

中图法分类号: TP751 文献标识码: A 文章编号: 1006-8961(2024)07-2046-17

论文引用格式: Zhou Q Z, Guo Q, Wang H R and Li A. 2024. Two-discriminators-deep residual GAN hyperspectral image pan-sharpening. Journal of Image and Graphics, 29(07):2046-2062(周庆泽, 郭擎, 王海荣, 李安. 2024. 双判别器深度残差 GAN 高光谱图像融合. 中国图象图形学报, 29(07):2046-2062)[DOI:10.11834/jig.220932]

双判别器深度残差 GAN 高光谱图像融合

周庆泽^{1,2}, 郭擎^{1*}, 王海荣^{1,2}, 李安¹

1. 中国科学院空天信息创新研究院, 北京 100094; 2. 中国科学院大学电子电气与通信工程学院, 北京 100049

摘要: 目的 为了解决高空间分辨率多光谱图像与高光谱图像融合时的多波段对多波段问题, 以及高空间分辨率多光谱图像波谱范围不能完全涵盖高光谱图像波谱范围而导致的光谱失真问题, 本文利用深度学习的数据驱动优势, 基于高分 5 号 (GF-5) 高光谱数据和 Sentinel-2 多光谱数据, 提出一种基于生成对抗网络 (generative adversarial network, GAN) 的高光谱图像空谱融合方法——双判别器深度残差 GAN 网络 (two discriminator deep residual GAN, 2DDRGAN)。方法 考虑待融合图像间的波谱范围关系, 采用分组融合策略, 利用波段间的相关性, 将多对多的融合问题转变为多个一对多的融合问题。使用深度残差模块深度提取图像的光谱和空间特征, 用两个判别网络对融合图像的空间和光谱质量分别进行判断, 改善生成网络生成的融合图像质量。另外, 本文的深度学习网络不需要制作额外的融合结果标签, 待融合图像本身就是标签, 这大大降低了高光谱融合的工作量, 也是目前深度学习遥感图像融合的根本改变。结果 与常用传统空谱融合方法和经典深度学习方法比较的实验结果表明, 对于不同地物类型数据, 该网络得到的融合结果在提升空间分辨率的同时, 有较高的光谱保真度。光谱曲线评价也验证了该网络对于高空间分辨率图像波谱范围以外的高光谱图像波段进行融合时有良好的光谱保真度。结论 本文方法通过深度残差模块提取高光谱图像光谱特征和高空间分辨率图像空间特征, 同时引入双判别网络, 使得融合结果在保持光谱信息的同时更好地提升空间信息。

关键词: 高光谱图像空谱融合; 高空间分辨率多光谱图像; 光谱失真; 融合策略; 生成对抗网络 (GAN); 光谱曲线评价

Two-discriminators-deep residual GAN hyperspectral image pan-sharpening

Zhou Qingze^{1,2}, Guo Qing^{1*}, Wang Hairong^{1,2}, Li An¹

1. Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China;

2. School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China

Abstract: Objective Hyperspectral image (HS) pan-sharpening can obtain the fused image with both high spatial resolution and high spectral resolution by using the complementary information between the high spectral resolution HS and the high spatial resolution multi-spectral image (MS). However, HS and MS images both have multi-bands, which is different from the traditional MS pan-sharpening where one band panchromatic image (PAN) with multi-band MS image. This many-to-many band relationship is not convenient for the implementation of pan-sharpening method. Moreover, for those bands

收稿日期: 2022-10-09; 修回日期: 2023-10-19; 预印本日期: 2023-10-26

* 通信作者: 郭擎 guoqing@aircas.ac.cn

基金项目: 国家自然科学基金项目 (61771470)

Supported by: National Natural Science Foundation of China (61771470)

where HS exceeds the spectral range of MS, there will be obvious spectral distortion in the fused image due to the lack of strict physical complementary information. In order to solve the above problems, this paper exploits the data-driven advantages of deep learning and proposes the two discriminators deep residual generative adversarial network (2DDRGAN) for hyperspectral image pan-sharpening based on Gaofen-5 (GF-5) HS image and Sentinel-2 MS image. **Method** Considering the spectral range relationship between HS and MS, this paper adopts the grouping pan-sharpening strategy to transform the many-to-many problem into multiple one-to-many problems by using the correlation between bands of HS and MS. For the HS band within the spectral range of the high spatial resolution MS image, the band of HS image is directly assigned to the group corresponding to the high spatial resolution MS image. For the HS bands outside the spectral range, the correlation coefficient is used as the band grouping standard to group the bands of HS and MS. It solves the problem that two multi-band images are not easy to be fused directly, and indirectly improves the spectral fidelity of the fused image. The proposed 2DDRGAN consists of one generation network and two discrimination networks. The generation network mainly extracts the deep spectral and spatial features by establishing the deep residual module. Two discrimination networks judge the spatial quality and the spectral quality of the fused image, respectively, to improve the quality of the output fused image from the generation network. The main task of the spatial discrimination network is to compare the results of the generated network with the high spatial resolution MS image, so as to ensure that the generated fused images have high spatial resolution. The main task of spectral discrimination network is to compare the results of the generated network with the HS images, to ensure the generated fused images have high spectral resolution. Moreover, the deep learning pan-sharpening methods do not have real high spatial resolution and high spectral resolution images as the fusion result labels. Most of them currently are based on the simulated data made by Wald protocol. In this paper, the deep learning network does not need to create additional fusion result labels. The images to be fused themselves are labels, which greatly reduces the workload of hyperspectral fusion with a huge amount of data, and is also a fundamental change in the current deep learning fusion.

Result The fusion results of the proposed 2DDRGAN method in different scenarios are compared with traditional methods and existing deep learning methods. The experimental results show that the 2DDRGAN method has high spectral fidelity while improving the spatial resolution. The spectral curve evaluation also verifies shows that the 2DDRGAN network has good spectral fidelity for hyperspectral image bands beyond the spectral range of high spatial resolution images. **Conclusion** This method extracts the spectral features of hyperspectral images and the spatial features of high spatial resolution images through the depth residual module, and introduces the double discrimination network, so that the fusion results can better improve the spatial information while maintaining the spectral information.

Key words: spatial-spectral fusion of hyperspectral images; high spatial resolution multispectral image; spectral distortion; fusion strategy; generative adversarial network (GAN); spectral curve evaluation

0 引言

高光谱遥感技术是在可见光、近红外、中红外甚至远红外波谱范围内,获取近乎光谱连续的遥感数据的技术。高光谱遥感由于其光谱分辨率高、图谱合一的特点,近年来在矿物勘探、农业估产和环境监测等领域发挥着越来越重大的作用(童庆禧等, 2016; 聂江涛等, 2023)。不过,受到卫星平台高光谱成像硬件设备的限制以及现实物理条件与成本的制约,高光谱图像的空间分辨率相对较低,很难满足地物检测、地物分类等实际应用的需求(崔艳荣等, 2015, 方帅和许漫, 2023)。高空间分辨率的多光谱

图像是最常见的遥感数据之一。利用高光谱图像(hyperspectral image, HS)与多光谱图像(multispectral image, MS)之间的信息互补性,通过图像融合可以得到同时具有高空间分辨率和高光谱分辨率的融合图像,为高光谱图像准确解译和广泛应用奠定基础。

当前高光谱图像融合的方法可分为4类:1)将传统多光谱图像与全色图像(panchromatic image, PAN)的空谱融合方法进行优化,使其适用于高光谱图像与高空间分辨率的全色图像或多光谱图像;2)基于矩阵的高光谱图像融合方法;3)基于张量的融合方法;4)基于深度学习的高光谱图像融合方法。

基于传统空谱融合的高光谱图像融合方法主要分为成分替换法(component substitution, CS)(Shah等, 2008)和多分辨率分析法(multiresolution analysis, MRA)(Mallat, 2006)两大类。成分替换法主要的操作是对MS图像进行一些变换,然后用PAN替换掉MS变换后的空间分量,再与光谱分量一起进行逆变换得到融合结果。当前基于CS融合方法扩展到高光谱图像的方法主要有:基于主成分分析(principal component analysis, PCA)改进的融合方法(Vivone等, 2014)和基于GS(Gran-Schmidt)变换改进的融合方法(Dalla Mura等, 2015)。多分辨率分析法先将多幅图像分解到不同尺度、不同分辨率上,再对低频轮廓信息和高频细节信息分别按照一定的规则进行融合,常见的方法有基于小波变换改进的融合方法、基于亮度调节的平滑滤波方法(smoothing filter-based intensity modulation, SFIM)(王密等, 2018)以及基于拉普拉斯金字塔变换方法(generalized Laplacian pyramid, GLP)(Vivone等, 2019)。基于传统空谱融合改进的高光谱图像融合方法很少真正考虑待融合图像间波谱范围是否重叠的问题:与高分辨率PAN融合时,将HS当做原来的MS直接和PAN融合,对于HS超出PAN波谱范围的波段,由于没有严格物理意义上的互补信息,融合图像会有明显的光谱失真。

基于矩阵的高光谱图像空谱融合方法主要有两类:统计估计法(Wei等, 2014, 2015)和混合光谱分解法(刘洋等, 2016)。统计估计法假设低空间分辨率HS是高空间分辨率HS在空间域上的降采样,而高空间分辨率MS是高空间分辨率HS在光谱域的降采样。基于以上假设,可以把HS与MS的空谱融合看做低空间分辨率的HS和高空间分辨率的MS融合重建的反问题。在此基础上,Xie等人(2020)基于最大后验概率估计(maximum a posteriori, MAP)和随机混合模型(stochastic mixing model, SMM),提出了MAPSMM方法。混合光谱分解法假设高空间分辨率的高光谱图像HS可以通过低空间分辨率HS确定的端元和高空间分辨率多光谱图像MS确定的丰度共同表示。耦合非负矩阵分解法(coupled non-negative matrix factorization, CNMF)是实现这一方法的常见技术(Zhang等, 2013)。此类方法在光谱混合模型中,间接考虑了待融合图像波谱范围的问题,融合效果比传统空谱融合方法有一定提升,但依赖

于先验假设和知识。

基于张量的融合方法是通过空间光栅扫描,将HS图像和MS图像立方体投射为二维矩阵,通过对矩阵分解进行融合(Dong等, 2016)。Li等人(2018)提出了基于耦合稀疏张量分解的融合方法(coupled sparse tensor factorization, CSTF),将高光谱分辨率高空间分辨率的高光谱图像视为三维张量,将融合问题重新定义为3种模式的核心张量和字典估计。Liu等人(2021)提出了基于低阶张量近似的融合方法,该方法将高光谱分辨率的HS图像与高空间分辨率的MS图像耦合起来,利用张量在空间和频谱上直接施加低秩属性,避免了计算复杂的补丁聚类 and 字典学习。不过,基于张量的融合方法常将待融合图像转化为三维图像立方体,这种矩阵化方法易混淆图像立方体内的空间纹理结构和空间—光谱相关性,从而导致空间信息和光谱特征损失。

