

生成对抗网络在遥感图像融合中的研究进展

郑黄齐眉, 潘成毅, 金鑫, 王倩倩, 苗圣法, 江倩

1. 云南大学 软件学院, 昆明 650000;

2. 云南大学 跨境网络空间安全教育部工程研究中心, 昆明 650000

摘要: 遥感图像融合作为数据融合的一个重要分支, 对于地物研究具有重要意义, 有效选择合适的融合方法对提高图像精度尤为重要。随着遥感技术的发展, 传统的图像融合方法已经难以满足图像精度的要求, 不断有新的融合方法被提出。在遥感图像处理领域, 生成对抗网络 (GAN) 以其强大的生成能力和对复杂分布的建模能力, 迅速成为提升图像融合质量的重要技术。本文回顾了传统的遥感图像融合方法及其局限性, 并且分析了深度学习, 尤其是GAN在本领域中的优势。通过对多种GAN架构和损失函数的详细介绍, 揭示了其在提升融合图像空间和光谱分辨率方面的巨大潜力。此外, 还详细阐述了近年来各种基于GAN的遥感图像融合方法, 探讨了其在全色锐化以及高光谱全色锐化任务中的应用。本文总结了基于GAN的遥感图像融合方法的发展过程, 并从3个方面分析了目前技术存在的问题和未来发展的方向。

关键词: 遥感图像, 多源数据融合, 深度学习, 生成对抗网络, 全色遥感图像锐化, 多光谱图像, 高光谱图像

中图分类号: TP751.2/P2

引用格式: 郑黄齐眉, 潘成毅, 金鑫, 王倩倩, 苗圣法, 江倩. 2025. 生成对抗网络在遥感图像融合中的研究进展. 遥感学报, 29(10): 2859-2872

Zheng H Q M, Pan C Y, Jin X, Wang Q Q, Miao S F and Jiang Q. 2025. Research progress of generative adversarial networks in remote sensing image fusion. National Remote Sensing Bulletin, 29(10): 2859-2872 [DOI: 10.11834/jrs.20254439]

1 引言

遥感图像是记录了丛林、地面设施、矩形农田、房屋建筑、山地等各种地物的电磁波反射照片, 凭借信息量大、宏观、和动态监测等特点可有效帮助人们观察了解地球信息。而多传感器技术的发展十分迅速, 使得信息多样化 (Sun 和 Deng, 2004; Khaleghi 等, 2013)。但由于传感器的物理限制, 光学遥感卫星通常多在同一覆盖区域捕获全色图像和多光谱图像。全色图像空间分辨率较高, 包含丰富的空间纹理信息, 但光谱信息有限。而多光谱图像 (MS) 具有更多的波段, 包含更多的光谱信息 (Achalakul 和 Taylor, 2003)。将具有高质量空间分辨率的全色图像与具有高光谱分辨率的多光谱图像 (LRMS) 结合是整合空间

和光谱信息的重要技术, 将两种信息融合生成高空间分辨率的多光谱图像 (HRMS) 的过程就称为遥感图像融合中的 Pan-sharpening (全色锐化)。在1986年, 地球观测系统 (SPOT-1) 提供了两张 LRMS 图像和一张 PAN 图像 (Cliche 等, 1985), 此后30多年时间, 遥感图像融合技术得到了快速发展。

近年来, 遥感图像融合的技术不断更新演化, 大量的融合方法被提出, 并且这些方法可以分为基于组件替换的方法、基于多分辨率分析的方法, 以及近年来提出的新的基于深度学习的融合方法。深度学习方法, 通过多层网络结构, 可以在不同的感受域内提取丰富的图像特征, 并通过端到端的训练优化特征融合策略, 从而有效减少传统方法带来的图像失真问题。比如, 卷积神经网络

收稿日期: 2024-10-10; 预印本: 2025-06-27

基金项目: 国家自然科学基金(编号: 62101481, 62261060); 云南省基础研究计划(编号: 202301AW070007, 202201AU070033, 202201AT070112, 202301AU070210)

第一作者简介: 郑黄齐眉, 研究方向为遥感图像融合。E-mail: zhenghuangqimei@stu.ynu.edu.cn

通信作者简介: 金鑫, 研究方向为人工智能、图像处理。E-mail: xinjin@ynu.edu.cn

(CNN) 和自编码器网络 (AE), 在特征提取和多尺度信息处理方面展现了强大的能力, 推动了遥感图像融合技术的发展。尽管 CNN 在一定程度上改善了图像融合的质量, 但如何在减小对参考图像的依赖的同时, 进一步提升融合图像的光谱保真度和空间分辨率, 仍然是一个亟待解决的问题。

生成对抗网络 (GAN) 的引入为解决上述问题提供了新的解决思路。GAN 通过生成器和判别器之间的对抗博弈, 能够在没有明确参考图像的情况下, 生成具有高度真实性的融合图像。与传统的深度学习方法相比, GAN 不仅能够更好地模拟复杂的图像分布, 还能够处理未对齐的多源数据时展现出强大的鲁棒性。

GAN 模型在 2014 年由 Goodfellow 等 (2020) 开创性地提出, 此后在图像生成、图像超分辨率、图像融合等领域大放异彩。在遥感图像融合中, GAN 的表现也十分出色。不同的 GAN 架构如 PSGAN (Liu 等, 2021)、FusGAN (Luo 等, 2021)、Pan-GAN (Ma 等, 2020a) 等, 在处理全色锐化和高光谱全色锐化等任务时, 均展示了其在空间和光谱信息平衡方面的卓越表现。这些方法通过多尺度特征提取、光谱和空间信息的分支处理以及对抗性损失的优化, 使得融合图像的质量显著提升。同时, GAN 和无监督学习的高度适配使其能够在没有高质量参考图像的情况下进行训练, 这对于遥感数据中普遍存在的参考图像缺乏问题提供了有效的解决方案。

本文旨在系统地探讨基于 GAN 的遥感图像融合方法, 从传统遥感图像融合技术出发, 介绍了各种方法和技术, 随后总结了 GAN 在多模态图像融合中的多种常见架构, 进一步讨论了其在全色锐化和高光谱全色锐化中的具体应用, 分析了目前技术存在的问题并展开了一些对未来发展的拓展思考。

2 遥感图像融合技术

本章将探讨遥感图像融合的主要方法, 分为传统方法和基于深度学习的方法, 并从监督学习和无监督学习两个方向来简要介绍深度学习方法。最后, 我们会介绍遥感图像融合相关的一些参数。

2.1 传统方法

基于分量替换 (CS) 方法通过将多光谱图像

转换到变换域并用全色图像替换其中的特定分量, 从而生成高空间分辨率的图像。CS 方法包括基于 IHS (Intensity-Hue-Saturation) 的图像融合方法、基于 PCA (Principal Component Analysis) 变换的图像融合方法和基于 GS (Gram-Schmidt) 的方法。

IHS 图像融合方法 (Tu 等, 2007; Carper 等, 1990) 通过融合 I 分量和全色图像提高空间分辨率, 但易导致光谱失真, 且仅适用于三波段图像。PCA 方法通过线性变换提取特征并去除噪声, 也可能导致光谱失真 (Bai 等, 2015), 特别是在第一分量与全色图像相关性低的时候。GS 方法通过最佳匹配波段, 改善了光谱响应不一致的问题。

基于多分辨率分析 (MRA) 方法通过多分辨率分解将全色图像的细节注入多光谱图像, 避免了 CS 方法的光谱失真问题 (Hu 等, 2021)。其中, 小波变换最早被引入图像融合, 尽管存在伪影问题, 但仍广泛应用, UDWT 方法 (Garzelli 和 Nencini, 2007) 尤其受关注。多尺度几何变换方法 (Mehra 和 Nishchal, 2014), 克服了小波在二维方向信息捕捉中的局限性, 提升了对复杂图像结构的处理能力 (Balakrishnan 等, 2012)。

以上方法中, CS 方法的光谱失真问题在多波段融合时表现得尤其明显, 而 MRA 方法在图像分解和重构过程中容易产生伪影。可以发现, 传统方法缺乏对复杂光谱与空间信息的自适应建模能力, 难以同时兼顾高空间分辨率和高光谱保真度。

2.2 基于深度学习的方法

随着传统方法局限性的显现 (Zhang 等, 2021), 研究人员开始探索深度学习技术在遥感图像融合中的应用。深度学习方法通过端到端的特征提取和优化策略, 显著改善了融合图像的空间和光谱表现 (Dong 等, 2016)。深度学习的优势在于可以使用不同的网络分支来实现不同感受域的特征抽取, 从而获得更加丰富的特征; 在损失函数的指导下, 优化特征融合策略, 从而实现自适应特征融合。深度学习在遥感图像融合领域取得了远超传统方法的性能表现。

