基于扩散模型的高效遥感图像可控生成和应用

陈振源
指导老师：刘仁义, 张丰

浙江大学地球科学学院

June 19, 2025

大纲

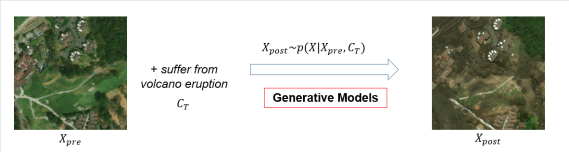
背景介绍与研究目标

背景介绍

- 自然灾害不可预测，遥感影像对救灾关键
- 现有高分辨率双时相灾害遥感图像稀缺
- 传统与文本引导方法难以生成真实、语义一致的遥感影像

研究目标

- 利用扩散模型进行图像生成（编辑）
- 使用灾前图像和文本描述进行可控生成



研究目标示意图

数据集

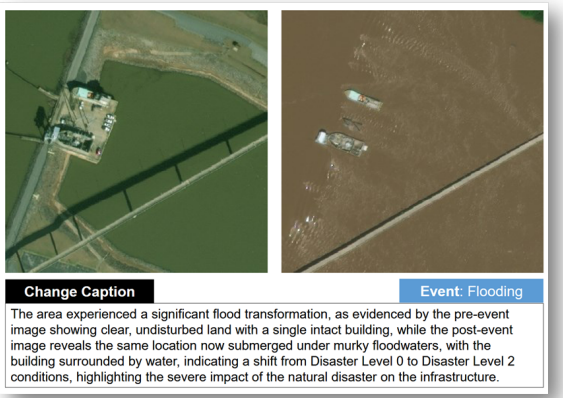


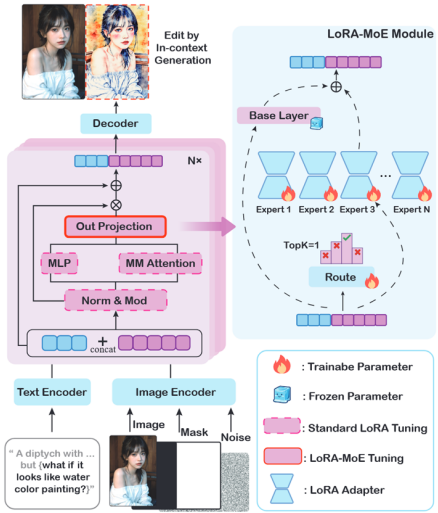
Table 4: The 31 disaster events from RSCC dataset.

Source	Disaster type	Disaster event	Event date
xBD	Earthquake	Mexico City earthquake	Sep 19, 2017
	Wildfire	Portugal wildfires	Jun 17-24, 2017
	Wildfire	Santa Rosa wildfires	Oct 8-31, 2017
	Wildfire	Carr wildfire	Jul 23-Aug 30, 2018
	Wildfire	Woolsey fire	Nov 9-28, 2018
	Wildfire	Pinery fire	Nov 25-Dec 2, 2018
	Volcano	Lower Puna volcanic eruption	May 23-Aug 14, 2018
	Volcano	Guatemalu Fuego volcanic eruption	Jun 3, 2018
	Storm	Tuscaloosa, AL tornado	Apr 27, 2011
	Storm	Joplin, MO tornado	May 22, 2011
	Storm	Moore, OK tornado	May 20, 2013
	Storm	Hurricane Matthew	Sep 28-Oct 10, 2016
	Storm	Hurricane Florence	Sep 10-19, 2018
	Flooding	Monsoon in Nepal, India, Bangladesh	Jul-Sep, 2017
	Flooding	Hurricane Harvey	Aug 17-Sep 2, 2017
	Flooding	Hurricane Michael	Oct 7-16, 2018
	Flooding	Midwest US floods	Jan 3-May 31, 2019
	Tsunami	Indonesia tsunami	Sep 18, 2018
	Tsunami	Sunda Strait tsunami	Dec 22, 2018
EBD	Hurricane	Hurricane Delta	Oct 8, 2020
	Hurricane	Hurricane Dorian	Sep 1, 2019
	Hurricane	Hurricane Ida	Oct 29, 2021
	Hurricane	Hurricane Laura	Aug 26, 2020
	Hurricane	Hurricane Irma	Sep 6, 2017
	Hurricane	Hurricane Ian	Sep 26, 2022
	Tornadoes	Texas Tornadoes	Mar 23, 2022
	Volcanic Eruption	Mount Semeru Eruption	Dec 4, 2021
	Volcanic Eruption	ST. Vincent Volcano	Apr 9, 2021
	Volcanic Eruption	Tonga Volcano	Jan 15, 2022
	Earthquake	Turkey Earthquake	Feb 6, 2023
	Flooding	Pakistan Flooding	Jul 26, 2022

RSCC 数据集 **[Chen et al., 2025]**。该数据集包含 **62,351** 对事件前后遥感图像（**512×512** 像素）及详细的变化描述文本。包含 **31** 次不同灾害事件。

