THE UNIVERSITY of TENNESSEE KNOXVILLE

Lecture 35: VAE Applications

Instructor: Sergei V. Kalinin

VAE applications for real problems

- We often deal with complex data sets containing the information on physics of objects we seek to understand
- This can be spectral data sets (EELS in STEM, CITS in STM, complex spectroscopies in PFM) or single, multimodal, or hyperspectral images
- Often, we seek to reduce dimensionality and explore similarities in these data sets.

Two things matter: descriptors and ML method

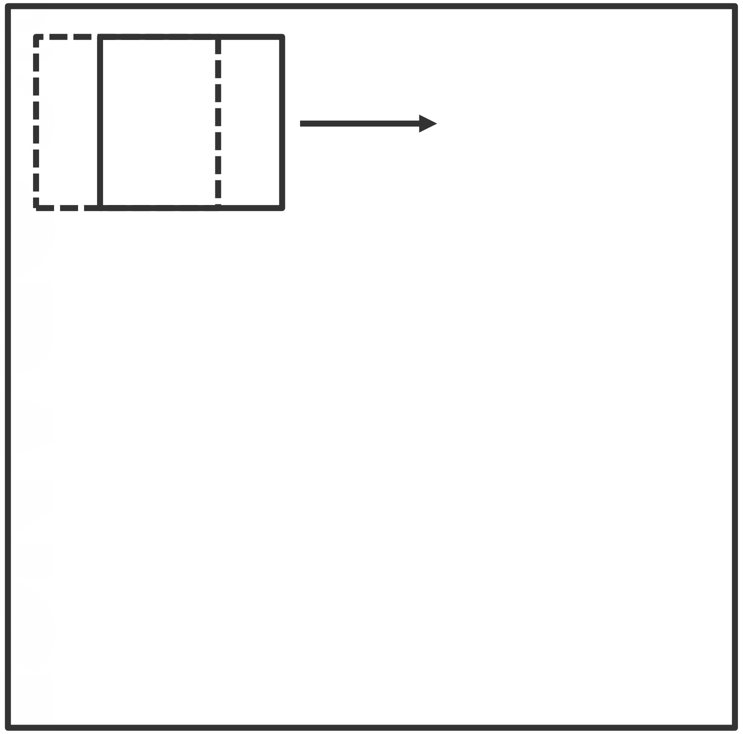
- In analysis of EELS or CITS data, very often our descriptor is the spectrum at each pixel. Typical analysis will be either linear or non-linear dimensionality reduction or clustering:
  - Linear dimensionality reduction: PCA, NMF, BLU
  - Clustering: k-means, GMM
  - Manifold learning: ISO, UMAP, tSNE, DBSCAN
  - Neural nets: SOFM, AEs, VAEs

- Typical result will be the components (representing behavior), and loading maps representing spatial variability of these behaviors. By construct, components will not depend on the relative spatial positions of pixel.

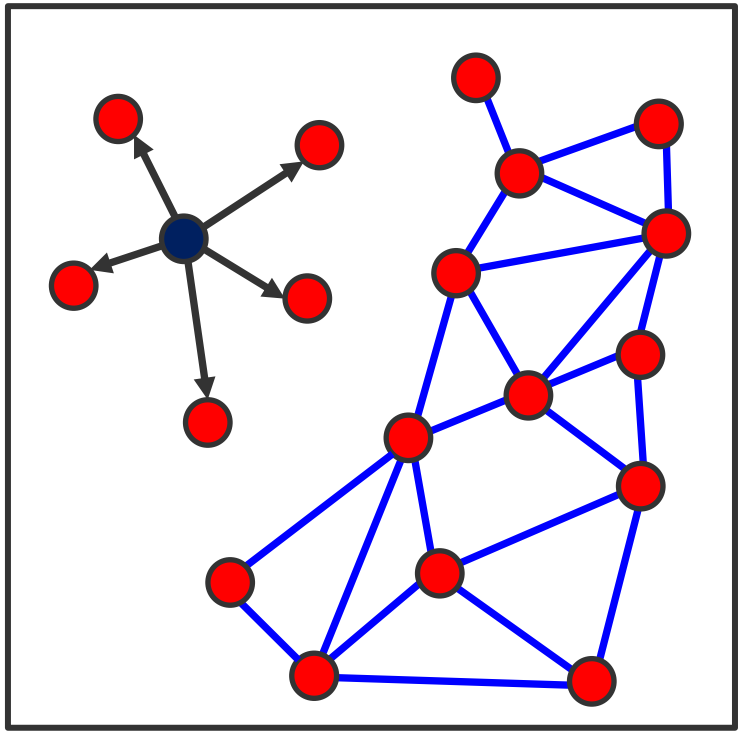
Describing the building blocks

- The classical physical descriptions (symmetry, etc) can be defined locally only in Bayesian sense
- We can argue that local descriptors are simple, if not necessarily known
- And the rules that guide their emergence are also simple, if not known

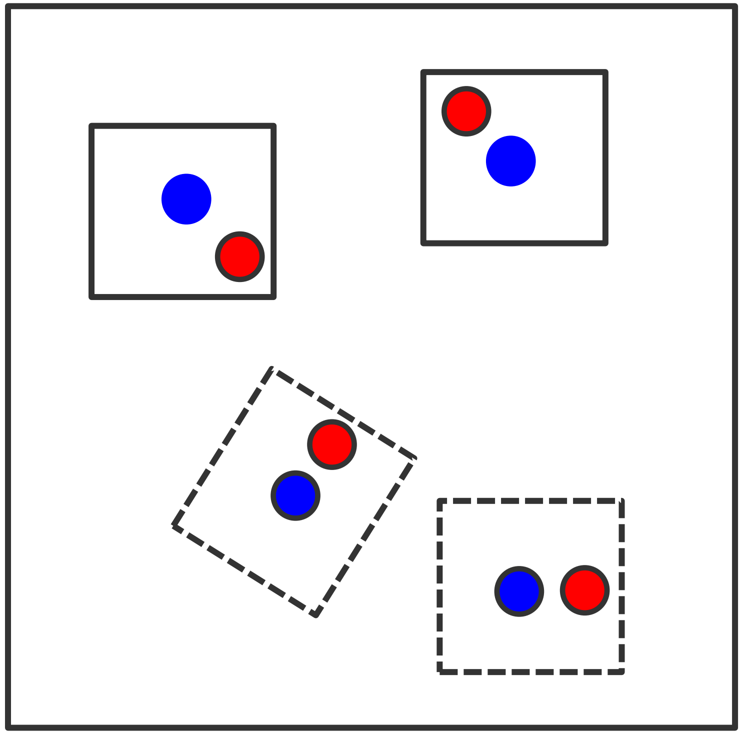
Continuous translational symmetry



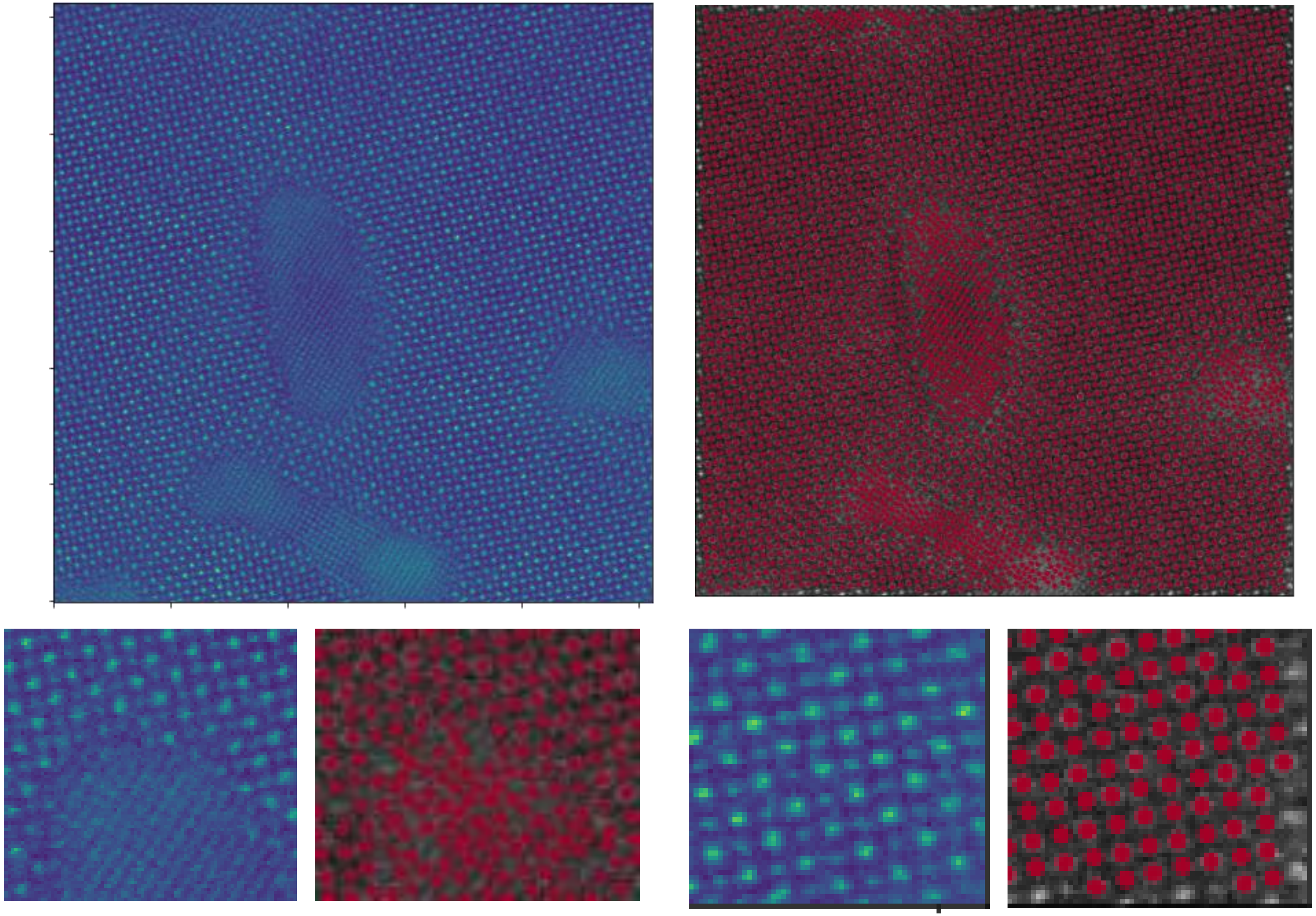
Atom based descriptions



Localized sub-images

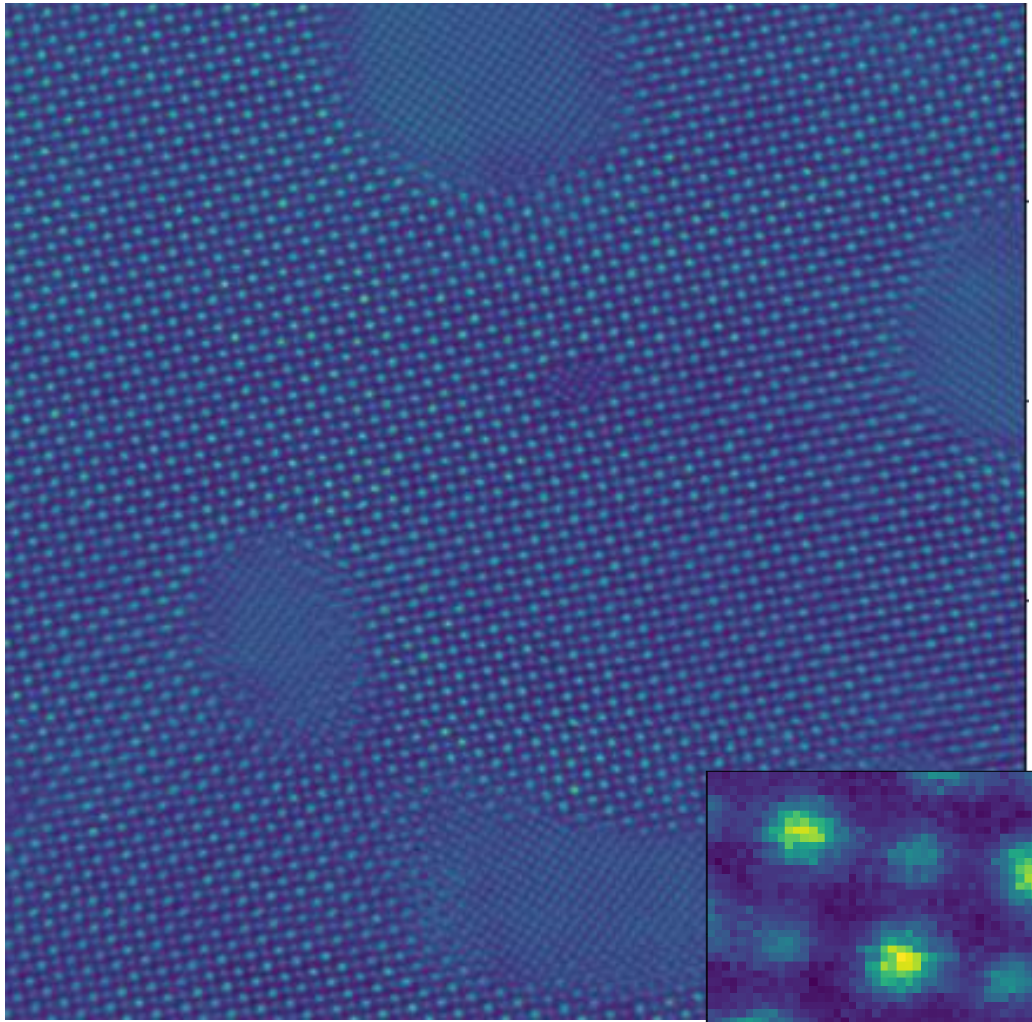


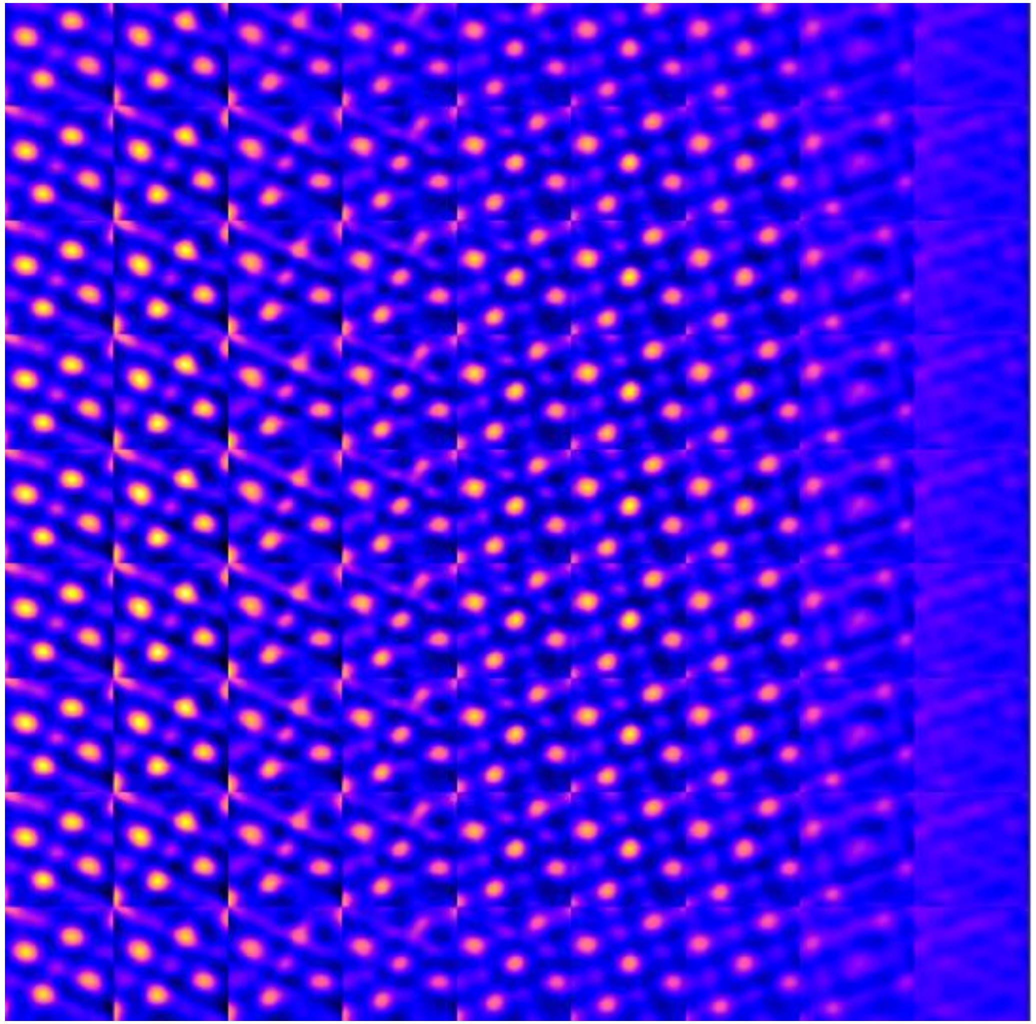
Let’s put it all together!



