引文格式：靳道明,李路沙. 双通道卷积神经网络在影像融合中的应用[J]. 地理空间信息, 2023, 21(11): 1-4. doi:10. 3969/j.issn. 1672-4623.2023.11.001

双通道卷积神经网络在影像融合中的应用

靳道明[1]，李路沙[1]
（1. 中国铁路设计集团有限公司，天津 300142）

摘　要：利用Landsat8遥感卫星影像数据制作影像融合数据集，提出了一种双通道融合网络，并利用结果影像的质量指数对网络融合性能进行评估，分析与双三次卷积插值和GS影像融合方法的差异。结果表明，该网络加强了对高频空间信息的提取，在更高效提取空间特征的同时，减弱了融合过程中对多光谱影像光谱特征的影响，从而提高了融合影像的综合影像质量（QNR = 0.885 2）。

关键词：深度学习；遥感影像融合；双通道卷积神经网络；多尺度特征
中图分类号：P237　　**文献标志码：**B　　**文章编号：**1672-4623（2023）11-0001-04

Application of Dual-channel Concerlutional Neural Network in Image Fusion

JIN Daoming[1], LI Lusha[1]
(1. China Railway Design Corporation, Tianjin 300142, China)

Abstract: We used Landsat 8 remote sensing satellite image data to make the image fusion dataset, put forward a dual-channel fusion network, evaluated the performance of this network by the quality index of remote sensing images, and analyzed the differences between the network and bicubic interpolation and GS image fusion methods. The results show that the network proposed in this paper can strengthen the extraction of high-frequency spatial information. While extracting spatial features more efficiently, it also weakens the influence on the spectral features of multi-spectral images during the fusion process, thus improving the comprehensive image quality of fusion images(QNR = 0.885 2).

Key words: deep learning, remote sensing image fusion, dual-channel convolutional neural network, multi-scale feature

高分辨率遥感影像具有丰富的细节与地物信息，广泛应用于地物分类、目标识别、变化检测等领域[1]。多光谱遥感卫星通过搭载不同波段的传感器针对不同地物的光谱反射特性获取相应的影像，但受限于技术或成本，通常提供高分辨率的全色影像和低分辨率的多光谱影像。低分辨率影像会出现像元混淆，无法反映地面真实情况，因此对于高分辨率多光谱影像的需求不断增加。遥感影像融合技术旨在结合全色影像的分辨率优势与多光谱影像的光谱优势得到高分辨率多光谱影像。传统遥感影像融合方法主要分为分量替换和多分辨率分析[2]，分量替换是将低分辨率多光谱影像上采样后转换至特定的维度，再将全色影像替换转换后影像的某一分量，最后对融合影像进行逆转换得到最终影像，包括强度—亮度—饱和度变换法（将影像中强度分量替换为全色影像）、主成分分析法、Gram-Schmidt分解等，但由于全色波段的光谱范围和光谱特征与多光谱波段存在差异，引入空间细节的同时会引起光谱失真；多分辨率分析是将提取的全色影像中的高频特征注入到上采样后的多光谱影像中，包括抽取或非抽取小波变换[3]、拉普拉斯金字塔等。上述方法均在视觉上有较好的效果，但算法较固定，且在融合过程中易产生光谱信息损失，影响遥感影像质量。相较于传统方法，机器学习方法具有灵活性和可学习性的优点[4]，如压缩感知、字典学习等方法在遥感领域的成功应用。自2014年有学者利用CNN结构解决遥感影像融合问题[5]以来，CNN在遥感融合领域的发展不断加快，Masi G[6]等设计了3层卷积结构的网络，通过变换不同波段组合的方式提升网络的训练效果；Yang J F[7]等构建了分别涵盖光谱特征和空间特征的双通道网络；Scarpa G[8]等通过CNN结构中损失函数、网络结构和训练策略的LI-RES-微调最优组合，建立了影像融合网络A-PNN结构。

