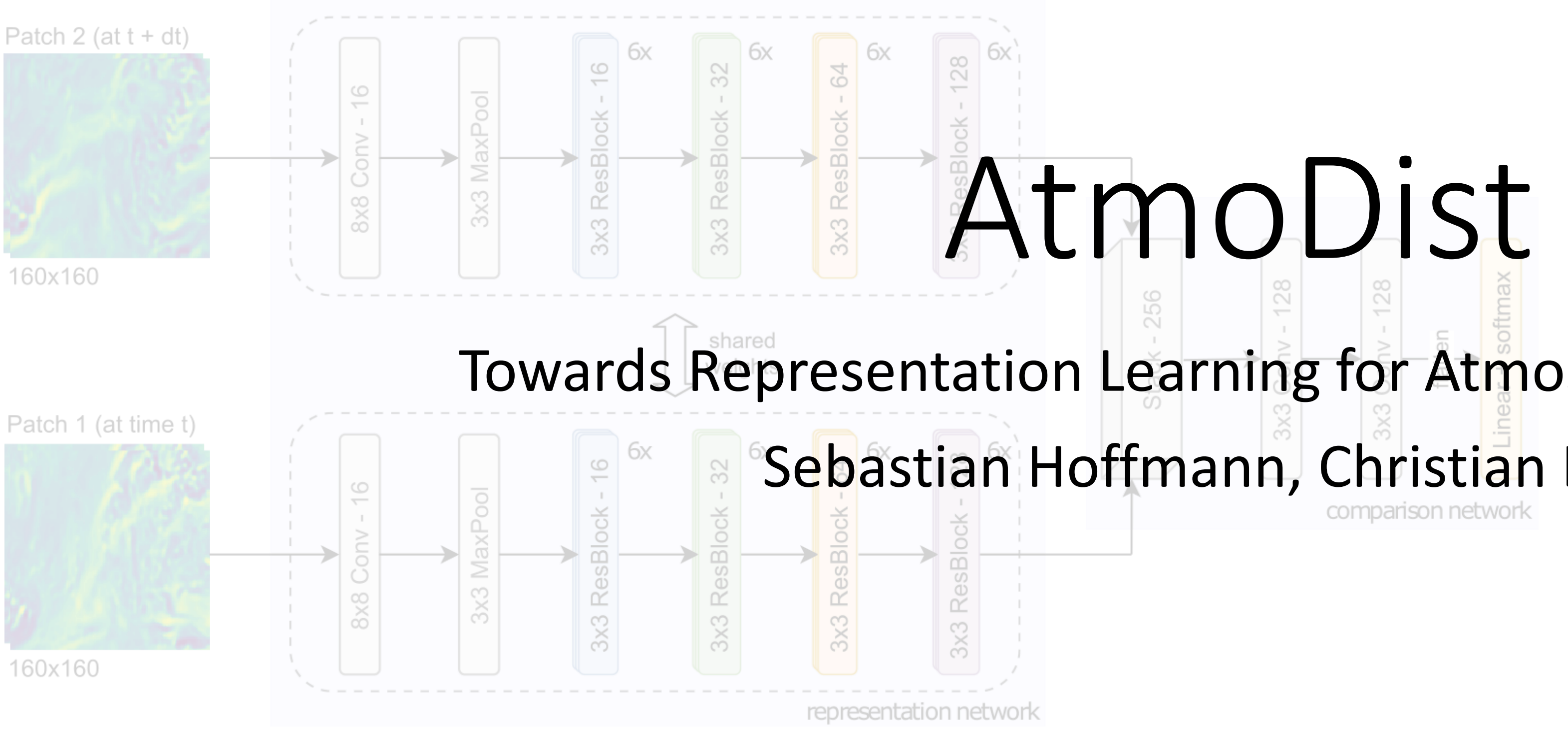
AtmoDist

Towards Representation Learning for Atmospheric Dynamics

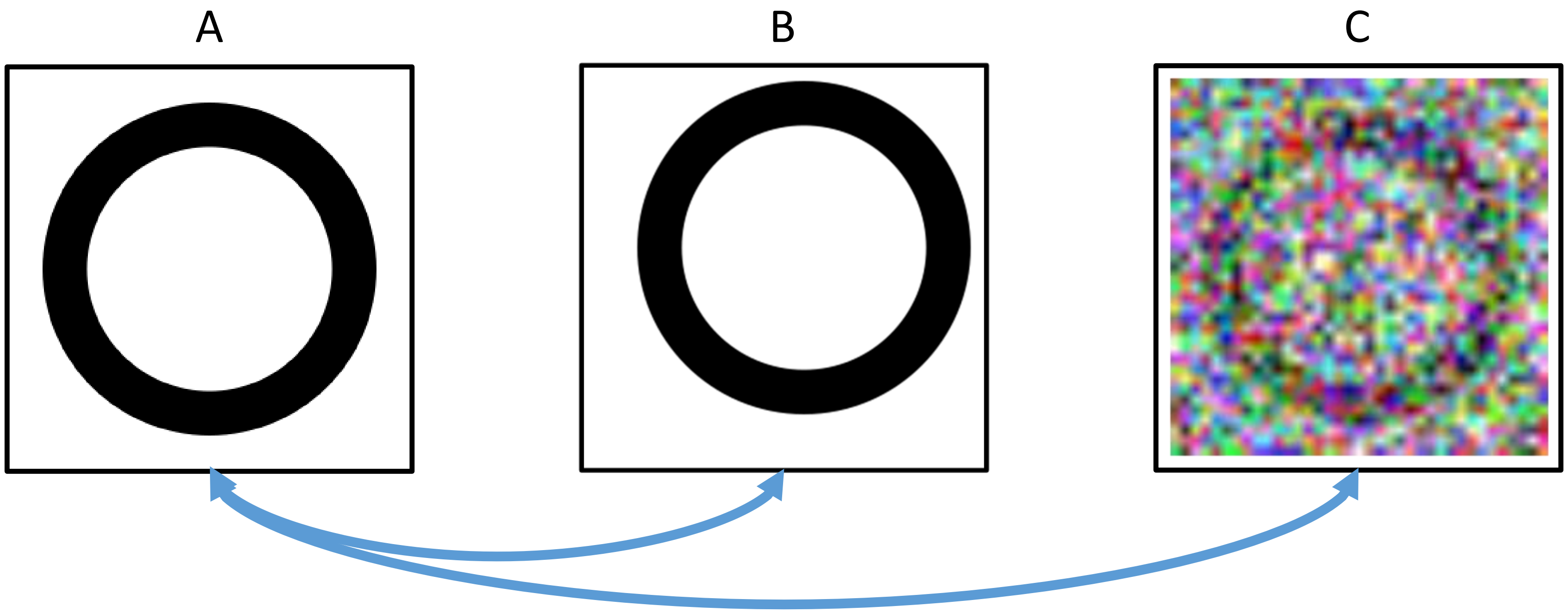
Sebastian Hoffmann, Christian Lessig



Overview

- AtmoDist is a novel representation learning technique for atmospheric dynamics
- Learned representations can be utilized in downstream applications, e.g. for Super-Resolution / Downscaling (this work)
- Demonstrates usefulness of representation learning for atmospheric dynamics

