

基于时谱域融合与时序自注意力增强的无监督遥感云层遮挡图像修复^①



邱涛¹, 吴倩², 张艳艳¹

¹(南京信息工程大学 电子与信息工程学院, 南京 210044)

²(金陵科技学院 智能科学与控制工程学院, 南京 211169)

通信作者: 邱涛, E-mail: 202283270477@nuist.edu.cn

摘要: 光学遥感图像中的云遮挡问题是遥感数据处理中的核心挑战之一, 针对目前云层去除技术在处理云层边缘信息及保留图像细节方面的缺陷, 本文提出基于时谱域融合与时序自注意力增强的生成对抗网络 (TGAN). TGAN 通过其两阶段模块化设计, 实现了遥感图像修复质量与处理效率的双重提升. 第 1 阶段, 基于时序自注意力机制的特征提取模块, 用线性升维层捕获时域、谱域特征, 以一维线性降维层弥补传统最大池化不足, 增强时间序列位置特征建模能力, 还设计含权重分配策略的多头自注意力机制精准捕捉边缘信息; 第 2 阶段为自适应图像修复模块, 由随机噪声消除与局部对比增强子模块协同改善图像细节、抑制噪声. 此外, TGAN 的鉴别器采用多尺度模块, 这一策略实现了全局一致性与局部细节之间的平衡. 通过生成器与鉴别器之间的交互博弈, 生成器持续优化修复图像, 以提高修复效果. 这种动态的博弈过程推动了生成器在图像修复任务中的迭代优化. 为了验证 TGAN 的有效性, 本文在 Sen2_MTC 数据集上进行了实验. 实验结果表明, TGAN 在峰值信噪比 (PSNR) 和主观视觉评估方面均显著优于现有方法, 在训练集和测试集的 PSNR 分别达到了 21.547 dB 和 20.206 dB, 表明该方法在遥感云层图像修复任务中具有较好的性能与应用潜力.

关键词: 时谱域融合; 云层检测; 多头自注意力机制; 云层图像修复

引用格式: 邱涛, 吴倩, 张艳艳. 基于时谱域融合与时序自注意力增强的无监督遥感云层遮挡图像修复. 计算机系统应用, 2025, 34(7): 59-71. <http://www.c-s-a.org.cn/1003-3254/9908.html>

Unsupervised Remote Sensing Cloud Occlusion Image Restoration Based on Temporal-spectral Domain Fusion and Temporal Self-attention Enhancement

QIU Tao¹, WU Qian², ZHANG Yan-Yan¹

¹(School of Electronics & Information Engineering, Nanjing University of Information Science & Technology, Nanjing 210044, China)

²(College of Intelligent Science and Control Engineering, Jinling Institute of Technology, Nanjing 211169, China)

Abstract: Cloud occlusion in optical remote sensing images is one of the core challenges in remote sensing data processing. To address the limitations of current cloud removal technologies in handling cloud edge information and preserving image details, a generative adversarial network (TGAN) based on temporal-spectral domain fusion and temporal self-attention enhancement is proposed. Through its two-stage modular design, TGAN simultaneously improves the quality of remote sensing image restoration and processing efficiency. In the first stage, the feature extraction module, based on a temporal self-attention mechanism, uses a linear expansion layer to capture temporal and spectral domain features, compensating for the limitations of traditional maximum pooling with a one-dimensional linear dimensionality reduction layer, thus enhancing the modeling capability of time-series positional features. A multi-head self-attention

① 基金项目: 国家自然科学基金面上项目 (62272234); 国家重点研发计划 (2022YFB4401301)

收稿时间: 2024-12-16; 修改时间: 2025-01-07, 2025-02-11; 采用时间: 2025-02-18; csa 在线出版时间: 2025-05-16

CNKI 网络首发时间: 2025-05-19

mechanism with a weight allocation strategy is designed to accurately capture edge information. The second stage is an adaptive image restoration module, composed of a random noise cancellation submodule and a local contrast enhancement submodule, which collaboratively enhances image details and suppresses noise. In addition, TGAN's discriminator incorporates multi-scale modules, a strategy that balances global consistency and local detail. Through the interactive game between the generator and discriminator, the generator continuously optimizes the restored image, improving restoration performance. This dynamic adversarial process drives iterative optimization of the generator in the image restoration task. To verify the effectiveness of TGAN, experiments are conducted on the Sen2_MTC dataset. The results show that TGAN significantly outperforms existing methods in terms of peak signal-to-noise ratio (PSNR) and subjective visual evaluation, with PSNR values of 21.547 dB and 20.206 dB for the training and test sets, respectively, indicating that TGAN demonstrates strong performance and application potential in remote sensing cloud image restoration.

Key words: temporal-spectral domain fusion; cloud detection; multi-head self-attention mechanism; cloud image restoration

在光学遥感图像处理中, 云层遮挡现象是遥感数据处理的一项核心挑战之一。云层遮挡导致地表信息的显著缺失与质量恶化, 进而影响地物识别、变化监测及土地覆盖分类等后续应用任务的精确性和实效性。云遮挡问题研究的焦点集中于恢复被云层遮蔽区域的地表特征, 以提升遥感数据的完整度和可用性, 现有云去除技术主要包括多时相法、多光谱法和基于深度学习的方法。

多时相法的云层去除效果受时序数据获取的准确性以及地表动态变化的影响较大; 多光谱法则对高光谱分辨率数据的依赖性较强, 在云层较厚的情况下去除效果受限; 而基于深度学习的云去除技术性能则依赖于大规模标注数据集, 效果容易受到训练数据质量及其分布特征的制约。因此, 如何有效融合多时相与多光谱方法的物理优势以及深度学习的强大建模能力, 同时降低对高质量标注数据的依赖, 已成为当前研究的重要课题与发展方向。

在此背景下, 2017年, Yu 等人^[1]针对交通数据非线性和复杂性导致中长期预测难的问题, 提出 STGCN 结构。该模型基于卷积, 利用谱图卷积提取空间特征, 用门控 CNN 提取时间特征, 通过时空卷积块堆叠构建网络, 能并行处理, 有更少参数和更快训练速度, 有效捕捉交通数据的时空特征以进行预测。2023年, Wang 等人^[2]提出 TF-GridNet 这一用于无回声条件下单通道语音分离的时频域多路径深度神经网络, 通过堆叠包含帧内频域、子带时域和全带自注意力模块的多路径块, 利用局部和全局时频信息分离, 采用复数谱映射,

还添加新损失函数, 在 WSJ0-2mix 数据集上取得良好效果。2024年, Zhang 等人^[3]针对建筑设计信息不全影响结构地震响应估计的问题, 从城市尺度建筑群的有限建筑信息出发, 结合深度学习, 借鉴自编码器模块设计, 搭建自适应多尺度融合输出的端到端网络, 输入结构参数和地震动时程, 输出建筑各层时程响应, 经超参数优化后模型预测性能良好、计算效率高。以上研究成果表明, 多时相多光谱融合的图像修复方法在提升遥感图像质量与增强信息获取能力上具有一定的效果, 借助融合不同时间点的多光谱数据, 可消除云层以及其他干扰因素, 使地面信息得以清晰呈现, 同时强化了空间和光谱特征, 进而提高了检测与分类的精准度, 减少了数据丢失的风险。

