Generative Adversarial Networks for Climate Data Field Generation

Jussi Leinonen
Environmental Remote Sensing Laboratory, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
jussi.leinonen@epfl.ch

With contributions from:
Tianle Yuan (UMBC/NASA-GSFC), **Alexandre Guillaume** (Caltech-JPL), **Alexis Berne** (EPFL-LTE),
Christophe Praz (EPFL-LTE), **Daniel Wolfensberger** (EPFL-LTE), **Marco Gabella** (MeteoSwiss)

Summary

Generative Adversarial Networks (GANs) are neural networks that learn to map a simple probability distribution to the training data distribution. Through adversarial training, they learn to generate realistic artificial examples of their training dataset.

GANs, especially their variants such as the conditional GAN, are applicable to many Earth Science data problems:

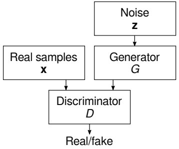
- GANs are inherently probabilistic and therefore a good fit for tasks requiring infererence from incomplete data
- Using convolutional neural networks, GANs can generate complex spatial patterns, commonly encountered in many Earth Science applications

---

Generative Adversarial Networks (GANs)

Basic GAN: Adversarial training between two neural networks

- The *discriminator* learns to distinguish real samples from generated ones
- The *generator* learns to fool discriminator as much as possible, transforming random noise into samples that resemble those in the training set



Noise
z

Real samples
x

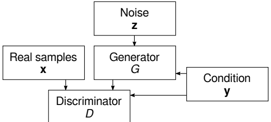
Generator
G

Discriminator
D

Real/fake

Conditional GAN (CGAN): Learns a conditional data distribution rather than the global distribution

- In addition to the random noise, a conditioning variable is passed to the generator and the discriminator
- Generator produces realistic examples



Noise
z

Real samples
x

Generator
G

Condition
y

Discriminator
D

In this poster, we demonstrate proofs of concept for several applications of GANs to problems related to clouds and precipitation.

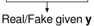
Reconstructing Cloud Vertical Profiles

Download a copy of the poster here:

Data: Cloud vertical profiles (2D) from the CloudSat cloud radar and retrieved cloud properties (1D) from the MODIS optical spectrometer.

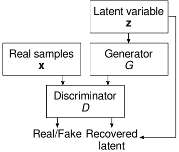
Question: Can we use CGAN to generate the cloud vertical profile using only the MODIS variables as input?

subject to the condition



InfoGAN: Recovery of latent variables

- Discriminator also predicts the latent variable passed to the generator
- Once trained, this can be applied to the real samples to extract their essential features

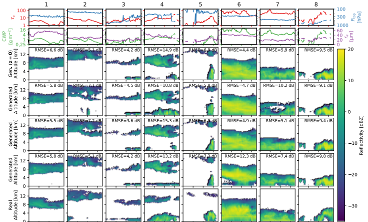


Downscaling Weather Radar Data

Classifying Snowflakes

Data: Images of snowflakes from the Multi-Angle Snowflake Camera (MASC) deployed in Davos,

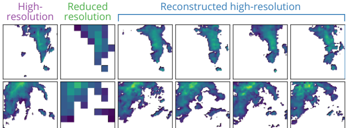
Result: The CGAN can generate realistic cloud profiles and also estimate the uncertainty of its predictions through sample diversity.



Data: Precipitation fields from a composite of the MeteoSwiss weather radar network.

Question: Can we recover high-resolution precipitation from reduced-resolution data?

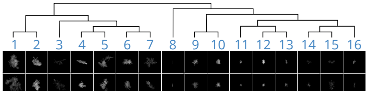
Result: The CGAN can generate realistic solutions and can often infer the type of precipitation from the low-resolution data.

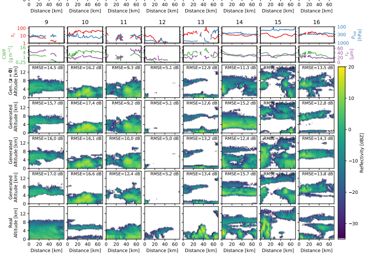


Switzerland in 2015-16.

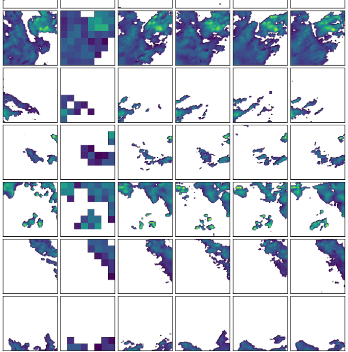
Question: Can we classify the snowflakes in an unsupervised way using InfoGAN?

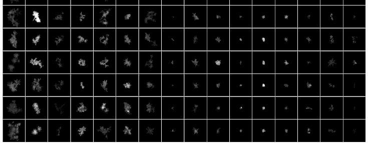
Result: We can classify snowflakes and organize them as a "family tree" using InfoGAN-derived latent variables in combination with the K-medoids algorithm and hierarchical clustering. Data augmentation is used to make the classification insensitive to some features such as rotation.





Paper: Leinonen et al. (2019), GRL,
https://doi.org/10.1029/2019GL082532





More Information

My blog on GANs, atmospheric science et al., with a GAN tutorial series in the making:
https://jleinonen.github.io/

Code repository for the localized StyleGAN for precipitation data:
https://github.com/jleinonen/geogan