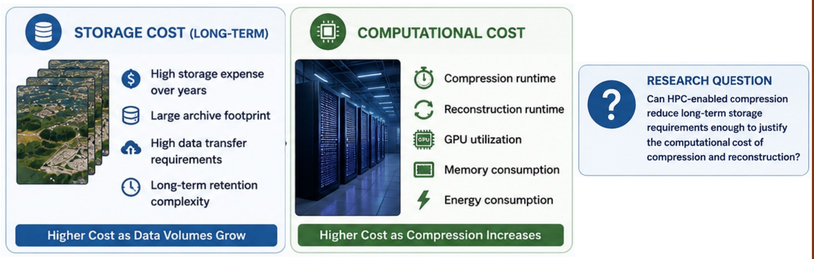
Optimizing HPC Performance for CDC-Based Storage and Reconstruction of High-Resolution Imagery

Jooho Kim[1], Yifan Yang[2], Anish Shakya[3], Jacob Kelly[4]

[1] *Institute for Disaster Resilient Texas (IDRT), Texas A&M University, College Station, TX*
[2] *Department of Geography, Texas A&M University, College Station, TX*
[3] *Department of Marine and Coastal Environmental Science, Texas A&M University, Galveston, TX*
[4] *Department of Computer Science and Engineering, Texas A&M University, College Station, TX*

Introduction

Advances in drone technology have enabled the collection of ultra-high-resolution imagery, but a single mission can generate hundreds of gigabytes of data, creating challenges for storage, transfer, and processing. This study evaluates a conditional diffusion compression (CDC) [1] workflow for drone imagery on the NCSA DeltaAI GH200 system [7], examining how compression ratio and reconstruction tile size affect runtime, GPU memory, image quality, and downstream object detection. The goal is to identify configurations that maximize storage savings and computational efficiency while preserving image quality and computer vision performance.

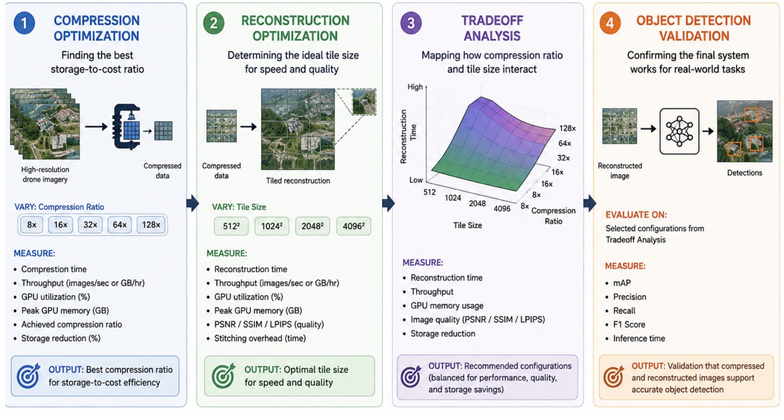


Data Description

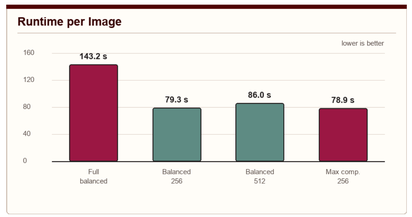
- Study area: Galveston, Texas
- 100 RGB drone images acquired at altitude of approx. 47 m
- Average original size of 8.34 MB, 24-bit color depth
- Spatial resolution: 5472 × 3648 pixels

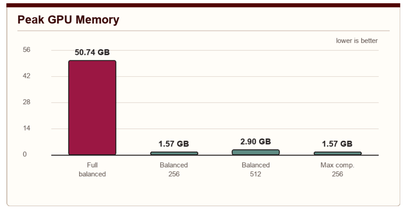


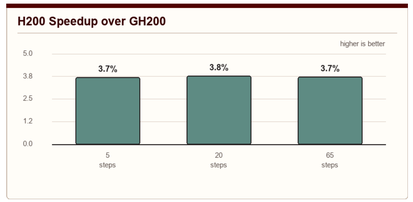
Experiment Design

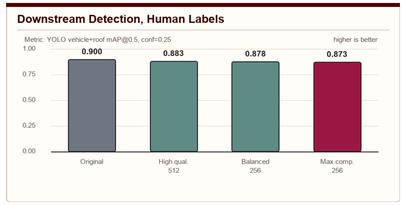


Resource Efficiency and Bottleneck Summary









Measured Proxies, Not Hardware Telemetry

	Full balanced	Balanced 256	Balanced 512	Max comp. 256
Runtime pressure	100%	55%	60%	55%
GPU memory pressure	50.7 GB	1.6 GB	2.9 GB	1.6 GB
Non-GPU overhead	0.2%	5.9%	5.0%	6.0%
Quality-loss proxy	33.29 dB	33.14 dB	33.23 dB	30.89 dB
Detection retention	0.900	0.878	0.883	0.873
SAM roof IoU	1.000	0.899	0.900	0.896

Results

- **Recommended setting:** balanced checkpoint b00064 with 256 x 256 tiling. 512 x 512 is the quality-safe backup.
- **Tiling wins on cost:** 256 x 256 tiling cuts peak GPU memory from ~52 GB to ~1.6 GB and runtime from ~143 s to ~79 s per image.
- **Detection holds up:** YOLOv8s [2] loses only ~1.4% F1 at balanced_256; SAM [3] masks usable on all 50 images.