然而, 尽管基于深度学习的多时相多光谱融合图像修复方法在提取云层多维特征方面展现出一定优势, 但在实际应用中仍面临诸多挑战。首先, 这些方法通常依赖深层卷积网络进行特征提取, 但在复杂的遥感场景中可能引入噪声, 从而降低复原图像的质量。此外, 现有技术普遍缺乏针对性的细节增强机制, 难以有效修复云层遮挡导致的局部细节损失, 无法满足高精度复原的要求。同时, 传统注意力机制 (Gulati 等人^[4]提出的 Conformer 模型) 在处理云遮挡图像时对边缘信息的关注不足, 限制了修复效果的一致性和全面性。这些问题表明, 云遮挡图像修复在高质量复原和信息保真度提升方面仍有显著改进空间。因此, 亟需开发一种结合全局与局部特征的协同建模方案, 以强化边缘信息的捕捉, 并通过自适应细节增强技术实现高精度修复,

从而提升图像的视觉质量与信息完整性,为后续遥感应用提供更为可靠的数据支持。

在遥感图像处理中,云遮挡问题严重影响了图像的质量和信​​息提取。为了解决这一问题,本文提出了一种结合时谱域融合与时序自注意力增强的无监督精细修复技术。该方法充分整合了云层的时域与谱域特征,显著降低了对完整配对图像数据的依赖。该方法通过结合全局特征建模与局部细节优化,设计了自适应处理框架,以增强修复图像的全局一致性和细节保真度。在GAN的交互博弈中,生成器与鉴别器相互作用,推动生成器不断优化修复结果,从而在无监督条件下显著提升遥感图像的修复性能。相较于现有云去除技术,TGAN在视觉质量和信息保真度方面均展现出显著优势。

1 相关工作

遥感图像修复技术在应对云层遮挡、噪声干扰及其他环境因素导致的图像质量下降问题上发挥着重要作用,已成为遥感图像处理领域的研究热点之一。近年来,基于深度学习的无监督图像修复方法因其无需依赖大量标注数据的显著优势,吸引了广泛关注。借助深度神经网络强大的特征学习能力,这些方法能够从大规模未标注数据中自动提取和建模复杂特征,快速生成高质量修复结果。与传统方法相比,无监督学习在处理复杂的图像缺陷时展现出更强的灵活性和适应性,不仅显著提升了修复效果,还提高了处理效率,适用于多种复杂场景。因此,基于无监督学习的图像修复技术是目前遥感图像处理的主要研究方向。

生成对抗网络 (generative adversarial network, GAN) 作为一种无监督学习框架,已成为生成模型领域的关键研究方向。其核心机制通过生成器和判别器的对抗训练实现:生成器从随机噪声中生成样本,判别器负责区分生成样本与真实样本,以不断提升生成质量。自2014年Goodfellow等人^[5]提出GAN网络以来,该模型在图像生成、文本建模和语音合成等领域得到了广泛应用。与其他生成模型(如自回归模型和变分自编码器)相比,GAN无需显式建模数据分布,通过生成器与判别器的动态博弈生成高质量样本,展现出简单高效的特点。以上优势使得GAN网络在遥感图像修复中的应用备受关注,其强大的生成能力为解决云遮挡、噪声干扰和数据缺失等问题提供了有效途径。通过生成器与判别器的对抗训练,GAN能够从受损图像中恢

复复杂的结构和纹理特征,实现高质量的图像修复。为增强生成器的表达能力,自注意力机制的引入使得模型能够有效地捕捉输入数据中的全局特征及其复杂的内部关系,从而显著提升生成效果。尤其是自注意力生成对抗网络 (self-attention GAN, SAGAN)^[6]等变体的引入,使得模型能够更高效地捕捉多尺度结构和全局上下文信息,进一步提升修复效果。基于GAN网络的遥感图像数据修复,为复杂场景下遥感数据的精细化处理开辟了新方向。

尽管GAN在遥感图像修复中展现出显著优势,但仍面临多项技术挑战。首先,遥感图像的高分辨率和复杂地物分布使得传统GAN在生成细节时容易模糊,难以满足高质量修复要求。其次,遥感数据的多源异构性导致不同图像在光谱特性上存在较大差异,增加了数据融合的难度。此外,GAN的生成能力依赖于数据分布,受损图像与参考图像的统计特性差异可能导致伪影生成,影响修复效果。GAN训练需要大量高质量标注数据,而遥感领域的配对样本极为有限,限制了模型性能提升。

因此,如何利用GAN精确建模遥感图像数据分布,解决训练中的不稳定性、模式崩溃和伪影问题,成为重要挑战。特别是多光谱和多时相图像处理中,GAN在捕捉多尺度特征和保持时空一致性方面仍面临瓶颈。此外,从未标注数据中自动学习复杂特征分布以应对数据稀缺性和多源异构性,对GAN的结构设计和训练策略提出了更高要求。接下来将介绍无监督学习、时谱域信息融合策略、GAN网络结构及自注意力机制的工作原理。

1.1 无监督学习

无监督学习为解决现实场景中对无标记数据的需求提供了解决思路,并推动了相关研究领域的发展^[7]。根据训练数据中标签的有无以及数量,学习方法通常分为有监督学习、半监督学习和无监督学习。不同于有监督学习和半监督学习,无监督学习完全摆脱标签依赖,通过探索数据内在规律进行分析,利用聚类、降维及关联规则挖掘等技术揭示数据的潜在模式,从而降低数据标注成本并支持后续分析和决策,尤其在标注困难或昂贵的领域展现出独特优势。

当前,深度学习的数据修复方法通常依赖完整数据集进行有监督训练,但在实际场景中,获取完整数据集存在挑战,限制了有监督方法的性能。相比之下,基

于无监督学习的数据修复技术,特别是生成对抗网络(GAN),在处理缺失值问题上展现出巨大潜力。GAN通过生成器和判别器的对抗训练,能够捕获不完整或异构数据的分布特征,实现对缺失值的高效估计,无需依赖标注数据,生成高质量修复结果,提升数据修复的准确性与鲁棒性,为解决缺失数据问题提供新思路。

1.2 时谱域信息融合策略

云修复技术的目标在于恢复云遮挡区域的地表信息,并尽可能还原真实的空间和光谱特性。然而,由于云遮挡对遥感图像的多重影响,其修复过程极具挑战性。传统的单一信息修复方法难以应对云层干扰的复杂性,因此引入时谱域信息融合策略已成为提升云修复效果的关键研究方向。

首先,时谱域信息融合能够充分利用多维数据特性,增强了云修复的全面性与鲁棒性。在时域上,多时相遥感数据捕捉了同一区域在不同时刻的动态变化,为云遮挡区域的恢复提供了时间冗余信息,通过分析这些变化,可以更准确地推断被云层遮挡的地表特征。同时,遥感图像的光谱域提供了丰富的物理信息,不同波段反映了地物的材质和反射特性,有助于地表特征的识别与还原。将时域动态关联与谱域物理特性结合,可以更有效地恢复云遮挡区域的细节和一致性。例如,2024年,Li等人^[8]提出一种双融合注意力机制的多模态网络,通过伪点云对齐二维图像与三维点云数据,结合图像特征与点云信息。该方法在 Semantic-KITTI 等基准测试中表现优异,通过谱域特征对齐与多模态融合,有效解决了点云数据缺失问题,提升分割精度与鲁棒性。

其次,时谱域融合策略在云修复中的多尺度特征建模具有显著优势,能够应对云层遮挡的复杂形态,包括边缘过渡和核心厚重区域。该策略结合时域信息提供的长时间序列上下文关系与谱域信息的波段协同作用,提升了对地物细节和纹理的重建能力。通过深度结合时谱域,修复模型实现了从全局到局部的多尺度特

征建模,有效解决了传统方法在复杂场景中细节丢失的问题。例如,2022年,Xu等人^[9]提出了一种互惠引导的半监督薄云去除框架,利用卷积块注意力模块(CBAM)实现薄云图像的空间和光谱定位,并通过跨尺度融合结构整合多尺度上下文信息,增强土地纹理的空间连续性和去云结果的高分辨率细节。