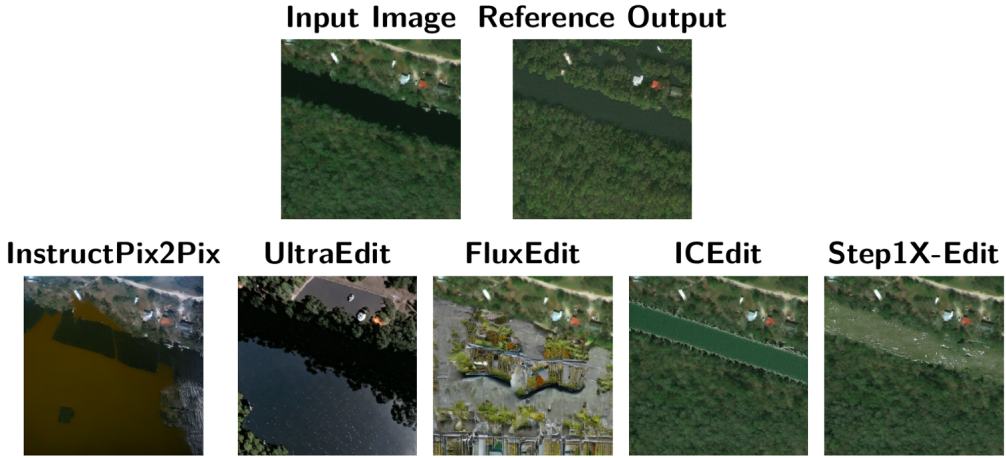
基线模型框架

- InstructPix2Pix [Brooks et al., 2023]
- UltraEdit [Zhao et al., 2024]
- Step1X-Edit [Liu et al., 2025]
- ICEdit [Zhang et al., 2025]
- ICEdit 框架：通过上下文提示词实现零样本指令遵循，无需结构性修改。LoRA-MoE 混合微调策略：高效适应与动态专家路由，提升灵活性，无需大规模再训练。推理时早期筛选方法：利用视觉-语言模型（VLM）在早期选择更优初始噪声，提升编辑质量。



现有图像编辑框架 **ICEdit** 架构示意图 **[Zhang et al., 2025]**

初步实验结果（基线模型）



不同图像编辑模型（未经微调）的结果。

提示词：“在最近的卫星图像对比中，观测到显著变化：水位明显上升，淹没了之前可见的部分陆地，极大地改变了景观外观。”

下游任务应用：数据增广方法对比

传统数据增强方法

- **主要技术**
  - 几何变换, 颜色扰动, 噪声注入
  - Cutout [DeVries and Taylor, 2017]
  - CutMix [Yun et al., 2019]
  - Copy-Paste [Ghiasi et al., 2021]
- **相关研究 [Steiner et al., 2022] 讨论了各种数据增广方式对 ViT 分类精度提升效果。**

合成数据增广在图像分类中的应用

前人研究发现 [He et al., 2023] 通过生成合成遥感图像及其变化描述，可用于扩充有限的真实数据集，提升下游图像分类模型的泛化能力。

当前进展

当前完成项 ✓

- 数据集准备完成
- 基线模型训练代码开发
- 基线模型推理代码开发
- 单模型最小化训练测试验证

进行中任务

- 评估指标开发（部分完成）
- 大规模消融实验设计
- 下游任务应用（基于数据增强的图像分类）

参考文献 I

Brooks, T., Holynski, A., and Efros, A. A. (2023).
InstructPix2Pix: Learning to follow image editing instructions.
In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 18392–18402.

Chen, Z., Wang, C., Zhang, N., and Zhang, F. (2025).
RSCC: A Large-Scale Remote Sensing Change Caption Dataset for Disaster Events.

DeVries, T. and Taylor, G. W. (2017).
Improved Regularization of Convolutional Neural Networks with Cutout.

Ghiasi, G., Cui, Y., Srinivas, A., Qian, R., Lin, T.-Y., Cubuk, E. D., Le, Q. V., and Zoph, B. (2021).
Simple Copy-Paste Is a Strong Data Augmentation Method for Instance Segmentation.
pages 2918–2928.

参考文献 II

He, R., Sun, S., Yu, X., Xue, C., Zhang, W., Torr, P., Bai, S., and Qi, X. (2023).
Is synthetic data from generative models ready for image recognition?

Liu, S., Han, Y., Xing, P., Yin, F., Wang, R., Cheng, W., Liao, J., Wang, Y., Fu, H., Han, C., Li, G., Peng, Y., Sun, Q., Wu, J., Cai, Y., Ge, Z., Ming, R., Xia, L., Zeng, X., Zhu, Y., Jiao, B., Zhang, X., Yu, G., and Jiang, D. (2025).
Step1X-Edit: A Practical Framework for General Image Editing.

Steiner, A. P., Kolesnikov, A., Zhai, X., Wightman, R., Uszkoreit, J., and Beyer, L. (2022).
How to train your ViT? Data, Augmentation, and Regularization in Vision Transformers.

参考文献 III

Yun, S., Han, D., Oh, S. J., Chun, S., Choe, J., and Yoo, Y. (2019).
CutMix: Regularization Strategy to Train Strong Classifiers With Localizable Features.
pages 6023–6032.

Zhang, Z., Xie, J., Lu, Y., Yang, Z., and Yang, Y. (2025).
In-Context Edit: Enabling Instructional Image Editing with In-Context Generation in Large Scale Diffusion Transformer.

Zhao, H., Ma, X., Chen, L., Si, S., Wu, R., An, K., Yu, P., Zhang, M., Li, Q., and Chang, B. (2024).
UltraEdit: Instruction-based Fine-Grained Image Editing at Scale.