



# UNCERTAINTY QUANTIFICATION THROUGH METRIC-AWARE SAMPLING USING RIEMANNIAN MANIFOLDS

Trym Hamer Gudvangen

# TABLE OF CONTENTS

## Background

- VAEs
- Manifold Hypothesis
- Differential Geometry – Riemannian Manifolds
- Geodesics and Interpolation
- Models Used

## Sampling Methods

- Gaussian Sampling
- Geodesic Sampling

## Uncertainty

- Metrics
- From Geodesic Sampling
- Log of Volume Metric
- RBFNet

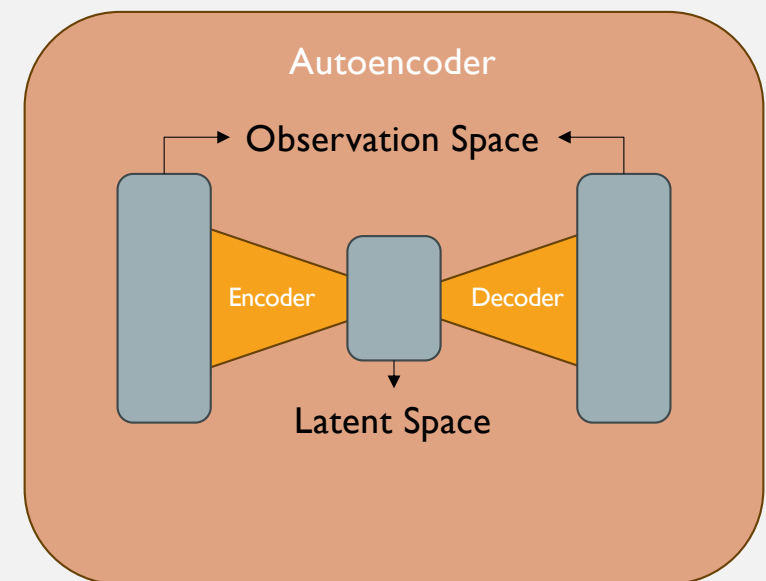
BACKGROUND

# VARIATIONAL AUTOENCODER

**Goal:** Encode data from observation space to a lower dimensional latent distribution and decode latent point to observation space.

**How:** Special Loss + Reparameterization Trick

**Why:** Stability, Simplicity, Baseline in differential geometry papers, Probabilistic



# VARIATIONAL AUTOENCODER

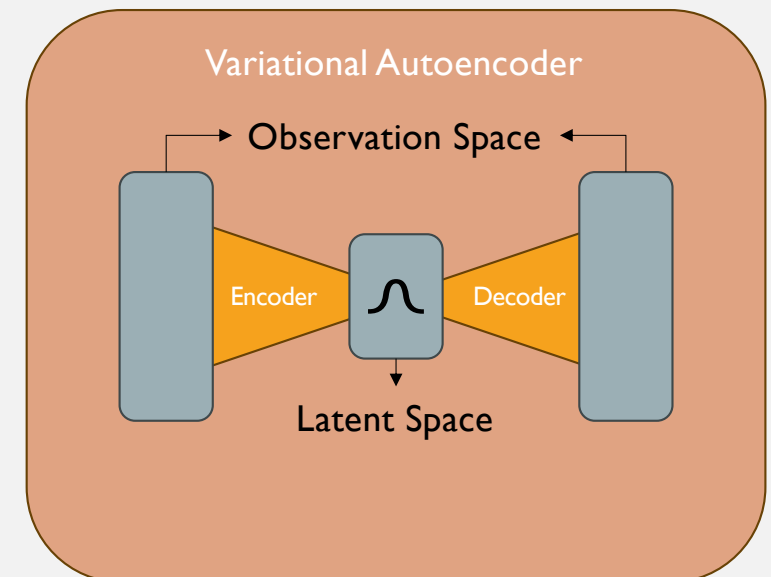
**Loss** = Evidence Lower Bound = Reconstruction Loss + KL-Divergence

$$L(\phi, \theta; x) = \ln p_{\theta}(x) - D_{KL}(q_{\phi}(z|x) || p_{\theta}(z|x))$$

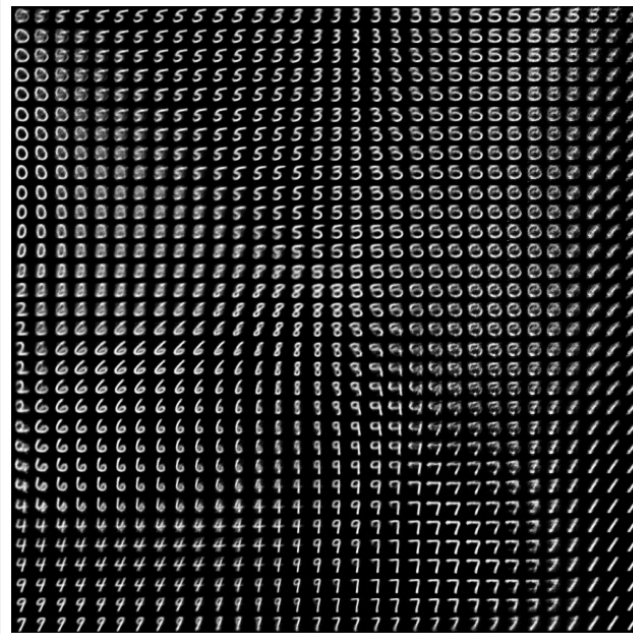
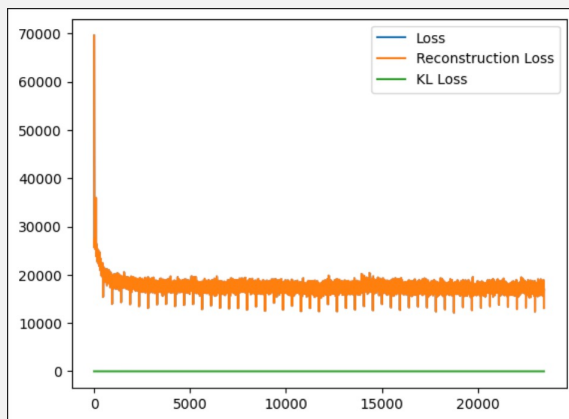
$$\ln p_{\theta}(x) \geq \mathbb{E}_{z \sim q_{\phi}(\cdot|x)} \left[ \ln \frac{p_{\theta}(x, z)}{q_{\phi}(z|x)} \right]$$

Jensen's Inequality

$$\ln p_{\theta}(x) = \ln \mathbb{E}_{z \sim q_{\phi}(\cdot|x)} \left[ \frac{p_{\theta}(x, z)}{q_{\phi}(z|x)} \right] \geq \mathbb{E}_{z \sim q_{\phi}(\cdot|x)} \left[ \ln \frac{p_{\theta}(x, z)}{q_{\phi}(z|x)} \right]$$



