CuMoLoS-MAE

A Masked Autoencoder for Remote Sensing Data Reconstruction

NYU ARTS & SCIENCE

Anurup Naskar, Nathanael Wong Zhixin, Sara Shamekh
Courant Institute of Mathematical Sciences, New York University

Accurate atmospheric profiles are often corrupted by low SNR gates, range folding, and discontinuities. Traditional gap filling blurs fine-scale structure, and most deep models lack calibrated uncertainty. We present **CuMoLoS-MAE**, a curriculum-guided Monte Carlo stochastic-ensemble masked autoencoder that ***(i) restores fine-scale features*** such as updraft and downdraft cores, shear lines, and small vortices, ***(ii) learns a data-driven prior*** over atmospheric vertical profiles, and ***(iii) quantifies pixel-wise uncertainty***. We believe that our high-fidelity, uncertainty-aware workflow improves convection diagnostics, supports real-time data assimilation, and strengthens long-term climate reanalysis.

Model Architecture

Micro-patchified MAE:

- Each 64×64 velocity slice is tokenized into 2×2 micro-patches.
- 12-layer ViT encoder operates on the visible tokens
- 4-layer decoder reconstructs the full field.

Simple Curriculum Training:

- Masking ratio *r* is set to 50% for the first 5 epochs
- *r* increased via a cosine ramp to 70% by epoch 30 and held fixed
- Optimisation uses MSE error computed on the hidden pixels.

Monte-Carlo Ensembling:

- Random masks generated for each forward pass at inference time
- Masking, encoding, and decoding process is repeated *N = 50* times
- Ensemble is aggregated to compute the mean and pixel-wise σ

Training Procedure

Preprocessing:

- Apply an SNR filter (intensity >= 0.005)
- Clamp valid velocities to the range [-5, 5] m s^{-1}
- Extract non-overlapping 64×64 patches.

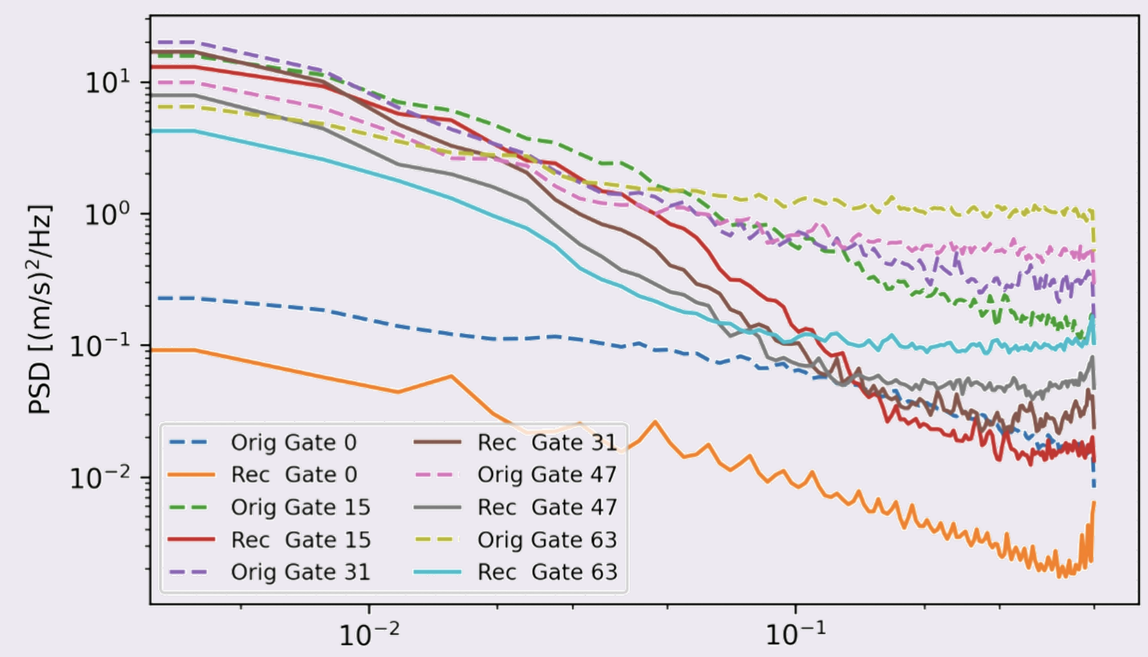
Training runs for 500 epochs on a single NVIDIA A100 GPU using AdamW (base learning rate 1.5e-4 * 32/256, weight decay 0.05) with a batch size of 32. We use a cosine learning-rate schedule with warmup aligned to the mask-ratio curriculum period.

