

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(XXXX)XX-0001-17

论文引用格式: Zhang Wenbo, Zhang Guangyang, Huang Yuanwan, Huang Weiquan, Li Xiaohui, Zhou Zhilin, Yang Pinglv, Zhou Zeming, Chen Qiang. Construction and evaluation of a dataset for completing missing nighttime satellite band information[J/OL]. Journal of Image and Graphics, XXXX:1-17. DOI: 10.11834/jig.260366. (张文博, 张光阳, 黄渊万, 黄伟权, 李孝辉, 周稚麟, 杨平吕, 周则明, 陈强. 夜间卫星缺失波段信息补全数据集构建与评估[J/OL]. 中国图象图形学报, XXXX:1-17. DOI: 10.11834/jig.260366.) [DOI: 10.11834/jig.260366]

夜间卫星缺失波段信息补全数据集构建与评估

张文博¹, 张光阳², 黄渊万¹, 黄伟权¹, 李孝辉¹, 周稚麟¹, 杨平吕², 周则明², 陈强^{1*}

1. 南京理工大学计算机科学与工程学院, 南京 210094; 2. 国防科技大学气象海洋学院, 长沙 410083

摘要: 目的 针对夜间缺乏太阳短波辐射导致气象卫星可见光与近红外(visible and near-infrared, VIS-NIR)波段观测缺失、相关遥感产品难以连续获取的问题,以构建面向夜间缺失波段信息补全的大规模数据集为主要目标,同时解决相关研究生成光谱波段覆盖不足与公开可用的夜间缺失波段信息补全数据集稀缺问题。方法 基于Himawari-8/9多波段观测数据,构建包含多波段红外亮温观测(multi-band infrared brightness temperature, MIBT)、VIS-NIR、观测几何角度和水陆掩膜信息的VIS-NIR-70K数据集,并选取代表性生成模型进行定量与定性评估。结果 实验表明,MIBT能够为补全VIS-NIR缺失信息提供有效条件。扩散模型的生成性能优于基于生成对抗网络(generative adversarial network, GAN)的模型,其中,原始图像空间扩散模型在结构保持、纹理细节和分布一致性方面表现更优,多波段VIS-NIR的平均结构相似性指数(structural similarity index measure, SSIM)提升至0.864。同时,其在真彩色(true-color imagery, TCI)观测生成、低云区域、台风场景和夜间全图补全中展现出较好的重建性能,TCI波段的平均SSIM提升至0.881。结论 VIS-NIR-70K缓解了夜间卫星缺失波段信息补全领域的资源不足,为全天候天气系统观测和生成式遥感研究提供了数据基础。完整的MIBT多波段是VIS-NIR生成质量的关键信息来源,卫星天顶角(satellite zenith angle, SAZ)与水陆掩膜信息(land-sea mask, Landsea)组合可以提高生成结果的稳定性。数据集链接: (ScienceDB: <https://doi.org/10.57760/sciencedb.j00240.00227>)。

关键词: 多波段卫星图像; 缺失波段补全; 多波段数据集; VIS-NIR生成; 生成式算法; Himawari-8/9

Construction and evaluation of a dataset for completing missing nighttime satellite band information

Zhang Wenbo¹, Zhang Guangyang², Huang Yuanwan¹, Huang Weiquan¹, Li Xiaohui¹, Zhou Zhilin¹, Yang Pinglv², Zhou Zeming², Chen Qiang^{1*}

1. School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China; 2. Meteorology and Oceanography College of National University of Defense Technology, Changsha 410083, China

Abstract: Objective Geostationary meteorological satellites provide observations of clouds, the atmosphere, and Earth's surface, but their visible and near-infrared (VIS-NIR) reflectance channels cannot provide usable measurements at night because reflected solar shortwave radiation is unavailable. Consequently, true-color imagery (TCI) and other products

收稿日期: 2026-06-25; 修回日期: 2026-10-08

* 通信作者: 陈强 chen2qiang@njust.edu.cn

基金项目: 国家自然科学基金重点研发计划(项目编号: 92370109); 国家自然科学基金(项目编号: 42305159; 62172223; 62672244); 江苏省自然科学基金(项目编号: BK20252033)

Supported by: the Major Research Plan of the National Natural Science Foundation of China (Grant No. 92370109), National Natural Science Foundation of China (Grant No. 42305159; 62172223; 62672244), and Natural Science Foundation of Jiangsu Province (Grant No. BK20252033)

©中国图象图形学报版权所有

derived from these channels have gaps that hinder continuous monitoring of clouds, fog, and severe weather. Missing reflectance observations also affect products describing aerosols, cloud type, cloud optical thickness, and shortwave radiation, extending the importance of reconstruction beyond visual interpretation. Existing nighttime image-generation studies have generally targeted only one or a few visible bands, while public paired datasets spanning multiple reflective bands and diverse scenes remain limited. This study develops a large-scale benchmark, VIS-NIR-70K, for learning the relationship between infrared observations and missing VIS-NIR information. It also evaluates whether representative generative models can reconstruct six reflective bands from inputs available at night, thereby establishing a basis for comparing cross-modal generation methods and examining their potential for nighttime satellite applications. **Methods** VIS-NIR-70K was constructed from the Advanced Himawari Imager aboard the Himawari-8/9 satellites. Daytime observations provide simultaneous multi-band infrared brightness temperatures (MIBT) and measured VIS-NIR reflectances, which form paired training and evaluation samples even though nighttime reflectance ground truth is unavailable. The six prediction targets are AHI albedo bands 1-6, spanning visible, near-infrared, and shortwave-infrared wavelengths. Conditional inputs consist of brightness-temperature bands 7-16; one mid-wave infrared band, three water-vapor bands, and six additional thermal infrared bands. Satellite zenith angle (SAZ) and a land-sea mask (Landsea) supply observation-geometry and underlying-surface information. Solar zenith angle (SOZ) is used during preprocessing to exclude poorly illuminated regions with SOZ greater than 85° , rather than serving as a model input. Band-specific extreme values are clipped using the first and 99th percentiles, and coregistered observations are divided into nonoverlapping 512×512 -pixel patches; 256×256 -pixel versions are used for model training and inference. Samples are separated by more than one hour. The training set contains 62,290 paired patches derived from 690 observation times in 2016~2024, whereas the independent test set contains 5,446 pairs from 80 observation times in December 2025~January 2026. In total, 67,736 paired patches yield 1,219,248 individual multi-band images. Cloudy patches account for approximately 23% of the dataset, and both land and ocean are represented. The year-separated split reduces leakage from temporally adjacent scenes, although the test period is concentrated in winter. A linear baseline, CycleGAN, Pix2Pix, three DiT variants, a DDPM, and RefDiff are evaluated under a shared data split. Assessment combines pixel and radiometric errors, including mean squared error (MSE), root mean squared error, and mean bias error, with peak signal-to-noise ratio, structural similarity index measure (SSIM), correlation coefficient, spectral angle mapper, Fréchet Inception Distance, and Kernel Inception Distance. Visual comparisons examine individual bands, TCI, clouds, sea fog, a typhoon, and nighttime full-disk reconstructions. **Results** The paired-data experiments show that MIBT contains useful information for estimating VIS-NIR reflectance, although performance varies across target bands and scene types. Diffusion-based approaches generally outperform the adversarial models on structural similarity and perceptual distribution measures. The unpaired CycleGAN has difficulty preserving local structures, while paired Pix2Pix improves structural fidelity but tends to smooth complex cloud textures. DiT variants represent large-scale patterns and offer faster inference than the image-space diffusion approaches, yet their results retain less fine detail; increasing DiT model size produces only modest changes in reconstruction quality. DDPM and RefDiff, which perform diffusion in the original image space, recover cloud boundaries, spatial patterns, and subtle brightness variation more effectively. Across all six predicted bands, DDPM attains an average SSIM of 0.864, an MSE of 0.003, an FID of 11.178, and a KID of 0.007. RefDiff achieves a slightly higher average peak signal-to-noise ratio of 29.143 and a lower spectral angle mapper value of 0.0579, illustrating that no single model leads on every criterion. These differences show why pixel accuracy, feature-space similarity, and multi-band spectral agreement should be considered together when comparing generators for remote sensing. For TCI assembled from the three visible channels, DDPM obtains an average SSIM of 0.881 and an MSE of 0.002. Qualitative results show that the original-image-space diffusion models better preserve low-cloud textures and the organization of the typhoon-eye region. Experiments across cloudy and clear skies, land, ocean, and sea-fog scenes also indicate broad scene coverage. Nevertheless, albedo bands 4 and 5 remain harder to reconstruct than the visible channels, reflecting information about surface reflectance and cloud microphysics that infrared brightness temperatures do not fully determine. Input ablations identify the complete set of bands 7~16 as the main source of performance improvement; adding SAZ and Landsea together provides complementary constraints and improves overall stability. A nighttime full-disk example illustrates spatially continuous generated bands and a plausible TCI across the day-night boundary. An

approximate scaling calculation suggests that DDPM could generate a full disk in about 57 seconds, within AHI's ten-minute observation interval, although this is not a measured deployment time. Because concurrent nighttime VIS-NIR measurements are unavailable, this example supports qualitative interpretation rather than direct quantitative validation. **Conclusion** VIS-NIR-70K provides a multi-band, multi-scene resource for evaluating nighttime missing-band completion using daytime paired observations. The benchmark reveals the value of complete infrared inputs and the complementary roles of observation geometry and surface-type information. Original-image-space diffusion yields strong structural and spectral reconstruction, especially for complex clouds, while latent-space approaches offer a useful quality-efficiency tradeoff. The generated nighttime imagery could support continuous meteorological analysis, but daytime-to-nighttime domain shifts, limited independent seasonal testing, and the lack of nighttime reflectance ground truth constrain present claims about operational accuracy. Future work should incorporate physical knowledge of radiation and clouds, develop task-specific evaluation measures, and reduce diffusion sampling costs to improve the reliability and practicality of multi-band generation. Dataset access link: (ScienceDB: <https://doi.org/10.57760/sciencedb.j00240.00227>).