# TRAINING THE VAE



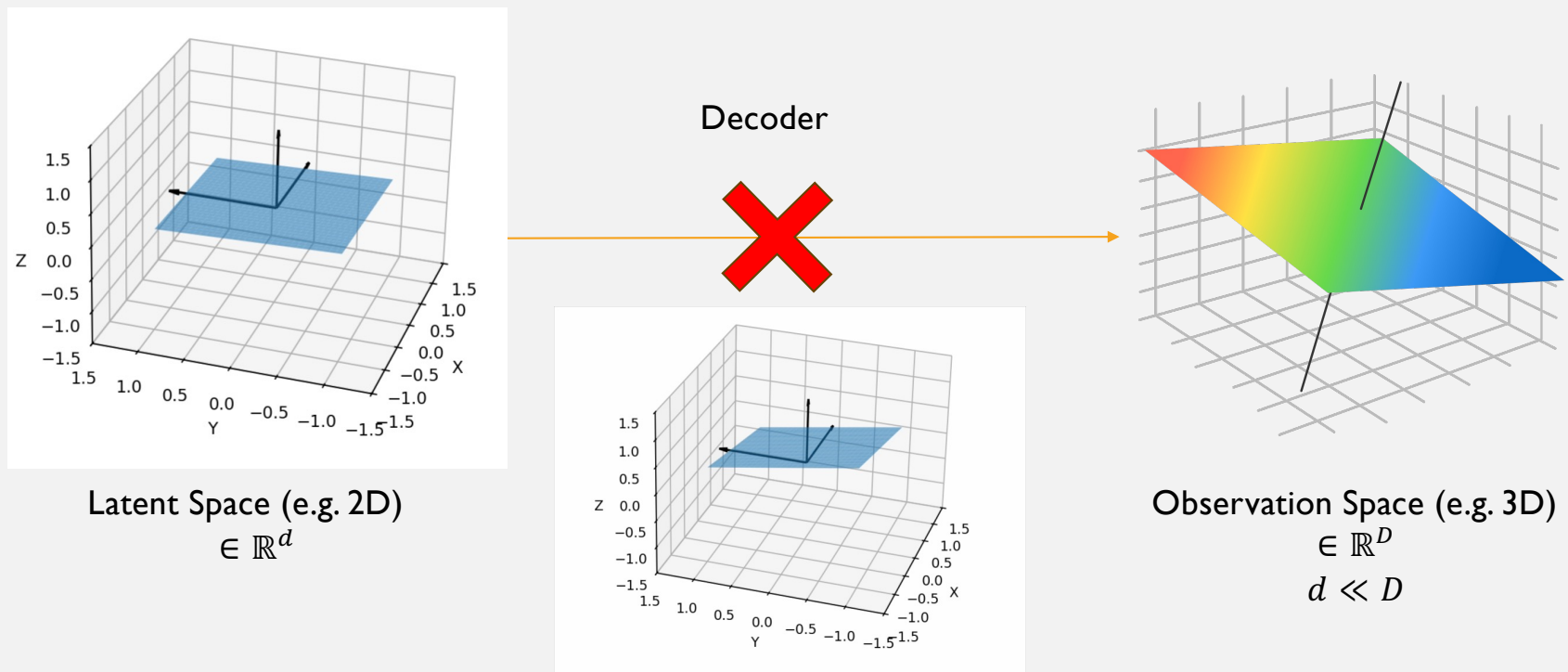
Attempts to make reconstruction better (less blurry)

- Changing activation functions
- Normalization
- BCE instead of MSE for reconstruction loss
- Using Beta VAEs (essentially weighted loss)
- Hierarchy/Depth

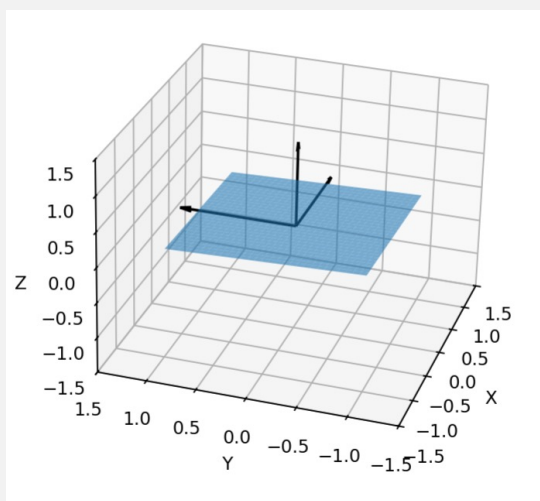
# MANIFOLD HYPOTHESIS

- **Assumes:** Real-life high dimensional datasets exist as embedded data manifolds.
- Underlying assumption for generative models
- Important for approach going forward

# DIFFERENTIAL GEOMETRY

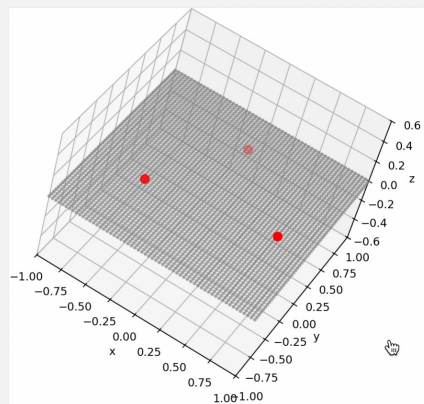


# DIFFERENTIAL GEOMETRY



Latent Space (e.g. 2D)  
 $\in \mathbb{R}^d$

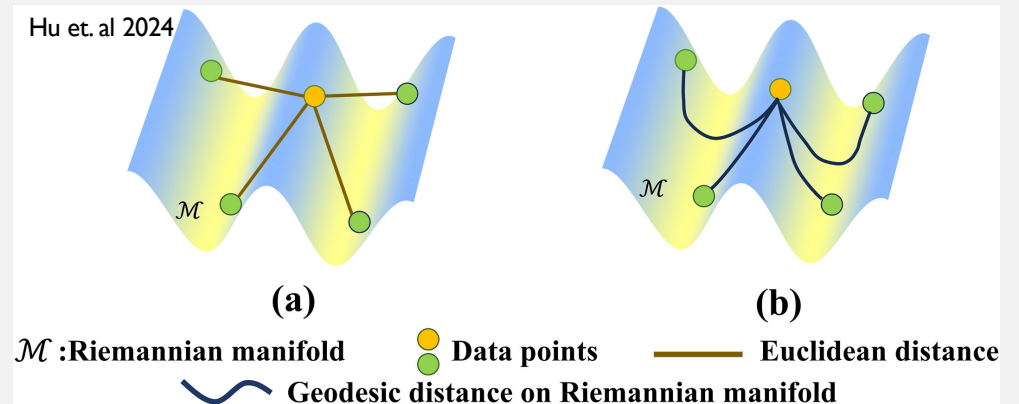
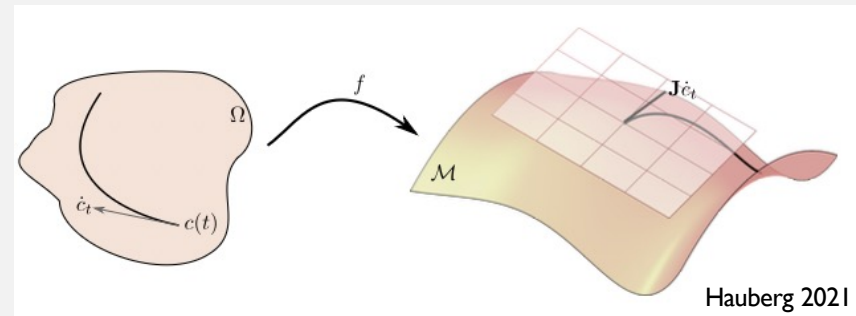
Decoder