Reconstruction tradeoff matrix (N50, LPIPS)

Config	Sec/img	GPU GB	Ratio	PSNR	SSIM
High quality 512	88.5	2.96	29.7x	34.19	0.938
Balanced 256 ★	**79.3**	**1.57**	**80.0x**	**33.14**	**0.877**
Balanced 512	86.0	2.90	78.5x	33.23	0.878
Max compression 256	78.9	1.57	139.9x	30.89	0.819

Object detection impact — YOLOv8s vehicle + roof (N50)

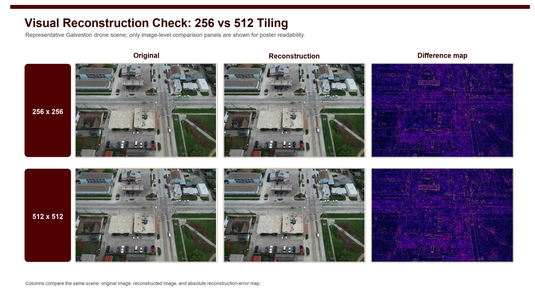
Image set	mAP@.5	mAP@.5:.95	Precision	Recall	F1
Original	0.864	0.607	0.884	0.882	0.883
High quality 512	0.845	0.488	0.893	0.869	0.880
Balanced 256 ★	**0.839**	**0.481**	**0.877**	**0.864**	**0.870**
Max compression 256	0.833	0.475	0.879	0.857	0.868

Baseline compression matrix (full-image)

Role	Ckpt	Ratio	PSNR	SSIM
High quality	b00512	32.67x	34.84	0.9423
Balanced	**b00064**	**93.11x**	**33.73**	**0.8803**
High compression	b00128	165.73x	31.52	0.8281
Baseline	b02048	71.20x	29.45	0.8727

SAM mask stability (vehicle / roof prompts)

Prompt	Config	IoU	Dice	Fail%
Vehicle	HighQual 512	0.735	0.838	0%
Vehicle	**Balanced 256**	**0.735**	**0.838**	**0%**
Vehicle	MaxComp 256	0.733	0.837	0%
Roof	HighQual 512	0.900	0.945	0%
Roof	**Balanced 256**	**0.899**	**0.945**	**0%**
Roof	MaxComp 256	0.896	0.943	0%



Qualitative panel - 256 x 256 vs 512 x 512 tiling

Discussion

- **Why CDC over JPEG2000/WebP:** classical codecs plateau at modest ratios; CDC reaches 80–140× at >30 dB PSNR, at a GPU cost this study quantifies.
- **Why tiling works:** memory scales with tile area, runtime with total pixels, so 256 tiles cut peak memory 33× but runtime only ~45%.
- **Detection is robust:** F1 within ~2% of original (YOLOv8s > YOLO-World [4]); mAP@.5:.95 drops ~21%, so use 512 tiles for fine localization.
- **Scope:** one site, 100 RGB images, one GPU; other scenes, sensors, and scales remain untested.

Conclusion / Future Direction

- Full-image CDC is feasible but memory-heavy; balanced b00064 with 256 tiling is the recommended operating point on DeltaAI.
- **Next (systems):** multi-GPU/multi-node scaling, A100/H100 vs GH200, and a cost model predicting optimal tile size from a GPU memory budget.
- **Next (data):** multispectral and thermal imagery, multi-site datasets, and integration into the IDRT disaster-response pipeline.

References

[1] Yang & Mandt, Lossy image compression with conditional diffusion models, NeurIPS 2023. [2] Jocher et al., Ultralytics YOLOv8, 2023.
[3] Kirillov et al., Segment Anything, ICCV 2023. [4] Cheng et al., YOLO-World: real-time open-vocabulary object detection, CVPR 2024.
[5] Zhang et al., The unreasonable effectiveness of deep features as a perceptual metric (LPIPS), CVPR 2018.
[6] Wang et al., Image quality assessment: from error visibility to structural similarity (SSIM), IEEE TIP 2004. [7] NCSA, DeltaAI user documentation, 2024.