CHULA ENGINEERING
Foundation toward Innovation

Chula
Chulalongkorn University

COLLEGE OF COMPUTING
KHON KAEN UNIVERSITY

Asia Oceania Geosciences Society
www.asiaoceania.org

IG18-A001

Generative AI with Kernel-adaptive Diffusion for High-resolution Satellite Image Inpainting Across Urban, Forest, and Agricultural Landscapes

Teerapong Panboonyuen

https://kaopanboonyuen.github.io/KAO

CONTEXT

AI in Geospatial Applications

- Earth-observation archives grow by petabytes a year, from optical and radar satellites alike
- Generative models increasingly condition directly on geospatial metadata — location, time, resolution
- This unlocks tasks beyond classification: temporal synthesis, super-resolution, and inpainting

Temporal Generation

Synthesize imagery across time from metadata alone

Super-Resolution

Sharpen low-resolution sensor bands into fine detail

Inpainting (our focus)

Recover missing regions from clouds, gaps, occlusion

Satellite Imagery Is Often Incomplete

- Clouds, haze, and sensor faults leave persistent gaps — e.g., the LANDSAT 7 scan-line corrector failure
- Missing pixels break downstream pipelines: road extraction, land-use classification, change detection

Goal: recover missing regions while preserving geometric structure and semantic consistency.

Every Landscape Fails Differently

Urban

Sharp edges, building boundaries

Forest

Stochastic canopy texture

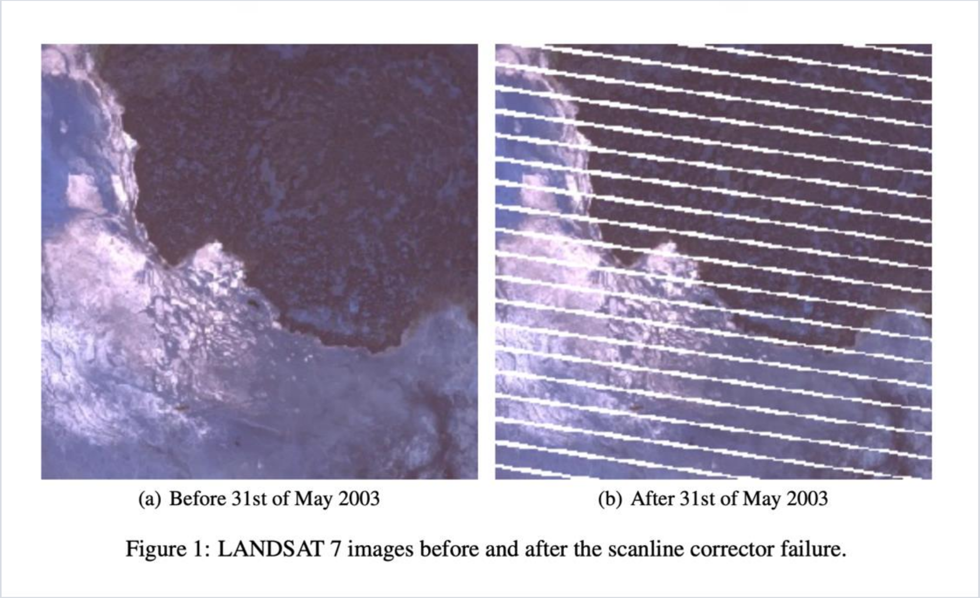
Agriculture

Repetitive field patterns

Convolutional Neural Processes for Inpainting

Pondaven et al., 2022 · arXiv:2205.12407

- Meta-learning approach: treats every image as a new task, adapting to its own spatial pattern
- Targets missing data from scan gaps (LANDSAT 7) and cloud cover
- Strong uncertainty modeling — but relies on global statistical aggregation



LANDSAT 7 imagery before / after the scan-line corrector failure

Diffusion Foundation Models for Satellite Imagery

Khanna et al., 2023 · DiffusionSat · arXiv:2312.03606

- Conditions a latent diffusion model on freely available metadata: coordinates, time, resolution
- One foundation model, many tasks: generation, super-resolution, temporal inpainting
- State-of-the-art image quality — but applies the same diffusion kernel everywhere