2.2.1 基于监督学习的方法

监督学习框架通过构建输入图像与参考图像之间的映射关系进行模型训练。有监督的模型一般都会遵循 Wald 协议 (Wald 等, 1997), 就是将

原始 *MS* 图像和 *PAN* 图像进行下采样，而原始 *MS* 图像作为参考图，最终将下采样后的低分辨率 *MS* 图像、*PAN* 图像以及原始 *MS* 作为输入供网络模型学习。其训练过程依赖成对数据集 (*PAN*, *LRMS*), *HRMS*，其中 *HRMS* 作为监督信号指导参数优化。现有方法主要沿着两个技术路线演进。

第一，基于 CNN 的监督框架早期采用级联卷积结构实现端到端映射。受到超分辨率方法 SRCNN (Dong 等, 2016) 的启发，PNN 模型 (Masi 等, 2016) 建立了 3 阶段处理流程：浅层特征提取、非线性映射、图像重建。随后发展的 PanNet (Yang 等, 2017) 在特征提取阶段引入残差连接，将浅层空间细节直接传递至重建层。DRPNN (Wei 等, 2017) 则通过堆叠残差块形成深层网络，其递进式特征精炼机制显著提升了光谱保真度。此外，双分支结构和注意力机制也被引入 CNN 架构，代表性模型包括 MSDCNN (Yuan 等, 2018)、TFNet (Liu 等, 2020)、Attention-FPNet (Zhong 等, 2021) 等。

第二，基于 GAN 的监督框架通过对抗训练提升生成质量。图 1 展示了用 GAN 实现遥感图像融合的基本模型结构。其中的创新主要体现在如何构建更好的生成器和判别器。比如，PSGAN (Liu 等, 2021) 采用双流 CNN 作为生成器，并行处理 *PAN* 和 *MS* 模态数据；采用卷积—全连接混合结构作为判别器，自适应学习多尺度判别特征。而 FusGAN (Luo 等, 2021) 提出残差密集块构建生成器，设计了“退化—复原”训练流程，通过局部特征复用增强细节重建能力。

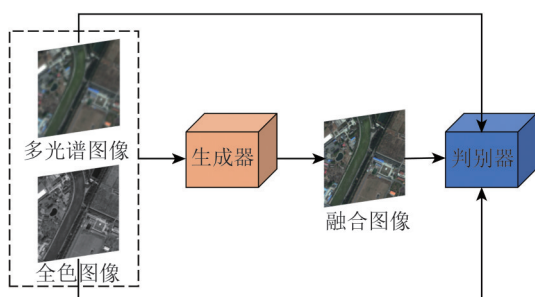


图1 基于GAN的遥感图像融合模型的基本结构
(以全色锐化为例)

Fig.1 Basic structure of GAN-based remote sensing image fusion model (Using pan-sharpening as an example)

2.2.2 基于无监督学习的方法

由于在实际应用中，高分辨率的参考图像往

往难以获得，而监督学习方法却通常依赖这些参考图像进行模型训练，因此无监督学习应运而生。如前文所述，传统的遥感融合方法解决这一难题通常是使用退化图像来训练模型，再将训练好的模型应用于原始图像的融合。然而由于退化图像与原始图像在分辨率和尺度上的不一致，融合结果往往难以真实反映原始图像的特征。无监督的遥感图像融合技术主要包括基于 GAN 和基于无参考损失函数的技术。进入深度学习时代以后，研究人员发现，GAN 的独特机制使它在无监督和半监督学习上的应用更加广泛。与监督学习相反，基于 GAN 的无监督学习框架突破了参考图像的依赖，其核心创新在于构建自洽的优化目标。

这类方法通常采用双鉴别器双重约束机制，即空间鉴别器评估生成图像与 *PAN* 图像的空间一致性，光谱鉴别器保障生成图像与 *MS* 图像的光谱保真度。Pan-GAN (Ma 等, 2020a) 是采用这种设计的经典方法。在此基础上，后来的方法还通过设计物理约束损失更好地实现无监督。有些方法引入循环一致性损失来强制生成图像下采样后与原始 *MS* 图像对齐，例如 PGMAN (Zhou 等, 2021)；再引入结构相似性损失，通过 *MS* 图像的超像素分割结果，约束生成图像的光谱分布，例如 UCGAN (Zhou 等, 2022)。

一些融合方法也采用自监督和半监督的监督范式。自监督学习也属于无监督学习的一种分支，模型会在在没有外部标签的情况下，自己生成“伪标签”或预测丢失的部分数据来指导学习。半监督学习则结合了少量标注数据和大量未标注数据来进行训练，以减少对标注数据的依赖。PanColorGAN (Ozcelik 等, 2021) 就是一种基于自监督的方法，该模型采用任意比例的下采样作为噪声干扰，以提高模型的泛化性能，同时通过基于着色的生成器，将多光谱图像转换为灰度图像，并在训练过程中注入颜色信息，提高了图像的色彩还原度。除此之外，Jin 等 (2021) 提出了基于半监督学习的 GAN 模型，通过将输入图像转换为 HSV 色彩空间，采用亮度分量进行训练，并且将 WGANs 的损失函数与 PSNR 结合构成一种新的损失函数，避免了对参考图像的依赖。

2.3 图像评价指标

在遥感图像融合的研究中，图像评价指标是

评估融合效果的重要手段。这些指标通常可以分为全参考指标和无参考指标两大类,用于衡量融合图像的质量。

全参考指标要求有一个高质量的参考图像作为标准。常用的全参考指标包括均方根误差(RMSE)、峰值信噪比(PSNR)、结构相似性(SSIM)、相关系数(CC)、光谱角度映射(SAM)等。RMSE用于衡量融合图像与参考图像之间的像素级距离,反映图像在空间细节上的差异。PSNR则评估图像中有效信息与噪声的比值,值越高表示图像质量越好。SSIM通过亮度、对比度和结构

的相似性来评价图像的质量,CC用于评估融合图像与参考图像之间的线性相关性,而SAM通过计算光谱矢量之间的角度来衡量光谱失真程度。此外,ERGAS和Q指数则提供了综合的质量评价。

无参考指标在没有参考图像的情况下使用,常用于实际遥感应用中。常见的无参考指标包括光谱失真测量(D_λ)、空间失真测量(D_s)以及基于它们的综合评价指标QNR。无参考指标能够在没有理想参考图像的情况下,提供对融合效果的有效评估,特别适用于真实的遥感场景。详细公式在表1中展示。

表1 图像评价指标

Table 1 Image evaluation metrics

指标名称	公式	备注
RMSE	$\text{RMSE} = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (F(i,j) - R(i,j))^2}$	F 和 R 分别表示融合图像和参考图像的像素值, M 和 N 分别表示图像的行数和列数
PSNR	$\text{PSNR} = 10 \log_{10} \left(\frac{G^2}{\text{MSE}} \right)$	G 表示图像的灰度级别,MSE是均方误差
SSIM	$\text{SSIM} = \frac{(2\mu_F \mu_R + C_1)(2\sigma_{FR} + C_2)}{(\mu_F^2 + \mu_R^2 + C_1)(\sigma_F^2 + \sigma_R^2 + C_2)}$	μ_F 和 μ_R 分别表示融合图像和参考图像的平均值, σ_F^2 和 σ_R^2 分别表示它们的方差, σ_{FR} 表示协方差, C_1 和 C_2 表示稳定常数
CC	$\text{CC} = \frac{\sum_{i=1}^M \sum_{j=1}^N (F(i,j) - \mu_F)(R(i,j) - \mu_R)}{\sqrt{\sum_{i=1}^M \sum_{j=1}^N (F(i,j) - \mu_F)^2} \sqrt{\sum_{i=1}^M \sum_{j=1}^N (R(i,j) - \mu_R)^2}}$	式中变量的意义同上
SAM	$\text{SAM} = \arccos \left(\frac{\sum_{i=1}^L F_i R_i}{\sqrt{\sum_{i=1}^L F_i^2} \sqrt{\sum_{i=1}^L R_i^2}} \right)$	L 表示光谱波段数, F_i 和 R_i 分别表示融合图像和参考图像在第 i 个波段的像素值
ERGAS	$\text{ERGAS} = 100 \times \frac{1}{h} \sqrt{\frac{1}{L} \sum_{l=1}^L \left(\frac{\text{RMSE}_l}{\mu_{R,l}} \right)^2}$	h 表示全色图像和多光谱图像的空间分辨率, RMSE_l 是第 l 个波段的均方根误差, $\mu_{R,l}$ 表示参考图像第 l 波段的平均值
D_λ	$D_\lambda = \left(\frac{1}{L(L-1)} \sum_{i=1}^L \sum_{j=1}^L Q(M_i, M_j) - Q(F_i, F_j) ^p \right)^{\frac{1}{p}}$	$Q(\cdot, \cdot)$ 表示UIQI计算函数, p 为表征光谱差异的正整数, M_i 和 M_j 分别表示多光谱图像在第 i 个和第 j 个波段的像素值
D_s	$D_s = \left(\frac{1}{L} \sum_{i=1}^L Q(F_i, P_{NL}) - Q(M_i, P_N) ^q \right)^{\frac{1}{q}}$	P_N 表示全色图像, P_{NL} 表示降采样后的全色图像, q 表示范数
QNR	$\text{QNR} = (1 - D_\lambda)^\alpha \cdot (1 - D_s)^\beta$	α 和 β 表示用于调整光谱和空间失真对整体质量影响的权重参数