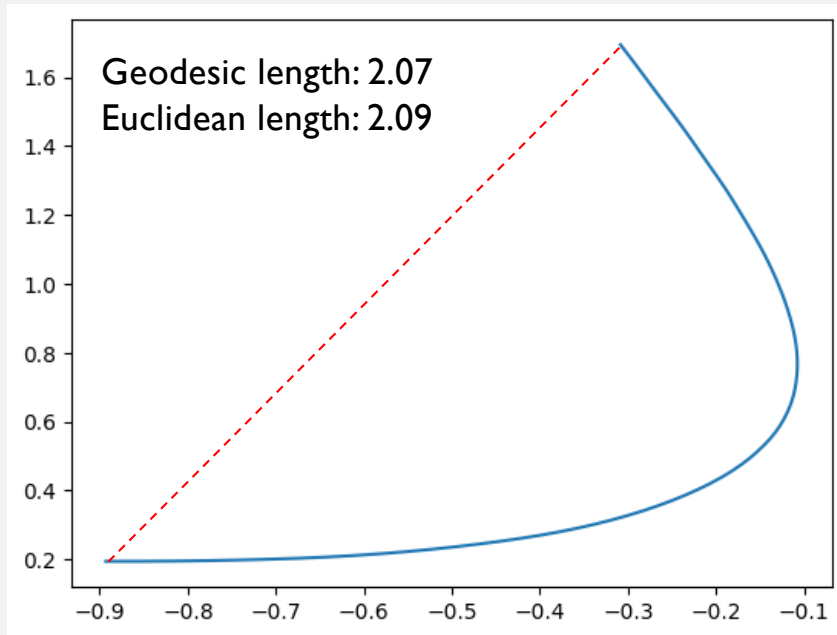
Observation Space (e.g. 3D)  
 $\in \mathbb{R}^D$   
 $d \ll D$

# RIEMANNIAN MANIFOLD

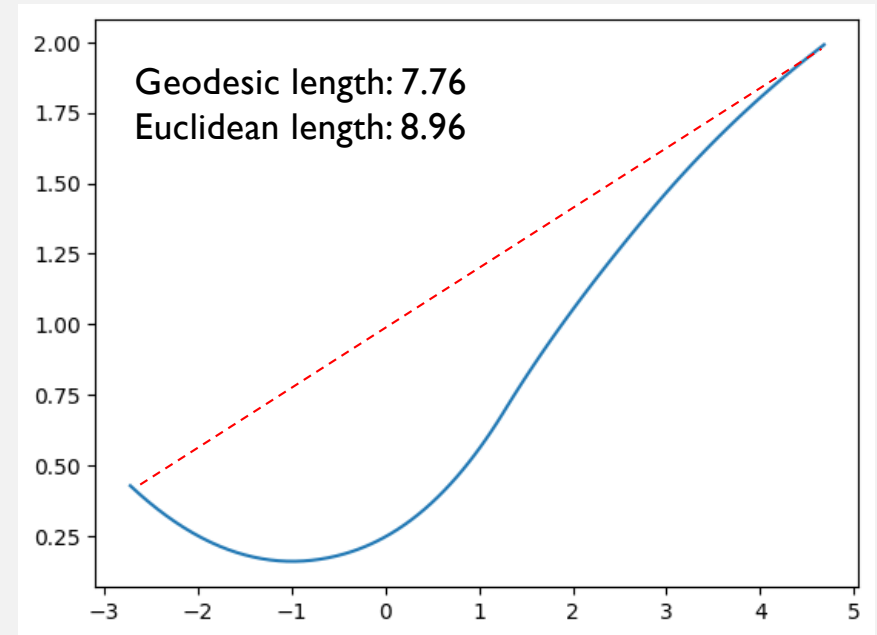
- Curved surface with local Euclidean geometry
- Metrics on Riemannian manifold:
  - Riemannian Metric  $\mathbf{M}_z = \mathbf{J}_z^\top \mathbf{J}_z$
  - Geodesics
  - Volume Metric  $\sqrt{\det(\mathbf{J}^\top \mathbf{J})}$



# GEODESIC



Two Latent Points from 5-digit inputs

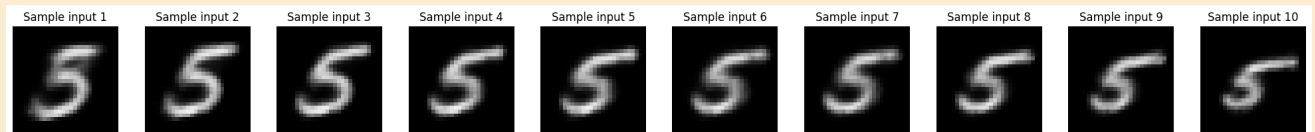


Two Latent Points from 0- and 1-digit inputs

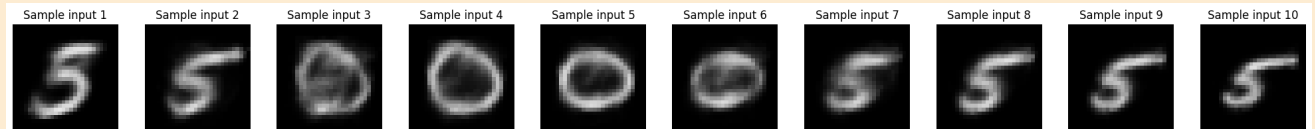
# INTERPOLATION

## Same Class

Geodesic Interpolation

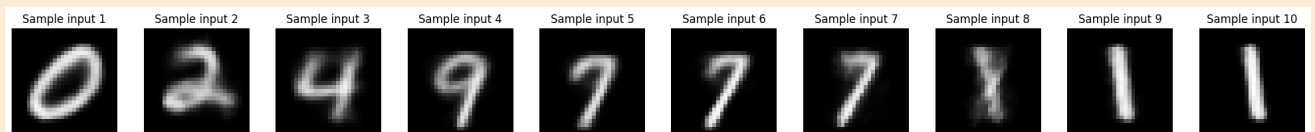


Euclidean Interpolation

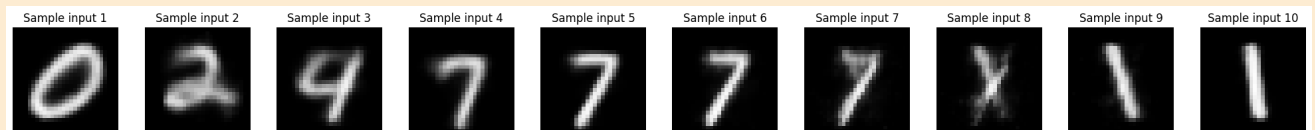


## Different Class

Geodesic Interpolation



Euclidean Interpolation

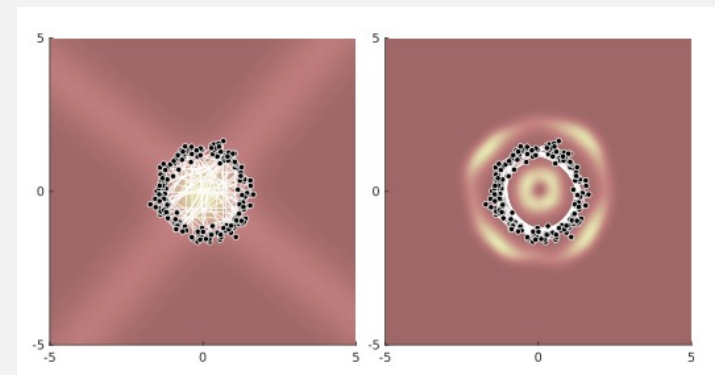


# RANDOM MANIFOLDS

**Goal of regularization:** To smooth the function and ease optimization but with the side effect of creating smooth manifolds, lacking information about holes and such.