最后,时谱域融合为无监督学习框架中的遥感云修复提供了重要理论支持。由于高质量配对标注数据获取成本高且难以覆盖多样化场景,无监督学习依赖于模型对数据分布的自动建模。整合时谱域信息后,无监督模型能够捕捉云层遮挡区域的时空一致性和光谱特征,从未标注数据中挖掘规律,减少对人工标注的依赖,并缓解多源异构性问题,提升模型的泛化性能。例如,2023年,Wang等人^[10]构建无监督域分解网络,其中时谱域融合起关键指导作用。它融合不同时间序列与光谱特征,为无监督学习提供丰富且具代表性的特征空间,引导模型精准识别厚云与地物,实现高精度厚云去除,大幅提升多时相遥感图像质量与可用性。

因此,时域动态关联与谱域物理特性的联合建模对无监督遥感云修复至关重要,能够捕捉地表特征的时序变化,确保修复结果的一致性,同时增强模型对多源数据的适应能力,为高质量云修复奠定技术基础。

1.3 生成式对抗网络

GAN网络由两个核心模块构成:生成器和判别器。生成器接收随机噪声输入,通过学习数据分布生成假样本,目标是使生成样本在特征上尽可能接近真实数据,从而“欺骗”判别器。判别器则通过生成对抗损失,评估生成样本与真实样本的相似性,其目标是准确区分样本来源,即判断数据是来自真实数据集还是由生成器生成。这种对抗性博弈机制驱动两者在训练过程中相互优化:生成器不断提升生成能力以生成更真实的数据,而判别器则持续增强判别能力以提高识别精度。通过这种动态平衡,GAN系统逐步提升生成样本的质量和真实性。模型结构如图1所示。

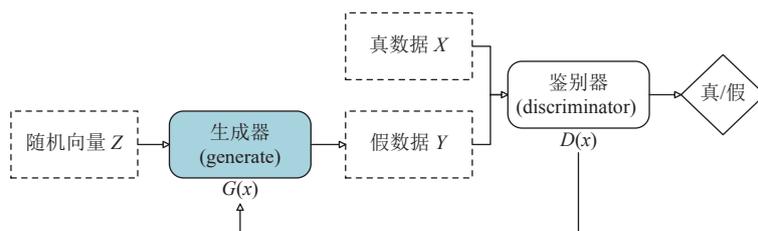


图1 GAN的模型结构

随着对 GAN 网络的不断研究改进, 研究者们探索更高效的模型结构, 以捕捉复杂数据特征并提升生成质量. 2015 年, Shen 等人^[11]提出了时空生成对抗网络 (STGAN), 针对视频生成和时序数据合成问题, 通过融合空间和时间信息, 并引入自注意力机制捕捉长距离依赖关系, 从而生成具有高视觉真实感和时间连贯性的时空数据. STGAN 采用生成器与判别器的对抗训练策略, 使生成结果在视觉质量和时间序列一致性方面均表现优异. 2023 年, Yu 等人^[12]进一步改进 GAN 结构, 提出结合生成对抗网络 (GAN) 与卷积神经网络 (CNN) 的深度卷积生成对抗网络 (DCGAN). 该方法通过引入步幅卷积 (strided convolution) 替代传统池化操作, 实现自学习下采样, 从而在保留图像细节的同时, 避免了池化过程中的信息丢失. DCGAN 的生成器采用 tanh 激活函数, 判别器则使用 LeakyReLU 激活函数, 有效提升了高分辨率图像生成任务的性能. 实验表明, DCGAN 不仅显著缩短了训练时间, 还提高了生成质量与稳定性, 为大尺寸高质量图像的生成提供了精准、高效的解决方案, 其优势在图像生成与处理等领域得到了广泛验证和应用.

虽然 GAN 网络在深度学习领域取得了显著进展, 但应用于实际的遥感图像修复中仍存在诸多挑战, 如模型易陷入崩溃以及生成样本缺乏多样性. 这些问题不仅限制了 GAN 模型的训练效果, 也对实际应用构成了一定障碍. 因此, 分析 GAN 模型的核心问题及其根源, 并探索有效的改进策略, 已成为 GAN 研究的关键方向之一. 相比传统 GAN, 基于 Wasserstein 距离的生成对抗网络 (WGAN)^[13,14]通过优化 Wasserstein 距离, 有效缓解了模型崩溃和样本退化现象, 从而显著提升了训练的稳定性和效果. WGAN 通过引入梯度削弱和梯度惩罚等技术, 不仅增强了生成样本的多样性, 还进一步提高了生成模型的逼真度和可靠性. 实验表明, WGAN 在评估真实数据与生成数据差异方面的改进, 使其在生成任务中展现出更强的性能和更广泛的适用性, 为 GAN 研究和应用提供了全新的方向.

1.4 自注意力机制

近年来, 自注意力机制在图像修复任务中得到了广泛应用, 展现出优越的性能, 能够有效捕捉图像中的关键特征, 增强对局部和全局特征的敏感性, 并建模长距离依赖关系. 其灵活性和可扩展性使其能够与卷积神经网络 (CNN) 和生成对抗网络 (GAN) 等模型结合, 推动图像处理技术的发展. 这种跨领域的融合不仅显

著提升了模型的表达能力, 还为应对复杂实际应用中的挑战提供了创新思路.

2017 年, Vaswani 等人^[15]提出的 Transformer 模型完全基于自注意力机制, 摒弃了传统的循环神经网络 (RNN) 和卷积神经网络 (CNN), 通过建模输入序列间的全局依赖, 提高了并行计算能力和训练效率; 2018 年, Yu 等人^[16]提出了一种基于自注意力的图像修复框架, 有效捕捉长距离依赖关系, 大幅提升了修复效果; 2020 年, Gulati 等人^[4]设计了结合卷积层与 Transformer 架构的“卷积增强变换器 (Conformer)”, 在语音信号处理中通过卷积提取局部特征并利用自注意力捕获全局上下文, 显著提升了语音识别任务的性能.

然而, 现有自注意力机制在处理云层图像时对边缘区域的重要信息关注不足, 限制了修复效果. 为此, 本文基于 Conformer 模块设计了新型的注意力权重张量, 构成 ConformerPlus 模块, 该模块继承了 Conformer 模块的优点, 并通过动态调整注意力分数, 增强对边缘区域的关注, 从而有效引导去云过程, 提升整体修复效果和生成图像质量.

2 模型结构

针对以上分析, 本文提出了一种融合时域动态关联与谱域物理特性的时序自注意力增强 GAN 框架 (TGAN), 以解决遥感图像中因云层遮挡导致的信息缺失问题并实现高精度图像修复. 模型整体架构如图 2 所示, 从图中可以看出, 该修复模型由两个阶段组成, 以生成对抗网络 (GAN) 为框架, 结合多模块协同建模实现云遮挡区域的精细修复, 具体如下.

第 1 阶段由 GAN 网络核心组件生成器构成, 生成器的核心结构由特征提取器、ConformerPlus 模块和云检测模块组成. 通过有效融合时域与谱域特征, 该阶段能够准确捕捉遥感图像的时序变化规律, 提取全局信息, 并初步生成完整图像, 重点解决因云层遮挡导致的信息丢失问题.

第 2 阶段基于第 1 阶段的训练成果, 采用自适应图像修复模块对生成的初步图像进行进一步优化. 自适应模块由随机噪声消除和局部对比度增强子模块构成, 分别针对噪声抑制与细节提升设计, 显著改善图像的细节质量与整体视觉一致性. 此外, 鉴别器中引入多尺度分析模块, 从全局结构与局部纹理两方面对修复图像的真实性进行严格判断, 确保生成图像在视觉质量和信息保真度上的卓越表现.