3 GAN在遥感图像融合中的技术演进

在本节中,我们将深入探讨GAN在遥感图像融合中的技术演进,首先从GAN在多源图像融合的架构演变入手,分析GAN在遥感图像融合中网络结构上的改进。在3.3节中,我们分别对GAN在两种遥感图像融合领域的应用进行了介绍。最后介绍在基于GAN的方法中常见的损失函数设计形式。

3.1 网络结构上的改进

通过对GAN架构的优化和发展,研究人员逐渐构建出适用于不同应用场景的多源图像融合模型。对于解决遥感图像融合任务,大多数研究是在已有的多源图像融合模型的基础上做出改进。传统GAN模型是在没有外部条件约束的情况下,通过单个生成器和判别器之间对抗博弈生成图像。随着GAN在多源图像融合领域的深入应用,研究

人员发现改变生成器和判别器的数量对融合图像的质量有着很大影响。如果一些外部条件约束,也可以进一步控制和提升生成图像的特定属性。下面我们介绍了5种常用于图像融合的GAN架构,分别是基础的GAN、单生成器—双判别器架构(SGDD GAN)、多生成器—单判别器架构(MGSD GAN)、多生成器—多判别器架构(MGMD GAN),

以及条件约束GAN架构(CGAN)。其中着重介绍了常用于遥感图像融合任务的架构。图2总结了一些基于GAN的图像融合方法,包括经典的多模态图像融合方法和近期的遥感图像融合方法,图2中可以观察到多模态图像融合方法的发展给遥感图像融合方法带来的影响,力求梳理清楚GAN在本领域的发展趋势。

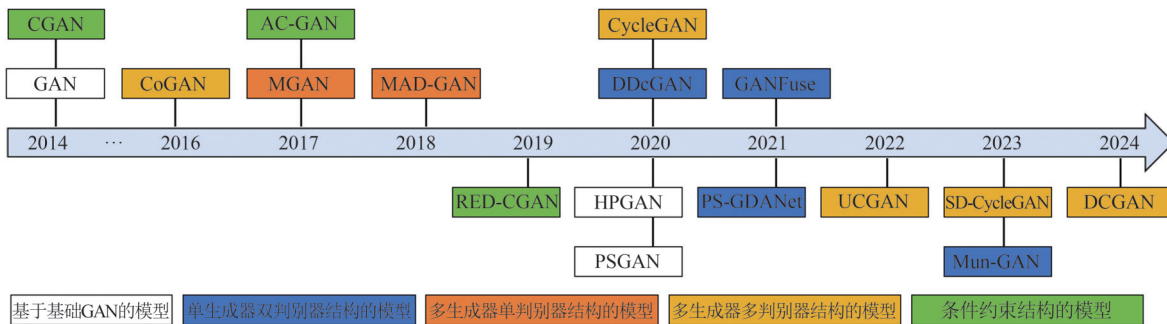


图2 基于GAN的图像融合模型发展历程(时间轴上方是经典的多模态图像融合方法,时间轴下方是经典的和近期的遥感图像融合方法。可以发现,多模态图像融合方法是遥感图像融合方法发展的风向标)

Fig. 2 Development timeline of GAN-based image fusion models (With classical multimodal image fusion methods above the timeline and classical and recent remote sensing image fusion methods below the timeline. It can be found that the multimodal image fusion method is the wind vane for the development of remote sensing image fusion methods)

3.1.1 基础的GAN模型

GAN最初是一个用于生成任务的网络模型。通过将生成任务设计为生成器和判别器之间的对抗博弈,实现数据的自动生成。GAN的设计灵感来自博弈论中的零和博弈思想,生成器和判别器在训练过程中不断竞争,从而优化生成的样本质量。GAN的基本模型如图3所示,其中实线代表前向传播,即生成伪样本和真实数据一起输入判别器的信息流;虚线代表后向传播,即生成器根据判别器的反馈调整参数的信息流,用于优化生成器以提高生成图像的质量。

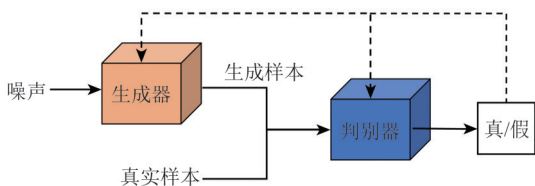


图3 GAN的基本模型

Fig. 3 Basic model of GAN

在标准的GAN模型中,生成器和判别器通常采用多层感知机结构。GAN的训练过程可以通过以下的最小最大化损失函数来描述:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (1)$$

式中, x 表示从真实数据分布 $p_{\text{data}}(x)$ 中采样的真实数据, z 是从噪声分布 $p_z(z)$ 中采样的随机噪声向量。

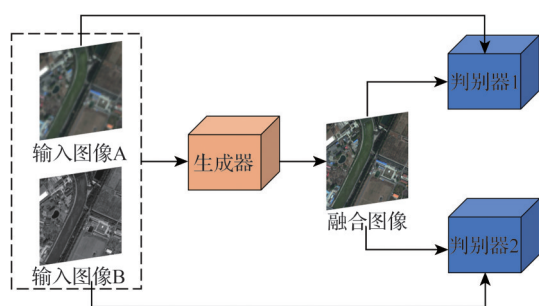
生成器的目标是生成与真实样本 $p_{\text{data}}(x)$ 分布相似的样本 $G(z)$ 。它的输入就是服从简单分布(如高斯分布)的随机噪声 z 。判别器则作为二分类器,输入可以是真实样本 x 或生成样本 $G(z)$,输出表示输入样本是真实还是生成的。判别器旨在尽可能准确地区分真实样本和生成样本,即对真实样本输出接近1,对生成样本输出接近0,而生成器则努力生成逼真的样本,以迷惑判别器,使其难以正确区分。基础GAN的独特机制导致它难以达到收敛状态和纳什均衡(Ratliff等, 2013),甚至可能出现振荡或模式崩溃。并且在基础GAN模型中,生成器的输入仅仅为随机噪声,无法通过约束条件来生成具有指定特征的样本。

尽管如此,相比其他生成模型,GAN还是具有一些优势,促使研究者进行改进(Wang等, 2021; Gui等, 2023)。例如,在生成的图像清晰度方面与变分自编码器相比,不需要显式地建模

数据的概率分布,更利于生成清晰的图像。在计算速度方面也有优势,因为GAN不需要复杂的近似推理,只需使用反向传播算法计算梯度。与独立成分分析(ICA)相比,GAN的生成器对输入数据的大小没有限制,并且可以训练各种生成网络。

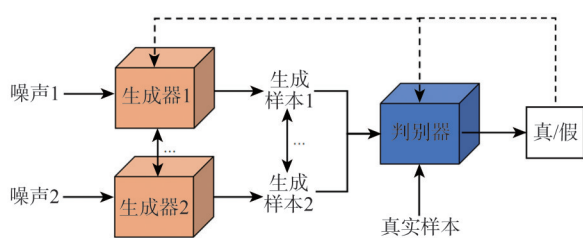
3.1.2 单生成器双判别器生成对抗网络

单生成器双判别器生成对抗网络(SGDD GAN)是通过一个生成器和两个判别器来构建的网络架构。该架构的设计目标是增强生成图像的细节和质量,尤其是在处理复杂的多源图像融合任务时表现突出。两个判别器分别针对不同的数据源进行判别,使生成器能够更好地学习和融合多源信息。SGDD GAN的代表模型包括DDcGAN(Ma等,2020b)(如图4(a))和GANFuse(Yang等,2021),这些模型通过引入双判别器机制,有效提升了图像的融合效果和视觉质量。