**Solution:** Random manifolds.

$$\overline{\mathbf{M}}_{\mathbf{z}} = \mathbb{E}_{p(\epsilon)}[\mathbf{M}_{\mathbf{z}}] = \left(\mathbf{J}_{\mathbf{z}}^{(\mu)}\right)^{\top} \left(\mathbf{J}_{\mathbf{z}}^{(\mu)}\right) + \left(\mathbf{J}_{\mathbf{z}}^{(\sigma)}\right)^{\top} \left(\mathbf{J}_{\mathbf{z}}^{(\sigma)}\right)$$



Deterministic  
Embedded Manifold

Stochastic/Random  
Embedded Manifold

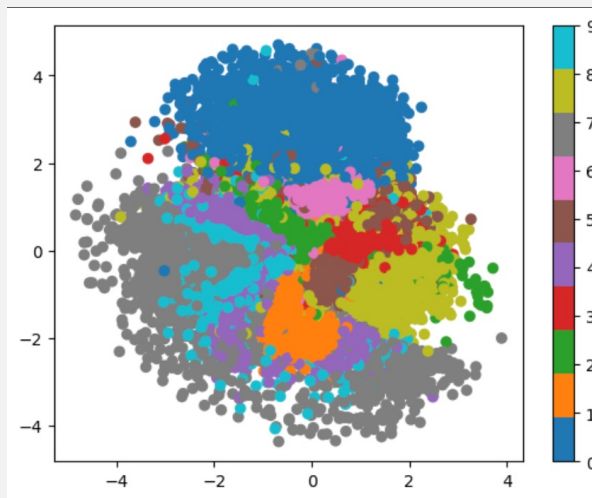
# ISOMETRIC REPRESENTATION VAE (IRVAE)

**Goal:** Preserve Euclidean Relationship

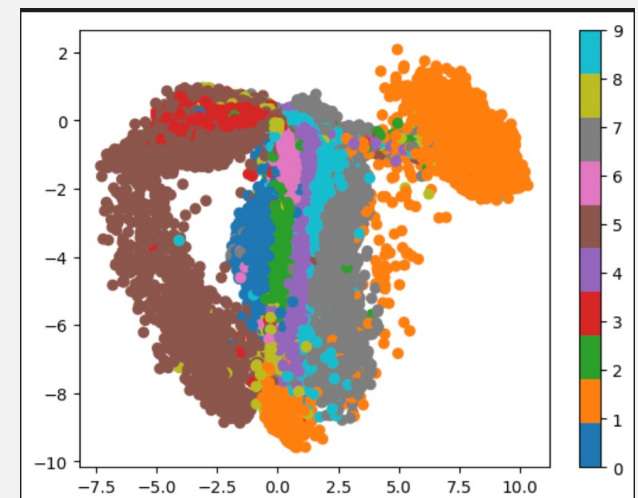
**How:** Isotropic Learnable Gaussian + Scaled  
Isometric Regularization Term in Loss Function

**Paper:**

Regularized Autoencoders for Isometric  
Representation Learning – Lee et. al 2022

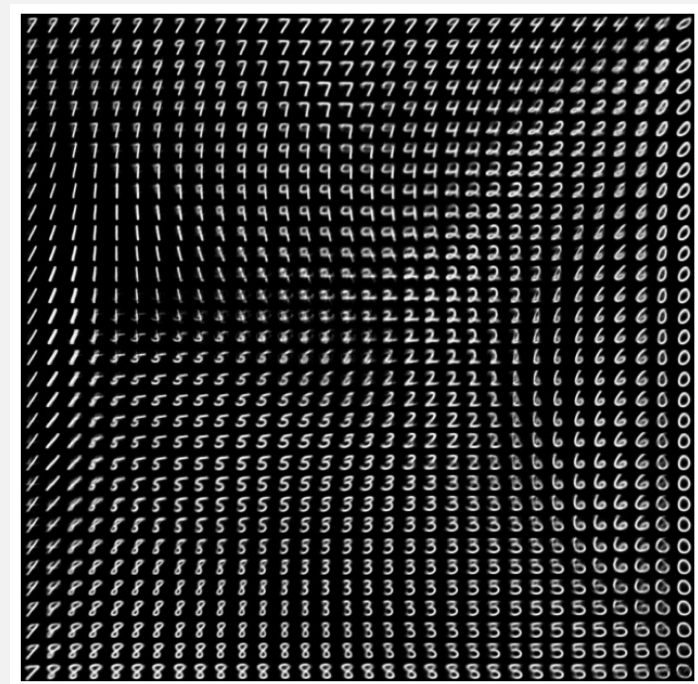
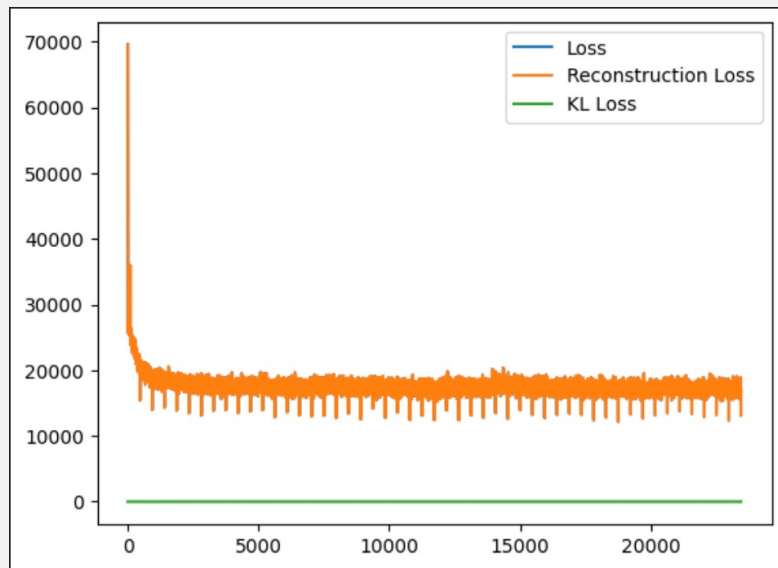


Simple VAE



IRVAE

# TRAINING IRVAE



## RADIAL BASIS FUNCTION NETWORK (RBFNET)

**What:** Network that combines radial kernels to approximate a target function.

### How

Initialize the centers of radial kernels using k-means. Then, bandwidths ("decay rate") for the radial kernels are generated. Finally, inference involves calculating the distance between a data point and the centers and weighting it using the bandwidths.

