

doi: 10.6046/zrzyyg.2021433

引用格式: 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 J W, Wang Z P, Hu P. A review of pansharpening methods based on deep learning[J]. Remote Sensing for Natural Resources, 2023, 35(1): 1-14.)

基于深度学习的空谱遥感图像融合综述

胡建文, 汪泽平, 胡 佩

(长沙理工大学电气与信息工程学院, 长沙 410114)

摘要: 随着遥感技术的快速发展与广泛应用,对获取的遥感图像质量有了更高的要求。但是,难以直接获得高空间分辨率多光谱遥感图像。为了结合不同成像传感器的信息,获得高质量的图像,图像融合技术应运而生。空谱遥感图像融合是一种获取高空间分辨率多光谱图像的有效方法,目前已有许多学者针对空谱遥感图像融合展开研究,取得了较多成果。近年来,深度学习理论得到了快速发展,广泛应用于空谱遥感图像融合。为了让学者们能够更系统地了解空谱遥感图像融合的现状,推动空谱遥感图像融合的发展,首先对常用的遥感卫星作了介绍,并简单总结了传统的经典空谱图像融合算法;其次,从监督学习、无监督学习和半监督学习 3 个方面,重点对基于深度学习的空谱图像融合算法进行了阐述,还对损失函数进行了描述与分析;然后,为了证明基于深度学习方法的优越性以及分析损失函数的影响,开展了遥感图像融合实验;最后,对基于深度学习的空谱图像融合方法进行了展望。

关键词: 遥感图像; 空谱图像融合; 深度学习; 卷积神经网络

中图法分类号: TP 391.4 **文献标志码:** A **文章编号:** 2097-035X(2023)01-0001-14

0 引言

传统的卫星所获得的单传感器图像特征比较单一,限制了其应用。因此人们用卫星搭载多源传感器来获取丰富多样的信息,这些多源信息既冗余又互补^[1]。由于传感器成像原理的限制,遥感图像的空间分辨率与光谱分辨率相互制约,单一传感器不能获得高空间分辨率的多光谱图像,满足不了现代社会多样化的需求^[2]。为了合理利用和整合多源卫星图像的信息,获得高空间分辨率的多光谱图像,空谱遥感图像融合技术应运而生^[3-4]。

图像融合是一种将多个不同成像传感器对同一物体或场景所获取的信息进行处理并融合的技术。图像融合方法不仅能合理有效地利用图像的有用信息,在提高原始图像空间分辨率的同时也提高了其光谱分辨率,同时也减小了单一成像传感器对目标物体描述的偏差与不确定性^[5]。图像融合技术不仅深化了信息的使用,也优化了图像的质量。在对高分辨率多光谱图像需求日益增多的时代,图像融合技术已广泛应用于交通^[6]、灾害监测^[7]和农业等领域^[8]。

现已提出许多空谱图像融合的方法,并有学者在综述中针对空谱图像融合算法进行了归纳整理。张良培等^[9]和李树涛等^[10]将空谱图像融合分为全色-多光谱融合、全色-高光谱融合和多光谱-高光谱融合 3 个方面进行概括与分析;张立福等^[11]在对现有的遥感图像融合研究成果的调研中,将应用于空谱图像融合的基于空间维提升的算法和基于光谱维提升的算法进行了归纳与总结;Meng 等^[12]基于元分析的思想,对 2000—2016 年间提出的不同类别的空谱图像融合方法的性能进行了评价,并阐述了空谱图像融合的发展过程;Javan 等^[13]介绍了 41 种空谱图像融合方法,将其分为成分替换法、多尺度分解法、基于变分优化法和混合方法 4 类,分析了各类方法的融合性能;Vivone 等^[14]研究了不同类别最先进的空谱图像融合算法,将经典的空谱图像融合算法与第三代空谱图像融合算法进行了比较,并为空谱图像融合算法的定量评估提供了 MATLAB 工具箱。

上述综述主要是对传统空谱图像融合方法性能的比较和优缺点的分析,如成分替换法的实现较为简单,但是融合的图像质量欠佳,光谱扭曲较严重。多尺度分解法在光谱保持方面优于成分替换法,但

收稿日期: 2021-12-13; 修订日期: 2022-05-09

基金项目: 国家自然科学基金项目“高效多任务高光谱遥感图像超分辨率及质量评价研究”(编号: 62271087)和湖南省自然科学基金项目“基于动态卷积神经网络的遥感图像融合”(编号: 2021JJ40609)共同资助。

第一作者: 胡建文(1985-),男,副教授,研究方向为图像处理、深度学习和稀疏表示。Email: hujianwen1@163.com。

是空间结构还欠佳。近年来,随着深度学习理论的发展,学者们对基于深度学习的遥感图像融合进行了大量的研究。Huang 等^[15]所提出的基于深度神经网络的遥感图像融合算法; Masi 等^[16]所提出的简单有效的基于三层网络的遥感图像融合方法 (pansharpening by convolutional neural networks, PNN)都是较早利用深度学习实现空谱图像融合的方法。相比于传统的空谱图像融合方法,基于深度学习的方法无论是光谱保持还是空间细节锐化程度都有较大提高。虽然已提出了许多基于深度学习的空谱图像融合方法,但缺乏对这些方法的全面、系统的综合分析。因此,本文将现有的基于深度学习的空谱图像融合方法分为监督学习、半监督学习和无监督学习 3 大类进行了归纳,并阐述了各类方法的

优点以及所存在的一些问题。最后,对基于深度学习的空谱图像融合的发展趋势和未来可能的研究方向进行了展望。

1 常用遥感卫星

遥感图像融合所需数据集由多种遥感卫星提供的全色图像和多光谱图像组成。常用的卫星包括国内的高分系列,国外的 GeoEye1, IKONOS 和 World-View 系列等。全色图像的空间分辨率通常是多光谱图像的 4 倍,能够提供丰富的空间信息。多种遥感卫星不仅为图像融合技术提供了必要的数据支持,并且在工业和农业等方面都发挥了巨大价值。表 1 给出了常用卫星的参数信息。

表 1 常用卫星基本参数
Tab. 1 Basic parameters of common satellites

卫星名称及图像		波段数	空间分辨率/m	光谱范围/nm	重访周期/d
GaoFen - 1	PAN	1	2	450 ~ 900	4
	MS	4	8	蓝光: 450 ~ 520, 绿光: 520 ~ 590, 红光: 630 ~ 690, 近红外: 770 ~ 890	
GaoFen - 2	PAN	1	1	450 ~ 900	5
	MS	4	4	蓝光: 450 ~ 520, 绿光: 520 ~ 590, 红光: 630 ~ 690, 近红外: 770 ~ 890	
GeoEye - 1	PAN	1	0. 46	450 ~ 800	3
	MS	4	1. 84	蓝光: 450 ~ 510, 绿光: 510 ~ 580, 红光: 655 ~ 690, 近红外: 780 ~ 920	
IKONOS	PAN	1	1	526 ~ 929	3
	MS	4	4	蓝光: 445 ~ 516, 绿光: 506 ~ 595, 红光: 632 ~ 698, 近红外: 757 ~ 853	
QuickBird	PAN	1	0. 61	405 ~ 1 053	1 ~ 6
	MS	4	2. 44	蓝光: 430 ~ 545, 绿光: 466 ~ 620, 红光: 590 ~ 710, 近红外: 715 ~ 918	
WorldView - 2	PAN	1	0. 5	450 ~ 800	1. 1
	MS	8	2	蓝光: 450 ~ 510, 绿光: 510 ~ 580, 红光: 630 ~ 690, 近红外: 770 ~ 895, 海岸: 400 ~ 450, 黄色: 585 ~ 625, 红色边缘: 705 ~ 745, 近红外 2: 860 ~ 1 040	
WorldView - 3	PAN	1	0. 31	450 ~ 800	1
	MS	8	1. 24	蓝光: 445 ~ 517, 绿光: 507 ~ 586, 红光: 626 ~ 696, 近红外: 765 ~ 899, 海岸: 397 ~ 454, 黄色: 580 ~ 629, 红色边缘: 698 ~ 749, 近红外 2: 857 ~ 1 039	

2 传统空谱遥感图像融合算法

2.1 成分替换法

成分替换法一般包括以下 3 步: 首先, 将多光谱图像变换到另一个空间, 获得多光谱图像的各种分量; 然后, 用全色图像替代其中一个分量; 最后, 通过相应的反变换得到融合图像。经典的成分替换方法包括主成分分析^[17]、Gram - Schmidt (GS) 变

换^[18]和 IHS(强度、色调、饱和度)变换^[19]等。经典成分替换法一般是全局的, 它在整个图像上以相同的方式运行。成分替换法快速简单, 由于用全色图像直接替换其中一个分量, 而全色图像包含丰富的细节, 所以这类方法融合图像的空间细节通常较好。但是成分替换法只适用于全色图像与多光谱图像之间高度相关的情况, 否则容易因图像之间光谱不匹配而产生局部差异, 从而引起显著的光谱失真。为了克服这个问题, 学者们提出了一些改进方法。如