Step 1: Find all atoms (or all that you can) – use maximum finders, blob-log, or DCNNs
Step 2: Create descriptors – patches centered on atoms. Keep track on what part of image (or stack) it has came from

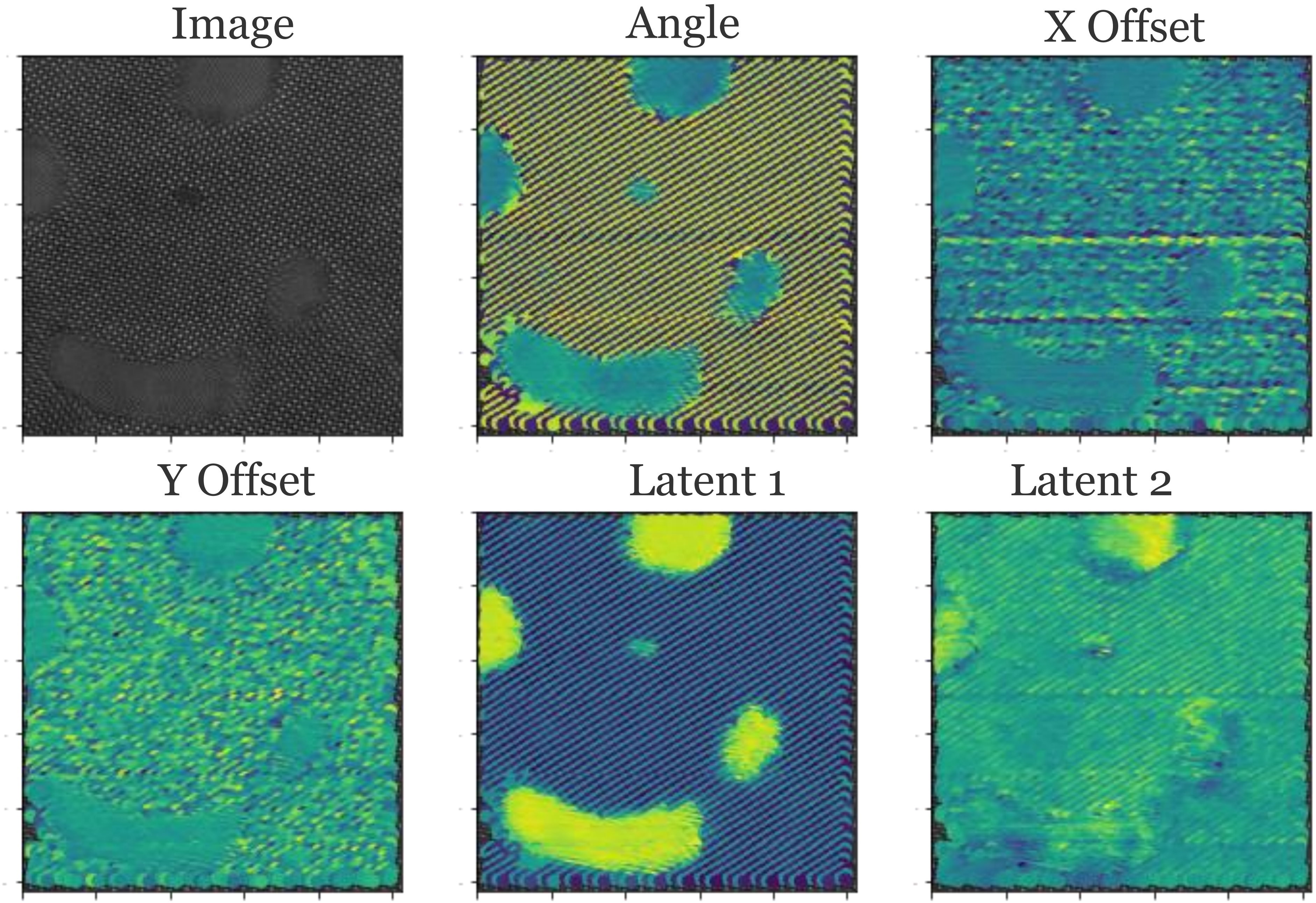
Analysis of the NiO-LSMO



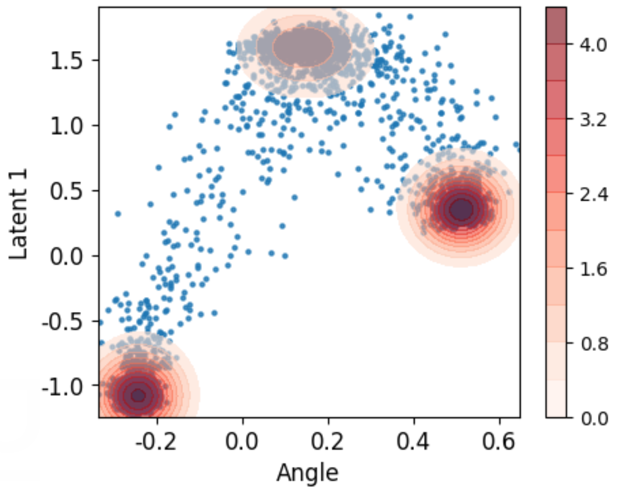


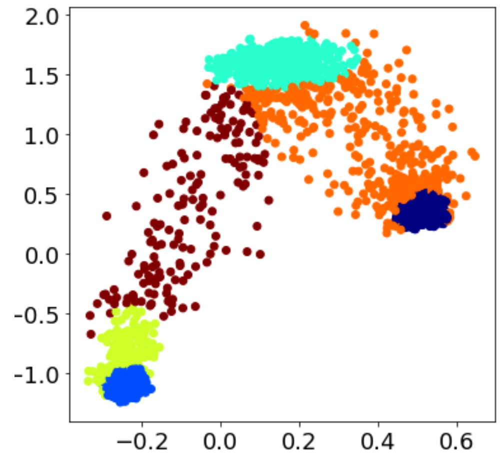
Data – M. Chisholm

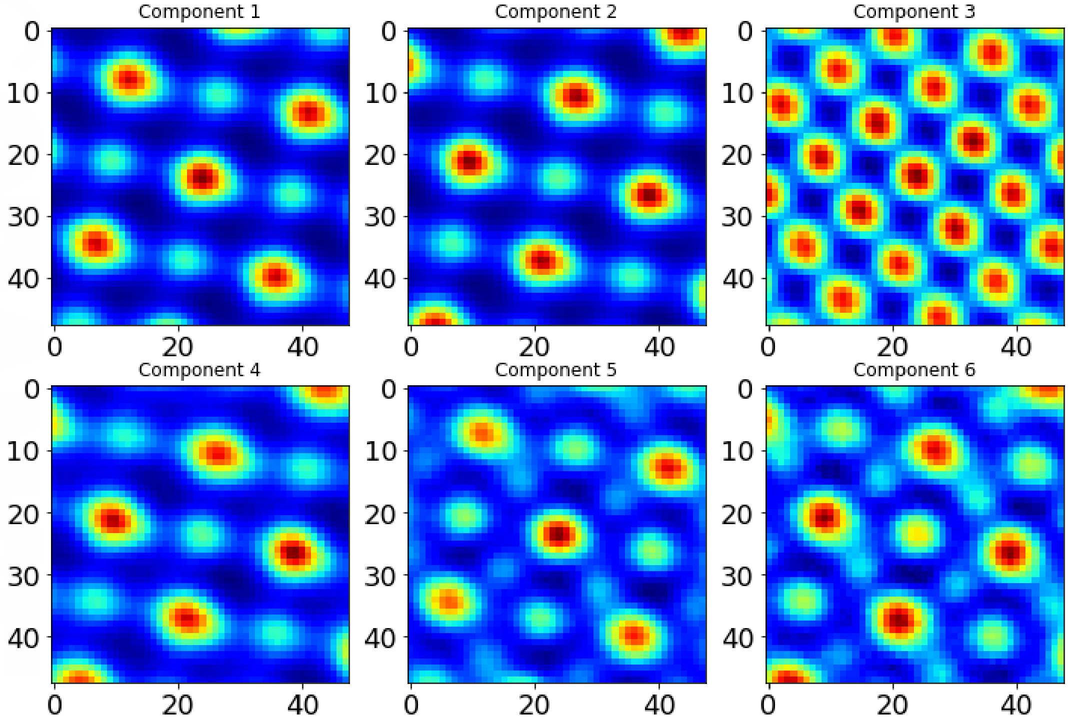
Let’s look at latent space



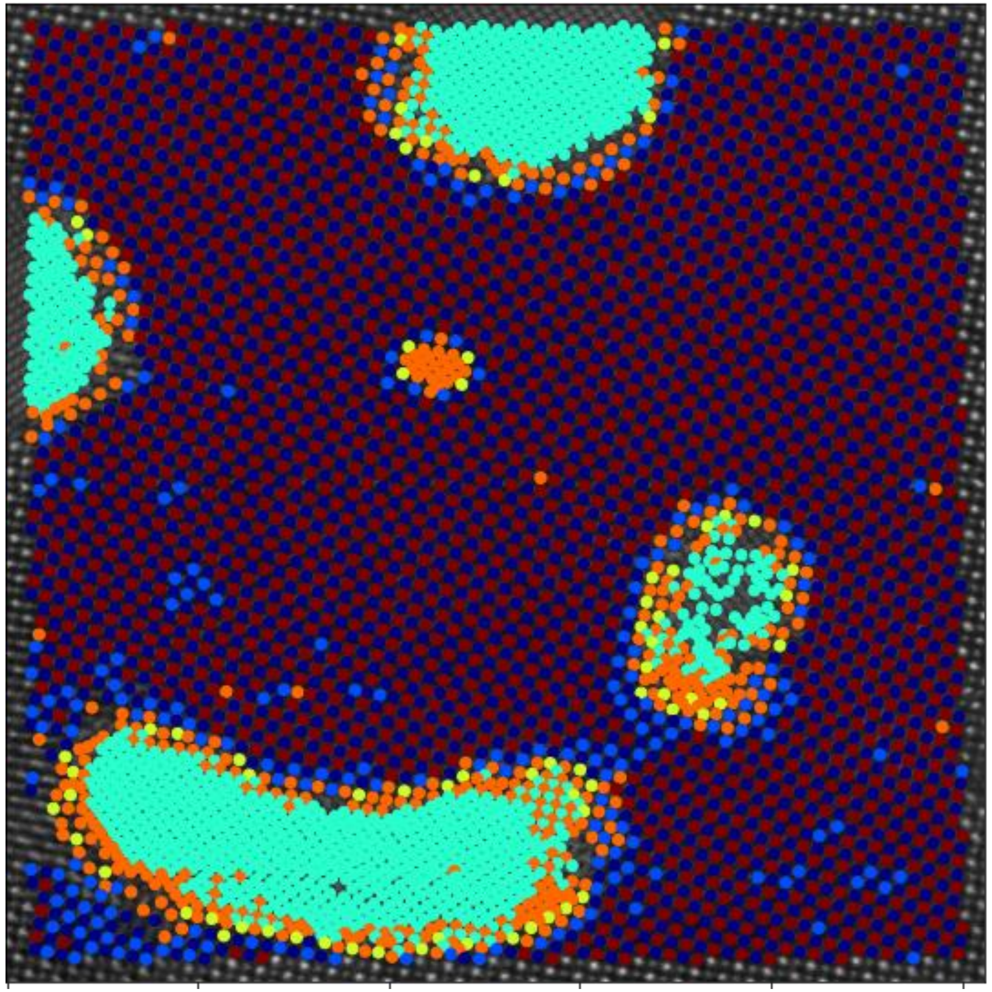
Exploring latent distributions





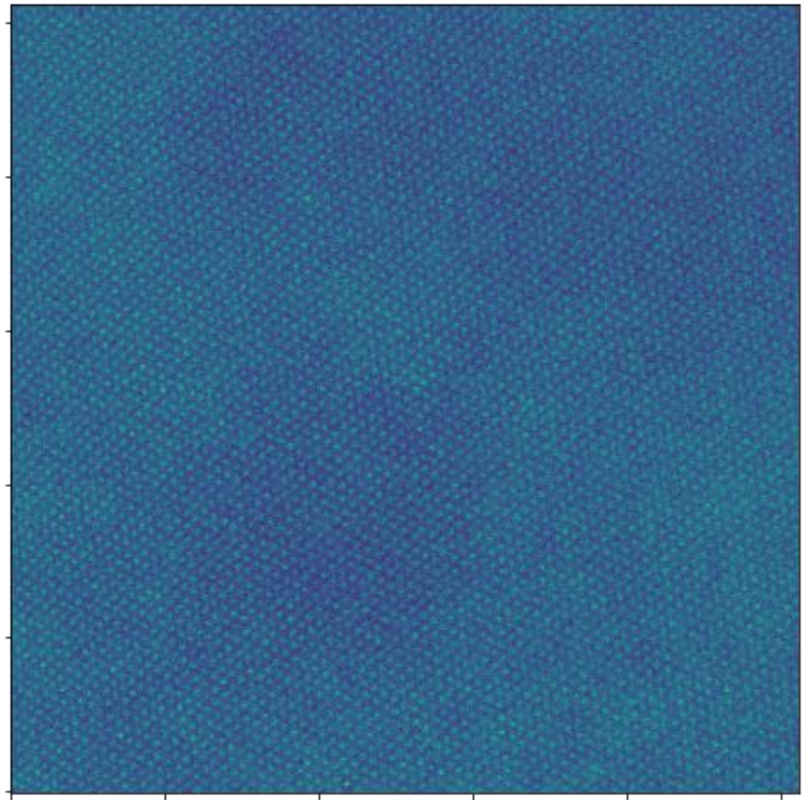


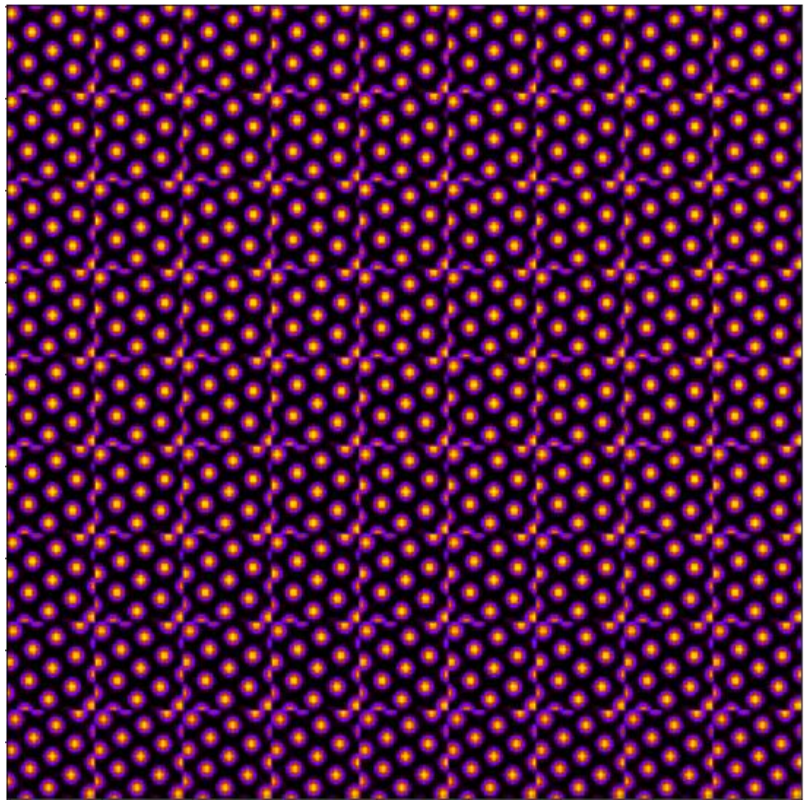
Labeled image



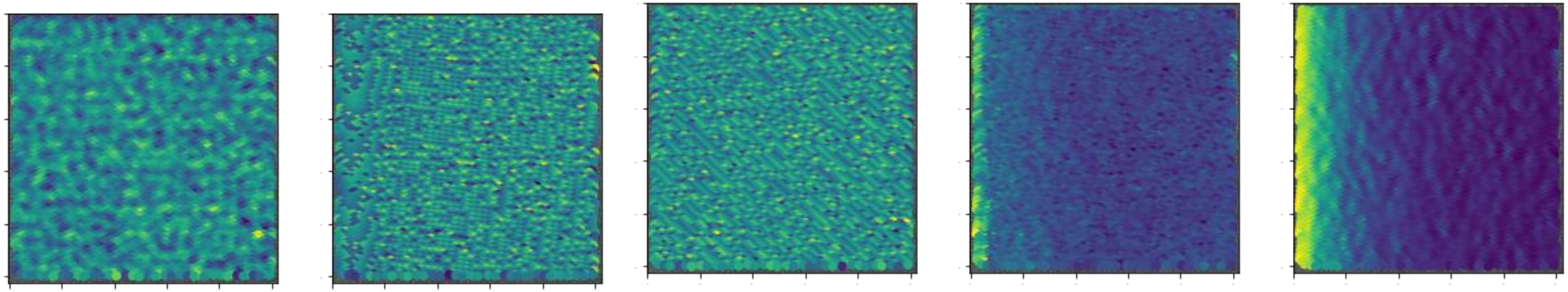
- Classes and variability are mixed in latent space
- Disentangling of representation

Out of curiosity: single crystal?



