GFM4D: Conditional Diffusion Model for Post-event Satellite Image Generation

Sakura Chen, *Student Member, IEEE,* Zhenyuan Chen, *Member, IEEE,*

***Abstract*—Satellite images play a crucial role in humanitarian assistance and disaster response. With the advancement of powerful computer vision algorithms, disaster evaluation can be performed automatically and remotely. However, disasters are unpredictable, creating challenges for data collection. Additionally, accurate analysis requires high-resolution images, increasing the pressure on data acquisition. To address these issues, we present the Generative Foundation Model for Disaster (GFM4D), which is pretrained on existing pre- and post-event remote sensing images, enabling it to synthesize satellite images of events as desired by users. Our model demonstrates strong robustness and reliability in image generation. Notably, we also introduce the Global-scale Building Damage (GBD) dataset, comprising over 1 million images worldwide. Pre-trained segmentation models further trained on our dataset show improved performance.**

***Index Terms*—Object detection, building damage assessment, post-event, pre-trained model, generative model**

I. Introduction

NATURAL disasters such as hurricanes, tornadoes, earthquakes, and floods cause extensive damage to infrastructure and lead to significant loss of life and property, amounting to billions of dollars [1]. With the increasing frequency and severity of these events, it is essential to swiftly create maps and assess the extent of destruction to effectively allocate resources and direct first responders [1]. Satellite images are vital in humanitarian assistance and disaster response efforts [2]. The use of advanced computer vision algorithms enables the automatic and remote evaluation of disaster impacts [3], [4], [5], [6]. This technological advancement significantly enhances the ability to respond promptly and accurately to natural disasters, thereby improving the overall efficiency of disaster management operations.

However, the availability of remote sensing image datasets for disasters is limited. High-resolution images are essential to accurately reflect the affected areas, but there are few satellites capable of capturing such detailed images [7], [4], and their long cycle times hinder timely acquisition of disaster scenarios. An alternative approach is to use unmanned aerial vehicles (UAVs) [6], [8]. Nevertheless, the unpredictable and sudden nature of disasters makes it challenging to deploy UAVs in time to capture every event. The combination of limited satellite availability and the logistical difficulties of UAV deployment underscores the need for more efficient and reliable methods of acquiring disaster imagery. Additionally, existing datasets for post-event satellite images are geographically biased. Most of the available data is concentrated in regions with higher economic resources and advanced infrastructure, leaving many disaster-prone areas underrepresented. This geographic bias limits the ability to develop and validate models that are globally applicable and effective in diverse environments.

Manuscript last updated 25 June 2024

To address these challenges, the advancement of image generation technologies, particularly diffusion models, has shown great promise in various fields. These models excel in creating high-quality, detailed images, which are critical for accurate analysis and response. However, the application of diffusion models in remote sensing, specifically for generating post-event satellite images, remains underexplored.

To facilitate the shortage of high-resolution post-event satellite images, we propose a specialized diffusion model tailored for remote sensing applications. This model is designed to generate high-fidelity synthetic images that accurately depict post-disaster scenarios, addressing the limitations of existing data sources.

The primary contributions of our research are as follows:

1) We propose the first-ever satellite image generation model specifically designed for disaster scenarios, based on an off-the-shelf diffusion model.
2) Our model exhibits strong capabilities in both temporal generation and temporal inpainting, accurately reconstructing and predicting post-disaster scenarios based on pre-event conditions.
3) Using our model, we have created a synthesized global-scale dataset of high-resolution images, which can be widely utilized by the AI community for training and validation purposes.
4) Experimental results indicate that training on our synthesized dataset significantly improves model performance, demonstrating the robustness and effectiveness of our dataset.

II. Related Work

A. Remote Sensing Image Generation

In the field of remote sensing, image generation models have recently garnered increasing attention.

B. Change Detection for Natural Disaster

References

[1] "Flood Vulnerability Detection Challenge Using Multiclass Segmentation," Feb. 2022.

[2] R. Gupta, B. Goodman, N. Patel, R. Hosfelt, S. Sajeev, E. Heim, J. Doshi, K. Lucas, H. Choset, and M. Gaston, "Creating xBD: A Dataset for Assessing Building Damage from Satellite Imagery," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2019, pp. 10–17.

[3] D. Zhao, J. Lu, and B. Yuan, “See, Perceive, and Answer: A Unified Benchmark for High-Resolution Postdisaster Evaluation in Remote Sensing Images,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1–14, Apr. 2024.

[4] J. Doshi, S. Basu, and G. Pang, “From Satellite Imagery to Disaster Insights,” in *Neural Information Processing Systems*, 2018.

[5] T. G. J. Rudner, M. Rußwurm, J. Fil, R. Pelich, B. Bischke, V. Kopačková, and P. Biliński, “Multi3Net: Segmenting Flooded Buildings via Fusion of Multiresolution, Multisensor, and Multitemporal Satellite Imagery,” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, pp. 702–709, July 2019.

[6] M. Rahnemoonfar, T. Chowdhury, A. Sarkar, D. Varshney, M. Yari, and R. R. Murphy, “FloodNet: A High Resolution Aerial Imagery Dataset for Post Flood Scene Understanding,” *IEEE Access*, vol. 9, pp. 89 644–89 654, 2021.

[7] G. Christie, N. Fendley, J. Wilson, and R. Mukherjee, “Functional Map of the World,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6172–6180.

[8] C. Kyrkou and T. Theocharides, “Deep-Learning-Based Aerial Image Classification for Emergency Response Applications Using Unmanned Aerial Vehicles,” in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, June 2019, pp. 517–525.