Key words: multiband satellite imagery; missing-band completion; multiband dataset; VIS-NIR generation; generative algorithms; Himawari-8/9

论文引用格式: Zhang Wenbo, Zhang Guangyang, Huang Yuanwan, Huang Weiquan, Li Xiaohui, Zhou Zhilin, Yang Pinglv, Zhou Zeming, Chen Qiang. Construction and evaluation of a dataset for completing missing nighttime satellite band information [J/OL]. Journal of Image and Graphics, XXXX: X-XX. DOI: 10.11834/jig.260366. (张文博, 张光阳, 黄渊万, 黄伟权, 李孝辉, 周稚麟, 杨平吕, 周则明, 陈强. 夜间卫星缺失波段信息补全数据集构建与评估[J/OL]. 中国图象图形学报, XXXX: X-XX. DOI: 10.11834/jig.260366.) [DOI: 10.11834/jig.260366.]

0 引言

地表与大气目标在不同波长电磁辐射作用下具有差异显著的光谱反射与辐射特性(Hong等, 2026)。这一特性为卫星遥感识别地表覆盖类型、云系结构以及大气参数反演提供了物理基础(Xu等, 2025)。快速发展的气象卫星遥感技术已经成为全天候监测地球系统的重要手段,在大气参数反演、灾害性天气监测和数值天气预报中发挥着重要作用(Wang等, 2025, 张本硕等, 2026)。气象卫星搭载多个波段的成像载荷,其中可见光(visible, VIS)与近红外(near-infrared, NIR)波段对云顶纹理、云体边缘、低云、海雾及部分强对流系统的精细结构具有更高的刻画能力(Liu等, 2026, 赵忠明等, 2019)。重要的是, VIS波段包含的红、绿、蓝通道合成的真

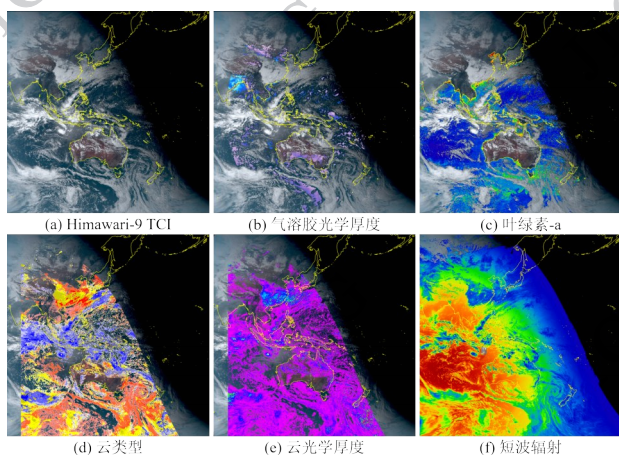
彩色图像(true-color imagery, TCI)不仅符合人眼视觉认知,也能够为天气分析和目标识别提供更加直观的观测依据(Zeng等, 2026)。

由于夜间缺乏太阳短波辐射,气象卫星在VIS-NIR波段难以获得有效的反射信息,导致依赖VIS-NIR的观测产品和反演参数在夜间无法连续提供(Zhang等, 2025c, Kim等, 2025)。图1以Himawari-9号卫星(Bessho等, 2016)在2026年1月1日7时采集的数据为例,展示了依赖VIS-NIR波段的相关气象产品在夜间部分内容缺失的示例图。这种夜间观测的缺失会限制卫星对云雾演变与灾害性天气的连续监测,同时也会降低全天候遥感分析的应用价值。因此,研究如何利用夜间可以稳定获取的多波段红外亮温(multi-bands infrared brightness temperature, MIBT)观测重建对应时刻的VIS-NIR缺失波段信息具有重要的科学意义与实际应用价值。

由于传统图像补全算法主要利用领域相似性完成对单幅图像中局部缺失像素的补全,而对于夜间气象卫星观测而言,利用夜间稳定获取的MIBT观测数据生成对应大范围缺失的VIS-NIR信息,本质上是一种具有物理约束的跨模态条件生成问题。将生成式人工智能方法引入夜间缺失波段信息补全,不仅有助于突破传统观测手段在夜间VIS-NIR波段获取上的物理限制,还可以显著提升气象卫星的全天时连续观测能力,也为多源遥感产品协同生成提供了新的技术路线和解决方案。

现有相关研究可以基于夜间稳定获取的MIBT观测生成不可直接观测的单波段或VIS图像。Kim

等人(2019)提出了条件生成对抗网络,利用单一长波红外图像重建夜间可见光图像。在此基础上, Kim 等人(2020)进一步提出多波段条件生成对抗网络,克服了单一红外波段在表征流云和台风细节结构方面的不足。Han 等人(2022)提出一种条件生成对抗网络方法,实现由 MIBT 到 VIS 波段的转换,解决了静止气象卫星在夜间无法直接获取 VIS 图像的问题。Li 等人(2020)基于生成对抗网络提出了一种红外图像生成模型,可将已有 VIS 图像快速转换为红外图像。Yuan 等人(2020)系统性从方法论层面揭示了深度生成模型在遥感信息重建与预测生成中的应用潜力。随着扩散模型的发展, Zhou 等人(2025)提出反射扩散模型(reflectance diffusion, Ref-Diff),基于风云四号 B 星 MIBT 数据生成夜间气象卫星 VIS 观测,从而缓解了现有生成模型在复杂云结构恢复精度方面的不足。



((a) Himawari-9 TCI; (b) Aerosol optical thickness; (c) Chlorophyll-a; (d) Cloud type; (e) Cloud optical thickness; (f) Short wave radiation)

图1 Himawari-9 卫星中依赖太阳反射波段的产品示例, (观测时刻: 2026/01/01 07:00)

Fig. 1 Examples of Himawari-9 satellite products dependent on solar reflectance bands, (Observation time: 2026/01/01 07:00)

现有研究主要面向单一或少量可见光波段,生成光谱波段覆盖不足,此外,面对复杂云系、低云海雾等细粒度结构场景恢复稳定性不足。同时,当前该方向仍面临基础数据资源匮乏的关键瓶颈,公开可用的高质量、多波段、多场景的夜间遥感数据集资源不足,严重制约了相关方法的系统发展与公平评估。因此,如何构建高质量夜间多波段信息补全数

据集成为该领域亟待解决的关键问题。

针对上述存在的问题,本文基于 Himawari-8/9 号静止气象卫星,系统性构建了一个大规模 VIS-NIR 夜间缺失波段补全数据集,命名为 VIS-NIR-70K。该数据集涵盖 2016 年至 2026 年间多区域、多气象条件下的观测样本,总规模约 7 万对数据、近 121 万幅多波段图像,实现了跨时间、跨场景的高质量数据覆盖。本文选取了多种具有代表性的生成模型对所构建数据集进行了系统性评估与对比分析。实验结果表明,基于 MIBT 多波段观测能够有效重建夜间缺失的 VIS-NIR 多波段信息,验证了该数据集在跨模态生成任务中的有效性。本文所构建的数据集不仅缓解了该研究方向公开数据集较少的问题,还为推动夜间遥感信息重建与全天地气系统监测提供了数据基础和技术路线。

1 VIS-NIR-70K 数据集

1.1 数据来源与数据详情

本研究使用的数据来源于日本气象厅(japan meteorological agency, JMA)运行的 Himawari-8/9 号静止气象卫星所搭载的先进葵花成像仪(advanced himawari imager, AHI) (Bessho 等, 2016)。Himawari-8 号于 2015 年 7 月 7 日正式投入业务运行, 2022 年 12 月 13 日, JMA 完成了业务卫星由 Himawari-8 号向 Himawari-9 号的切换。由于两者搭载的 AHI 在参数和通道配置上完全一致,这保证了多波段观测数据在时间跨度上的连续性与一致性。

AHI 传感器每 10 分钟全圆盘观测一次。经度范围为东经 70°至西经 150°,纬度范围为南纬 60°至北纬 60°。AHI 传感器包含 16 个观测通道,涵盖了 3 个 VIS、3 个 NIR 和 10 个热红外观测数据。除此之外,还包含了 4 个观测角度信息,即太阳天顶角(solar zenith angle, SOZ),太阳方位角(solar azimuth angle, SOA),卫星天顶角(satellite zenith angle, SAZ)和卫星方位角(satellite azimuth angle, SAA),一个水陆掩膜信息(land-sea mask, Landsea)用于表征下垫面类型(葛邦宇等, 2018)。Himawari-8/9 号卫星的 TCI 由 albedo 1~3 三个观测波段合成。表 1 详细介绍了 Himawari-8/9 号卫星提供的观测信息,这些观测信息共同构成了夜间 VIS-NIR 缺失波段补全的数据基础,为跨模态生成模型学习 MIBT 与 VIS-

NIR 观测之间的映射关系提供了丰富的参考信息。

表 1 Himawari-8/9 卫星 AHI 传感器波段与辅助信息说明
Table 1 Himawari-8/9 AHI sensor bands and auxiliary information

名称	波段	类型	物理信息	主要应用
albedo 1	0.47	可见光(蓝)	气溶胶、海洋水体和云层信息	云层识别、海洋观测、气溶胶检测、水体边界区分
albedo 2	0.51	可见光(绿)	植被与地表反射信息	植被检测、云检测、地表分类
albedo 3	0.64	可见光(红)	云系纹理、低云/雾和地表反照率特征	标准可见光成像、云图制作、台风监测
albedo 4	0.86	近红外	植被识别、水陆边界和地表覆盖判别	云检测、植被分析、海陆区分
albedo 5	1.61	短波近红外	云顶相态(水滴、冰晶)	冰云识别、雪云区分、云相态检测
albedo 6	2.26	短波红外	云粒子半径、火点和冰晶特征	云粒径反演、火点检测、冰云分析
tbb 7	3.9	中波红外	夜间云层与地表热辐射	夜间云检测、火灾监测、低云识别
tbb 8	6.2	水汽波段	高层大气水汽分布	高层大气湿度分布、水汽输送、云形成过程
tbb 9	6.9	水汽波段	中层大气水汽分布	中层大气湿度分布、水汽输送、云形成过程
tbb 10	7.3	水汽波段	低层大气水汽分布	低层大气湿度分布、水汽输送、云形成过程
tbb 11	8.6	红外	云相态、冰云信息	云顶温度、云顶高度、地表温度
tbb 12	9.6	红外	臭氧吸收信息	云顶温度、云顶高度、地表温度
tbb 13	10.4	红外窗口	云顶温度、地表温度	云顶温度、云顶高度、地表温度
tbb 14	11.2	红外窗口	云顶高度与温度	云顶温度、云顶高度、地表温度
tbb 15	12.4	分裂窗	薄云、沙尘和雾检测	云顶性质分析、水汽校正、识别沙尘、薄云、雾
tbb 16	13.3	CO ₂ 吸收波段	云顶高度、云顶气压反演	云顶高度估计、高云检测、云层敏感性增强
SOZ	--	几何参数	太阳天顶角	太阳光线与地面法线夹角, 决定反照率大小
SOA	--	几何参数	太阳方位角	太阳照射的方向
SAZ	--	几何参数	卫星天顶角	卫星观测方向与地面法线夹角
SAA	--	几何参数	卫星方位角	卫星观测方向
Landsea	--	掩膜参数	陆地与海洋掩膜	表征下垫面类型, 区分陆地、海洋、山脉、河流