Published as a conference paper at ICLR 2024

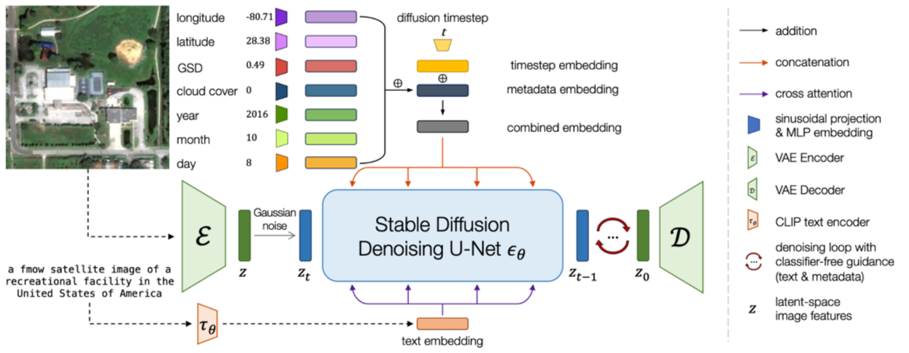


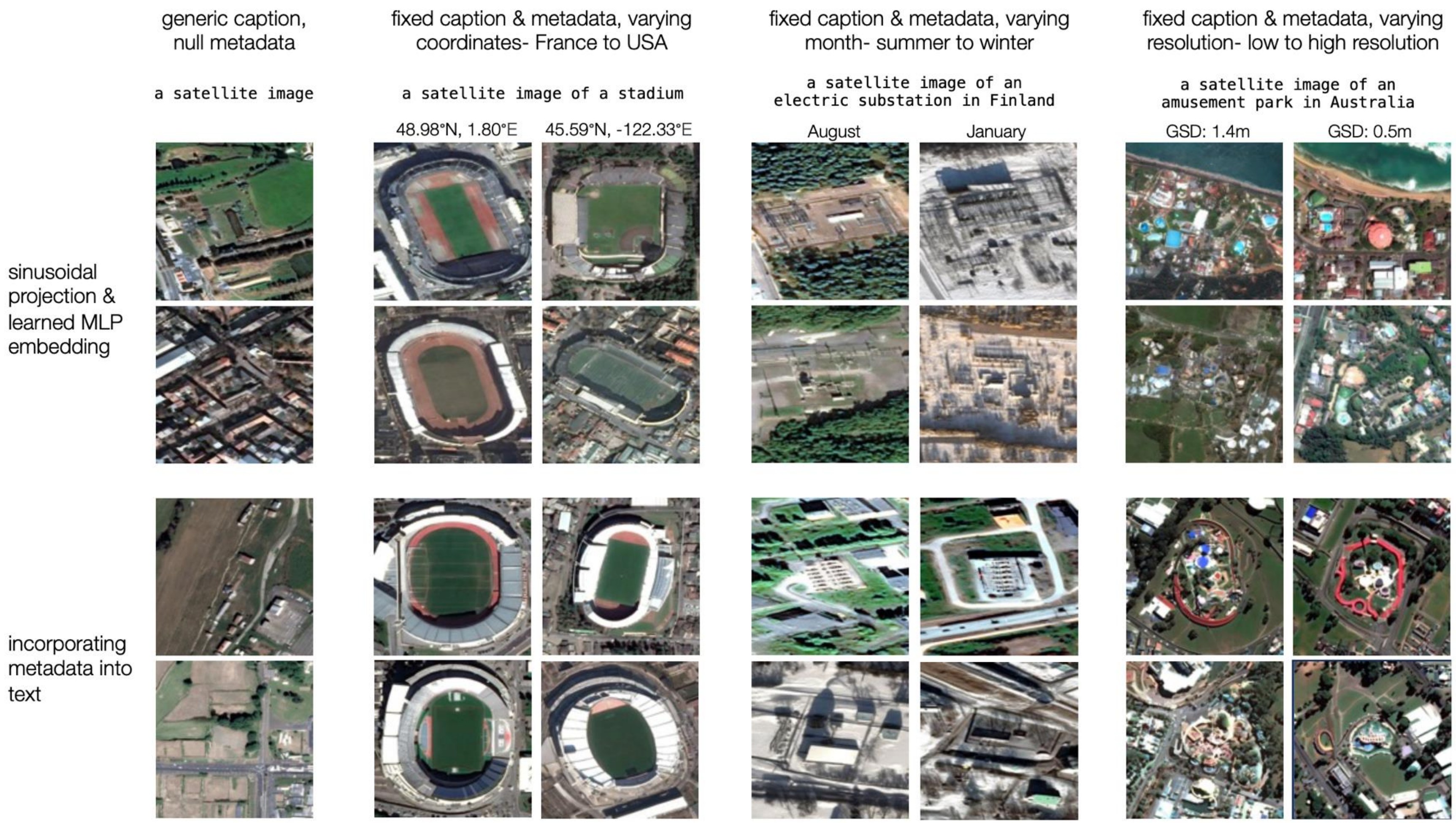
Figure 1: Conditioning on freely available metadata and using large, publicly available satellite imagery datasets shows DiffusionSat is a powerful generative foundation model for remote sensing data.

