye-1 dataset

表 3 GeoEye-1 数据集上 7 种融合方法的实验结果评价指标

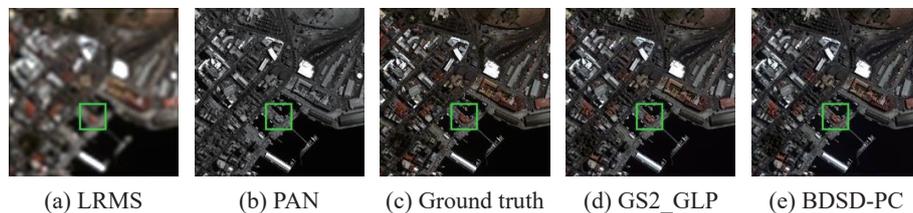
Table 3 Evaluation indexes for experimental results of seven fusion methods on the GeoEye-1 dataset

方法	SAM↓	ERGAS↓	RASE↓	SCC↑	$Q\uparrow$	SSIM↑
GS2_GLP	3.2744	2.8539	13.3318	0.8168	0.8206	0.8929
BSDS-PC	3.0881	2.5465	11.8565	0.8794	0.8601	0.9222
AWLP-H	2.4347	2.3470	10.8151	0.9124	0.8789	0.9427
MTF-GLP	2.9326	2.7003	12.2147	0.8766	0.8431	0.9186
TFNet	1.6539	<b>1.6880</b>	7.6106	0.9499	<b>0.9137</b>	0.9633
PanNet	1.6587	1.7261	7.7400	0.9468	0.9070	0.9621
本文方法	<b>1.6339</b>	1.6998	<b>7.1722</b>	<b>0.9571</b>	0.9083	<b>0.9655</b>

注: 最佳数值用黑体标出。

IKONOS 数据集的融合结果如图 5 所示, 其中绿框所标区域能够更好地展示结果差异。GS2\_GLP 方法的融合图像产生了较为严重的光谱失真, 红色建筑物区域尤为明显。同时部分边缘和纹理信息很模糊, 表明融合图像存在空间失真。BSDS-PC 方法的融合图像同样产生了较严重的光谱失真和空间失真。AWLP-H 和 MTF-GLP 算法的融合图像存在轻微模糊且光谱失真较严重。相较于传统方法, 基于 CNN 的算法在光谱特

征和空间信息上更接近于参考图像, 但本文算法具有最小的光谱失真和空间失真。表 4 给出了 IKONOS 数据集 25 幅测试图像在不同融合方法下的客观评价指标平均值。从表中可以看出, 传统方法相比基于 CNN 的方法明显处于劣势。本文方法的评价指标在所有算法中具有最佳数值结果, 其中 RASE 指标下降尤为明显。实验表明, 本文方法有效地提升了融合图像的光谱质量和空谱质量。



(a) LRMS (b) PAN (c) Ground truth (d) GS2\_GLP (e) BSDS-PC

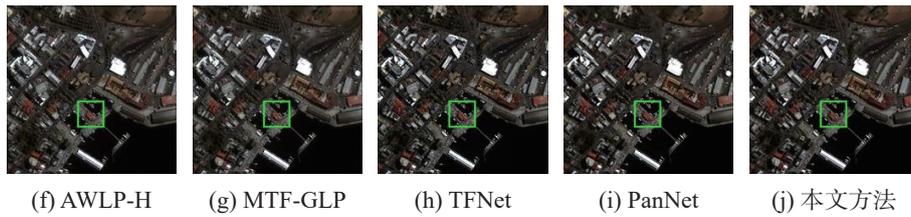


图 5 IKONOS 数据集上 7 种融合方法的融合结果

Fig. 5 Fusion results of seven fusion methods on the IKONOS dataset

表 4 IKONOS 数据集上 7 种融合方法的实验结果评价指标

Table 4 Evaluation indexes for experimental results of seven fusion methods on the IKONOS dataset