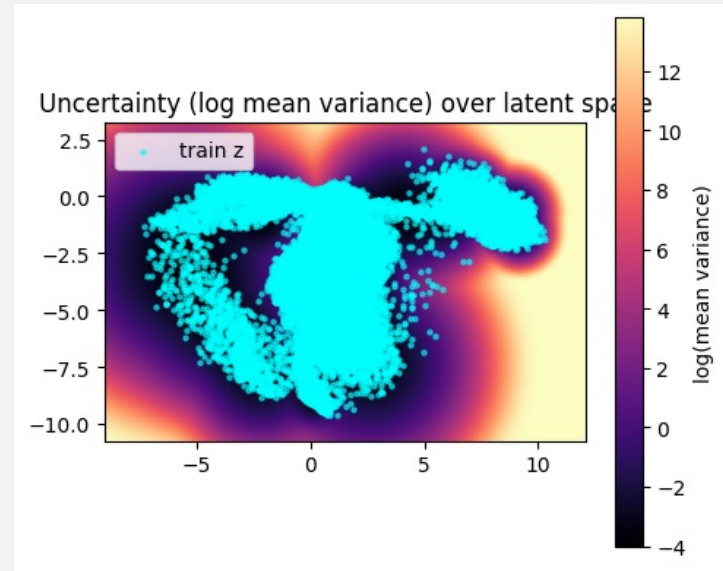
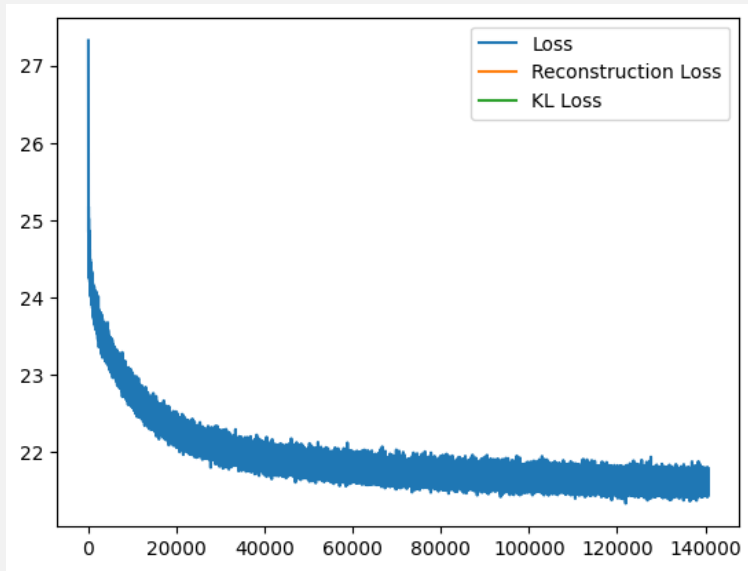
### Why

- Controllable decay rate and tolerance towards out-of-distribution data.
- Often used to model VAE uncertainty

### Paper Inspiration:

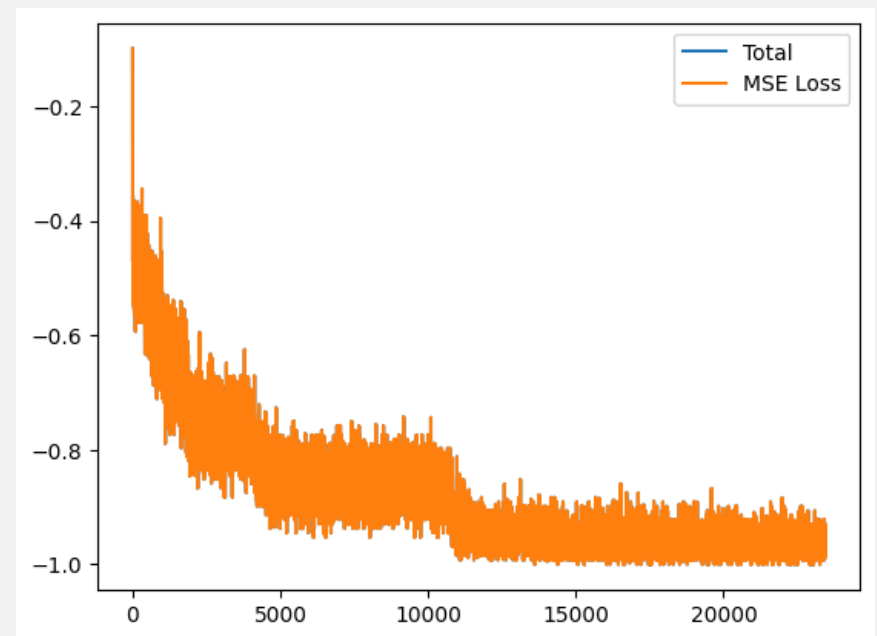
Back to the Future: Radial Basis Function Network Revisited – Que et. al 2020

## TRAINING RBFNET



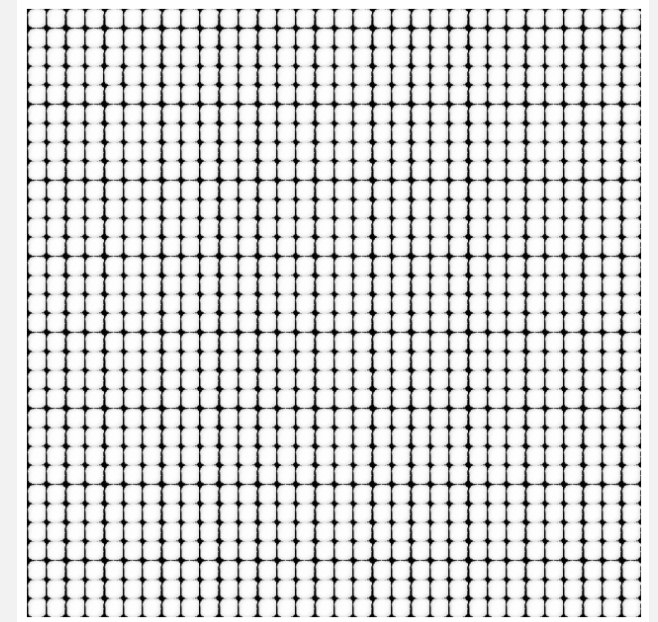
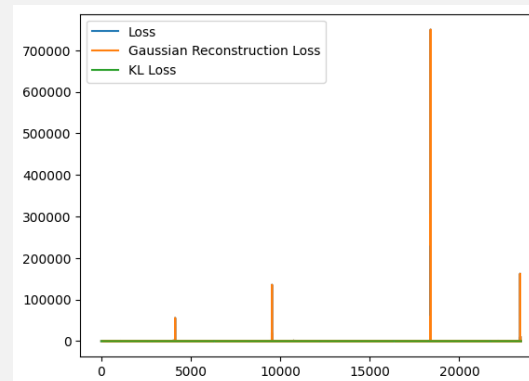
## DISCRIMINATIVE CNN MODEL

```
CustomCNN(  
  (layers): Sequential(  
    (0): Conv2d(1, 16, kernel_size=(3, 3), stride=(1, 1))  
    (1): LayerNorm((26,), eps=1e-05, elementwise_affine=True)  
    (2): ReLU(inplace=True)  
    (3): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)  
    (4): Flatten(start_dim=1, end_dim=-1)  
    (5): LazyLinear(in_features=0, out_features=128, bias=True)  
    (6): ReLU(inplace=True)  
    (7): Linear(in_features=128, out_features=10, bias=True)  
    (8): Softmax(dim=1)  
  )  
  (loss_func): NLLLoss()  
)
```



## OTHER (FAILED) EXPERIMENTS

- Gaussian Uncertainty VAE
- Hierarchical VAE
- Gaussian Process Variational Latent Model (GPVLM)

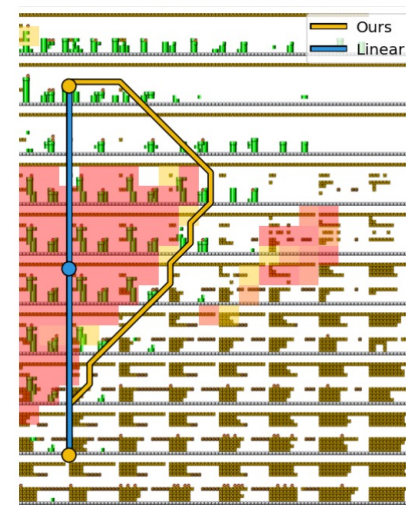


## COOL RELATED WORK

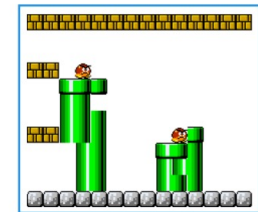
Mario Plays on a Manifold: Generating Functional Content in Latent Space through Differential Geometry - Gonzalez-Duque et. al 2022