空谱融合算法的目标是生成一个空间分辨率和光谱分辨率都高的图像。在融合的过程中,为了保持这些图像的光谱保真度,就需要高度非线性函数进行模拟。然而传统方法多使用线性模型对高维特征进行描述(余典等, 2023)。深度学习方法具有非线性、深度特征自学习能力,可以打破线性模型的局限,学习到丰富的深层特征,从而能更好地表达图像的高维度属性。

生成对抗网络(generative adversarial network, GAN)(Goodfellow等, 2014)相比于其他神经网络,不需要大量带标签数据就能学习到深层特征,而且生成的数据更符合真实数据分布。针对训练过程中没有任何控制导致结果不理想的问题,Mirza和Osindero(2014)提出了条件生成对抗网络(conditional GAN, CGAN),将无监督学习的GAN转变成有监督学习的GAN。为解决GAN训练不稳定、生成过程不可控等问题,Radford等人(2016)提出将卷积神经网络(convolutional neural network, CNN)和GAN结合起来构建DCGAN(deep convolutional GAN)模型。但是现有DCGAN的生成网络一般是只有几个网络层的浅层模型,提取能力较弱,无法充分提取高光谱图像和高空间分辨率图像丰富的光谱和空间特征。此外,原始GAN的单判别网络,使其生成网络很难在提升空间分辨率和保持光谱保真度之间取得平衡。Ma等人(2020)将GAN应用于多光谱融合,提出了一种基于生成对抗网络的无监督遥感图像融合

的网络 PAN-GAN。Xiao 等人(2021)提出了 physics-based GAN 网络,该网络利用物理约束的生成对抗网络来融合高光谱图像和多光谱图像,从而提高图像的空间和光谱分辨率。该网络通过引入物理约束和递归模块,有效地提升了图像的空间细节和光谱保真度,并在多个数据集上所体现的性能优于现有方法,但其没有从图像的光谱一致性去设计判别器,这可能导致生成图像在光谱上存在偏差或者噪声,影响图像的光谱质量。

考虑到上述问题,以及高空间分辨率图像波段范围以外的高光谱图像波段融合时,没有严格遥感物理意义的互补信息带来的光谱失真问题和多光谱图像与高光谱图像融合时的多波段对多波段问题,本文发挥深度学习的驱动优势,采用分组融合策略,提出一种基于 GAN 的高光谱图像空谱融合方法—双判别器深度残差 GAN(two discriminator deep residual GAN, 2DDRGAN)。另外,目前应用于空谱融合的深度学习由于没有真实的高空间分辨率高光谱分辨率图像作为融合结果标签,大多是基于 Wald 协议制作的模拟数据进行的,对于真实数据的融合效果往往不佳。本文改变思路,让待融合数据本身做标签,不再需要另外制作模拟融合结果标签,这是对目前深度学习融合的根本改变,也大大降低了标签制作的烦琐工作量。

1 数据介绍

1.1 使用数据

本文使用的数据主要来自于 2 颗卫星:中国的高分 5 号(GF-5)卫星和欧空局的 Sentinel-2 卫星,其中选用 2019 年 5 月 1 日获取的 GF-5 卫星影像和 2019 年 5 月 28 日获取的 Sentinel-2 卫星影像作为实验数据。

GF-5 搭载的可见短波红外高光谱相机主要用于获取从紫外到短波红外谱段(400~2 500 nm)的高光谱分辨率数据产品,其光谱通道数有 330 个,可见近红外光谱分辨率为 5 nm,短波红外光谱分辨率为 10 nm,幅宽达 60 km,空间分辨率为 30 m(刘银年等,2020;孙允珠 等,2018)。

Sentinel-2 卫星的多光谱成像仪,波谱范围从可见光到近红外再到短波红外,共计 13 个波段,其中 10 m 分辨率的波段有 4 个,20 m 分辨率的波段有

6 个,60 m 分辨率的波段有 3 个。其幅宽达到 290 km,且双星组合可以实现 5 天的重访周期。

GF-5 可见短波红外高光谱相机和 Sentinel-2 卫星多光谱成像仪的主要技术指标如表 1(Hoersch, 2017)所示。

本文选取 Sentinel-2 的 4 个 10 m 分辨率的波段 2、波段 3、波段 4 和波段 8 作为高空间分辨率多光谱数据与 GF-5 高光谱数据进行融合。

表 1 GF-5 可见短波红外高光谱相机和 Sentinel-2 卫星多光谱成像仪的主要技术指标

Table 1 The main specifications of GF-5 visible short wave infrared hyperspectral camera and Sentinel-2 satellite multispectral image

技术指标	GF-5	Sentinel-2
光谱范围	0.400 ~ 2.500 μm	0.443 ~ 2.190 μm
空间分辨率	30 m	10 m, 20 m, 60 m
幅宽	60 km	290 km
光谱分辨率	可见近红外 VNIR: 5 nm 短波红外 SWIR: 10 nm	> 5 μm

1.2 数据预处理

考虑到待融合的高光谱图像和高空间分辨率多光谱图像分别来自于 GF-5 和 Sentinel-2 两个传感器,同时图像的获取时间也有差异,融合前需要对 GF-5 高光谱图像和 Sentinel-2 多光谱图像进行几何校正和图像配准预处理、重采样、裁剪以及数据格式转换的操作,以获取精确配准后的融合数据集,便于后续的融合操作。融合数据集的制作流程如图 1 所示。

2 分组融合策略

高空间分辨率图像与高光谱图像的波谱范围通常有一定的重叠。在高空间分辨率图像和高光谱图像的波谱范围内,高空间分辨率图像的一个波段对应着多个高光谱图像波段。Sentinel-2 四个波段波谱范围内分别对应着 GF-5 的多个波段,这种多对多的关系不利于后续融合操作。Zhang 等人(2017)采用分组策略的思想,根据高光谱图像的波段相关性将波段分成几个组,而后利用群体谱嵌入正则项来约束高光谱图像的波段重构,并通过求解优化问题,利用多光谱图像和高光谱图像的信息,恢复出高分

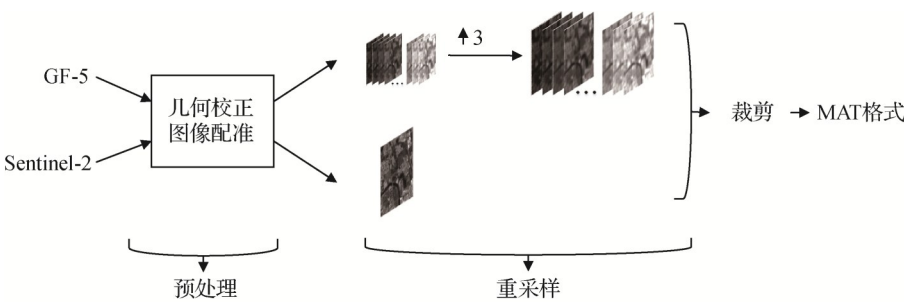


图1 GF-5和Sentinel-2进行预处理、重采样、裁剪以及格式转换的融合数据集的制作过程
Fig. 1 The GF-5 and Sentinel-2 fusion dataset manufacturing process of preprocessing, resampling, cropping, and format converting

分辨率高光谱图像。虽然利用了高光谱图像自身的各波段之间的相关性去分组实现融合,但没有考虑高光谱波段和多光谱波段之间的相关性。因此,本文提出一种分组融合策略,即在融合前先对待融合的高光谱图像和高空间分辨率图像按波段范围进行分组。

对于多光谱波段范围内的高光谱波段,它们之间的分组关系是确定的,直接将高光谱图像的波段分配到对应多光谱图像的组中。但是,相邻多光谱波段的波谱范围并不是连续的,对于它们之间的高光谱波段,通过相关系数求解其对应的分组会更精确。对于从近红外到短波红外范围的高光谱波段,一般没有多光谱波段的覆盖(高光谱卫星的波段范围一般是400~2 500 nm,多光谱卫星的波段范围一般是400~1 000 nm),这导致二者之间没有可直接采用的分组依据,所以采用相关系数作为它们的分组依据。首先,对高空间分辨率多光谱图像按照高光谱图像的尺寸大小进行下采样,得到下采样后的多光谱数据;然后,计算高光谱图像每个波段和下采样后多光谱数据各个波段之间的相关系数;最后,得到一个HS每个波段与MS所有波段的相关系数表,根据表中最大值将HS波段与对应的MS波段分为一组。本文中,MS有4个波段,HS有330个波段,得到一张330×4的相关系数表,根据每一行的最大值将HS所有波段与对应的MS波段分为一组。由于相关系数表过大,用折线图展示,如图2所示,横轴为HS的各个波段,纵轴为单个HS波段与4个MS波段的相关系数计算结果。图2中3个标记点是HS单个波段与4个MS波段相关系数最大值变化的节点,与表2示例分组结果对应。利用相关系数将高光谱图像与多光谱图像的波段进行分组,既解决了两个多波段图像不便于直接融合的问题,同时也间接提升了融合图像的光谱保真度。

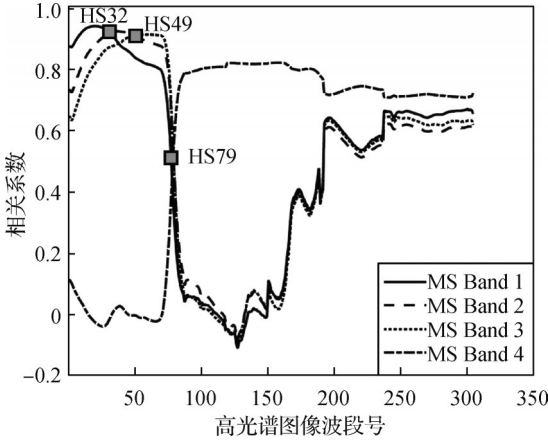


图2 高光谱图像与多光谱图像波段间相关系数
Fig. 2 The band correlation coefficients between hyperspectral image and multispectral image

表2 示例分组结果		
Table 2 The results of grouping		
分组名	MS波段	HS对应波段
Group 1	MS_1	$HS_1 \sim HS_{32}$
Group 2	MS_2	$HS_{33} \sim HS_{49}$
Group 3	MS_3	$HS_{50} \sim HS_{79}$
Group 4	MS_4	$HS_{80} \sim HS_{330}$

3 双判别器深度残差GAN网络

3.1 双判别器深度残差GAN模型

相比于其他神经网络,GAN生成的数据更符合真实数据分布,但原始GAN直接应用于高光谱图像融合仍有不少问题。首先,原始GAN网络过于“自由”,网络输出不受控制。本文将参考CGAN (Mirza和Osindero, 2014)的数据输入形式,通过标签引入额外信息,控制GAN生成网络的输出,加速网络收敛。其次,现有的DCGAN一般是只有几个卷积层的