方法	SAM↓	ERGAS↓	RASE↓	SCC↑	Q↑	SSIM↑
GS2_GLP	3.8824	2.6749	10.7669	0.8818	0.7867	0.9070
BDS-PC	3.8636	2.6820	10.7902	0.8885	0.7965	0.9117
AWLP-H	3.5207	2.6405	10.6203	0.8931	0.8098	0.9201
MTF-GLP	4.0959	2.7863	11.7074	0.8716	0.7747	0.8959
TFNet	3.0225	2.2030	8.9109	0.9298	0.8392	0.9401
PanNet	3.0675	2.3449	9.5166	0.9130	0.8287	0.9330
本文方法	<b>2.9291</b>	<b>2.1194</b>	<b>8.6041</b>	<b>0.9347</b>	<b>0.8414</b>	<b>0.9433</b>

注: 最佳数值用黑体标出。

### 2.5 消融实验

本文采用混合双分支网络和 GCNN 对图像特征进行提取。为了验证所提网络的有效性, 本文做了相关消融实验。在超参数相同的条件下, 分别去除本文网络中的 3D CNN 和 GCNN 两个模块进行对比。表 5 是 GeoEye-1 数据集 24 幅测试图

像的平均评价指标。从表中数据可以看出, 去除 GCNN 后, SAM、ERGAS 和 RASE 3 项指标明显变差, 说明 GCNN 在提升光谱质量上有显著效果。网络在去除 3D CNN 后, 各项指标明显变差, 表明 3D CNN 不仅改善了光谱特征, 更是增强了融合图像的空间结构信息。

表 5 GeoEye-1 数据集网络框架的消融实验结果评价指标

Table 5 Evaluation indexes for ablation experimental results of network framework on the GeoEye-1 dataset

2D CNN	3D CNN	GCNN	SAM↓	ERGAS↓	RASE↓	SCC↑	Q↑	SSIM↑
√	√	×	1.7514	1.7259	7.3086	0.9547	0.9019	0.9628
√	×	√	2.6823	2.6057	10.9639	0.8882	0.8457	0.9370
√	√	√	<b>1.6339</b>	<b>1.6998</b>	<b>7.1722</b>	<b>0.9571</b>	<b>0.9083</b>	<b>0.9655</b>

此外, 本文还针对损失函数进行了消融实验。在设计损失函数时, 在光谱和空间约束前加

了权重。如表 6 所示, 当光谱损失权重为 0.1, 空谱损失和梯度损失的权重分别为 1 时, 融合效果达到最好。

表 6 GeoEye-1 数据集损失函数权重的消融实验结果评价指标

Table 6 Evaluation indexes for ablation experimental results of loss function weights on the GeoEye-1 dataset

$L_{grad}$	$L_{Spatial}$	$L_{Spectral}$	SAM↓	ERGAS↓	RASE↓	SCC↑	Q↑	SSIM↑
1	1	1	3.1216	3.8354	18.1520	0.6975	0.6959	0.8382
1	0.1	1	3.5726	4.2870	18.3653	0.6977	0.6911	0.8349
1	1	0.1	<b>1.6339</b>	<b>1.6998</b>	<b>7.1722</b>	<b>0.9571</b>	<b>0.9083</b>	<b>0.9655</b>

### 3 结束语

本文结合 2D CNN、3D CNN 和 GCNN 3 种网

络结构, 针对 LRMS 和 PAN 图像的特性, 提出了一种基于混合双分支卷积神经网络和图卷积神经网络的全色锐化方法。混合双分支网络用于提取

图像的光谱特征与空间特征;其中,2D CNN 在提取空谱特征的同时降低了网络的参数量;3D CNN 有效地提取图像的光谱特征,增强了融合图像的色彩信息。同时,本文利用跳跃连接,将低层特征与高层特征相连接,更好地补充了高层特征损失的细节信息。在非欧氏空间中,GCNN 利用像素上下文语义关联,更好地联系了图像的全局信息,学习更丰富、更抽象的非结构化特征。损失函数的设计也充分考虑了光谱和空谱的一致性,有效地约束了模型的训练。在两个数据集上的实验表明,本文方法具有更好的融合性能。在未来工作中,将进一步探索图卷积神经网络在全色锐化领域中的应用。