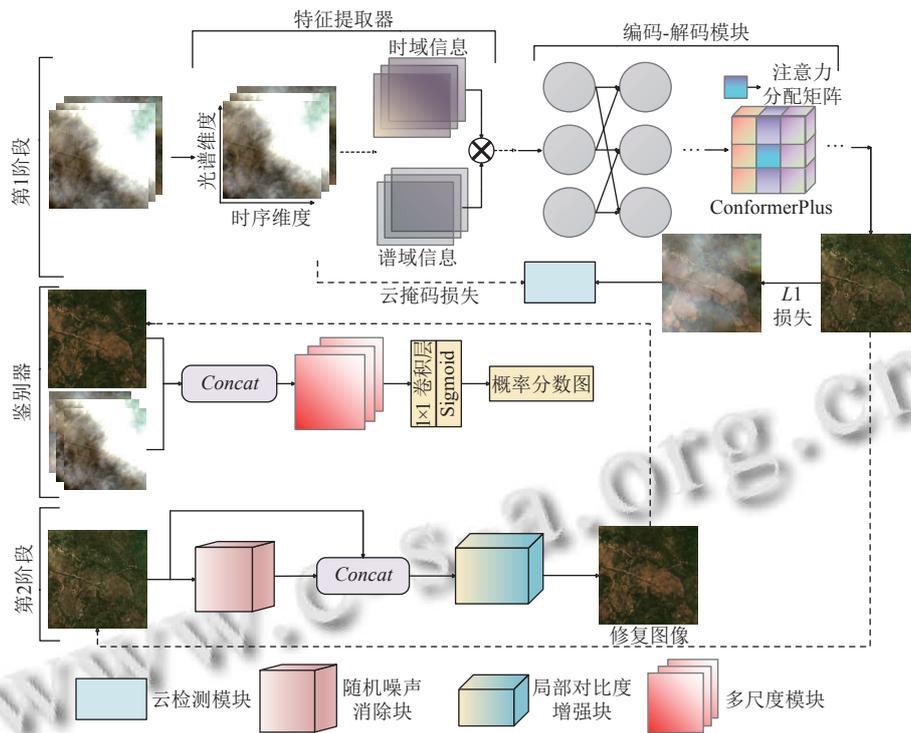


图2 TGAN 整体框架

以下部分将详细介绍模型的各个模块及其在遥感图像修复中的作用,并分析性能提升机制及优势。

2.1 特征提取器

自注意力机制的核心在于能对序列中每个位置的特征进行独立编码,同时捕获全局上下文信息,从而提升模型的特征表达能力,该机制在识别序列中的关键

特征方面具有较好的性能。

针对遥感图像数据的高复杂性和多维特征交互问题,本文采用基于时序自注意力的特征提取器.该提取器通过对图像中各位置的特征进行动态编码,全面整合全局语义信息与局部细节特征,以应对遥感场景中复杂地物分布及云层遮挡,设计架构如图3所示。

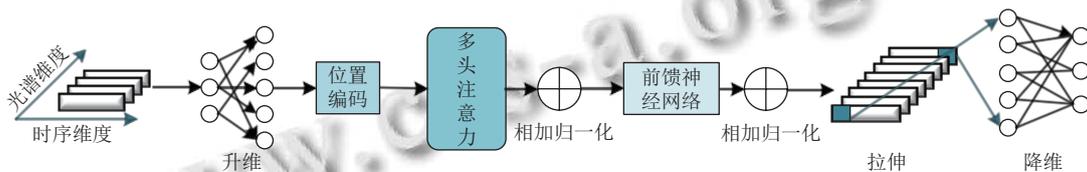


图3 特征提取器整体结构

图3 中所示的基于 Transformer 编码器的特征提取器旨在对遥感图像中的时域与光谱域特征进行上下文关系建模,可显著提升特征提取的精确性和鲁棒性.具体流程如下。

首先通过一维线性变换层对图像数据的时域与光谱特征进行扩展,以形成融合两者信息的综合特征向量。

随后,这些向量被输入包含单编码器的 Transformer 模型,其中通过残差连接和归一化操作增强特征的稳

定性及传递效率.多头注意力机制进一步捕捉特征之间的全局和局部依赖关系,而前馈神经网络则对高维特征表达进行精细化提炼.为了降低过拟合风险,编码过程中引入了 Dropout 操作,以提升模型的泛化能力。

最后,通过一维线性变换将高维特征投影回一维空间,生成3个代表性的时域-光谱特征向量,为后续处理步骤提供高质量的数据支持。

整体而言,该方法充分利用了 Transformer 在复杂场景的特征建模能力,实现了遥感图像多维信息提取

与融合任务.

2.2 生成器结构

生成器的内部架构借鉴了 STGAN^[11]的设计思路,并针对多时相遥感图像的复杂特征建模进行了改进.

我们对生成器中的特征提取器进行了深度定制,重点优化了多时相特征的下采样流程和注意力机制,以提升模型在捕捉全局与局部信息方面的表现.生成器的详细网络架构如图4所示.

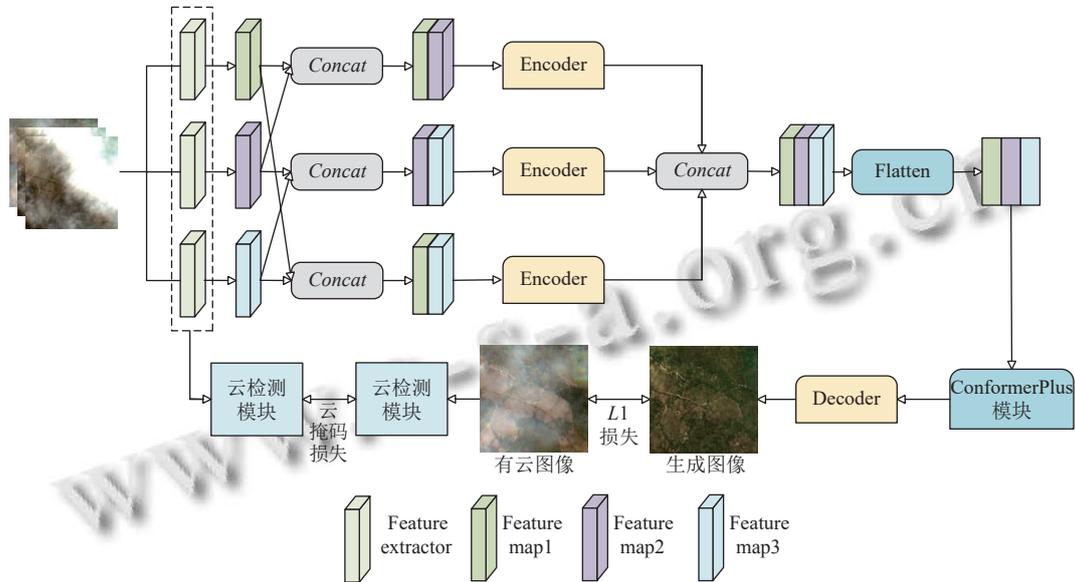


图4 生成器结构

从图4中可以看出,生成器的整体结构由特征提取器、ConformerPlus模块和云检测模块组成,形成一个高效协作的特征建模框架.特征提取器负责初步抽取和整合时域与光谱域信息,ConformerPlus模块进一步对多尺度特征进行增强以捕捉复杂的上下文关系,而云检测模块则专注于精准识别和处理云层遮挡区域.