注: --表示当前单元格没有对应数值。

1.2 数据预处理

由于 Himawari-8/9 卫星的 VIS-NIR 波段在夜间无法直接获取, 本文希望生成式模型通过学习白天 MIBT 与 VIS-NIR 波段之间的映射关系, 可以在夜间通过稳定观测的 MIBT 生成对应时刻的 VIS-NIR 缺失观测。因此, 构建的夜间卫星缺失波段信息补全数据集 VIS-NIR-70K 包含白天成对的 MIBT 与 VIS-NIR 数据, 以及对应的观测角度信息和水陆掩膜。

在时间维度上, 本文从 2016 年 1 月至 2024 年 12 月之间随机选取了 690 个 Himawari-8/9 卫星观测样本构建训练集。从 2025 年 12 月至 2026 年 1 月之间随机选取了 80 个 Himawari-9 卫星观测样本构建测试集。为降低样本之间的时序相关性, 任意两个被

选择的样本的时间间隔均大于 1 小时。为避免训练集与测试集之间由于相邻时刻观测造成的场景信息泄漏, VIS-NIR-70K 采用跨年份独立划分策略, 将 2016~2024 年数据用于模型训练, 而将 2025~2026 年数据作为独立测试集, 以评估模型在面对未来时段卫星观测时的泛化能力。

在空间维度上, 每个观测样本包含 16 个观测波段和 4 个观测角度信息。每个波段或观测角度信息的空间分辨率为 6001×6001 像素, 同一时刻的不同波段已经由数据发布者统一配准到同一像元网格。其中, albedo 1~albedo 6 的数值分布为 (0.0, 1.0), tbb 7 的数值分布为 (230, 400), tbb 8~tbb 10 的数值分布为 (180, 280), tbb 11~tbb 15 的数值分布为

(180, 330), tbb 16的数值分布为(180, 300), SOZ的数值分布为(0, 100), SOA的数值分布为(-180, 180), SAZ的数值分布为(0, 90), SAA的数值分布为(-180, 180)。为便于后续训练生成模型和可视化, 每个波段的数据通过公式(1)归一化至(0, 255), 并将.nc文件转换为.png图像文件。为降低极端异常值对数据分布及模型训练稳定性的影响, 采用基于分位数的异常值截断方法。即对于每个波段分别统计第1百分位数 P_1 和第99百分位数 P_{99} , 将高于 P_{99} 的像素值截断为 P_{99} , 将低于 P_1 的像素值截断为 P_1 。通过该处理可有效抑制极端离群值, 同时尽可能保持原始观测数据的主体分布特征。

$$\tilde{x}_i = \frac{x_i - x_i^{mn}}{x_i^{mx} - x_i^{mn}} \quad (1)$$

式中, $\tilde{x}_i$ 表示归一化后的第 i 个观测波段, x_i 表示第 i 个观测波段或观测角度信息, x_i^{mn} 与 x_i^{mx} 分别对应 x_i 的最小值和最大值, 观测角度的最小值和最大值分别为 -180° 和 180° 。

原始的Himawari-8/9卫星图像覆盖范围广、空间尺寸大, 部分时刻同时包含昼夜观测区域。由于VIS-NIR波段依赖太阳辐射, 夜间或低照度条件下的观测信息会显著减弱甚至缺失, 导致VIS-NIR波段的部分区域在夜间无法与稳定获取的MIBT观测建立准确的对应关系。因此, 如果直接利用原始观测数据构建样本, 不仅会引入大量低质量数据, 还会影响后续模型对VIS-NIR与MIBT映射关系的学习能力。

为了保证数据质量, 本文首先利用SOZ对原始观测数据进行筛选。SOZ直接反映太阳照射条件, 并决定地表及云层的反照率水平。基于原始SOZ图像, 通过设置阈值剔除低照度区域($SOZ > 85^\circ$), 仅保留太阳照射条件较好的区域($SOZ \leq 85^\circ$), 从而确保选择的样本对具有可靠的映射关系(Zhou等, 2025)。图2展示了基于SOZ构建有效数据对的数据预处理流程。

在获得满足光照条件的有效观测区域后, 考虑到原始卫星图像仍具有较大的空间尺寸以及复杂的场景分布特征, 直接使用原始数据集会导致样本规模不均衡、存储计算开销过大, 并增加模型训练难度。因此, 本文进一步采用无重叠裁剪策略, 将原始图像划分为 512×512 像素的图像块。该尺寸既能较好地保留云体形态、纹理细节及空间上下文信

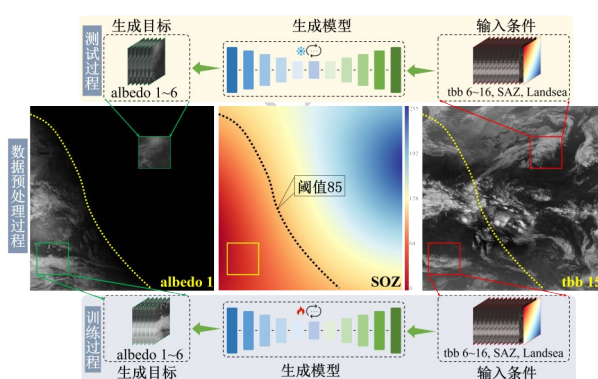


图2 数据预处理流程和模型训练与测试流程示例, (观测时刻: 2026/01/01 10:00)

Fig. 2 Examples of the data preprocessing workflow and model training and testing workflow, (Observation time: 2026/01/01 10:00)

息, 又能够提高样本利用率并降低计算成本, 从而在训练效率与空间信息完整性之间取得良好的平衡, 为后续VIS-NIR缺失波段补全任务提供高质量的数据基础。

1.3 输入条件与预测目标

本文选取了夜间观测性能较稳定的1个中波红外通道(tbb 7)、3个水汽通道(tbb 8~10)和6个长波红外通道(tbb 11~16)作为可控输入条件, 用于训练模型学习MIBT稳定观测与VIS-NIR之间的映射关系。同时, 为了减弱大范围观测下观测几何变化引起的大气路径效应, 本文还引入SAZ作为辅助输入变量。SAZ表征卫星观测方向与地表-大气系统之间的几何关系, 是影响遥感辐射响应的重要参数, 将其纳入模型有助于刻画不同视角下的辐射差异, 提升生成结果在空间纹理、区域亮度分布及复杂天气场景下的物理一致性。与此同时, 水陆掩膜被用于标识像元所属的陆地、海洋及海岸过渡区域, 可为模型提供下垫面类型先验信息, 增强其对异质地表场景的表征能力, 尤其有助于提升海陆交界区域及海雾、低云等复杂天气背景下的生成质量。

然而, 由于SOZ与太阳光反射有关, 夜间太阳光反射微弱, 无法稳定地表征夜间的VIS-NIR成像, 因此不纳入可控输入条件。SOA/SAA主要表示太阳/卫星方位方向, 不决定辐射反射值, 与夜间VIS-NIR响应关系微弱, 因此也不纳入可控输入条件。

综上所述, VIS-NIR-70K数据集共包含62,290对训练样本和5,446对测试样本。图3给出了本文构建的VIS-NIR-70K数据集的各波段示例结果, 表2

详细地展示了 VIS-NIR-70K 数据集的样本分布。从场景分布来看,训练集与测试集在多云、晴空及台风等主要天气类型上的比例总体接近,且均同时覆盖陆地和海洋下垫面,能够降低明显的场景分布失衡对测试结果的影响。当前测试样本主要集中于 2025~2026 年,因此测试结果可能会受到一定季节分布差异的影响,未来将构建跨季节独立测试集,增强生成模型在不同季节和气候背景下的泛化能力验证。

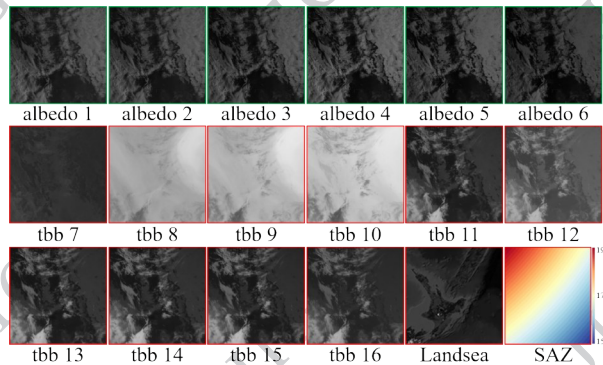


图3 VIS-NIR-70K 各波段示例, (观测时刻: 2025/12/15 18:00)
Fig. 3 Examples of different bands in VIS-NIR-70K, (Observation time: 2025/12/15 18:00)

1.4 相关数据集对比

虽然现有夜间可见光生成研究已经构建了一定规模的数据资源,但是,在光谱完整性、时间跨度与任务通用性方面仍存在局限。表3对比了 VIS-NIR-70K 与现有夜间可见光生成数据集。NightVision (Harder 等, 2020) 基于 GOES-16 建立约 8000 对红外-可见光样本,主要面向 RGB 可见图像生成,数据规模和空间覆盖范围有限。Yao 等人 (2025) 基于 Himawari-8/9 构建了大规模 AHI 数据集,引入多红外波段、观测几何与地表信息,但其预测目标仅为 0.645 μm 单一可见光波段,且研究场景重点面向热

