

文章编号: 2095-4980(2023)01-0088-07

联合辐射指数的多流融合网络空谱融合方法

雷大江^{a,b}, 王纪^a, 张策^a, 张莉萍^b, 李伟生^b, 于洪^a

(重庆邮电大学 a. 计算机科学与技术学院; b. 重庆市图像认知重点实验室, 重庆 400065)

摘要: 结合遥感图像融合的特点, 提出一种联合辐射指数的多流融合生成对抗网络的遥感图像融合方法。该方法从初始的多光谱图像和全色图像中按照特定的规则提取特征图, 利用生成器子网络分别提炼输入的多源图像, 在特征域上叠加特征后通过生成器的主网络进行融合。根据遥感领域中多光谱图像和全色图像在波段上的特性, 在判别器中引入调制传递函数(MTF)来判别融合图像的光谱信息和空间结构信息。为评估所提方法的有效性, 进行视觉分析, 并与其他算法进行客观评价的比较。实验结果表明, 该方法在视觉效果和客观评价上优于其他算法。

关键词: 遥感图像融合; 生成对抗网络; 辐射指数; 多流融合

中图分类号: TN911.73

文献标志码: A

doi: 10.11805/TKYDA2020526

Spatial-spectral fusion method of multi-stream fusion network based on joint radiometric indices

LEI Dajiang^{a,b}, WANG Ji^a, ZHANG Ce^a, ZHANG Liping^b, LI Weisheng^b, YU Hong^a

(a.College of Computer; b.Chongqing Key Laboratory of Image Cognition, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: A remote sensing image fusion method is proposed based on multi-stream fusion generative adversarial network with radiation index. This method extracts feature maps from the Multi-Spectral(MS) image and Panchromatic(PAN) image according to specific rules, and uses the generator sub-network to refine the input multi-source image, then fuses images through the main network of the generator after superimposing the features. According to the characteristics of multi-spectral and panchromatic images in the remote sensing field, the Modulation Transfer Function(MTF) is introduced into the discriminator to simultaneously discriminate the spectral information and spatial structure information of the fused images. In order to evaluate the effectiveness of the proposed method, its visual analysis is conducted and it is compared with other algorithms for objective evaluation. Experimental results show that this method is superior to other algorithms in visual effects and objective evaluation.

Keywords: remote sensing image fusion; generative adversarial network; radiation index; multi-stream fusion

目前遥感图像已广泛用于许多领域, 如图像分类^[1]、目标追踪、环境监测、土地覆盖制图^[2]等。由于传感器硬件条件和技术的限制, 现有遥感卫星仅能获取低空间分辨率的多光谱图像和高空间分辨率的全色图像, 单独使用其中一种图像并不能满足实际需求。为得到既满足高空间分辨率又满足高光谱分辨率的图像, 将全色(PAN)图像和多光谱(MS)图像进行了融合^[3]。近几十年, 虽然已有很多图像融合方法被提出, 但保持融合图像的光谱信息和结构信息仍是一项具有挑战性的任务。本文提出一种联合辐射指数的生成对抗网络多流融合方法, 该方法引入多光谱图像的辐射指数, 获取多源图像, 再将输入的多源图像利用子网络分别提取特征, 将多源图像逐个经过子网络处理后在特征域上叠加, 这一多流融合框架可以通过增加子网络的方式扩展更多的特征输入。鉴于生成对抗网络在图像处理时的优异表现, 本文利用生成对抗网络学习融合中的映射关系, 根据多源图像的特点,

收稿日期: 2020-10-14; 修回日期: 2020-11-19

基金项目: 国家自然科学基金资助项目(61972060; 61802148; U1401252); 重庆市杰出青年基金资助项目(cstc2014jcyj40001); 重庆市留学人员创新创业基金资助项目(cx2018120)