Garzelli 等^[20]提出了一种波段相互依赖的细节注入模型(band dependent spatial – detail, BDSD),在 4 波段的多光谱图像融合处理上有较好的性能。

2.2 多尺度分解法

基于多尺度分解的图像融合方法主要包含 3 步:首先,对多源图像进行多尺度分解;然后,融合不同源图像的分解系数;最后,在融合系数上进行多尺度逆变换得到融合图像。其中,多尺度分解将原始图像分解成多个不同尺度的高频和低频系数,系数融合是根据不同的融合规则对不同尺度分解层上的高频与低频系数分别进行融合。常用的多尺度分解方法包括拉普拉斯金字塔^[21]、广义拉普拉斯金字塔(generalized laplacian pyramid, GLP)^[22]、离散小波变换^[23]、Curvelet 变换^[24]、非下采样轮廓波变换^[25]和非下采样 Shearlet 变换^[26-27]等。与成分替换法相比,多尺度分解方法不仅注入了高频空间细节,而且还较好地保持了融合图像的光谱信息,从而解决一定的光谱失真问题,但是相应地会带来空间信息丢失、造成振铃现象等问题,因此在融合图像的结构细节方面还需加强。

2.3 稀疏表示法

除了成分替换法和多尺度分解法之外,利用稀疏表示理论进行遥感图像融合也是学者们广泛关注的方向。稀疏表示法利用图像块在过完备字典上表示的稀疏性对图像进行融合。基于稀疏表示理论, Li 等^[28]首次尝试从压缩感知的角度提出了遥感图像全色锐化方法。针对高低空间分辨率图像的稀疏

表示系数不同的问题, Yin^[29]将稀疏表示超分辨率和图像融合的细节注入集成到一个框架内,提出了基于局部自适应稀疏表示的遥感图像融合方法;李成铁等^[30]提出利用冗余字典对图像进行稀疏表示,进行初步的图像超分辨率,然后再利用 GS 方法进行融合。

在实际的应用中无法直接获取字典学习所需要的高空间分辨率多光谱图像,针对这一问题,学者们提出了一些解决方法。Yin^[31]通过学习字典提取全色图像的空间细节信息,然后将提取的信息与多光谱图像融合得到高质量的融合图像; Li 等^[32]提出了一种基于学习字典稀疏表示的遥感图像融合方法,该方法从源图像中自适应学习全色图像字典和多光谱图像字典。尽管稀疏表示法可以获得较好的融合性能,但由于在融合时需要进行稀疏优化求解,导致其效率较低,限制了稀疏表示融合方法的应用。

3 基于深度学习的空谱遥感图像融合

传统的空谱遥感图像融合方法虽然已取得了一些成果,但是仍存在较多问题,具有一定的局限性。与传统的空谱遥感图像融合方法不同,基于深度学习的空谱遥感图像融合直接将原始数据输入到网络中进行训练与学习,避免了人工提取信息过程中所产生的误差,使得其性能明显提高。本文将基于深度学习的空谱遥感图像融合方法分为监督学习、无监督学习和半监督学习 3 大类进行分析与研究,如图 1 所示。

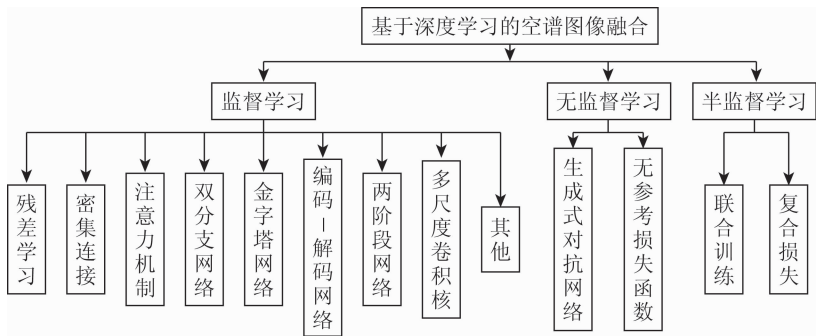


图 1 基于深度学习的空谱遥感图像融合方法分类

Fig. 1 Classification of pansharpening based on deep learning

3.1 监督学习空谱遥感图像融合

目前,大部分基于深度学习的空谱图像融合方法都采用监督学习的方式。监督学习的空谱遥感图像融合过程如图 2 所示。由于缺乏高空间分辨率的多光谱图像,先在退化的低空间分辨率图像和多光谱图像上,通过反向传播训练网络模型,然后再用训练好的模型融合原始分辨率的图像,得到最终的高空间分辨率的多光谱图像。

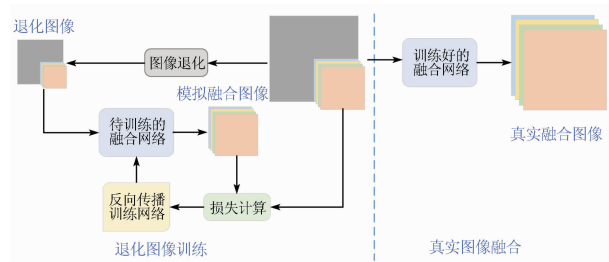
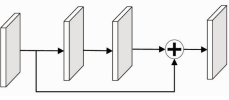
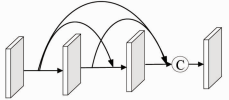
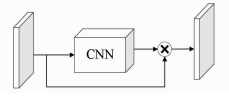
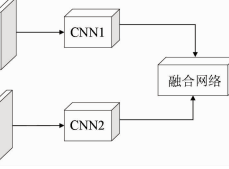
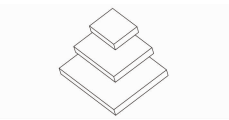
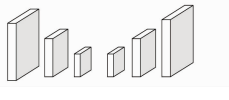
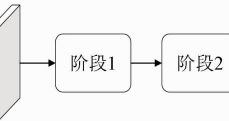
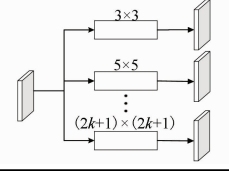


图 2 基于监督学习的空谱遥感图像融合

Fig. 2 Pansharpening based on supervised learning

目前大部分的深度学习空谱遥感图像融合方法都是基于监督学习,针对不同方法的特性,本文将其分为 9 大类:基于残差学习的方法、基于密集连接的方法、基于注意力机制的方法、基于双分支网络结构的方法、基于金字塔网络的方法、基于编码-解码网络的方法、基于两阶段网络的方法、基于多尺度卷积核的方法和其他方法。表 2 从各类方法的特点和主要网络结构进行了归纳总结。

表 2 各类监督学习空谱图像融合算法比较
Tab.2 Comparison among pansharpening methods based on supervised learning

方法类别	主要网络结构	特点
残差学习方法		采用残差连接,提高了信息的流通,避免了由于网络过深所引起的梯度消失问题和退化问题
密集连接方法		密集网络作为基础网络,采用密集连接来加强信息的传递
注意力机制方法		采用注意力机制自适应调节重要信息,提高特征提取的准确率
双分支网络方法		采用 2 个分支分别提取全色图像和多光谱图像的特征,然后利用融合网络融合所提取的特征
金字塔网络方法		利用金字塔网络对输入图像从低空间分辨率到高空间分辨率(或从高空间分辨率到低空间分辨率)提取不同尺度特征
编码-解码网络方法		由编码器和解码器构成的对称网络,编码器提取不同尺度特征,解码器还原各尺度信息
两阶段网络方法		2 个阶段都有着各自任务,发挥不同作用
多尺度卷积核方法		采用多个大小不同的卷积核对图像分别进行特征提取以获得不同感受野的特征信息

3.1.1 残差学习融合方法

基于残差学习的方法在网络中加入残差连接,使上一个残差块的信息能很好地流入到下一个残差块,提高了信息的流通,避免了由于网络过深所引起

的梯度消失和退化问题。

对于结构较深的网络模型,学者们都采用残差连接来进行信息之间的传递。早期的残差学习融合方法包括: Scarpa 等^[33]研究的利用目标自适应策略实现全色锐化的方法和 Wei 等^[34]提出的深度残差神经网络方法,都利用残差连接结构深化网络来提高网络性能。Yang 等^[35]提出了一种采用残差网络结构的深度网络模型(PanNet),该方法为了保持光谱,在网络输出中加入上采样的多光谱图像,直接将光谱信息传播到重建图像中。为了保持空间结构,在高通域而不是图像域训练网络参数。

针对现有基于深度学习的空谱图像融合方法空间增强不足的问题,Zhang 等^[36]提出一种基于梯度变换先验的残差学习网络,该方法在融合网络中采用了有效的残差学习设计。光谱信息主要由插值多光谱图像提供,空间信息由学习残差获取。不同于上述方法在融合过程中直接向上采样 4 倍,Yang 等^[37]提出了一种渐进级联深度残差网络,该网络包含 2 个残差子网络,该方法首先采用 2 次上采样操作,然后利用 2 个残差子网络学习 2 个尺度下源图像到参考图像的非线性特征映射,最终实现低空间分辨率多光谱图像和全色图像的逐步融合。

