

Improving Few-Shot learning through latent space topology of deep audio generative models

Nils Demerlé

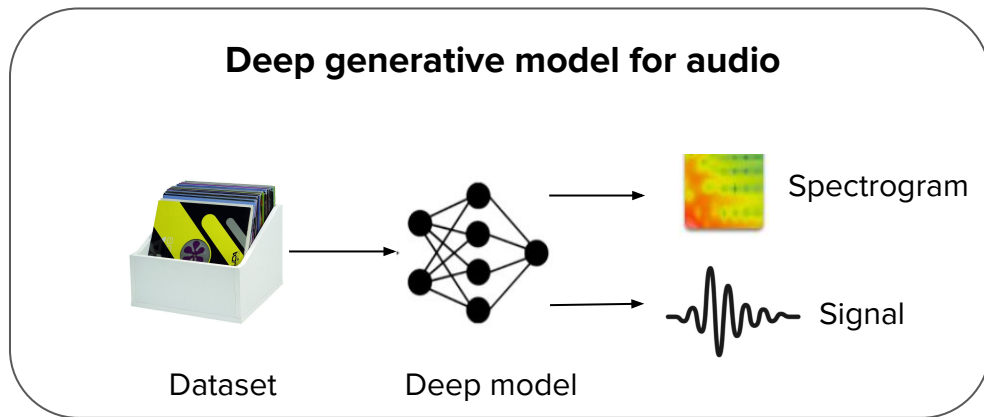


13.02.2022

Supervisors
Philippe Esling
Guillaume Doras

Context

Introduction



Recent advances in deep generative audio

- **Complex sounds**
- **Diverse and realistic**
- *Rely on **complex architectures***
- *Trained on **large datasets***

Major limitations

- **Time consuming** process
- Tasks with **limited data** available
- **Energy cost** of training

Objective

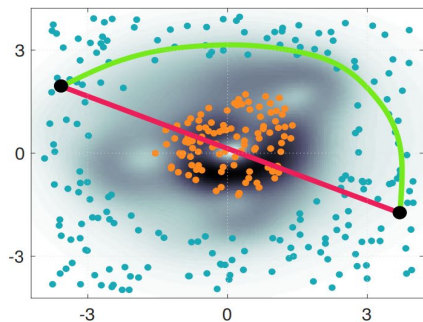
Improve generalisation of deep generative models trained on limited data

Latent Geometry

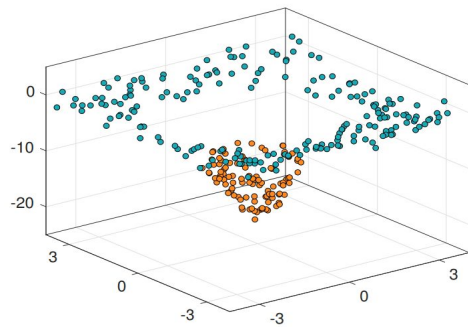
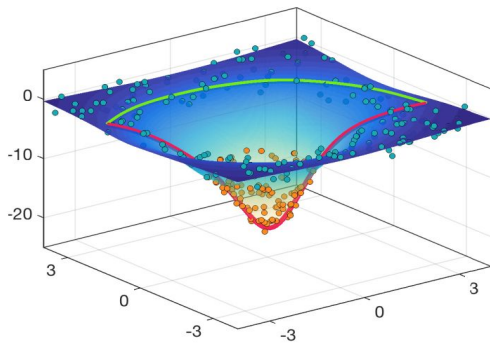
State of the art

Hypothesis : geometrical considerations are overlooked by most generative models

Latent space



Ambient space



$$\mathcal{Z} \xrightarrow{f} \mathcal{M} \subset \mathcal{X}$$

The manifold hypothesis

Most real-world high-dimensional datasets lie in the vicinity of a manifold of much lower dimension.

Euclidean latent space

- Interpretation of linear distances in \mathcal{Z}
- Interpolation / Extrapolation
- Mismatch Prior / Posterior

Geometry Aware VAE

Data augmentation in high dimensional low sample size setting using a geometry-based variational autoencoder. Chadebec, C., Thibeau-Sutre, E., Burgos, N., & Allasonnière, S. (2022). IEEE

State of the art

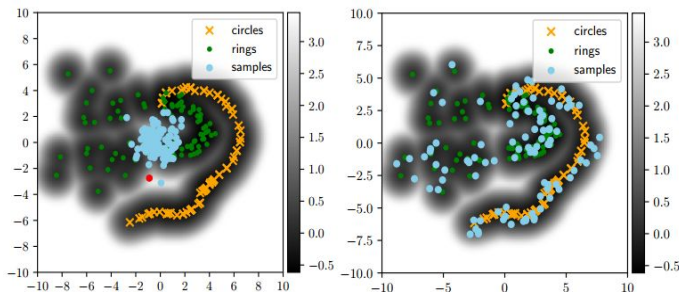
Modelling the latent space as a Riemannian manifold

- ❖ Learn a Riemannian metric from data that captures the geometry

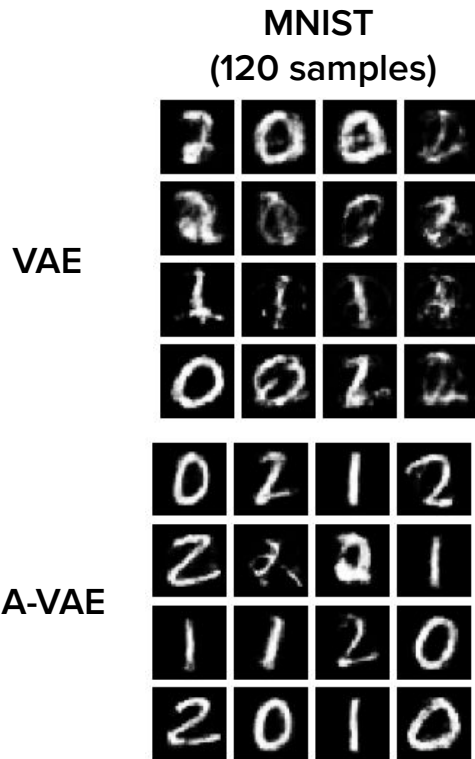
Exploit the learned geometry

- ❖ Interpolate with respect to geodesics in the ambient space
 - **Realistic and smooth** interpolations
- ❖ Geometry-aware sampling
 - Sample **onto the data manifold** using the learned metric
 - **Improved sample quality and diversity**

Gaussian prior



Geometric prior



Research Approach

Understand the topology of audio manifolds and its implications for deep audio models

- Modelling the latent space structure as a **Riemannian manifold** to characterize the learned geometry.
- Derive **geodesic interpolations and geometry aware sampling** techniques to improve quality of generated samples.
- Interpretability of geodesic interpolations / disentangling

Exploit the latent geometry during training

- Design **implicit data augmentation** strategies informed by the latent geometry
- **Regularize** the training based on geometrical constraints
- **Transfer the latent structure** from a different task to inform training with inductive bias