### 参考文献:

- [1] 杨勇, 苏昭, 黄淑英, 等. 基于深度学习的像素级全色图像锐化研究综述[J]. 遥感学报, 2022, 26(12): 2411–2432.  
YANG Yong, SU Zhao, HUANG Shuying, et al. Survey of deep-learning approaches for pixel-level pansharpening[J]. National remote sensing bulletin, 2022, 26(12): 2411–2432.
- [2] 赵其昌, 吴一全, 苑玉彬. 光学遥感图像舰船目标检测与识别方法研究进展[J]. 航空学报, 2024, 45(8): 029025.  
ZHAO Qichang, WU Yiquan, YUAN Yubin. Progress of ship detection and recognition methods in optical remote sensing images[J]. Acta aeronautica et astronautica sinica, 2024, 45(8): 029025.
- [3] LI Zhi, PENG Zhenming, CAO Siying, et al. High-resolution remote sensing change detection based on inverse correction and density peak clustering[C]// 2023 IEEE International Geoscience and Remote Sensing Symposium. Pasadena: IEEE, 2023: 3028–3031.
- [4] BAI Jing, LIU Ruotong, ZHAO Haisheng, et al. Hyperspectral image classification using geometric spatial-spectral feature integration: a class incremental learning approach[J]. IEEE transactions on geoscience and remote sensing, 2023, 61: 5531215.
- [5] 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 spatial-spectral remote sensing image fusion based on deep learning[J]. Remote sensing for natural resources, 2023, 35(1): 1–14.
- [6] 徐佳, 关泽群, 何秀凤, 胡俊伟, 等. 基于传感器光谱特性的全色与多光谱图像融合[J]. 遥感学报, 2009, 13(1): 97–102.  
XU Jia, GUAN Zequn, HE Xiufeng, et al. Panchromatic and multispectral image fusion based on sensor spectral characteristics[J]. Journal of remote sensing, 2009, 13(1): 97–102.
- [7] 肖亮, 刘鹏飞, 李恒. 多源空——谱遥感图像融合方法进展与挑战[J]. 中国图象图形学报, 2020, 25(5): 851–863.  
XIAO Liang, LIU Pengfei, LI Heng. Advances and challenges in multi-source spatial-spectral remote sensing image fusion methods[J]. Journal of image and graphics, 2020, 25(5): 851–863.
- [8] TU Teming, SU Shunchi, SHYU H C, et al. A new look at IHS-like image fusion methods[J]. Information fusion, 2001, 2(3): 177–186.
- [9] KWARTENG P, CHAVEZ A. Extracting spectral contrast in Landsat Thematic Mapper image data using selective principal component analysis[J]. Photogrammetric engineering & remote sensing, 1989, 55: 339–348.
- [10] MALLAT S G. A theory for multiresolution signal decomposition: the wavelet representation[J]. IEEE transactions on pattern analysis and machine intelligence, 1989, 11(7): 674–693.
- [11] LIU Pengfei. Pansharpening with spatial hessian non-convex sparse and spectral gradient low rank priors[J]. IEEE transactions on image processing, 2023, 32: 2120–2131.
- [12] ZHU Xiaoxiang, GROHNFELDT C, BAMLER R. Exploiting joint sparsity for pansharpening: the J-SparseFI algorithm[J]. IEEE transactions on geoscience and remote sensing, 2016, 54(5): 2664–2681.
- [13] LIU Pengfei, TANG Songze, HUANG Lili. Pansharpening with spatial hyper-Laplacian and spectral sparse constraints[C]// 2022 IEEE International Geoscience and Remote Sensing Symposium. Kuala Lumpur: IEEE, 2022: 3762–3765.
- [14] MASI G, COZZOLINO D, VERDOLIVA L, et al. Pansharpening by convolutional neural networks[J]. Remote sensing, 2016, 8(7): 594.
- [15] LIU Xiangyu, LIU Qingjie, WANG Yunhong. Remote sensing image fusion based on two-stream fusion network[J]. Information fusion, 2020, 55: 1–15.
- [16] YANG Junfeng, FU Xueyang, HU Yuwen, et al. PanNet: a deep network architecture for pan-sharpening[C]//2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 1753–1761.
- [17] LIU Xiangyu, WANG Yunhong, LIU Qingjie. Psgan: a generative adversarial network for remote sensing image pan-sharpening[C]//2018 25th IEEE International Conference on Image Processing. Athens: IEEE, 2018: 873–877.