浅层网络,特征提取能力较弱,无法充分提取高光谱图像和高空间分辨率图像丰富的光谱和空间特征。最后,原始 GAN 网络的单判别网络在反向传播之后,生成网络很难在提升空间分辨率和提高光谱保真度之间取得平衡。因此,本文提出双判别器深度残差 GAN 网络。

3.1.1 深度残差模块

基于 CNN 模型改进的残差网络可有效地解决传统卷积神经网络和全卷积神经网络随着网络加深优化效果变差、出现梯度消失和梯度爆炸的问题(He 和 Sun, 2015)。残差网络通过跳跃连接的方式来实现某一层的特征到另一层的恒等映射,一个残差结构可表示为

$$H(\mathbf{x}) = F(\mathbf{x}) + \mathbf{x} \quad (1)$$

式中, $H(\mathbf{x})$ 是残差网络的输出; $F(\mathbf{x})$ 是输入与输出之间的差距,即为残差; $\mathbf{x}$ 是残差网络的输入。没有添加跳跃连接的原始网络学习的特征是 $H(\mathbf{x})$,而残差网络学习的是输入数据的残差 $F(\mathbf{x}) = H(\mathbf{x}) - \mathbf{x}$,由于残差一般比较小,这使得残差网络学习的内容是稀疏的,因此网络无论是学习能力还是学习效率都大大提升,也更加容易收敛。此外,这种跳跃连接结构在反向传播中不需要经过那么多层网络,避免出现梯度更新信息以指数形式衰减或增加,很好地解决了梯度消失和梯度爆炸问题。

本文参考残差网络学习深层与浅层特征之间残差的思路,将残差结构应用到遥感图像空间和光谱特征的提取工作中,得到更丰富的空间和光谱特征,提升网络性能。本文堆叠多个卷积层构建深层卷积网络,同时将残差网络加入到生成网络中,建立深度残差模块(见图3(He 和 Sun, 2015))来提取图像的空间和光谱特征。

在深度残差模块中,残差子块与子块之间建立

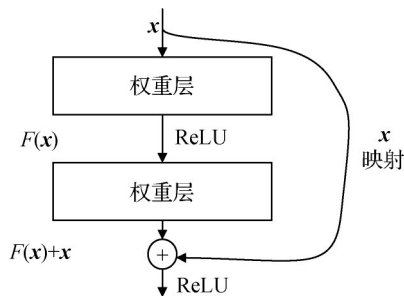


图3 残差网络的基本单元(He 和 Sun, 2015)

Fig. 3 The basic unit of residual network (He and Sun, 2015)

跳跃连接,这样可以将梯度信息传递到更深的网络层,避免梯度消失问题,使得特征信息的传递更加高效。每个残差子块由5个网络层构成,分别是卷积层、批量标准化层(batch normalization, BN)、带参数的修正线性单元层(parametric rectified linear unit, PReLU)、卷积层和BN层。利用多个残差子块进行堆叠形成最终的深度残差模块,如图4所示。

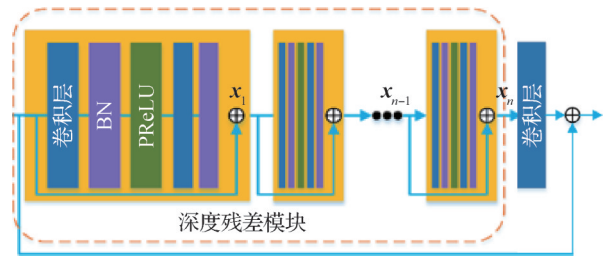


图4 深度残差模块

Fig. 4 Deep residual module

深度残差模块中第 n 层的输出可以表示为

$$\mathbf{x}_n = \max(0, \omega_n \times \mathbf{x}_{n-1}) + b_n, n = 1, 2, 3, \dots \quad (2)$$

式中, $\mathbf{x}_n$ 是第 n 层的输出, ω_n 是第 n 层的权重, $\mathbf{x}_{n-1}$ 是第 $n-1$ 层的输出, b_n 是第 n 层的调整参数。

3.1.2 双判别器深度残差 GAN 结构

本文设计的双判别器深度残差 GAN 网络如图5所示,包括一个以深度残差模块为主体的生成网络和两个以卷积模块为主体的判别网络。生成网络主要提取高光谱图像的光谱特征和高空间分辨率图像的空间特征,并融合得到高空间分辨率的高光谱图像。两个判别网络分别是空间判别网络和光谱判别网络。空间判别网络的主要任务是将生成网络的结果与高空间分辨率图像进行对比,比较生成结果和真实图像的空间分辨率。光谱判别网络的主要任务是将生成网络的结果与高光谱图像进行对比,比较生成结果和真实图像的光谱保真度。

在生成网络中,从输入到输出卷积模块之前的恒等分支使用全局残差,在深度残差模块堆叠多个局部残差单元进行图像的特征抽取;再将深度残差模块的输出通过一个卷积层来调整输出图像大小,从而生成高空间分辨率的高光谱图像。生成网络的主体是由卷积神经网络改进而来,主要分为3个部分:浅层特征提取层、深度残差模块以及特征融合层。首先,浅层特征提取层由一个 9×9 的卷积核构成。其次,对于深度残差模块,大的卷积核会导致计算量的增加,导致网络计算性能下降,因此不适用于

深层网络。Simonyan 和 Zisserman(2015)发现叠加使用多个小卷积核比单独使用大卷积核具有更好的性能。叠加使用多个小卷积核能减少参数的个数,降低计算复杂度。因此,对于深度残差模块中的残差子块,本文选用 3×3 的卷积核。最后,使用一个 9×9 的卷积核进行特征融合,获得最终的融合图像。

对于生成网络的输入,由于高光谱图像与待融合的高空间分辨率图像往往来自不同平台,这些数据的时间相往往差异也很大,不同源不同时间相的数据无法满足Wald协议的基本要求。因此,本文没有采用Wald协议构造参考图像及其下采样后的训练数

据集,而是通过上采样的高光谱图像 $\widetilde{HS}$ 和高空间分辨率图像 HR 组成数据集 $\hat{X} = \{\widetilde{HS}, HR\}$ 。以相关系数为分组标准对原始 HS 和原始 HR 进行分组操作,对每一个分组中的图像分别构建数据集 $\hat{X}$ 。

生成网络的特征提取主要分为浅层特征提取和深层残差特征提取。浅层特征提取由一个 3×3 的卷积层完成,激活函数为PReLU,该过程可以表示为

$$F_q = \max(0, \omega_q \times \hat{X}) + b_q \tag{3}$$

式中, F_q 是提取到的浅层特征, ω_q 是浅层特征提取过程中的权重参数, b_q 是对应的偏置项。深层残差特征提取是由10个残差子块构成的深度残差模块

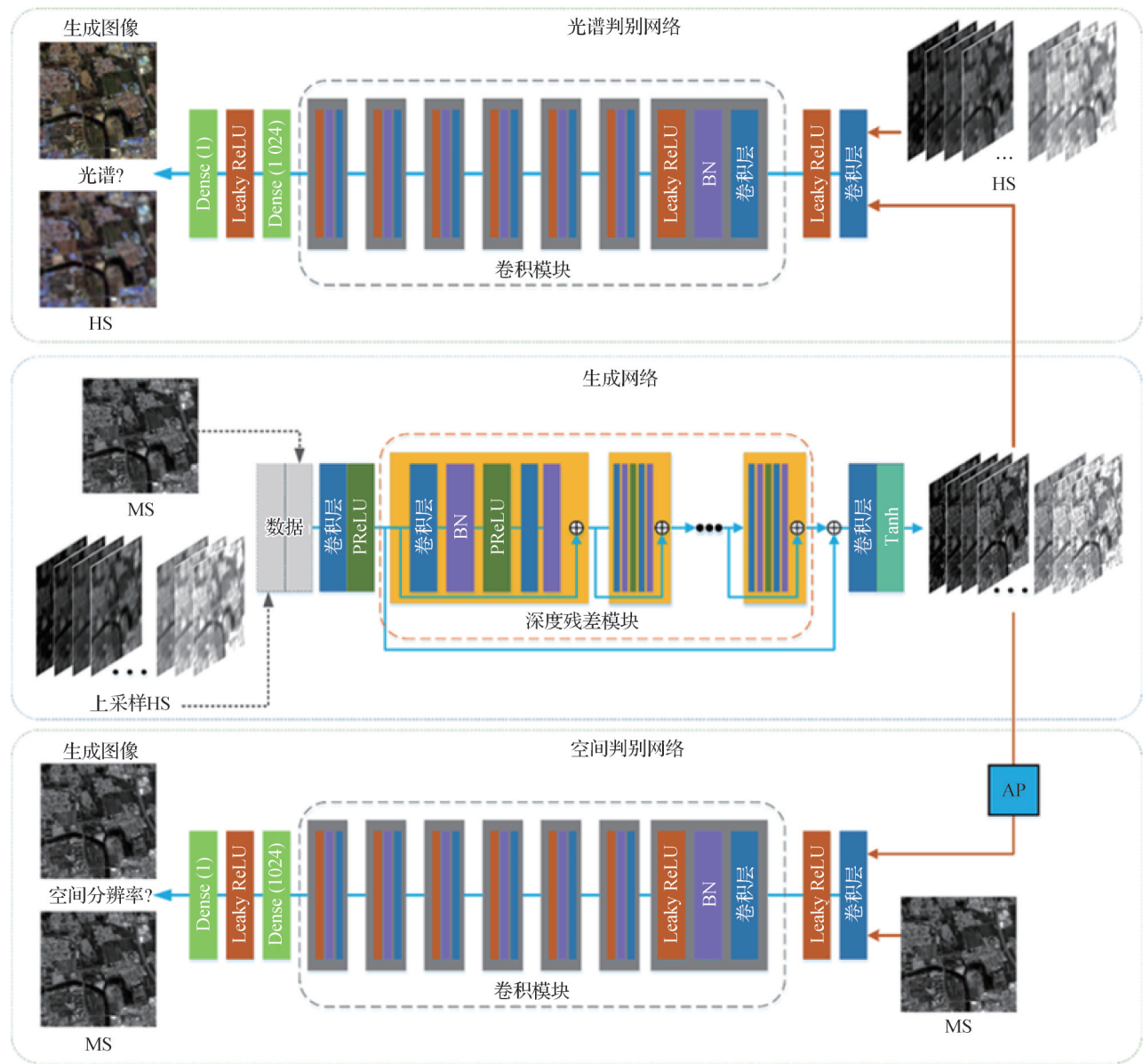


图5 双判别器深度残差GAN网络的结构
Fig. 5 The structure of the proposed 2DDRGAN

来完成。每个残差子块主要包含 2 个 3×3 的卷积层, 2 个 BN 层和 1 个 PReLU 激活函数层。整个深度残差特征提取的过程可以表示为

$$\mathbf{x}_n = \max(0, \omega_n \times \mathbf{x}_{n-1}) + b_n, n = 1, 2, \dots, 10 \quad (4)$$