残差学习方法缓解了深度网络中的梯度消失问题,使得能够设计和训练深度神经网络,从而通过增加网络深度来提高融合性能。但当模型深度达到一定程度以后,再增加深度能带来的性能增益较少,并且过深的网络具有较大的计算量和参数量,模型的训练时间更长。

3.1.2 密集连接融合方法

随着网络结构的加深,容易造成信息的流失,为了将信息传递给后面的所有层,基于密集连接的方法将所有层密集连接起来,保证了网络中各层之间最大信息的传递。

Huang 等^[38]提出了一种用于空谱图像融合的密集连接卷积网络结构(densely connected convolutional networks for pan-sharpening,DCCNP),采用了一种有效的密集连接网络结构来解决过拟合和梯度消失问题,同时为了降低网络体系结构的复杂性,密集连接网络的设计体系结构中去掉了批量归一化层; Peng 等^[39]提出了一种基于多尺度密集网络的端到端空谱图像融合方法,设计了一种多尺度密集块从网络的浅层特征中充分提取空间和光谱信息,并设计了全局密集连接,允许从当前多尺度密集块状态直接连接到所有后续多尺度密集块状态,以提取更高级的特征。

与残差学习通过求和来与前面的特征进行融合

不同,基于密集连接的网络模型通过串联能够重用前面所有层的特征,进行进一步的特征提取,从而改善特征信息和梯度的流通。由于密集连接网络的特殊结构,宽度不会设置很大,因此需要通过增加层数来保证性能,但过多的层数会使特征图的通道数线性增加,使显存消耗爆炸式增长。

3.1.3 注意力机制融合方法

大多数融合模型都平等地处理不同特征,忽略了不同特征的差异,从而阻碍了卷积神经网络的表示能力。针对该问题,Luo 等^[40]提出了一种通道相似性注意力融合网络,该网络通过堆叠多个通道注意力残差密集块来充分利用鉴别性特征,采用通道注意力使网络关注重要特征,并对其充分利用;孔爱玲等^[41]提出一种基于先验知识的融合方法,利用先验知识来对特征进行逐像素的约束,然后利用通道注意力对特征进行调整。

上述方法中所采用的注意力机制都是通道注意力,学者们还提出了空间注意力和非局部注意力等方法。Zhang 等^[36]采用空间注意力机制沿空间维度加强了对有效特征的关注。为了提高网络提取特征的能力,Jiang 等^[42]将通道注意力机制和空间注意力机制相结合,提出了联合空间-通道混合注意模块。由于卷积运算集中于局部区域,即使在深度网络中,也很难获得全局信息。为了捕获所有像素的相似上下文依赖关系,Lei 等^[43]提出了一种高效的非局部注意力残差网络(nonlocal attention residual network, NLRNet)。

基于注意力的方法能够灵活地利用图像不同区域、不同通道甚至不同像素点之间的相似性,自适应地根据重要程度调整特征信息,增强学习能力。但通道注意力将通道内的信息直接进行全局处理,忽略了空间信息交互;而空间注意力对每个通道进行同样的处理,忽略了通道间的信息交互。非局部注意力虽然能够捕获全局信息,但其计算量大,且对显存的要求高,为了降低计算量而刻意降低通道维度会使性能下降。

3.1.4 双分支网络融合方法

基于双分支网络结构的方法采用 2 个分支分别提取全色图像和多光谱图像的特征,然后融合所提取的特征重建融合图像。

Shao 等^[44]和 Liu 等^[45]分别提出的融合网络 RSIFNN(remote sensing image fusion with deep convolutional neural network)和 TFNet(two-stream fusion network)都采用 2 个分支网络来分别捕捉多光谱图像和全色图像的显著特征;Fu 等^[46]在双分支网络中加入反馈连接来充分利用深度特征的强大表达能

力,提出具有反馈连接的双路径网络,可以携带强大的深度特征,并以反馈的方式对浅层特征进行精化。上述方法设计 2 个网络结构相同的分支来分别提取全色图像和多光谱图像的特征。有些学者采用不同的网络结构对全色图像和多光谱图像进行处理。He 等^[47]提出的光谱感知卷积神经网络中构建了一个双分支结构,包括细节分支和近似分支,该网络采用二维卷积来提取全色图像细节,采用三维卷积来提取光谱信息,减少了二维卷积神经网络模型导致的光谱失真。

由于空谱图像融合需要将信息差别很大的多光谱图像和全色图像同时作为输入,因此双分支网络能够彼此独立地将多光谱图像和全色图像的特征进行提取,但普遍认为多光谱图像中存在空间信息,全色图像中也存在光谱信息,因此双分支网络虽然能够分别提取信息,但忽略了输入图像之间的互补性,导致特征信息提取可能不充分,重建后的融合图像仍可能有空间失真或光谱失真的情况。

3.1.5 金字塔网络融合方法

金字塔是一种典型的多尺度分解方法,广泛应用于图像融合。学者们将金字塔的思想引入到卷积神经网络中构建金字塔网络,一般通过逐层上采样或逐层下采样来实现金字塔的效果。

Zhang 等^[48]提出一种端到端双向金字塔网络,该网络在 2 个独立分支逐级处理多光谱图像和全色图像,在网络的每一层,将从全色图像中提取的空间细节注入上采样的多光谱图像中;Fang 等^[49]提出了一种平行金字塔网络来实现空谱遥感图像融合,该网络利用不同的分支处理全色图像细节和多光谱图像细节,在 2 个细节分支中引入金字塔网络结构,解决了尺度差异导致的弱相关性问题的;方帅等^[50]提出了一种基于深度金字塔网络的遥感图像融合方法,该方法设计两级金字塔网络,将网络分成细节提取和光谱重构 2 部分。

金字塔网络融合方法过程中会经过多次上、下采样,通过逐层采样能够提取不同尺度下的特征信息,在一定程度上使感受野增大,特征信息的提取更加充分,并且较之一直在同一尺度上训练,金字塔融合方法能够有效减小计算量和参数量。但是任何上采样或下采样的方法都会导致信息的丢失,产生负面影响,甚至采样方法不适当的话,容易造成失真。

3.1.6 编码-解码网络融合方法

编码-解码网络是一种由编码器和解码器构成的对称网络,利用编码器进行特征提取,解码器将提取的特征进一步优化和处理。U-Net 结构是一种典型的编码-解码网络^[51]。

为了满足像素级分析对定位精度和回归过程的抽象能力的要求,Yao 等^[52]采用 U 型结构来构建网络模型,通过在网络之间建立直接连接,可以保留不同空间分辨率级别对应的图像特征,然后得到这些多分辨率图像特征与目标图像像素值之间的回归关系;Wang 等^[53]提出了一种基于多尺度深度残差网络的空谱图像融合方法,该方法网络结构也是 U 型,通过构建多层次网络来更好地利用源图像的尺度信息;Lai 等^[54]提出了一种充分利用全色图像和多光谱图像层次互补特征的融合网络,通过编码-解码结构和粗细化策略,分别有效地提取全色图像和多光谱图像的多尺度特征。

大多数基于 U-Net 的编码-解码融合网络都需要多级缩放,编码器提取特征中容易丢失边界信息,虽然会利用跳跃连接来保留原始光谱和空间信息,难以直接通过解码器进行恢复和重建,可能造成空间失真。

3.1.7 两阶段网络融合方法

两阶段网络融合方法首先在第一阶段利用如细节注入、超分辨率等方法对全色图像或多光谱图像进行第一步优化;然后,在第二阶段进行信息整合,完成融合和重建,这样能够分层次多阶段的提取空间和光谱信息。

Benzenati 等^[55]提出了一种两阶段细节注入方法,第一阶段利用广义拉普拉斯金字塔在像素级预测基于卷积神经网络的注入权重,该阶段将全色图像提取的高频细节注入到多光谱图像中,使光谱信息得到了保存;第二阶段是对空间质量的提升,利用基于深度残差网络(ResNet)^[56]的能力来预测丢失的高频细节。这个阶段在保留光谱特性的同时,在空间细节方面对融合图像进行了改进。

Hu 等^[57]提出的两阶段网络方法由超分阶段和融合阶段构成,利用超分阶段对多光谱图像进行加强,然后通过融合阶段对超分的多光谱图像注入空间细节;Zhang 等^[58]将显著性分析与深度学习相结合,提出了一个显著性级联空谱图像融合神经网络,该网络包括 2 个阶段,首先采用空洞变形卷积网络来区分对空间和光谱质量有不同需求的区域以便进行显著性分析,然后用显著性级联残差密集网络实现空谱图像融合。

虽然基于两阶段的网络在训练上也是端到端的,但是却人为的分成了 2 个阶段,而第一阶段的设计需要较强的经验和领域知识,如超分时,需要有基于超分的先验知识来进行处理。

3.1.8 多尺度卷积核融合方法

多尺度卷积核特征提取是深度学习网络中常见

的一种方法,采用多个不同大小的卷积核对图像分别进行特征提取,以获得不同范围的特征信息。除了利用不同大小的卷积核来实现多尺度卷积核外,空洞卷积也可以在不增加参数负担的情况下有效地增加感受野。