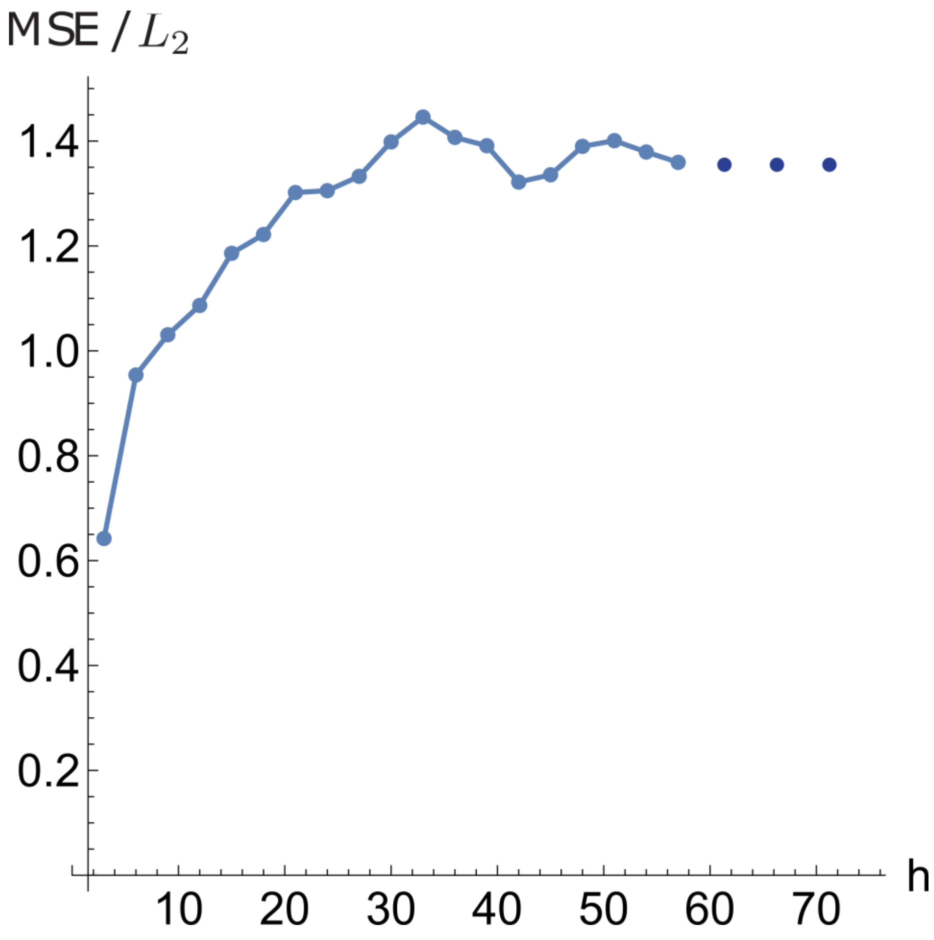
MSE / L2 is not an ideal metric



MSE (A, B) = MSE (A, C)

MSE / L2 is not an ideal metric

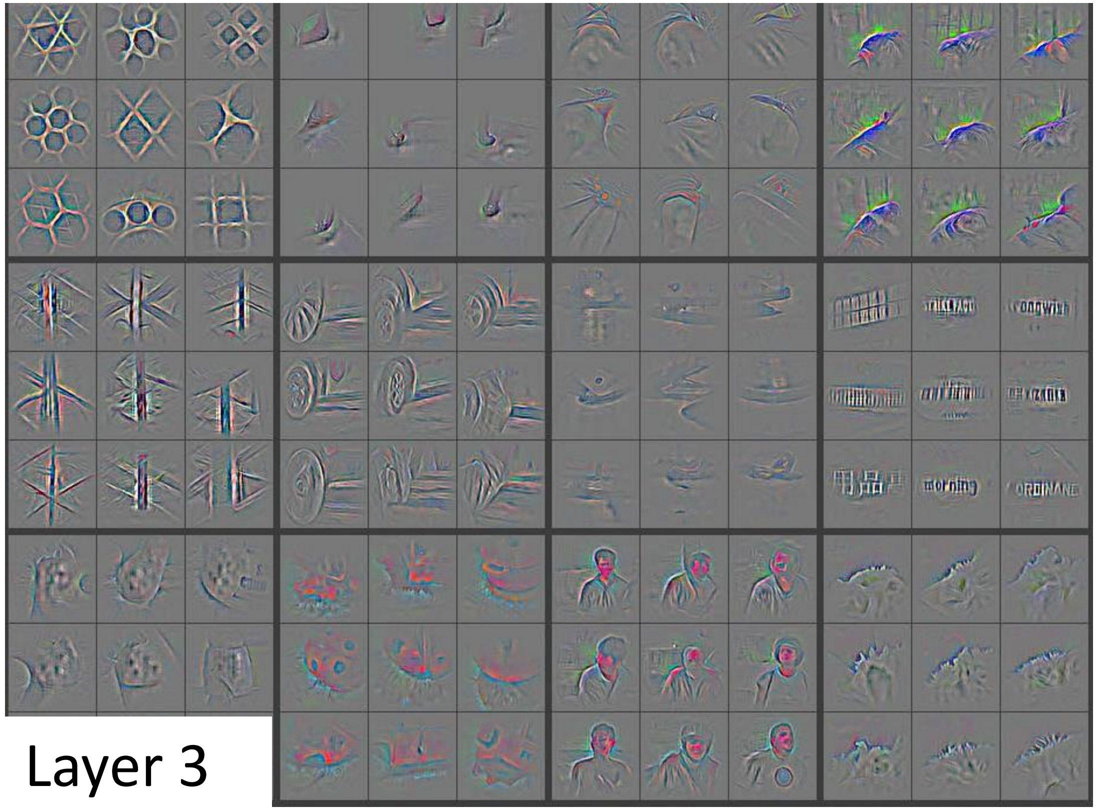
- Sharp edges with small offset can result in big distances
- Biased towards overly smooth images
- Converges quickly to equilibrium



Can we do better?

- During training, neural networks learn progressively more abstract features

- Thus, metrics based on these features measure semantic differences, not pixel-level perturbations



Layer 3

Visualizing and understanding convolutional networks, Zeiler & Fergus, 2014

Can we do better?

- Yes! Metrics based on neural network activations have been used for a long time in Computer Vision

- Examples: Super-resolution, Style-transfer, Fréchet Inception Distance

- However: Require pretrained neural network!

A Neural Algorithm of Artistic Style, Gatys et al., 2015

(Almost) no labelled datasets are available

- Labelling atmospheric data is tedious and requires domain experts

- Solution: Self-supervised learning
  - Train a network on a “pretext task” for which labels can be computed
  - Pretext-task forces network to capture semantic meaning in order to solve well