生成器的输入由3张存在云层遮挡的遥感图像组成,定义为 $x_c \in \{x_1, x_2, x_3\}$.由于难以收集与缺失图像完全对应的真实无云图像,本研究假设地物在时间尺度上保持相对不变,并选取其中一张有云图像作为参考图像,记为 y .在此基础上,给定 x_c ,模型学习生成一张无云图像 x_g ,该图像 x_g 在未缺失区域与参考图像 y 保持高相似度.通过融合输入图像中未缺失部分的信息,模型逐步优化特征表达以实现高质量复原.具体而言,输入图像 x_c 首先通过特征提取模块TempSA进行处理,提取出3个特征表示 F_{im1} 、 F_{im2} 、 F_{im3} .这些特征表示在时域和空间域中建模了输入图像的局部和全局信息,为后续无云图像生成奠定了特征基础,即:

$$\begin{cases} F_{im1} = \text{TempSA}(x_1) \\ F_{im2} = \text{TempSA}(x_2) \\ F_{im3} = \text{TempSA}(x_3) \end{cases} \quad (1)$$

3个特征表示 F_{im1} 、 F_{im2} 、 F_{im3} 均被Concat层相互融合起来得到3个并列的特征表示 F_{im1} 、 F_{im2} 、 F_{im3} ,即:

$$\begin{cases} F_{im12} = \text{Concat}(F_{im1}, F_{im2}) \\ F_{im13} = \text{Concat}(F_{im1}, F_{im3}) \\ F_{im23} = \text{Concat}(F_{im2}, F_{im3}) \end{cases} \quad (2)$$

其中, F_{im1} 、 F_{im2} 、 F_{im3} 分别被传递到单个编码器管道,编码器分别对 F_{im1} 、 F_{im2} 、 F_{im3} 进行编码,并通过Concat层融合编码器输出的特征表示 F_{im123} ,对图像进行解码,输出无云的图像修复结果 x_g ,其中编码器和解码器分别是步幅为2的卷积层.

此外,生成器还集成了云检测模块,用于识别并标记输入图像中云层的分布位置.该模块通过分析图像特征,生成对应的云掩模(cloud mask),以明确标识云覆盖区域和潜在信息缺失或遮挡区域.基于生成的云掩模,进一步实现了源图像(包含云层的原始图像)与目标图像(云层被去除后的图像)之间的选择性对比,仅关注未被云层遮挡的区域.这种差异化对比策略不仅能够有效引导生成器优化模型训练,还提高了在无云区域的重建精度,从而增强了生成器在复杂场景下

的泛化能力。

尽管级联自注意力机制在捕捉长距离依赖特征方面优势显著,但遥感图像数据具备丰富的时间序列和光谱特性,导致该机制对局部特征信息的提取能力欠佳.尤其是在处理含云层的遥感图像时,传统注意力机制往往对边缘区域重要信息关注不足,进而限制了图像修复效果.为此,本文在 Conformer 模块的基础上设计一种新型多头自注意力机制 (NMHSA), 构建了改进的 ConformerPlus 模块. 该机制通过生成新的注意力权重张量, 以提高对边缘信息的捕捉能力, 进而引导模型实现从外向内的递进式云层去除, 优化整体修复效果. 该模块融合了卷积神经网络、NMHSA 与 Transformer 架构, 从而有效强化了关键特征的表达能力. 其具体实现如图 5 所示.

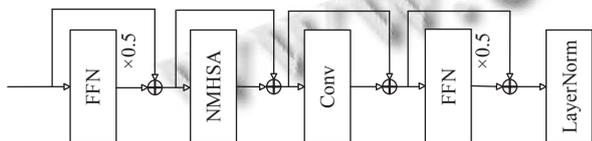


图 5 ConformerPlus 模块结构

从图 5 中可以看出, ConformerPlus 模型在 Conformer 架构^[5]的基础上通过整合 CNN 模块增强对序列中相对位置编码的局部特征提取能力. 此外, 该模型设计了 NMHSA, 相较于传统多头自注意力 (MHSA), 不仅能够更加高效地建模序列内元素间的长距离依赖关系, 还能优化对云层外围信息的关注, 弥补传统机制对关键边缘特征捕捉不足的问题. 在结构设计上, ConformerPlus 将 NMHSA 模块与卷积层嵌入到两个前馈神经网络 (FFN) 之间, 这种双 FFN 配置显著优于仅使用单个 FFN 的方式, 可有效提升模型的性能和鲁棒性. 前馈网

络包含两个线性变换和一个非线性激活函数, 全局特征提取网络由 L 个 ConformerPlus 组成, 对于第 i 个 ConformerPlus 模块的输入 y_i , 其中间变量和输出 h_i 可通过以下公式获得:

$$\begin{cases} y_i = y_i + \frac{1}{2}FFN(y_i) \\ y'_f = y_i + NMHSA(y_i) \\ y''_i = y'_f + Conv(y'_f) \\ h_i = LayerNorm\left(y''_i + \frac{1}{2}FFN(y''_i)\right) \end{cases} \quad (3)$$

其中, 输入张量 y_i 通过前馈神经网络 (FFN), 该网络包含两个全连接层、Swish 激活函数和 dropout 正则化, 以增强模型的表达能力. 同时, 采用新型多头自注意力机制 (NMHSA), 该机制结合了传统多头注意力与注意力自调整矩阵, 能够有效优化注意力分数. 此外, 卷积层 (Conv) 融合了门控线性单元 (GLU)、深度可分离卷积及其他激活和正则化技术, 以高效处理序列数据并提取特征. 最后, 使用 LayerNorm 进行输入数据的归一化处理, 从而提升模型的训练效果.

2.3 自适应图像修复模块

自适应图像修复模型核心由随机噪声消除模块和局部对比增强模块组成, 这两个模块均基于深层卷积神经网络架构, 并引入残差模块, 以强化特征提取和信息传递能力. 在模块间的设计中, 模型采用拼接 (Concat) 操作, 将输入图像与去噪后的图像特征融合. 这种融合策略不仅保留了去噪过程中的关键信息, 同时整合了原始图像中的细节特征, 从而提供更加丰富的上下文信息, 进一步优化了后续处理阶段的表现. 整体网络结构如图 6 所示.

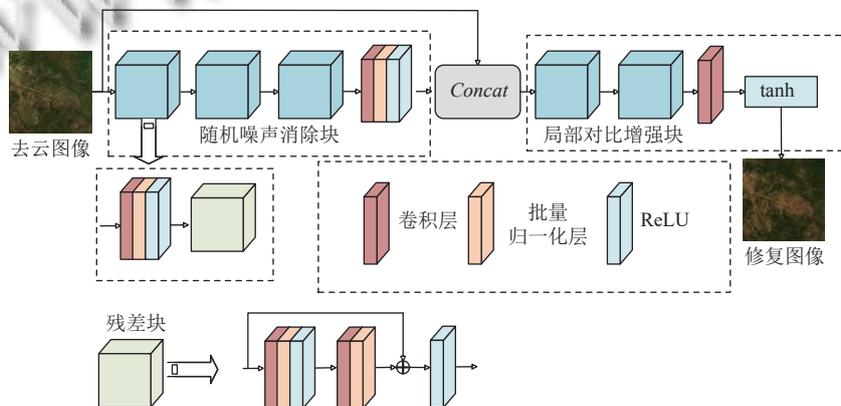


图 6 自适应模块结构

从图6中可以看出,自适应图像修复模块由随机噪声消除模块和局部对比增强模块构成.这种分离设计提高了系统的适应性和可扩展性,能够根据不同图像特性和噪声分布进行优化.随机噪声消除模块通过深度卷积网络、批量归一化层和残差块,增强了对噪声的感知与去除能力,缓解了深层网络的梯度消失问题,提升了训练稳定性和收敛效率.同时,逐步变化的通道数策略捕捉多尺度特征,实现对不同类型噪声的高效去除.局部对比增强模块则专注于提取高级特征,保留原始细节,避免过度增强导致的失真.该模块通过卷积网络提取特征,并利用 tanh 激活函数进行归一化,确保增强后的图像不过饱和或过暗.