带气旋。RefDiff (Zhou 等, 2025) 基于 FY-4B 多波段热红外、SAZ 和地表信息生成 0.47 μm 、0.65 μm 和 0.825 μm 三个反射波段,但数据时间跨度相对较短,且未覆盖短波红外波段。相比之下, VIS-NIR-70K 覆盖 2016~2026 年 Himawari-8/9 长期观测,包括完整的 tbb 7~16、SAZ 和水陆掩膜为条件,同时提供 albedo 1~6 六个对应的 VIS、NIR 目标波段,并覆盖多云、晴空、海雾、台风及不同下垫面场景,因此能够为夜间完整多光谱信息补全、跨波段物理关系研究及不同生成模型的统一评估提供更加系统的数据基础。

2 方法与实验

2.1 实验设计

2.1.1 实验细节

所有实验均基于 PyTorch-GPU 2.1.1 和 Python 3.8.20 开发环境实现,并在配备 24 GB 显存的 NVIDIA GeForce RTX 4090 GPU 平台上完成训练与测试。为保证实验结果的可重复性与公平性,所有对比模型均采用相同的软硬件环境、数据划分方式以及训练测试流程。

在模型训练过程中,统一采用 Adam 优化器进行参数更新,初始学习率设置为 1×10^{-4} 。为了提高模型收敛稳定性并获得更优的泛化性能,引入余弦退火学习率衰减策略,使学习率从 1×10^{-4} 逐步衰减至 1×10^{-5} 。训练批次大小设置为 8。考虑到不同模型的收敛特性存在差异, CycleGAN 和 Pix2Pix 训练 300 个 Epoch,其余生成模型训练 1000 个 Epoch,以确保各模型均达到稳定收敛状态。

VIS-NIR-70K 数据集中的原始样本尺寸为 512×512 像素。考虑到高分辨率输入会显著增加模型训练开销,同时现有主流图像生成模型普遍采用 256×256 分辨率作为标准输入尺寸,因此在不影

表2 VIS-NIR-70K数据集详情表
Table 2 Detailed information of the VIS-NIR-70K dataset

数据集	原始样本	图像对	多波段图像	多云(%)	晴空(%)	陆地(%)	海洋(%)	台风(%)	时间跨度
训练集	690	62,290	1,121,220	23.09	76.91	48.41	51.59	1.15	2016/01~2024/12
测试集	80	5,446	98,028	23.39	76.61	41.39	58.61	1.25	2025/12~2026/01
综合	770	67,736	1,219,248	23.11	76.89	47.85	52.15	1.16	2016/01~2026/01

响模型性能的前提下,将所有样本统一缩放至 256×256 像素后再输入网络进行训练和推理。该处理方

式既能够保留主要云系结构和空间纹理特征,又能够有效降低显存占用和训练时间,提高实验效率。

表3 VIS-NIR-70K与现有夜间可见光生成数据集对比

Table 3 Comparison of VIS-NIR-70K with existing nighttime visible image generation datasets

数据集名称	卫星/传感器	时间范围	输入数据	生成目标	规模(对)	公开	场景
NightVision(Harder 等, 2020)	GOES-16/ABI	2018~2020	Band 8~16	Band 1~3	8,000	是	多场景
AHI dataset(Yao 等, 2025)	Himawari-8/9/AHI	2015~2023	tbb 8~16, 方向参数, Landsea	albedo 3	246, 078	是	热带气旋
RefDiff(Zhou 等, 2025)	FY-4B/AGRI	2022~2024	Band 7~14, 方向参数, Landsea	Band 1~3	262, 947	否	多场景
VIS-NIR-70K	Himawari-8/9/AHI	2016~2026	tbb 7~16, 方向参数, Landsea	albedo 1~6	67, 736	是	多场景

2.1.2 评价指标

为全面评估不同生成模型在夜间 VIS-NIR 缺失波段补全任务中的性能,本文从图像结构、像素误差、辐射精度、空间相关性和光谱一致性等多个维度对生成结果进行综合评价。其中,均方误差(mean squared error, MSE)、峰值信噪比(peak signal-to-noise ratio, PSNR)和结构相似性指数(structural similarity index measure, SSIM)用于评估生成图像与真实图像之间的像素一致性和结构保持能力。fréchet inception distance (FID) 和 kernel inception distance (KID)用于评估生成样本与真实样本在特征空间中的分布差异,从而衡量生成结果的感知质量和真实性。平均偏差(mean bias error, MBE)用于评价生成波段与真实波段的系统性辐射偏差。均方根误差(root mean square error, RMSE)衡量整体辐射重建误差。相关系数(correlation coefficient, CC)用于刻画生成结果与真实观测在空间变化趋势上的一致性。光谱角制图(spectral angle mapper, SAM)通过计算生成与真实多波段光谱向量之间的夹角,评价 VIS-NIR 多波段光谱形状的一致程度。

MSE 用于衡量生成图像与真实图像之间的像素级误差, MSE 越小越好。

PSNR 用于评价真实图像与生成图像之间的误

差水平, PSNR 越大越好。

SSIM 从亮度、对比度和结构三个方面评价图像相似性, SSIM 越大越好。

RMSE 用于衡量生成结果与真实观测之间的整体辐射误差, 对较大误差更加敏感, 取值越小表示生成结果的辐射重建精度越高。

$$RMSE = \sqrt{\frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W (\mathbf{x}_{h,w} - \hat{\mathbf{x}}_{h,w})^2} \quad (2)$$

式中, H 与 W 分别表示图像的高度和宽度, $\mathbf{x}_{h,w}$ 与 $\hat{\mathbf{x}}_{h,w}$ 分别表示真实图像和生成图像在 (h, w) 处的像素值。

FID 用于衡量生成图像和真实图像在深度语义特征空间中的分布差异, FID 越小越好。

$$FID = \left\| \mu_{\mathbf{x}} - \mu_{\hat{\mathbf{x}}} \right\|^2 + \text{Tr}(\Sigma_{\mathbf{x}} + \Sigma_{\hat{\mathbf{x}}} - 2(\Sigma_{\mathbf{x}} \Sigma_{\hat{\mathbf{x}}})^{\frac{1}{2}}) \quad (3)$$

式中, $\mathbf{X}, \hat{\mathbf{X}}$ 分别表示真实图像特征集和生成图像特征集, $\mu_{\mathbf{x}}$ 表示真实图像的语义特征均值, $\mu_{\hat{\mathbf{x}}}$ 表示生成图像的语义特征均值, $\Sigma_{\mathbf{x}}$ 表示真实图像的协方差矩阵, $\Sigma_{\hat{\mathbf{x}}}$ 表示生成图像的协方差矩阵, $\text{Tr}(\cdot)$ 表示矩阵迹。

KID 基于最大均值差异衡量生成图像和真实图像深度语义特征分布差异, KID 越小越好。

$$KID = \frac{1}{m(m-1)} \sum_{i=1}^m \sum_{j=1, j \neq i}^m k(\mathbf{X}_i, \mathbf{X}_j) + \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j=1, j \neq i}^n k(\hat{\mathbf{X}}_i, \hat{\mathbf{X}}_j) - \frac{2}{mn} \sum_{i=1}^m \sum_{j=1}^n k(\mathbf{X}_i, \hat{\mathbf{X}}_j) \quad (4)$$

式中, m, n 表示真实样本和生成样本的数量, $k(\cdot, \cdot)$ 表示多项式核函数, $\mathbf{X}_t, \hat{\mathbf{X}}_t, t \in \{i, j\}$ 表示第 t 个真实样本和第 t 个生成样本的语义特征。该指标不受真实样本与生成样本数量不相等的限制, 可稳定评价图像生成质量。

MBE 用于衡量生成波段相对于真实观测的系统性辐射偏差, MBE 越接近 0 越好, MBE 大于 0 表示模型整体高估, 小于 0 表示整体低估。

$$MBE = \frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W (x_{h,w} - \hat{x}_{h,w}) \quad (5)$$

CC 采用 Pearson 相关系数衡量生成结果与真实观测在空间变化趋势上的一致性。CC 的取值范围为 $(-1, 1)$, 越接近 1 表示生成结果与真实观测具有越高的空间相关性。

$$CC = \frac{\sum_{h=1}^H \sum_{w=1}^W (x_{h,w} - \mu_x)(\hat{x}_{h,w} - \mu_{\hat{x}})}{\sqrt{\sum_{h=1}^H \sum_{w=1}^W (x_{h,w} - \mu_x)^2} \sqrt{\sum_{h=1}^H \sum_{w=1}^W (\hat{x}_{h,w} - \mu_{\hat{x}})^2}} \quad (6)$$

式中, μ_x 表示真实图像的均值, $\mu_{\hat{x}}$ 表示生成图像的均值。

SAM 评价生成结果与真实观测之间的多波段光谱一致性。 (h, w) 位置的 N 个波段构成的光谱向量为 $\mathbf{y}_{h,w} = [y_{h,w}^1, y_{h,w}^2, \dots, y_{h,w}^N]$ 和 $\hat{\mathbf{y}}_{h,w} = [\hat{y}_{h,w}^1, \hat{y}_{h,w}^2, \dots, \hat{y}_{h,w}^N]$ 。

$$SAM = \frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W \arccos \left(\frac{\hat{\mathbf{y}}_{h,w}^T \mathbf{y}_{h,w}}{\|\hat{\mathbf{y}}_{h,w}\|_2 \|\mathbf{y}_{h,w}\|_2} \right) \quad (7)$$

SAM 越接近 0, 说明生成光谱与真实光谱的形状越一致。

2.1.3 对比模型

传统图像补全方法主要解决单幅图像中局部区域缺失的问题, 通常依赖缺失区域周围已知像素的空间连续性、纹理相似性或邻域结构, 通过插值等方式恢复缺失内容。夜间缺失 VIS-NIR 生成则属于跨光谱、跨模态映射任务, 目标 VIS-NIR 波段在夜间表现为大区域、广范围的整体不可观测性, 不存在可直接利用的邻域像素。同时, MIBT 与 VIS-NIR 之间具有复杂的非线性辐射关系, 并受到云相态、水汽、地表类型及观测几何等因素共同影响。因此, 传统空间域图像补全方法无法直接适用于本研究的 VIS-NIR 波段生成任务。