DiffusionSat: latent diffusion conditioned on geospatial metadata

Before **KAO** —
What has generative AI already learned to do with satellite images?

Change the coordinates, change the world.

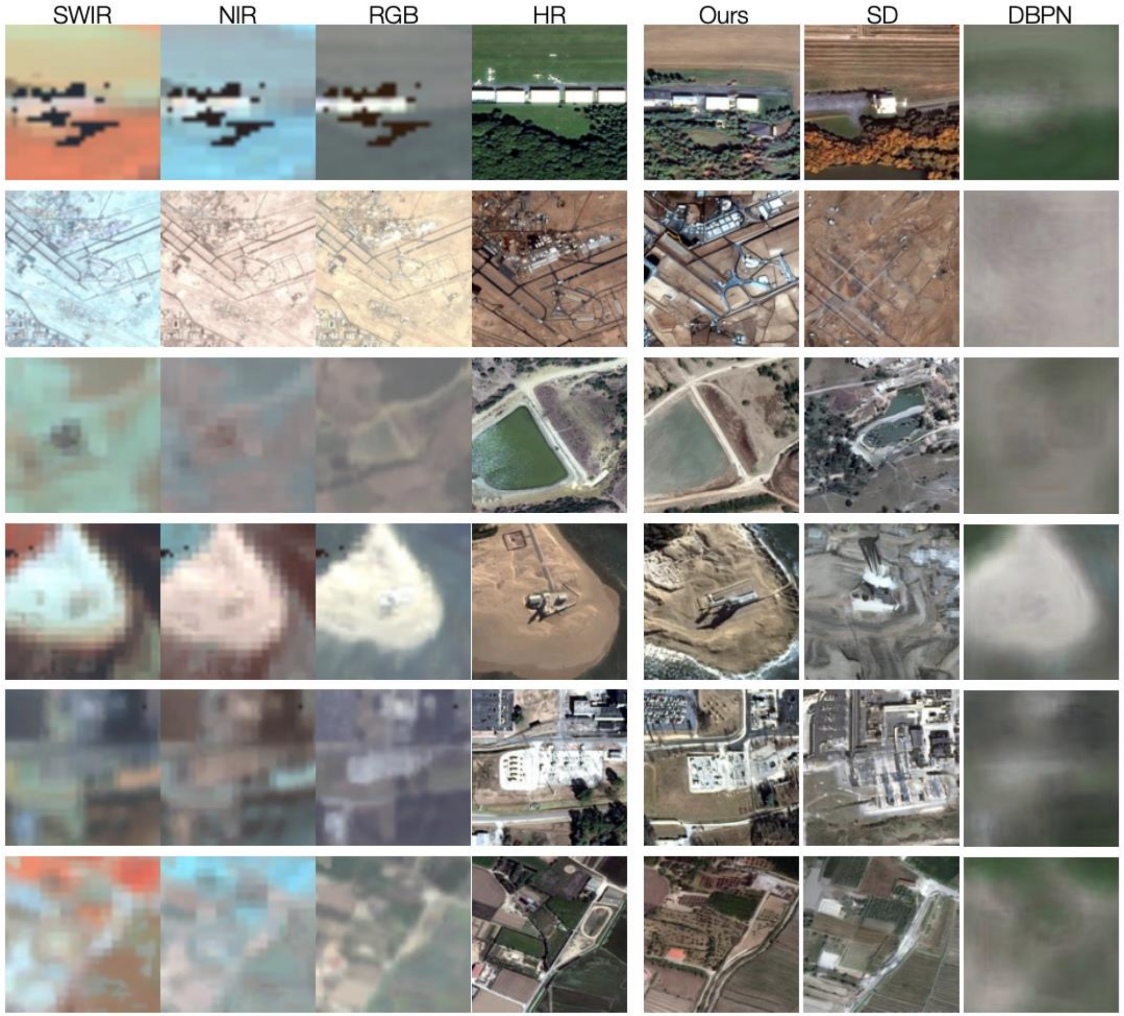
Move a stadium from Paris to the USA — DiffusionSat swaps football for baseball. Snowy latitudes get snow. But push it too far: winter and summer start blurring together.