Data Generation in Low Sample Size Setting Using Manifold Sampling and a Geometry-Aware VAE - Chadebec et. al 2021

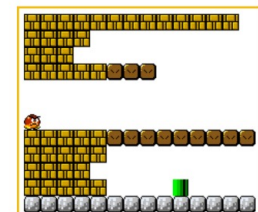
Latent space of Super Mario Bros



Not Functional

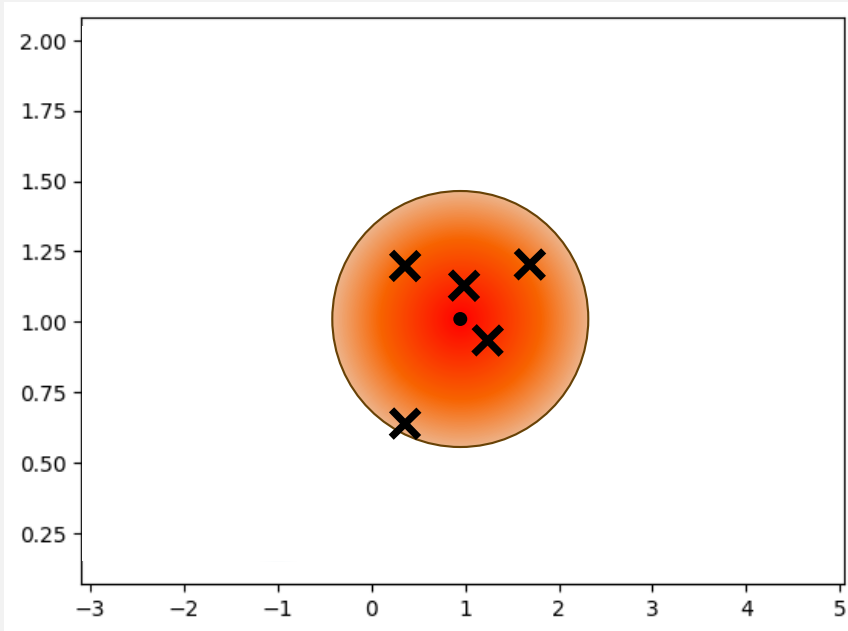


Functional



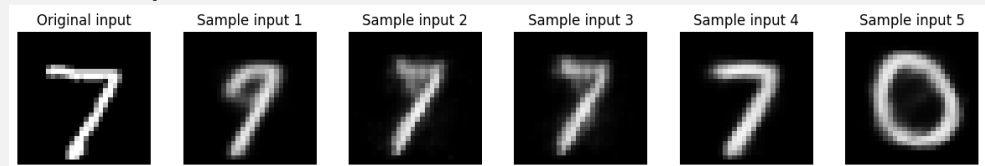
# SAMPLING

# GAUSSIAN SAMPLING

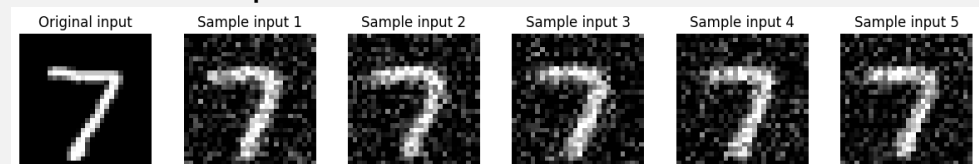


## VAE Examples

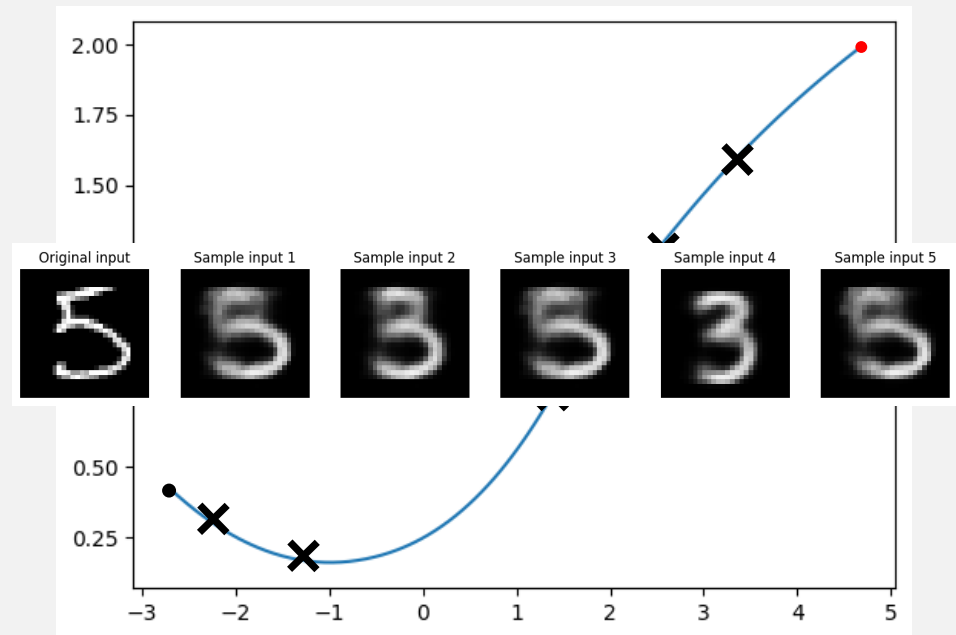
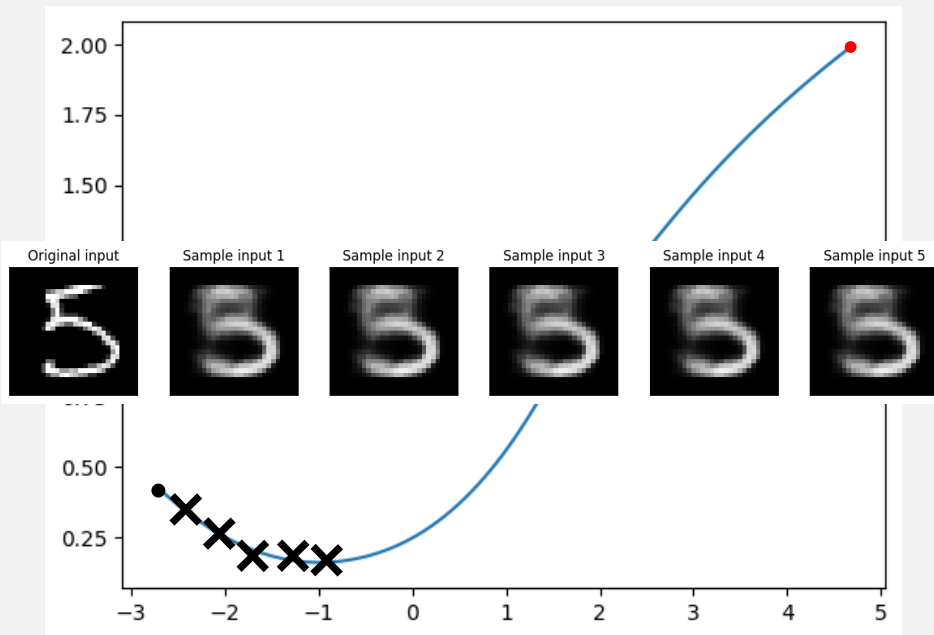
### Latent Space