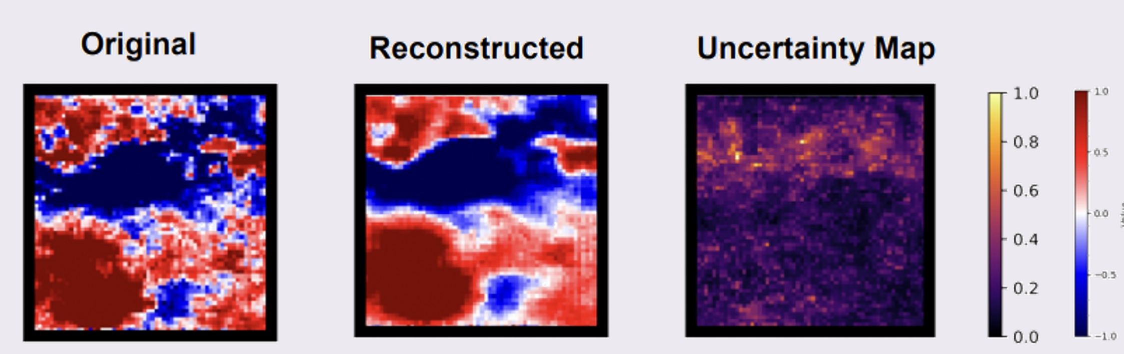
Results

Method	PSNR (dB) ↑	SSIM ↑	MSE ↓	FID ↓	Spectral Fidelity ↑
8x8 Mean-Filter	23.41	0.4950	0.5186	5.13	91.67%
CVAE	26.70	0.4190	0.4036	3.28	80.21%
DnCNN (Noise2Void)	23.09	0.6466	0.6232	0.12	36.46%
U-Net (Noise2Void)	27.70	0.7016	0.2581	0.44	49.48%
CuMoLoS-MAE	**29.45**	**0.7857**	**0.1854**	**1.87**	**93.75%**

Window	PSNR (dB) ↑	SSIM ↑	MSE ↓	FID ↓	Spectral Fidelity ↑
64 x 64	29.45	0.7857	0.1854	1.87	93.75%
128 x 64	30.11	0.7697	0.2253	3.73	87.50%
256 x 64	28.55	0.6103	0.3205	5.50	38.02%

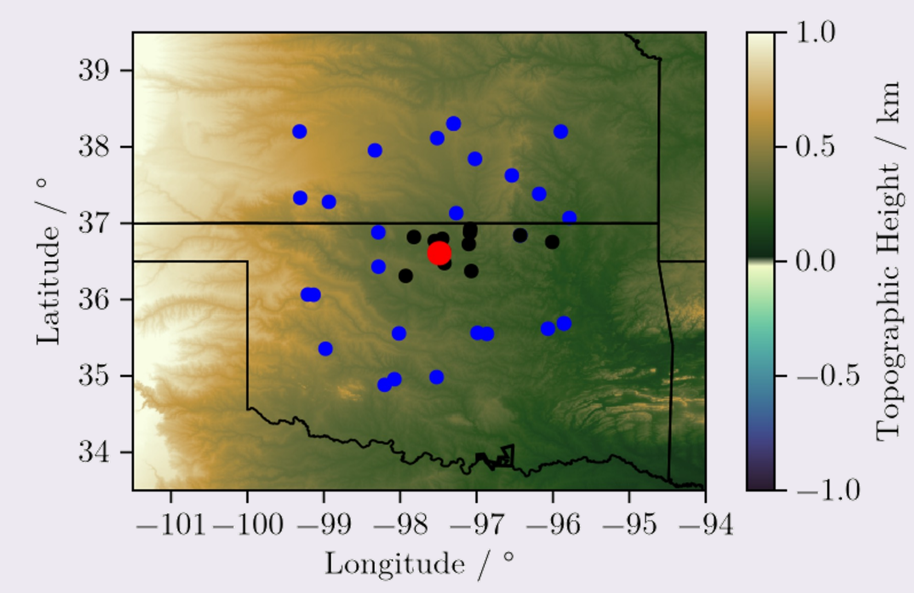
Configuration	PSNR (dB) ↑	SSIM ↑	MSE ↓	FID ↓	Spectral Fidelity ↑
Without Curriculum	29.45	0.7857	0.1854	1.87	93.75%
With Curriculum	28.90	0.7868	0.2106	1.88	93.23%