(a) DDcGAN的基本结构

(a) Basic structure of DDcGAN



(b) MAD-GAN的基本结构

(b) Basic structure of MAD-GAN

图4 SGDD和MGSD代表结构

Fig. 4 Representative structures of SGDD and MGSD

双判别器设计在无监督的遥感图像融合方法中运用广泛。继经典的Pan-GAN(Ma等,2020a)之后,还有很多双判别器的新方法表现出色。比如PS-GDANet(Dong等,2022)也采用双判别器结构分别针对光谱和空间信息进行优化,在较少迭代次数下生成高分辨率图像,表现出色。

在近期的SGDD结构中,最重要的创新是生成器的设计,有些方法会结合多尺度特征提取和自适应融合进行设计。Mun-GAN(Liu等,2023)是一种无监督网络,通过多尺度特征提取器、自适应加权融合模块和嵌套特征聚合模块的结合来组成生成器,实现了特征的自适应融合和多尺度信息的高效重构。同样是SGDD GAN结构,UPGAN(Jin等,2024)则以U形结构为骨干设计了一个生成器,提出了多尺度细节特征学习和跨尺度注意力融合模块,并通过设计混合损失函数来提升了生成图像的空间和光谱分辨率。

3.1.3 多生成器单判别器生成对抗网络

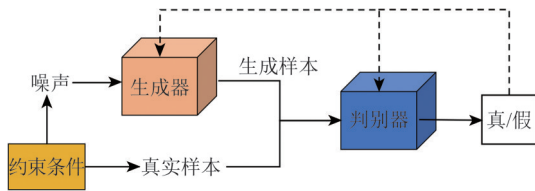
多生成器单判别器生成对抗网络(MGSD GAN)采用多个生成器和一个判别器的架构,旨在解决生成器单一输出导致的模式崩溃问题。通过多个生成器生成不同的图像样本,并由单一判别器进行评估,这种架构可以生成更具多样性和细节丰富的图像。典型的MGSD GAN架构如MAD-GAN(Ghosh等,2018)(图4(b)),通过多生成器的并行工作,显著提高了生成图像的多样性和适应性。

3.1.4 多生成器多判别器生成对抗网络

多生成器多判别器生成对抗网络(MGMD GAN)通过多个生成器和多个判别器的协同工作,进一步提升了图像的生成质量。MGMD GAN架构根据生成器和判别器之间的逻辑结构,又可细分为级联结构、并行结构和循环结构3种类型。级联结构(如StackGAN(Zhang等,2017))分阶段优化特征,从低分辨率到高分辨率逐步优化图像质量。并行结构(如CoGAN(Liu和Tuzel,2016))通过并行处理多源图像特征,实现无配对训练数据的图像融合。循环结构(如CycleGAN(Zhu等,2017),图5(b))通过双向映射,保证不同域之间的图像转换的一致性。这些子类型模型分别在处理复杂多源数据、无监督学习以及跨域图像翻译等任务中展示了强大的性能。

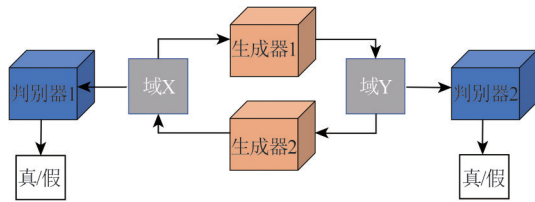
其中,以CycleGAN(Zhu等,2017)为代表的循环结构,在无监督的情况下,可以通过循环一致性保证图像转换的精确性。这一特性特别适用于不需要一一对应样本的遥感图像融合任务,因此在近年来被广泛运用。CycleGAN通过引入循

环一致性损失（Cycle Consistency Loss），确保生成的图像在双向映射后能够恢复到原始输入，从而保持图像的结构和内容一致性。CycleGAN的架构包括两个生成器和两个判别器，生成器A用于生成域A中的图像，生成器B用于生成域B中的图像。



(a) CGAN的基本结构

(a) Basic structure of CGAN



(b) CycleGAN的基本结构

(b) Basic structure of CycleGAN

图5 MGMD和CGAN代表结构

Fig. 5 Representative structures of MGMD and CGAN

CycleGAN的损失函数由对抗性损失和循环一致性损失组成，如下述公式所示：

$$L_{\text{CycleGAN}} = L_{\text{GAN}}(G_A, D_B) + L_{\text{GAN}}(G_B, D_A) + \lambda L_{\text{cyc}}(G_A, G_B) \quad (2)$$

式中， L_{cyc} 为循环结构特有的循环一致性损失，用于确保图像在双向转换后能够还原。CycleGAN通过学习域A与域B之间的映射关系，使得生成的图像在视觉上与目标域的真实图像高度一致。

近年来，CycleGAN架构颇受遥感图像融合领域的研究人员偏爱。基于CycleGAN架构，有很多创新的方法被提出。UCGAN（Zhou 等，2022）的做法是用双流生成器提取PAN和MS图像的模式特征，在特征域进行融合，然后重建全色锐化图像，为此引入基于循环一致性的混合丢失机制来提高性能。STP-SOM模型（Zhang 和 Ma，2023）利用循环对抗学习机制在光谱降解和插值之间进行对抗，学习了一个光谱观测模型，并通过该模型建立了一个基于空间保真度的尺度转移全色锐化范式，显著改善了融合图像的空间和光谱信息平衡问题。DCGAN（He 等，2024）则是基于双循环一

致性的无监督模型，通过跨域相关性信息交互网络实现了空间和光谱特征的深度融合，并引入双循环一致性损失以减少模拟退化过程中信息的损失，增强了模型的泛化能力。SD-CycleGAN（Diao 等，2023）也是基于单生成器循环结构GAN的一种方法，通过光谱和空间退化模块重构高分辨率多光谱图像，并通过循环一致性损失和边缘增强损失保证空间和光谱信息的一致性。

3.1.5 条件约束生成对抗网络

条件约束生成对抗网络（CGAN）通过引入外部条件（如辅助图像或标签等）来控制生成图像的属性，使得生成的图像更符合特定需求。条件约束不仅增加了模型的灵活性，还提高了生成图像的质量和精确性。典型的条件约束生成对抗网络模型包括CGAN（Mirza 和 Osindero，2014）（图5（a））和AC-GAN（Odena 等，2017）。这些模型通过条件输入，在生成过程中施加额外约束来进行引导，使得生成的图像具备更高的可控性和细节表现。CGAN通过条件约束生成高质量特定属性图像，也适用于遥感图像融合任务。对于遥感图像融合任务，条件信息通常是来自不同传感器或卫星的多源图像数据。CGAN的损失函数可以表示为

$$L_{\text{CGAN}}(G, D) = \left(\mathbb{E}_{x \sim p_{\text{data}}(x)} \log D(x|c) \right) + \mathbb{E}_{z \sim p_z(z)} \left(\log(1 - D(G(z|c))) \right) \quad (3)$$

式中， c 代表条件信息，如辅助图像。

在CGAN架构中，条件信息被输入到生成器和判别器中，使得生成的图像更接近参考图像的特征空间。这种方法在遥感图像融合中非常有效，尤其在融合多光谱或多传感器数据时，能够生成更具一致性的融合图像。CGAN架构代表性的遥感图像融合方法是RED-cGAN（Shao 等，2020）。该模型采用了残差编码—解码生成器和条件判别器网络，以确保生成的图像在保持空间信息的同时，减少光谱失真。

3.2 GAN在全色锐化和高光谱全色锐化中的应用

在本节中，我们分别对GAN在两种遥感图像融合领域的应用进行了介绍。全色锐化处理全色图像和多光谱图像，而高光谱全色锐化涉及全色图像、多光谱图像和高光谱图像。虽然二者很相似，但由于高光谱图像包含更多的光谱波段，融