Li 等^[59]设计了一个多尺度特征提取块来提取全色图像中的空间特征,用不同大小的卷积核对全色图像进行特征提取;Peng 等^[39]和 Yuan 等^[60]都在多尺度卷积核特征提取模块中加入了残差连接,残差连接可以使变换后的特征图与原始提取的特征保持一致,有效地避免部分失真。上述方法中的卷积核在训练后固定不变,而图像每个位置的空间细节不同,用同样的滤波器对不同的空间位置进行滤波,很难得到满意的细节图像,难以达到良好的锐化效果。Hu 等^[61]提出一种多尺度动态卷积神经网络(multiscale dynamic convolutional neural network, MD-CNN),根据输入自适应生成不同大小的卷积核与全色图像进行动态卷积,提取不同尺度的特征,有效提高了网络的融合性能。

上述方法是利用不同卷积核大小来实现多尺度卷积核,Guo 等^[62]引入了一种空洞多级结构,将不同接收域下的特征与局部级联层进行级联,可以充分利用多尺度特征,在不增加计算量的前提下扩大接收域。由于每个像素的空间结构和光谱特征不同,Hu 等^[63]提出了一种空间动态模块来动态提取像素特征,该模块设计了 2 个互补的分支来提取不同的特征,其中设计了空洞卷积支路来提取多尺度特征,该支路由空洞因子为 1,2 和 4 空洞卷积构成。

不同卷积核具有不同大小的感受野,因此利用多尺度卷积核能够整合不同感受野的信息,但是大卷积核相当于多层小卷积核,计算量和参数量会明显增加。利用空洞卷积实现多尺度卷积虽然不会增加额外的计算量和参数量,但是空洞卷积对图像中较小物体的特征提取困难,且可能由于空洞丢失空间连续信息。空洞卷积增大了感受野,但过大范围的信息可能不相关。

3.1.9 其他方法

除上述方法外,学者们还提出了一些其他的基于深度学习的空谱图像融合方法。Hu 等^[64]提出了一种用于自适应图像全色锐化的深度自学习网络,和一种点扩展函数估计算法来获取多光谱图像的模糊核,并设计了一种基于边缘检测的像素匹配方法来恢复图像之间所产生的局部误配准。

有学者结合深度学习方法与传统方法来实现空谱图像融合。考虑到成分替换法和多尺度分解法的互补性,以综合它们的优势为目标。Liu 等^[65]提出

一种全色锐化权值网络对成分替换法和多尺度分解法的融合结果进行自适应加权平均,得到最终融合图像;Deng 等^[66]提出一种基于细节注入的深度卷积神经网络方法,该方法利用传统的成分替换法和多尺度分解法框架进行细节提取,然后利用神经网络进行非线性注入模型估计。2 种方法都是深度学习方法与成分替换法和多尺度分解相结合,既利用了神经网络的数据适应性与灵活性,又结合了传统方法的优势。

3.2 无监督学习空谱遥感图像融合

监督学习方法训练过程都需要参考图像,而实际中,参考的高分辨率多光谱图像不存在。为了克制用退化图像训练模型,再用训练好的模型融合原始图像,而退化图像与原始图像的分辨率和尺度不一致,导致退化图像融合过程不能完全反映原始图像融合的问题。学者们提出了直接在原始图像上训练的无监督学习遥感图像融合方法,主要包括生成式对抗网络和无参考损失函数 2 大类。

生成式对抗网络是 Goodfellow 等^[67]提出的一种基于博弈论思想,由生成网络和判别器网络 2 个神经网络组成的网络。基于生成式对抗网络的无监督学习空谱图像融合的基本框架如图 3 所示。Ma 等^[68]提出了一种基于生成式对抗网络的无监督学习方法 Pan-GAN,在该方法中,生成器分别与光谱判别器和空间判别器建立对抗性博弈,以保持融合图像丰富的光谱信息和空间信息;Zhou 等^[69]提出了一个名为“感知全色锐化”的无监督学习方法,网络结构包括生成器、重构器和鉴别器组成,该方法基于自编码和感知损失,不需要退化图像的训练。

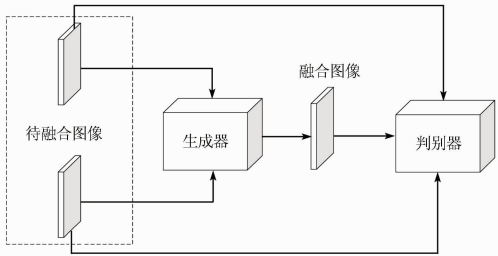


图 3 生成式对抗网络融合结构

Fig.3 Pansharpening based on generative adversarial network

学者们通过对融合图像构造不需要参考图像的损失函数,提出了一些融合方法。利用无参考图像的光谱扭曲评价指标 D_λ 和空间扭曲评价指标 D_s 作为损失函数,Xiong 等^[70]提出了一种无监督的空谱图像融合方法,该方法的损失函数不需要参考图像,因此不需要仿真数据进行训练;Luo 等^[71]提出了一种无监督训练损失函数,该函数分别利用输入多光

谱图像、全色图像以及融合图像之间的关系来设计空间约束和光谱一致性。该函数还添加了典型的无参考质量指标(quality with no reference,QNR),进一步调整光谱和空间质量。

3.3 半监督学习空谱遥感图像融合

无监督方法由于缺乏高分多光谱图像作为监督信号,导致其性能受限。为了克服该问题,学者们结合监督学习和无监督学习,提出了半监督学习融合方法,主要包括退化图像与原始图像联合训练和复合损失函数 2 大类。

杜晨光等^[72]在退化图像融合时采用常规监督训练,在真实图像融合时,由于没有高分辨率多光谱参考图像,设计光谱退化网络和空间退化网络对融合图像进行退化,再训练融合网络,该半监督方法改善了融合性能;Vitale 等^[73]提出了一种跨尺度学习策略,将在模拟图像上所获得的融合结果与真实图像一起输入网络中训练,提高了在真实图像上的融合性能。

除了真实图像和模拟图像联合训练网络来实现半监督学习融合,有学者利用复合损失函数来实现半监督学习。黄珊珊等^[74]提出了一种基于条件生成对抗网络的半监督学习图像融合方法,该方法通过设计结合 WGAN^[75]和 cGAN^[76]损失以及 PSNR 损失的复合损失函数实现半监督学习。

4 损失函数

损失函数(也称为代价函数)表示网络模型的预测值和真实值之间的差异,对模型的性能具有重要的影响。下面使用 y 来表示真实值,即参考图像; $f(x)$ 表示网络模型预测值,即融合图像; N,H 和 W 分别表示一个批次中的训练样本个数、图像的高和宽。损失函数最终的差异值是一个非负实数,通常情况下,其值越小,网络模型的性能越好。针对于遥感图像融合,已经提出了多种损失函数,如空间损失、光谱损失、空谱损失以及感知损失。

4.1 空间损失

均方差损失(mean squared error,MSE)也称为 L2 损失,是深度学习进行图像重建常用的损失函数,也广泛应用于遥感图像融合。计算预测值与真实值之间差值的平方的平均值 L_{MSE} ,即

$$L_{MSE} = \frac{1}{NHW} \sum_{n=1}^N \sum_{i=1}^H \sum_{j=1}^W |f(x_n,i,j) - y_n,i,j|^2, (1)$$

式中 i 和 j 分别为空间坐标。MSE 损失虽然简单常用,但可能会使得重建图像较模糊或平滑。

平均绝对误差损失 (mean absolute error, MAE) 也是一种常用的损失函数, 也称为 L1 损失, 是预测值与真实值之间差值的绝对值的平均, 即是绝对误差的平均值 L_{MAE} , 即

$$L_{MAE} = \frac{1}{NHW} \sum_{n=1}^N \sum_{i=1}^H \sum_{j=1}^W |f(x_n, i, j) - y_n, i, j| \quad (2)$$

式中: $\mu_{f(x_n)}$ 和 μ_{y_n} 分别表示融合图像 $f(x)$ 和参考图像 y 的局部像素的平均值; $\delta_{f(x_n)y_n}$ 表示 $f(x)$ 和 y 的协方差; $\delta_{f(x_n)}^2$ 和 $\delta_{y_n}^2$ 分别表示 $f(x)$ 和 y 的方差; C_1 和 C_2 分别是为了避免分母为 0 的常数。

为了减少全色图像与多光谱图像之间误配准的影响, Choi 等^[79] 提出了一种空间结构损失函数, 称

$$\text{grad}(X) = \frac{X - m(X)}{\sqrt{|m(X \odot X) - [m(X) \odot m(X)]| + \sigma}} \quad (4)$$

$$S = \left| \frac{m(y \odot P) - [m(y) \odot m(P)]}{\sqrt{|m(y \odot y) - [m(y) \odot m(y)]| + \sigma} \sqrt{|m(P \odot P) - [m(P) \odot m(P)]| + \sigma}} \right|^\gamma \quad (5)$$

$$L'_{\text{spatial}} = \frac{1}{N} \sum_{n=1}^N \|\{\text{grad}[f(x)_n] - \text{grad}(P)\} \odot (2 - S)\|_1 \quad (6)$$