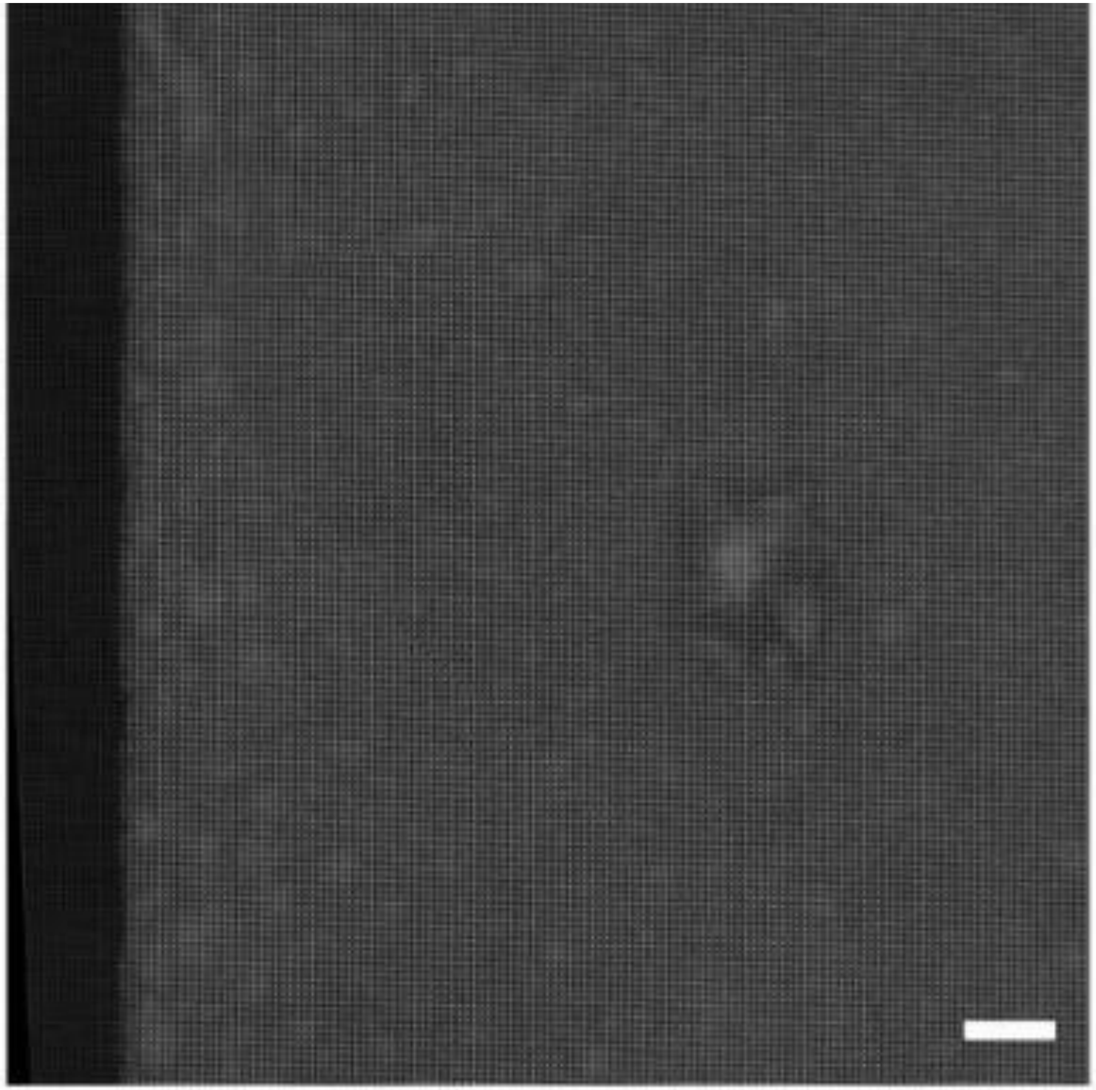


Rotation Offset X Offset Y Latent 1 Latent 2

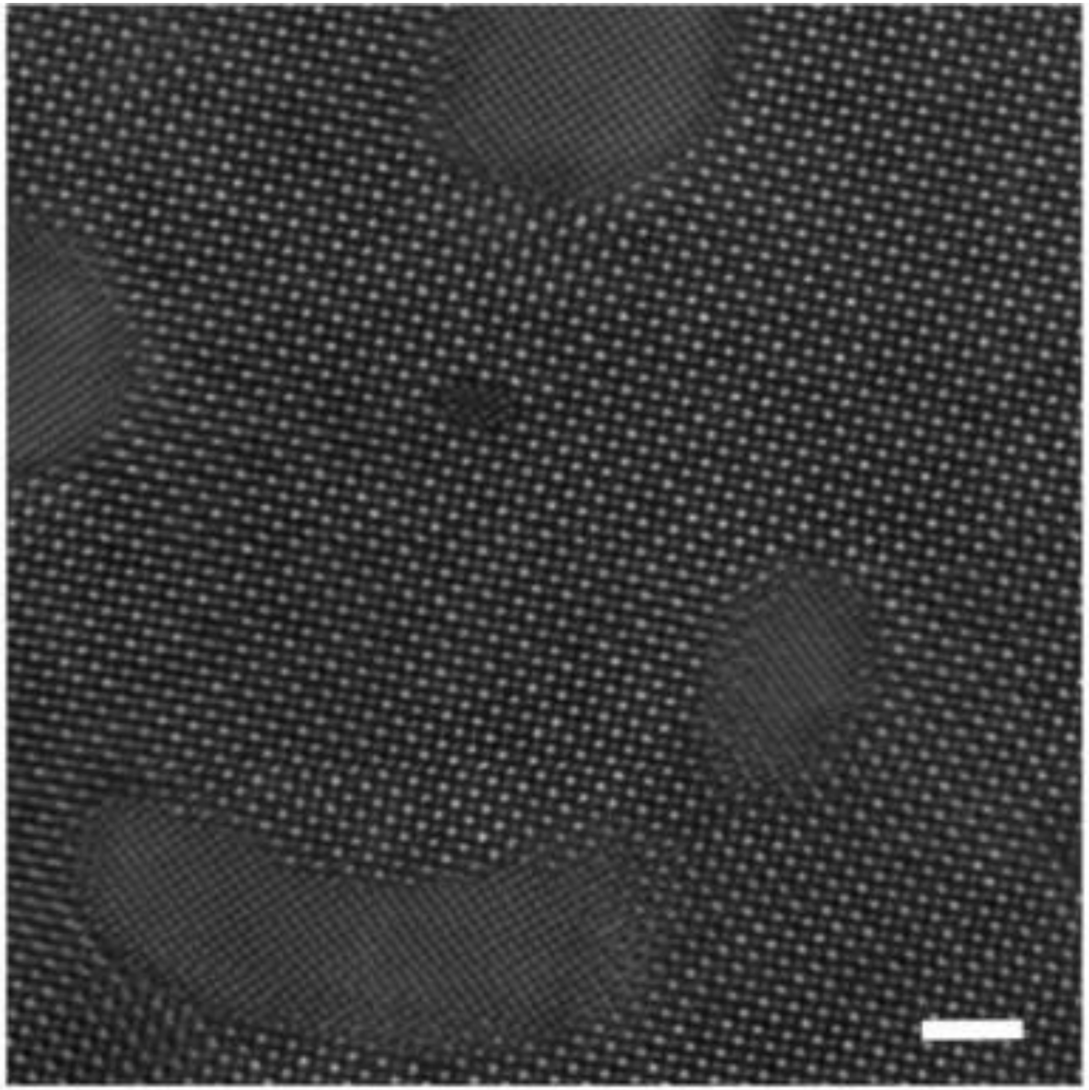


VAE without Atom Finding

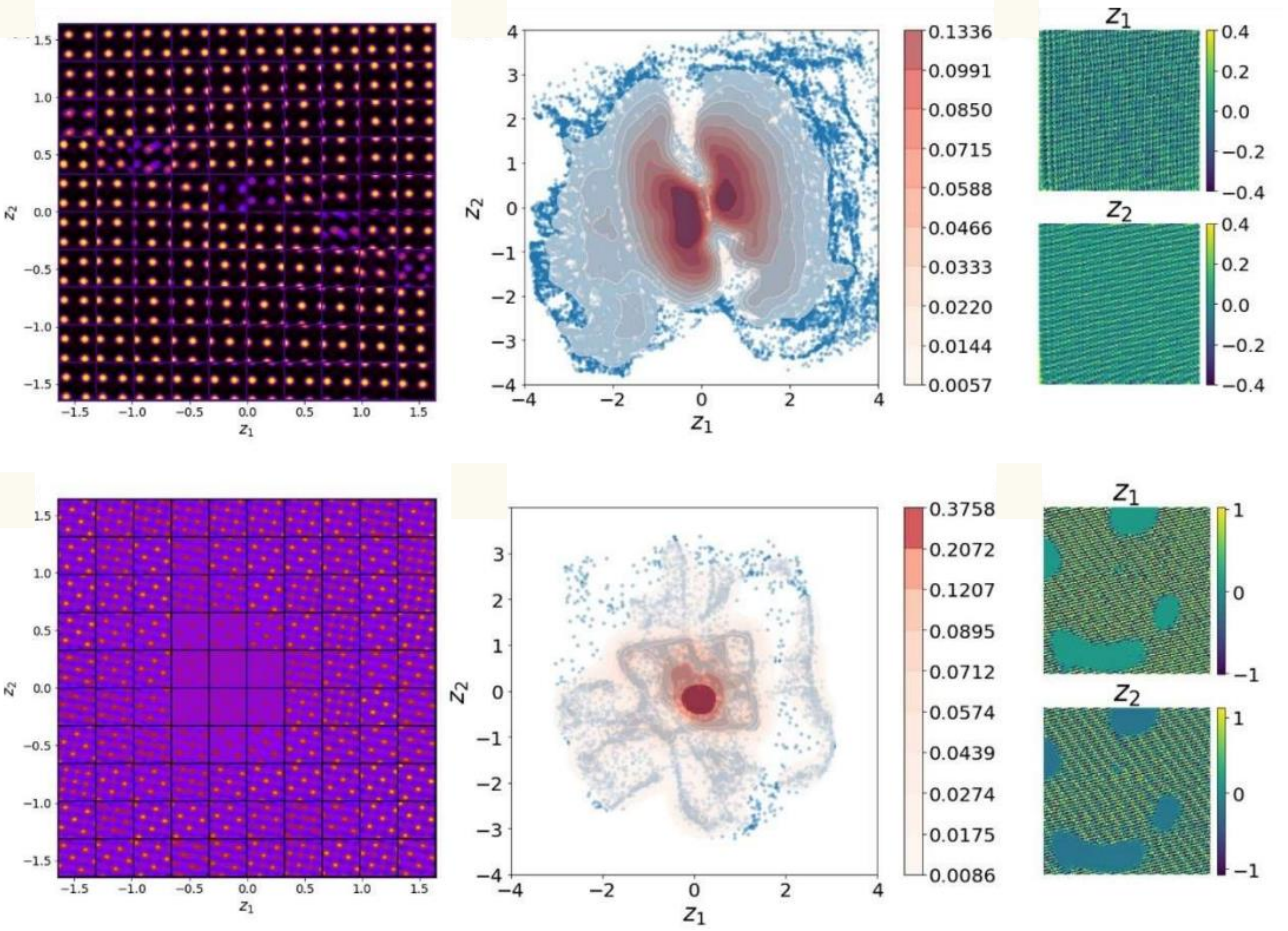
Ferroelectric $BiFeO_3$



$NiO - La_xSr_{1-x}MnO_3$

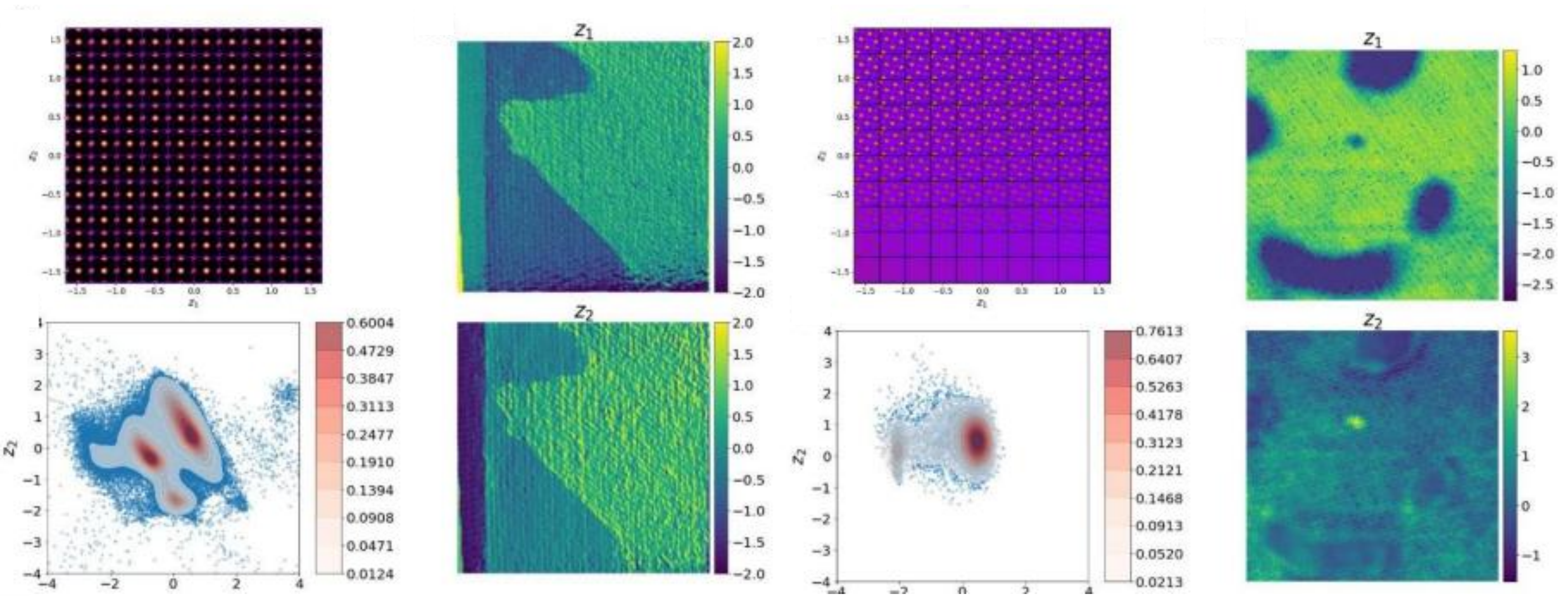


Simple VAE



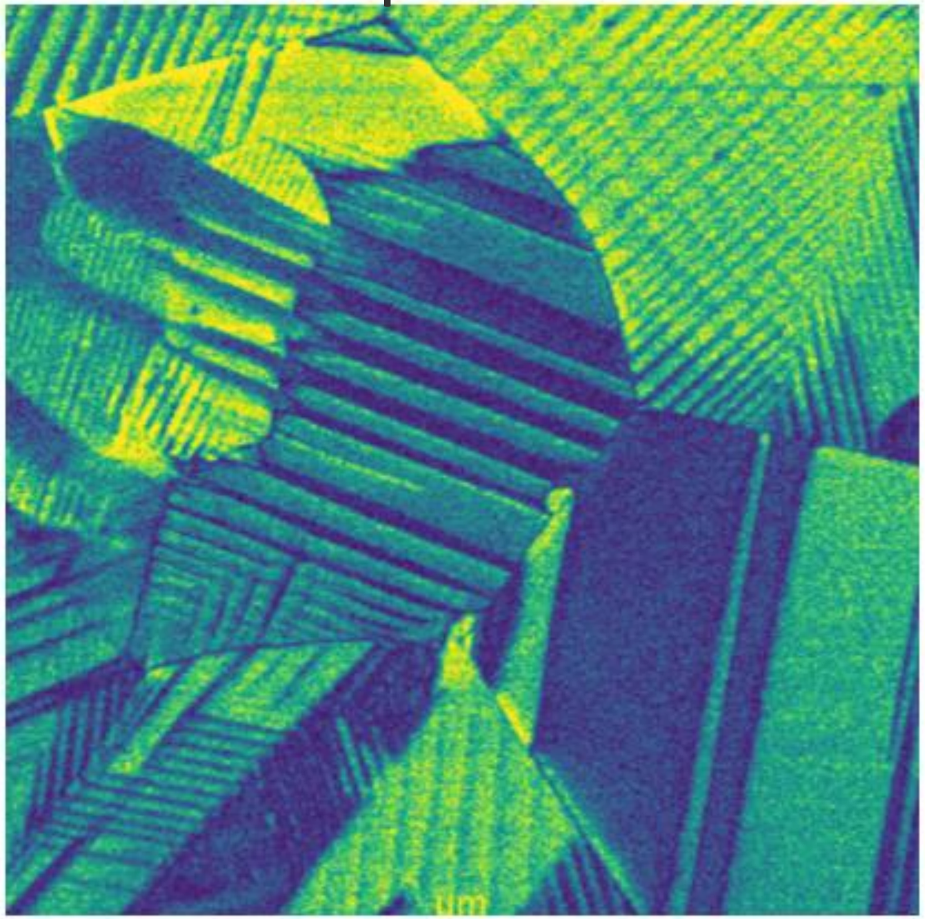
z2
z1
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
4
3
2
1
0
−1
−2
−3
−4
0.1336
0.0991
0.0850
0.0715
0.0588
0.0466
0.0333
0.0220
0.0144
0.0057
0.4
0.2
0.0
−0.2
−0.4
0.3758
0.2072
0.1207
0.0895
0.0712
0.0574
0.0439
0.0274
0.0175
0.0086
1
0
−1