合过程中需要处理更复杂的光谱信息，导致二者的最新方法整理如表2，由于篇幅原因省略了部分数的技术路线存在差异。本文中涉及的经典方法和数据集。

表2 基于GAN的遥感图像融合方法
Table 2 GAN-based remote sensing image fusion methods

方法	发表年份	应用场景	数据集	监督范式	发表会议或期刊
RED-cGAN	2019	全色锐化	WorldView-2, WorldView-3	有监督	LGRS
PSGAN	2020	全色锐化	QuicBird, GaoFen-2, WorldView-2	有监督	TGRS
FusGAN	2020	高光谱全色锐化	Landsat 8, Sentinel-2	有监督	TGRS
Pan-GAN	2020	全色锐化	WorldView-2, GaoFen-2	无监督	INFFUS
PanColorGAN	2020	全色锐化	WorldView-2, WorldView-3	自监督	TGRS
HPGAN	2020	高光谱全色锐化	CAVE, Chikusei, GaoFen-2	有监督	TGRS
ESGAN	2021	全色锐化	Maryland	半监督	JSTARS
PS-GDANet	2021	高光谱全色锐化	CAVE, Pavia Center	有监督	TNNLS
physics-based GAN	2021	高光谱全色锐化	CAVE, Pavia University	有监督	JSTARS
DI-GAN	2022	全色锐化	QuickBird, IKONOS, WorldView-3	有监督	ESWA
UCGAN	2022	全色锐化	GaoFen-2, WorldView-3	无监督	TGRS
MSGAN	2022	全色锐化	QuickBird, GaoFen-2	无监督	MMM
STP-SOM	2023	全色锐化	QuickBird, GaoFen-2	有监督	IJCV
SD-CycleGAN	2023	全色锐化	QuickBird, GeoEye-1, GaoFen-2	无监督	JSTARS
Mun-GAN	2023	全色锐化	QuickBird, GaoFen-2, WorldView-3	无监督	TGRS
CMAN	2023	高光谱全色锐化	Pavia University, Houston University	无监督	Remote Sensing
DCGAN	2024	全色锐化	QuickBird, GaoFen-2, WorldView-3	无监督	TGRS
UPGAN	2024	全色锐化	QuickBird, WorldView-2	无监督	IJGI
MFT-GAN	2024	高光谱全色锐化	CAVE, Chikusei, Bostwana	无监督	TGRS

3.2.1 GAN在全色锐化中的应用

如前文所介绍，全色锐化技术的目标是将高空间分辨率的全色图像与低空间分辨率的多光谱图像相结合，以生成同时具备高空间和光谱分辨率的遥感图像（图6）。这一技术在遥感领域具有重要意义，因为它能够弥补单一传

感器在空间和光谱分辨率方面的不足，从而提高遥感影像的解析度和应用价值。由于全色图像通常仅包含单一波段，而多光谱图像包含多个波段，因此全色锐化的主要挑战在于如何在增强空间分辨率的同时最大程度保持光谱信息。

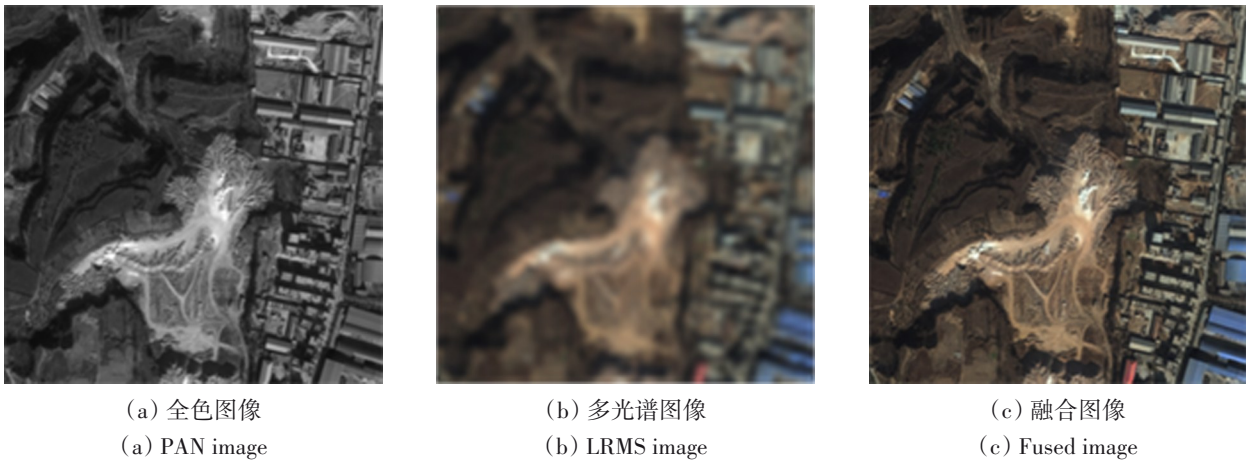


图6 全色锐化
Fig. 6 Pan-sharpening

传统的CNN方法虽然在全色锐化任务中取得了一定成效,但仍然面临光谱失真和细节丢失的问题,难以满足高精度遥感应用的需求。为了进一步提升全色锐化的性能,研究者引入GAN来解决这一问题。GAN能够更有效地模拟复杂的空间和光谱关系,减少信息丢失,提高融合图像的整体质量。GAN的优势在于通过构建生成器和判别器之间的对抗博弈,使得生成的融合图像在统计分布上与真实图像更为接近。特别是无监督GAN方法,比如Pan-GAN (Ma等, 2020a)和UCGAN (Zhou等, 2022),此处不再赘述。

这几年全色锐化中的GAN模型重点关注空间特征的增强,因此生成器通常采用较深的卷积结构,并引入多尺度特征提取模块,在不同尺度和层次上对图像进行细化处理,以进一步提升融合图像的细节表现力和光谱一致性。例如,DI-GAN (Benzenati等, 2022)引入了相对生成对抗网络来注入细节信息,从而更好地保留了空间结构。而MSGAN (Wang等, 2022)通过多尺度处理,增强了对不同分辨率特征的捕捉能力,使得生成的图像在各个尺度上都具备较高的清晰度和光谱保真度。

基于GAN的全色锐化方法在实际应用中显示出了巨大的潜力,不仅能够生成视觉上更为逼真的融合图像,还显著提高了图像在去云、目标检测、分类和变化检测等下游遥感任务中的表现。不过,如何进一步优化GAN的训练过程,提升生成器与判别器之间的平衡性,以及如何在更大规模的遥感数据集上实现高效、稳定地训练,仍然是未来研究的重要方向。

3.2.2 GAN在高光谱全色锐化中的应用

多光谱—高光谱图像融合,也称为高光谱全色锐化,旨在结合多光谱图像的高空间分辨率和高光谱图像的丰富光谱信息(即MS和HS图像),从而生成同时具备高空间和光谱分辨率的融合图像。这一过程类似于全色锐化,但在处理多源数据时,由于存在更大尺度空间分辨率、更多波段的差异,高光谱全色锐化面临着更加复杂的挑战,也需要更加复杂的特征提取和光谱保持策略。此外,高光谱数据更加难以获得的性质,也使得高光谱全色锐化对无监督技术的需求更高。

CNN也是这一领域应用最广泛的模型,目前

研究者们通过引入双分支结构、残差连接和多尺度卷积等技术,优化了CNN在高光谱图像锐化中的表现。虽然基于CNN的方法是高光谱全色锐化的主力,但是基于GAN的高光谱全色锐化研究也不在少数,在这一领域发展前景良好。

鉴于高光谱全色锐化的对分辨率的高要求,研究者们首先将超分辨率技术和GAN模型结合起来。比如SRGAN (Ledig等, 2017) (Super-Resolution GAN)和ESRGAN (Romero等, 2020; Wang等, 2018) (Enhanced Super-Resolution GAN)技术在高光谱全色锐化中就得到了应用。SRGAN最初用于超分辨率任务,但其图像生成能力被引入,以提高多光谱图像的空间分辨率。ESRGAN作为SRGAN的改进版本,通过密集残差连接和去除批量归一化,进一步增强了模型的性能。基于ESRGAN, Luo等(2021)提出了FusGAN,将其应用于Sentinel-2和Landsat 8图像的锐化,将分辨率从10 m提高到2.5 m,这也展现了GAN在多源卫星图像联合使用中的潜力。

近几年的工作展示了GAN在提升高光谱融合效果方面的显著潜力,特别是将GAN和其他神经网络相结合的复合模型。HPGAN (Xie等, 2021)是一个经典的基于GAN的高光谱全色锐化方法,其在GAN模型中应用了三维卷积,并设计了专门的提取网络来融合高频特征,将输出转化为伪PAN图像,并将真PAN图像作为判别器的输入。physics-based GAN (Xiao等, 2021)另辟蹊径,利用物理退化模型结合多尺度残差通道注意力模块,也提升了图像细节和光谱保真度。CMAN (Su等, 2023)通过耦合自编码网络和多注意力机制,在端元和丰度的基础上融合HSI和MSI,提升了重建图像的准确性和细节表现。MFT-GAN (Shang等, 2024)通过引入Transformer网络与多尺度交互技术,实现了光谱和空间信息的深度融合。

我们发现,高光谱全色锐化的GAN方法更重视高频特征,通常会在方法中引入多尺度技术和自编码技术。此外,高光谱数据存在较强的波段间冗余,因此很多近几年的高光谱全色锐化模型采用注意力机制,以增强关键波段的特征表达能力。目前的研究不仅表明,GAN在高光谱全色锐化任务中具有有良好的适应性,也为进一步探索其应用提供了重要的理论和实践依据。