- [18] SHI Cheng, PUN C M. Superpixel-based 3D deep neural networks for hyperspectral image classification[J]. *Pattern recognition*, 2018, 74: 600–616.
- [19] AHMAD M, KHAN A M, MAZZARA M, et al. A fast and compact 3-D CNN for hyperspectral image classification[J]. *IEEE geoscience and remote sensing letters*, 2020, 19: 5502205.
- [20] KIPF T N, WELLING M. Semi-supervised classification with graph convolutional networks[EB/OL]. (2016–09–09) [2024–01–01]. <https://arxiv.org/abs/1609.02907v4>.
- [21] 侯磊, 刘金环, 于旭, 等. 图神经网络研究综述[J]. *计算机科学*, 2024, 51(6): 282–298.  
HOU Lei, LIU Jinhuan, YU Xu, et al. Review of graph neural networks[J]. *Computer science*, 2024, 51(6): 282–298.
- [22] XU Boyan, YIN Hujun. Graph convolutional networks in feature space for image deblurring and super-resolution [C]//2021 International Joint Conference on Neural Networks. Shenzhen: IEEE, 2021: 1–8.
- [23] YAN Keyu, ZHOU Man, LIU Liu, et al. When pansharpening meets graph convolution network and knowledge distillation[J]. *IEEE transactions on geoscience and remote sensing*, 2022, 60: 5408915.
- [24] WANG Xiaolong, GIRSHICK R, GUPTA A, et al. Non-local neural networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 7794–7803.
- [25] PALSSON F, SVEINSSON J R, ULFARSSON M O, et al. Quantitative quality evaluation of pansharpened imagery: consistency versus synthesis[J]. *IEEE transactions on geoscience and remote sensing*, 2016, 54(3): 1247–1259.
- [26] WANG Wenqing, ZHOU Zhiqiang, ZHANG Xiaoqiao, et al. DiTBN: detail injection-based two-branch network for pansharpening of remote sensing images[J]. *Remote sensing*, 2022, 14(23): 6120.
- [27] VIVONE G, ALPARONE L, CHANUSSOT J, et al. A critical comparison among pansharpening algorithms[J]. *IEEE transactions on geoscience and remote sensing*, 2015, 53(5): 2565–2586.
- [28] KALLEL A. MTF-adjusted pansharpening approach based on coupled multiresolution decompositions[J]. *IEEE transactions on geoscience and remote sensing*, 2015, 53(6): 3124–3145.
- [29] VIVONE G. Robust band-dependent spatial-detail approaches for panchromatic sharpening[J]. *IEEE transactions on geoscience and remote sensing*, 2019, 57(9): 6421–6433.
- [30] ZHOU J, CIVCO D L, SILANDER J A. A wavelet transform method to merge Landsat TM and SPOT panchromatic data[J]. *International journal of remote sensing*, 1998, 19(4): 743–757.

#### 作者简介:



王文卿, 副教授, 主要研究方向为遥感影像处理与解译、智能信息处理、机器学习。主持国家自然科学基金项目 2 项, 发表学术论文 30 余篇。E-mail: [wangwenqing@xaut.edu.cn](mailto:wangwenqing@xaut.edu.cn)。



张小乔, 硕士研究生, 主要研究方向为遥感图像融合、深度学习、图像处理。E-mail: [2210320138@stu.xaut.edu.cn](mailto:2210320138@stu.xaut.edu.cn)。



刘涵, 教授, 主要研究方向为复杂工业过程建模与控制、机器学习、人工智能、智能信息处理。主持国家自然科学基金项目 3 项, 发表学术论文 160 余篇。E-mail: [liuhan@xaut.edu.cn](mailto:liuhan@xaut.edu.cn)。