在整体处理流程中,自适应图像修复模块首先执行去噪操作,将去噪后的图像与原始输入图像进行特征拼接后输入增强层.此流程不仅充分利用了去噪信息提升增强效果的质量与稳定性,还通过融合原始图像特征,使增强过程能够依据图像的固有特性进行自适应调整,最终实现更加自然且高效的增强效果.这种设计思路突出了对图像修复任务的全局与局部平衡的关注,同时增强了模型的鲁棒性和实用性.

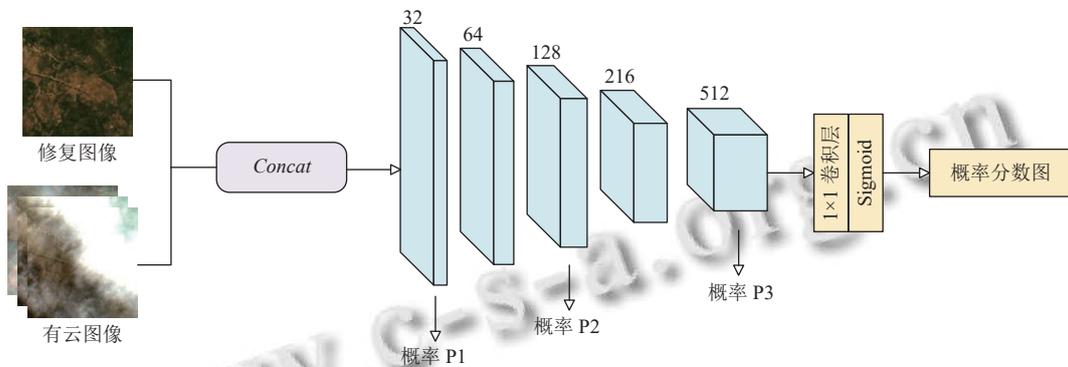


图7 多尺度模块结构

在特征提取过程中,所有卷积操作均采用 4×4 的卷积核,步长设定为2,并结合 LeakyReLU 激活函数引入非线性特性.特征图的维度采用 $H \times W \times C$ (图像高度 \times 图像宽度 \times 通道数)表示.通过这种设计,鉴别器能够高效提取不同尺度上的图像特征,全面评估输入图像的真实性,为生成器优化提供指导.

在获得 P1、P2 和 P3 后,数据进一步通过最后一层 1×1 的特征图处理,步长设定为1,并结合 Sigmoid 激活函数生成最终判别概率 P4.鉴别器将生成图像划分为多个独立的图像块,假设每个图像块内的像素相互

2.4 多尺度模块鉴别器

鉴别器在生成对抗网络(GAN)架构中扮演着至关重要的角色,其主要功能是通过多尺度感受野来捕捉图像的全局特征与局部细节.本文引入了多尺度模块鉴别器(multiscale module discriminator,简称D),其架构如图7所示,以构建最终的TGAN模型.该模块通过对全局图像和局部区域进行独立的真假判定,显著增强了生成器在全局信息与局部细节之间的协同表达能力.相较于传统鉴别器,该方法能够更为精准地评估图像的整体一致性与局部差异性,从而有效提升生成图像的质量.

D通过一系列特征图的处理,从全局到局部逐步提取特征并生成分层判别概率.首先,数据经过尺寸为 $256 \times 256 \times 32$ 的特征图处理,输出初始判别概率 P1.随后,数据依次通过 $128 \times 128 \times 64$ 和 $64 \times 64 \times 128$ 的特征图,生成第2层判别概率 P2.接着,特征图进一步压缩至 $32 \times 32 \times 256$ 和 $16 \times 16 \times 512$,最终输出第3层判别概率 P3.这种多尺度特征提取机制能全面捕捉输入数据的全局信息与局部细节,提高鉴别器的鲁棒性和判别精度.

独立,单独进行真假判定.最终,通过对所有图像块的判定结果取平均值,形成整幅图像的整体真假评估结果.

这种模块化设计不仅能够精细化评估图像局部真实性,还兼具对全局信息的综合分析能力,为生成器的优化提供了更加准确的监督信号.

2.5 损失函数设计

损失函数包括3部分,条件GAN损失函数、标准L1损失函数、云掩模损失函数,定义为:

$$L = \min_{GP} \max_D L_{cGAN}(GP, D) + \lambda_{GP} L_1(GP) + L_{mask} \quad (4)$$

其中, D 表示模型的鉴别器, 参数 G 和 P 分别表示生成器和自适应图像修复块。

其中, 第 1 部分为条件 GAN 损失函数, 定义为:

$$L_{\text{cGAN}}(GP, D) = E_{(g,y)} [\log D(x_g, y)] + E_{(x_c)} [\log (1 - D(y, GP(x_c)))] \quad (5)$$

第 2 部分为标准 L_1 损失函数, 定义为:

$$L_1(GP) = \frac{1}{CWH} \sum \|y^{c,w,h} - GP(x_c)^{c,w,h}\| \quad (6)$$

其中, $GP(x_c)^{c,w,h}$ 表示生成的输出图像在坐标 (c, w, h) 处的像素。

第 3 部分是云掩模损失函数, 定义为:

$$L_{\text{mask}} = \|M - M'\|_2^2 \quad (7)$$

其中, M 和 M' 分别为真实有云掩模和生成云掩模。

3 实验结果及分析

3.1 实验数据集

为验证本文提出的 TGAN 云去除模型的有效性, 将其应用于真实的 Sen2-MTC 云数据集。该数据集来自哥白尼数据空间生态系统 (Copernicus Data Space Ecosystem, Europe's Eyes on Earth), 包含 50 个不重复的类别, 每个类别包括 70 张图像, 包含 50 个不重复的类别, 每个类别含有 70 张图像, 共计 3500 张图像。每张图像的像素值范围为 $[0, 10000]$, 分辨率为 256×256 , 且具有 4 个光谱通道: 红色 (R)、绿色 (G)、蓝色 (B) 以及近红外 (NIR)。该数据集涵盖了不同地理区域和天气条件, 包含丰富的云层类型和变化特征, 具备较强的代表性和多样性, 非常适合用于遥感图像中的云去除任务。

为了确保模型的泛化能力和评估的客观性, 本文将数据集随机划分为训练集、验证集和测试集, 划分比例为 7:1:2。具体而言, 70 张图像中的 70% 用于训练, 10% 用于验证, 剩余 20% 用于测试。为了避免数据划分时的类别偏差, 来自同一类别的图像被分配到相同的子集, 确保每个子集内的图像在类别分布上的均匀性, 这种划分策略有助于保证模型能够在不同的数据集上进行有效的训练和评估, 从而提高在实际应用中的鲁棒性和可靠性。

同时在实际的实验中通过多种数据增广技术 (如旋转、翻转等) 扩展了训练数据集的规模, 从而进一步提升模型的泛化能力和稳定性。

3.2 实验环境

本文的实现基于 PyTorch 深度学习框架, 实验在配备 24 GB 显存的 NVIDIA GeForce RTX 3090 GPU 服务器上进行, 为高分辨率遥感图像处理提供了充足的计算资源。模型训练过程中采用 Adam 优化器, 通过自适应调整学习率加速模型收敛, 并在两个动量参数 ($\beta_1=0.5$ 和 $\beta_2=0.999$) 的控制下平衡梯度更新。批次大小 (batch size) 设定为 4, 以确保训练稳定性并有效利用 GPU 内存。初始学习率为 $5E-4$, 经过 100 个训练回合后采用线性衰减策略逐步降低至 $1E-6$, 以实现更细致的优化并防止后期过拟合。整个训练过程旨在平衡模型的训练速度与精度, 提升其在云去除任务中的表现。