式中: m 为均值滤波器; \odot 为逐元素相乘; y 和 P 分别为参考图像和全色图像; σ 为 10^{-10} 的常值; γ 为一个控制参数, 设置为 4; $\|\cdot\|_1$ 为 1 范数。

4.2 光谱损失

多光谱图像具有多个光谱通道, 而空间损失是基于逐点的损失, 不能反映光谱向量之间的关系。

$$L_{\text{SAM}} = \frac{1}{NHW} \sum_{n=1}^N \sum_{i=1}^H \sum_{j=1}^W \left| \frac{\langle f(x_{n,i,j}), y_{n,i,j} \rangle}{\|f(x_{n,i,j})\|_2 \|y_{n,i,j}\|_2} - 1 \right| \quad (7)$$

式中: $\langle \cdot, \cdot \rangle$ 表示内积; $\|\cdot\|_2$ 为 2 范数。

Eghbalian 等^[80] 也提出了一种光谱损失函数。该损失函数不光从 2 个向量之间的角度出发, 并且

$$L = \frac{1}{NHW} \sum_{n=1}^N \sum_{i=1}^H \sum_{j=1}^W \text{acosh} \{ \rho \text{corr}[y_{n,i,j}, f(x_{n,i,j})] - \rho \} - a + \frac{(1-a)}{4} \left(\frac{A^2}{A^2} - 1 \right)^2 \quad (8)$$

式中: A^2 和 \hat{A}^2 分别为 y 和 $f(x)$ 的大小, 第一项和第二项分别是从小角度和大小来进行设计的; a 为平衡角度和大小关系的常数; ρ 为可用于改变学习速率的常数; $\text{corr}[y_{n,i,j}, f(x_{n,i,j})]$ 是参考图像和输出图像之间的相关性, 范围在 $(-1, 1)$ 之间, 正值表示正相关, 负值表示负相关, 值越大相关性越强; \cosh 为双曲余弦函数。

Xu 等^[78] 同样提出了基于参考图像和全色图像之间的相关性来约束光谱信息的光谱损失函数

Liu 等^[77] 和 Xu 等^[78] 为了在融合过程中加强空间信息的保存, 提出了基于图像结构相似性的空间损失函数 L_{spatial} 。在此空间损失函数中, 提取局部亮度、对比度和结构对比来构造函数, 评价融合图像与参考图像之间的结构相关性。公式为:

$$L_{\text{spatial}} = \frac{1}{N} \sum_{n=1}^N \left\{ 1 - \frac{[2\mu_{f(x_n)}\mu_{y_n} + C_1][2\delta_{f(x_n)y_n} + C_2]}{[\mu_{f(x_n)}^2 + \mu_{y_n}^2 + C_1][\delta_{f(x_n)}^2 + \delta_{y_n}^2 + C_2]} \right\} \quad (3)$$

为 S3 空间损失, 该损失基于融合图像与全色图像之间的梯度差异, 并且考虑了参考多光谱图像与全色图像之间的相关性, 从而实现融合图像空间结构增强, 减少如伪影等异常空间部分。该方法中梯度 $\text{grad}(X)$ 、相关性 S 和空间损失 L'_{spatial} 的计算公式分别为:

为此, 学者们提出了一些光谱损失函数。

受评价指标光谱角映射 (spectral angle mapper, SAM) 的启发, Hu 等^[57] 提出了一种基于光谱向量相关性的损失函数 L_{SAM} , 通过计算融合图像和真实图像之间向量的余弦得到光谱相似性, 量化图像之间的光谱相似程度, 计算公式为:

利用了 2 个向量之间的大小的差异。其损失函数 L 计算公式为:

L_{spectral} , 主要考虑参考光谱图像和输出图像之间高度相关的像素区域, 再利用相关性与相关区域进行点乘得到需要进行优化的像素点。计算公式为:

$$L_{\text{spectral}} = \frac{1}{N} \sum_{n=1}^N \|[f(x_n) - y_n] \odot S\|_1 \quad (9)$$

4.3 空谱损失

为了结合空间结构特性和光谱向量相关性, 学者们提出了空谱损失函数, 最常见的方法是将上述空间损失与光谱损失加权求和, 计算公式为:

$$L = \alpha L_{\text{spatial}} + \beta L_{\text{spectral}} \quad , \quad (10)$$

式中： α 和 β 分别表示权重系数。Xu 等^[78] 提出的 S3 损失便是一个典型的空谱损失,将由式(6)所示的空间损失项和式(9)所示的光谱损失项利用都为 1 的权重系数进行相加。Hu 等^[57] 结合式(1)所示的 MSE 损失和式(7)所示的 SAM 损失构造了一种空谱损失。

4.4 感知损失

上述损失函数都是在原图像域进行处理,不能

$$L_{\text{perceptual}} = \frac{1}{N} \sum_{n=1}^N \sum_{k=1}^K \frac{1}{H_k W_k C_k} \| \varphi_k(y_n) - \varphi_k[f(x_n)] \|_F^2 \quad , \quad (11)$$

式中： φ_k 为在普通的网络模型中预训练的损失网络,一般采用 VGG19； $\| \cdot \|_F$ 为 F 范数； H_k , W_k 和 C_k 分别为第 k 层特征图的长、宽以及通道数。

5 实验结果及分析

为了对比传统方法和深度学习方法中的融合性能,以及损失函数对于基于深度学习方法的重要性,在本节给出了相应的实验。实验使用的是 8 通道 WorldView-3 卫星图像,14 400 幅大小为 32 像素 × 32 像素的图像对用于训练,36 对大小为 240 像素 × 240 像素的源图像用于测试。同时在模拟图像融合上采用空间相关系数 (spatial correlation coefficient, SCC)、SAM、相对整体维数综合误差 (erreur relative global adimensionnelle de synthèse, ERGAS)、通用图像质量指标 Q 和第 N 波段图像融合质量评价指标 Q_N 这 5 项指标来对融合结果进行评估^[83],在真实图像实验上则是利用 QNR 对融合结果进行客观定量评价^[84]。

5.1 融合方法比较

选取了传统方法中的 GS^[18], BDSD^[20] 和 GLP^[22],深度学习方法中选取典型的 PNN^[16], Pan-Net^[35], DCCNP^[38], NLRNet^[43], RSIFNN^[44], TF-Net^[45], MSDCNN^[60], MDCNN^[61] 和 SDS^[63] 一共 12 种方法进行了实验对比。表 3 是在 WorldView-3 数据集上的实验结果。从表 3 中可以看出,相比传统方法,基于深度学习的方法在整体性能上都有很大的提升。在基于深度学习的方法中,只有简单 3 层卷积的 PNN 性能上除了 SAM 和 SCC 指标比 RSIFNN 略好以外,其余都较差,说明神经网络的深度对于融合性能具有重要影响。对比 PNN,基于残差连接的 PanNet 以及基于密集连接的 DCCNP 都表现出更好的性能。基于双分支的 RSIFNN 由于其网络结构偏重对全色图像空间信息的提取,对于光谱

够捕捉到融合图像与参考图像之间的感知区别,导致缺失一些高级的空间语义信息,最终在视觉感知中会存在缺陷。因此,Johnson 等^[81] 最先提出了“感知损失”的概念。感知损失利用深度网络提取高层语义特征,通过计算高层特征上的差异作为损失,从而使生成图像能够保存高频信息。感知损失广泛应用于 GAN 中,目前一些遥感融合方法^[68,82] 中也采用了感知损失函数 $L_{\text{perceptual}}$,设损失网络共 K 层,其计算公式为:

信息的提取存在不足,因此在 SAM 指标上有所欠缺。而同样基于双分支的 TFNet 则对于多光谱图像和全色图像做同样的处理,在模拟图像实验上取得了更好的结果。说明对于多光谱图像和全色图像的信息提取同等重要。基于多尺度卷积核的 MSDC-NN 在并没有大幅度增加网络深度的基础上,通过加宽网络宽度来增强网络的表征能力。相较于采用改变网络深度以及宽度的方法来说,基于非局部注意力机制的 NLRNet 具有比较明显的优势,突出了非局部注意力机制的优势。相较于其他基于静态卷积的方法,MDCNN 是基于动态卷积的方法,从性能上看,多项指标优于其他方法。SDS 则是基于动态网络的方法,性能更是在 MDCNN 的基础上进一步提高,说明无论是动态卷积还是动态网络,动态的方法能够明显提升遥感图像融合性能。