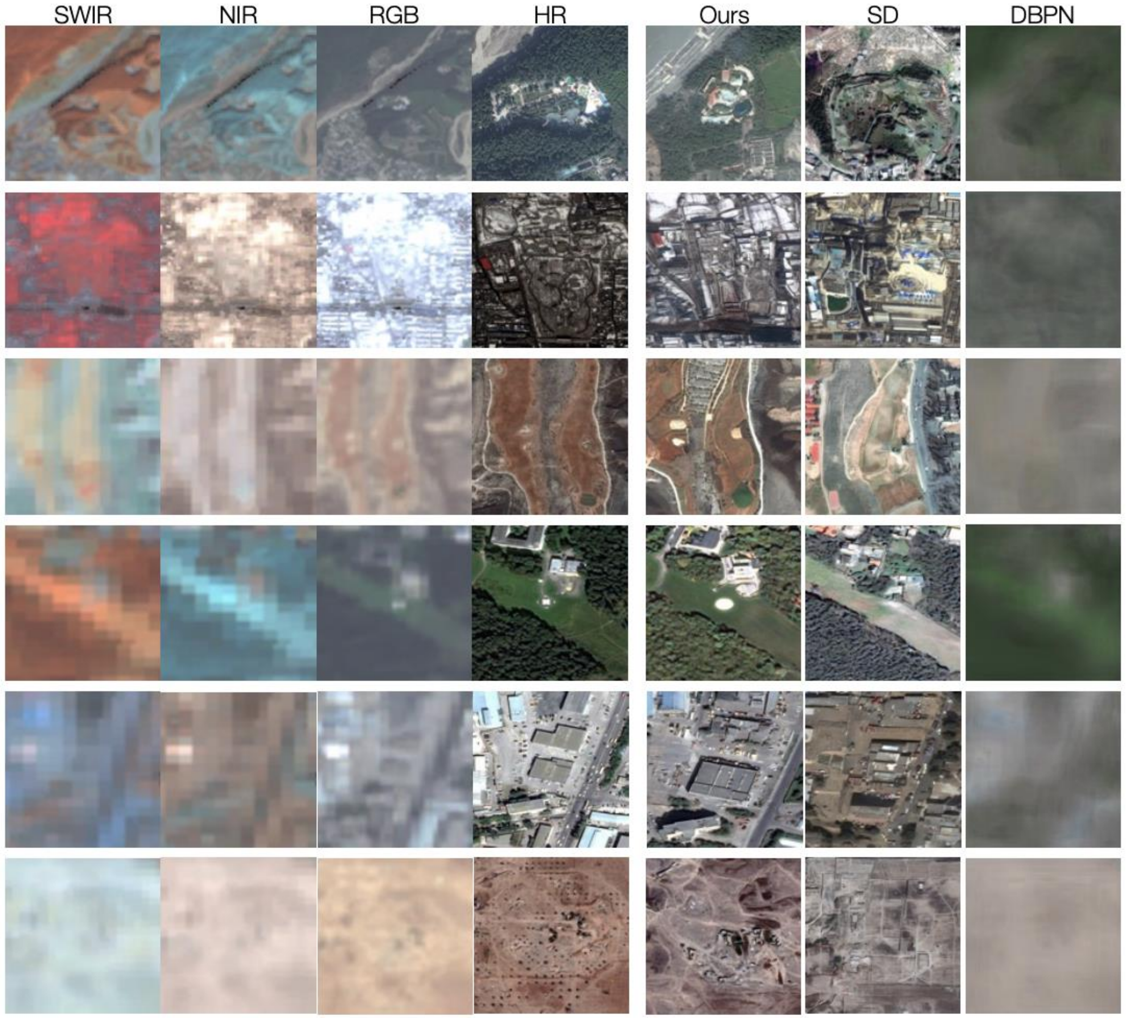


Credit: *Khanna et al., 2023 · DiffusionSat · arXiv:2312.03606*

Can you sharpen what was never sharp?

From coarse Sentinel-2 bands to fine-grained detail — DiffusionSat