相对于影像识别等复杂任务来说，分辨率重建任务的输出影像与原影像有很高的相似性，因此过于复杂的网络结构并不利于数据相似特征的传递，而仅保留卷积层和激活函数等基本模块的网络在影像融合领

收稿日期：2022-06-06。
第一作者简介：靳道明（1990—），硕士，现主要从事遥感数据处理与应用、深度学习方法在遥感中的应用等方面工作，E-mail：jindaoming@crdc.com。

域具有更强的适用性。本文在设计网络时，参考了PanNet结构的双通道特性，从光谱特征和空间特征两个角度出发搭建网络，并在空间通道中加入多尺度模块，以提高空间通道获取特征的能力。

1　网络结构

1.1　融合方法

CNN作为一种高度依赖于标签数据的全监督网络结构，在训练时需以标签数据（真值）作为评判网络输出结果质量的标准；但在影像融合领域，高分辨率多光谱影像是未知的，如何获取标签数据就显得尤为重要。根据Wald协议，假定在不同分辨率下进行影像融合过程具有自相似性，根据融合前后分辨率提升的倍率，对影像进行相同倍率的降采样处理，作为训练数据；再将原有数据看作标签数据，即在低分辨率下进行网络训练，训练完成后，在原始分辨率下对未知影像进行融合即可得到结果。在模拟融合过程中，通过损失函数评定网络输出值与真值的近似程度，再对网络中的参数进行调整，经过多次迭代直至网络收敛并趋于稳定后，网络训练完成；预测时将待融合影像输入网络，利用训练好的网络参数得到融合结果（图1）。

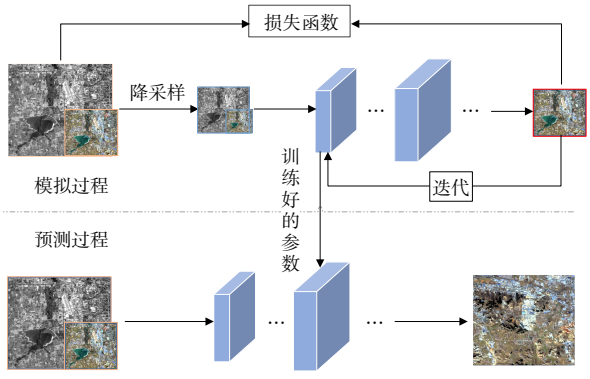


图1　深度学习方法实现遥感影像融合过程

1.2　双通道融合网络

遥感影像融合通常分为特征数据提取、非线性映射和影像重构3个步骤。由于遥感影像的光谱特征与空间特征差异较大，两种特征的提取过程应独立进行，因此为在引入高频空间信息的同时最大程度地保持原影像的光谱信息，本文提出了一种双通道CNN，分别处理多光谱数据和全色数据。考虑到地物的尺寸差异明显、空间信息与尺度大小关系密切，在搭建全色通道时引入多尺度模块，旨在从不同尺度全面获取全色波段中的高频空间信息；且高分辨率多光谱影像与低分辨率多光谱影像具有高度相似性，由上述复杂的网络结构对于光谱特征的传递不具有适用性。

本文的网络结构见图2，在光谱通道中引入残差结构，使网络训练的目的由学习输出值变为学习输出值与输出值差异，既能降低网络的学习难度，又能最大限度地避免输入影像的光谱信息产生损失。为了使光谱影像更好地与高频空间信息相结合，在光谱数据输入网络前对影像进行高斯滤波，剔除低频信息。本文的网络中多尺度模块为特征提取阶段；在后续的卷积层中以ReLU为激活函数获取影像中的非线性数据，最后结合两种数据对影像进行重构。