表 3 图像融合算法性能比较^①

Tab.3 Comparison of pansharpening methods

算法	Q	SAM	$ERGAS$	SCC	Q_N	QNR
GS	0.816 8	6.711 1	4.580 6	0.827 3	0.820 7	0.867 7
BDSD	0.853 8	8.081 8	4.733 2	0.830 7	0.858 3	0.896 5
GLP	0.866 9	6.432 2	4.125 0	0.849 9	0.872 1	0.800 4
PNN	0.931 1	5.089 0	2.979 8	0.930 6	0.929 5	0.913 0
PanNet	0.938 2	4.847 9	2.851 7	0.935 1	0.935 9	0.929 0
DCCNP	0.935 1	5.076 3	2.951 5	0.935 8	0.933 6	0.915 2
RSIFNN	0.934 7	5.132 3	2.917 3	0.930 1	0.933 4	0.928 8
TFNet	0.941 6	4.613 7	2.774 3	0.943 1	0.939 9	0.918 6
MSDCNN	0.935 1	4.939 9	2.891 9	0.936 3	0.933 2	0.928 7
NLRNet	<u>0.943 3</u>	<u>4.360 2</u>	2.929 6	<u>0.948 3</u>	0.939 9	0.943 3
MDCNN	0.941 6	4.365 5	<u>2.706 3</u>	0.947 9	<u>0.940 1</u>	<u>0.948 8</u>
SDS	0.948 3	4.348 1	2.596 9	0.951 7	0.946 4	0.955 6

①加粗代表最好,下划线次之。

5.2 不同损失函数融合结果及性能分析

以经典的 PNN 为网络结构,分析不同损失函数对融合性能的影响。比较的损失函数包括 MSE, MAE, SAM^[57], SSIM^[77-78], MSE + SSIM 和 MAE + SSIM 构成的空间损失函数以及不同组合构造的空谱损失函数。损失函数比较实验结果如表 4 所示。

表 4 不同损失函数性能比较^①

Tab. 4 Performance comparison of different loss function

融合网络	类别	损失函数	<i>Q</i>	<i>SAM</i>	<i>ERGAS</i>	<i>SCC</i>	<i>Q_N</i>	<i>QNR</i>
PNN	空间损失	MSE	0.931 1	5.089 0	2.979 8	0.930 6	0.929 5	0.913 0
		MAE	0.929 1	5.133 6	3.033 7	0.928 7	0.927 6	0.895 6
		SSIM	0.821 6	15.413 1	9.500 0	0.923 6	0.675 7	0.814 5
		MSE + SSIM	0.938 1	5.131 1	2.984 0	<u>0.931 0</u>	<u>0.934 6</u>	0.932 9
		MAE + SSIM	0.937 2	5.066 7	2.964 4	0.930 9	0.934 8	0.926 6
	光谱损失	SAM	0.566 2	4.712 8	8.928 9	0.503 4	0.539 0	0.641 0
		MSE + SAM	0.878 7	5.088 3	3.594 5	0.916 6	0.877 6	0.867 1
		MAE + SAM	0.928 1	5.061 6	3.035 3	0.928 3	0.927 1	0.900 7
		MSE + SAM + SSIM	<u>0.937 7</u>	<u>5.015 0</u>	<u>2.977 7</u>	0.931 4	0.933 9	<u>0.928 2</u>
		MAE + SAM + SSIM	0.934 1	5.146 0	3.020 2	0.925 5	0.931 5	0.923 4

①加粗代表最好,下划线次之。

从表 4 中可以看出,MSE 损失函数在整体上给出了比 MAE 更好的性能。SAM 和 SSIM 损失函数都明显较差,两者都不适宜单独作为网络的损失函数。单独使用时,SAM 损失函数虽然比其他差,但 SAM 指标达到了最优,因此,SAM 损失函数有利于保持光谱信息。类似地,SSIM 损失函数的其他指标较差,但 SCC 指标较好,因此,SSIM 损失函数有利于空间结构细节重建。MSE + SSIM 对比 MSE 以及 MAE + SSIM 对比 MAE,SCC 指标都有一定的提升,真实图像实验的 QNR 指标提升更明显,进一步证明了 SSIM 损失函数能够促进空间细节的注入。

在空谱损失函数的实验中,为了简单起见,本文采取的权重系数都为 1。相比 MSE + SAM,MAE + SAM 在性能上具有明显优势,原因可能是 MSE + SAM 之间直接取权重系数为 1 并不合适。对比 MAE 的结果,MAE + SAM 的 SAM 和 QNR 指标也有一定的提升,表明 SAM 损失函数对融合网络提取光谱信息具有重要意义。

根据由 3 个损失函数构成的空谱损失函数的实验结果,MSE + SAM + SSIM 方法比 MAE + SAM + SSIM 方法的性能更好,但是对比只由 2 个损失函数构成的空谱损失函数以及单独一个损失函数的方法,简单组合损失函数并不一定能改善网络性能,原因可能是权重系数不合适或者不同损失函数之间可能存在不适配。

根据实验结果,不同损失函数对融合结果都有影响。MSE 和 MAE 作为最常用的损失函数,都给出了较好的融合性能,光谱损失函数 SAM 和空间结构相似性损失函数 SSIM 也分别针对光谱和空间有一定的优化作用,空谱损失函数也表现出比单独只使用一种损失函数更好的性能。

6 结论与展望

本研究首先介绍了 3 种传统的经典全色锐化方

法,然后介绍和分析了基于深度学习的全色锐化的关键技术,其次着重综述了深度学习在空谱图像融合研究中所获得的成果,分为监督学习、无监督学习和半监督学习 3 大类进行梳理和总结,并介绍了目前基于深度学习的全色锐化方法中所使用的损失函数,最后对多种经典融合方法以及损失函数做了定量实验和总结分析。尽管基于深度学习的空谱图像融合取得了不错的成果,但仍有许多方向值得进一步研究与探索。

1)图像迁移。现有的空谱图像融合方法的迁移能力普遍较差,不能将在一个卫星数据集上训练的模型迁移到另一个卫星数据集上,也就是需要针对每个卫星分别训练网络,如何在多个卫星图像上同时训练网络,或如何构建一个具有迁移能力的融合模型是一个值得探索的方向。

2)轻量化网络。为了获得更好的融合效果,学者们设计的模型都具有较深的网络结构,这大大提高了模型的复杂度,增加了计算量,并且对于实验设备的要求也比较高。为了在获得更好性能的同时减少计算量,利用动态等思想设计轻量化网络来实现空谱图像融合有着重要意义。

3)神经网络结构搜索。网络结构、深度和宽度等对融合性能具有重要影响,为了获得好的融合性能,学者们往往需要大量的实验才能选择和确定性能较优的网络结构和超参数,并且一般网络模型都具有较深的结构,模型复杂度和计算量都很大。神经网络结构搜索可以实现网络空间搜索,自动生成性能最优的网络结构,可以有效地提高网络性能。因此,将神经网络结构搜索应用于空谱图像融合是一个值得研究的方向。

4)模型与数据驱动结合。目前基于卷积神经网络的融合方法都是数据驱动型,依赖大量的数据进行训练,属于黑箱算法,对其内部的融合过程不甚了解。而模型驱动型方法通过从目标、机理以及先

验知识出发设计一个融合过程透明的模型,能够很好地解释模型的可信性。所以将模型与数据驱动结合进行图像融合是一个值得研究的方向。

5) 面向高级与多任务的融合。融合图像通常要用于进一步的图像分类和目标检测等高级任务,目前的融合方法主要是基于重构性能设计,没有结合高级任务的特点。因此,结合高级任务设计融合方法具有较强的实际意义。此外,多任务学习能够联合多个相关任务进行学习训练,并且任务之间共享信息,利用多个任务的相关性来提升最终单个任务的性能。在遥感图像融合任务中,可以联合分类和目标检测等高级任务进行多任务学习,例如利用分类任务将遥感图像分成城市、农田等类别和目标检测任务检测桥梁、公路等获取的独特的光谱和空间信息共享到图像融合任务中,从而提高融合任务的性能和泛化能力。因此利用多任务学习来进行遥感图像融合也是一个值得深入研究的问题。

参考文献 (References):

- [1] 肖 亮,刘鹏飞,李 恒. 多源空一谱遥感图像融合方法进展与挑战[J]. 中国图象图形学,2020,25(5):851-863.
Xiao L,Liu P F,Li H. Progress and challenges in the fusion of multisource spatial-spectral remote sensing images[J]. Journal of Image and Graphics,2020,25(5):851-863.
- [2] Ghassemian H. A review of remote sensing image fusion methods [J]. Information Fusion,2016,32:75-89.
- [3] 张丽霞,曾广平,宣兆成. 多源图像融合方法的研究综述[J]. 计算机工程与科学,2021:1-14.
Zhang L X,Zeng G P,Xuan Z C. A survey of fusion methods for multi-source image [J]. Computer Engineering and Science, 2021:1-14.
- [4] 杨丽萍,马 孟,谢 巍,等. 干旱区 Landsat8 全色与多光谱数据融合算法评价[J]. 国土资源遥感,2019,31(4):11-19. doi: 10.6046/gtzyyg. 2019. 04. 02.
Yang L P,Ma M,Xie W, et al. Fusion algorithm evaluation of Landsat8 panchromatic and multispectral images in arid regions [J]. Remote Sensing for Land and Resources,2019,31(4):11-19. doi:10.6046/gtzyyg. 2019. 04. 02.
- [5] Weiss M,Jacob F,Duveiller G. Remote sensing for agricultural applications: A meta-review [J]. Remote Sensing of Environment, 2020,236:111402-111420.
- [6] 刘万军,高健康,曲海成,等. 多尺度特征增强的遥感图像舰船目标检测[J]. 自然资源遥感,2021,33(3):97-106. doi:10.6046/zrzyyg. 20200372.
Liu W J,Gao J K,Qu H C, et al. Ship detection based on multi-scale feature enhancement of remote sensing images [J]. Remote Sensing for Natural Resources, 2021,33(3):97-106. doi:10.6046/zrzyyg. 20200372.
- [7] 王 琳,李 迅,包云轩,等. 遥感技术在交通气象灾害监测中的应用进展[J]. 国土资源遥感,2018,30(4):1-7. doi:10.6046/gtzyyg. 2018. 04. 01.