### Observation Space



# GEODESIC SAMPLING



Red – Latent sample from  
predicted class

# UNCERTAINTY

## METRICS

- Variance
- Coefficient of Variation (Relative STD)
- Entropy
- Percent Entropy

$$CV = \frac{\sigma}{\mu}$$

$$H(X) := - \sum_{x \in \mathcal{X}} p(x) \log p(x)$$

# UNCERTAINTY FROM GEODESIC SAMPLING

## Mispredictions

```
Prediction on samples: [3, 5, 3, 3, 0]
Variance of prediction: 2.56
Coefficient of variation: 0.5714285693877552
Entropy of prediction: 2.270611524581909
Entropy as percent: 98.6%
```

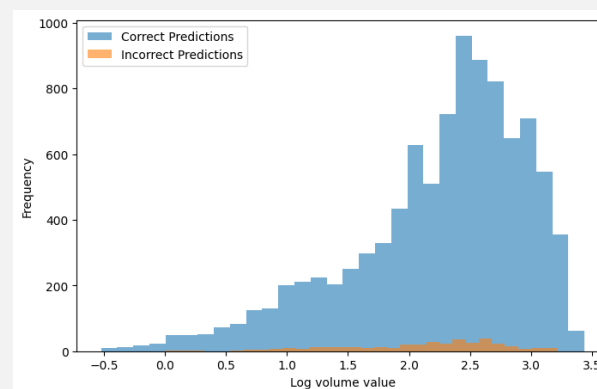
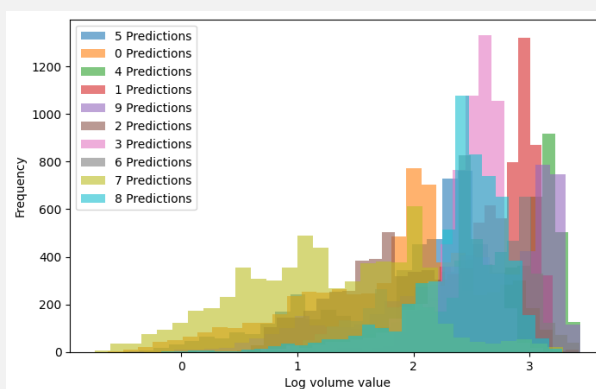
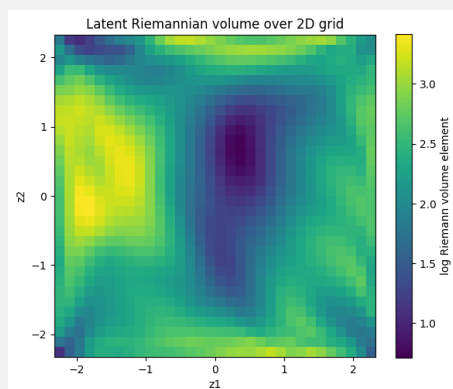
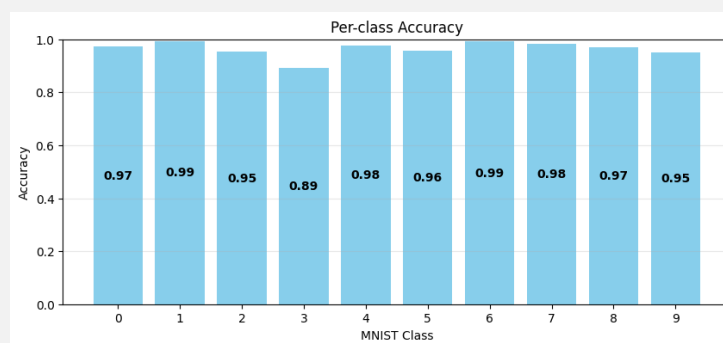
```
----- Average uncertainty values for mispredictions -----
Variance of prediction: 1.0560000000000003
Coefficient of variation: 0.09364875324704096
Entropy of prediction: 2.2392616271972656
Entropy as percent: 97.2%
```

## Correct Predictions

```
Prediction on samples: [7, 7, 7, 7, 7]
Variance of prediction: 0.0
Coefficient of variation: 0.0
Entropy of prediction: 2.2291808128356934
Entropy as percent: 96.8%
```

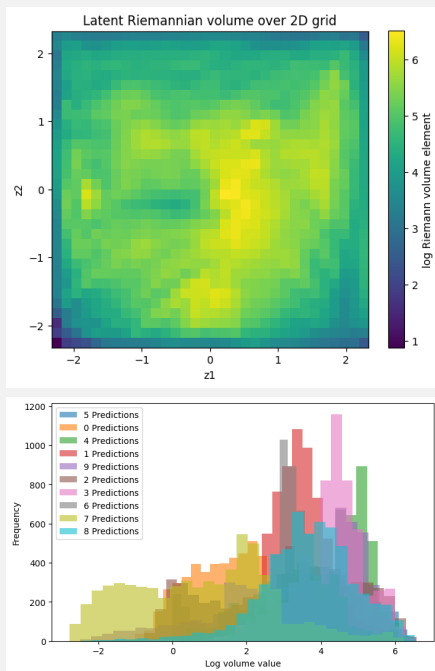
```
----- Average uncertainty values for correct predictions -----
Variance of prediction: 0.09600000000000002
Coefficient of variation: 0.0544331050927756
Entropy of prediction: 2.2332890033721924
Entropy as percent: 97.0%
```