式中, $\mathbf{x}_n$ 是深度残差模块提取到的特征, ω_n 和 b_n 分别为各残差子块的权重参数和偏置。

整个残差模块学习到的特征 $\mathbf{F}_T$ 由全局残差来实现, 全局残差由学习到的深层特征 $\mathbf{F}_s$ 和浅层特征 $\mathbf{F}_q$ 构成。整个深度残差模块的融合过程表示为

$$\mathbf{F}_T = \mathbf{F}_q + \mathbf{F}_s \quad (5)$$

式中, $\mathbf{F}_s = \mathbf{x}_{10}$ 。

通过一个 3×3 的卷积层重构学习到的空间和光谱特征, 得到融合图像, 该层的输出结果就是生成网络预测的融合图像 $\mathbf{Y}$, 即

$$\mathbf{Y} = \omega_Y \times \mathbf{F}_T + b_Y \quad (6)$$

式中, ω_Y 和 b_Y 分别为融合图像重构过程中的权重参数和偏置。

在判别网络中, 其主体由 DCGAN 改进而来, 由一系列卷积层组成的卷积模块构成。由 7 个卷积层组成的卷积模块和两个全连接层构成, 每个卷积块由 1 个卷积层、1 个 BN 层和 1 个带泄露修正线性单元的激活函数层 (Leaky ReLU) 构成。此外, 传统 DCGAN 网络使用一个判别器, 然而对于高光谱图像融合任务而言, 通过一个判别网络同时判断融合图像的空间和光谱质量的好坏是不够的, 这样会造成在反向传播之后生成网络新生成的融合图像很难在提升空间分辨率和提高光谱保真度之间取得平衡, 要么空间分辨率很高但是光谱失真, 要么光谱保真度很高但是空间分辨率提升不明显。因此, 本文使用两个判别网络来对融合图像质量进行判断, 将融合图像整体质量的判断这一复杂任务分解为对空间质量的判断和对光谱质量的判断两个任务。

光谱判别网络的目的是使生成图像的光谱信息与原始高光谱图像保持一致。光谱判别网络的输入是上采样的高光谱图像和生成网络输出的融合图像, 光谱判别网络通过对输入图像进行特征提取, 最终输出一个概率值, 表示该输入图像的光谱保真度高的概率。空间判别网络的目的是使得生成图像的空间信息与高空间分辨率图像保持一致。空间判别网络的输入是高空间分辨率图像和光谱维平均池化 (average pooling, AP) 后的单波段融合图像, 空间判别网络通过对输入图像进行特征提取, 最终输出一个

概率值, 表示该输入图像的空间分辨率高的概率。

3.2 损失函数

本文提出的双判别器深度残差卷积生成对抗网络包含一个生成网络 G 和两个判别网络 (光谱判别网络 D_1 和空间判别网络 D_2)。其损失函数和网络优化算法具体介绍如下。

3.2.1 生成网络的损失函数

生成网络 G 总的损失函数 L_G 可以定义为

$$L_G = \alpha L_{\text{spatial}} + \beta L_{\text{spectral}} \quad (7)$$

式中, L_{spectral} 表示生成的融合图像与原始高光谱图像之间的光谱信息损失, L_{spatial} 表示生成的融合图像与原始高空间分辨率图像之间的空间信息损失, α 和 β 为平衡空间损失和光谱损失的权重系数。

对于第 1 项光谱损失 L_{spectral} , 本文利用图像融合评价方法中的光谱角度映射表 (spectral angle mapper, SAM) 评价指标, 将生成的融合图像与原始高光谱图像之间的光谱信息损失表示为

$$L_{\text{spectral}} = \frac{1}{N} \sum_{n=1}^N \text{SAM}(\mathbf{I}_F^{(n)}, \mathbf{I}_{\text{HS}}^{(n)}) \quad (8)$$

式中, $\mathbf{I}_F^{(n)}$ 为生成的融合图像, $\mathbf{I}_{\text{HS}}^{(n)}$ 为原始高光谱图像, 对于损失函数中的 SAM 指标可以进一步写成

$$\text{SAM}(\mathbf{I}_F^{(n)}, \mathbf{I}_{\text{HS}}^{(n)}) = \arccos \left(\frac{\sum_{j=1}^n \mathbf{F}_{i,j} \mathbf{HS}_{i,j}}{\sqrt{\sum_{j=1}^n \mathbf{F}_{i,j}^2} \sqrt{\sum_{j=1}^n \mathbf{HS}_{i,j}^2}} \right) \quad (9)$$

式中, $\mathbf{F}_{i,j}$ 表示融合图像 $\mathbf{I}_F^{(n)}$ 第 i 个像素第 j 波段的像素值, $\mathbf{HS}_{i,j}$ 表示光谱参考图像 $\mathbf{I}_{\text{HS}}^{(n)}$ 第 i 个像素第 j 波段的像素值, n 表示波段数。

对于第 2 项空间损失 L_{spatial} , 将均方误差作为生成的融合图像与原始高空间分辨率图像之间的空间信息损失, 表示为

$$L_{\text{spatial}} = \frac{1}{N} \sum_{n=1}^N \left\| \text{AP}(\mathbf{I}_F^{(n)}) - \mathbf{I}_{\text{HR}}^{(n)} \right\|^2 \quad (10)$$

式中, $\mathbf{I}_{\text{HR}}^{(n)}$ 为高空间分辨率图像, $\text{AP}(\cdot)$ 运算表示对输入图像的光谱维度进行平均池化操作得到一个单波段图像。

由以上损失函数建立过程可知, 本文整个损失过程不需要额外的融合结果参考标签, 而是以网络的输入本身为参考, 这不但大大降低了标签制作的繁重工作量, 也让网络学习到真正的融合关系, 而非目前大多数深度学习利用模拟数据做参考的上下采样关系。

3.2.2 判别网络的损失函数

本文提出的双判别器深度残差卷积生成对抗网络中包含两个判别网络,一个用来判别融合图像的光谱保真度,一个用来判别融合图像的空间信息提升程度。它们的损失函数可以统一定义为

$$L_D = \frac{1}{N} \sum_{n=1}^N (D(I^{(n)}) - b)^2 + \frac{1}{N} \sum_{n=1}^N (D(I_F^{(n)}) - a)^2 \quad (11)$$

式中, $I^{(n)}$ 表示参考图像(在光谱判别网络中是高光谱图像,在空间判别网络中是高空间分辨率图像), $I_F^{(n)}$ 表示生成网络得到的融合图像, a 和 b 分别表示融合图像 $I_F^{(n)}$ 和参考图像 $I^{(n)}$ 的标签。 $D(I^{(n)})$ 和 $D(I_F^{(n)})$ 分别表示判别网络对于输入图像和参考图像的分类结果。本文用均方误差作为判别网络输出结果和其对应标签之间的损失函数。

对于光谱判别网络 D_1 , 将 $D = D_1$ 以及 $I^{(n)} = \uparrow I_{HS}^{(n)}$ 代入到式(11)中,即可得到光谱判别网络的损失函数, $\uparrow$ 代表上采样。光谱判别网络主要是判别生成的融合图像与上采样高光谱图像的光谱保真情况。当光谱判别网络无法判别输入的融合图像与上采样高光谱图像的光谱保真度孰优孰劣时,认为生成网络对于光谱保真度达到最优状态。

对于空间判别网络 D_2 , 将 $D = D_2$ 和 $I_F^{(n)} = AP(I_F^{(n)})$ 代入到式(11)中,即可得到空间判别网络的损失函数。空间判别网络主要是判别生成的融合图像与高空间分辨率图像空间分辨率的优劣。当空间判别网络无法判别输入的融合图像与高空间分辨率图像的空间细节信息孰优孰劣时,认为生成网络对于空间细节信息的提升达到最优状态。

4 实验结果与分析

4.1 实验设置

生成网络分为3个部分:浅层特征提取层、深度残差模块以及特征融合层。首先,浅层特征提取层由一个 9×9 的大卷积核构成。其次,深度残差模块中的每个残差子块都使用了 3×3 的小卷积核。最后,用一个 9×9 的大卷积核进行特征融合,生成最终的融合图像。

两个判别网络由7个卷积块组成的卷积模块和两个全连接层构成,每个卷积块由1个卷积层、1个BN层和1个Leaky ReLU激活函数层构成。卷积层

的卷积核大小都是 3×3 , 7个卷积块的特征图数目分别是16、32、64、128、256、512和1 024,两层全连接层的神经元个数分别为1 024和1。

本次实验迭代了400个周期,网络的批量输入大小为32,学习率设置为0.000 2,优化器选用Adam。网络输入的GF-5训练图像为 64×64 像素大小, Sentinel-2 图像为 192×192 像素大小。生成网络的空间损失和光谱损失的平衡权重系数 α 设置为0.7, β 设置为0.3。判别网络的融合图像标签 a 设置为一个 $0 \sim 0.1$ 的随机数,参考图像标签 b 设置为一个 $0.9 \sim 1$ 的随机数。

双判别器深度残差卷积生成对抗网络的训练在Linux发行版本Ubuntu 16.04系统中的PyTorch深度学习框架下执行,处理器为Intel(R) Xeon(R) CPU E5620 @ 2.40 GHz,内存大小为64 GB。程序代码使用Python语言实现,网络模型训练使用NVIDIA TITAN Xp显卡加速。

4.2 图像融合质量评价方法

本文通过主观视觉评价和客观定量评价对融合图像进行综合评估。此外,本文还采用光谱曲线评价的方法对所有波段的融合结果进行更直观的整体评价。其中,客观定量评价包括:1)空间质量评价(标准差(standard deviation, STD));2)光谱质量评价(相关系数(correlation coefficient, CC)、相对整体维数综合误差(error relative global adimensionnelle de synthèse, ERGAS)、光谱角度映射表(SAM)、通用图像质量指数(universal image quality index, UIQI));3)总体质量评价(信息熵(entropy, EN))。

光谱曲线评价可以将所有波段的融合结果放在一起进行整体质量评价。将融合后的高光谱图像的光谱曲线与融合前的低空间分辨率高光谱图像进行比较,即将原始图像某一像素点和融合后图像对应位置 k^2 (k 为融合前后图像空间分辨率之比) 个像素点的光谱曲线取平均后进行比对。