- Wang L,Li X,Bao Y X, et al. Research progress of remote sensing application on transportation meteorological disasters [J]. Remote Sensing and for Land and Resources,2018,30(4):1-7. doi:10.6046/gtzyyg. 2018. 04. 01.
- [8] 王洪斌,肖 嵩,曲家慧,等. 基于多分支 CNN 高光谱与全色影像融合处理[J]. 光学学报,2021,41(7):55-63.
Wang H B,Xiao S,Qu J H, et al. Pansharpening based on multi-branch CNN [J]. Acta Optica Sinica,2021,41(7):55-63.
- [9] 张良培,沈焕烽. 遥感数据融合的进展与前瞻[J]. 遥感学报, 2016,20(5):1050-1061.
Zhang L P,Shen H F. Progress and future of remote sensing data fusion [J]. Journal of Remote Sensing,2016,20(5):1050-1061.
- [10] 李树涛,李聪好,康旭东. 多源遥感图像融合发展现状与未来展望[J]. 遥感学报,2021,25(1):148-166.
Li S T,Li C Y,Kang X D. Development status and future prospects of multi-source remote sensing image fusion [J]. National Remote Sensing Bulletin,2021,25(1):148-166.
- [11] 张立福,彭明媛,孙雪剑,等. 遥感数据融合研究进展与文献定量分析(1992—2018) [J]. 遥感学报,2019,23(4):603-619.
Zhang L F,Peng M Y,Sun X J, et al. Progress and bibliometric analysis of remote sensing data fusion methods (1992—2018) [J]. Journal of Remote Sensing,2019,23(4):603-619.
- [12] Meng X,Shen H,Li H, et al. Review of the pansharpening methods for remote sensing images based on the idea of meta-analysis: Practical discussion and challenges [J]. Information Fusion,2019, 46:102-113.
- [13] Javan F D,Samadzadegan F,Mehravar S, et al. A review of image fusion techniques for pan-sharpening of high-resolution satellite imagery [J]. ISPRS Journal of Photogrammetry and Remote Sensing,2021,171:101-117.
- [14] Vivone G,Mura M D,Garzelli A, et al. A new benchmark based on recent advances in multispectral pansharpening: Revisiting pansharpening with classical and emerging pansharpening methods [J]. IEEE Geoscience and Remote Sensing Magazine, 2020,9(1):53-81.
- [15] Huang W,Xiao L,Wei Z, et al. A new pan-sharpening method with deep neural networks [J]. IEEE Geoscience and Remote Sensing Letters,2015,12(5):1037-1041.
- [16] Masi G,Cozzolino D,Verdoliva L, et al. Pansharpening by convolutional neural networks [J]. Remote Sensing,2016,8(7):594-615.
- [17] Shahdoosti H R,Ghassemian H. Combining the spectral PCA and spatial PCA fusion methods by an optimal filter [J]. Information Fusion,2016,27:150-160.
- [18] Laben C A,Brower B V. Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening; U. S. ,US09/069232 [P]. 2000-01-04.
- [19] Gahremani M,Ghassemian H. Nonlinear IHS: A promising method for pan-sharpening [J]. IEEE Geoscience and Remote Sensing Letters,2016,13(11):1606-1610.
- [20] Garzelli A,Nencini F,Capobianco L. Optimal MMSE pan sharpening of very high resolution multispectral images [J]. IEEE Transactions on Geoscience and Remote Sensing,2007,46(1):228-236.
- [21] Aiazzi B,Alparone L,Baronti S, et al. MTF-tailored multiscale fu-

- sion of high-resolution MS and pan imagery[J]. Photogrammetric Engineering and Remote Sensing, 2006, 72(5): 591-596.
- [22] Vivone G, Restaino R, Chanussot J. Full scale regression-based injection coefficients for panchromatic sharpening[J]. IEEE Transactions on Image Processing, 2018, 27(7): 3418-3431.
- [23] Khan M M, Chanussot J, Condat L, et al. Indusion: Fusion of multispectral and panchromatic images using the induction scaling technique[J]. IEEE Geoscience and Remote Sensing Letters, 2008, 5(1): 98-102.
- [24] Dong L, Yang Q, Wu H, et al. High quality multi-spectral and panchromatic image fusion technologies based on curvelet transform[J]. Neurocomputing, 2015, 159: 268-274.
- [25] El-Mezouar M C, Kpalma K, Taleb N, et al. A pan-sharpening based on the non-subsampled contourlet transform: Application to WorldView-2 imagery[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2014, 7(5): 1806-1815.
- [26] 曹义亲, 杨世超, 谢舒慧. 基于 NSST 的 PCNN-SR 卫星遥感图像融合方法[J]. 航天控制, 2020, 38(2): 44-50.
Cao Y Q, Yang S C, Xie S H. Satellite remote sensing image fusion method based on NSST and PCNN-SR[J]. Aerospace Control, 2020, 38(2): 44-50.
- [27] 吴一全, 王志来. 混沌蜂群优化的 NSST 域多光谱与全色图像融合[J]. 遥感学报, 2017, 21(4): 549-557.
Wu Y Q, Wang Z L. Multispectral and panchromatic image fusion using chaotic Bee Colony optimization in NSST domain[J]. Journal of Remote Sensing, 2017, 21(4): 549-557.
- [28] Li S, Yang B. A new pan-sharpening method using a compressed sensing technique[J]. IEEE Transactions on Geoscience and Remote Sensing, 2010, 49(2): 738-746.
- [29] Yin H. PAN-guided cross-resolution projection for local adaptive sparse representation-based pansharpening[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(7): 4938-4950.
- [30] 李成轶, 田淑芳. 基于字典学习的遥感影像超分辨率融合方法[J]. 国土资源遥感, 2017, 29(1): 50-56. doi:10.6046/gtzyyg.2017.01.08.
Li C Y, Tian S F. Super-resolution fusion method for remote sensing image based on dictionary learning[J]. Remote Sensing and for Land and Resources, 2017, 29(1): 50-56. doi:10.6046/gtzyyg.2017.01.08.
- [31] Yin H. Sparse representation based pansharpening with details injection model[J]. Signal Processing, 2015, 113: 218-227.
- [32] 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): 4779-4789.
- [33] Scarpa G, Vitale S, Cozzolino D. Target-adaptive CNN-based pansharpening[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(9): 5443-5457.
- [34] Wei Y, Yuan Q, Shen H, et al. Boosting the accuracy of multispectral image pansharpening by learning a deep residual network[J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(10): 1795-1799.
- [35] Yang J, Fu X, Hu Y, et al. PanNet: A deep network architecture for pan-sharpening[C]//Proceedings of the IEEE International Conference on Computer Vision. IEEE, Venice, Italy, 2017: 5449-5457.
- [36] Zhang H, Ma J. GTP-PNet: A residual learning network based on gradient transformation prior for pansharpening[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2021, 172: 223-239.
- [37] Yang Y, Tu W, Huang S, et al. PCDRN: Progressive cascade deep residual network for pansharpening[J]. Remote Sensing, 2020, 12(4): 676.
- [38] Huang W, Feng J, Wang H, et al. A new architecture of densely connected convolutional networks for pan-sharpening[J]. ISPRS International Journal of Geo-Information, 2020, 9(4): 242.
- [39] Peng J, Liu L, Wang J, et al. PSMD-Net: A novel pan-sharpening method based on a multiscale dense network[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 59(6): 4957-4971.
- [40] Luo S, Zhou S, Qi Y. CSAFNET: Channel similarity attention fusion network for multispectral pansharpening[J]. IEEE Geoscience and Remote Sensing Letters, 2020.
- [41] 孔爱玲, 张承明, 李峰, 等. 基于知识引导的遥感影像融合方法[J]. 自然资源遥感, 2022, 34(2): 47-55. doi:10.6046/zrzyyg.2021179.
Kong A L, Zhang C M, Li F, et al. Knowledge-based remote sensing imagery fusion method[J]. Remote Sensing for Natural Resources, 2022, 34(2): 47-55. doi:10.6046/zrzyyg.2021179.
- [42] Jiang M, Shen H, Li J, et al. A differential information residual convolutional neural network for pansharpening[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 163: 257-271.
- [43] Lei D, Chen H, Zhang L, et al. NLRNet: An efficient nonlocal attention resnet for pansharpening[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5401113.
- [44] Shao Z, Cai J. Remote sensing image fusion with deep convolutional neural network[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(5): 1656-1669.
- [45] Liu X, Liu Q, Wang Y. Remote sensing image fusion based on two-stream fusion network[J]. Information Fusion, 2020, 55: 1-15.
- [46] Fu S, Meng W, Jeon G, et al. Two-path network with feedback connections for pan-sharpening in remote sensing[J]. Remote Sensing, 2020, 12(10): 1674.
- [47] He L, Xi D, Li J, et al. A spectral-aware convolutional neural network for pansharpening[J]. Applied Sciences, 2020, 10(17): 5809.
- [48] Zhang Y, Liu C, Sun M, et al. Pan-sharpening using an efficient bidirectional pyramid network[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(8): 5549-5563.
- [49] Fang S, Wang X, Zhang J, et al. Pan-sharpening based on parallel pyramid convolutional neural network[C]//IEEE International Conference on Image Processing (ICIP), 2020: 453-457.
- [50] 方帅, 方赛华, 姚宏亮. 基于深度金字塔网络的 Pan-Sharpening 算法[J]. 计算机辅助设计与图形学学报, 2019, 31(10): 1831-1837.
Fang S, Fang S H, Yao H L. Pan-sharpening based on a deep pyramid network[J]. Journal of Computer-Aided Design and Computer Graphics, 2019, 31(10): 1831-1837.
- [51] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks

- for biomedical image segmentation[C]//International Conference on Medical Image Computing and Computer - Assisted Intervention. Springer, Cham, 2015:234 - 241.
- [52] Yao W, Zeng Z, Lian C, et al. Pixel - wise regression using U - Net and its application on pansharpening[J]. *Neurocomputing*, 2018, 312:364 - 371.
- [53] Wang W, Zhou Z, Liu H, et al. MSDRN: Pansharpening of multispectral images via multi - scale deep residual network[J]. *Remote Sensing*, 2021, 13(6):1200.
- [54] Lai Z, Chen L, Jeon G, et al. Real - time and effective pan - sharpening for remote sensing using multi - scale fusion network[J]. *Journal of Real - Time Image Processing*, 2021:1 - 17.
- [55] Benzenati T, Kallel A, Kessentini Y. Two stages pan - sharpening details injection approach based on very deep residual networks[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 59(6):4984 - 4992.
- [56] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016:770 - 778.
- [57] Hu J, Du C, Fan S. Two - stage pansharpening based on multi - level detail injection network[J]. *IEEE Access*, 2020, 8:156442 - 156455.
- [58] Zhang L, Zhang J, Ma J, et al. SC - PNN: Saliency cascade convolutional neural network for pansharpening[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021:1 - 19.
- [59] Li W, Liang X, Dong M. MDECNN: A multiscale perception dense encoding convolutional neural network for multispectral pan - sharpening[J]. *Remote Sensing*, 2021, 13(3):535.
- [60] Yuan Q, Wei Y, Meng X, et al. A multiscale and multidepth convolutional neural network for remote sensing imagery pan - sharpening[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2018, 11(3):978 - 989.
- [61] Hu J, Hu P, Kang X, et al. Pan - sharpening via multiscale dynamic convolutional neural network[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 59(3):2231 - 2244.
- [62] Guo Y, Ye F, Gong H. Learning an efficient convolution neural network for pansharpening[J]. *Algorithms*, 2019, 12(1):16.
- [63] Hu J, Hu P, Wang Z P, et al. Spatial dynamic selection network for remote - sensing image fusion[J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19:8013205.
- [64] Hu J, He Z, Wu J. Deep self - learning network for adaptive pansharpening[J]. *Remote Sensing*, 2019, 11(20):2395.
- [65] Liu J, Feng Y, Zhou C, et al. PWnet: An adaptive weigh network for the fusion of panchromatic and multispectral images[J]. *Remote Sensing*, 2020, 12(17):2804.
- [66] Deng L J, Vivone G, Jin C, et al. Detail injection - based deep convolutional neural networks for pansharpening[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2020:1 - 16.
- [67] Goodfellow I, Pouget - Abadie J, Mirza M, et al. Generative adversarial networks[J]. *Communications of the ACM*, 2020, 63(11):139 - 144.
- [68] Ma J, Yu W, Chen C, et al. Pan - GAN: An unsupervised pan - sharpening method for remote sensing image fusion[J]. *Information Fusion*, 2020, 62:110 - 120.
- [69] Zhou C, Zhang J, Liu J, et al. PercepPan: Towards unsupervised pan - sharpening based on perceptual loss[J]. *Remote Sensing*, 2020, 12(14):2318.
- [70] Xiong Z, Guo Q, Liu M, et al. Pan - sharpening based on convolutional neural network by using the loss function with no - reference[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020, 14:897 - 906.
- [71] Luo S, Zhou S, Feng Y, et al. Pansharpening via unsupervised convolutional neural networks[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020, 13:4295 - 4310.
- [72] 杜晨光, 胡建文, 胡 佩. 半监督卷积神经网络遥感图像融合[J], *电子测量与仪器学报*, 2021, 35(6):63 - 70.
- Du C G, Hu J W, Hu P. Semi - supervised convolutional neural network remote sensing image fusion[J]. *Journal of Electronic Measurement and Instrumentation*, 2021, 35(6):63 - 70.
- [73] Vitale S, Scarpa G. A detail - preserving cross - scale learning strategy for CNN - based pansharpening[J]. *Remote Sensing*, 2020, 12(3):348.
- [74] 黄珊珊, 江 倩, 金 鑫, 等. 结合双胞胎结构与生成对抗网络的半监督遥感图像融合[J]. *计算机辅助设计与图形学学报*, 2021, 33(1):92 - 105.
- Huang S S, Jiang Q, Jin X, et al. Semi - supervised remote sensing image fusion method combining siamese structure with generative adversarial networks[J]. *Journal of Computer - Aided Design and Computer Graphics*, 2021, 33(1):92 - 105.
- [75] Arjovsky M, Chintala S, Bottou L. Wasserstein generative adversarial networks[C]//International Conference on Machine Learning. PMLR, 2017:214 - 223.
- [76] Mirza M, Osindero S. Conditional generative adversarial nets[J]. *Computer Science*, 2014:2672 - 2680.
- [77] Liu X, Deng C, Zhao B, et al. Feature - level loss for multispectral pan - sharpening with machine learning[C]//IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, 2018:8062 - 8065.
- [78] Xu H, Ma J, Shao Z, et al. SDPNet: A deep network for pan - sharpening with enhanced information representation[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 59(5):4120 - 4134.
- [79] Choi J S, Kim Y, Kim M. S3: A spectral - spatial structure loss for pan - sharpening networks[J]. *IEEE Geoscience and Remote Sensing Letters*, 2019, 17(5):829 - 833.
- [80] Eghbalian S, Ghasseman H. Multi spectral image fusion by deep convolutional neural network and new spectral loss function[J]. *International Journal of Remote Sensing*, 2018, 39(12):3983 - 4002.
- [81] Johnson J, Alahi A, Li F F. Perceptual losses for real - time style transfer and super - resolution[C]//European Conference on Computer Vision. Springer, Cham, 2016:694 - 711.
- [82] Bello J L G, Seo S, Kim M. Pan - sharpening with color - aware perceptual loss and guided re - colorization[C]//IEEE International Conference on Image Processing (ICIP). IEEE, 2020:908 - 912.
- [83] Vivone G, Alparone L, Chanussot J, et al. A critical comparison among pansharpening algorithms[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2014, 53(5):2565 - 2586.

[84] 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.

A review of pansharpening methods based on deep learning

HU Jianwen, WANG Zeping, HU Pei

(School of Electrical and Information Engineering, Changsha University of Science and Technology, Changsha 410114, China)

Abstract: With the fast development and wide application of remote sensing technology, remote sensing images with higher quality are needed. However, it is difficult to directly acquire high – resolution, multispectral remote sensing images. To obtain high – quality images by integrating the information from different imaging sensors, pansharpening technology emerged. Pansharpening is an effective method used to obtain multispectral images with high spatial resolution. Many scholars have studied this method and obtained fruitful achievements. In recent years, deep learning theory has developed rapidly and has been widely applied in pansharpening. This study aims to systematically introduce the progress in pansharpening and promote its development. To this end, this study first introduced the traditional, classical pansharpening methods, followed by commonly used remote sensing satellites. Then, this study elaborated on the pansharpening methods based on deep learning from the perspective of supervised learning, unsupervised learning, and semi – supervised learning. After that, it described and analyzed loss functions. To demonstrate the superiority of the pansharpening methods based on deep learning and analyze the effects of loss functions, this study conducted remote sensing image fusion experiments. Finally, this study presented the future prospects of the pansharpening methods based on deep learning.

Keywords: remote sensing image; pansharpening; deep learning; convolutional neural network

(责任编辑: 陈 理)