Shift VAE: Translational Invariance

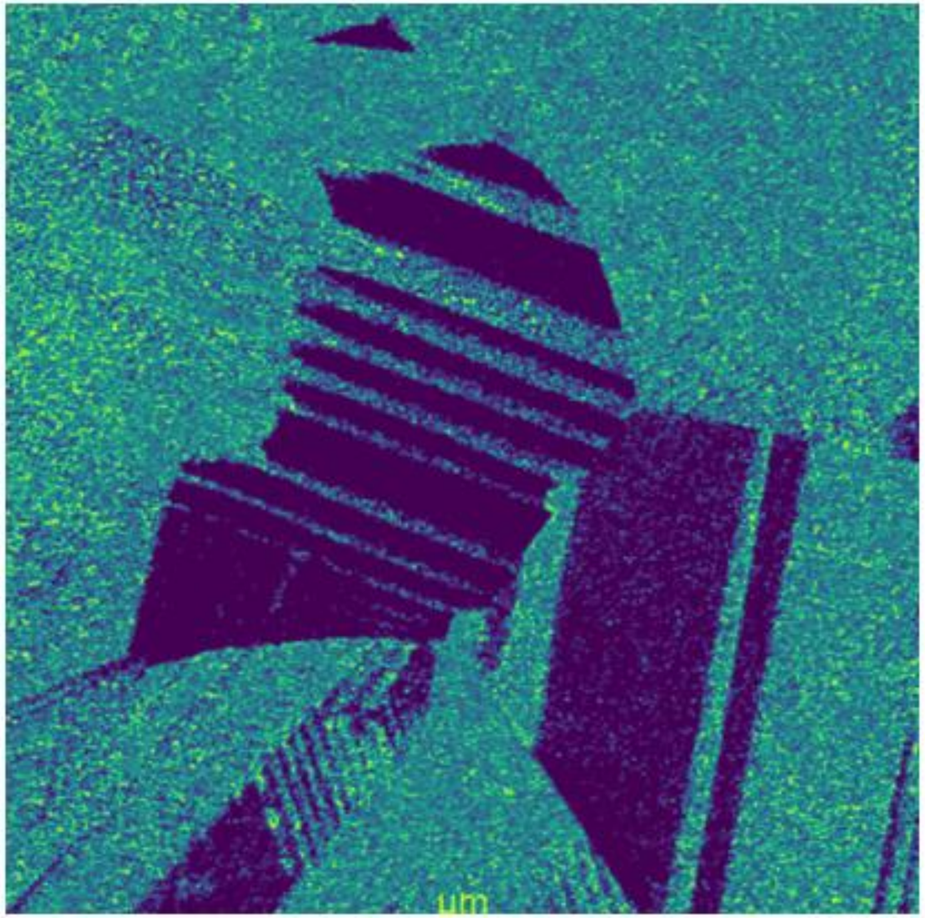


Ferroelectric domain and domain walls

Amplitude

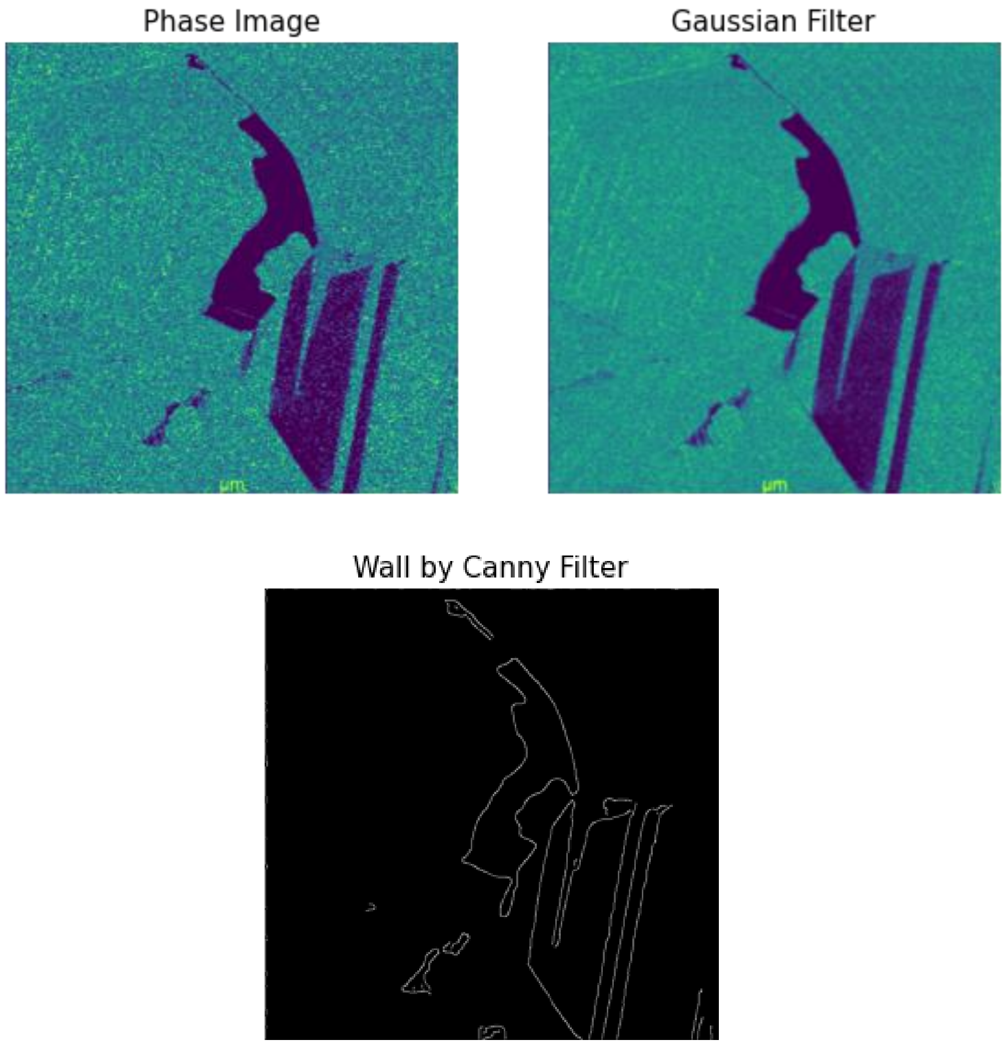


Phase

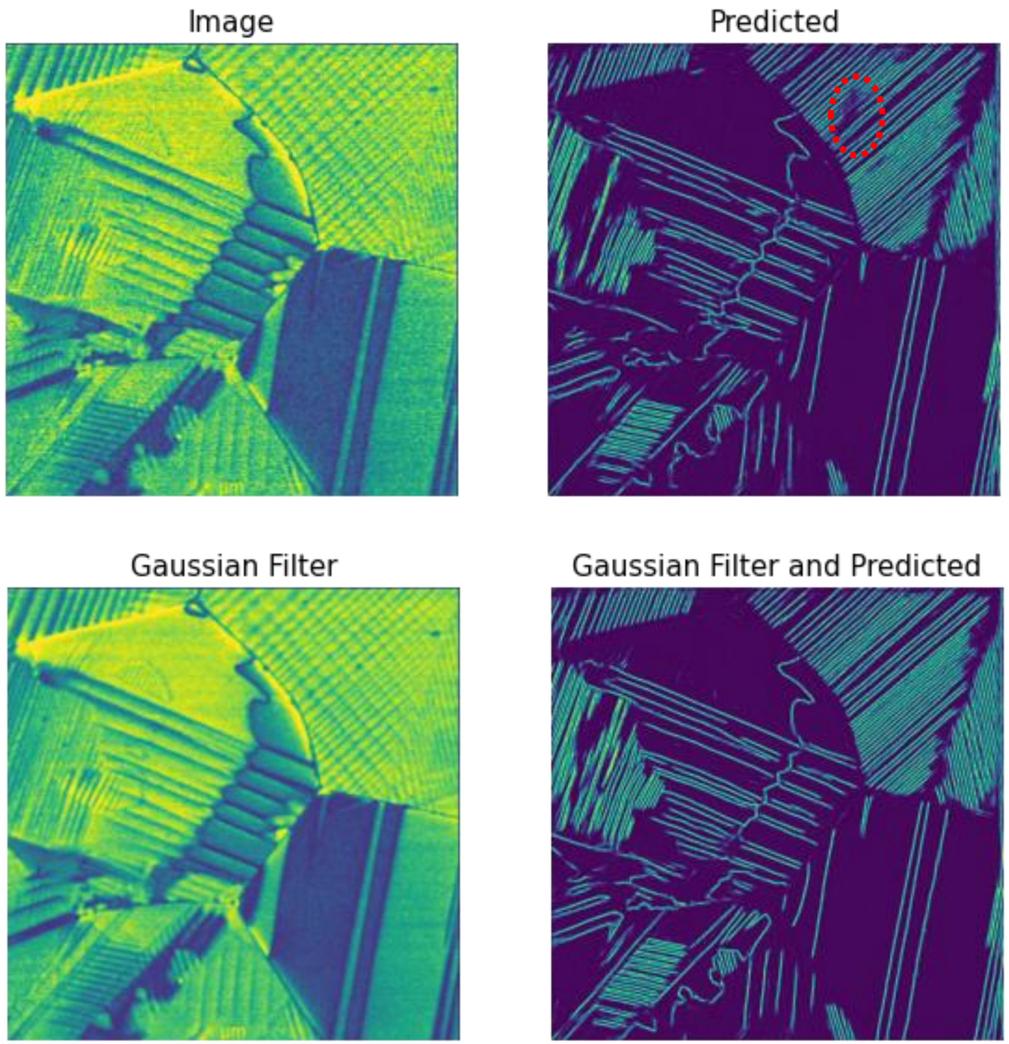


Detecting domain walls

Canny filter

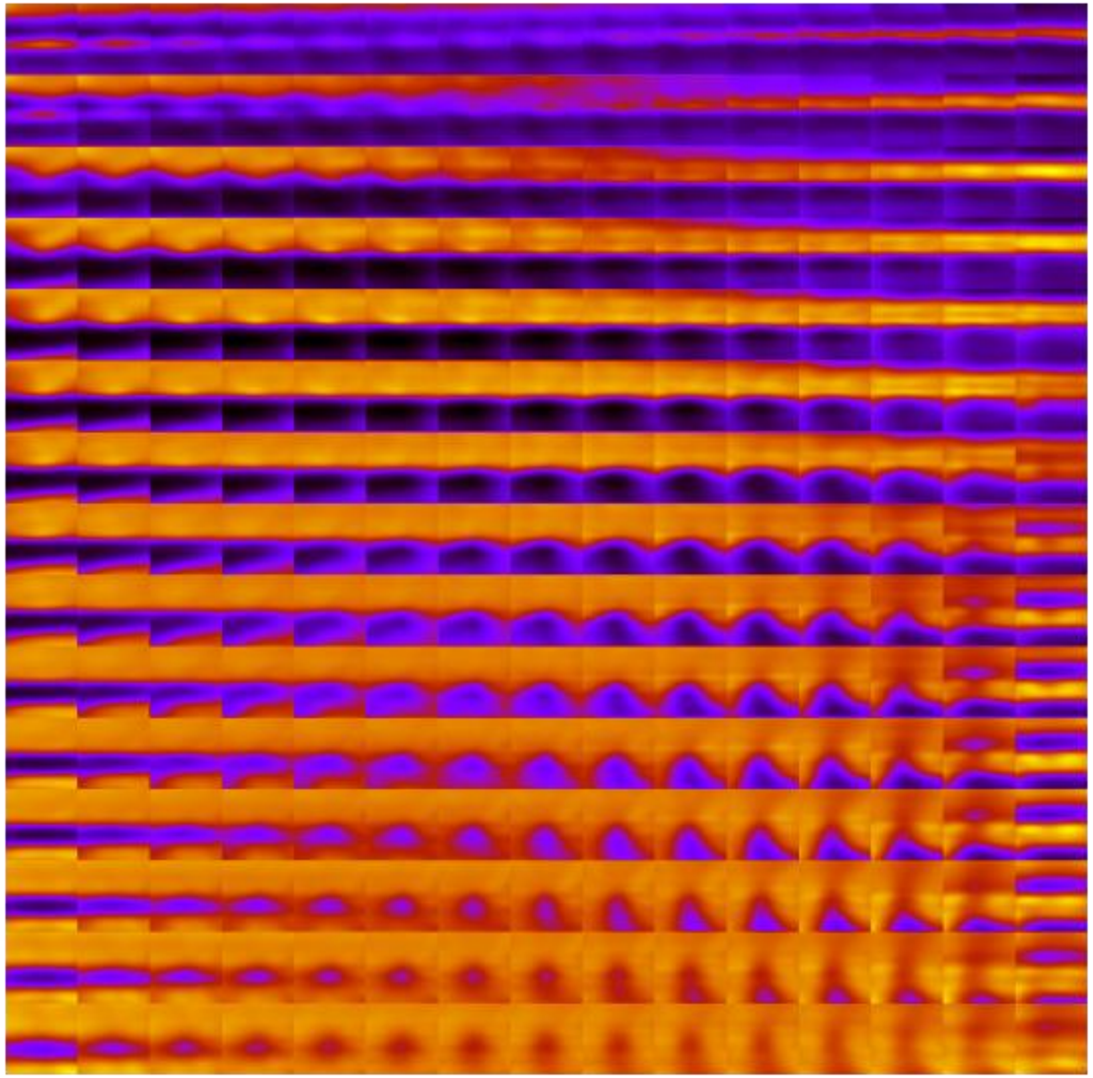


DCNN Prediction

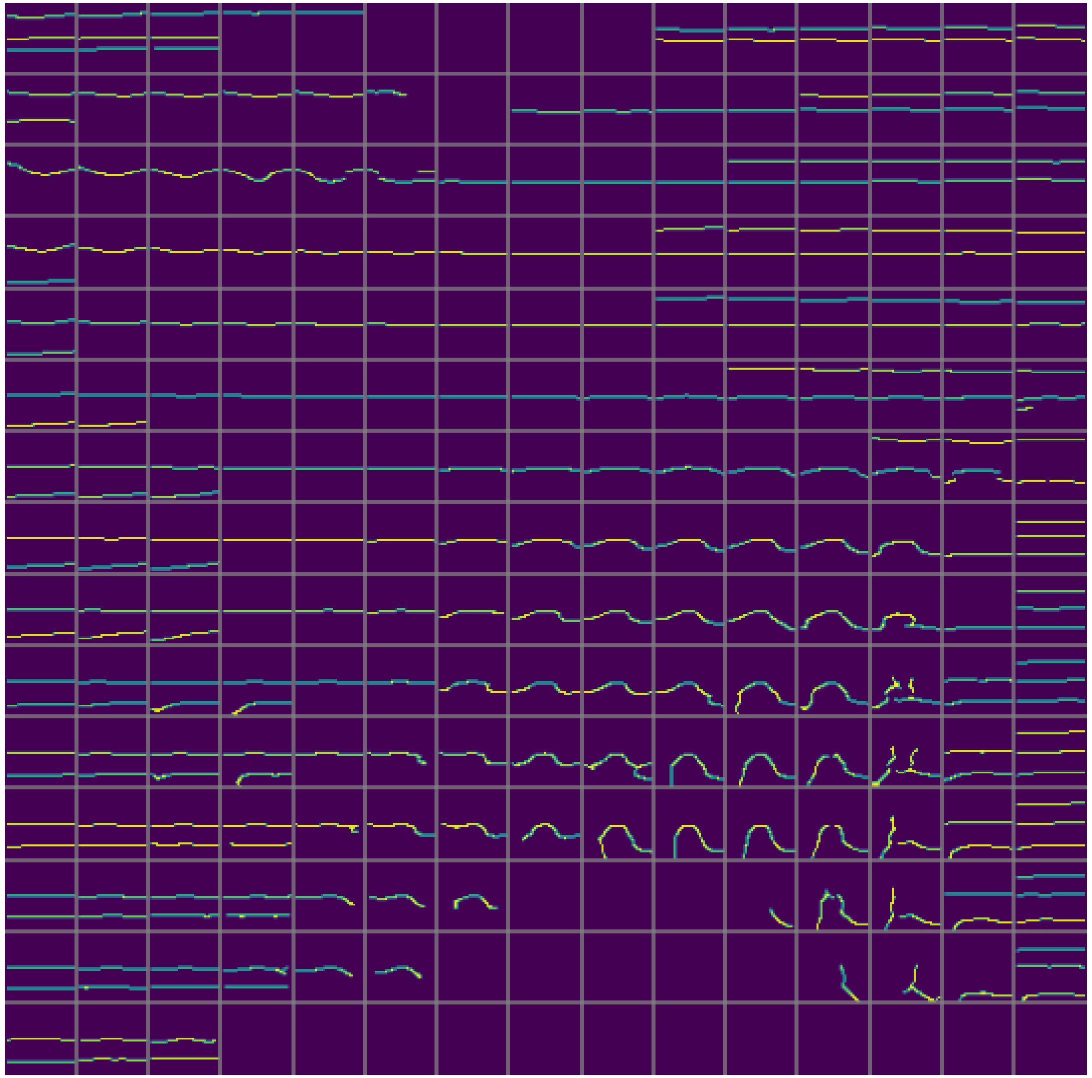


rVAE analysis

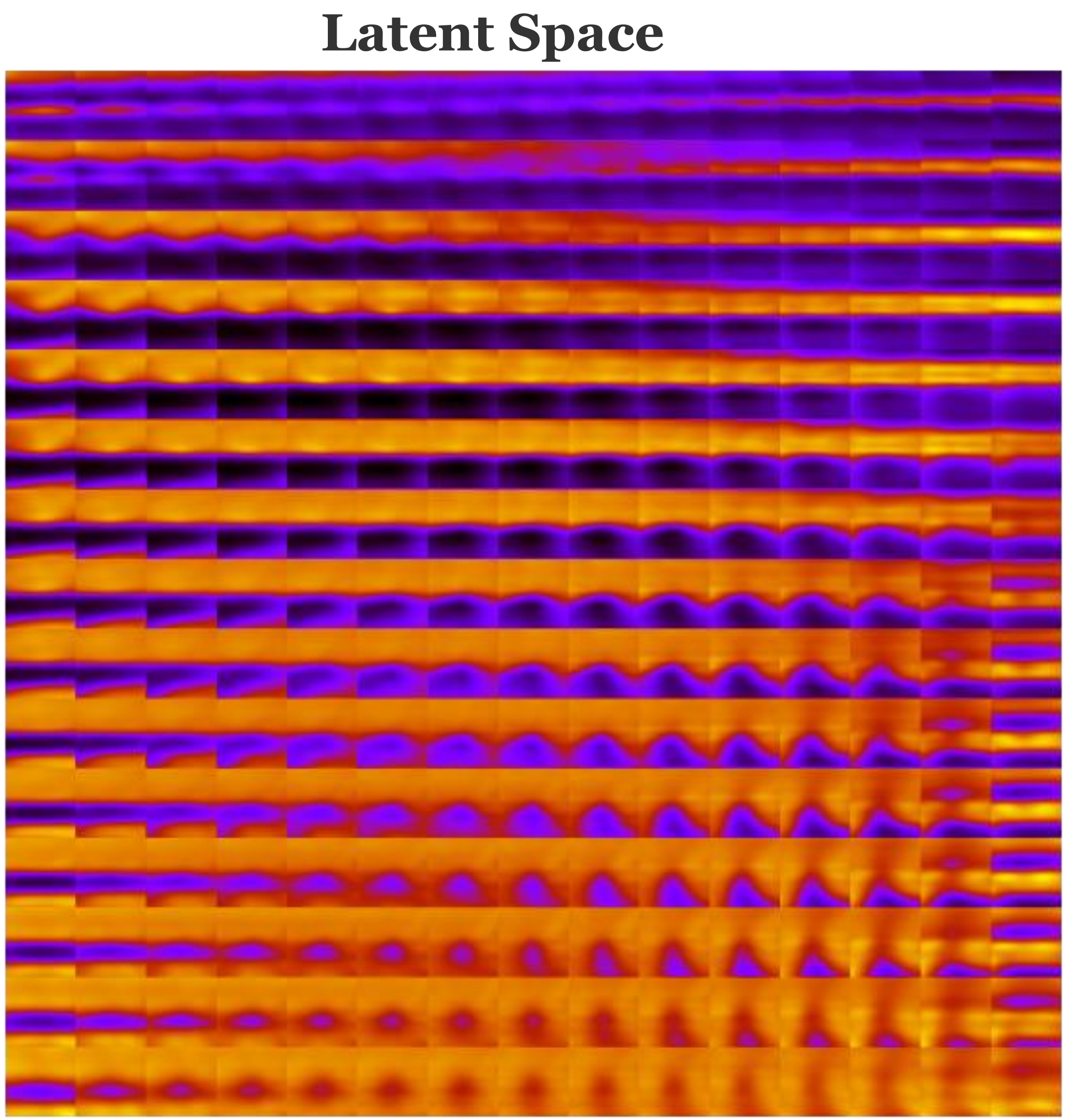
Latent Space

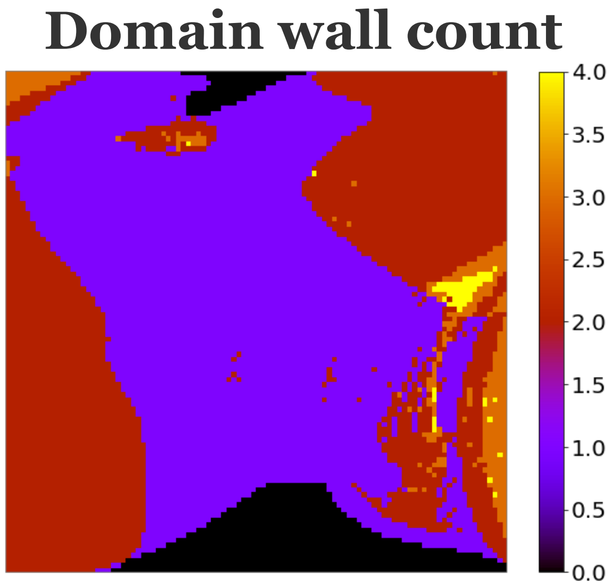


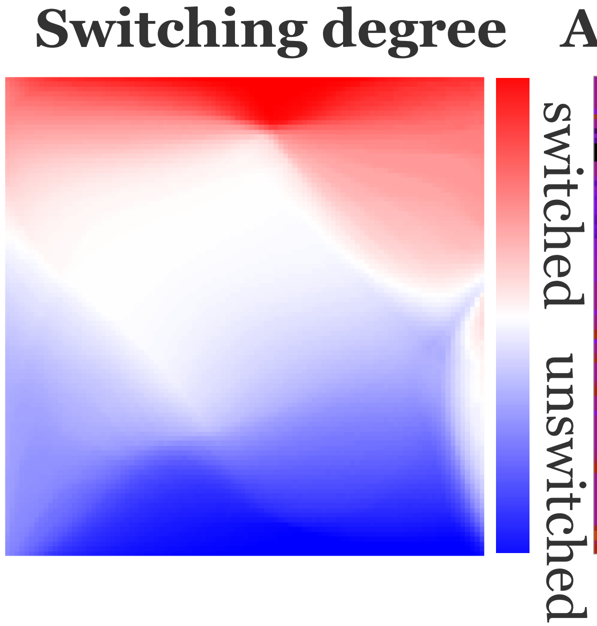
Domain Walls

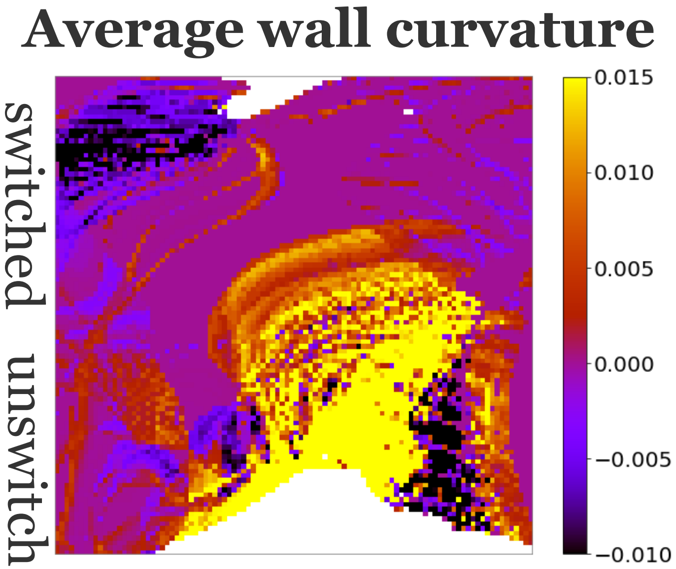


rVAE latent space



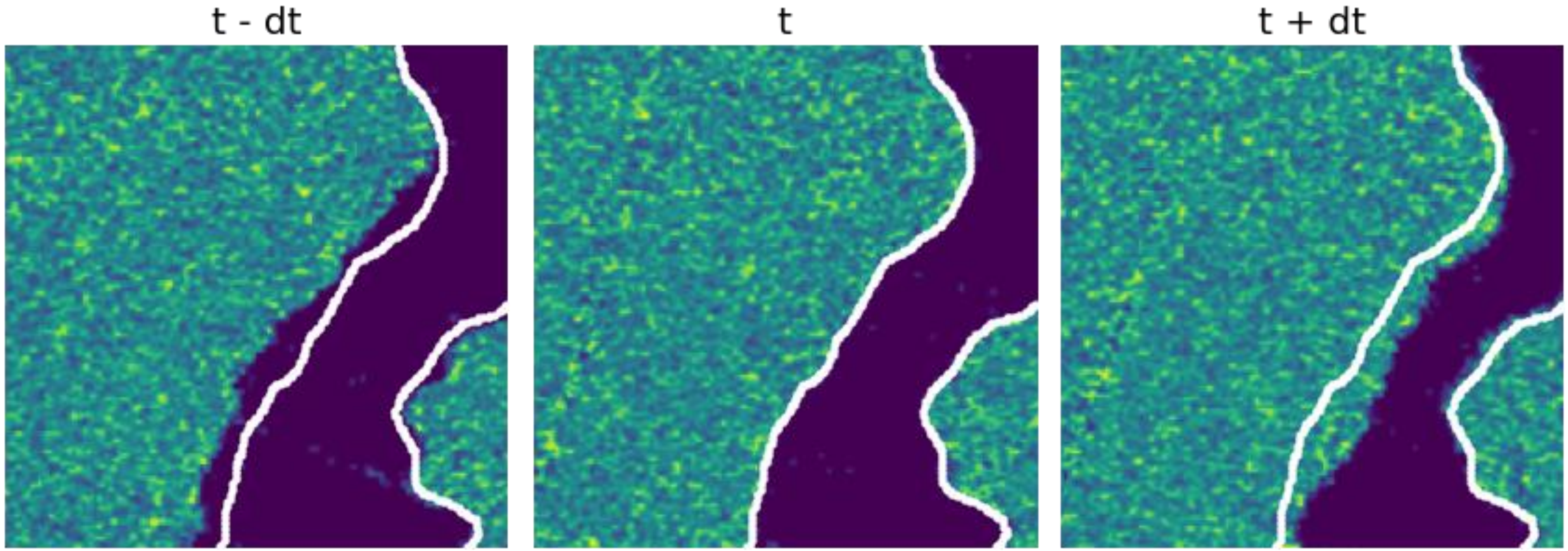






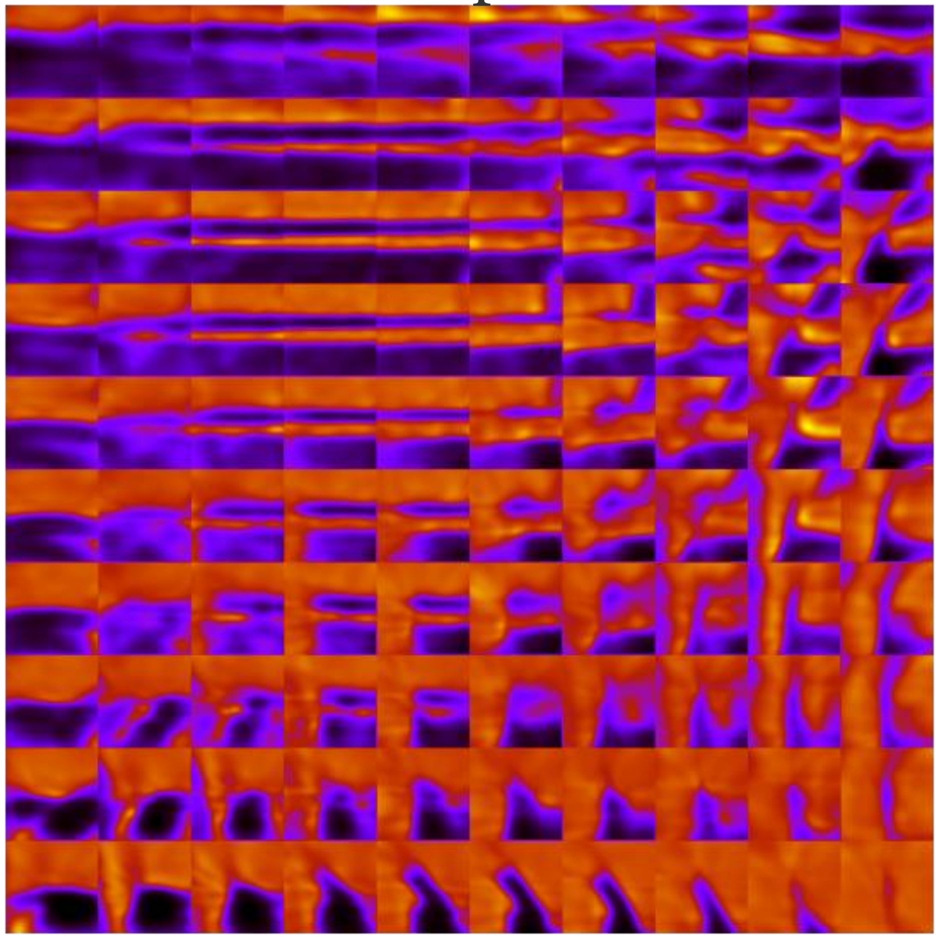
rVAE with time delay

Training dataset

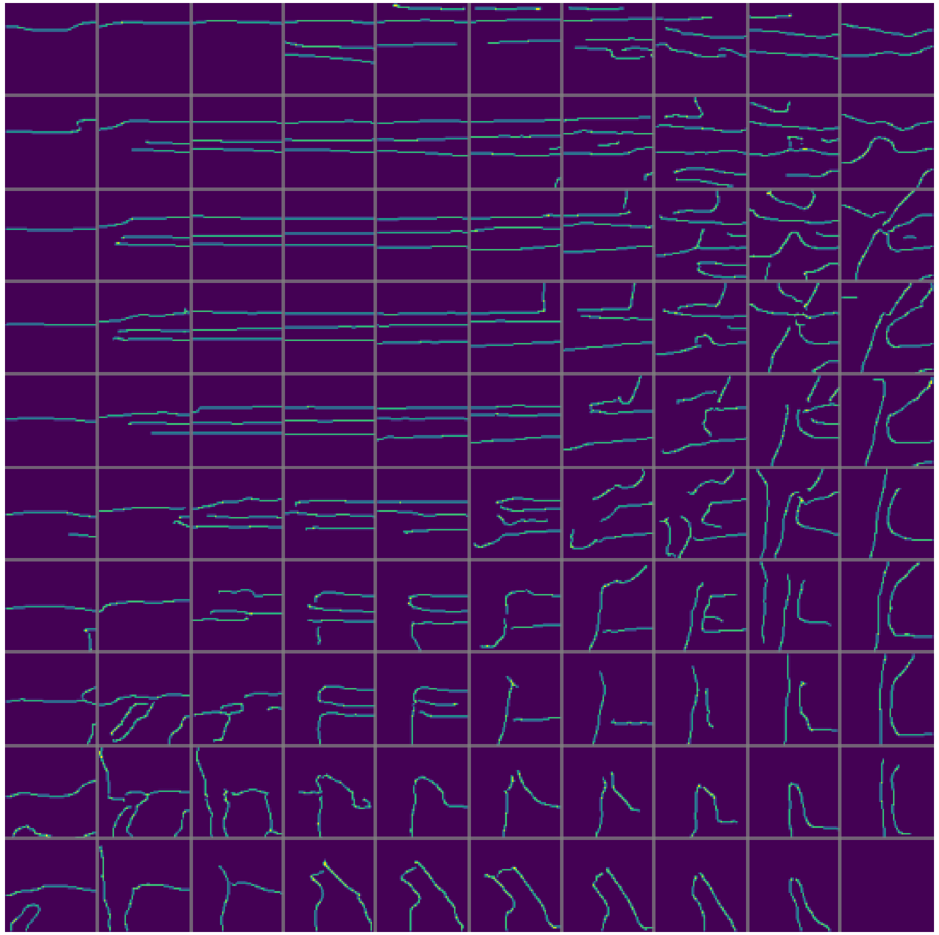