4.3 图像融合对比方法

为了评价本文提出的双判别器深度残差GAN网络的有效性与可靠性,将融合结果与GS、SFIM、GLP、CNMF、MAPSMM以及经典的空谱融合深度学习算法PNN(pan-sharpening neural network)(Masi等, 2016)、PAN-GAN(Ma等, 2020)进行比较。本文选取农田、裸地、建筑、河流和山脉等多种地物类型的区域对8种融合方法进行比较,结果如图6—图8

所示。融合图像使用《GF-5 产品使用说明》中推荐的波段 62、波段 38 和波段 15 进行真彩色展示。

4.4 定性与定量评价

对于农田、裸地和建筑的区域,如图 6 所示,从融合结果整体上看,本文方法、PAN-GAN、GS 和 CNMF 的空间分辨率提升很明显;GLP、MAPSMM、SFIM 和 PNN 生成的图像较为模糊,特别是左下角的密集建筑区较为模糊,农田的轮廓出现了虚影;此外,本文方法的色调明显与高光谱图像更接近;PNN 方法的色调稍逊色于本文方法,但整体也接近于高光谱图像;CNMF、GLP、MAPSMM、SFIM 和 GS 方法融合图像的色调介于多光谱图像和高光谱图像之间,光谱失真较为明显,PAN-GAN 融合图像的色调有所偏离原始高光谱图像,虽然有着较高的空间分辨率,但光谱上有着明显的失真。从局部细节来看,以图 6 红框中的农田及裸地为例,PAN-GAN、GS 和本文方法的农田轮廓较为清晰,空间细节信息提升较为明显;而 GLP、CNMF、MAPSMM 和 SFIM 方法的农田轮廓也较为清晰,但是出现了一些方块状的斑块以及某些不规则的斑块;PNN 方法对于地物细节信息的提升最差,农田的轮廓仍然较为模糊。对于 8 种方法的光谱保真度来说,本文方法光谱保真度较高,融合图像色调基本与高光谱图像一致;CNMF、PAN-GAN、GLP、MAPSMM 和 SFIM 的色调更接近于多光谱图像,光谱失真较为明显;GS 方法红框中的一块农田出现了明显的光谱失真。此外,只

有本文方法和 GS 方法把左下角密集建筑群中的红顶建筑和蓝顶建筑正确显示出来,其他方法皆失真为白色或者蓝灰色。

对于表 3 中的定量指标,GS 方法的 STD 和 EN 好于其他 7 种方法,与定性评价中 GS 方法地物最清晰、空间细节信息丰富的分析相一致。本文方法的 STD 和 EN 仅次于 GS 方法,说明本文方法在细节信息注入方面取得了较好的效果。本文方法的 CC、ERGAS 和 SAM 指标均优于其他对比方法,与定性评价中本文光谱保真度最好的分析相符合。此外,本文方法的 UIQI 指标优于其他方法,表明本文方法

表 3 区域 1 (农田、裸地和建筑)融合结果的数值评估
Table 3 Numeric assessment of Region 1 (farmland, bare land and buildings) fusion results

方法	STD ↑	EN ↑	UIQI ↑	CC ↑	ERGAS ↓	SAM ↓
GS	443.41	8.87	0.64	0.77	5.00	4.46
PNN	431.82	8.56	0.60	0.79	2.67	2.27
GLP	432.08	8.68	0.86	0.88	2.64	2.08
MAPSMM	430.64	8.67	0.81	0.88	2.84	2.00
SFIM	432.65	8.69	0.83	0.87	2.85	2.25
CNMF	435.99	8.64	0.70	0.79	3.61	2.58
PAN-GAN	423.17	7.86	0.79	0.88	3.06	3.85
本文	437.76	8.72	0.89	0.91	2.25	1.79

注:加粗字体表示各列最优结果,↑表示值越高越好,↓表示值越低越好。

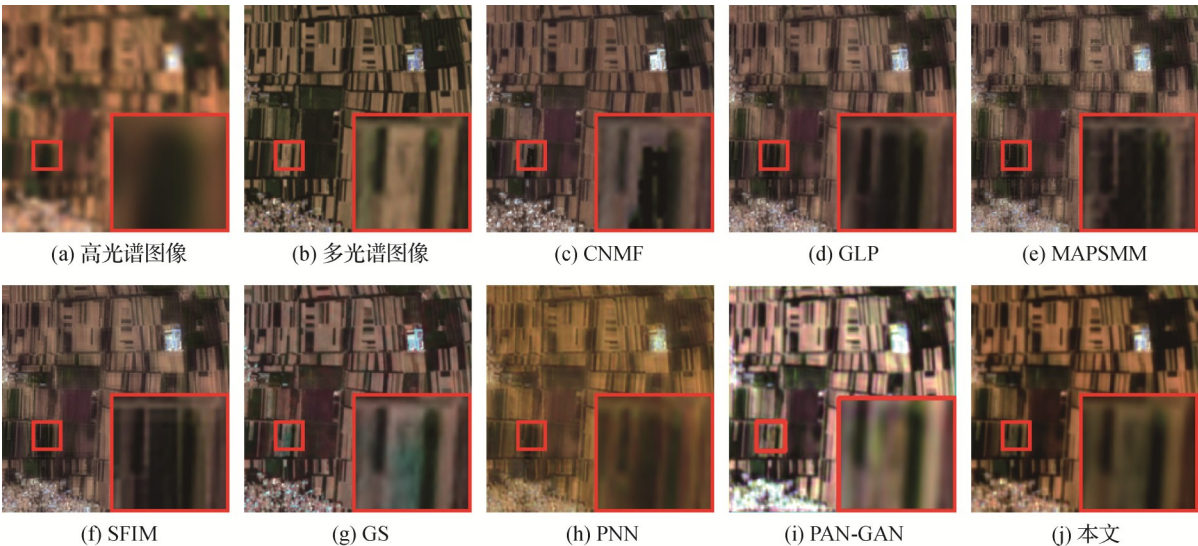


图 6 区域 1 (农田、裸地和建筑)的融合结果

Fig. 6 The fusion results of Region 1 (farmland, bare land and buildings) ((a) hyperspectral image; (b) multispectral image; (c) CNMF; (d) GLP; (e) MAPSMM; (f) SFIM; (g) GS; (h) PNN; (i) PAN-GAN; (j) ours)

的空间细节提升较为明显且具有良好的光谱保真度。PAN-GAN方法有着较好的空间定量指标评价结果,但是光谱保持方面表现得不够理想。可见,本文提出的双判别器深度残差卷积生成对抗网络在提高融合图像空间分辨率的同时还具有良好的光谱保真度。

对于河流、山脉和建筑的区域,如图7所示,从融合结果整体上看,8种方法对空间分辨率都有提升,其中CNMF、PAN-GAN、GS和本文方法的地物细节更加清晰,而GLP、MAPSMM、SFIM和PNN右上角的地物出现了明显的虚影。此外,本文方法的色调整体与高光谱图像更接近,CNMF、GLP、MAPSMM、SFIM和GS方法的色调介于高光谱图像和多光谱图像之间,而PNN方法整体偏红色,融合图像光谱失真较为明显。从局部细节来看,以图7红框中的河岸为例,8种方法融合图像的空间分辨率相比于高光谱图像均有明显的提升,都把高光谱图像中模糊的河岸轮廓信息进行了增强:CNMF、GLP和GS方法的河岸边缘最为锐利;SFIM和MAPSMM方法的细节信息较为模糊,呈现出方块状斑块;PAN-GAN以及本文方法的河岸边缘稍逊色于CNMF、GLP和GS方法,但是整个河岸细节信息过渡较为自然,没有多余的斑块信息;而PNN方法的细节信息明显不如其他7种方法,河岸边缘很模糊。对于8种方法的光谱保真度来说,本文方法的河岸及周围山脉、水体的色调与高光谱图像最为接近,光

谱保真度最好;CNMF方法稍逊色于本文方法;GS、MAPSMM、PAN-GAN、SFIM和GLP方法河岸及其周围地物的颜色介于高光谱图像和多光谱图像之间,有一定光谱失真;PNN方法光谱保真度最差,河岸及周围地物的色调完全失真。此外,只有本文方法把右上角和左下角红顶建筑正确地显示出来,其他方法皆失真为白色和灰色。

对于表4中的定量指标,GS方法的STD和EN均好于其他7种方法,与定性分析中GS方法地物最清晰、空间细节信息最丰富相吻合。本文方法的STD和EN值与除了GS方法以外的其他方法相差不大,但是细节信息注入仍有提升空间。虽然GLP方法的UIQI指标优于其他方法,但是从定性评价中可以看出,GLP方法光谱有一定失真,且细节信息注入逊色于GS方法。此外,GLP方法的ERGAS指标最好,但是定性评价中可以看到较为明显的光谱失真。本文方法的CC和SAM均优于其他对比方法,与定性评价中本文光谱保真度最好相符合,说明本文提出的双判别器深度残差GAN网络在以河流、山脉和建筑为主的场景中具有良好的光谱保真度。

对于建筑、农田和裸地的区域,如图8所示,从融合结果整体上看,GS和CNMF的空间分辨率提升很明显;GLP、MAPSMM和SFIM与之相比较为模糊,特别是右上方的密集建筑区较为模糊;而本文方法、PAN-GAN和PNN空间细节表现不够好,右上方密集建筑区稍模糊,但是本文方法比PNN和PAN-

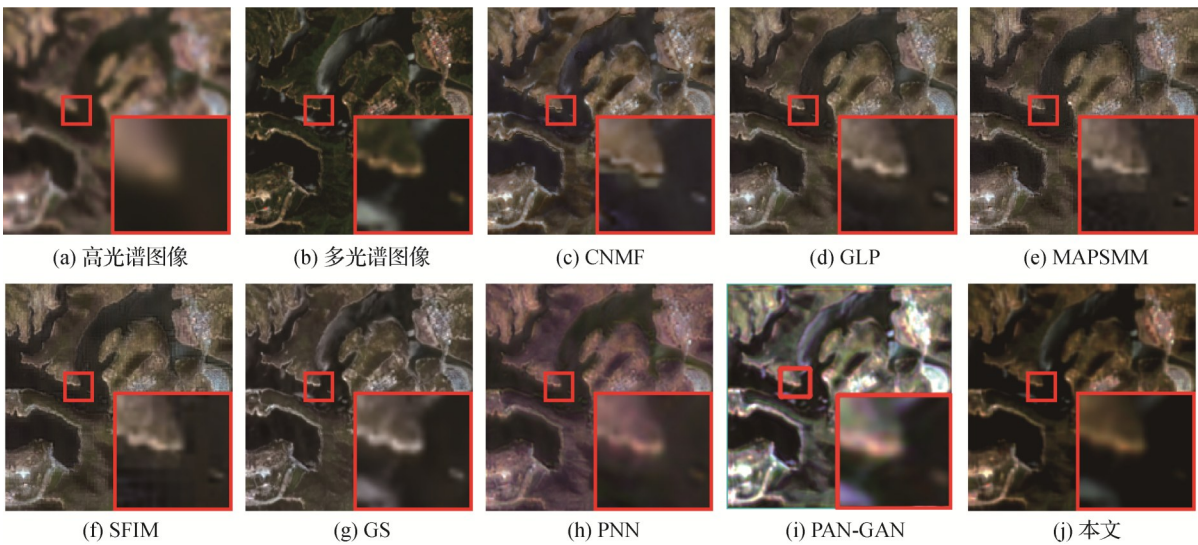


图7 区域2(河流、山脉和建筑)的融合结果

Fig. 7 The fusion results of Region 2 (rivers, mountains and buildings) ((a) hyperspectral image; (b) multispectral image; (c) CNMF; (d) GLP; (e) MAPSMM; (f) SFIM; (g) GS; (h) PNN; (i) PAN-GAN; (j) ours)

