

# 基于参数高效扩散微调的少样本参考图驱动图像补全

杨诗远<sup>1</sup>, 顾崢<sup>2</sup>, 郝文月<sup>1</sup>, 汪毅<sup>1</sup>, 蔡怀宇<sup>1</sup>, 陈晓冬<sup>1</sup>

<sup>1</sup>天津大学精密仪器与光电子工程学院光电信息技术教育部重点实验室, 中国天津市, 300072

<sup>2</sup>南京大学计算机软件新技术国家重点实验室, 中国南京市, 210008

**摘要:** 文本到图像的扩散模型在图像生成方面展现了卓越的能力, 并已广泛应用于图像补全任务。尽管文本提示能够为有条件的图像补全提供直观指导, 但用户往往希望通过提供参考图像为特定对象补全个性化外观。然而, 现有的参考图驱动图像补全方法难以实现高保真度的补全效果。为解决这一问题, 我们基于预训练的文本驱动图像补全模型提出一种即插即用的低秩适配 (LoRA) 模块。该模块通过少样本微调学习参考图像的特定特征, 显著提升了对自定义参考图像的拟合能力, 并且无需在大规模数据集上进行大量训练。此外, 引入 GPT-4V 提示词和先验噪声初始化技术, 进一步提升补全结果的保真度。简而言之, 去噪扩散过程首先从由复合参考-背景图像派生的初始噪声开始, 进而由 GPT-4V 从参考图中生成的丰富提示词引导后续生成过程。大量实验表明, 我们的方法在定性和定量指标上都达到目前最高水平, 为用户提供了一个具有更强定制化能力的参考图驱动图像补全工具。

**关键词:** 扩散模型; 图像补全; 参考图驱动; 少样本微调  
<https://doi.org/10.1631/FITEE.2400395>