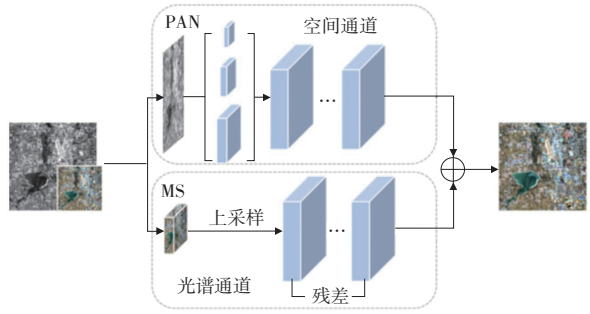


图2　双通道影像融合网络结构示意图

1.3　评价指标

分辨率重建后的影像质量评价包括全参考和无参考两种，全参考指数需与参考影像作为比较。均方根误差（RMSE）能反映融合影像与真值之间的的离散程度，值越小代表越接近真值：

$$\mathrm{RMSE}\triangleq\frac{\sqrt{\sum_{i}^{N}(x_i-y_i)}}{N} \tag{1}$$

式中，x、y分别为测试影像和参考影像；N为像素数量。

相对全局误差（ERGAS）是从整幅影像的角度对影像质量进行评价，值越小代表越接近真值：

$$\mathrm{ERGAS}\triangleq 100\frac{h}{l}\sqrt{\frac{1}{N}\sum_{i=1}^{N}\frac{\mathrm{RMSE}^2 B_i}{M_i^2}} \tag{2}$$

式中，N为波段数；M为辐射均值；$\frac{h}{l}$为重建倍率。

Q值通过计算融合结果与真值之间的相关系数判断影像质量，越接近于1，表示二者差异越小：

$$Q\triangleq\frac{4\sigma_{xy}\bar{x}\bar{y}}{\left(\sigma_x^2+\sigma_y^2\right)\left[\bar{x}^2+\bar{y}^2\right]} \tag{3}$$

式中，σ_x、σ_y为x、y的标准差。

光谱角测度（SAM）通过计算融合影像与真值像素向量的绝对差值判断影像质量，越接近于0，表示二者向量差值越小：

$$\mathrm{SAM} \triangleq \sum_{i=1}^{N} \cos^{-1}\left(\frac{\boldsymbol{v}_i^{\mathrm{T}} \boldsymbol{v}_i^{*}}{\left(\left\|\boldsymbol{v}_i\right\|\right)\left(\left\|\boldsymbol{v}_i^{*}\right\|\right)}\right) \tag{4}$$

式中，N 为像素数量；$\boldsymbol{v}$ 为对应像元的列向量。

D_λ 通过计算融合影像与原多光谱影像差值的1范数获取融合影像在光谱分量上的质量，越接近于0，表示影像的光谱质量越好：

$$D_\lambda = \frac{1}{B(B-1)} \sum_{i=1}^{B} \sum_{j=1(j\neq i)}^{B} \left| Q\left(X_i, X_j\right) - Q(x_i, x_j) \right| \tag{5}$$

D_s 与 D_λ 相似，通过计算融合影像与全色影像差值的1范数评价影像在空间分量上的质量，越接近于0，表示影像的空间质量越好：

$$D_s = \frac{1}{B} \sum_{i=1}^{B} \left| Q\left(X_i - P\right) - Q(x_i - p) \right| \tag{6}$$

式中，x、X 分别为结果影像和LRMS影像；p、P 分别为PAN影像和降采样后的PAN影像。

无参考质量评价（QNR）综合光谱与空间质量得到融合影像的综合影像质量评价指数[9]，越接近于1，表示二者差异越小：

$$\mathrm{QNR} \triangleq (1 - D_\lambda)(1 - D_s) \tag{7}$$

2　数据集与实验

本文以Landsat8遥感影像为数据源[10]，共选取9景影像作为实验数据，获取时间均匀分布在一年的4个季度中；包含森林、城市、盐碱地、农田等多种地物种类；所有数据均为标准1级产品，已经过辐射校正和几何校正（表1）。数据集由18 000个影像对组成，每个影像对由光谱影像和与之对应的全色影像组成，其中全色影像尺寸为256×256，多光谱影像尺寸为128×128，均由原始分辨率下的影像降采样而来。网络训练时，每批包含64幅影像，设置学习率为0.001，利用SGD优化器对网络参数进行迭代更新。