为全面评估 VIS-NIR-70K 数据集在夜间缺失波

段补全任务中的有效性与挑战性, 本文选取生成式人工智能不同发展阶段具有代表性的图像生成模型作为对比方法, 涵盖生成对抗网络 (generative adversarial network, GAN)、卷积扩散模型 (diffusion model) 以及 Transformer 扩散模型 (diffusion transformer) 三类主流技术路线。具体包括 Linear Regression、CycleGAN (Zhu 等, 2017)、Pix2Pix (Isola 等, 2017)、DDPM (Ho 等, 2020)、DiT-2S (Peebles 等, 2023)、DiT-2B (Peebles 等, 2023)、DiT-2L (Peebles 等, 2023) 以及 RefDiff (Zhou 等, 2025)。训练上述所有模型的损失函数都使用原始论文提供的基础损失。

2.2 VIS-NIR 生成对比实验

表 4 和表 5 给出了每个模型生成的 VIS-NIR 六个波段和 TCI 真彩色波段 (albedo 1~3) 上的平均定量评价结果。从整体趋势来看, 扩散模型普遍优于基于 GAN 的模型, 说明夜间缺失波段补全任务本质上需要学习复杂的跨模态映射关系, 而扩散模型在建模高维数据分布和恢复细粒度纹理方面具有更强的优势。MIBT 到 VIS-NIR 并非确定性的像素映射, 相同或相近的红外亮温状态可能对应不同的地表反射率、云光学性质和云微物理状态, 具有明显的跨模态非线性和一对多映射特征, 因此简单的多波段线性映射模型 (Linear) 难以重建高质量的 VIS-NIR 波段。

从模型结构角度来看, Cycle-GAN 和 Pix2Pix 是基于 GAN 的模型。Cycle-GAN 由于缺乏严格的配对监督, 仅依赖循环一致性约束学习跨域映射, 因此在多数评价指标上表现较差。Pix2Pix 利用配对样本进行监督训练, 生成的 VIS-NIR 观测和 TCI 观测的 SSIM 分别提升至 0.746 和 0.770, 证明配对监督能够有效增强生成结果与真实图像之间的结构一致性。基于 GAN 的生成模型容易受到对抗训练不稳定和模式崩塌问题的影响, 导致 FID 和 KID 指标明显高于扩散模型, 说明生成结果在整体分布和细节真实性方面存在局限性。

从扩散模型的尺度空间来看, DiT 系列模型采用预训练变分自编码器 (variational autoencoder, VAE) 构建潜在空间扩散框架, 通过 Transformer 实现全局特征建模。相比基于 GAN 的模型, DiT 系列在 FID 和 KID 指标上取得明显提升, 表明潜在空间扩散能够学习更合理的图像分布特征。然而, DiT-2S、

DiT-2B 和 DiT-2L 之间性能差异较小,且 SSIM 均维持在 0.66~0.70 之间。这说明虽然 Transformer 具有较强的全局依赖建模能力,但预训练的 VAE 编码器性能会影响扩散模型生成的图像质量。相比之下,DDPM 和 RefDiff 均在原始图像空间执行扩散与去噪过程,能够直接利用完整的空间信息进行建模精确的映射关系。因此,在所有模型中取得了最优性能。其中,DDPM 在 VIS-NIR 任务中获得最高 SSIM (0.864)、最低 FID(11.178)和最低 KID(0.007),在 TCI 任务中同样保持领先优势,表明原始空间扩散对于云体纹理、边缘细节及空间结构的恢复能力更强。RefDiff 作为针对夜间卫星可见光生成设计的专用扩散模型,在 PSNR 指标上略优于 DDPM,但 FID

和 KID 略有升高,说明其能够获得较高的像素重建精度,但在整体特征分布一致性方面仍存在一定优化空间。原始图像空间扩散模型的优势与其直接在像素空间建模有关,避免了潜在空间编码压缩过程中部分弱辐射差异和细粒度云纹理信息的丢失,因此更有利于保持遥感图像中的云边缘及局部亮度变化。

综上所述,VIS-NIR-70K 数据集能够有效区分不同技术路线模型的性能差异,其中,原始图像空间扩散模型在细节恢复能力、结构保持能力以及分布一致性方面均展现出显著优势,为夜间 VIS-NIR 缺失波段补全提供了可靠的解决方案。

表 4 生成的 VIS-NIR 图像平均定量对比结果

Table 4 Average quantitative comparison of VIS-NIR image generation results

模型名称	SSIM ↑	PSNR ↑	MSE ↓	FID ↓	KID ↓	MBE	RMSE ↓	CC ↑	SAM ↓
Linear	0.669	22.532	0.010	34.330	0.028	-0.0057	0.0855	0.7373	0.2253
Cycle-GAN(Zhu 等, 2017)	0.531	20.139	0.017	26.560	0.019	0.0017	0.1126	0.6411	0.2632
Pix2Pix(Isola 等, 2017)	0.746	24.646	0.006	22.490	0.015	0.0069	0.0682	0.8373	0.1667
DiT-2S(Peebles 等, 2023)	0.666	25.697	0.005	17.618	0.011	0.0031	0.0593	0.8576	0.1736
DiT-2B(Peebles 等, 2023)	0.666	25.686	0.005	16.367	0.010	0.0003	0.0600	0.8533	0.1766
DiT-2L(Peebles 等, 2023)	0.669	26.001	0.005	17.596	0.011	0.0062	0.0581	0.8643	0.1774
DDPM (Ho 等, 2020)	0.864	28.891	0.003	11.178	0.007	-0.0078	0.0448	0.9566	0.1387
RefDiff(Zhou 等, 2025)	0.860	29.143	0.003	13.593	0.009	-0.0091	0.0397	0.9532	0.0579

注: 黑色字体表示最优结果,↑表示值越大越好,↓表示值越小越好。

表 5 生成的 TCI 图像平均定量对比结果

Table 5 Average quantitative comparison of TCI image generation results

模型名称	SSIM ↑	PSNR ↑	MSE ↓	FID ↓	KID ↓	MBE	RMSE ↓	CC ↑	SAM ↓
Linear	0.7304	22.5535	0.009	40.6222	0.0342	-0.0069	0.0837	0.7605	0.0885
Cycle-GAN(Zhu 等, 2017)	0.482	20.141	0.018	26.361	0.019	-0.0039	0.1103	0.6713	0.0988
Pix2Pix(Isola 等, 2017)	0.770	24.502	0.006	23.816	0.016	0.0011	0.0672	0.8351	0.0644
DiT-2S(Peebles 等, 2023)	0.694	25.898	0.004	19.060	0.012	-0.0006	0.0563	0.8595	0.0649
DiT-2B(Peebles 等, 2023)	0.695	25.866	0.004	18.027	0.011	-0.0023	0.0570	0.8552	0.0680
DiT-2L(Peebles 等, 2023)	0.700	26.096	0.004	19.128	0.012	-0.0020	0.0553	0.8672	0.0659
DDPM (Ho 等, 2020)	0.881	29.435	0.002	12.515	0.008	-0.0091	0.0397	0.9532	0.0579
RefDiff(Zhou 等, 2025)	0.877	29.483	0.002	15.240	0.011	0.0049	0.0398	0.9496	0.0632

注: 黑色字体表示最优结果,↑表示值越大越好,↓表示值越小越好。

表 6 对比了不同模型的参数量与计算时间复杂度。以生成性能较优的 DDPM 和 RefDiff 为例,两者

具有较高的单步计算量,且完整生成过程需要执行 1000 次迭代采样,因此在全圆盘 VIS-NIR 补全任务

中会带来较大的累计计算开销。然而,结合 Himawari AHI 约 10min 一次的全圆盘观测更新周期来看,相邻时刻发生显著变化的 VIS-NIR 像元约为 30×6001 个,占全图约 0.49%,表明连续观测之间具有较强的时空冗余。进一步地,DDPM 生成 100 幅图像平均耗时约 45.6s,对应覆盖约 80% 的全圆盘有效区域;若近似按照计算量线性扩展,则完成一次全圆盘补全约需 57s,约为 10min 观测更新周期的 9.5%,即模型处理速度相对于卫星观测更新需求具有约 10 倍的时间剩余量。由此说明,尽管原始空间扩散模型计算复杂度较高,但其推理效率仍具备在 Himawari 高时间分辨率连续观测条件下开展实时 VIS-NIR 补全的潜力。

图 4 展示了不同模型在 VIS-NIR 补全任务中的生成结果和对应误差热力图。整体来看,每个模型都能够恢复云体的主要结构,但在纹理细节、云层边缘以及亮度分布方面存在明显差异。虽然 CycleGAN 生成结果恢复云系的大致轮廓,但由于缺乏严

格的配对监督约束,生成的图像整体存在明显的纹理缺失和结构模糊现象。对应的高误差区域集中分布在云体边缘和高云层区域,表明其对细节特征和亮暗变化较大区域的恢复能力有限。Pix2Pix 依托配对样本监督学习,生成图像的整体结构与真实图像较为接近。但其在复杂云纹理区域仍存在一定程度的平滑现象,误差主要集中于云层内部的细粒度纹理区域。DiT 系列模型表现出较好的全局结构建模能力,生成结果在云系空间分布和全局亮暗分布方面具有较高一致性。但部分局部细粒度区域仍存在纹理细节缺失和边缘模糊问题。这说明潜在空间扩散虽然能够有效学习全局语义特征,但对高频细节信息的恢复仍受到 VAE 压缩过程的影响。相比之下,DDPM 和 RefDiff 均表现出更加优异的生成质量。两种模型能够较好地恢复云层纹理、边缘轮廓以及局部亮暗变化特征。但是 RefDiff 会存在个别通道生成亮度分布差异较大的问题,这说明其对多波段 VIS-NIR 生成仍然存在局限性。

表 6 不同模型计算复杂度对比
Table 6 Comparison of computational complexity across different models