适当调整判别器和生成器的输入结构，使整体网络框架更适用于多源图像融合，更符合遥感图像融合特性。

1 相关工作

1.1 遥感图像辐射指数

辐射指数由遥感图像光谱波段间的线性或非线性组合构成，能够作为反映特定目标的相对丰度和活性的辐射量值，常见的有植被指数、水域指数、土壤指数等。Masi 等利用先验假设的非线性辐射指数归一化以及卷积神经网络进行遥感图像融合^[4]，验证了网络中的特征图与某些特定辐射指数存在显著相关性。在输入多光谱图像及全色图像的同时，加入一些针对水域、土壤、植被等相关的辐射指数，可以很好地提升融合图像的质量。

以 WV-3(WorldView-3)卫星为例，一些非线性辐射指数的计算方式如下：

水域指数：

$$NDWI_{wv-3} = \frac{Coastal - Nir2}{Coastal + Nir2} \quad (1)$$

植被指数：

$$NDVI_{wv-3} = \frac{Nir2 - Red}{Nir2 + Red} \quad (2)$$

土壤及非齐次特征指数(只适用于 8 层光谱波段的卫星遥感图像)：

$$\begin{cases} NDSI_{wv-3} = \frac{Green - Yellow}{Green + Yellow} \\ NHFD_{wv-3} = \frac{Rededge - Coastal}{Rededge + Coastal} \end{cases} \quad (3)$$

式(1)~(3)中各物理量的具体含义见参考文献[4]。

1.2 调制传递函数

调制传递函数(MTF)是一种评价影像对景物细部还原能力的方法，可以全面且客观地表征影像的明锐程度。任何光学影像的光量分布都可看作无数空间频率正弦波分布的线性组合，正弦分布的调制度是其振幅与平均值的比值，用正弦分布的极大与极小来计算。调制度 M 定义如下：

$$M = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \quad (4)$$

式中 I 表示亮度。正弦信号通过镜头后，它的调制度的变化是正弦信号空间频率的函数，这个函数称为 MTF。对于原来调制度为 M 的正弦光栅，如果经过镜头到达像平面的像的调制度为 M' ，则 MTF 函数值为：

$$MTF = \frac{M'}{M} \quad (5)$$

通过调制传递函数的测定，可具体分析成像系统的影像还原能力。对不同卫星图像数据进行分析，可得到多光谱图像各波段的 MTF 数值。

2 本文方法

受生成对抗网络和遥感图像辐射指数的启发，本文提出了联合辐射指数的多流融合生成对抗网络的遥感图像融合方法。对于 PAN 图像和原始的 MS 图像，首先将其同比例下采样，得到 PAN_{LR} 图像及相应的 MS_{LR} 图像；然后将 MS_{LR} 图像上采样，得到 $MS_{LR} \uparrow$ 图像。根据式(1)~(2)，得到相应的 $NDWI_{LR}$ 和 $NDVI_{LR}$ 图像；使用一阶前向差分算子得到 PAN_{LR} 图像的结构信息 ∇PAN_{LR} 图像。将 $MS_{LR} \uparrow, PAN_{LR}, NDWI_{LR}$ 和 $NDVI_{LR}$ 图像输入生成器，得到融合图像；将 $MS_{LR} \uparrow$ 和 ∇PAN_{LR} 图像输入判别器，根据 MTF 系数，将原始的 MS 图像和其结构信息输入判别器作为真实的参照，将生成的融合图像和其结构信息输入判别器作为生成图像的判别。判别器和生成器之间根据生成对抗网络的规则相互博弈，改进学习。最后将原始的 MS 图像和 PAN 图像经过同样的处理，利用学习好的生成器得到最终的融合图像。图 1 为提出的联合辐射指数的多流融合生成对抗网络的基本框架。