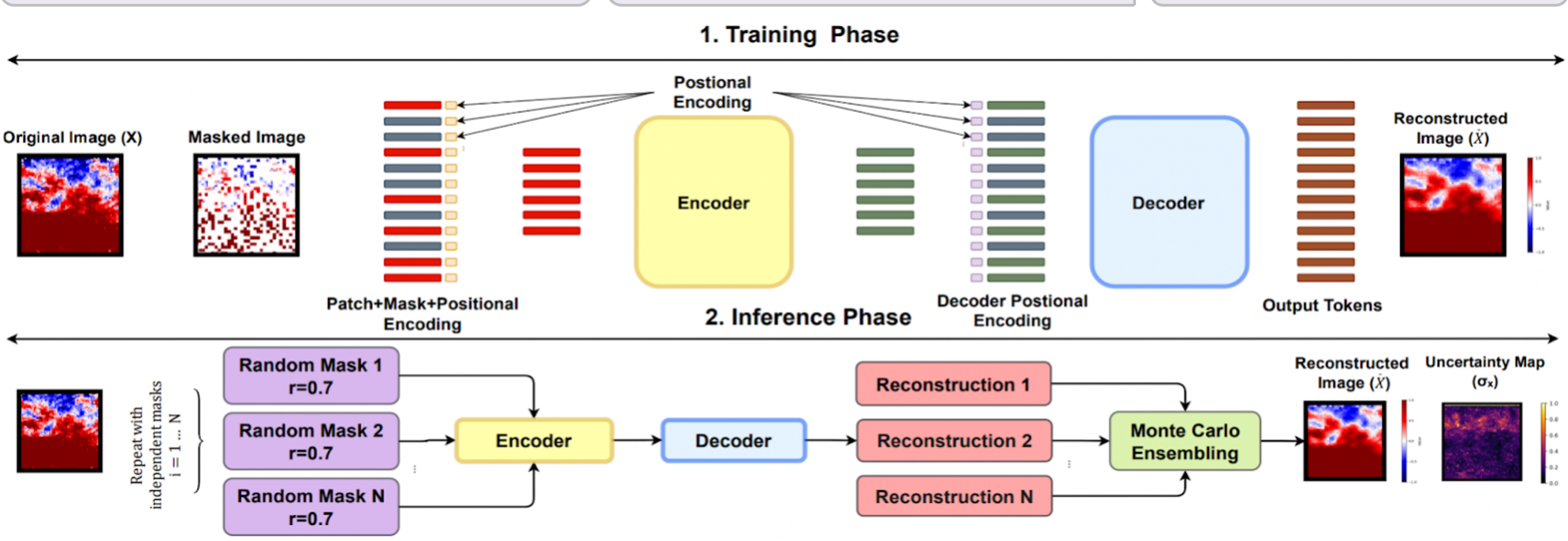




Datasets

Location: Southern Great Plains
Instrument: Doppler-Lidar
Variable: Vertical Velocity (*w*) vertical profiles
Timespan: June 2011
Frequency: 1s