表1　Landsat8卫星影像波段信息

波段	中心波长/nm	频带宽/nm	分辨率/m
B1	443	20	30
B2	482.5	65	30
B3	562.5	75	30
B4	655	50	30
B5	865	40	30
B6	1 610	100	30
B7	2 200	200	30
B8	590	180	15

3　研究结果与分析

训练和测试部分是以原始影像为标签在低分辨率下进行模拟融合，预测部分则是在原始尺度上进行的。本文采用Python3.6语言和Tensorflow 1.8.0框架。在计算测试影像质量时，选取2 000对测试影像用于验证网络的性能，取平均数作为最后结果；并与双三次卷积插值、GS融合和PanNet进行比较。测试结果见表2，在4种全参考指数中，本文方法均表现最好，其次是PanNet，GS融合在数值上表现最差；在3种无参考指数中，上采样方法与原始影像具有相同的光谱信息，因此光谱指数的数值最小，而在空间指数中本文方法表现最好，其次是GS融合；在整体的QNR指数中，本文方法表现最佳，PanNet次之。

表2　模拟分辨率下不同方法的结果比较

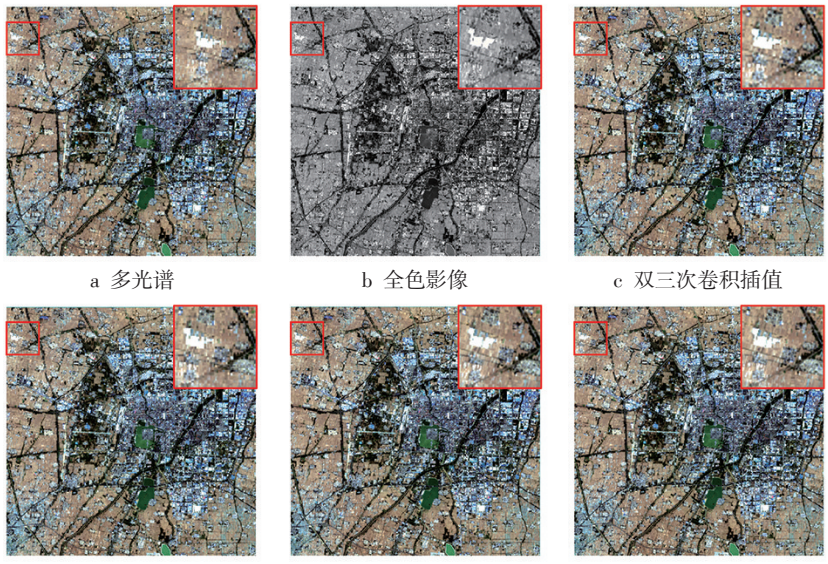
评价方式	评价指数	理想值	双三次卷积插值	GS	PanNet	本文网络
全参考	RMSE	0	0.075 0	0.120 2	0.052 3	**0.046 6**
	ERGAS	0	9.419 1	15.348 8	6.251 3	**5.873 0**
	Q	1	0.846 2	0.809 6	0.905 9	**0.912 6**
	SAM	0	1.499 6	1.430 1	1.176 5	**1.131 7**
无参考	D_λ	0	**0.005 4**	0.115 5	0.013 1	0.011 7
	D_s	0	0.125 0	0.037 1	0.038 4	**0.032 1**
	QNR	1	0.870 2	0.851 7	0.949 0	**0.956 6**

注：加粗数值为表现最好，下划线为次好。

预测阶段本文选取一幅原分辨率影像（影像尺寸为800×800）进行预测，以验证方法对真实影像的有效性。由于可视化一次只显示3个波段，选择3个波段（B4、B3和B2）进行真彩色合成，结果见图3，可以看出，双三次卷积插值的结果影像在视觉上与LRMS相同，即单纯的差值并未改善空间信息；GS方法具有良好的视觉性能，但细节图中白色建筑物在全色图中是分离式结构，在GS融合结果中体现并不明显；两种深度学习方法不仅有更好的视觉效果，细节图中的白色建筑物也表现出分离趋势。