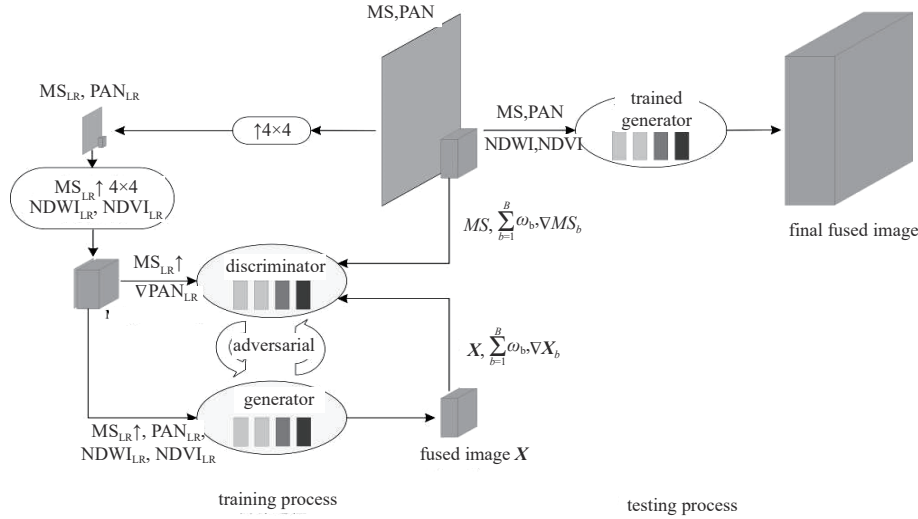


Fig.1 Generative adversarial network framework using MS_{LR} image and PAN image fusion
图 1 利用 MS_{LR} 图像与 PAN 图像进行融合的生成对抗网络框架

2.1 损失函数

用 $P \in \mathbf{R}^{H \times W}$ 表示 PAN 图像大小, $Y \in \mathbf{R}^{(H/4) \times (W/4) \times B}$ 表示低分辨率多光谱 (Low Resolution Multi Spectral, LRMS) 图像大小, $Y \uparrow$ 表示上采样后的 LRMS 图像。 $X_b = (X_1, X_2, \dots, X_B)$ 表示融合图像, 其中 $b=1, 2, \dots, B$, 表示 LRMS 图像的通道数, $\hat{X}_b = (\hat{X}_1, \hat{X}_2, \dots, \hat{X}_B)$ 表示真实的高分辨率多光谱 (High Resolution Multi Spectral, HRMS) 图像, 两者大小均为 $\mathbf{R}^{H \times W \times B}$ 。 $\nabla = (\nabla_h, \nabla_v)$ 分别表示在水平和垂直方向的一阶差分算子, $\nabla_h P, \nabla_v P$ 分别表示 PAN 图像在水平和垂直方向上的梯度信息。 W 和 V 分别表示 NDWI 和 NDVI 图像。融合过程的目标函数的一般形式如下:

$$X = f(Y \uparrow, P, W, V; \Theta) \quad (6)$$

式中: $f(\cdot)$ 表示将 $Y \uparrow, P, W, V$ 作为输入, 输出为融合图像 X 的生成对抗网络; Θ 为模型中的参数集。在式(6)中, 可以将 $f(\cdot)$ 看作一个从 $(Y \uparrow, P, W, V)$ 到 X 的映射函数。定义一个生成器 G , 它将联合分布 $p_{\text{data}}(Y \uparrow, P, W, V)$ 映射到目标分布 $p_r(X)$ 。生成器 G 和判别器 D 之间通过互相对抗产生无法与参考图像 \hat{X} 区分的融合图像 X 。

令 $M = [Y \uparrow; \nabla_h P; \nabla_v P]$; 生成器和判别器利用全色图像和多光谱图像相互博弈学习的过程, 可以表示为如下最大最小化问题:

$$\min_{\Theta_G} \max_{\Theta_D} E_{(S) \sim p_{\text{data}}(S)} [\log D_{\Theta_D}(M, Z)] + E_{(T) \sim p_{\text{data}}(T), X \sim p_r(X)} [\log(1 - D_{\Theta_D}(M, F))] \quad (7)$$