holds structure where Stable Diffusion just guesses.

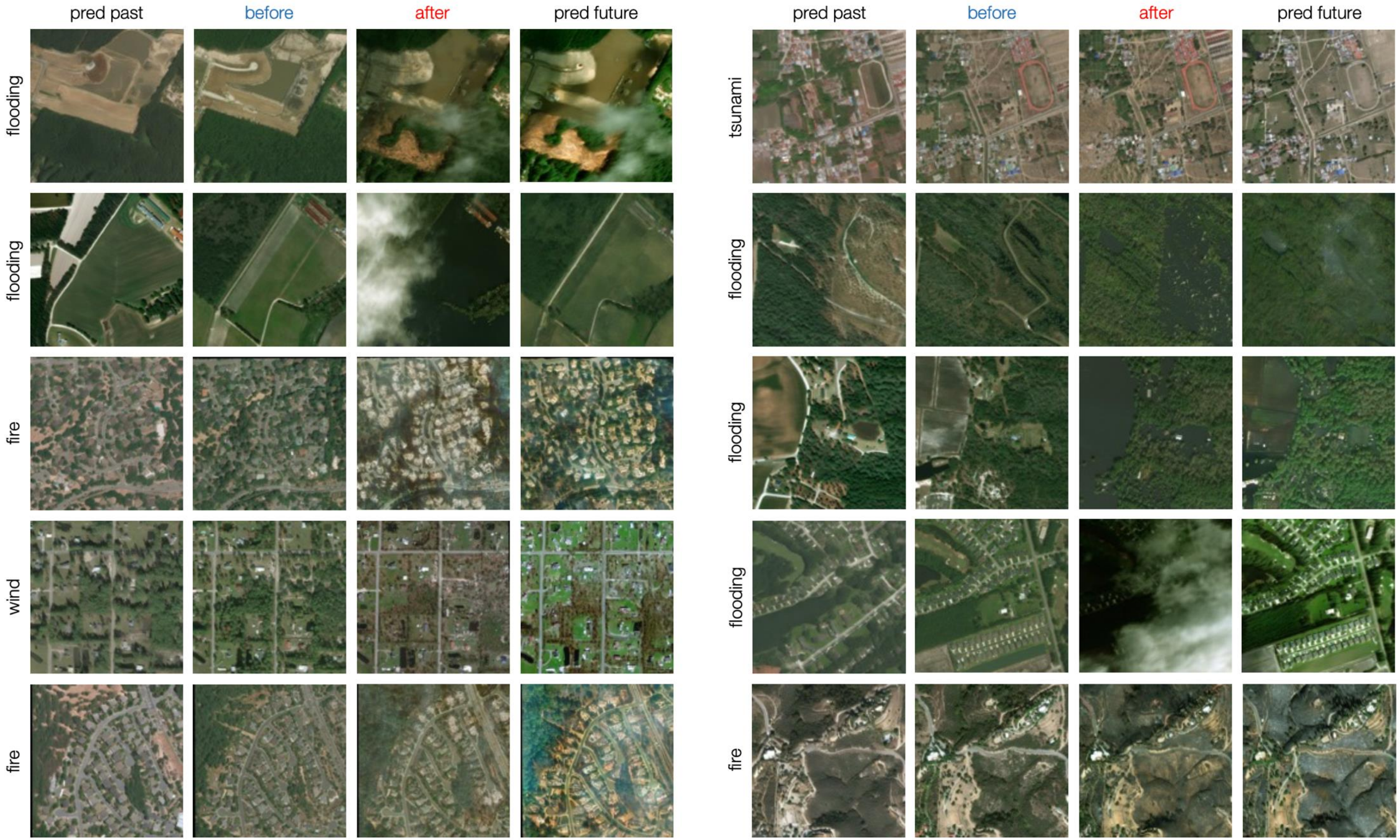




Credit: Khanna et al., 2023 · DiffusionSat · arXiv:2312.03606

Floods. Fires. Missing roofs and roads.