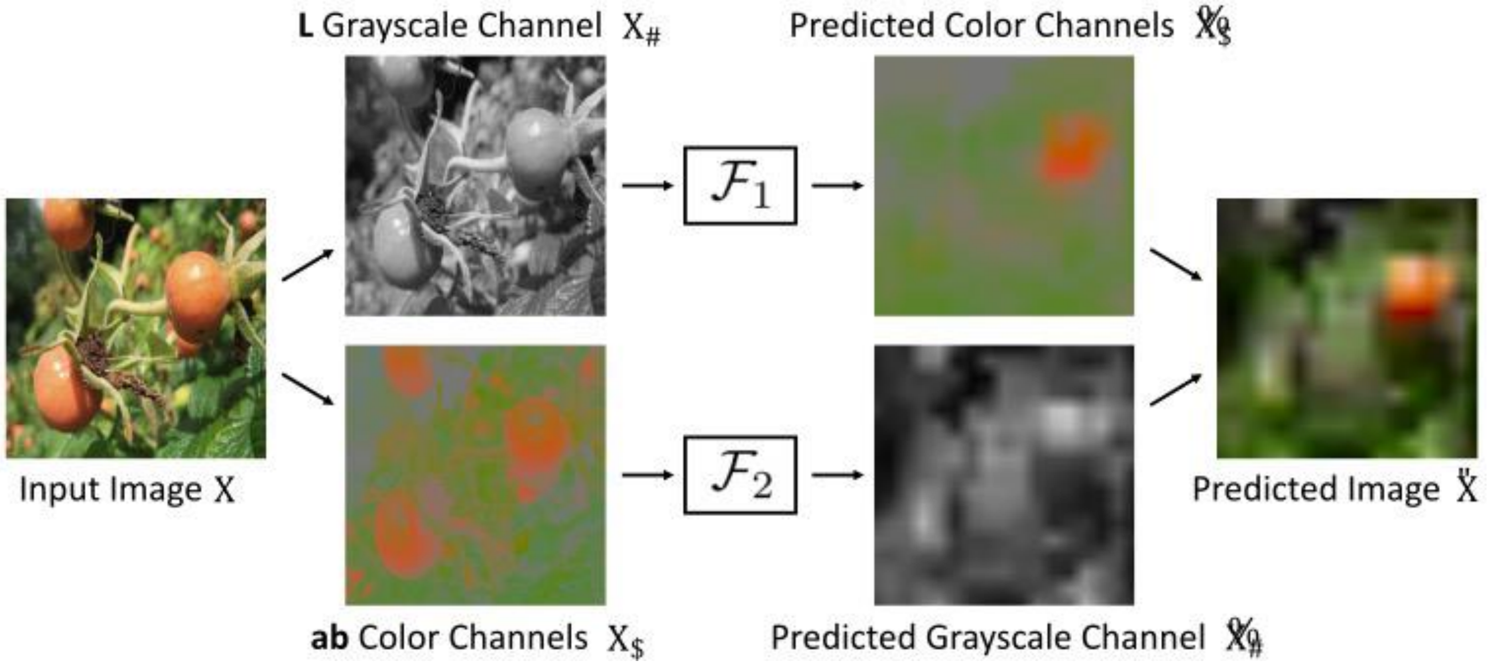
Pretext tasks

- Example: Inpainting of removed regions

Context Encoders: Feature Learning by Inpainting, Pathak et al., 2016

Pretext tasks

- Example: Predicting color from grayscale and vice-versa

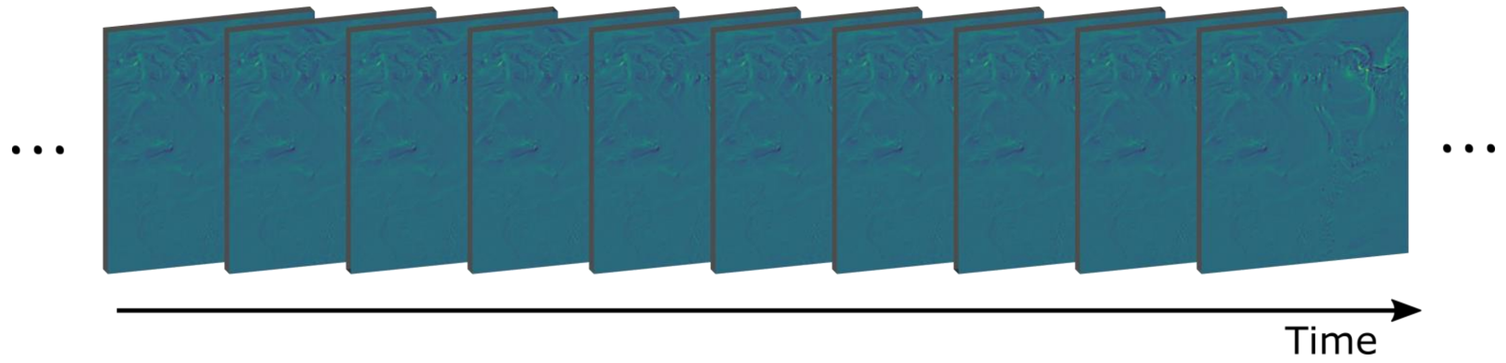


Split-Brain Autoencoders: Unsupervised Learning by Cross-Channel Prediction, Richard et al., 2017

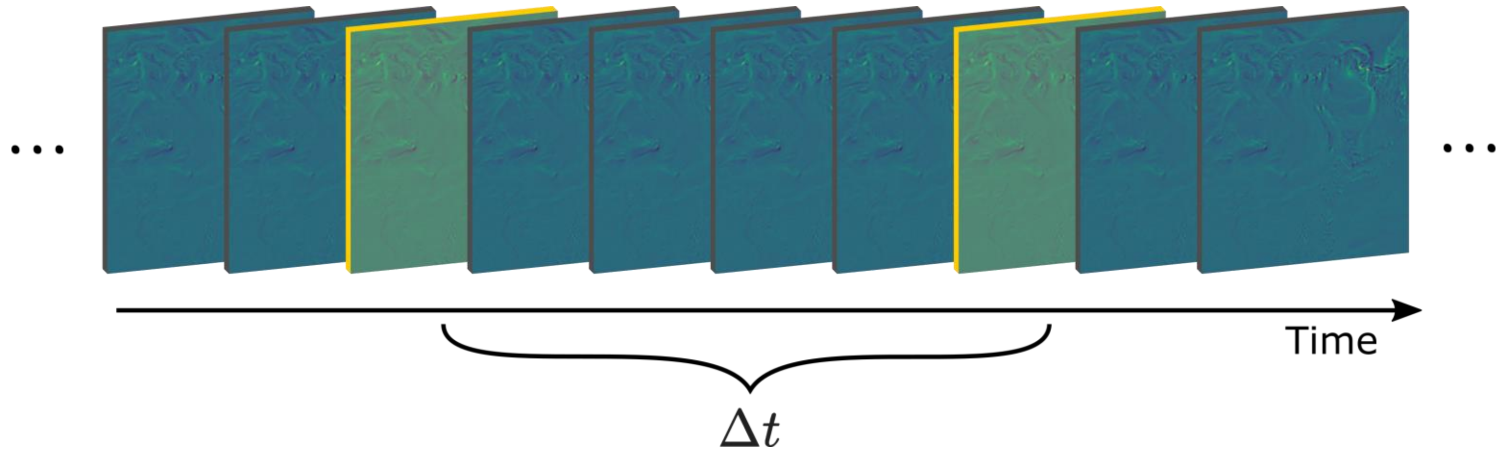
Pretext tasks

- What would be a suitable task for atmospheric dynamics?
- Not every task for CV is directly suited for atmospheric dynamics (e.g. rotations or jigsaw-puzzles)
- Predicting the temporal distance between two atmospheric states requires understanding how these systems evolve over time