表 4 区域 2(河流、山脉和建筑)融合结果的数值评估
Table 4 Numeric assessment of Region 2 (rivers, mountains and buildings) fusion results

方法	STD ↑	EN ↑	UIQI ↑	CC ↑	ERGAS ↓	SAM ↓
GS	425.59	9.37	0.77	0.87	5.43	5.32
PNN	414.67	9.30	0.80	0.82	4.07	2.92
GLP	416.48	9.25	0.91	0.91	3.27	2.40
MAPSMM	415.33	9.24	0.90	0.90	3.46	2.47
SFIM	417.47	9.26	0.87	0.90	3.65	2.53
CNMF	417.66	9.25	0.83	0.92	4.17	2.93
PAN-GAN	409.75	8.66	0.82	0.89	4.30	3.95
本文	416.58	9.26	0.86	0.93	3.43	2.35

注:加粗字体表示各列最优结果,↑表示值越高越好,↓表示值越低越好。

GAN 更清晰。此外,本文方法的色调整体与高光谱图像接近;PNN 方法的色调稍逊色于本文方法,但整体也接近于高光谱图像;而 CNMF、GLP、MAPSMM、SFIM、PAN-GAN 和 GS 方法融合图像的色调明显偏向于多光谱图像,光谱失真较为明显。从局部细节来看,以图 8 红框中的红顶建筑为例,CNMF、GLP、MAPSMM、SFIM 和 GS 方法建筑轮廓较为清晰,空间细节信息提升较为明显;而本文方法、PAN-GAN 和 PNN 方法融合图像的空间分辨率虽然有提升,但是建筑物边界仍然较为模糊。对于 8 种方法的光谱保真度来说:本文方法和 PNN 方法光谱保真度较高,

融合图像色调基本与高光谱图像一致,特别是红顶建筑及其周围地物的色调基本与高光谱图像一致;而 CNMF、GLP、MAPSMM、SFIM、PAN-GAN 和 GS 方法红顶建筑的色调更接近于多光谱图像,光谱失真较为严重。

对于表 5 中的定量指标,GS 方法的 STD 和 EN 好于其他 7 种方法,与定性分析中 GS 方法地物最清晰,空间细节信息丰富相一致。本文方法的 STD 和 EN 低于除 PNN 外的其他几种方法,说明本文方法对于以建筑物为主的场景,在细节信息注入方面仍有提升空间。由于空间细节信息方面提升不佳,本文方法的 UIQI 指标弱于 GLP、MAPSMM 和 SFIM 方法。对于 CC 指标,本文方法优于其他 7 种方法,说明本文方法与待融合图像之间有着较高的相关性,与定性评价中的分析相符合。而 CNMF 的 ERGAS 和 SAM 指标均取得了最优的评价,但是从定性评价中可以看出,CNMF 的光谱与高光谱图像仍然有差距,但是好于 GLP、MAPSMM、SFIM、PAN-GAN 和 GS 方法。因此,本文方法对于以建筑为主的场景,在空间细节注入方面仍有提升空间,但光谱保真度很好。

4.5 光谱曲线评价

光谱曲线评价是评价高光谱图像融合质量的一个重要指标。本文选取研究区中有代表性的地物进行光谱曲线评价。研究区内主要地物(建筑、道路、山体、植被、水体和裸地)的光谱曲线如图 9 所示。通过比较融合图像和原始高光谱图像的光谱曲线形状以

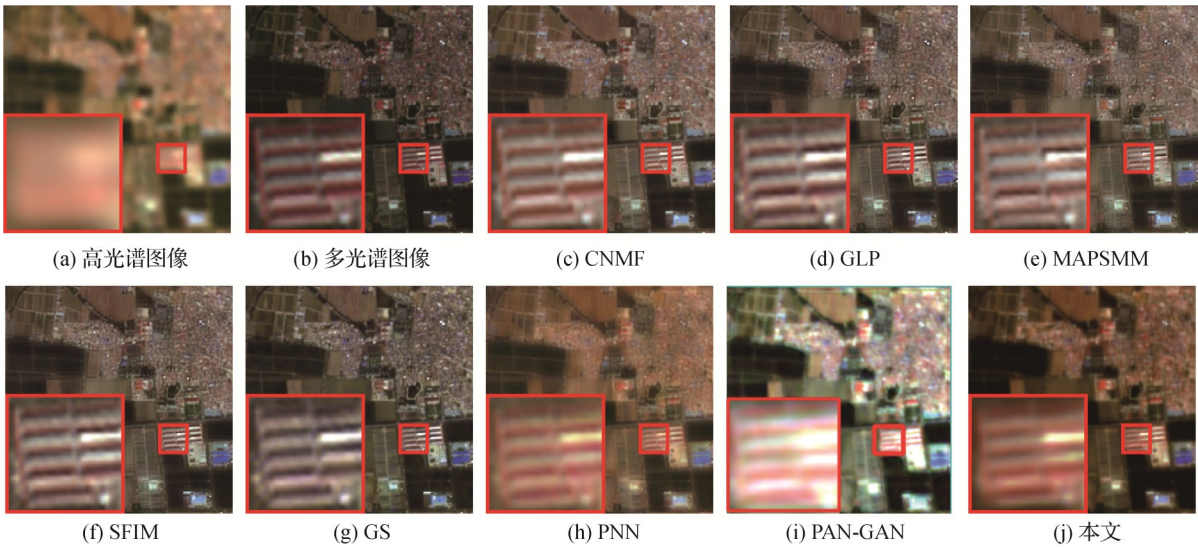


图 8 区域 3(建筑、农田和裸地)的融合结果

Fig. 8 The fusion results of Region 3 (buildings, farmland and bare land) ((a) hyperspectral image; (b) multispectral image; (c) CNMF; (d) GLP; (e) MAPSMM; (f) SFIM; (g) GS; (h) PNN; (i) PAN-GAN; (j) ours)

表5 区域3(建筑、农田和裸地)融合结果的数值评估
Table 5 Numeric assessment of Region 3 (buildings, farmland and bare land) fusion results

方法	STD ↑	EN ↑	UIQI ↑	CC ↑	ERGAS ↓	SAM ↓
GS	514.40	9.31	0.70	0.87	4.88	3.88
PNN	497.93	9.21	0.75	0.93	3.46	2.71
GLP	507.12	9.24	0.85	0.91	3.08	2.86
MAPSMM	504.04	9.18	0.89	0.92	2.77	2.95
SFIM	507.17	9.24	0.84	0.92	3.16	2.84
CNMF	504.74	9.14	0.78	0.91	2.56	2.54
PAN-GAN	498.93	8.91	0.90	0.89	3.15	3.57
本文	501.23	9.14	0.80	0.95	2.94	2.73

注:加粗字体表示各列最优结果,↑表示值越高越好,↓表示值越低越好。

及光谱曲线的数值差异来评价融合图像的光谱质量。

从以上定性、定量和光谱曲线三方面的评价可见,无论是波段的平均值还是各波段的结果,本文方

法在空间细节提升和光谱保真之间取得较好平衡。本文方法利用了空间与光谱双判别器网络,其中,光谱判别网络根据输入的融合图像和原始高光谱图像判断图像的光谱保真度,并反向传播更新生成网络的参数;空间判别网络根据输入的融合图像和原始高空间分辨率图像判断图像的空间细节情况,并反向传播更新生成网络的参数。本文方法综合融合图像的空间信息损失和光谱信息损失,提高融合图像空间细节信息的同时保持高的光谱保真度。

从图9可以看出,本文方法在各种地物中的光谱曲线(深蓝色)与原始高光谱图像的光谱曲线(红色)十分接近,每个波段的数值与高光谱图像相差很小,说明融合结果保留了原始高光谱图像的光谱特性。本文方法相比于其他6种方法,特别是在红外波段更接近于高光谱图像的光谱曲线。以建筑、道路和水体地物为例,本文方法的光谱曲线在近红外至短波红外波段与高光谱图像的光谱曲线差距最小,展现了很好的光谱保真度。

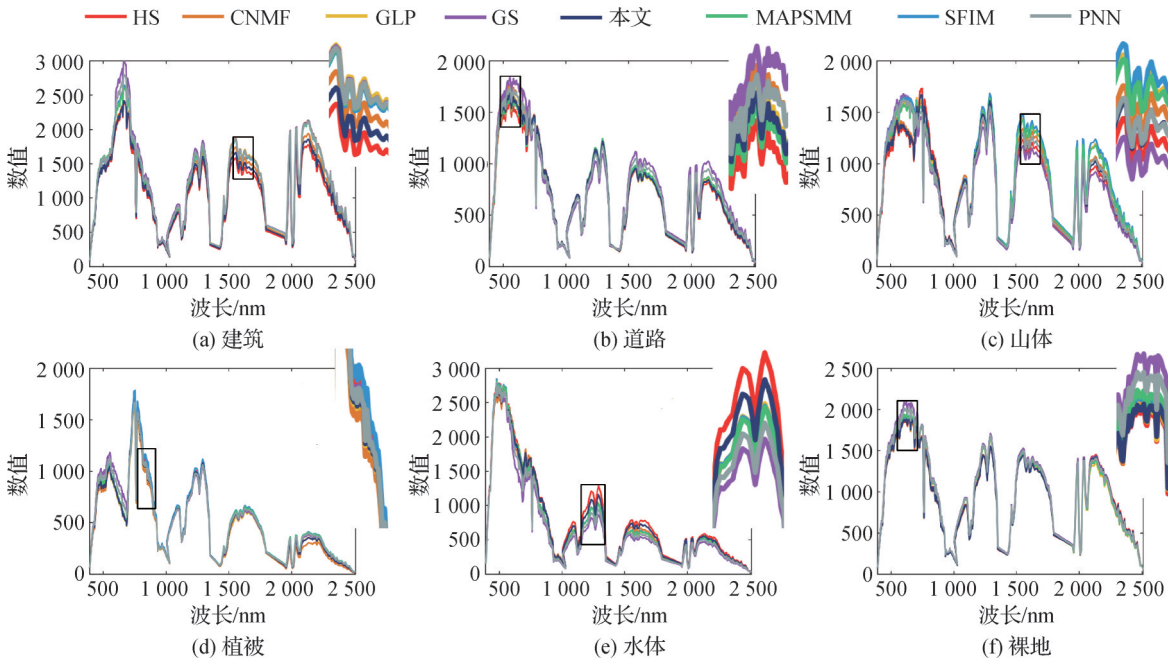


图9 各研究区主要地物的光谱曲线
Fig. 9 The spectral curves of main ground objects in each studyarea
((a) architectures; (b) roads; (c) mountains; (d) plants; (e) water; (f) bare land)