# UNCERTAINTY FROM LOG OF VOLUME METRIC

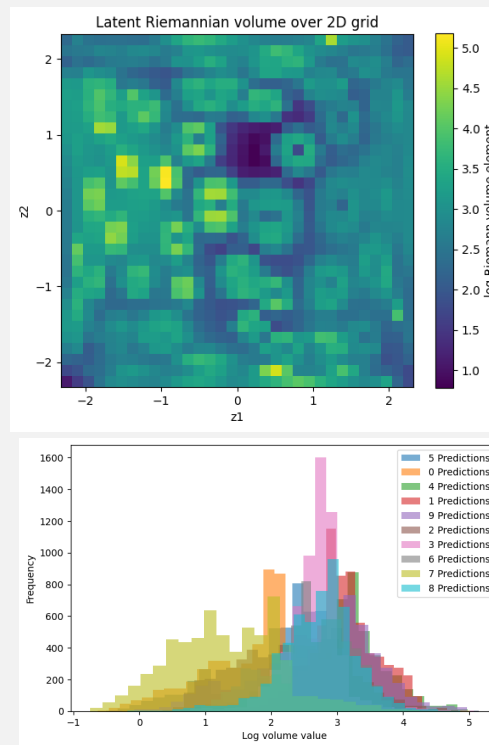


# UNCERTAINTY FROM LOG OF VOLUME METRIC

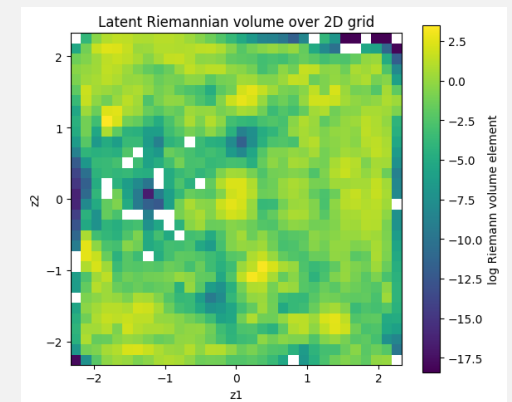
IRVAE



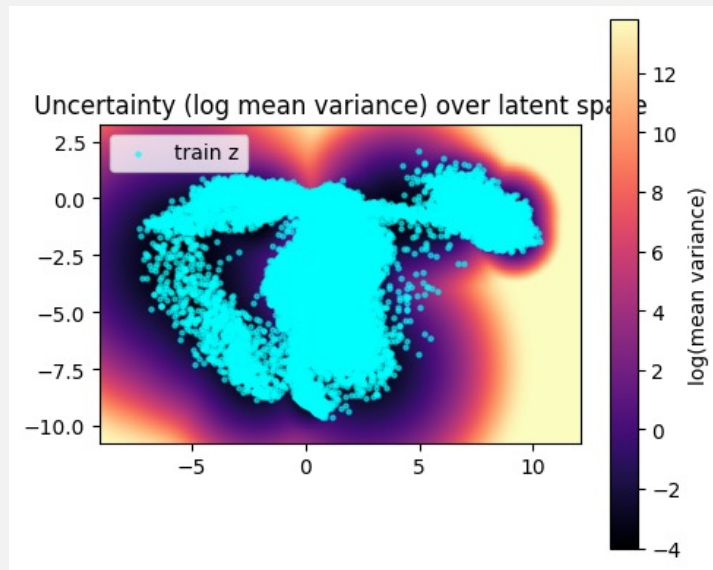
VAE+RBF (Random Manifold)



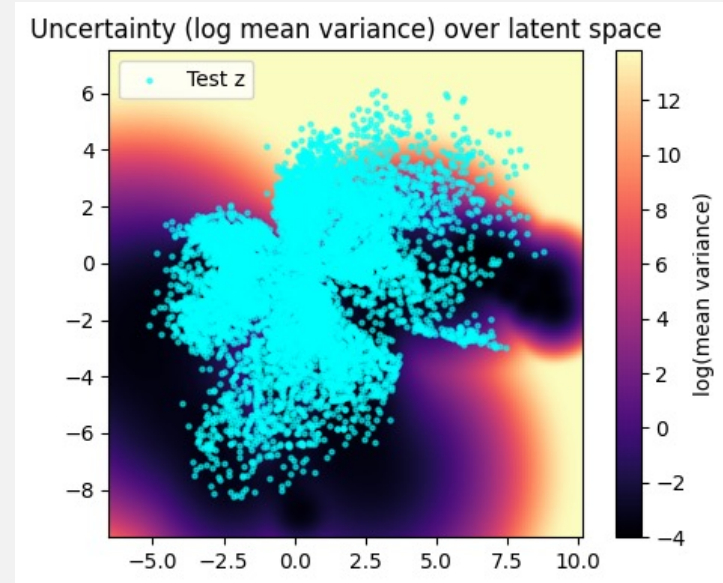
RBFNet



## UNCERTAINTY FROM TRAINED RBFNET



Train Dataset



Test Dataset

## FUTURE DIRECTIONS

- Experiment with using the average intraclass geodesic length as a metric for uncertainty
- GPLVM Experimentation
- Look at higher dimensionality latent spaces

## REFERENCES

- Q. Que and M. Belkin, "Back to the Future: Radial Basis Function Network Revisited," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 8, pp. 1856-1867, 1 Aug. 2020, doi: 10.1109/TPAMI.2019.2906594.
- Lee, Y., Yoon, S., Son, M., & Park, F. C. (2022). *Regularized Autoencoders for Isometric Representation Learning*. In *Proceedings of the International Conference on Learning Representations (ICLR)*. Retrieved from <https://openreview.net/forum?id=mQxt8I7jL04>
- Arvanitidis, G., Hansen, L. K., & Hauberg, S. (2021). Latent space oddity: On the curvature of deep generative models. arXiv. <https://arxiv.org/abs/1710.11379>
- Hauberg, S. (2021, November). Differential geometry for generative modeling [Tutorial, ACML 2021]. The Asian Conference on Machine Learning. [https://www2.compute.dtu.dk/~sohau/weekendwithbernie/Differential\\_geometry\\_for\\_generative\\_modeling.pdf](https://www2.compute.dtu.dk/~sohau/weekendwithbernie/Differential_geometry_for_generative_modeling.pdf)
- Hauberg, S. (2018). Only Bayes should learn a manifold: On the estimation of differential geometric structure from data [Preprint]. arXiv. <https://arxiv.org/abs/1806.04994>
- Chen, N., Klushyn, A., Kurle, R., Jiang, X., Bayer, J., & van der Smagt, P. (2018). Metrics for deep generative models. In A. Storkey & F. Perez-Cruz (Eds.), *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics (AISTATS)*, Vol. 84 (pp. 1540-1550). PMLR. <https://proceedings.mlr.press/v84/chen18e/chen18e.pdf>
- Arvanitidis, G., Hauberg, S., Hennig, P., & Schober, M. (2019). Fast and robust shortest paths on manifolds learned from data. In K. Chaudhuri & M. Sugiyama (Eds.), *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics (AISTATS)*, Vol. 89 (pp. 1506-1515). PMLR. <https://proceedings.mlr.press/v89/arvanitidis19a/arvanitidis19a.pdf>
- Arvanitidis, G., Hauberg, S., & Schölkopf, B. (2021). Geometrically enriched latent spaces. In A. Banerjee & K. Fukumizu (Eds.), *Proceedings of the Twenty-Fourth International Conference on Artificial Intelligence and Statistics (AISTATS)*, Vol. 130 (pp. 631-639). PMLR. <https://proceedings.mlr.press/v130/arvanitidis21a/arvanitidis21a.pdf>
- M. González-Duque, R. B. Palm, S. Hauberg and S. Risi, "Mario Plays on a Manifold: Generating Functional Content in Latent Space through Differential Geometry," 2022 IEEE Conference on Games (CoG), Beijing, China, 2022, pp. 385-392, doi: 10.1109/CoG51982.2022.9893612.
- Pouplin, A., Eklund, D., Ek, C. H., & Hauberg, S. (2023). Identifying latent distances with Finslerian geometry. *Transactions on Machine Learning Research (TMLR)*. <https://arxiv.org/abs/2212.10010>
- Syrota, S., Moreno-Muñoz, P., & Hauberg, S. (2024). Decoder ensembling for learned latent geometries. <https://arxiv.org/abs/2408.07507>
- Chadebec, C., & Allasonnière, S. (2021). Data generation in low sample size setting using manifold sampling and a geometry-aware VAE (arXiv:2103.13751). arXiv. <https://doi.org/10.48550/arXiv.2103.13751>
- Kingma, D. P., & Welling, M. (2013). Auto-Encoding Variational Bayes. arXiv. <https://arxiv.org/abs/1312.6114>
- Q. Que and M. Belkin, "Back to the Future: Radial Basis Function Network Revisited," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 8, pp. 1856-1867, 1 Aug. 2020, doi: 10.1109/TPAMI.2019.2906594
- Hu, Y., Zhou, Z., Yang, P., Zhao, X., & Zhang, P. (2024). Classification of ground-based auroral images by learning deep tensor feature representation on Riemannian manifold. *Journal of Geophysical Research: Machine Learning and Computation*, 1, e2023JH000109. <https://doi.org/10.1029/2023JH000109>
- <https://github.com/Gabe-YHLee/IRVAE-public>
- Oskar Jørgensen – Visualiseringskode
- <https://github.com/MachineLearningLifeScience/stochman>