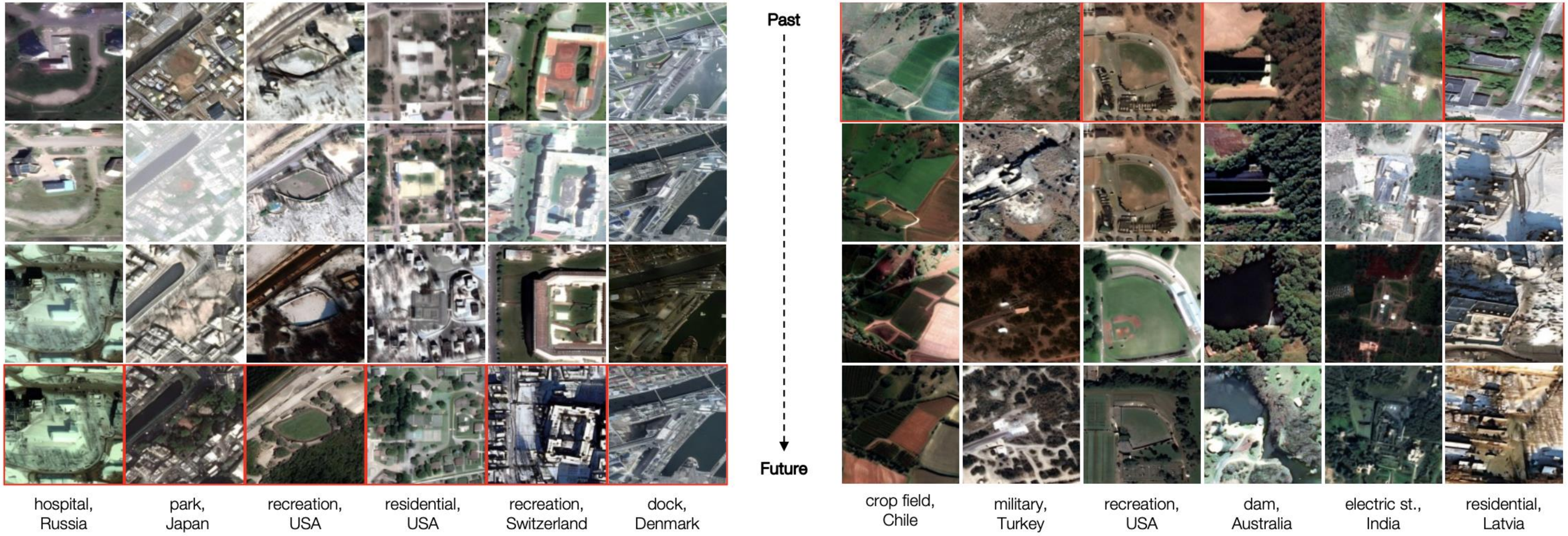
DiffusionSat rebuilds what clouds and disasters erase — conditioning on "before" to predict "after," and back again.



Credit: Khanna et al., 2023 · DiffusionSat · arXiv:2312.03606

What did this place look like a decade ago? What will it look like next?

Given one image and a timestamp, DiffusionSat generates entire histories — forward and backward in time.



Credit: Khanna et al., 2023 · DiffusionSat · arXiv:2312.03606

THE GAP

One Blind Spot Both Approaches Share

Neural Processes

✓ Strong uncertainty modeling

✓ Good for irregular missing data

BUT Relies on global statistical aggregation → over-smoothing, loss of fine structure

Diffusion Foundation Models

✓ Powerful learned image priors

✓ Flexible metadata conditioning

BUT Applies uniform diffusion kernels everywhere → blurred high-frequency regions

Neither approach adapts its process to local spatial complexity.

Why KAO?