在原始分辨率下评价影像质量是没有高分辨率多光谱影像作为真值的，因此在预测阶段利用无参考指数评价融合影像的质量，由表3可知，在原始分辨率下的结果与模拟分辨率下的结果具有相似性，说明降低分辨率获取标签数据的训练方法具有适用性，即在不同分辨率下同倍率影像融合具有相似性。此外，双三次卷积插值方法在光谱指数中取得最佳表现，归因于其不引入新的数据也未改变原始影像的光谱特征，而GS融合在空间分量上表现最好，但在光谱指数中的表现最差，也能体现出该方法会过度引入全色数据的特点。值得注意的是，两种深度学习方法在光谱和空间分量上的表现虽都未取得最好表现，但在综合影

像质量上的表现往往更好，相较于两种传统方法，深度学习能在空间和光谱两个分量上找到一个平衡，且本文方法在空间和光谱分量上均比PanNet具有更好的表现，说明加入多尺度模块可提高网络提取全色影像特征的能力，也能减弱对融合后影像光谱特征的影响。



a 多光谱　b 全色影像　c 双三次卷积插值

d GS融合　e PanNet　f 本文网络

图3　影像质量主观评价图

表3　原始分辨率下不同方法的结果比较

评价指数	双三次卷积插值	GS融合	PanNet网络	本文方法
D_λ	**0.003 8**	0.119 6	0.031 0	0.029 8
D_s	0.141 4	**0.080 6**	0.091 3	0.087 7
QNR	0.855 3	0.809 4	0.880 5	**0.885 2**

注：加粗数值为表现最好，下划线为次好。

4　结　语

本文以Landsat8卫星影像为例，制作影像融合据集；根据遥感影像融合过程，提出了一种双通道网络结构，独立处理多光谱和全色数据，并加入多尺度模块更加全面地提取全色波段中的高频空间信息。结果表明，本文中涉及的两种深度学习网络两个分量指数的表现不是最高，而综合质量表现更好；相较于PanNet，本文网络结构加强了对高频空间信息的提取，使得光谱和空间分量上的影像质量都有所改善，从而提高了融合影像的综合影像质量。

参考文献

[1] 韩志晟,孙丕川,唐超. 面向单幅遥感影像的生成对抗网络超分辨率重建[J]. 测绘通报,2021(8):106-110

[2] Li S, Yin H, Fang L. Remote Sensing Image Fusion via Sparse Representations Over Learned Dictionaries[J]. IEEE Transactions on Geoscience and Remote Sensing, 2013,51(9):4 779-4 789

[3] Palsson F,Sveinsson J R,Ulfarsson M O. A New Pansharpening Algorithm Based on Total Variation[J]. IEEE Geoscience and Remote Sensing Letters,2014,11(1):318-322

[4] 夏乐,李长安,江涛,等. 基于深度卷积神经网络的遥感影像船只识别[J]. 地理空间信息,2021,19(9):7-9

[5] Ledig C,Theis L,Huszar F,et al. Photo-realistic Single Image Super-resolution Using a Generative Adversarial Network[J]. IEEE Computer Society,2016(9):238-245

[6] Masi G,Cozzolino D,Verdoliva L,et al. Pansharpening by Convolutional Neural Networks [J]. Remote Sensing, 2016,8(7):594

[7] Yang J F,Fu X,Hu Y,et al. PanNet: A Deep Network Architecture for Pan-sharpening[C]. IEEE International Conference on Computer Vision (ICCV),2017

[8] Scarpa G,Vitale S,Cozzolino D. Target-adaptive CNN-based Pansharpening[C]. IEEE Transactions on Geoscience and Remote Sensing,2017:1-15

[9] Alparone L, Aiazzi B, Baronti S, et al. Multispectral and Panchromatic Data Fusion Assessment Without Reference [J]. Photogrammetric Engineering and Remote Sensing,2008,74(2):193-200

[10] 林娜,王伟,王斌. 基于随机森林和Landsat8 OLI影像的脐橙果园种植信息提取[J]. 地理空间信息,2021,19(11):96-100