

Representation Learning of Near-Surface Atmospheric Fields



Martin Vozár^{1,2,3}

¹Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava ²IBL Software Engineering ³Slovak Hydrometeorological Institute

This study adapts the Single-Head Vision Transformer (SHViT) architecture for representation learning of near-surface atmospheric fields within an autoencoder framework. Generator variants incorporating spectral filtering are developed, and a hybrid adversarial autoencoder / GAN optimization scheme is explored. Evaluated on the CERRA reanalysis (5.5 km) and Met Office UKV (2.2 km) datasets, the proposed architectures achieve compression ratios of $256\times$ to $1024\times$ with mean absolute errors of $\sim 0.3\text{--}0.5\text{ K}$ for temperature and $10\text{--}25\text{ Pa}$ for mean sea level pressure. Adversarial training improves capture of small-scale features at the cost of increased reconstruction error, while geographic generalization to unseen regions remains an open challenge.

1 Methodology

SHViT Autoencoder Architecture

- **Encoder:** SHViT [1] with an additional attention operation in the first stage to capture global dependencies in atmospheric fields.
- **Generator variants:**
 - SHViTGAN – direct inverse of the encoder.
 - SHhGAN – spectral-hierarchical design inspired by StyleGAN2 [2], enforcing multi-scale consistency via frequency-domain filtering.

Hybrid Optimization

- Alternating adversarial autoencoder (AAE) and GAN steps at a 4:1 ratio combines latent-space regularization with reconstruction stability.
- Softplus-based adversarial loss is selected for training robustness.

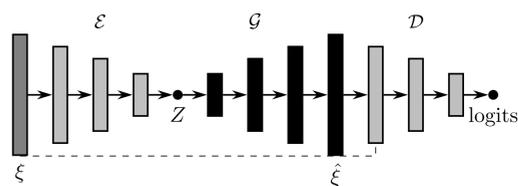


Figure 1: Encoder (\mathcal{E}), Generator (\mathcal{G}), and Discriminator (\mathcal{D}). Dashed line: direct input to \mathcal{D} .

2 Data & Preprocessing

A. CERRA Reanalysis [3]. European domain at 5.5 km resolution (2000–2019); regions span maritime, alpine, continental, and Mediterranean climates; geographically split into training/validation (black) and test (white) areas.

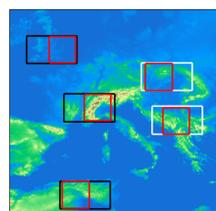


Figure 2: CERRA regions.

B. UKV (Met Office) [4]. UK domain at 2.2 km (non-hydrostatic), 2-year rolling archive; suitable for small-scale convective events.

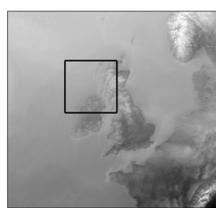


Figure 3: UKV domain.

Variables: temperature (2 m), relative humidity (2 m), MSLP, wind speed (10 m).

Normalization: sample-wise, channel-wise robust scaling (median / IQR).

3 Quantitative Results

- **Compression:** $256\times$ to $1024\times$ ratios achieved (latent sizes 1024 and 256, respectively).
- **Accuracy:** MAE $\approx 0.3\text{--}0.5\text{ K}$ (temperature), $10\text{--}25\text{ Pa}$ (MSLP).
- PCA baselines yield lower pixel-wise error overall, but deep learning models provide flexible, nonlinear backbones adaptable to downstream tasks.

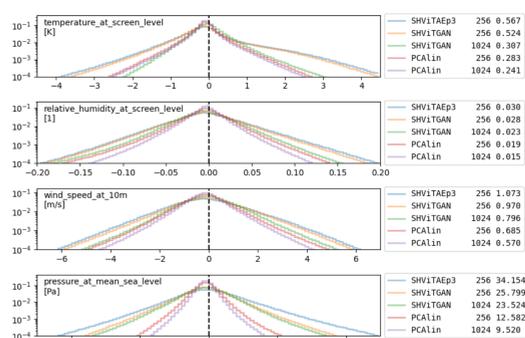


Figure 4: Error distribution histogram comparing the methods across used variables. Legend contains MAE value.

4 Visual Reconstruction

UKV – High-Resolution Fixed Region

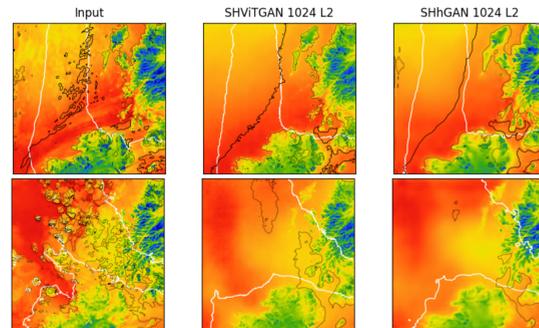


Figure 5: UKV (2.2 km) reconstruction. Large- and meso-scale structure is well preserved; convective-scale detail is smoothed.

5 Discussion & Limitations

Key Observations

- **Generalization:** a significant performance gap persists between validation (seen) and test (geographically distinct) regions, suggesting partially region-specific representations.
- **Small scales:** convective phenomena below $\sim 30\text{ km}$ remain challenging; models tend to smooth extreme local values such as gust fronts and cold pools.

Limitations

- Experiments restricted to four surface-level variables; vertical atmospheric structure is needed for full convective prediction.
- The hybrid AAE-GAN scheme is experimental and sensitive to hyperparameter choices; reproducibility requires careful tuning.

6 Conclusions & Future Work

- SHViT provides a computationally efficient backbone for atmospheric representation learning, achieving $256\text{--}1024\times$ compression with acceptable reconstruction quality at synoptic and mesoscales.
- Spectral-hierarchical generators (SHhGAN) show promise for improved domain generalization.

Future directions include integration with Graph Neural Network architectures [5] for hybrid grid-graph representations, ensemble-based probabilistic training on convection-permitting NWP data [6], and transfer learning across different NWP model domains.

References

- [1] Seokju Yun and Youngmin Ro. *SHViT: Single-Head Vision Transformer with Memory Efficient Macro Design*. 2024. arXiv: 2401.16456 [cs.CV]. URL: <https://arxiv.org/abs/2401.16456>.
- [2] Tero Karras et al. *Analyzing and Improving the Image Quality of StyleGAN*. 2020. arXiv: 1912.04958 [cs.CV]. URL: <https://arxiv.org/abs/1912.04958>.
- [3] S. Schimanke et al. *CERRA sub-daily regional reanalysis data for Europe on single levels from 1984 to present*. Accessed on 26-03-2025. 2021. DOI: 10.24381/cds.622a565a. URL: <https://doi.org/10.24381/cds.622a565a>.
- [4] Met Office. *Met Office UK Deterministic (UKV) 2 km*. <https://registry.opendata.aws/met-office-uk-deterministic>. 2-year rolling archive. Accessed: 2025-05-01. 2025.
- [5] Ryan Keisler. *Forecasting Global Weather with Graph Neural Networks*. 2022. arXiv: 2202.07575 [physics.aos-ph]. URL: <https://arxiv.org/abs/2202.07575>.
- [6] Clément Brochet et al. "Enriching Operational High-Resolution Ensemble Forecasts with Generative Deep Learning". In: *Artificial Intelligence for the Earth Systems* 4.1 (2025). DOI: 10.1175/AIES-D-24-0058.1. URL: <https://doi.org/10.1175/AIES-D-24-0058.1>.