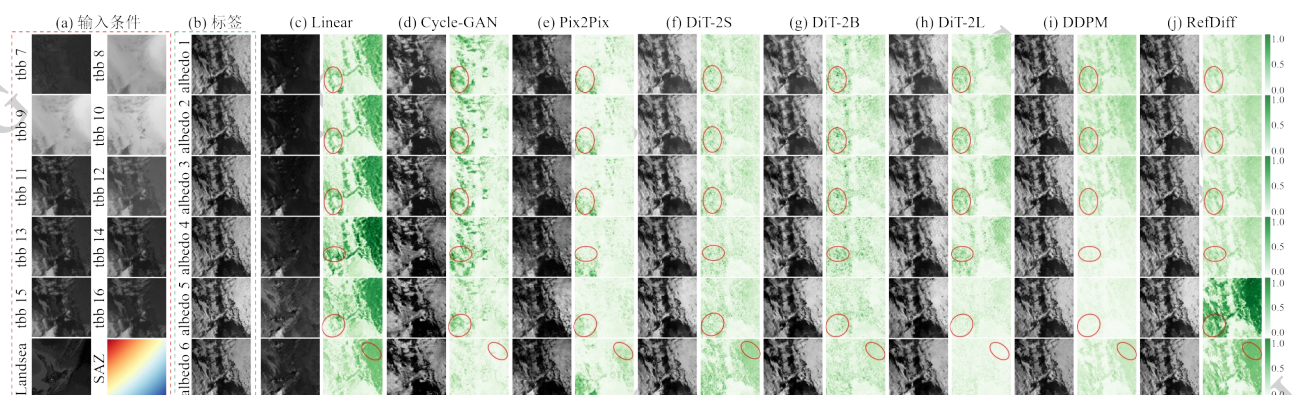
模型名称	参数 (M)	时间 (100 张图像/s)	采样步数 (步)
Linear	0.0099	0.000	1
Cycle-GAN(Zhu 等, 2017)	28.3741	0.022	1
Pix2Pix(Isola 等, 2017)	57.2104	0.003	1
DiT-2S(Peebles 等, 2023)	55.2938	0.290	1000
DiT-2B(Peebles 等, 2023)	218.2304	0.305	1000
DiT-2L(Peebles 等, 2023)	764.4571	0.554	1000
DDPM(Ho 等, 2020)	57.8715	45.600	1000
RefDiff(Zhou 等, 2025)	57.8823	909.740	1000

2.3 多波段消融实验

为进一步分析不同输入波段对 VIS-NIR 波段生成性能的贡献,本文以 DDPM 为基础设计了如图 5 所示的六组消融实验。实验结果表明,输入红外波段的完整性是影响 VIS-NIR 生成性能的主要因素。

Ablation1 采用 tbb 7、tbb 15 和 tbb 16 三个波段,可以利用短波红外、热红外窗口及云顶相关信息建立基本的场景热辐射表征,但对于大气垂直结构和云微物理特征的描述有限。在 Ablation1 的基础上,Ablation2 加入 tbb 8~10 的水汽波段后,各输出波段

总体仅表现出小幅变化,说明高、中层水汽信息能够提供一定的大气状态约束,但其与 VIS-NIR 地表反射及云光学特性的联系较为局限。在 Ablation2 的基础上,Ablation3 的输入扩展至完整的 tbb 7~16 波段后,生成的 VIS-NIR 波段的 SSIM、PSNR 和 MSE 均呈现明显的整体改善,表明连续红外波段提供的互补信息是生成性能提升的主要来源。其原因在于 tbb 11~14 进一步包含云相态、大气吸收以及云顶和地表热状态等信息,与 tbb 7~10、tbb 15 和 tbb 16 联合后能够更完整地刻画水汽、云、地表及大气垂直结



((a) Input condition; (b) Labels; (c) Linear; (d) Cycle-GAN; (e) Pix2Pix; (f) DiT-2S; (g) DiT-2B; (h) DiT-2L; (i) DDPM; (j) RefDiff)

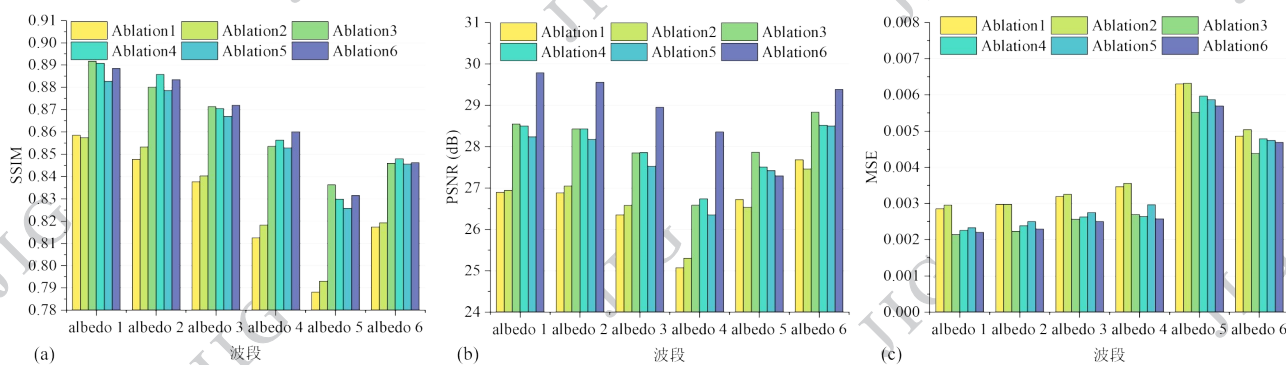
图4 不同算法在不同观测波段中的生成结果与误差热力图对比, (观测时刻: 2025/12/15 18:00)

Fig. 4 Comparison of generated results and error heatmaps of different algorithms across various observation bands, (Observation time: 2025/12/15 18:00)

构。Ablation1~3的结果表明,生成VIS-NIR的质量受云状态、地表类型及大气条件等潜在物理状态控制,同时,完整多波段热红外观测能够为VIS-NIR波段提供更充分的物理约束。

在Ablation3基础上单独添加SAZ (Ablation4)后,部分波段的结构一致性得到改善,说明观测几何信息能够帮助模型补偿不同卫星观测角度造成的光程及观测差异,单独添加Landsea (Ablation5)的改善相对有限,说明水陆掩膜主要提供地表类别先验,而无法替代连续红外观测中包含的丰富辐射信息。但

是,同时将SAZ与Landsea加入(Ablation6)后,生成的VIS-NIR整体取得较优且更稳定的结果,尤其在albedo 3~4等波段表现突出。这表明几何信息与地表类型信息具有一定互补作用。对于albedo 5~6,完整热红外波段本身已提供较强约束,说明近红外云相态和粒子特征的恢复更加依赖完整的云-大气热辐射信息。综上所述,完整tbw 7~16是VIS-NIR生成的核心信息来源,而SAZ和Landsea主要发挥辅助物理先验和条件约束作用。



((a) SSIM; (b) PSNR; (c) MSE)

图5 不同输入波段组合对VIS-NIR波段生成性能的消融实验

Fig. 5 Ablation study of different input bands for VIS-NIR band generation

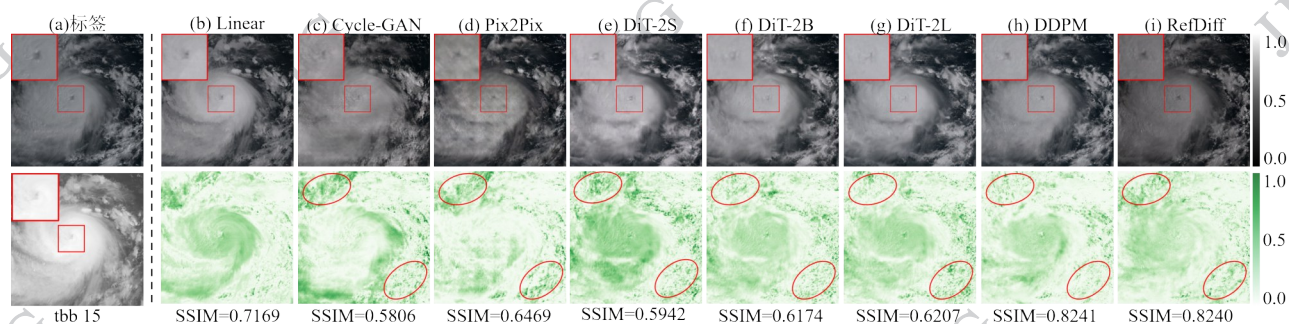
2.4 多场景生成对比实验

灾害性天气通常伴随着复杂的云系结构变化,因此,如果夜间也可以准确获取VIS-NIR将对于灾害监测、预警分析及应急决策具有重要意义(Zhang等, 2025a)。为了评估基于VIS-NIR-70K数据集训

练的模型在灾害性天气场景下的生成性能与鲁棒性,本文选取2025年11月6日第2526号台风“凤凰”作为典型案例,该案例独立于VIS-NIR-70K数据集,并在图6中展示了不同模型生成的台风场景中的TCI结果。其中台风眼是表征台风强度、发展阶段

及潜在危害的重要特征,生成质量能够较为直观地反映模型对灾害性天气细节结构的刻画能力(Jung等, 2025)。定量结果表明,DDPM和RefDiff模型取得了较高的SSIM指标,说明其生成结果在结构相似性方面与真实图像更加接近,能够更好地保持云系纹理与局部细节信息。图6进一步对比了不同模型生成的台风眼区域。相较于其他模型,DDPM与RefDiff生成的台风眼在形状轮廓、边界结构及灰度

分布方面与真实标签更为一致,表现出更强的细节恢复能力和结构表征能力。由于夜间缺少VIS-NIR观测,Zhang等人(2025b)仅依赖红外图像预测台风的发展趋势。如果基于扩散模型利用红外亮温观测生成夜间VIS-NIR观测,那么在台风发展趋势预测估计中,引入包含云体形态变化和云层纹理信息的观测波段将有可能提升台风发展趋势预测估计的准确性与稳定性。



((a) Labels; (b) Linear; (c) Cycle-GAN; (d) Pix2Pix; (e) DiT-2S; (f) DiT-2B; (g) DiT-2L; (h) DDPM; (i) RefDiff)

图6 不同算法在台风天气场景中的TCI生成结果和误差热力图对比,(观测时刻: 2025/11/06 01:00)

Fig. 6 Comparison of TCI generation results and error heatmaps of different algorithms under typhoon conditions, (Observation time: 2025/11/06 01:00)