3.3 损失函数

损失函数的设计对网络模型的性能往往有着重要的影响。由于遥感图像融合通常涉及空间分辨率和光谱信息的优化,设计合理的损失函数是确保生成图像在各个方面取得平衡的关键。在基于GAN完成的遥感图像融合任务中,损失函数通过衡量生成图像与参考图像之间的差异,指导生成器和判别器的优化过程,从而影响最终融合图像的质量。在生成任务中,生成器的损失函数 L_{ADV} 是作为一个整体的统计期望来设计的。但对于多源图像融合任务来说是不够的,为了使融合图像在视觉和信息保留方面更接近理想状态,需要引入多种其他类型的损失函数。以下将对各类型损失函数作出详细介绍。

3.3.1 对抗性损失

对抗性损失 (Adversarial Loss) 是GAN的核心组成部分,也是其名称的由来。它通过生成器和判别器之间的对抗性训练来优化生成结果。在遥感图像融合任务中,对抗性损失使得生成器生成的融合图像更加逼真,确保它能够逐步逼近真实图像的分布。对抗性损失函数 L_{ADV} 即前文中提到过的最小最大化损失函数。

3.3.2 像素内容损失

像素内容损失 (Pixel Content Loss) 用于提高生成图像的空间分辨率和整体清晰度。通过计算生成图像与参考图像之间的像素差异,试图令生成的图像在像素级别上与目标图像保持一致。使用该损失函数的方法一般是监督学习方法。在遥感图像融合中,像素内容损失通常采用均方差损失MSE (Mean Squared Error) 和平均绝对误差损失MAE (Mean Absolute Error),也叫L2损失和L1损失,来衡量生成图像与参考图像之间的距离。

均方差损失的公式为

$$L_{MSE} = \frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2 \quad (4)$$

平均绝对误差损失的公式为

$$L_{MAE} = \frac{1}{n} \sum_{i=1}^n |x_i - y_i| \quad (5)$$

以上两个公式中, x_i 是真实图像的像素值, y_i 是生成图像的像素值, n 是像素的总数。

3.3.3 光谱损失

光谱损失 (Spectral Loss) 在遥感图像融合中尤为重要,尤其是在处理多光谱或高光谱图像时。该损失尽量提高生成图像在光谱维度上的保真度,使得不同波段的光谱信息在融合后得以准确保留。上述的L2损失也可以用作光谱损失。

常用的光谱损失方法是光谱角度映射 (SAM) (Hu等, 2020),其公式为

$$L_{SAM} = \arccos \left(\frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} \right) \quad (6)$$

式中, x 和 y 分别表示参考图像和生成图像的光谱向量,这个公式通过计算光谱向量之间的夹角来衡量光谱相似性,使融合图像的光谱信息尽可能接近真实数据,从而减少光谱失真。

3.3.4 感知损失

感知损失 (Perceptual Loss) 基于预训练的深度学习卷积神经网络 (如VGG网络),用于在高层次语义特征上衡量生成图像与参考图像的相似性 (Johnson等, 2016)。与直接比较像素值不同,感知损失关注图像的视觉感知质量,使得生成图像在视觉上尽量与真实图像一致 (Bello等, 2020)。感知损失的公式为

$$L_{Perceptual} = \sum_i \|\phi_i(x) - \phi_i(y)\|_2 \quad (7)$$

式中, $\phi_i(\cdot)$ 表示在预训练网络的第 i 层提取的特征图。通过计算损失,生成器能够捕捉图像中的高级语义信息。

3.3.5 循环损失

循环损失由CycleGAN (Zhu等, 2017) 提出,通过循环一致性,可以在无监督的情况下尽可能地保证图像的精度,近年来特别受到研究者的欢迎,在前文中已经介绍过,因此不再赘述。

3.3.6 综合损失函数

GAN在遥感图像融合的实际应用中,通常将上述多种损失函数结合起来,以实现生成图像在不同维度上的优化。综合损失函数的设计需要权衡不同损失项的相对重要性,并根据具体任务调整各个损失项的权重。

4 未来挑战

遥感图像融合方法经过三十多年的显著发展,

并且已经发展了许多方法。尽管GAN是其中很有前景的一类方法，目前仍然存在许多悬而未决的问题。在这一部分中，我们提出了关于这个主题的一些未来挑战。

4.1 数据驱动和模型驱动

目前，大多数基于GAN的方法属于数据驱动类型。将图像融合过程建模为观测方程，可以看作一个反问题，反问题的特点是不适定性（Liu等，2022）。数据驱动解决不适定性的方法是通过大规模的数据集进行特征学习。它的优点是不依赖于传统的经验模型，通常被看作黑箱模型，即研究人员并不需要清晰地理解模型内部的具体过程（胡建文等，2023）。然而数据驱动方法的局限性在于对大数据集和强大计算能力的依赖。相比之下，模型驱动方法能够提供更加透明的融合过程并解释模型的可信性，对于特定的融合任务至关重要。因此，如何将模型驱动与数据驱动相结合，设计出具有更强解释能力、且能够在不同场景下适应性更好的融合方法，也许会成为未来遥感图像融合研究的重要方向。

4.2 面向应用的扩展方法

迄今为止，研究者们已经开发了许多遥感图像融合的方法。可以应用于硬件设备上的算法却并不多。虽然少部分基于CS和基于MRA的融合方法已经应用于一些专业软件。但这些方法通常也会引入光谱失真的问题。尽管基于深度学习的方法不在少数，大多数方法对来自不同场景的遥感图像却没有很好的鲁棒性。在农业、城市规划、环境监测等领域的实际应用中，对遥感图像处理的需求各不相同。如何针对不同应用场景设计合适的架构，并将其与现有的遥感图像融合流程无缝结合，是未来研究的重要方向。

相比于其他模型，GAN模型的对抗性损失由竞争性的训练进行表达，能够在不依赖参考图像的情况下实现高质量的图像融合。因此，GAN在处理多源异构数据方面的潜力尚未充分挖掘，这为其在复杂应用中的扩展提供了广阔的研究空间。

4.3 模型训练的通用性与开销

目前大部分基于GAN的模型还是依赖于大规模标注的数据进行训练，这在实际应用中，特别是在高分辨率参考图像难以获取的情况下，往往

难以实现，无监督学习和半监督学习为这一问题提供了部分解决方案，但如何进一步提高这些模型的可靠性和在不同场景下的通用性，仍然需要研究者不断探索。

此外，目前绝大部分模型还是需要样本众多、体量较大的数据集来进行训练，这也导致了很大的训练开销，需要如何降低现有GAN模型的计算复杂度，使其在有限计算资源下也能高效运行，比如引入少量学习（Few-shot Learning）甚至零次学习（Zero-shot Learning），探讨其在遥感图像融合领域的适用性，也可以是未来研究的一个方向。

5 结 语

遥感图像融合技术为提高多源遥感数据的空间和光谱分辨率发挥着至关重要的作用，对去云、遥感目标检测和分类等下游任务的意义重大。目前在遥感图像融合领域，基于GAN的方法展示了提升融合图像空间和光谱分辨率的巨大潜力。与传统方法相比，GAN不仅在处理复杂数据的同时减少了光谱失真，还能够在没有高分辨率参考图像的情况下实现高质量的图像融合。本文探讨了适用于遥感图像融合的GAN框架演进，并从不同角度对近几年基于GAN的遥感图像融合方法进行了详细介绍。尽管这项技术已经有了很多应用，目前也取得了显著进展，但还是面临一些挑战，本文也针对一些扩展性问题展开了讨论。未来的研究应继续优化GAN的架构和学习策略，克服目前存在的挑战，从而使其在更广泛的遥感应用中发挥出应有的作用。

参考文献(References)

- Achalakul T and Taylor S. 2003. A distributed spectral-screening PCT algorithm. *Journal of Parallel and Distributed Computing*, 63(3): 373-384 [DOI: 10.1016/S0743-7315(03)00017-0]
- Bai L Y, Xu C M and Wang C. 2015. A review of fusion methods of multi-spectral image. *Optik*, 126(24): 4804-4807 [DOI: 10.1016/j.ijleo.2015.09.201]
- Balakrishnan S, Cacciola M, Udpa L, Rao B P, Jayakumar T and Raj B. 2012. Development of image fusion methodology using discrete wavelet transform for eddy current images. *NDT and E International*, 51: 51-57 [DOI: 10.1016/j.ndteint.2012.06.006]
- Bello J L G, Seo S and Kim M. 2020. Pan-sharpening with color-aware perceptual loss and guided re-colorization//*Proceedings of the 2020 IEEE International Conference on Image Processing*