由图 8 可知, 通过训练模型, 损失函数逐步收敛, 当学习率 lr 为 0.00001 时, 损失值最小, 网络模型训练收敛得最好, 因此实验具体参数如表 1 所示。

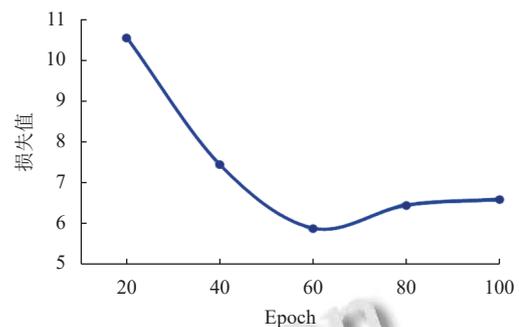


图 8 损失函数值

表 1 实验环境参数

名称	参数
GPU	NVIDIA RTX 3090 24 GB
Batch size	4
训练回合	100
动量参数	$\beta_1=0.5$ 、 $\beta_2=0.999$
学习率	$5E-4$ 至 $1E-6$
优化器	Adam

3.3 评价指标

本文采用峰值信噪比 (PSNR) 和结构相似性指数 (SSIM) 作为评估指标, 以全面衡量所提薄云去除方法的性能。PSNR 越高表示图像信号与噪声比率越好, 是经典的图像质量评价指标; 而 SSIM 则从亮度、对比度和结构等多个方面衡量图像相似度, 广泛应用于神经网络生成图像及深度学习图像处理任务^[17]。通过这两个指标, 研究能够对网络生成图像进行全面和客观的评价, 为模型优化和图像生成技术的进一步改进提供

理论支持和实验依据。

PSNR 计算最大信号和背景噪声之间的比值,是衡量图像质量的质量指标,定义为:

$$PSNR = 10 \times \log_{10} \left(\frac{(2^n - 1)^2}{MSE} \right) \quad (8)$$

其中, n 为每个采样值的比特数; MSE 为均方误差值; $PSNR$ 值越大,图像去云的效果越好。

SSIM 是用来衡量两幅图像相似度的指标,结构相似度从亮度、对比度、结构 3 个方面进行度量,均值作为亮度的估计,标准差作为对比度的估计,协方差作为结构相似程度的度量,定义为:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C)(2\sigma_{xy} + C)}{(\mu_x^2 + \mu_y^2 + C)(\sigma_x^2 + \sigma_y^2 + C)} \quad (9)$$

其中, μ_x 和 μ_y 分别为 x 、 y 的平均值; σ_x 和 σ_y 分别为 x 、 y 的标准差; σ_{xy} 为 x 、 y 的协方差; C 为一个较小的常数,避免分母为 0 导致计算错误。

3.4 实验结果

为验证所提出的 TGAN 模型在遥感图像数据修复任务中的性能,本研究基于 Sen2-MTC 数据集进行了全面实验并将实验结果分别与 TSGAN、STGAN^[11]、ST_net^[18]、AE^[19]和 TGAN 模型结果进行了对比分析,结果如表 2 和表 3 所示。

表 2 在 Sen2_MTC 数据集上 *PSNR* 指标性能 (dB)

模型	验证集	测试集
AE	6.010	15.251
ST_net	17.741	16.206
STGAN	20.612	18.152
TSGAN	21.259	18.308
TGAN	21.547	20.206

表 3 在 Sen2_MTC 数据集上 *SSIM* 指标性能

模型	验证集	测试集
AE	0.431	0.412
ST_net	0.467	0.427
STGAN	0.613	0.587
TSGAN	0.662	0.609
TGAN	0.621	0.576

实验结果表明, TGAN 模型在峰值信噪比 (*PSNR*) 指标上的表现显著优于其他模型。具体而言,目前最先进的 TSGAN 模型在验证集和测试集上的 *SSIM* 分别为 0.662 和 0.609, *PSNR* 分别为 21.259 dB 和 18.308 dB,而本文的 TGAN 模型在 *PSNR* 指标上实现了验证集和测试集分别提升 0.288 dB 和 1.898 dB 的显著增益。这

表明, TGAN 模型在还原图像整体质量和减少噪声干扰方面具备更高的鲁棒性。虽然 TGAN 的结构相似性指数 (*SSIM*) 略有下降,在一定程度上表明生成图像与参考图像的相似度减小,但主观视觉评估结果显示, TGAN 在生成图像的细节保真度和全局对比度方面表现出更高的质量。

图 9 展示了 TGAN 和 TSGAN 在图像去云任务上的可视化对比,从上至下依次为 3 张多云缺失图像、TSGAN 修复图像、TGAN 修复图像及对应的无云参考图像。从图 9 中可以观察到,经 TGAN 修复后的缺失图像在对比度、纹理细节以及整体视觉感官上更接近于真实图像,而 TSGAN 修复图像在某些细节和对比度方面略显不足。以上结果表明, TGAN 模型不仅在定量指标上表现优越,在主观感知质量上也展现出更强优势,为高质量遥感图像修复提供了可靠的解决方案。

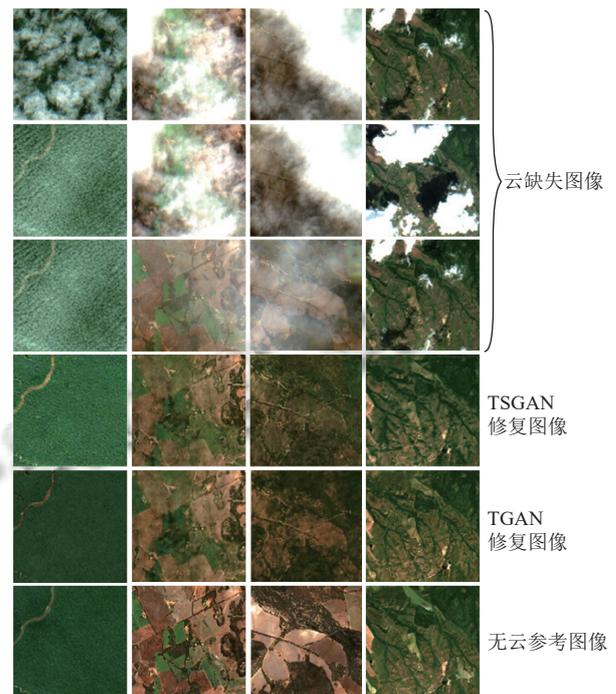


图 9 图像去云效果

3.5 消融实验

以 Sen2-MTC 数据集为基础,利用 TSGAN 网络进行多组消融实验研究。通过在 TSGAN 网络分别加入自适应图像修复模块 P (TSGAN+ P) 和 Conformer 模块中加入注意力分数调整矩阵 M (TSGAN+ M) 这两个关键模块,和在 TSGAN 网络中同时加入 P 和 M 模块 (TSGAN+ P + M),对模型进行消融分析,运用 *SSIM* 和

PSNR 评价指标测试不同模块下模型的有效性。研究结果表明, 加入自适应图像修复模块 P 和在 Conformer 模块中加入注意力分数调整矩阵 M 可以有效提高模型的表现力和性能。同时对损失函数进行消融实验, 通过分别去除条件 GAN 损失函数 G 、标准 $L1$ 损失函数 L 和云掩模损失函数 Y , 以对损失函数对模型性能影响进行定性分析。表 4 中为消融实验的评价指标结果。