So diffusion can already imagine the past, the future, the fine detail. What's missing?

Every one of these results still uses **one fixed kernel for the whole image** — the same denoising behavior for a runway and a rice field.

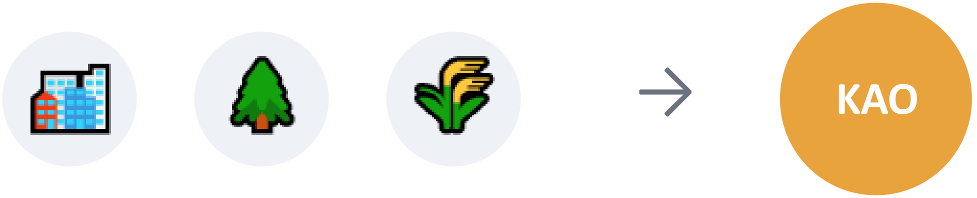
KAO asks a different question: what if the kernel adapted to *where* it's denoising, not just *what* it's denoising?

MOTIVATION

Our Key Insight

What if the diffusion process itself could adapt — region by region — to how complex each part of the scene really is?

That question is the starting point for **KAO** — Kernel-Adaptive Optimization in Diffusion.



KAO in One Picture

Masked Satellite Image → Encoder → Kernel-Adaptive Diffusion Denoising → Latent Conditioning + Explicit Propagation → Decoder → Reconstructed Image

KAO reweights the denoising kernel per region, then fuses known and inferred latents through explicit forward–backward propagation.

CORE INNOVATIONS

Three Building Blocks of KAO

1

Kernel-Adaptive Optimization

Dynamically reweights the diffusion kernel per region: stronger updates on edges, stable denoising on smooth areas.

2

Latent-Space Conditioning

Blends known and inferred latents directly — no retraining, no expensive iterative sampling.

3

Explicit Propagation

A forward–backward fusion step that carries structure from valid regions into occluded ones, even under heavy cloud cover.

WHY IT MATTERS FOR GEODATA

Every Landscape Demands a Different Kernel

Urban

Sharp edges and building boundaries.

Strong local updates preserve straight lines and corners.

Forest

Stochastic, high-frequency canopy texture.

Adaptive smoothing avoids speckle and false detail.