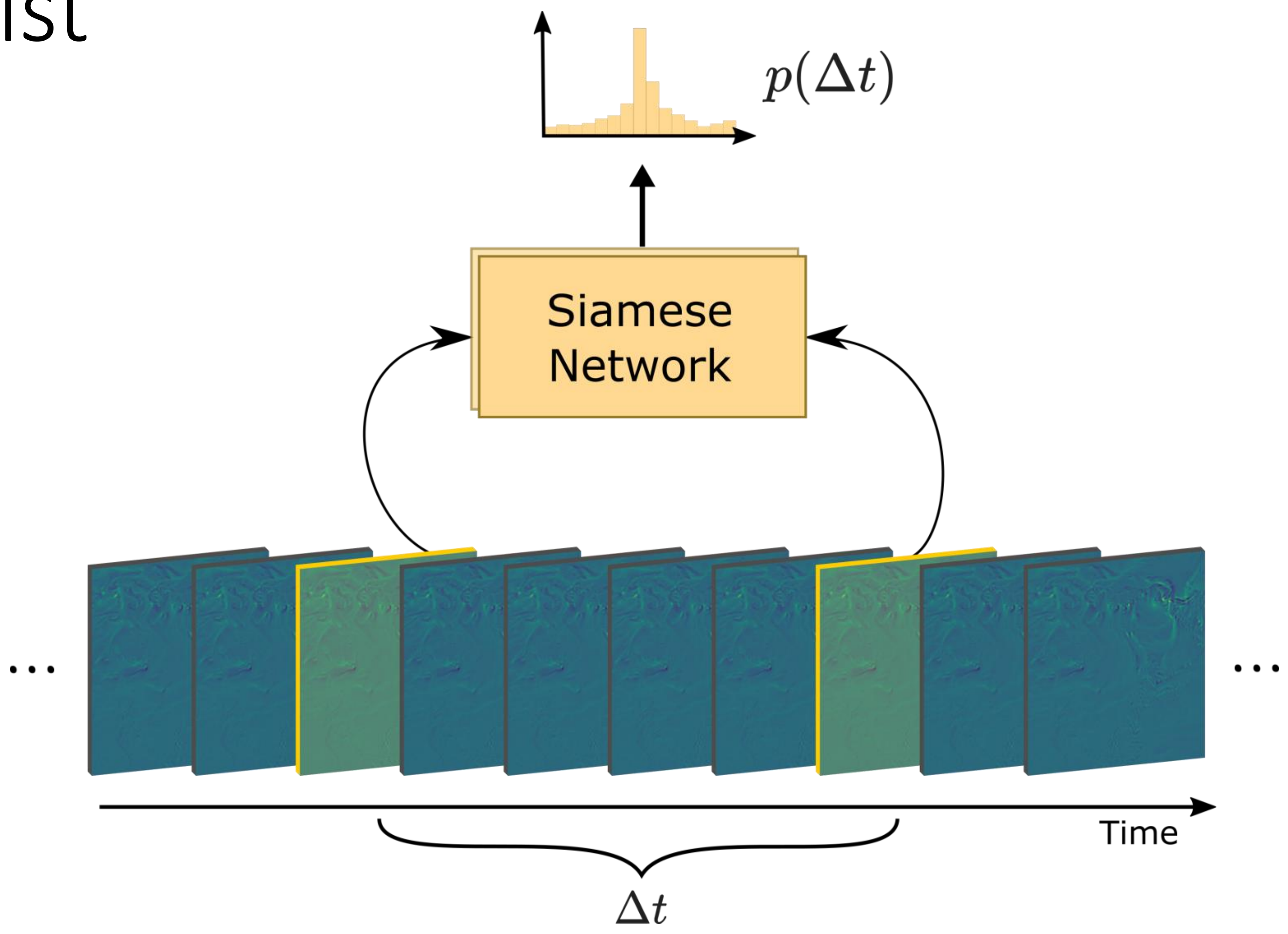
AtmoDist



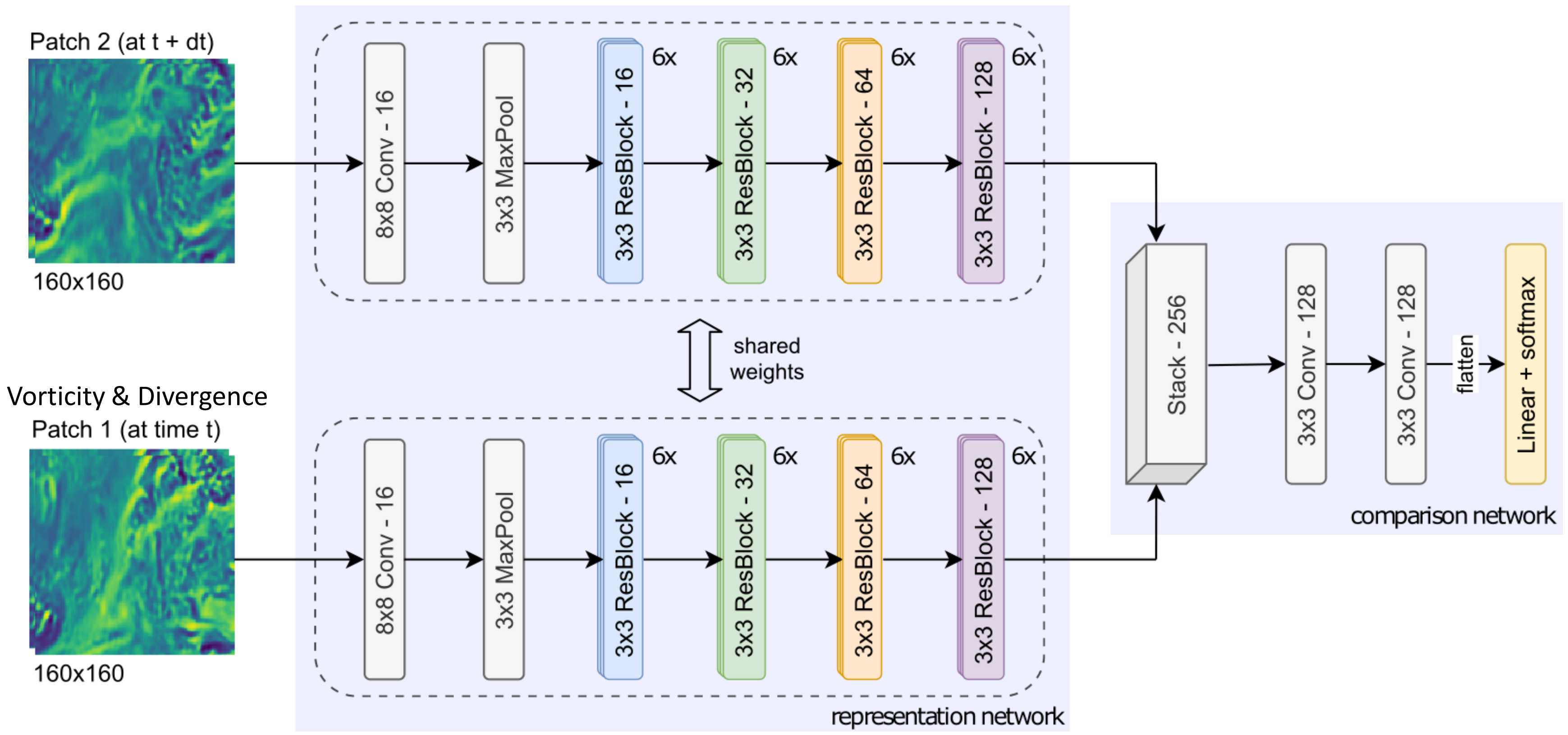
AtmoDist



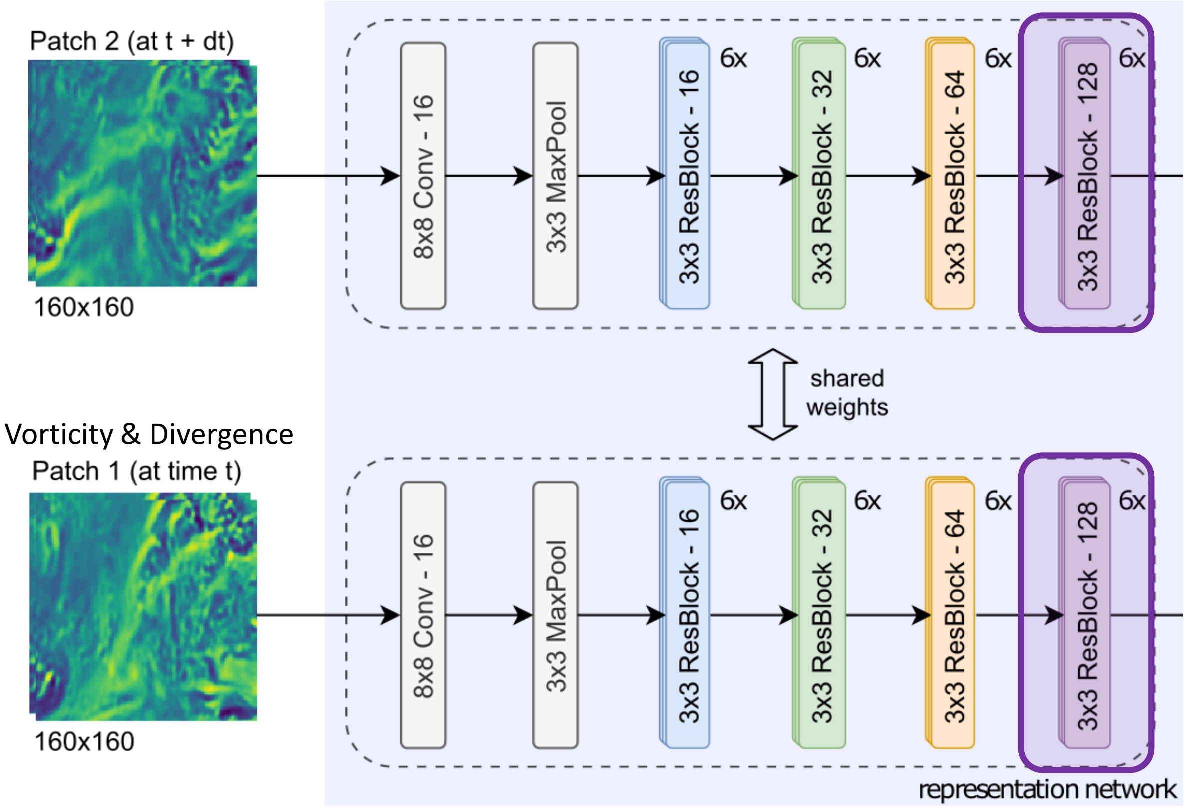
AtmoDist



Network architecture



Network architecture

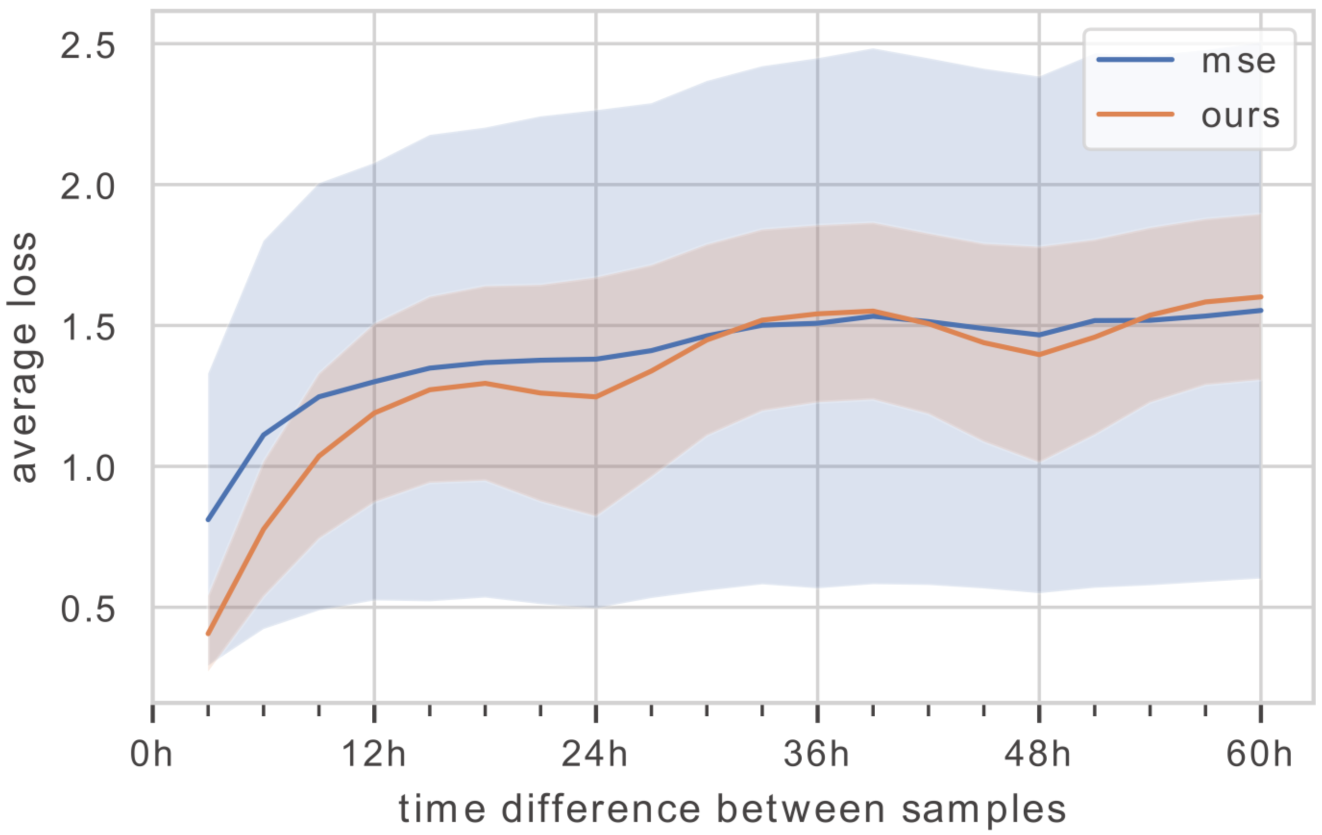


$$d(x_1, x_2) = |F(x_1) - F(x_2)|_2^2$$

Dataset

- ERA5 reanalysis data
- Vorticity and Divergence (potentials of the wind velocity field)
- Height: approx. 883 hPa (one level)
- 160 x 160 patches sampled from 1280 x 2560 lat-lon grids
- Training: 1979 - 1998, Evaluation: 2000 - 2006

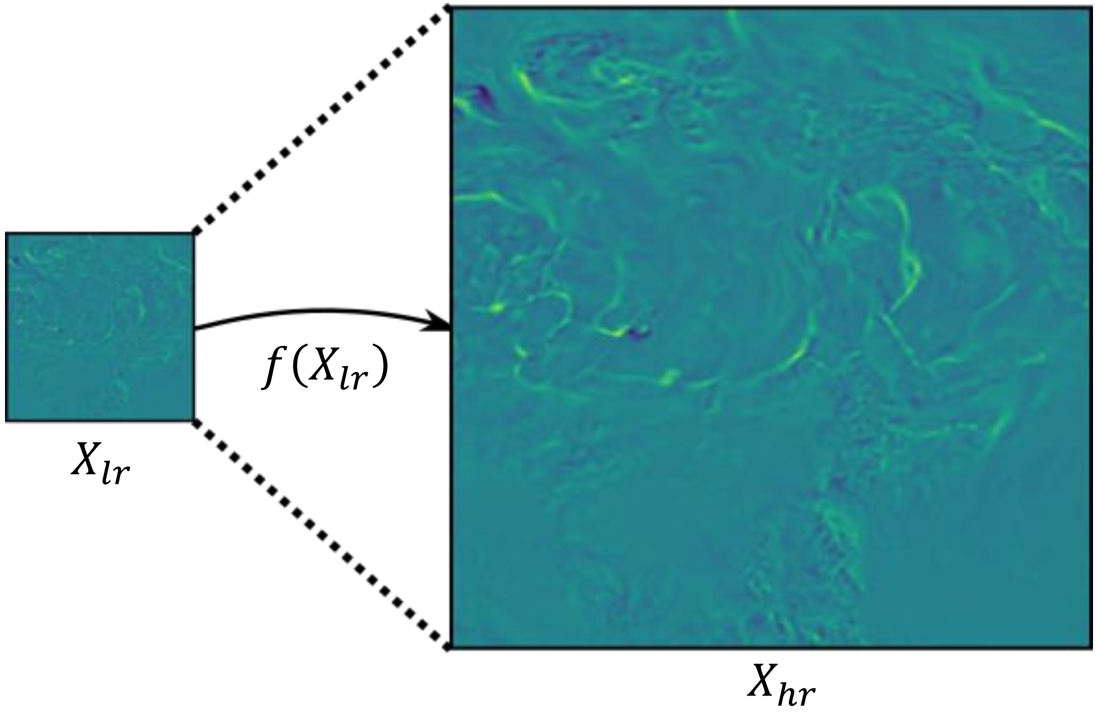
Evaluation



2.5
2.0
1.5
1.0
0.5
average loss
0h
12h
24h
36h
48h
60h
time difference between samples
mse
ours

Application: Super-Resolution

- Task: Given a low-resolution image X_{lr} find a high-resolution image X_{hr}
- In general, this is ill-posed
- But: A good matching image can often be found heuristically



SRGAN

- Jointly minimizes a content loss and adversarial loss

- Content-loss based on a VGG network pre-trained on ImageNet

$$\boxed{\mathcal{L}(X_{sr}, X_{hr}) = |F(X_{sr}) - F(X_{hr})|_2^2 + \ell_{adv}(X_{sr})}$$