rVAE with time delay

Latent space

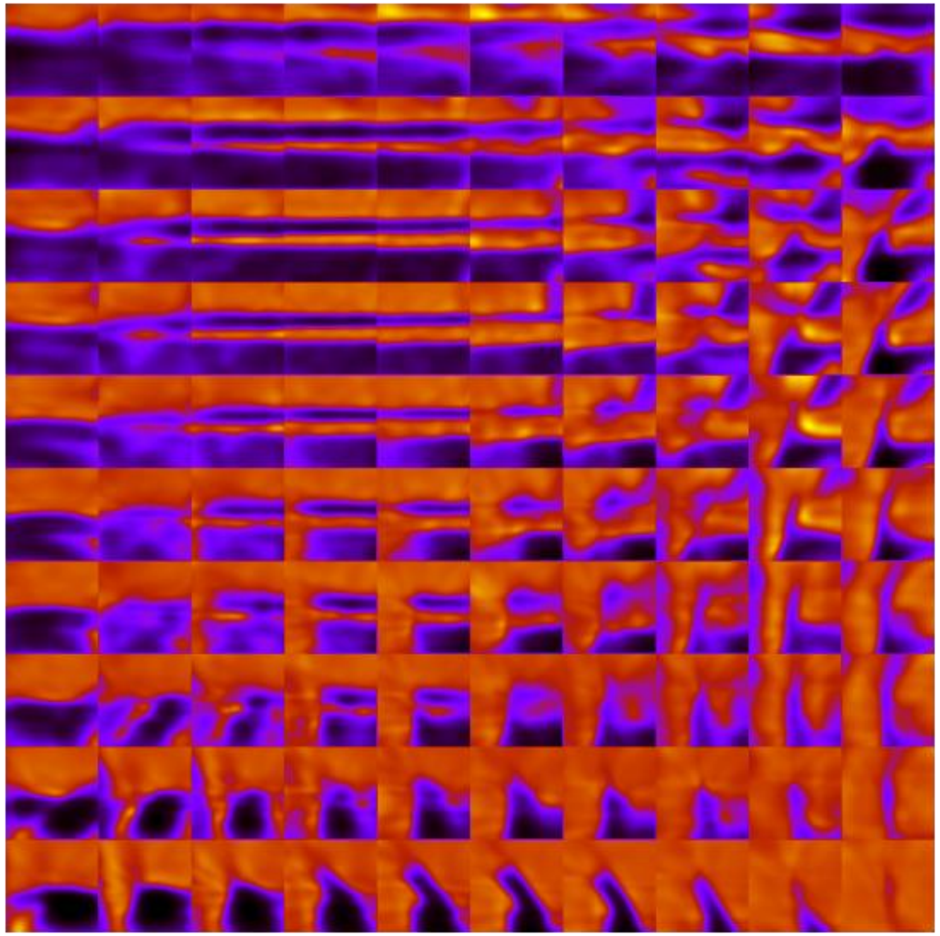


Domain wall

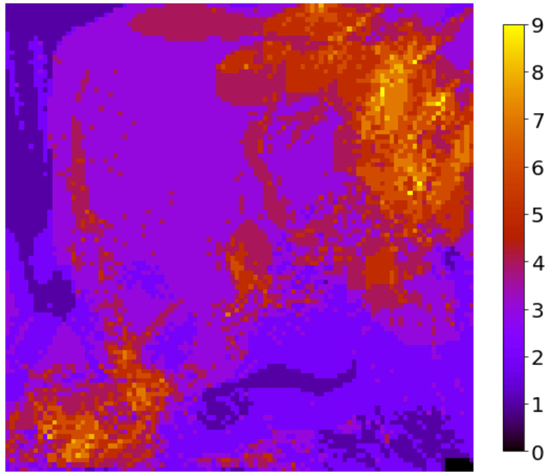


rVAE with time delay

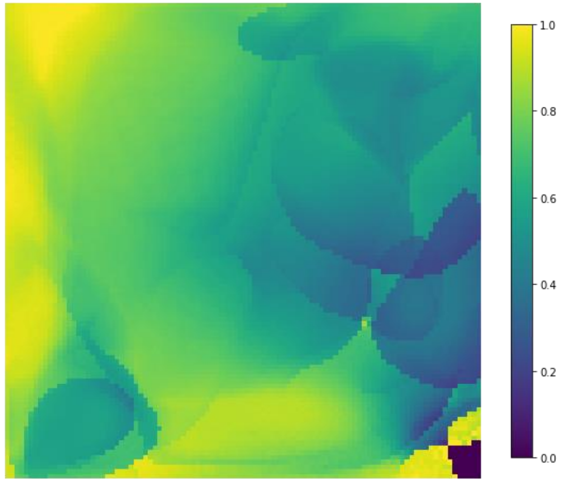
Latent space



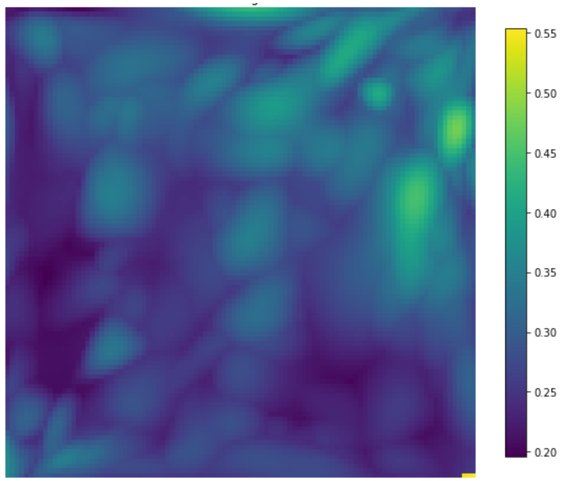
Wall count



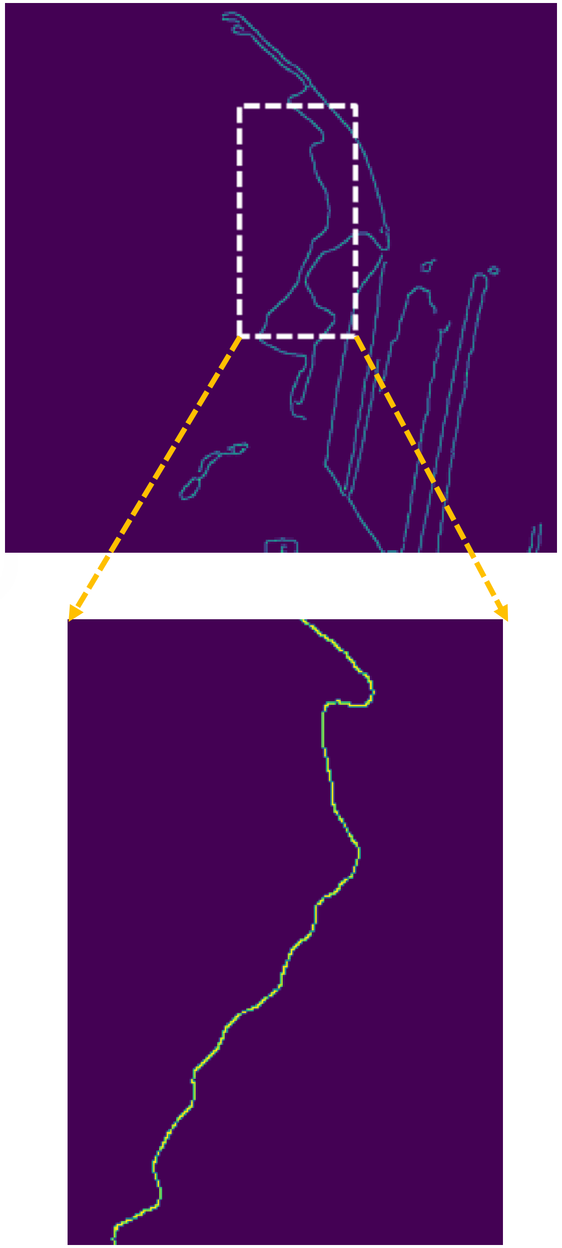
Domain convex



Switch significance

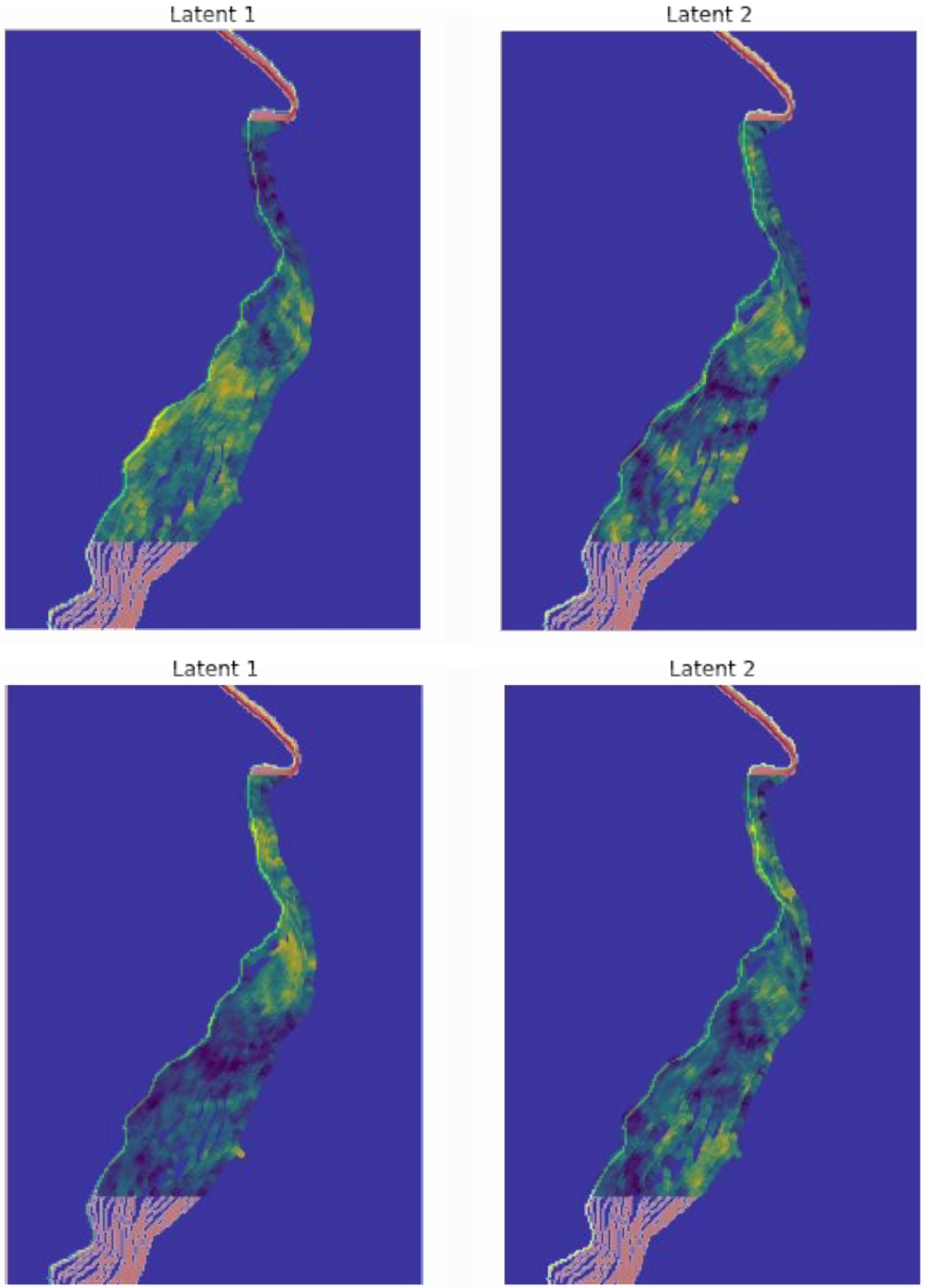


Domain wall evolution

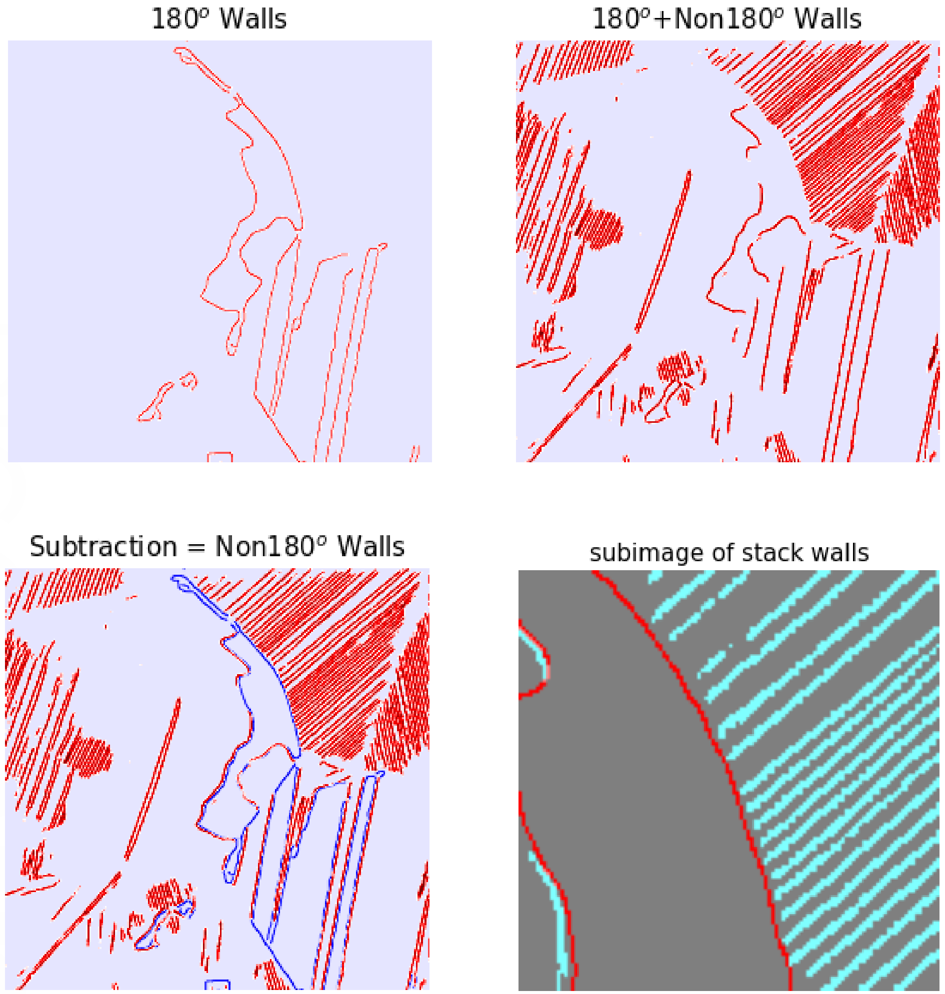


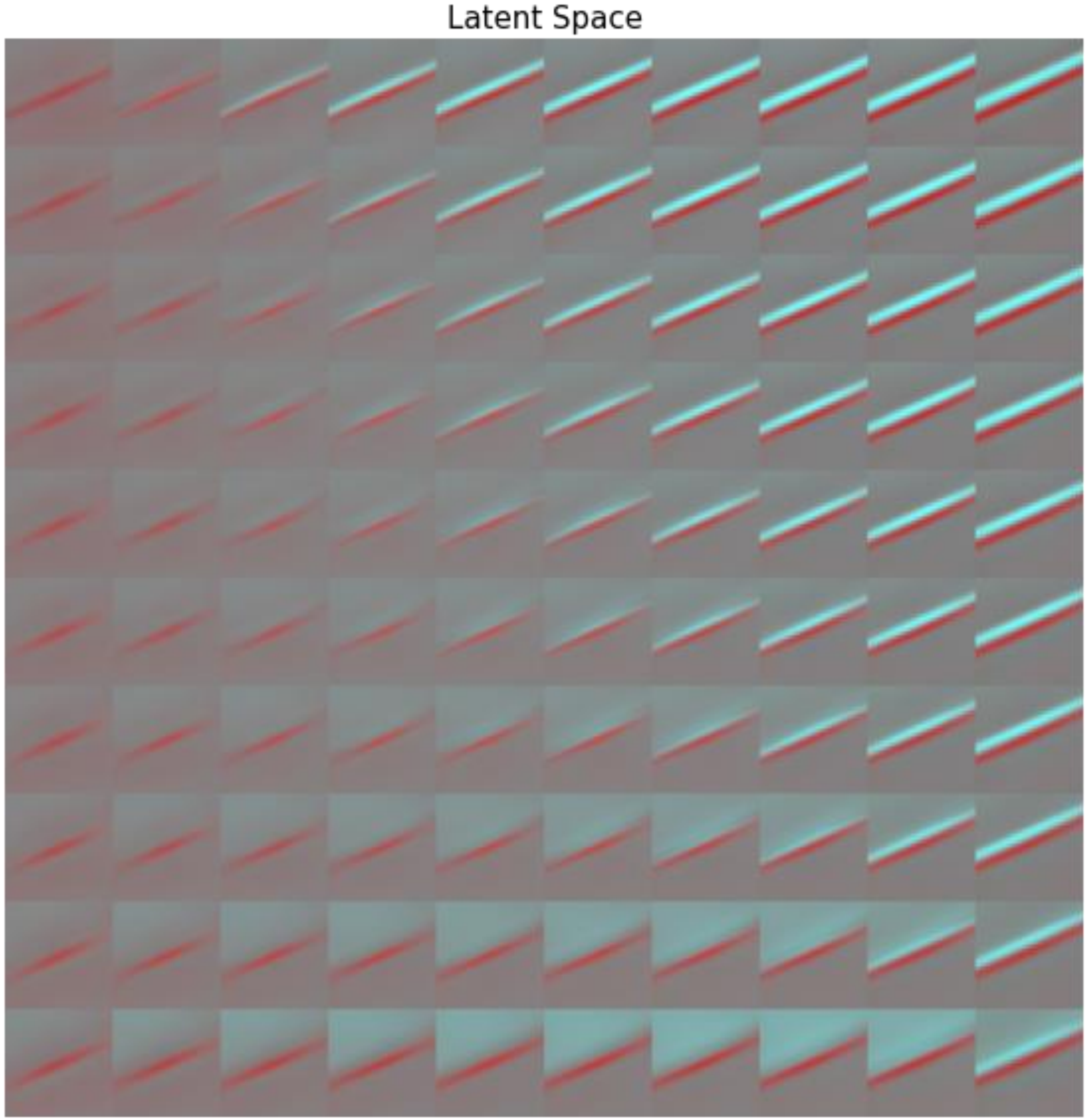
Forward:
t vs t+1

Reverse:
t vs t+1

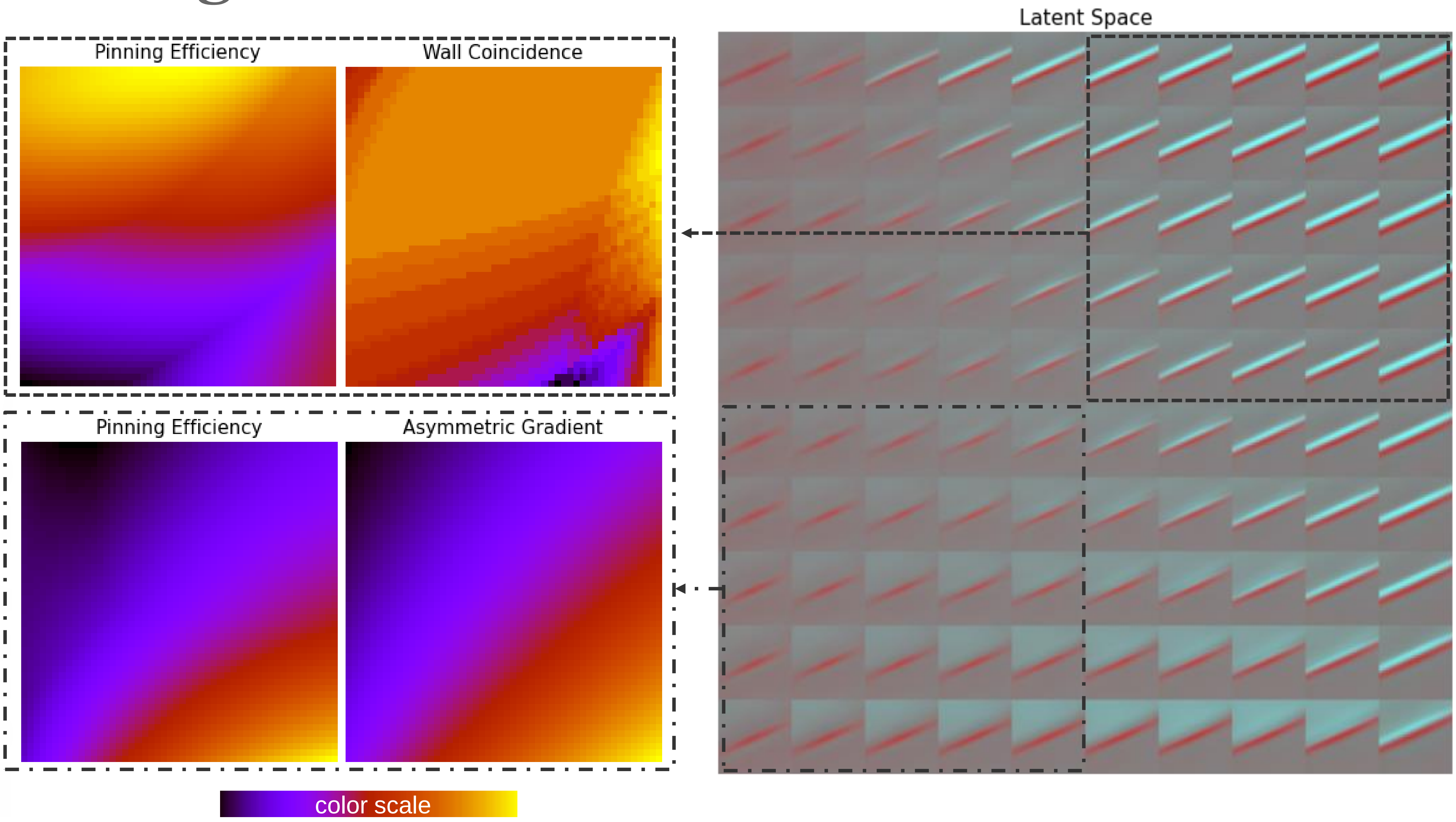


Multilayer rVAE





Pinning mechanism



What if we have multiple classes?

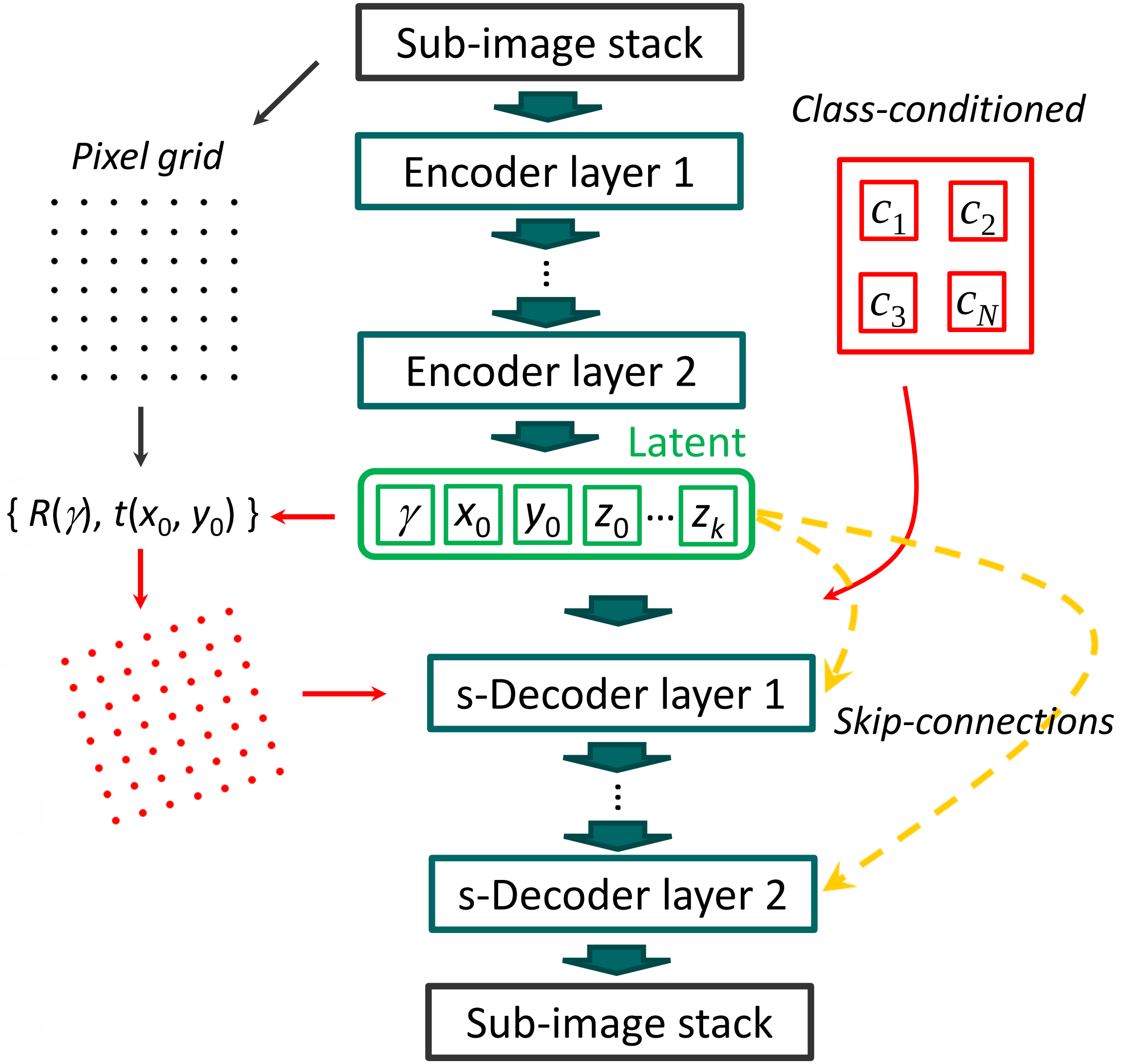
1. Classes are known: conditional (discrete) VAE

2. Factors of variability are known: conditional (continuous) VAE

3. Some classes are known: semi-supervised VAE

4. Number of classes are known: joint VAE

Conditional VAE



- Generative model is a function of spatial coordinate
- 3 additional latent variables to absorb rotations and shifts
- Disentangles rotations and translations from image content
- Ideal for analyzing microscopy sub-images on atomic level

$$\left.\begin{array}{l} ELBO \\ = Reconstruction\ Loss \\ -\ D_{KL}\big(q(z|x)\|\mathcal{N}(0,I)\big) \end{array}\right\} \text{Regular VAE}$$

$$-\ D_{KL}\left(q(\gamma|x)\|\mathcal{N}\left(0,s_\gamma^2\right)\right) \quad \text{Rotation}$$

$$-\ D_{KL}\left(q(\Delta r|x)\|\mathcal{N}\left(0,s_{\Delta r}^2\right)\right) \quad \text{Translation}$$

$+\ D_{KL}$(physics-based “priors”) ?