式中: $Z = [\hat{X}; \nabla_h \sum_{b=1}^B \omega_b \hat{X}_b; \nabla_v \sum_{b=1}^B \omega_b \hat{X}_b]$; $F = [G_{\Theta_G}(Y \uparrow, P, W, V); \nabla_h \sum_{b=1}^B \omega_b G_{\Theta_G}(Y \uparrow, P, W, V)_b; \nabla_v \sum_{b=1}^B \omega_b G_{\Theta_G}(Y \uparrow, P, W, V)_b]$,

ω_b 表示不同卫星传感器的 MTF 系数。

在网络学习阶段, 一旦生成器生成判别器无法判别的融合样本, 即为算法预期希望的融合图像。生成器的损失函数为:

$$L(G) = \frac{1}{N} \sum_{n=1}^N [-\log D_{\Theta_D}(M, F)] + \lambda \frac{1}{N} \sum_{n=1}^N \|\hat{X} - G_{\Theta_G}(Y \uparrow, P, W, V)\|_1 \quad (8)$$

式中: N 为训练样本的数量; λ 为一个参数, 用以平衡生成器与判别器之间的对抗损失和 ℓ_1 损失。通过分析对比 ℓ_1 范数和 ℓ_2 范数的优劣, 最终选择 ℓ_1 范数。因为 ℓ_1 范数计算的是 2 个图像各像素值的距离, 具有更好的稀疏性, 最终融合图像中的模糊会更多。判别器用于调整生成器中的融合规则, 其具体的损失函数定义如下:

$$L(D) = \frac{1}{N} \sum_{n=1}^N [\log(1 - D_{\Theta_D}(M, F))] + \frac{1}{N} \sum_{n=1}^N [\log D_{\Theta_D}(M, Z)] \quad (9)$$

2.2 多流融合的生成器结构

本文提出的多流输入的生成器体系结构如图 2(a) 所示。生成器将全色图像、上采样的多光谱图像以及辐射指数图像作为输入, 在网络底层用子网络分别对各图像提取分层特征, 再将提取的特征通过映射联合进行后续的