为进一步验证数据在不同典型场景下的适用性,本文在图7中补充了多云、晴空、海雾以及陆地/海洋等场景的生成结果对比实验。结果表明,在多云场景下,模型能够较好恢复大范围云团的空间形态、边缘结构及云系纹理,说明其能够从多热红外信息中学习云顶热状态与VIS-NIR反射特征之间的映射关系。对于碎云和不同云量条件,生成结果仍能保持较好的层次感与连续性,表明模型对复杂云分布具有较强鲁棒性。晴空海洋场景中,生成结果能够保持海面低反射、纹理平滑的物理特征,背景噪声较少。晴空陆地场景则能够恢复更复杂的地表反射差异和空间纹理,说明模型不仅适用于云区,也能够适应不同下垫面的辐射特性差异。对于海雾场景,生成结果能够较好重建近海连续分布的海雾层亮区及其与周围海面的过渡关系,反映出模型能够利用热红外中蕴含的低层云雾、水汽及海表热状态信息推断VIS-NIR反射特征。对于陆地场景,模型能够保持不同下垫面的辐射差异,呈现更丰富的纹理与反射变化,恢复地表异质性与边界结构。总体而言,多场景对比实验表明VIS-NIR-70K数据集覆盖了不同天气、云量与下垫面条件,具有较强的场景多样

性。同时,基于VIS-NIR-70K训练的生成模型在多种典型气象场景下均表现出稳定的生成能力,体现了较好的泛化性与业务应用潜力。

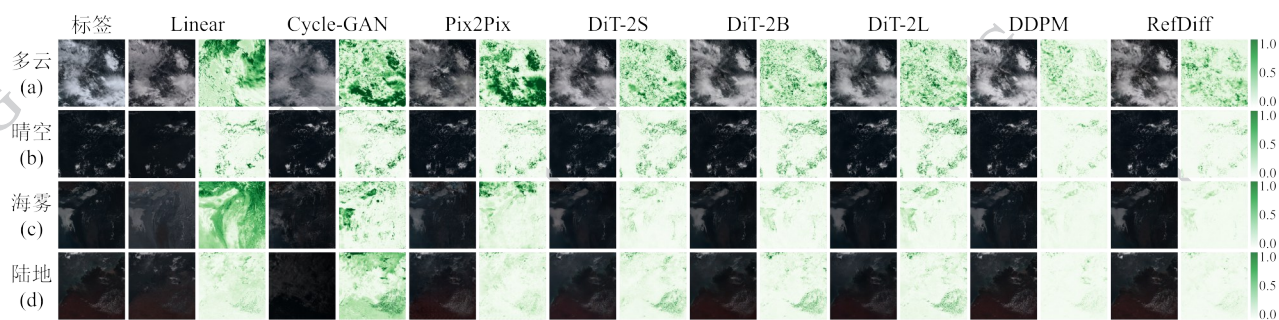
2.5 夜间VIS-NIR全图补全分析

图8是DDPM对葵花9号卫星全图VIS-NIR缺失区域补全对比结果。实验结果表明,DDPM能够对葵花9号缺失的VIS-NIR波段进行较高质量补全。单波段的albedo 1~albedo 6在云系纹理、旋涡结构及海陆边界等细节上保持了良好的空间连续性、灰度一致性和层次一致性。合成的TCI图像整体色调自然、云体形态清晰,与原始可见区域衔接平滑。该结果说明基于VIS-NIR-70K训练的扩散模型能有效恢复卫星夜间无法获取的关键波段的观测信息,其中,合成的TCI为夜间连续观测、灾害监测和气象业务化应用提供重要支撑(Ma等, 2026)。

3 存在的挑战与未来展望

3.1 VIS-NIR-70K域差异与场景局限性

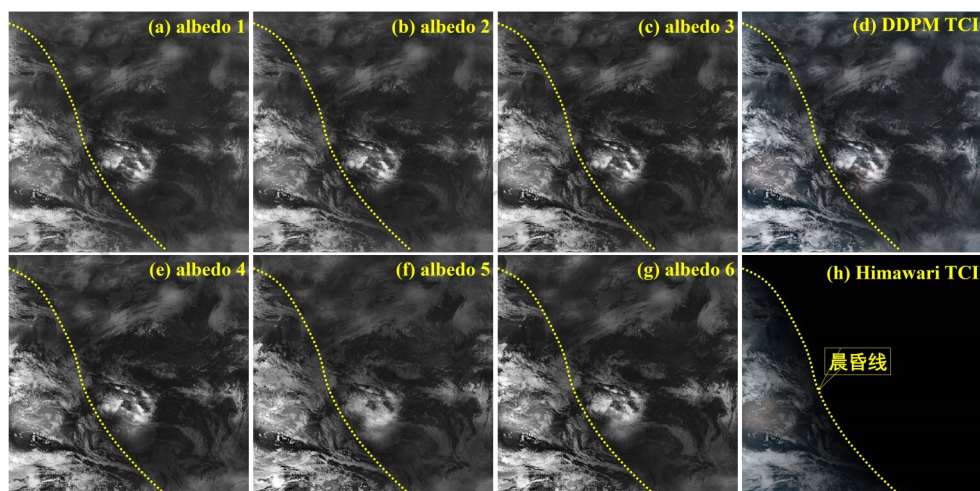
VIS-NIR-70K利用白天同时存在MIBT与VIS-NIR观测的样本建立配对关系,并将模型学习到的



((a) Cloudy; (b) Clear sky; (c) Sea fog; (d) Land)

图7 不同算法在多场景分布中的TCI生成结果和误差热力图对比, (观测时刻: (a) 2026/01/05 02:00; (b) 2025/12/15 23:00; (c) 2026/01/01 05:00; (d) 2026/01/05 05:00)

Fig. 7 Comparison of TCI generation results and error heatmaps of different algorithms across multiple scenarios, (Observation time: (a) 2026/01/05 02:00; (b) 2025/12/15 23:00; (c) 2026/01/01 05:00; (d) 2026/01/05 05:00)



((a) albedo 1; (b) albedo 2; (c) albedo 3; (d) DDPM TCI; (e) albedo 4; (f) albedo 5; (g) albedo 6; (h) Himawari TCI)

图8 基于DDPM的夜间VIS-NIR全图补全结果, (观测时刻: 2026/01/01 10:00)

Fig. 8 DDPM-based full-image nighttime VIS-NIR completion results, (Observation time: 2026/01/01 10:00)

跨波段映射用于夜间VIS-NIR缺失信息补全。该构建方式解决了夜间不能直接获得VIS-NIR真实配对观测的问题,但也带来了昼夜分布差异这一固有限制。昼夜条件下地表温度、大气热状态及云体辐射特征存在变化,尤其部分中波红外通道的辐射组成具有昼夜差异,因此模型从白天数据学习到的MIBT-VIS-NIR映射关系在夜间应用时可能产生一定域偏移。目前夜间全图实验能够验证生成结果的空间连续性和定性合理性,但由于缺少同一时刻真实VIS-NIR观测,尚难开展与白天测试集等价的夜间定量验证。此外,虽然VIS-NIR-70K覆盖晴空、多云、陆地、海洋、海雾和台风等典型场景,并在多场景实验中表现出一定适用性,但是对于沙尘、强对流、冰雪覆盖等低频复杂场景,目前尚未形成具有明确

标签和充足样本量的独立子集。因此,后续将进一步按照天气类型和下垫面建立场景子集,增加少数复杂天气样本比例,并探索昼夜域适应、分布对齐以及如何利用其他夜间观测资料进行交叉验证,以降低昼夜辐射差异对模型泛化能力的影响。

3.2 稳定观测信息挖掘不充分,缺乏物理先验知识引导

现有方法主要依赖深度神经网络从MIBT观测数据中自动学习跨模态映射关系,但对不同红外波段所蕴含的物理意义利用不足。例如,水汽波段反映大气湿度垂直分布,窗口红外波段包含云顶温度和云顶高度信息,而CO₂吸收波段能够表征高层云特征。波段对比实验中albedo 4和albedo 5的重建性能低于TCI波段,进一步说明仅依赖数据驱动映

射难以完整恢复植被、水体、地表反射及云微物理相关信息。这些信息与VIS-NIR观测存在明确的物理关联,但现有模型并未显式引入相关先验约束。因此,未来可结合辐射传输理论、云微物理参数以及气象动力学知识,构建物理引导的生成模型,将数据驱动与机理驱动相结合,提高生成结果的物理一致性、可解释性和泛化能力。

3.3 缺乏夜间遥感生成验证和气象任务专用评价指标体系

目前VIS-NIR补全任务主要采用MSE、RMSE、MBE、PSNR、SSIM、FID和KID等通用图像生成评价指标。这些指标能够从像素误差、结构相似性和特征分布等方面评价生成质量,但难以准确反映遥感图像中特定目标的重建效果。即使本文引入了SAM、CC等指标从辐射误差和光谱一致性等方面评价生成质量,但是这些指标并不能完全等价于实际气象业务可用性评价。例如,对于云体边缘清晰度、纹理连续性、灰度分布一致性以及台风云系等复杂天气结构的恢复能力,现有指标缺乏足够的敏感性。因此,不同模型即使获得相近的定量结果,其视觉质量和实际应用价值仍可能存在明显差异。未来有必要结合遥感业务需求,构建融合边缘保持、纹理统计、频谱分布及气象特征一致性的专用评价体系,为模型性能评估提供更加全面和可靠的依据。

3.4 扩散模型生成效率较低,算力能耗限制业务应用部署

扩散模型在VIS-NIR-70K评估中表现出较高的生成质量,但其性能优势随之带来显著的计算开销,逐步去噪的生成机制需要执行数百次迭代计算,导致推理时间显著高于GAN等单次前向传播模型。在大规模卫星数据处理和实时气象监测场景下,较高的计算开销限制了扩散模型的实际部署能力。虽然潜在空间扩散模型在一定程度上提高了生成效率,但仍存在生成质量与推理速度之间的权衡问题。因此,未来研究可围绕少步采样扩散模型、一致性蒸馏、流匹配等方向展开,进一步降低计算复杂度,实现高质量与高效率兼顾的夜间VIS-NIR补全框架。