表 4 消融实验结果

方法	<i>PSNR</i> (dB)	<i>SSIM</i>
TSGAN	21.259	0.662
TSGAN+ M	21.333	0.620
TSGAN+ P	21.411	0.603
TGAN (TSGAN+ P + N)	21.547	0.621
TGAN (G + L)	21.407	0.606
TGAN (G + Y)	18.887	0.502
TGAN (L + Y)	21.328	0.618
TGAN (G + L + Y)	21.547	0.621

从表 4 中可知, 相比于原始 TSGAN 模型, 对于自适应图像修复模块 P , *PSNR* 提高了 0.152 dB, 可见通过自适应图像修复模块 P , 能够有效地对图像进行去噪和细节增强。对于在 Conformer 模块中加入注意力分数调整矩阵 M , *PSNR* 提高了 0.074 dB, 通过引入新型权重分配策略动态调整注意力分数, 能够增强模型对边缘信息的捕捉能力, 以提高修复图像的质量。对于同时加入 P 和 M 的 TGAN 模型 *PSNR* 提高 0.288 dB, 且主观视觉评估方面显著优化, 这证明了引入这些新模块有助于提高模型的精度。

从表 4 可知, 在 TGAN 网络模型中, 分别采用不同的损失函数组合进行实验。使用生成损失函数 G 和重建损失函数 L (G + L) 时, *PSNR* 和 *SSIM* 分别为 21.407 dB 和 0.606; 使用生成损失函数 G 和约束损失函数 Y (G + Y) 时, *PSNR* 和 *SSIM* 分别为 18.887 dB 和 0.502; 使用重建损失函数 L 和约束损失函数 Y (L + Y) 时, *PSNR* 和 *SSIM* 分别为 21.328 dB 和 0.618; 而联合使用 3 种损失函数 (G + L + Y) 时, *PSNR* 和 *SSIM* 分别达到了 21.547 dB 和 0.621。实验结果表明, 与单独使用部分损失函数相比, 联合 3 种损失函数能够有效提升模型性能。具体而言, 使用 3 种损失函数的组合能够优化模型参数, 显著提高云缺失图像修复效果, 进而提升 *PSNR* 和 *SSIM* 值, 验证了这一组合对改善修复图像质量的有效。

4 结论

遥感图像数据具有丰富的光谱信息和时序特性,

但其高维复杂性在深度学习处理中常导致噪声引入和细节信息的丢失。针对上述问题, 本文提出了一种基于时谱域融合与时序自注意力增强的无监督遥感图像修复模型 (TGAN), 该模型在架构设计和特征提取机制上进行改进, 以提升遥感图像修复的质量和效率。TGAN 模型由两个阶段组成, 以模块化设计策略分别解决特征提取与图像修复中的关键问题。第 1 阶段引入了基于时序自注意力机制的特征提取器和新型多头自注意力机制, 用于生成对数据的全局与局部特征的高效表示。在该阶段, 特征提取器引入一维线性升维层以全面捕获时域和谱域特征, 并用一维线性降维层替代传统最大池化操作, 从而更精确地建模数据在不同时间序位置上的重要性。新型多头自注意力机制中引入新型权重分配策略动态调整注意力分数, 以增强对边缘信息的捕捉能力。第 2 阶段通过自适应图像修复模块进一步优化生成的图像, 模块由随机噪声消除和局部对比增强子模块构成, 以提升图像细节质量并降低噪声干扰。此外, 模型的鉴别器采用多尺度模块以同时兼顾图像修复的全局一致性和局部细节精度, 从而有效解决传统方法在平衡全局与局部信息方面的不足。在公开的遥感图像云去除数据集上的实验表明, TGAN 模型在 *PSNR* 和 *SSIM* 指标上均优于当前主流的无监督去云方法, 验证了其在遥感图像修复任务中的有效性和优越性。

参考文献

- 1 Yu B, Yin HT, Zhu ZX. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecast. Proceedings of the 27th International Joint Conference on Artificial Intelligence. Stockholm: IJCAI, 2017. 3634–3640.
- 2 Wang ZQ, Cornell S, Choi S, et al. TF-GRIDNET: Making time-frequency domain models great again for monaural speaker separation. Proceedings of the 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Rhodes Island: IEEE, 2023. 1–5.
- 3 Zhang CY, Wen WP, Zhai CH, et al. Structural nonlinear seismic time-history response prediction of urban-scale reinforced concrete frames based on deep learning. Engineering Structures, 2024, 317: 118702.
- 4 Gulati A, Qin J, Chiu CC, et al. Conformer: Convolution-augmented Transformer for speech recognition. Proceedings of the 21st Annual Conference of the International Speech Communication Association. Shanghai: Interspeech, 2020.

- 5036–5040.
- 5 Goodfellow IJ, Pouget-Abadie J, Mirza M, *et al.* Generative adversarial nets. Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal: MIT Press, 2014. 2672–2680.
 - 6 Zhang H, Goodfellow I, Metaxas D, *et al.* Self-attention generative adversarial networks. Proceedings of the 36th International Conference on Machine Learning. Long Beach: PMLR, 2019. 7354–7363.
 - 7 周卓浩. 基于半监督深度学习的带钢表面缺陷分类方法研究 [硕士学位论文]. 太原: 中北大学, 2023.
 - 8 Li MJ, Wang GH, Zhu MH, *et al.* DFAMNet: Dual fusion attention multi-modal network for semantic segmentation on LiDAR point clouds. Applied Intelligence, 2024, 54(4): 3169–3180.
 - 9 Xu ZX, Wu K, Wang WL, *et al.* Semi-supervised thin cloud removal with mutually beneficial guides. ISPRS Journal of Photogrammetry and Remote Sensing, 2022, 192: 327–343. [doi: [10.1016/j.isprsjprs.2022.08.026](https://doi.org/10.1016/j.isprsjprs.2022.08.026)]
 - 10 Wang JL, Zhao XL, Li HC, *et al.* Unsupervised domain factorization network for thick cloud removal of multitemporal remotely sensed images. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 5405912.
 - 11 Shen HF, Li XH, Cheng Q, *et al.* Missing information reconstruction of remote sensing data: A technical review. IEEE Geoscience and Remote Sensing Magazine, 2015, 3(3): 61–85.
 - 12 Yu GY, Wu P, Lv Z, *et al.* Few-shot fault diagnosis method of rotating machinery using novel MCGM based CNN. IEEE Transactions on Industrial Informatics, 2023, 19(11): 10944–10955.
 - 13 Jia CF, Wei PT, Lu ZH, *et al.* A novel prediction approach of polymer gear contact fatigue based on a WGAN-XGBoost model. Fatigue & Fracture of Engineering Materials & Structures, 2023, 46(6): 2272–2283.
 - 14 高闻霏, 杨波, 闵琳孜. 低分辨率红外感知系统的人体行为数据增强. 仪器仪表学报, 2023, 44(1): 163–171.
 - 15 Vaswani A, Shazeer N, Parmar N, *et al.* Attention is all you need. Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017. 6000–6010.
 - 16 Yu JH, Lin Z, Yang JM, *et al.* Generative image inpainting with contextual attention. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 5505–5514.
 - 17 肖昌城, 吴锡, 何妍. 基于稠密残差网络的多序列卫星图像去云. 计算机应用研究, 2022, 39(1): 303–307.
 - 18 Chen Y, Weng QH, Tang LL, *et al.* Thick clouds removing from multitemporal Landsat images using spatiotemporal neural networks. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 4400214.
 - 19 Sintarasirikulchai W, Kasetkasem T, Isshiki T, *et al.* A multi-temporal convolutional autoencoder neural network for cloud removal in remote sensing images. Proceedings of the 15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON). Chiang Rai: IEEE, 2018. 360–363.

(校对责编: 张重毅)