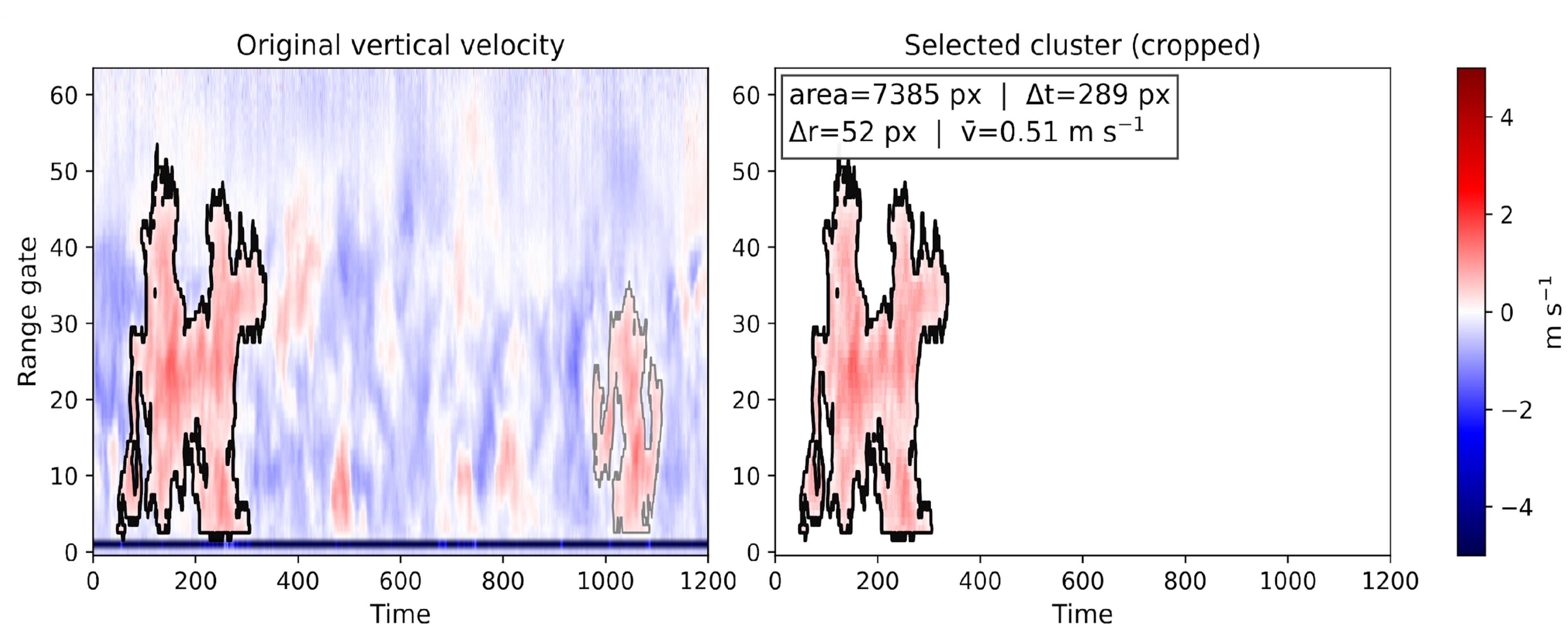




K-Means Clustering and Object Identification

On our denoised/smoothed output from CuMuLoS, we perform K-means clustering to classify regions of positive, neutral and negative vertical velocities, and using 4-connected-components methodology to identify updrafts and downdrafts.

- **Updraft-Downdraft-Neutral labeling:** Run K-means with K=3 on the reconstructed velocity field and map to following: down (negative), neutral (near zero), up (positive).
- **Objects & metrics:** For each class, label 4-connected components and compute the following: (a) duration, (b) height and (c) area
- **Meteorology gates & summaries:** Keep downdraft if duration ≥ 60 px, height ≥ 3 px, area ≥ 150 px; keep updraft if duration ≥ 120 px, height ≥ 4 px, area ≥ 300 px.



Conclusion and Future Work

CuMoLoS-MAE reconstructs Doppler-lidar vertical velocity with high fidelity while providing a per-pixel uncertainty map that reliably predicts where errors will be larger. By restoring fine-scale cores and preserving low-frequency energy, the method improves convection diagnostics and is suitable for confidence-aware ingestion into weather and climate pipelines. A simple mask-ratio curriculum speeds convergence, and Monte Carlo masking yields calibrated uncertainty without extra labels. In future, we will expand training to larger multi-season datasets, evaluate real-time operational use, and use uncertainty to weight observations in operational assimilation.

NEURAL INFORMATION PROCESSING SYSTEMS

AGU25
New Orleans, LA | 15–19 December 2025

This work is supported by the Department of Energy, Grant No. DE-SC0025456 under the **UNSHADE** (Understanding and Modelling of Shallow to Deep Convection Transition) project.
The NYU Greene cluster provided the computing resources for this work.