卷积操作。为保留完善的特征，同时兼顾内存的限制，在联合各子网络之前需将特征进行降采样操作，网络自身可通过卷积操作学到图像上采样和下采样的特点。

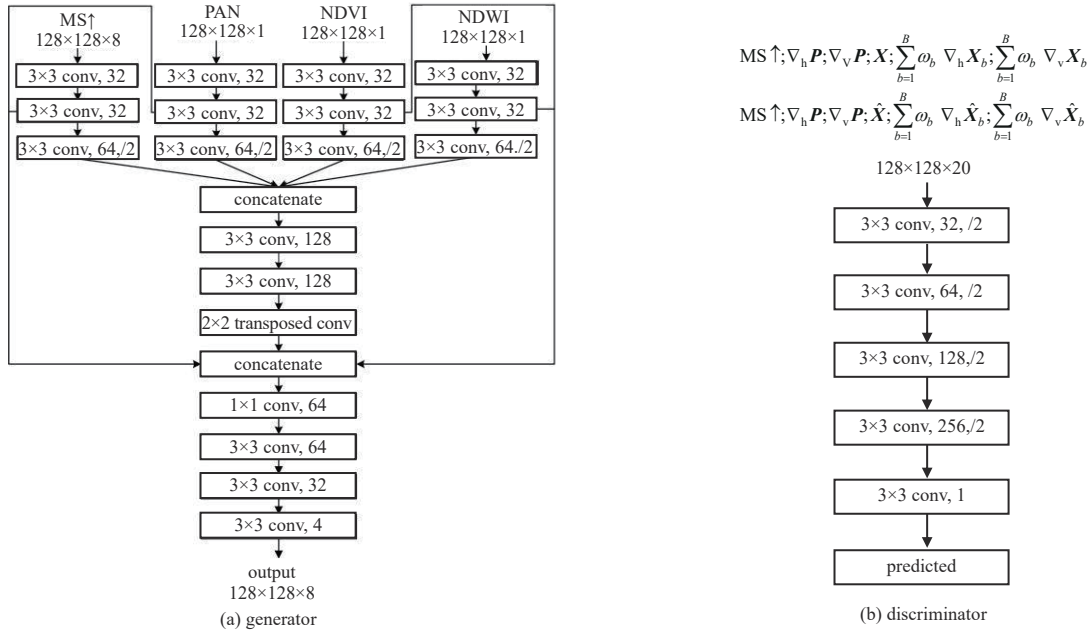


Fig.2 Detailed structure of multi-stream fusion framework
图2 多流融合框架详细结构

2.3 多流融合的判别器结构

设计的判别器网络结构是一个简单的5层卷积神经网络，详细的架构如图2(b)所示。判别器的本质是一个分类器，从第1层到第5层使用大小为3×3的、步长为2的卷积核提取特征，将提取的特征图进行分类，除了最后一层，其余卷积层都使用Leaky ReLU函数激活，最后一层使用sigmoid函数。在输入时，对于真实图像的判定，使用 $MS_{LR} \uparrow$ 图像、 ∇PAN_{LR} 图像、真实多光谱图像和根据MTF系数得到真实图像的梯度信息叠加成的分量；对于融合图的判定，相应使用 $MS_{LR} \uparrow$ 图像、 ∇PAN_{LR} 图像、融合后多光谱图像和根据MTF系数得到融合图像的梯度信息叠加成的分量。

3 实验结果分析

3.1 实验设置

将提出的方法在WV-3数据集上进行验证。WV-3数据集着重提供房屋和树木等特性的信息，且具有8个波段的光谱信息。实验包含仿真实验和真实数据实验两部分，将提出的方法与基于不同方法的经典算法和神经网络算法进行比较。所有深度学习的实验均在具有IntelXeon Gold 6148 CPU 2.40 GHz和NVIDIA Tesla V100-SXM2-16 GB的GPU服务器上完成。网络将学习来自WV-3卫星数据集的50 000组HRMS/PAN/LRMS图像，其中HRMS和PAN图像大小为128×128，LRMS图像大小为32×32。本文算法使用Tensorflow框架实现，损失函数的优化求解器使用Adam算法，批量大小设置为32，初始学习率设为0.000 2，Adam算法动量项设置为0.05，权重参数λ设置为100，针对WV-3数据集的训练时间大概在10 h左右，每次测试进行2 000×3 000的遥感图像融合生成，花费时间1.2 s~1.5 s。

3.2 仿真实验

仿真实验中，利用已有的HRMS图像作为参考，将PAN图像下采样到和HRMS图像相同尺寸，将HRMS图像下采样同等比例作为LRMS图像。利用LRMS和PAN图像进行图像融合，将融合后的图像与HRMS图像进行比较。对比分析算法包括：IHS、Brovey、ATWT-M3^[5]、动态梯度稀疏融合(Dynamic Gradient Sparse Fusion, DGSF)^[6]和分段卷积神经网络(Pansharpening by Convolutional Neural Networks, PCNN)^[4]。客观评价上选取的指标包括：光谱角度映射(Spectral Angle Mapper, SAM)^[11]、相对整体维数综合误差(Erreur Relative Global Adimensionnelle de Synthèse, ERGAS)^[7]、空间相关系数(Spatial Correlation Coefficient, SCC)^[8]、广义图像质量指标(Universal Image