4.6 泛化实验

为了验证本文方法的泛化性能,利用公开的意大利Pavia大学高光谱数据集进行泛化实验,该数据集是由ROSIS传感器在意大利北部Pavia地区收集的。在实验中去除了该数据集12个噪声带,使用了

103个谱带。该数据集的空间分辨率为1.3 m,图像大小为610×340像素。为了使泛化结果更具有说服力,还与PNN、PAN-GAN这两种深度学习方法的结果进行比较,分析本文方法的性能。

本文方法、PAN-GAN和physics-based GAN方法

在 Pavia 大学数据集的泛化结果如图 10 所示。可以看出,本文方法、PAN-GAN 和 physics-based GAN 方法在空间分辨率方面都表现出明显的提升。尽管 PNN 方法相较待融合高光谱图像有一定的空间分辨率提升,但与其他方法相比,提升较为有限。此外,本文方法在色调方面明显接近高光谱图像,而 PAN-GAN 和 PNN 方法的色调稍显逊色,光谱失真较为明显。

对于表 6 中的定量指标,本文方法与 PAN-GAN 的 STD 和 EN 都呈现出较好的结果,这与定性分析中 3 种方法所展现的地物清晰度和空间细节信息丰富的特点是一致的。然而,由于 PNN 方法的结构只包含几层卷积层,其泛化性能较差,难以有效地适应其他数据。本文方法在 UIQI、CC、EN、ERGAS、SAM 方面表现最佳,这表明本文方法在图像的光谱保持和

空间分辨率提升两方面都具有较好的泛化性,能够较好地应用于其他数据集上。因此,从定性评价和定量指标分析来看,本文方法在非训练数据集上也能保持较好的融合效果,这表明其具有良好的泛化性能和稳定性。

表 6 Pavia 大学高光谱数据集泛化结果的数值评估
Table 6 Numeric assessment of the Pavia University hyperspectral dataset generalization results

方法	STD ↑	EN ↑	UIQI ↑	CC ↑	ERGAS ↓	SAM ↓
PNN	33.87	6.81	0.76	0.83	7.23	6.73
PAN-GAN	56.40	7.57	0.91	0.89	6.53	6.14
本文	54.11	7.84	0.96	0.93	5.38	5.09

注:加粗字体表示各列最优结果,↑表示值越高越好,↓表示值越低越好。

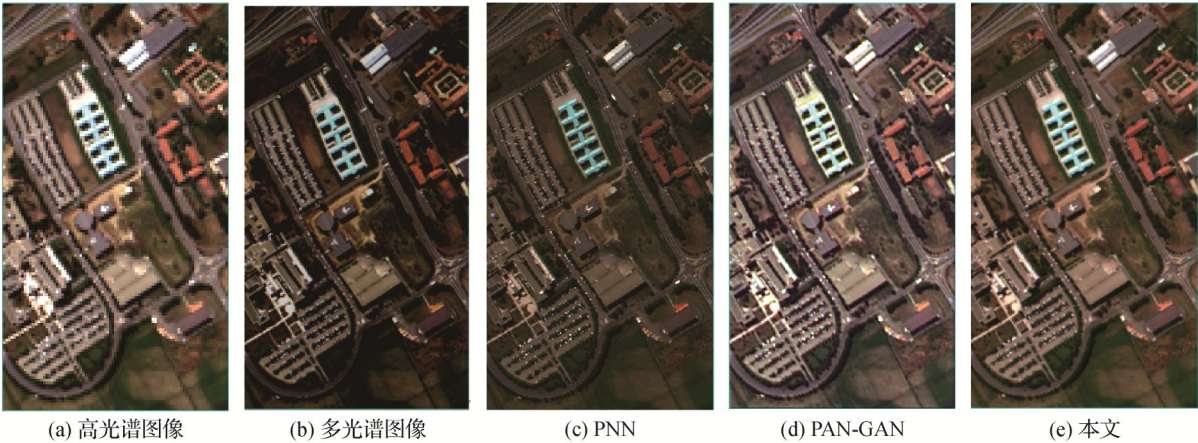


图 10 Pavia 大学高光谱数据集的泛化结果

Fig. 10 The generalization results of the Pavia University hyperspectral dataset
((a) hyperspectral image; (b) multispectral image; (c) PNN; (d) PAN-GAN; (e) ours)

4.7 消融实验

为了验证本文模型设计的有效性,进行了消融实验。使用 Sentinel-2 和 GF-5 卫星数据集中具有较为复杂地形的建筑、农田和裸地区域进行实验,以本文提出的 2DDRGAN 方法为基础架构,通过分别只保留空间判别器 (single space discriminator deep residual GAN network, S-space DDRGAN) 和只保留光谱判别器 (single spectral discriminator deep residual GAN network, S-spectral DDRGAN) 来验证模型设计的有效性。

从图 11 中可以看出,本文方法 (2DDRGAN) 在空间分辨率和光谱分辨率上都表现出较高的水平,显示出较好的图像质量和细节提升能力。而 S-space DDRGAN 方法在空间保持上表现较好。但在

空间分辨率的提升上有限。

Sentinel-2 和 GF-5 数据集的消融实验定量评价结果如表 7 所示。

表 7 Sentinel-2 和 GF-5 数据集消融结果的数值评估
Table 7 Numeric assessment of Sentinel-2 and GF-5 datasets ablation results

方法	STD ↑	EN ↑	UIQI ↑	CC ↑	ERGAS ↓	SAM ↓
S-space DDRGAN	488.24	8.65	0.65	0.72	13.45	9.90
S-spectral DDRGAN	470.21	7.90	0.78	0.93	3.14	3.81
本文	501.23	9.14	0.80	0.95	2.94	2.73

注:加粗字体表示各列最优结果,↑表示值越高越好,↓表示值越低越好。

根据表 7 可知, S-spectral DDRGAN 方法在 UIQI、CC、ERGAS 和 SAM 指标上表现较好, 证明了仅有光谱判别网络时光谱信息的良好保留。然而, 其对空间信息的提升相对较差。同时, S-space DDRGAN 方法在 STD 和 EN 方面表现优秀, 证明了仅有空间判别网络时在提升图像空间分辨率方面的有效性。相较之下, 本文方法在大多数评价指标上

均表现出优越性能, 具有较高的 STD、EN、UIQI、CC 值和较低的 ERGAS 和 SAM 值。这与定性评价中所观察到的实验结果是一致的, 同时也证实了在提升图像细节和光谱信息保持方面, 比单一判别器的方法更为出色。因此, 本文方法综合了两个判别器的优势, 能够更好地提高图像的空间分辨率并保持光谱信息的完整性。

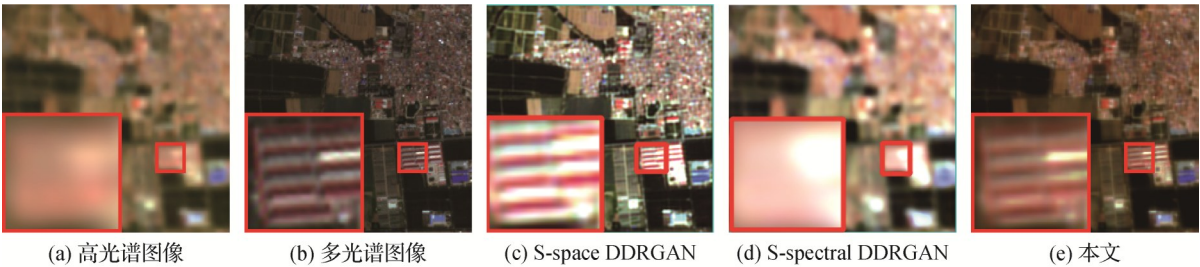


图 11 Sentinel-2 和 GF-5 数据集的消融结果

Fig. 11 The ablation results for Sentinel-2 and GF-5 datasets

((a) hyperspectral image; (b) multispectral image; (c) S-space DDRGAN; (d) S-spectral DDRGAN; (e) ours)

4.8 模型分析

本文提出了一种基于双判别器深度残差 GAN 网络的高光谱图像空谱融合方法 2DDRGAN, 该方法与对比方法 PAN-GAN 都是从生成对抗网络 GAN 衍生出来的网络, 因此, 本文使用 Sentinel-2 和 GF-5 卫星数据对这两种网络模型进行了对比分析, 同时计算了模型参数量与所需计算力, 计算过程中采用的训练集尺寸为 64×64 像素和 192×192 像素的图像。两种深度学习的模型参数量与所需计算力如表 8 所示。

从表 8 可以看出, 本文方法的模型参数量与所需计算力 FLOPS (floating point operations per second) 比 PAN-GAN 方法都要大, 这是由于 PAN-GAN 本身只有一个判别器, 而本文方法为了使融合图像同时具有更高的空间分辨率和光谱分辨率加入了双判别器。本文模型相对于对比方法在图像的空间和光谱分辨率上有所提高, 但是以增加模型的复杂度为代价。后续将进一步探究在保证模型融合效果的同时, 降低模型复杂度的方法。

表 8 两种深度学习的模型的参数量与所需计算力
Table 8 Two types of deep learning model parameters and required computational power

方法	参数量	FLOPS
PAN-GAN	875 855	3.23×10^{10}
本文	1 259 337	4.65×10^{10}

5 结 论

本文对多光谱图像波段范围之外的高光谱波段, 采用波段自适应的分组融合策略, 即利用相关系数作为波段分组标准, 将高光谱图像与多光谱图像的波段进行分组, 既解决了两个多波段图像不便于直接融合的问题, 同时也间接提升了融合图像的光谱保真度。