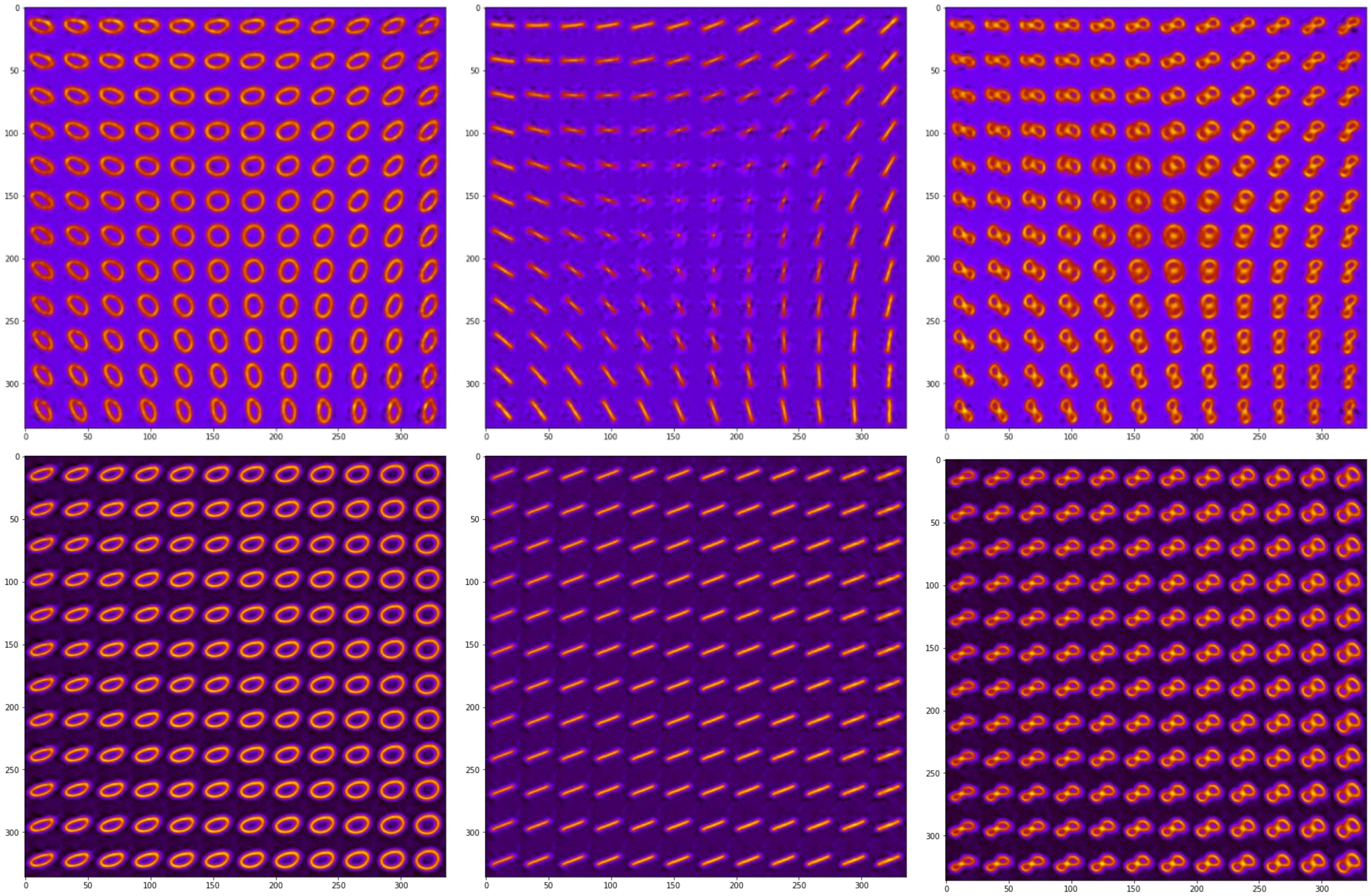
$+\ D$ (physics) ?

Bepler *et al.*, NeurIPS 2019

MNIST cVAE

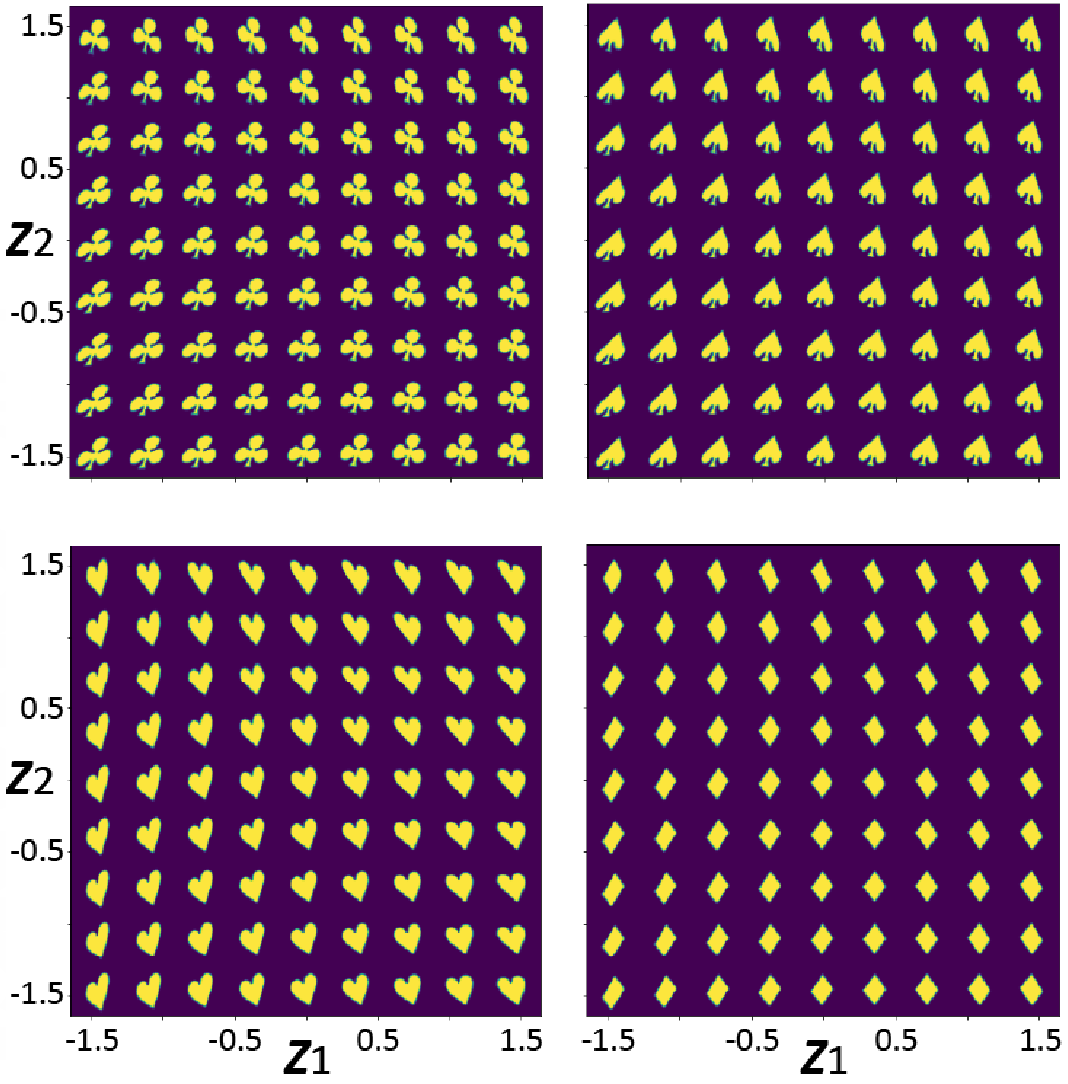
No rotations

With rotations

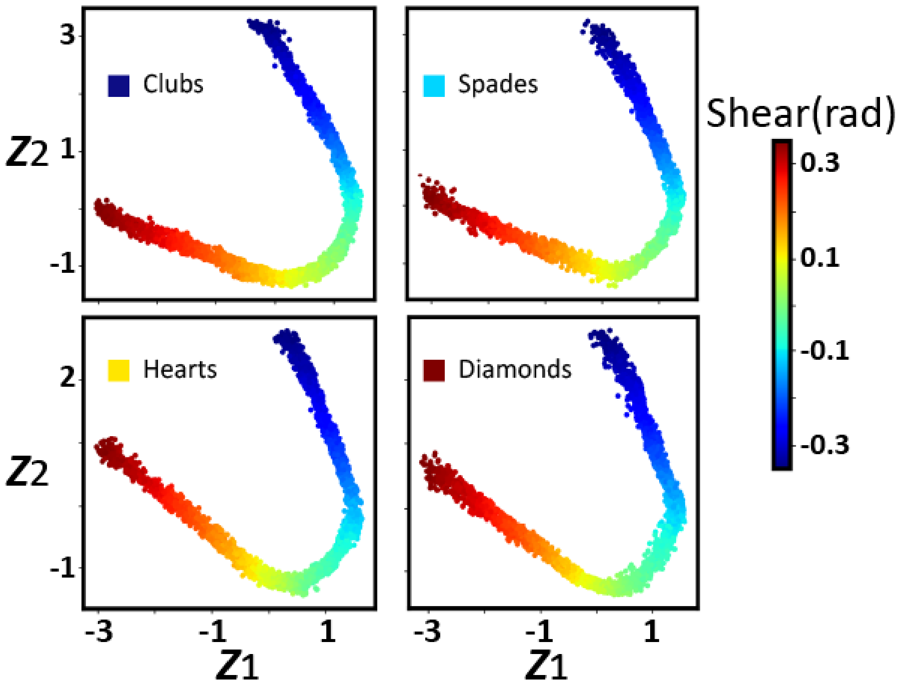


cVAE on cards data set

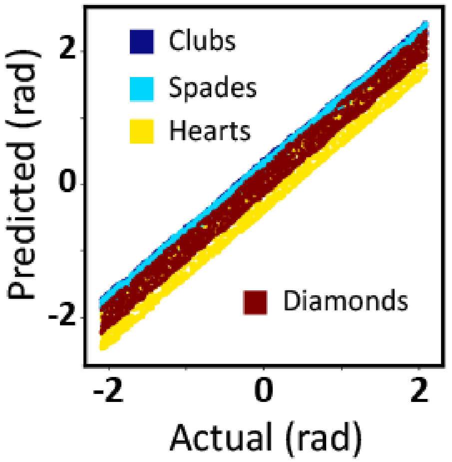
One independent latent space per class

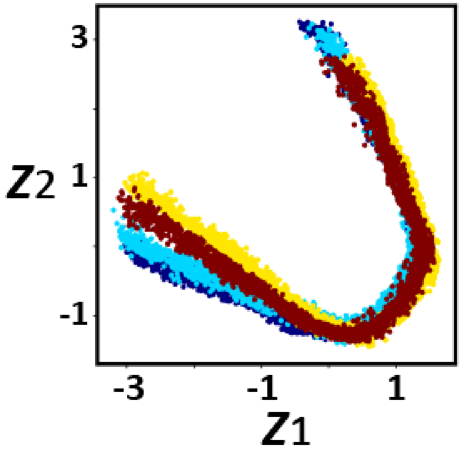


Shear distribution in latent space

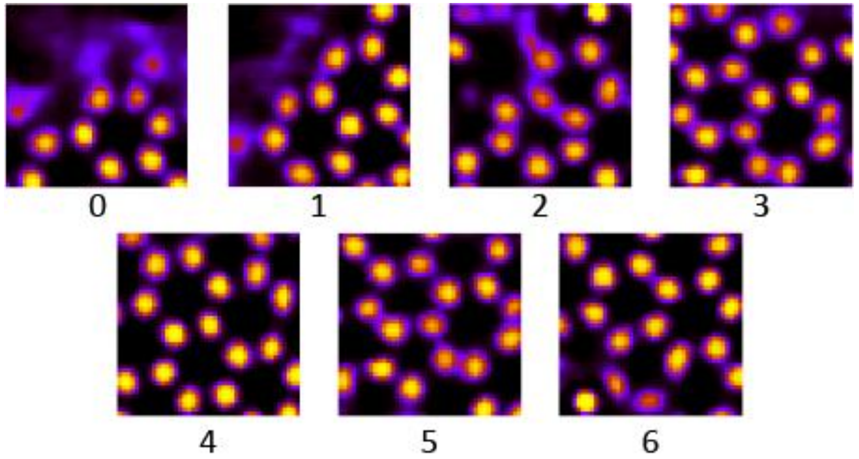


All four plotted jointly





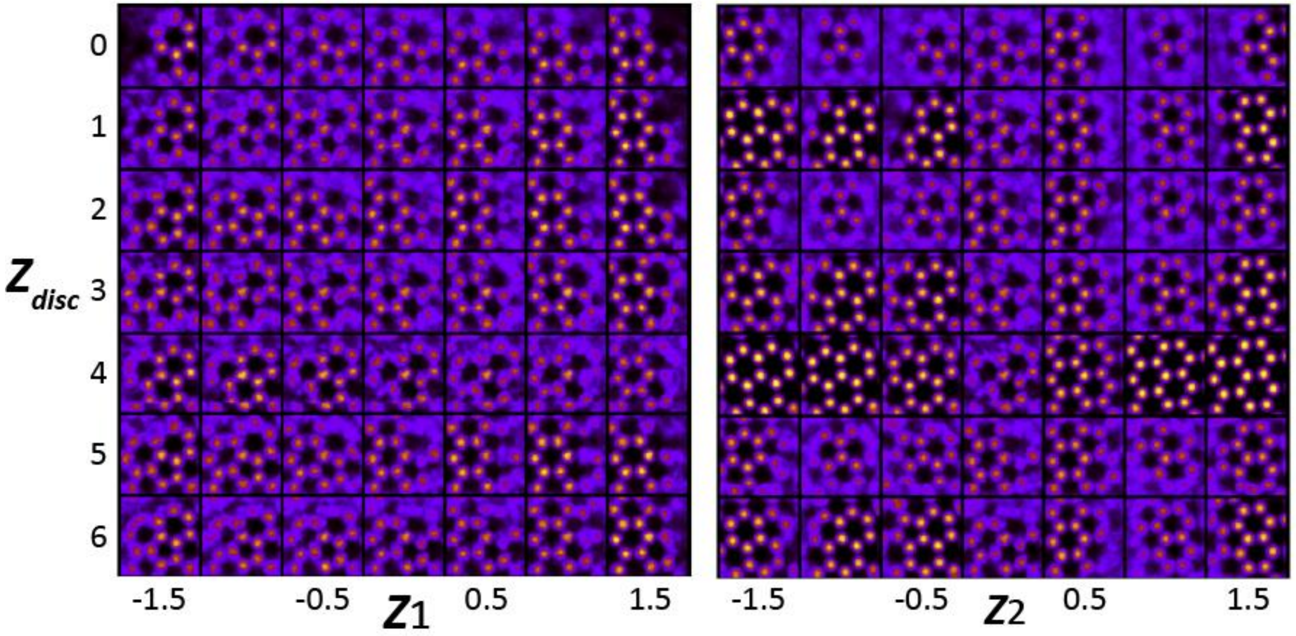
cVAE on Graphene



Autonomous sub-image labeling

0: Fewer than 2 rings
1: two rings of size [6, 6]
2: three rings of size [5, 6]
3: three rings of size [6, 6, 7]
4: three rings of size [6, 6, 6]
5: three rings of size [5, 6, 7]
6: three rings of size [5, 6, 6]