4 结 论

本文围绕夜间气象卫星VIS-NIR波段观测缺失问题,构建了一个面向夜间卫星缺失波段信息补全

的VIS-NIR-70K数据集,并基于代表性的生成式模型评估了该数据集的有效性。针对夜间缺乏太阳反射辐射导致VIS-NIR产品无法连续获取的问题,本文基于Himawari-8/9 AHI多波段观测数据,筛选白天成对的MIBT与VIS-NIR样本,并结合观测几何角度和水陆掩膜信息,构建了包含62,290对训练样本和5,446对测试样本的大规模多波段配对数据集,为夜间跨波段补全提供了标准化数据基础。在实验评估方面,选取CycleGAN、Pix2Pix、DDPM、DiT-2S、DiT-2B、DiT-2L和RefDiff等代表性方法进行系统性评估。结果表明,MIBT观测能够为VIS-NIR缺失波段重建提供有效条件信息,其中原始图像空间扩散模型在结构保持、纹理恢复和分布一致性方面表现更优。在TCI生成、低云区域和台风场景中,扩散模型也展现出更好的复杂云系刻画能力。多波段消融实验表明,完整tbb 7~16是VIS-NIR生成性能提升的主要信息来源,SAZ与Landsea联合输入可以确保生成结果的稳定性。不同生成模型的对比实验表明,原始图像空间扩散能够更充分保留云边缘、局部亮度和细粒度纹理。夜间缺失波段全图补全实验进一步说明,基于VIS-NIR-70K训练的模型具备夜间多波段信息恢复的应用潜力。总而言之,VIS-NIR-70K能够有效支撑不同生成范式的统一训练与评估,缓解了夜间卫星缺失波段信息补全领域的的数据资源不足,为全天时地气系统观测和生成式遥感研究提供了数据基础。

同时,现有生成式算法仍存在一定局限,多波段生成框架以单波段独立建模为主,没有考虑不同波段物理的先验差异知识。模型对MIBT中水汽、窗口红外和CO₂吸收等先验信息挖掘存在局限性。其次,现有评价指标难以精确衡量云边缘、纹理和气象结构一致性。后续研究可从波段自适应建模、物理先验融合、遥感专用评价指标和高效生成框架等方向进一步推进。

参考文献(References)

- Bessho K, Date K, Hayashi M, Ikeda A, Imai T, Inoue H, et al. 2016. An introduction to himawari-8/9-japan's new-generation geostationary meteorological satellites. *Journal of the Meteorological Society of Japan. Series II*, 94 (2) : 151-183 [DOI: 10.2151/jmsj.2016-009]

- Ge B Y, Yang L K, Chen X F, Li Z Q, Mei X D and Liu L. 2018. Study on aerosol optical depth retrieval over land from Himawari-8 data based on dark target method. *Journal of Remote Sensing*, 22(01): 38-50 (葛邦宇, 杨磊库, 陈兴峰, 李正强, 梅笑冬, 刘李. 2018. 暗目标法的himawari-8静止卫星数据气溶胶反演. *遥感学报*, 22(01): 38-50) [DOI: 10.11834/jrs.20187033]
- Han K-H, Jang J-C, Ryu S, Sohn E-H and Hong S. 2022. Hypothetical visible bands of advanced meteorological imager onboard the geostationary korea multi-purpose satellite-2a using data-to-data translation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15: 8378-8388 [DOI: 10.1109/JSTARS.2022.3210143]
- Harder P, Jones W K, Lguensat R, Bouabid S, Fulton J, Quesada-Chac'on D, et al. 2020. Nightvision: Generating nighttime satellite imagery from infra-red observations. *arXiv*, abs/2011.07017: [DOI: 10.48550/arXiv.2011.07017]
- Ho J, Jain A and Abbeel P. 2020. Denoising diffusion probabilistic models//*Proceedings of the International Conference on Neural Information Processing Systems (NeurIPS)*. : 6840-6851
- Hong D, Li C, Yokoya N, Zhang B, Jia X, Plaza A, et al. 2026. Hyperspectral imaging. *Nature Reviews Methods Primers*, 6: 19 [DOI: 10.1038/s43586-026-00470-x]
- Isola P, Zhu J-Y, Zhou T and Efros A A. 2017. Image-to-image translation with conditional adversarial networks//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. : 1125-1134
- Jung H, Moon I J, Kim D H and Lee J. 2025. Automated tropical cyclone intensity estimation through integration of historical information and self-diagnosis using ai models. *IEEE Transactions on Geoscience and Remote Sensing*, 63: 1-12 [DOI: 10.1109/TGRS.2025.3632317]
- Kim J-H, Ryu S, Jeong J, So D, Ban H-J and Hong S. 2020. Impact of satellite sounding data on virtual visible imagery generation using conditional generative adversarial network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13: 4532-4541 [DOI: 10.1109/JSTARS.2020.3013598]
- Kim K, Kim J-H, Moon Y-J, Park E, Shin G, Kim T, et al. 2019. Nighttime reflectance generation in the visible band of satellites. *Remote Sensing*, 11(18): 2087 [DOI: 10.3390/rs11182087]
- Kim Y, Han K-H, Park J and Hong S. 2025. Nighttime ground radar-like rainfall estimation using virtual visible and near-infrared and real infrared bands of advanced meteorological imager. *IEEE Transactions on Geoscience and Remote Sensing*, 63: 1-13 [DOI: 10.1109/TGRS.2025.3559686]
- Li C, Cao J and Zhang X. 2020. Design and implementation of an infrared image generative model//*Proceedings of the International Conference on Artificial Intelligence and Computer Applications (ICAICA)*. : 1338-1345 [DOI: 10.1109/ICAICA50127.2020.9182563]
- Liu S, Yu Z, Liu Y, Zhang Z, Shi C and Xiao B. 2026. Cafrans: Content-aware fusion transformer for ground-based remote sensing cloud detection. *International Journal of Applied Earth Observation and Geoinformation*, 146: 105044 [DOI: 10.1016/j.jag.2025.105044]
- Ma R and Letu H. 2026. Cloud, atmospheric radiation and renewal energy application (care) version 1.2 shortwave cloud radiative forcing from himawari-8 and fy-4a: Algorithm development and preliminary validation. *IEEE Transactions on Geoscience and Remote Sensing*, 64: 1-10 [DOI: 10.1109/TGRS.2026.3669180]
- Peebles W and Xie S. 2023. Scalable diffusion models with transformers//*Proceedings of the IEEE/CVF International Conference on Computer Vision ICCV*. : 4195-4205
- Wang M and Jiang L. 2025. Recovery of pixels with extremely turbid waters and intensive floating algae from false cloud masking in satellite ocean color remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, 137: 104408 [DOI: 10.1016/j.jag.2025.104408]
- Xu Z, Li J, Cheng S, Rui X, Zhao Y, He H, et al. 2025. Deep learning for wildfire risk prediction: Integrating remote sensing and environmental data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 227: 632-677 [DOI: 10.1016/j.isprsjprs.2025.06.002]
- Yao J, Du P, Zhao Y, and Wang Y. 2025. Simulating nighttime visible satellite imagery of tropical cyclones using conditional generative adversarial networks. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 18: 12616-12633 [DOI: 10.1109/JSTARS.2025.3567074]
- Yuan Q, Shen H, Li T, Li Z, Li S, Jiang Y, et al. 2020. Deep learning in environmental remote sensing: Achievements and challenges. *Remote Sensing of Environment*, 241: 111716 [DOI: 10.1016/j.rse.2020.111716]
- Zeng Z, Zhang S, Ke L, Sheng H, Liu D, Liu S, et al. 2026. Calipso-guided dual-branch network with spectral - spatial attention for himawari-8 sea fog detection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 19: 11027-11041 [DOI: 10.1109/JSTARS.2026.3673119]
- Zhang B S, Ke F, An Z Y, Yu X N, Han Z W and Zhao F. 2026. Himawari-8/9 satellite fire detection integrating multi-scale context and Mamba temporal modeling. *Journal of Image and Graphics*, 1-19 (张本硕, 柯枫, 安志勇, 于晓宁, 韩忠伟, 赵峰. 2026. 融合多尺度上下文与 Mamba 时序建模的 Himawari-8/9 卫星火点检测. *中国图象图形学报*: 1-19) [DOI: 10.11834/jig.260158]
- Zhang C J, Geng B F, Ma L M and Lu X Q. 2025a. Estimation of tropical cyclone size by combining sequential infrared satellite images with multitask deep learning. *IEEE Transactions on Geoscience and Remote Sensing*, 63: 1-15 [DOI: 10.1109/TGRS.2024.3522582]
- Zhang C J, Xu L Y, Ma L M, Lu X Q, He P, Wei X, et al. 2025b. Tropical cyclone intensity forecasting using multimodal data fusion guided by sequential infrared satellite images. *IEEE Transactions*

on Geoscience and Remote Sensing, 63: 1-14 [DOI: 10.1109/TGRS.2025.3616026]

Zhang L, Gao L, Ye Q, Liu N, Zhang J, Li S, et al. 2025c. Assessment of the himawari-9 downward surface shortwave radiation (dssr) product in china under different cloud and aerosol scenarios. Solar Energy, 292: 113429 [DOI: 10.1016/j.solener.2025.113429]

Zhao Z M, Gao L R, Chen D, Yue A Z, Chen J B, Liu D S, et al. 2019. Development of satellite remote sensing and image processing platform. Journal of Image and Graphics, 24(12): 2098-2110 (赵忠明, 高连如, 陈东, 岳安志, 陈静波, 刘东升, 等. 2019. 卫星遥感及图像处理平台发展. 中国图象图形学报, 24(12): 2098-2110) [DOI: 10.11834/jig.190450]

Zhou T, Zhang F, Fu H, Pan B, Zhang R, Lu F, et al. 2025. Lighting the night with generative artificial intelligence. arXiv: 2506.22511 [DOI: 10.48550/arXiv.2506.22511]

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 IEEE/CVF International Conference on Computer Vision (ICCV). : 2223-2232

作者简介

张文博,男,博士,主要研究方向为图像可控生成与语义分割。E-mail: zhangwb227@njust.edu.cn

陈强,通信作者,男,教授,主要研究方向为医学图像处理与智能影像分析。E-mail: chen2qiang@njust.edu.cn

张光阳,男,硕士,主要研究方向为卫星遥感。E-mail: zhang-guangyang17@nudt.edu.cn

黄渊万,男,硕士,主要研究方向为医学图像可控生成。E-mail: huangyuanwan@njust.edu.cn

黄伟权,男,硕士,主要研究方向为扩散模型效率优化。E-mail: 1244118349@qq.com

李孝辉,男,博士,主要研究方向为图像可控生成与术后预测。E-mail: xiaohuilai@163.com

周稚麟,男,博士,主要研究方向为影像智能分析与语义分割。E-mail: zzlbook@163.com

杨平吕,男,副教授,主要研究方向为卫星遥感。E-mail: ypl@nudt.edu.cn

周则明,男,教授,主要研究方向为大气物理学与大气环境。E-mail: zhou_zeming@nudt.edu.cn