- (ICIP). Abu Dhabi: IEEE: 908-912 [DOI: 10.1109/ICIP40778.2020.9190785]
- Benzenati T, Kessentini Y and Kallel A. 2022. Pansharpening approach via two-stream detail injection based on relativistic generative adversarial networks. *Expert Systems with Applications*, 188: 115996 [DOI: 10.1016/j.eswa.2021.115996]
- Carper W J, Lillesand T M and Kiefer R W. 1990. The use of intensity-hue-saturation transformations for merging SPOT panchromatic and multispectral image data. *Photogrammetric Engineering and Remote Sensing*, 56(4): 459-467
- Cliche G, Bonn F and Teillet P. 1985. Integration of the SPOT panchromatic channel into its multispectral mode for image sharpness enhancement. *Photogrammetric Engineering and Remote Sensing*, 51(3): 311-316
- Diao W X, Jin M Y, Zhang K and Xiao L. 2023. Unsupervised single-generator CycleGAN-based pansharpening with spatial-spectral degradation modeling. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 16: 10246-10263 [DOI: 10.1109/JSTARS.2023.3327169]
- Dong C, Loy C C, He K M and Tang X O. 2016. Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2): 295-307 [DOI: 10.1109/TPAMI.2015.2439281]
- Dong W Q, Hou S X, Xiao S, Qu J H, Du Q and Li Y S. 2022. Generative dual-adversarial network with spectral fidelity and spatial enhancement for hyperspectral pansharpening. *IEEE Transactions on Neural Networks and Learning Systems*, 33(12): 7303-7317 [DOI: 10.1109/TNNLS.2021.3084745]
- Garzelli A and Nencini F. 2007. Panchromatic sharpening of remote sensing images using a multiscale Kalman filter. *Pattern Recognition*, 40(12): 3568-3577 [DOI: 10.1016/j.patcog.2007.05.002]
- Ghosh A, Kulharia V, Namboodiri V P, Torr P H S and Dokania P K. 2018. Multi-agent diverse generative adversarial networks//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE: 8513-8521 [DOI: 10.1109/CVPR.2018.00888]
- Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A and Bengio Y. 2020. Generative adversarial networks. *Communications of the ACM*, 63(11): 139-144 [DOI: 10.1145/3422622]
- Gui J, Sun Z N, Wen Y G, Tao D C and Ye J P. 2023. A review on generative adversarial networks: algorithms, theory, and applications. *IEEE Transactions on Knowledge and Data Engineering*, 35(4): 3313-3332 [DOI: 10.1109/TKDE.2021.3130191]
- He L J, Ren Z H, Zhang W Y, Li F and Mei S H. 2024. Unsupervised pansharpening based on double-cycle consistency. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5613015 [DOI: 10.1109/TGRS.2024.3371971]
- Hu J W, Du C G and Fan S S. 2020. Two-stage pansharpening based on multi-level detail injection network. *IEEE Access*, 8: 156442-156455 [DOI: 10.1109/ACCESS.2020.3019201]
- Hu J W, Hu P, Kang X D, Zhang H and Fan S S. 2021. Pan-sharpening via multiscale dynamic convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 59(3): 2231-2244 [DOI: 10.1109/TGRS.2020.3007884]
- Hu J W, Wang Z P and Hu P. 2023. A review of pansharpening methods based on deep learning. *Remote Sensing for Natural Resources*, 35(1): 1-14 (胡建文, 汪泽平, 胡佩. 2023. 基于深度学习的空谱遥感图像融合综述. *自然资源遥感*, 35(1): 1-14) [DOI: 10.6046/zrzyg.2021433]
- Jin X, Feng Y T, Jiang Q, Miao S F, Chu X, Zheng H Q M and Wang Q Q. 2024. UPGAN: an unsupervised generative adversarial network based on U-shaped structure for pansharpening. *ISPRS International Journal of Geo-Information*, 13(7): 222 [DOI: 10.3390/ijgi13070222]
- Jin X, Huang S S, Jiang Q, Lee S J, Wu L W and Yao S W. 2021. Semisupervised remote sensing image fusion using multiscale conditional generative adversarial network with siamese structure. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14: 7066-7084 [DOI: 10.1109/JSTARS.2021.3090958]
- Johnson J, Alahi A and Fei-Fei L. 2016. Perceptual losses for real-time style transfer and super-resolution//Proceedings of the 14th European Conference on Computer Vision - ECCV 2016. Amsterdam: Springer: 694-711 [DOI: 10.1007/978-3-319-46475-6_43]
- Khaleghi B, Khamis A, Karray F O and Razavi S N. 2013. Multisensor data fusion: a review of the state-of-the-art. *Information Fusion*, 14(1): 28-44 [DOI: 10.1016/j.inffus.2011.08.001]
- Ledig C, Theis L, Huszar F, Caballero J, Cunningham A, Acosta A, Aitken A, Tejani A, Totz J, Wang Z H and Shi W Z. 2017. Photo-realistic single image super-resolution using a generative adversarial network//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE: 105-114 [DOI: 10.1109/cvpr.2017.19]
- Liu M Y and Tuzel O. 2016. Coupled generative adversarial networks//Proceedings of the 30th International Conference on Neural Information Processing Systems. Barcelona: Curran Associates Inc.: 469-477 [DOI: 10.48550/arXiv.1606.07536]
- Liu P, Li J, Wang L Z and He G J. 2022. Remote sensing data fusion with generative adversarial networks: state-of-the-art methods and future research directions. *IEEE Geoscience and Remote Sensing Magazine*, 10(2): 295-328 [DOI: 10.1109/MGRS.2022.3165967]
- Liu Q J, Zhou H Y, Xu Q Z, Liu X Y and Wang Y H. 2021. PSGAN: a generative adversarial network for remote sensing image pansharpening. *IEEE Transactions on Geoscience and Remote Sensing*, 59(12): 10227-10242 [DOI: 10.1109/TGRS.2020.3042974]
- Liu X B, Liu X, Dai H R, Kang X D, Plaza A and Zu W J. 2023. MUNGAN: a multiscale unsupervised network for remote sensing image pansharpening. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 5404018 [DOI: 10.1109/TGRS.2023.3288073]
- Liu X Y, Liu Q J and Wang Y H. 2020. Remote sensing image fusion based on two-stream fusion network. *Information Fusion*, 55: 1-15 [DOI: 10.1016/j.inffus.2019.07.010]
- Luo X, Tong X H and Hu Z W. 2021. Improving satellite image fusion