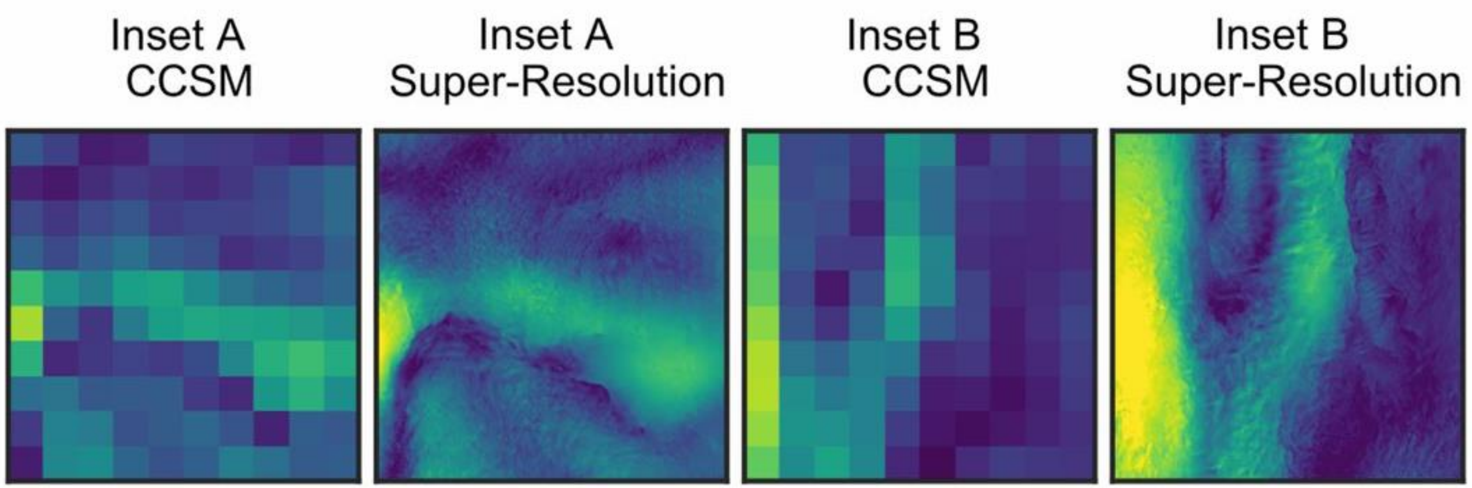
content loss

adversarial loss

Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network, Ledig et al., 2017

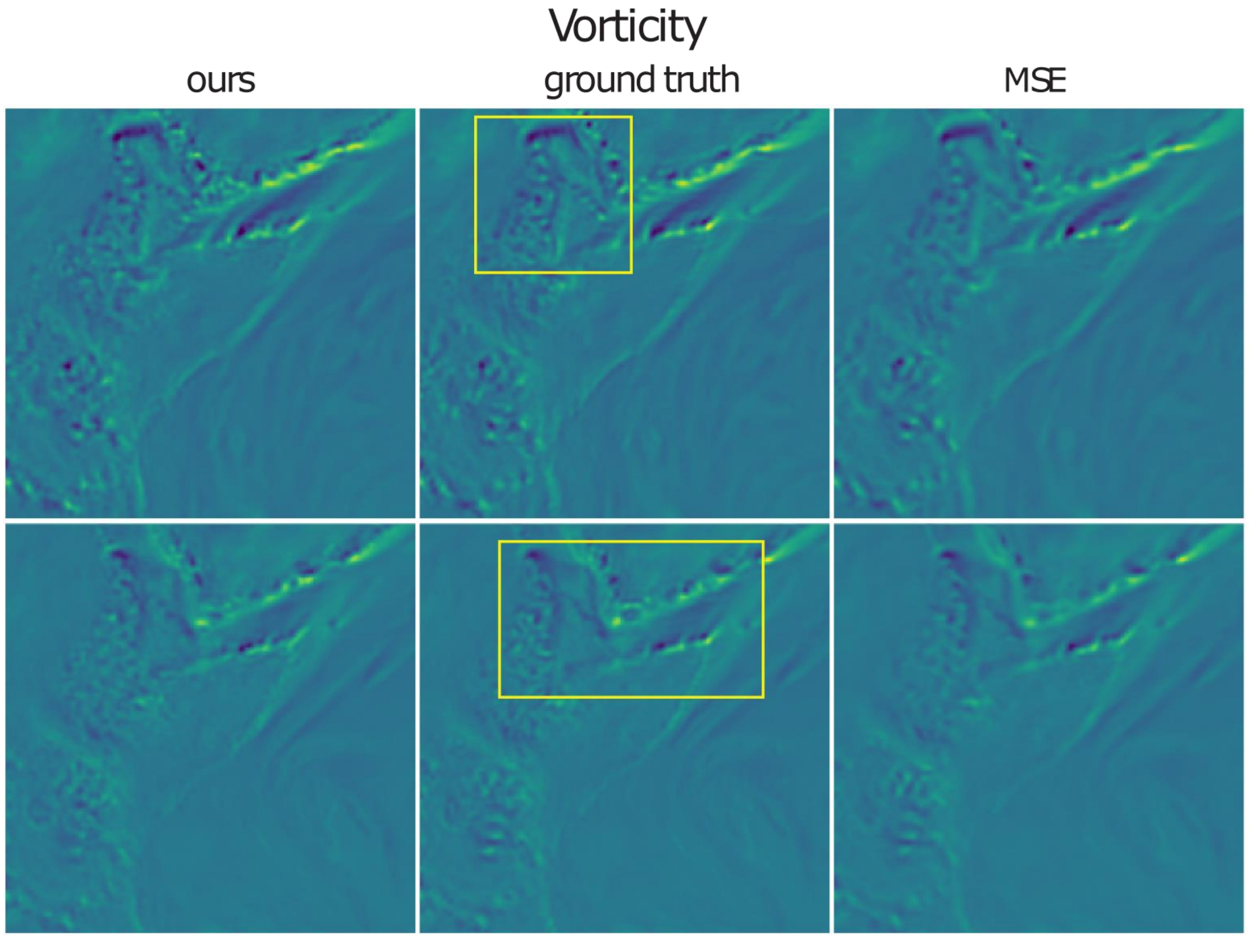
SRGAN

- Used for SOTA super-resolution of wind and solar data
- But: MSE as content loss

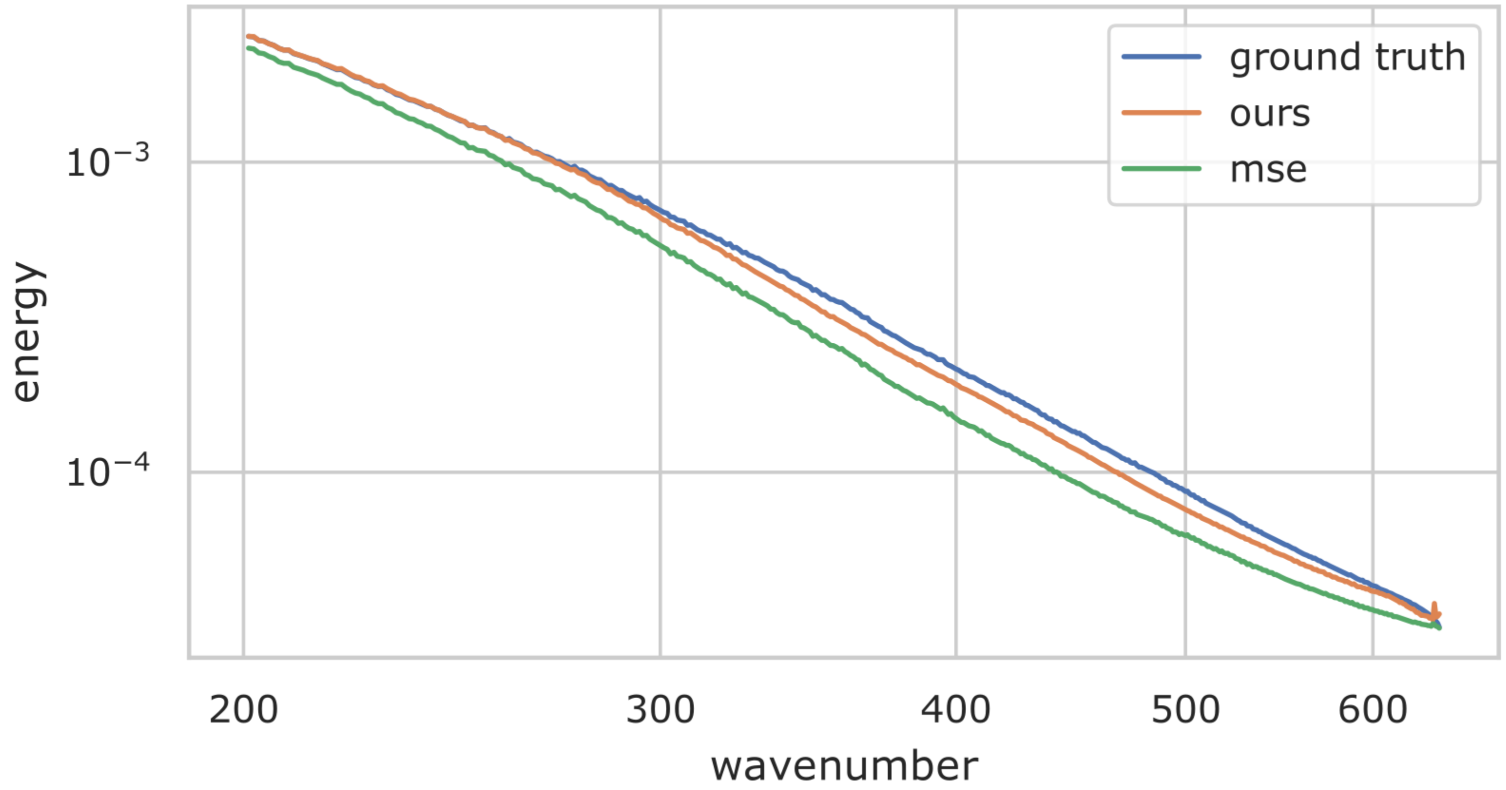


Adversarial super-resolution of climatological wind and solar data, Stengel et al., 2020