Quality Index, UIQI)向 n 波段的扩展(Q_n)^[7]。图 3 为各种算法利用 WV-3 卫星数据集的仿真实验融合结果。图 3(a)和图 3(b)分别为对应的 HRMS 图像和 PAN 图像,其分辨力大小为 512×512 。从图 3 可以清晰地看出, IHS (Intensity, Hue, Saturation)算法和 Brovey 转换算法存在较为明显的光谱失真, ATWT-M3 算法的图像存在模糊和伪影。相较而言, DGSF, PCNN 和本文的多流融合框架拥有更好的视觉效果,但 DGSF 算法存在部分区域引入了多余的噪声现象, PCNN 在细节部分,如屋顶的光谱信息保留,不如本文算法。

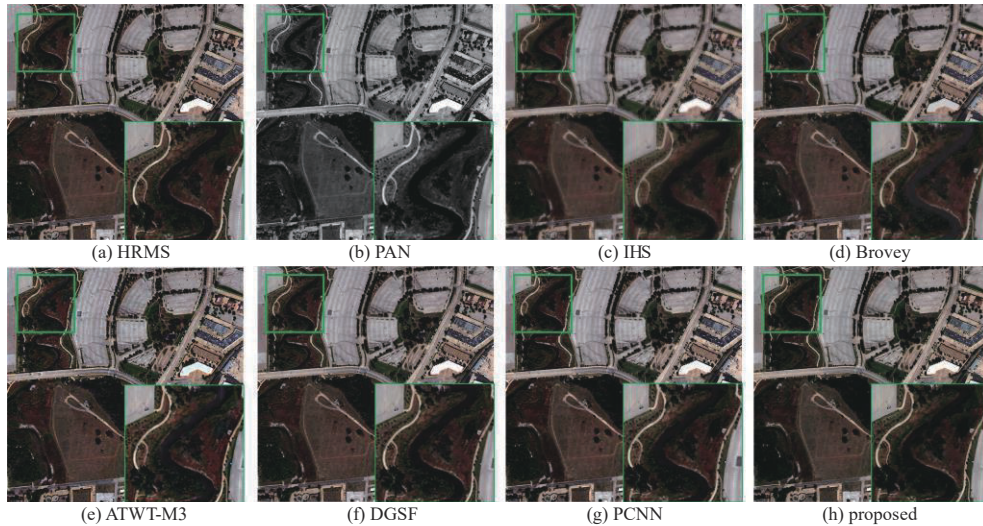


Fig.3 Comparison of fusion results of WV-3 satellite image simulation experiments
图 3 WV-3 星图像仿真实验融合结果对比

采用客观指标评价进行更加详细的对比,如表 1 所示。从各项指标上可以看出,各种融合方法的图像效果与指标反映出的信息基本保持一致,基于神经网络的 PCNN 方法和本文算法的融合效果远远优于其他传统算法。本文方法和 PCNN 方法都利用了非线性辐射指数来增加网络的输入,且本文方法的指标是各方法中最优的,说明了本文设计的多流融合网络的有效性。

表 1 基于 WV-3 卫星仿真实验融合结果

Table1 Fusion results of simulation experiments based on WV-3 dataset

method	SAM	ERGAS	Q_s	SCC
IHS	7.523 1	7.186 2	0.794 0	0.865 9
Brovey	7.389 5	6.219 9	0.782 5	0.869 9
ATWT-M3	8.216 1	7.576 9	0.726 4	0.855 4
DGSF	5.834 3	5.734 4	0.826 4	0.893 4
PCNN	5.675 7	5.132 7	0.881 0	0.926 2
proposed	5.597 0	5.077 7	0.886 4	0.934 2
reference	0	0	1	1

表 2 基于 WV-3 真实数据集实验融合结果

Table2 Fusion results of real data experiments based on WV-3 dataset

method	D_s	D_λ	QNR
IHS	0.076 3	0.077 3	0.852 3
GSA	0.082 6	0.072 3	0.851 0
BDSB	0.069 1	0.062 7	0.872 5
DGSF	0.073 9	0.063 1	0.867 7
PCNN	0.073 7	0.058 6	0.873 0
proposed	0.073 5	0.047 5	0.882 5
reference	0	0	1