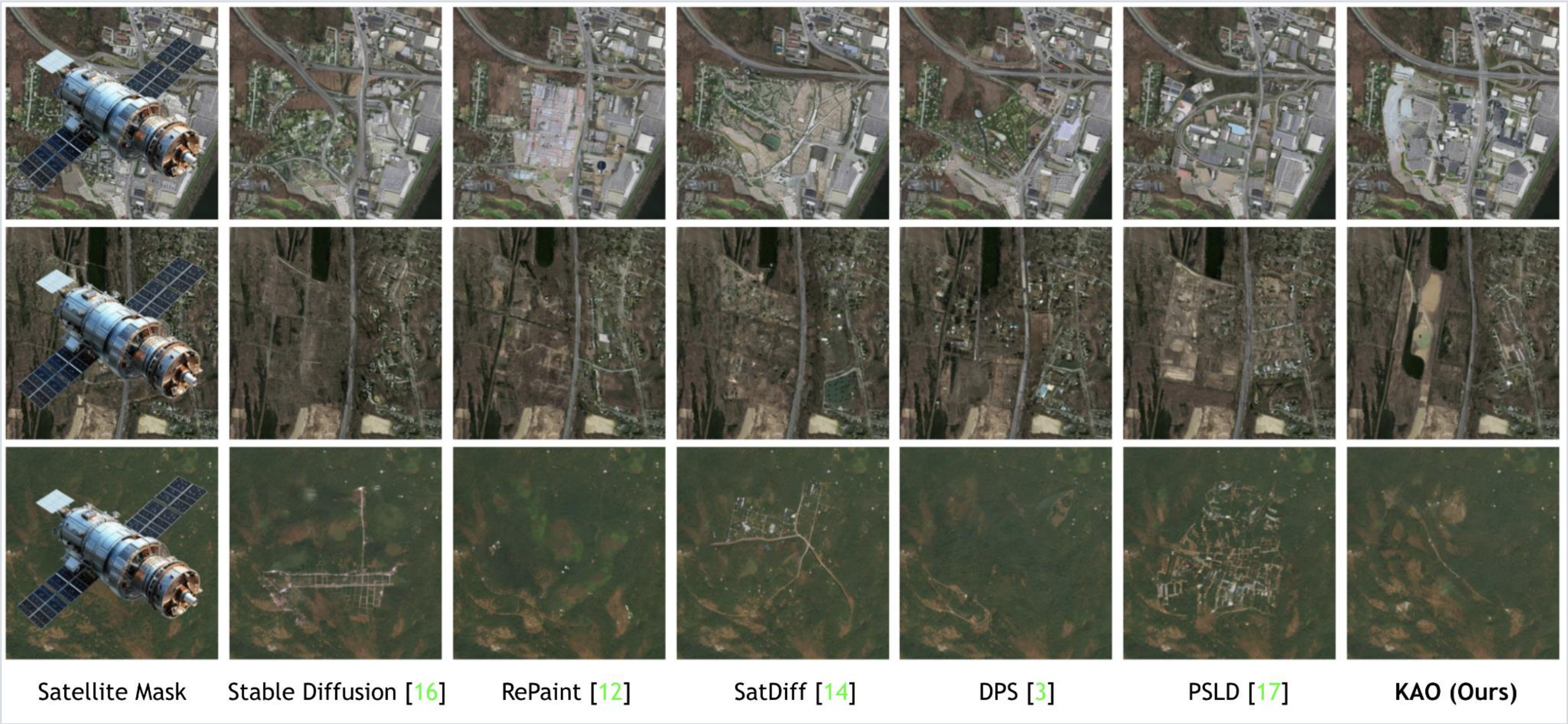
Agriculture

Repetitive, periodic field patterns.

Kernel locks onto periodicity instead of blurring it away.

One fixed kernel cannot serve all three — KAO doesn't need to choose.

Sharper Structure, Fewer Artifacts



Across urban, forest, and agricultural scenes, KAO preserves roads, field boundaries, and canopy texture that baseline diffusion methods blur or hallucinate.

Consistent Gains Across Benchmarks

TABLE I
EVALUATION METRICS FOR SATELLITE IMAGE INPAINTING ON THE MASSACHUSETTS ROADS DATASET AND DEEPGLOBE 2018 DATASET.

	Massachusetts Roads Dataset			DeepGlobe 2018 Dataset			
Method	**FID ↓**	**Prec. ↑**	**Recall ↑**	**Method**	**FID ↓**	**Prec. ↑**	**Recall ↑**
Stable Diffusion [25]	6.98	0.59	0.69	Stable Diffusion [25]	5.62	0.51	0.44
RePaint [16]	6.12	0.65	0.71	RePaint [16]	5.19	0.59	0.47
LatentPaint [13]	4.44	0.71	0.81	LatentPaint [13]	2.55	0.61	0.51
SatDiff [1]	3.99	0.88	0.86	SatDiff [1]	1.98	0.80	0.55
DPS [26]	3.67	0.89	0.87	DPS [26]	1.76	0.82	0.56
PSLD [27]	3.42	0.91	0.89	PSLD [27]	1.65	0.84	0.58
KAO (Ours)	**3.11**	**0.93**	**0.91**	**KAO (Ours)**	**1.42**	**0.88**	**0.63**

Massachusetts Roads & DeepGlobe 2018 — vs. Stable Diffusion, RePaint, LatentPaint, SatDiff, DPS, PSLD

1.42

FID ↓ (best baseline: 1.65)

0.63

Recall ↑ (best baseline: 0.58)

DeepGlobe 2018

TAKEAWAY

KAO: Local Adaptation Meets Efficient, Structure-Aware Diffusion

- Kernel-adaptive diffusion + latent-space conditioning + explicit propagation
- Higher-fidelity, structure-preserving inpainting across heterogeneous landscapes
- Efficient by design: no retraining, no heavy iterative sampling

"From uniform diffusion to spatially intelligent diffusion."

Thank You

Questions & Discussion

Teerapong Panboonyuen

Project page: **https://kaopanboonyuen.github.io/KAO/**

Paper: **arXiv:2511.02462**

Satellite Mask Stable Diffusion [16] RePaint [12]