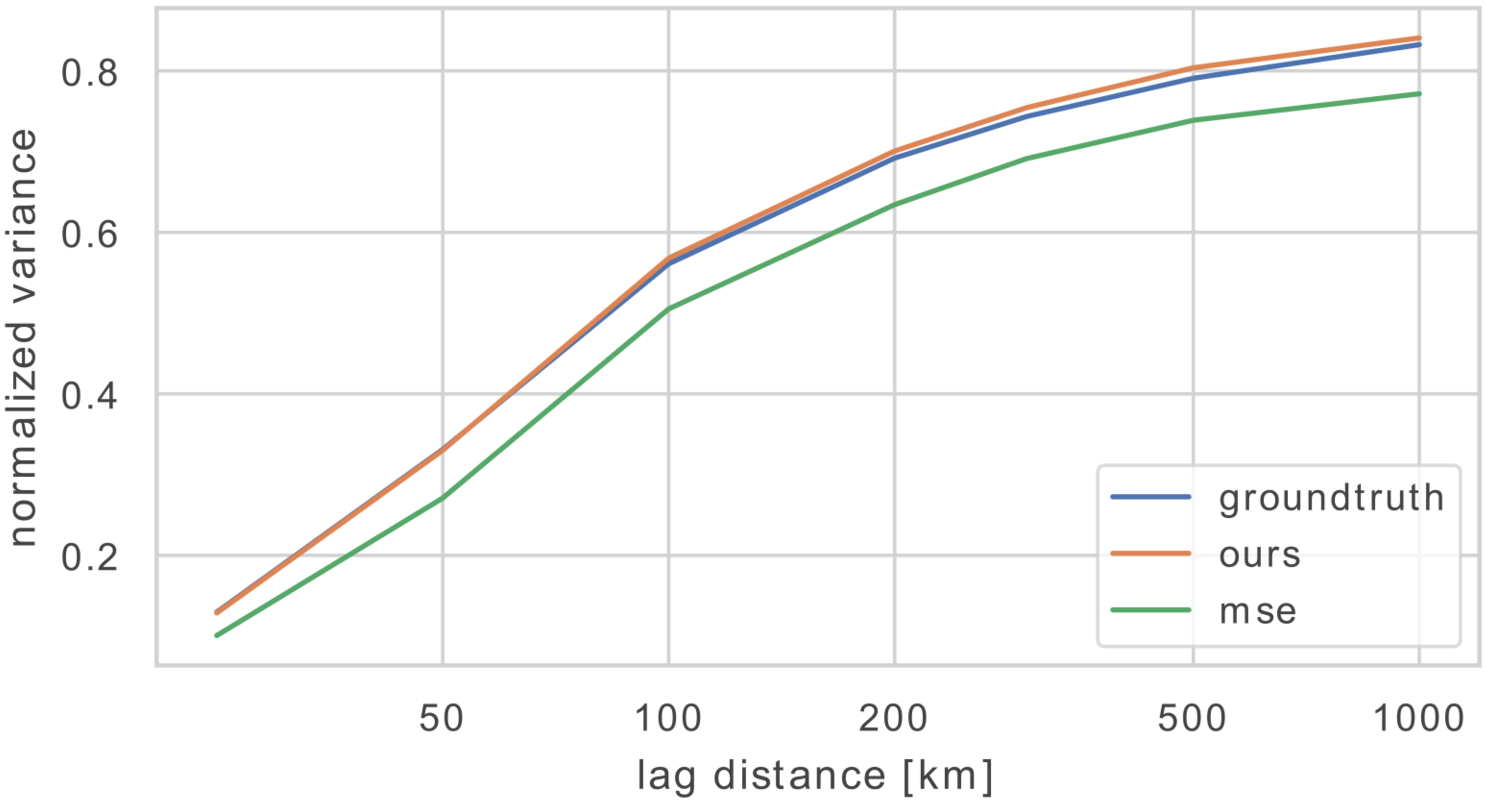
Evaluation on Super-Resolution



Energy spectrum

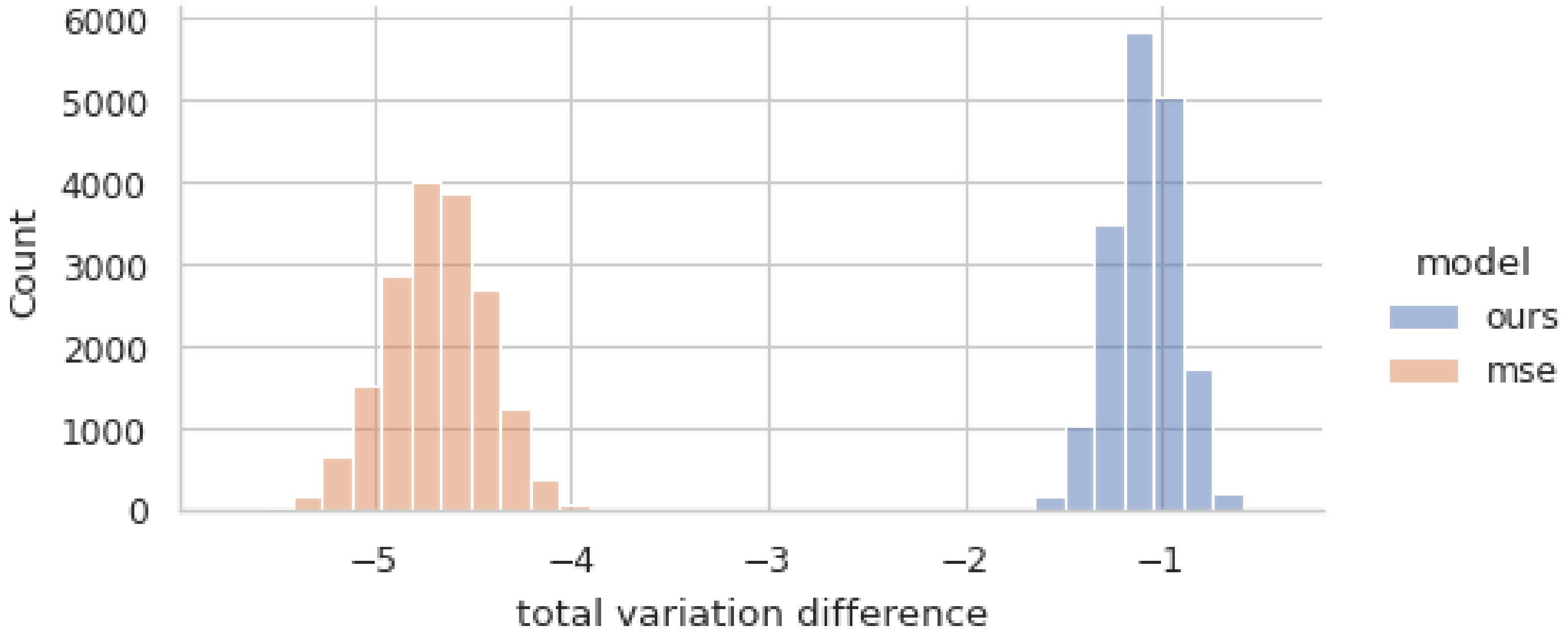


Semivariogram



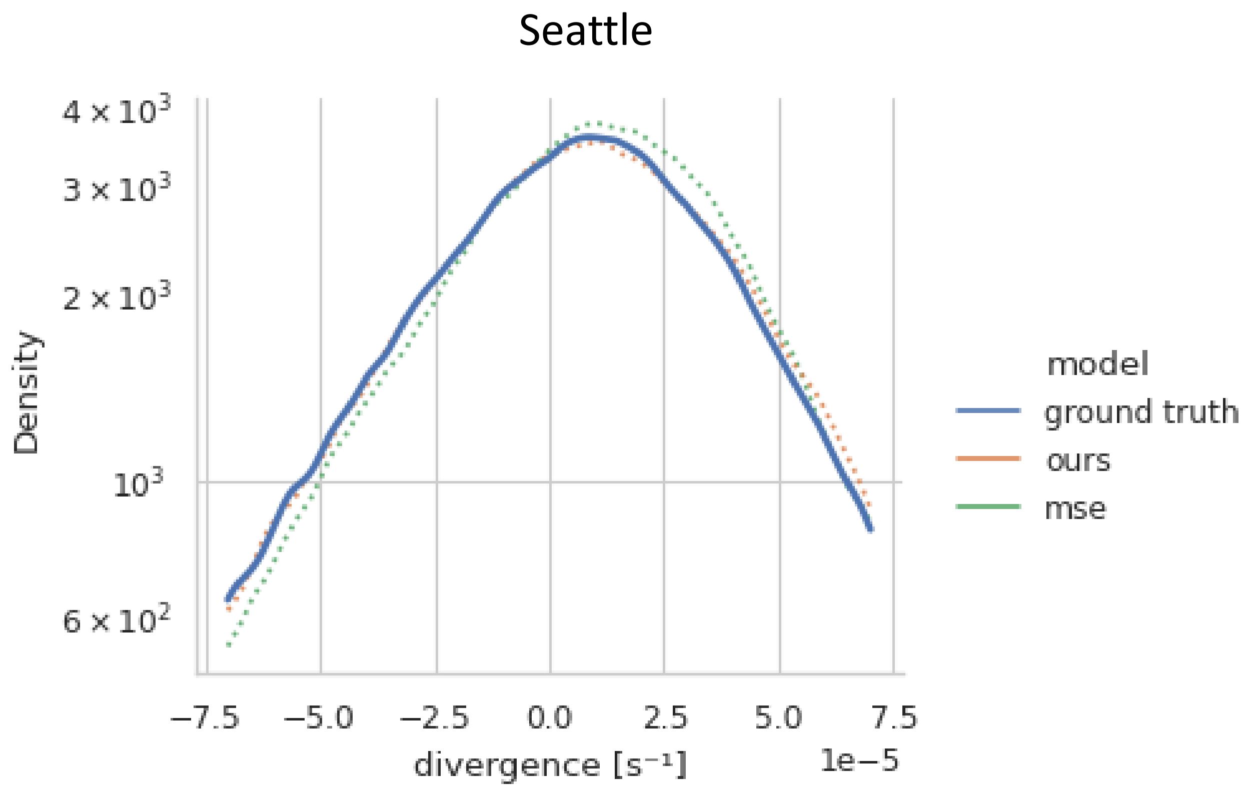
Total Variation

- Standard functional in Computer Vision
- Distribution of (absolute) errors:



Pointwise statistics

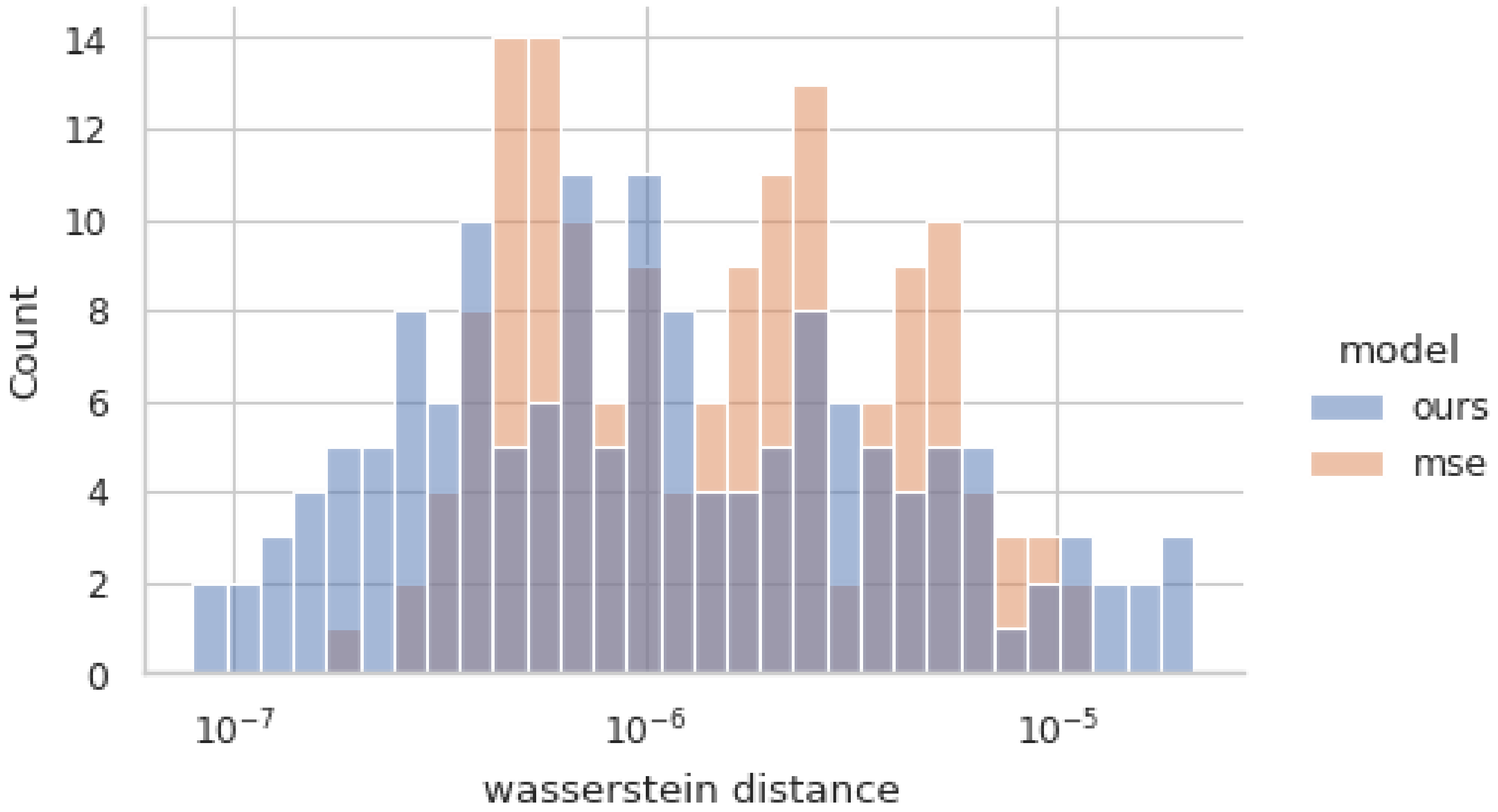
- Calculated across time
- 150 big cities (evenly distributed)
- Systematic evaluation with Wasserstein distance



Pointwise statistics

	Better	Equal	Worse
Divergence	**102**	12	36
Vorticity	**90**	11	49

Pointwise statistics



Future Work

- Upscaling: More data, levels, variables
- Availability to other researchers: How can we account for different data / variable choices?
- Incorporate spatial dependencies as well: spatio-temporal pretext task
- How to deal with climate change / distribution shift?
  - Domain Adaptation could help here

Future Work

- Detection of atmospheric phenomena
  - Tropical cyclones & atmospheric rivers (ClimateNet, Prabhat et al., 2021)
  - Blocking events
- Forecasting applications: AtmoDist as loss-function
- Evaluation of generative models or simulations (e.g. analogous to Fréchet Inception Distance; Heusel et al., 2017)

Summary

- AtmoDist provides a simple but effective pretext task for atmospheric data

- As unlabelled data is abundant, but labelled data is scarce, unsupervised / semi-supervised learning is a promising direction for atmospheric machine learning

- We demonstrated the usefulness of these ideas by improving upon the previous SOTA for SR

Code available at

https://github.com/sehoffmann/AtmoDist