3.3 真实数据实验

真实数据集实验是利用仿真实验训练好的参数在原始数据集的尺寸上进行融合,真实数据集只有原始的多光谱图像和全色图像,全色图像的分辨力是多光谱图像的 4 倍。由于没有参考图像,融合图像的效果评定通常取

决于融合图像与多光谱图像和全色图像之间的差距，因此在真实数据集实验中使用无参考质量评价指标(Quality with No Reference index, QNR)^[9]进行评价，其中 QNR 指标包含评价光谱信息损失的指标 D_λ 和评价空间信息损失的指标 D_s 。

图 4 为 WV-3 卫星图像真实数据的融合结果，图 4(a) 和图 4(b) 分别对应上采样的 LRMS 图像和 PAN 图像，其中原始 LRMS 图像分辨力大小为 128×128 ，PAN 图像分辨力大小为 512×512 。从图 4 中可以看出，IHS 方法生成的图像较为模糊，并存在光谱失真现象；GSA(Gram Schmidt Adaptive)^[10]方法生成的图像无法清晰显示出树木的轮廓；DGSF 方法和 BDSD(Band-Dependent Spatial-Detail)^[11]生成的图像可以区分出树木的轮廓，但树木之间的层次并不清晰。相较而言，PCNN 算法和本文算法有更好的视觉效果。更加详细的客观指标评价对比如表 2 所示，在 WV-3 卫星的真实数据融合实验中，在评判结构信息损失的指标上，BDSD 方法取得最优结果，本文方法次优。本文算法由于利用了辐射指数更好地学习植被和水域区域的特征，相应的光谱信息损失的指标上，本文算法取得最优结果。同时，本文算法在综合图像质量指标 QNR 上也优于其他算法，这些指标与生成图像的视觉效果表现基本一致，结合仿真实验中的各项指标，充分说明本文方法在遥感图像融合中起到了正向作用。

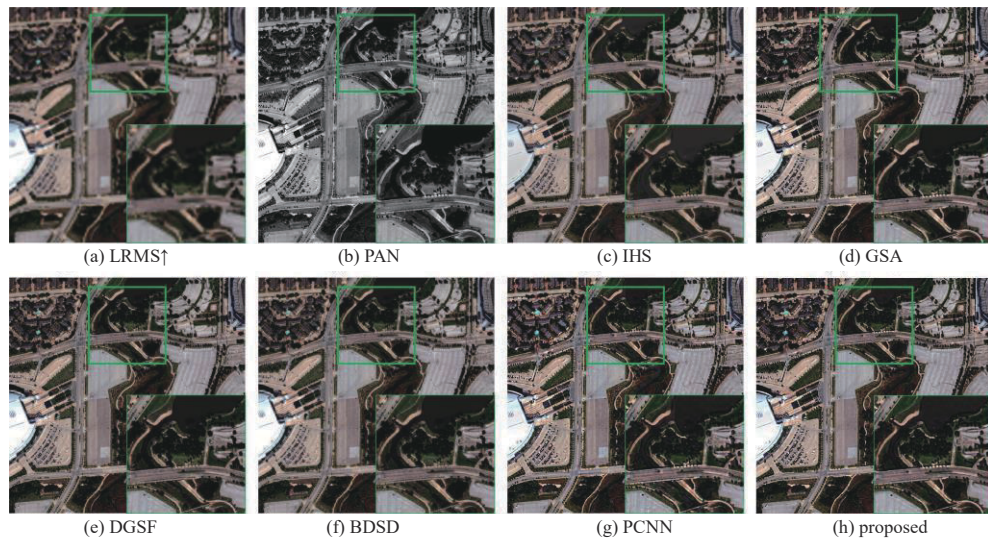


Fig.4 Comparison of real data fusion results of WV-3 satellite image
图 4 WV-3 卫星图像真实数据融合结果对比