结合生成对抗网络和残差网络的特性, 本文提出了一种基于生成对抗网络的高光谱图像空谱融合方法—双判别器深度残差 GAN (2DDRGAN)。通过生成网络、光谱判别网络和空间判别网络的设计, 该方法无需进行大量的数据标签制作, 待融合数据本身就是参考标签, 利用双判别器来同时控制融合图像的空间信息损失和光谱信息损失, 平衡了图像的空间和光谱特性。从定性、定量以及融合图像的光谱曲线 3 个角度, 与其他 7 种融合方法进行比较评价。不同地物类型的实验结果表明本文方法得到的融合结果在空间分辨率方面有一定的提升, 不同地物的光谱保真度要优于所对比的传统方法和深度学习算法。此外, 从光谱曲线的角度来看, 本文方法融合结果的光谱曲线与原始高光谱图像接近, 是一种光谱信息高保真的空谱融合深度学习算法。最后,

本文的泛化实验验证了该网络在未训练过的异源数据上表现良好,消融实验也证明了本文方法能够充分综合两个判别器网络的优势,在图像细节信息提升和光谱信息保持方面呈现出优越的性能。

参考文献(References)

- Cui Y R, He B B, Zhang Y and Li M. 2015. Fusion of hyperspectral and multispectral data using nonnegative matrix factorization. *Remote Sensing Technology and Application*, 30(1): 82-91 (崔艳荣, 何彬彬, 张瑛, 李蔓. 2015. 非负矩阵分解融合高光谱和多光谱数据. *遥感技术与应用*, 30(1): 82-91) [DOI: 10.11873/j.issn.1004-0323.2015.1.0082]
- Dalla Mura M, Vivone G, Restaino R, Addesso P and Chanussot J. 2015. Global and local Gram-Schmidt methods for hyperspectral pansharpening//2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Milan, Italy: IEEE: 37-40 [DOI: 10.1109/IGARSS.2015.7325691]
- Dong W S, Fu F Z, Shi G M, Cao X, Wu J J, Li G Y and Li X. 2016. Hyperspectral image super-resolution via non-negative structured sparse representation. *IEEE Transactions on Image Processing*, 25(5): 2337-2352 [DOI: 10.1109/TIP.2016.2542360]
- Fang S and Xu M. 2023. Hyperspectral and multispectral image fusion focused on error compensation. *Journal of Image and Graphics*, 28(1): 277-289 (方帅, 许漫. 2023. 面向误差补偿的高光谱与多光谱图像融合. *中国图象图形学报*, 28(1): 277-289) [DOI: 10.11834/jig.220568]
- Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A and Bengio Y. 2014. Generative adversarial nets//Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, Canada: MIT Press: 2672-2680 [DOI: 10.48550/arxiv.1406.2661]
- He K M and Sun J. 2015. Convolutional neural networks at constrained time cost//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE: 5353-5360 [DOI: 10.1109/CVPR.2015.7299173]
- Hoersch B. 2017. Sentinel-2 mission status//Proceedings of the 19th EGU General Assembly Conference. Vienna, Austria: EGU
- Li S T, Dian R W, Fang L Y and Bioucas-Dias J M. 2018. Fusing hyperspectral and multispectral images via coupled sparse tensor factorization. *IEEE Transactions on Image Processing*, 27(8): 4118-4130 [DOI: 10.1109/TIP.2018.2836307]
- Liu N, Li L, Li W, Tao R, Fowler J E and Chanussot J. 2021. Hyperspectral restoration and fusion with multispectral imagery via low-rank tensor-approximation. *IEEE Transactions on Geoscience and Remote Sensing*, 59(9): 7817-7830 [DOI: 10.1109/TGRS.2020.3049014]
- Liu Y, Xu H P, Yi H, Shi Q P, Xia W Q and Kang J. 2016. Hyperspectral and multi-spectral data fusion based on constraint CNMF. *Journal of Applied Sciences*, 34(6): 651-660 (刘洋, 徐洪平, 易航, 施清平, 夏伟强, 康健. 2016. 基于有约束CNMF的遥感高光谱多光谱图像融合. *应用科学学报*, 34(6): 651-660) [DOI: 10.3969/j.issn.0255-8297.2016.06.001]
- Liu Y N, Sun D X, Hu X N, Liu S F, Cao K Q, Chai M Y, Liao Q J, Zuo Z Q, Hao Z Y, Duan W B, Zhou W Y N, Zhang J and Zhang Y. 2020. Development of visible and short-wave infrared hyperspectral imager onboard GF-5 satellite. *Journal of Remote Sensing*, 24(4): 333-344 (刘银年, 孙德新, 胡晓宁, 刘书锋, 曹开钦, 柴孟阳, 廖清君, 左志强, 郝振贻, 段微波, 周魏乙诺, 张静, 张营. 2020. 高分五号可见短波红外高光谱相机设计与研制. *遥感学报*, 24(4): 333-344) [DOI: 10.11834/jrs.20209196]
- Ma J Y, Yu W, Chen C, Liang P W, Guo X J and Jiang J J. 2020. Pan-GAN: an unsupervised pan-sharpening method for remote sensing image fusion. *Information Fusion*, 62: 110-120 [DOI: 10.1016/j.inffus.2020.04.006]
- Mallat S G. 2006. A theory for multiresolution signal decomposition: the wavelet representation//Heil C and Walnut D F, eds. *Fundamental Papers in Wavelet Theory*. Princeton, USA: Princeton University Press [DOI: 10.1515/9781400827268.494]
- Masi G, Cozzolino D, Verdoliva L and Scarpa G. 2016. Pansharpening by convolutional neural networks. *Remote Sensing*, 8(7): #594 [DOI: 10.3390/rs8070594]
- Mirza M and Osindero S. 2014. Conditional generative adversarial nets [EB/OL]. [2023-09-29]. <https://arxiv.org/pdf/1411.1784.pdf>
- 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]
- Radford A, Metz L and Chintala S. 2016. Unsupervised representation learning with deep convolutional generative adversarial networks//Proceedings of the 4th International Conference on Learning Representations. San Juan, Puerto Rico: [s.n.]
- Shah V P, Younan N H and King R L. 2008. An efficient pansharpening method via a combined adaptive PCA approach and contourlets. *IEEE Transactions on Geoscience and Remote Sensing*, 46(5): 1323-1335 [DOI: 10.1109/TGRS.2008.916211]
- Simonyan K and Zisserman A. 2015. Very deep convolutional networks for large-scale image recognition//Proceedings of the 3rd International Conference on Learning Representations. San Diego, USA: [s.n.]
- Sun Y Z, Jiang G W, Li Y D, Yang Y, Dai H S, He J, Ye Q H, Cao Q, Dong C Z, Zhao S H and Wang W H. 2018. GF-5 satellite: overview and application prospects. *Spacecraft Recovery & Remote Sensing*, 39(3): 1-13 (孙允珠, 蒋光伟, 李云端, 杨勇, 代海

- 山, 何军, 叶擎昊, 曹琼, 董长哲, 赵少华, 王维和. 2018. “高分五号”卫星概况及应用前景展望. 航天返回与遥感, 39(3): 1-13 [DOI: 10.3969/j.issn.1009-8518.2018.03.001]
- Tong Q X, Zhang B and Zhang L F. 2016. Current progress of hyperspectral remote sensing in China. Journal of Remote Sensing, 20(5): 689-707 (童庆禧, 张兵, 张立福. 2016. 中国高光谱遥感的前沿进展. 遥感学报, 20(5): 689-707) [DOI: 10.11834/jrs.20166264]
- Vivone G, Addesso P, Restaino R, Mura M D and Chanussot J. 2019. Pansharpening based on deconvolution for multiband filter estimation. IEEE Transactions on Geoscience and Remote Sensing, 57(1): 540-553 [DOI: 10.1109/TGRS.2018.2858288]
- Vivone G, Restaino R, Licciardi G, Dalla Mura M and Chanussot J. 2014. MultiResolution analysis and component substitution techniques for hyperspectral pansharpening//2014 IEEE Geoscience and Remote Sensing Symposium. Quebec City, Canada: IEEE: 2649-2652 [DOI: 10.1109/IGARSS.2014.6947018]
- Wang M, He L X, Cheng Y F and Chang X L. 2018. Panchromatic and multi-spectral fusion method combined with adaptive Gaussian filter and SFIM model. Acta Geodaetica et Cartographica Sinica, 47(1): 82-90 (王密, 何鲁晓, 程宇峰, 常学立. 2018. 自适应高斯滤波与SFIM模型相结合的全色多光谱影像融合方法. 测绘学报, 47(1): 82-90 [DOI: 10.11947/j.AGCS.2018.20170421]
- Wei Q, Bioucas-Dias J, Dobigeon N and Tourneret 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]
- Wei Q, Dobigeon N and Tourneret J Y. 2014. Bayesian fusion of hyperspectral and multispectral images//Proceedings of 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Florence, Italy: IEEE: 3176-3180 [DOI: 10.1109/icassp.2014.6854186]
- Xiao J J, Li J, Yuan Q Q, Jiang M H and Zhang L P. 2021. Physics-based GAN with iterative refinement unit for hyperspectral and multispectral image fusion. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14: 6827-6841 [DOI: 10.1109/JSTARS.2021.3075727]
- Xie W Y, Lei J, Cui Y H, Li Y S and Du Q. 2020. Hyperspectral pansharpening with deep priors. IEEE Transactions on Neural Networks and Learning Systems, 31(5): 1529-1543 [DOI: 10.1109/TNNLS.2019.2920857]
- Yu D, Li K, Zhang W, Li D D, Tian X and Jiang H. 2023. Deep network-interpreted multispectral image fusion in remote sensing. Journal of Image and Graphics, 28(1): 290-304 (余典, 李坤, 张玮, 李对对, 田昕, 江昊. 2023. 可解译深度网络的多光谱遥感图像融合. 中国图象图形学报, 28(1): 290-304 [DOI: 10.11834/jig.220575]
- Zhang K, Wang M and Yang S Y. 2017. Multispectral and hyperspectral image fusion based on group spectral embedding and low-rank factorization. IEEE Transactions on Geoscience and Remote Sensing, 55(3): 1363-1371 [DOI: 10.1109/TGRS.2016.2623626]
- Zhang Z, Shi Z W and An Z Y. 2013. Hyperspectral and panchromatic image fusion using unmixing-based constrained nonnegative matrix factorization. Optik, 124(13): 1601-1608 [DOI: 10.1016/j.ijleo.2012.04.022]

作者简介

周庆泽,男,博士研究生,主要研究方向为遥感图像处理。

E-mail: qingzez@126.com

郭擎,通信作者,女,研究员,主要研究方向为遥感图像处理、遥感信息提取与应用。E-mail: guoqing@aircas.ac.cn

王海荣,男,硕士研究生,主要研究方向为遥感图像处理。

E-mail: wanghairong@aircas.ac.cn

李安,男,正高级工程师,主要研究方向为数字信号处理与信息提取、遥感卫星数据处理。E-mail: lian@aircas.ac.cn