- via generative adversarial training. *IEEE Transactions on Geoscience and Remote Sensing*, 59(8): 6969-6982 [DOI: 10.1109/TGRS.2020.3025821]
- Ma J Y, Xu H, Jiang J J, Mei X G and Zhang X P. 2020b. DDcGAN: a dual-discriminator conditional generative adversarial network for multi-resolution image fusion. *IEEE Transactions on Image Processing*, 29: 4980-4995 [DOI: 10.1109/TIP.2020.2977573]
- Ma J Y, Yu W, Chen C, Liang P W, Guo X J and Jiang J J. 2020a. 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]
- Masi G, Cozzolino D, Verdoliva L and Scarpa G. 2016. Pansharpening by convolutional neural networks. *Remote Sensing*, 8(7): 594 [DOI: 10.3390/rs8070594]
- Mehra I and Nishchal N K. 2014. Image fusion using wavelet transform and its application to asymmetric cryptosystem and hiding. *Optics Express*, 22(5): 5474-5482 [DOI: 10.1364/OE.22.005474]
- Mirza M and Osindero S. 2014. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784* [DOI: 10.48550/arXiv.1411.1784]
- Odena A, Olah C and Shlens J. 2017. Conditional image synthesis with auxiliary classifier GANs//*Proceedings of the 34th International Conference on Machine Learning*. Sydney: JMLR.org: 2642-2651
- Ozcelik F, Alganci U, Sertel E and Unal G. 2021. Rethinking CNN-based pansharpening: guided colorization of panchromatic images via GANs. *IEEE Transactions on Geoscience and Remote Sensing*, 59(4): 3486-3501 [DOI: 10.1109/TGRS.2020.3010441]
- Ratliff L J, Burden S A and Sastry S S. 2013. Characterization and computation of local Nash equilibria in continuous games//*Proceedings of the 2013 51st Annual Allerton Conference on Communication, Control, and Computing*(Allerton). Monticello: IEEE: 917-924 [DOI: 10.1109/Allerton.2013.6736623]
- Romero L S, Marcello J and Vilaplana V. 2020. Super-resolution of sentinel-2 imagery using generative adversarial networks. *Remote Sensing*, 12(15): 2424 [DOI: 10.3390/rs12152424]
- Shang Y L, Liu J J, Zhang J Y and Wu Z B. 2024. MFT-GAN: a multi-scale feature-guided transformer network for unsupervised hyperspectral pansharpening. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5518516 [DOI: 10.1109/TGRS.2024.3402058]
- Shao Z M, Lu Z X, Ran M S, Fang L Y, Zhou J L and Zhang Y. 2020. Residual encoder - decoder conditional generative adversarial network for pansharpening. *IEEE Geoscience and Remote Sensing Letters*, 17(9): 1573-1577 [DOI: 10.1109/LGRS.2019.2949745]
- Su L J, Sui Y X and Yuan Y. 2023. An unmixing-based multi-attention GAN for unsupervised hyperspectral and multispectral image fusion. *Remote Sensing*, 15(4): 936 [DOI: 10.3390/rs15040936]
- Sun S L and Deng Z L. 2004. Multi-sensor optimal information fusion Kalman filter. *Automatica*, 40(6): 1017-1023 [DOI: 10.1016/j.automatica.2004.01.014]
- Tu T M, Cheng W C, Chang C P, Huang P S and Chang J C. 2007. Best tradeoff for high-resolution image fusion to preserve spatial details and minimize color distortion. *IEEE Geoscience and Remote Sensing Letters*, 4(2): 302-306 [DOI: 10.1109/LGRS.2007.894143]
- Wald L, Ranchin T and Mangolini M. 1997. Fusion of satellite images of different spatial resolutions: assessing the quality of resulting images. *Photogrammetric Engineering and Remote Sensing*, 63(6): 691-699
- Wang X T, Yu K, Wu S X, Gu J J, Liu Y H, Dong C, Qiao Y and Loy C C. 2018. ESRGAN: enhanced super-resolution generative adversarial networks//*Proceedings of the European Conference on Computer Vision (ECCV) Workshops*. Munich: Springer: 63-79 [DOI: 10.1007/978-3-030-11021-5_5]
- Wang Y J, Xie Y Y, Wu Y Y, Liang K and Qiao J L. 2022. An unsupervised multi-scale generative adversarial network for remote sensing image pan-sharpening//*Proceedings of the 28th International Conference on Multimedia Modeling*. Phu Quoc: Springer: 356-368 [DOI: 10.1007/978-3-030-98355-0_30]
- Wang Z W, She Q and Ward T E. 2021. Generative adversarial networks in computer vision: a survey and taxonomy. *ACM Computing Surveys*(CSUR), 54(2): 37 [DOI: 10.1145/3439723]
- Wei Y C, Yuan Q Q, Shen H F and Zhang L P. 2017. Boosting the accuracy of multispectral image pansharpening by learning a deep residual network. *IEEE Geoscience and Remote Sensing Letters*, 14(10): 1795-1799 [DOI: 10.1109/LGRS.2017.2736020]
- 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, Cui Y H, Li Y S, Lei J, Du Q and Li J J. 2021. HPGAN: hyperspectral pansharpening using 3-D generative adversarial networks. *IEEE Transactions on Geoscience and Remote Sensing*, 59(1): 463-477 [DOI: 10.1109/TGRS.2020.2994238]
- Yang J F, Fu X Y, Hu Y W, Huang Y, Ding X H and Paisley J. 2017. PanNet: a deep network architecture for pan-sharpening//*Proceedings of the 2017 IEEE International Conference on Computer Vision*. Venice: IEEE: 1753-1761 [DOI: 10.1109/ICCV.2017.193]
- Yang Z G, Chen Y P, Le Z L and Ma Y. 2021. GANFuse: a novel multi-exposure image fusion method based on generative adversarial networks. *Neural Computing and Applications*, 33(11): 6133-6145 [DOI: 10.1007/S00521-020-05387-4]
- Yuan Q Q, Wei Y C, Meng X C, Shen H F and Zhang L P. 2018. A multiscale and multidepth convolutional neural network for remote sensing imagery pan-sharpening. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(3): 978-989 [DOI: 10.1109/JSTARS.2018.2794888]
- Zhang H and Ma J Y. 2023. STP-SOM: scale-transfer learning for pansharpening via estimating spectral observation model. *International Journal of Computer Vision*, 131(12): 3226-3251 [DOI: 10.1007/s11263-023-01840-8]
- Zhang H, Xu H, Tian X, Jiang J J and Ma J Y. 2021. Image fusion meets deep learning: a survey and perspective. *Information Fusion*, 76: 323-336 [DOI: 10.1016/j.inffus.2021.06.008]

- Zhang H, Xu T, Li H S, Zhang S T, Wang X G, Huang X L and Metaxas D. 2017. StackGAN: text to photo-realistic image synthesis with stacked generative adversarial networks//Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice: IEEE: 5908-5916 [DOI: 10.1109/ICCV.2017.629]
- Zhong X W, Qian Y R, Liu H, Chen L, Wan Y L, Gao L, Qian J and Liu J. 2021. Attention_FPNNet: two-branch remote sensing image pansharpening network based on attention feature fusion. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14: 11879-11891 [DOI: 10.1109/JSTARS.2021.3126645]
- Zhou H Y, Liu Q J and Wang Y H. 2021. PGMAN: an unsupervised generative multiadversarial network for pansharpening. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14: 6316-6327 [DOI: 10.1109/JSTARS.2021.3090252]
- Zhou H Y, Liu Q J, Weng D W and Wang Y H. 2022. Unsupervised cycle-consistent generative adversarial networks for pan sharpening. IEEE Transactions on Geoscience and Remote Sensing, 60: 5408814 [DOI: 10.1109/TGRS.2022.3166528]
- Zhu J Y, Park T, Isola P and Efros A A. 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks//Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice: IEEE: 2242-2251 [DOI: 10.1109/ICCV.2017.244]

Research progress of generative adversarial networks in remote sensing image fusion

ZHENG Huangqimei, PAN Chengyi, JIN Xin, WANG Qianqian, MIAO Shengfa, JIANG Qian

1.School of Software, Yunnan University, Kunming 650000, China;

2.Engineering Research Center of Cyberspace, Yunnan University, Kunming 650000, China

Abstract: Remote sensing image fusion, as an important branch of data fusion, has attracted considerable attention in recent years. Images with high spatial and spectral resolutions can be generated by fusing remote sensing data from different sources, improving the ability to capture detailed information about ground objects. In the field of remote sensing image processing, the application of Generative Adversarial Networks (GANs) has developed rapidly, demonstrating significant advantages. This paper starts by introducing traditional remote sensing image fusion methods, followed by an analysis of GAN-based fusion techniques, with a detailed explanation of GAN's applications in tasks such as pan-sharpening and hyperspectral pan-sharpening.

Although traditional remote sensing image fusion methods can improve either spatial or spectral resolution to some extent, they often introduce spectral distortions, limiting their effectiveness. In recent years, various deep learning-based image fusion models have emerged. Among them, GANs, with their unique adversarial training mechanism, have demonstrated a remarkable ability to generate fused images that effectively balance spatial and spectral information while significantly reducing distortion. Numerous GAN-based remote sensing image fusion models, which leverage multiscale feature extraction, separate processing of spectral and spatial information, and optimized adversarial losses to substantially enhance the quality of the fused images, have been proposed. GANs are especially suitable for unsupervised and semisupervised learning, enabling the generation of high-quality fused images without reference images and thus addressing the common issue of the lack of reference data in remote sensing.

This study analyzes common architectures of GANs used in multisource image fusion, including single-generator—double-discriminator, multigenerator—single-discriminator, multigenerator—multidiscriminator, and conditional GAN architectures. These architectures, using different training strategies, effectively handle a range of task scenarios. This study also explores the application of GANs in remote sensing image fusion. In the pan-sharpening task, GAN-based models not only improve spatial resolution but also enhance the fidelity of image details while preserving spectral information. Similarly, GANs show great potential in hyperspectral pan-sharpening, especially when combined with super-resolution techniques and adversarial learning, allowing for the effective fusion of multisource images and generating high-resolution multispectral and hyperspectral images. This paper discusses the importance of loss functions such as adversarial loss, spectral loss, and perceptual loss in GAN training, all of which ensure balanced optimization across different dimensions of the fused images.

Finally, this paper summarizes the current state of GAN-based remote sensing image fusion methods and explores future research directions, including the integration of data- and model-driven approaches. It also expands on the challenges of generalization and computational complexity of GAN models in various remote sensing applications, proposing potential avenues for further optimization.

Key words: remote sensing image, multi-source data fusion, deep learning, generative adversarial network, pan-sharpening, multispectral image, hyperspectral image

Supported by National Natural Science Foundation of China (No. 62101481, 62261060); Yunnan Provincial Basic Research Program Projects (No. 202301AW070007, 202201AU070033, 202201AT070112, 202301AU070210)