4 结论

本文提出了一种联合辐射指数的多流融合生成对抗网络的遥感图像融合方法。引入多光谱图像辐射指数，结合生成对抗网络在图像处理中的优异表现，设计适用于遥感图像融合的生成对抗网络框架。由于加入了植被和水域辐射指数作为网络输入，生成图像融合效果在植被和水域的表现上更加优异。通过对实验结果分析和其他算法的对比，可以得出本文算法在主观视觉表现和客观指标的评比中都超越大多数先进的遥感图像融合算法。未来的工作中，所提出的方法将在两个方向上扩展：一方面，本文提出的方法理论上能适用于其他方面的多源图像融合，如红外图像和可视图像的融合等，只需要根据领域知识设计新的生成器和判别器；另一方面，在本文方法基础上，使用 2 个判别器分别判别光谱信息和空间信息的保留程度，可能会进一步提升融合图像的质量。

参考文献：

- [1] VIVONE G,ALPARONE L,CHANUSSOT J,et al. A critical comparison among pansharpening algorithms[J]. IEEE Transactions on Geoscience Remote Sensing, 2014,53(5):2565–2586.
- [2] WANG Peng,ZHANG Lei,ZHANG Gong,et al. Superresolution land cover mapping based on pixel-,subpixel-,and superpixel-scale spatial dependence with pansharpening technique[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2019,12(10):4082–4098.
- [3] GHASSEMIAN H. A review of remote sensing image fusion methods[J]. Information Fusion, 2016(32):75–89.
- [4] MASI G,COZZOLINO D,VERDOLIVA L,et al. Pansharpening by convolutional neural networks[J]. Remote Sensing, 2016, 8(7):594.
- [5] RANCHIN T,WALD L. Fusion of high spatial and spectral resolution images:the ARSIS concept and its implementation[J].

- Photogrammetric Engineering and Remote Sensing, 2000,66(2):49-61.
- [6] CHEN Chen, LI Yeqing, LIU Wei, et al. Image fusion with local spectral consistency and dynamic gradient sparsity[C]// 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE, 2014:2760-2765.
- [7] WALD L. Data fusion: definitions and architectures: fusion of images of different spatial resolutions[M]. Paris, France: Presses des mines, 2002.
- [8] 张新曼, 韩九强. 基于视觉特性的多尺度对比度塔图像融合及性能评价[J]. 西安交通大学学报, 2004, 38(4):380-383. (ZHANG Xinman, HAN Jiuqiang. Image fusion of multiscale contrast pyramid based vision feature and its performance evaluation[J]. Journal of Xi'an Jiaotong University, 2004, 38(4):380-383.)
- [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] AIAZZI B, BARONTI S, SELVA M. Improving component substitution pansharpening through multivariate regression of MS+Pan data[J]. IEEE Transactions on Geoscience Remote Sensing, 2007, 45(10):3230-3239.
- [11] GARZELLI A, NENCINI F, CAPOBIANCO L. Optimal MMSE pan sharpening of very high resolution multispectral images[J]. IEEE Transactions on Geoscience Remote Sensing, 2008, 46(1):228-236.

作者简介:

雷大江(1979-), 男, 博士, 教授, 主要研究方向为机器学习、模式识别. email:leidj@cqupt.edu.cn.

王纪(1993-), 男, 在读硕士研究生, 主要研究方向为遥感图像处理、模式识别.

张策(1994-), 男, 在读硕士研究生, 主要研究方向为遥感图像处理、机器学习.

张莉萍(1983-), 女, 博士, 讲师, 主要研究方向为遥感图像处理、模式识别.

李伟生(1975-), 男, 博士, 教授, 主要研究方向为模式识别、智能信息处理.

于洪(1972-), 女, 博士, 教授, 主要研究方向为可信人工智能、数据挖掘和智能信息处理.