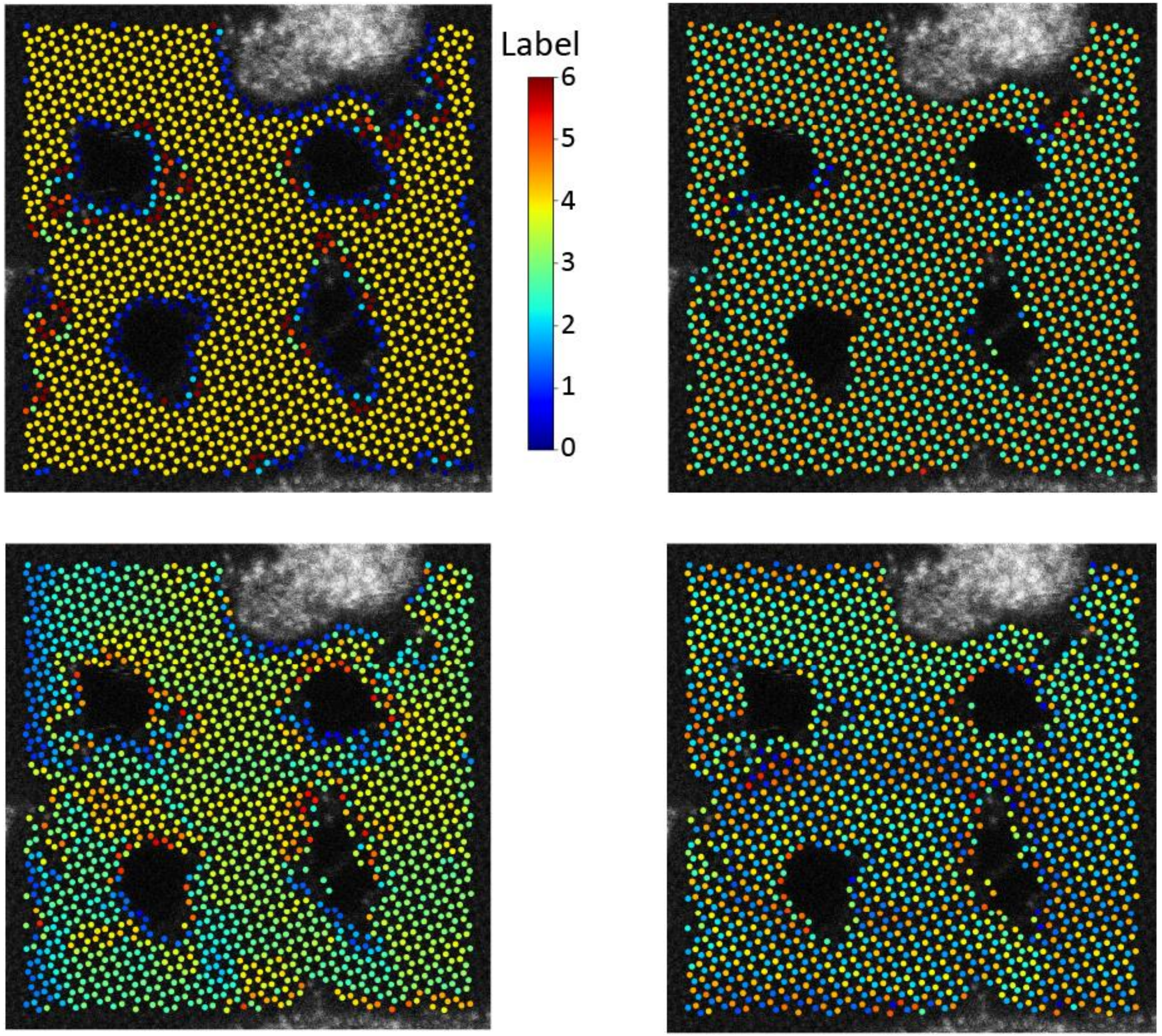
Traversal manifolds



cVAE on graphene

Class labels

Encoded Angle



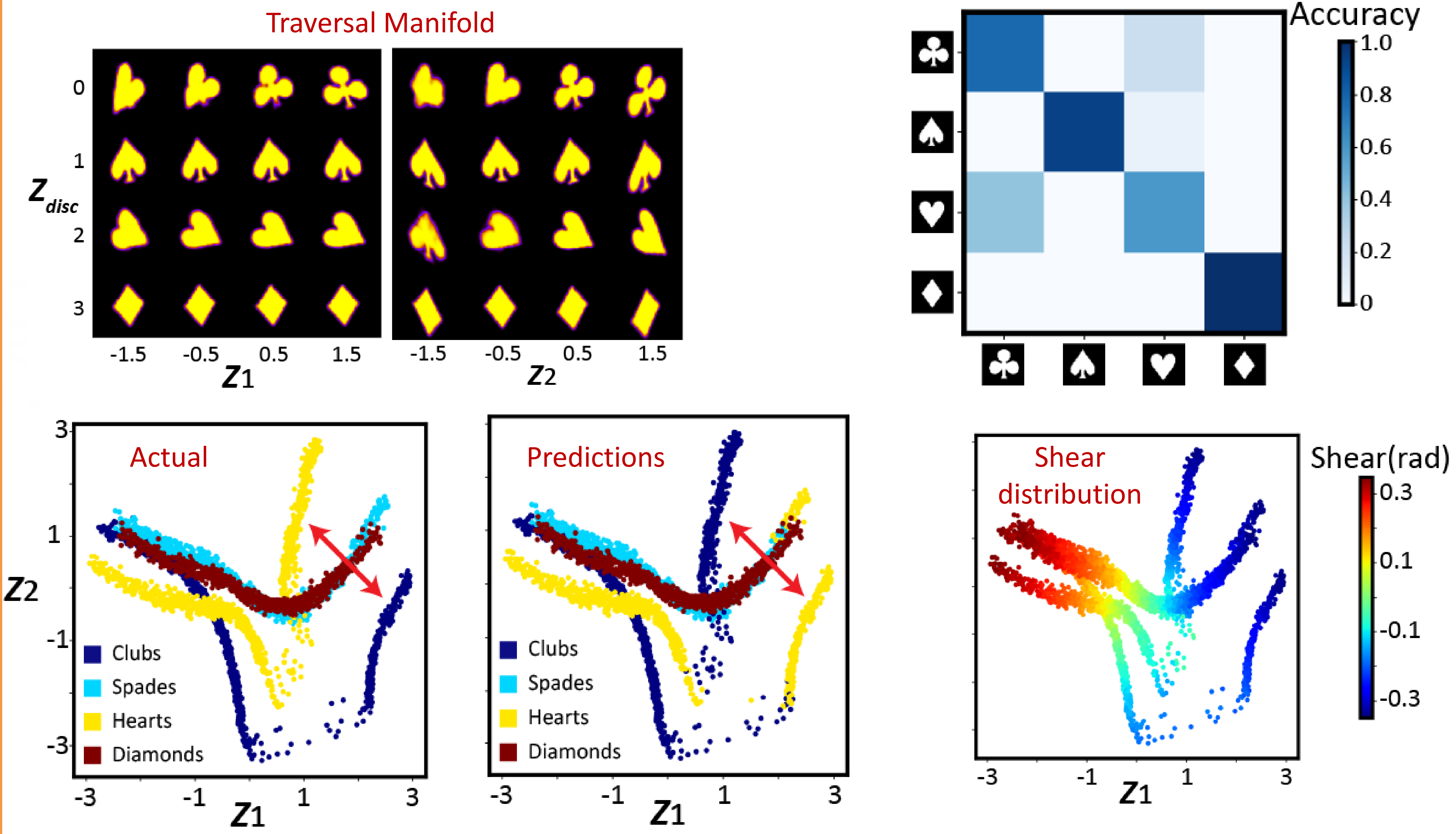
Z_1

Z_2

Correlation

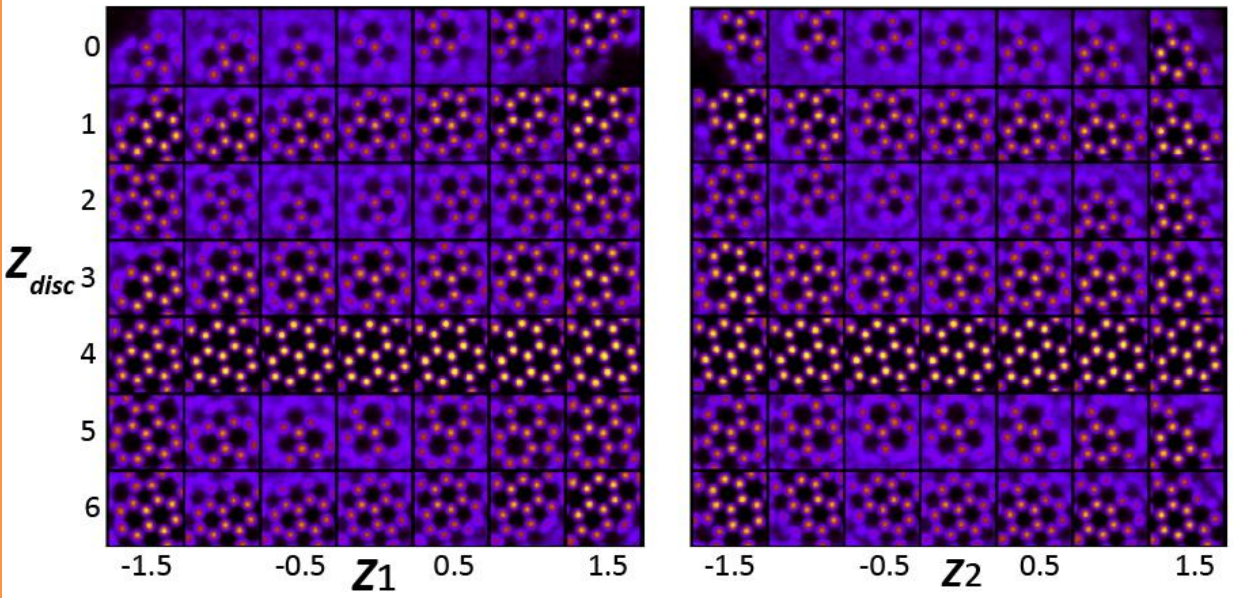
Hypothesis: Abundance of latent variables

Semi-supervised VAE

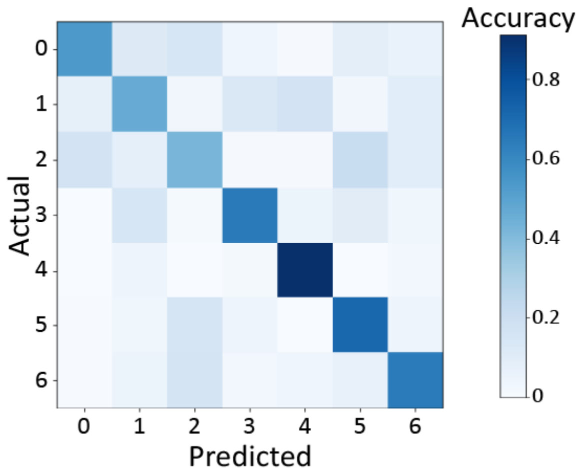


Semi-supervised VAE (graphene)

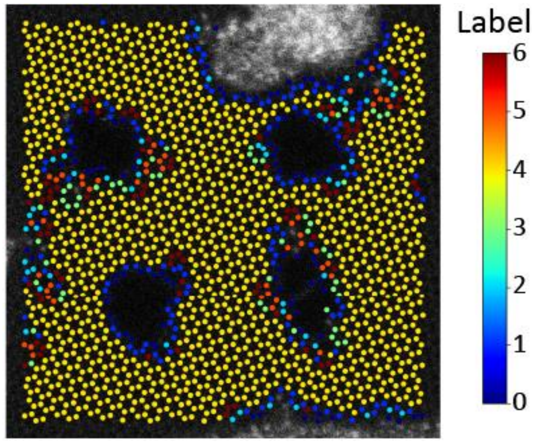
Traverse manifolds



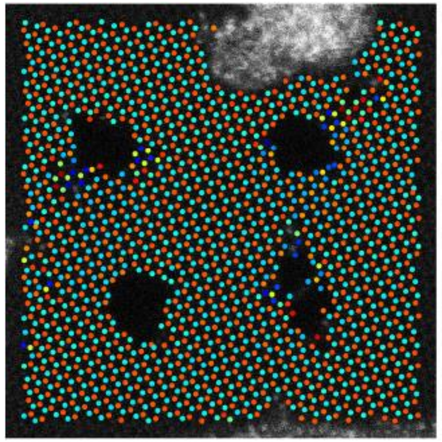
Confusion matrix



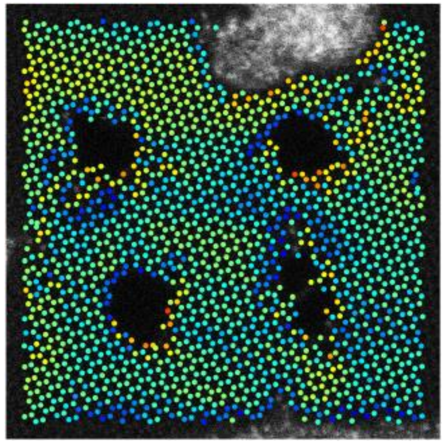
Class labels



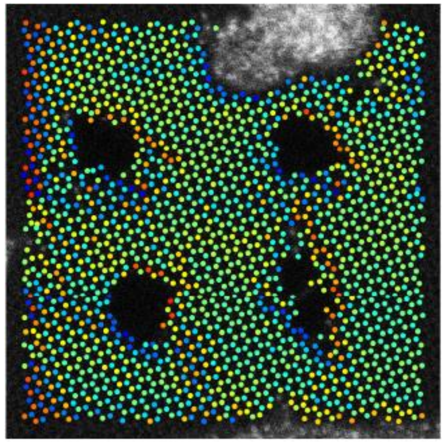
Angle Encoding



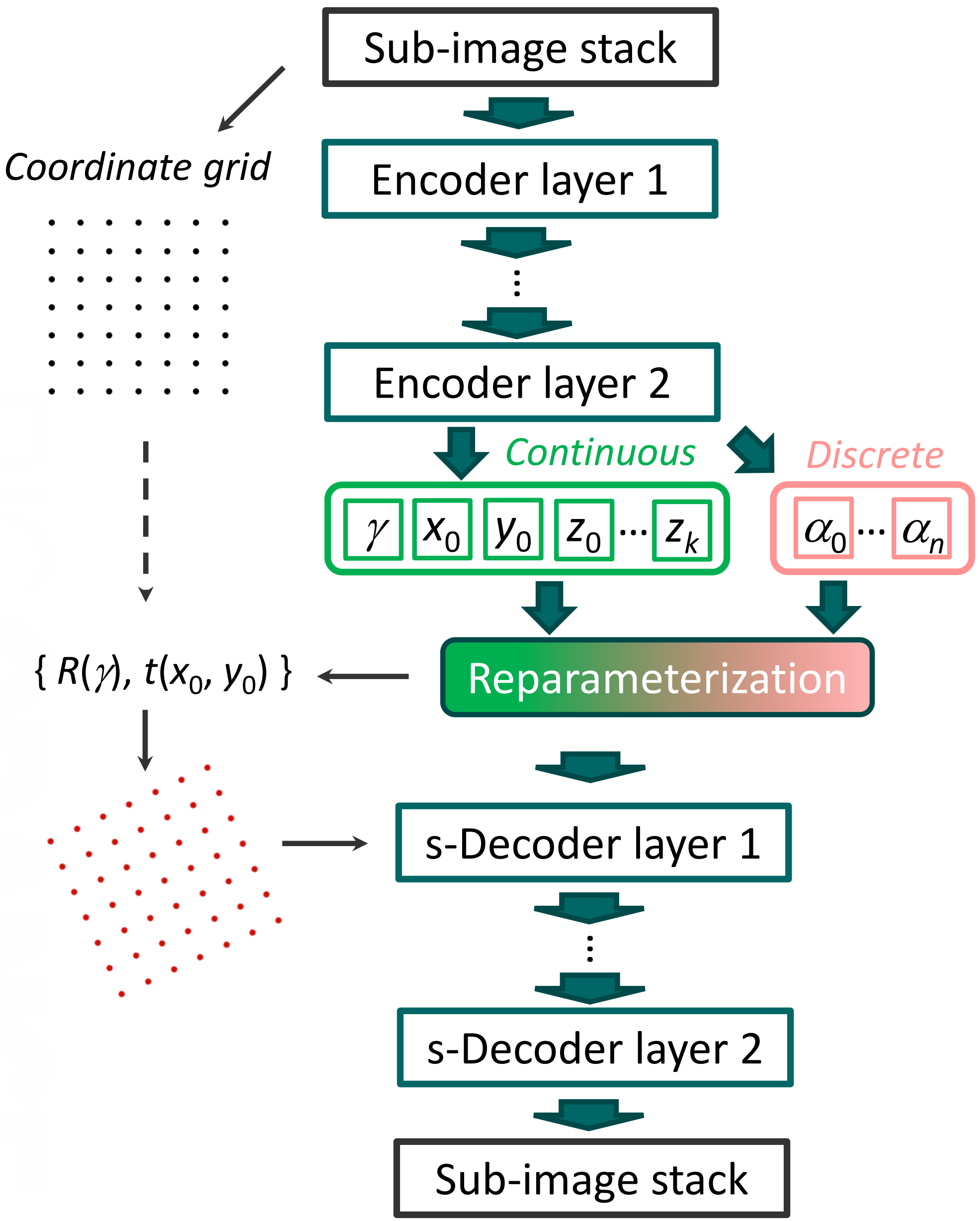
Z_1



Z_2



Joint VAe



- Generative model is a function of spatial coordinate (*e.g.*, via spatial broadcasting)
- 3 additional latent variables to absorb rotations and shifts
- Disentangles rotations and translations from image content
- Learns discrete classes in unsupervised fashion
- Well-suited for analyzing microscopy (sub-)images on atomic and molecular levels

$$ELBO =$$

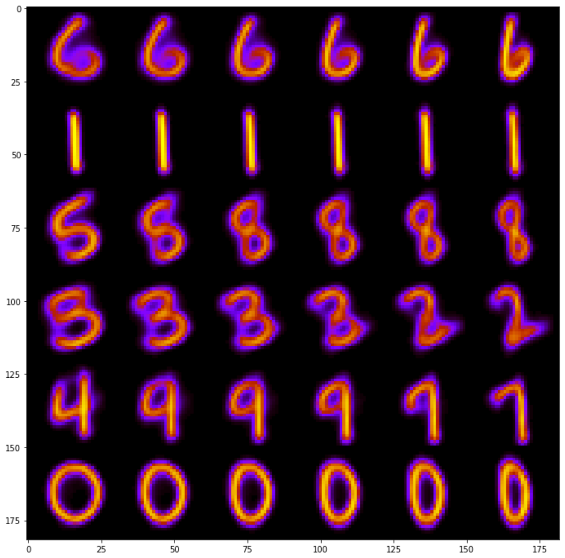
$$-\ Reconstruction\ Loss$$

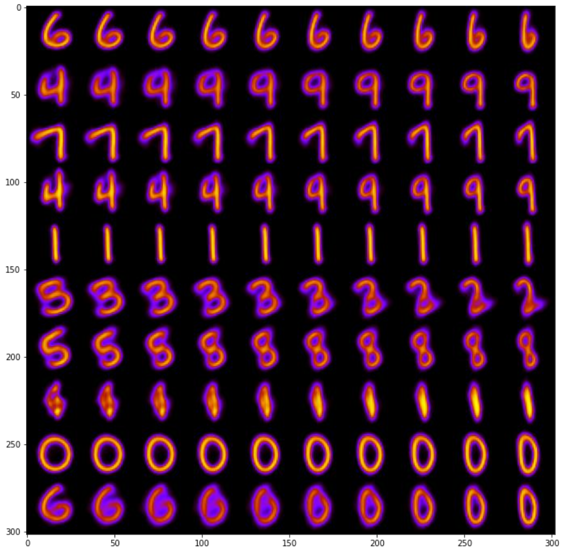
$$-\beta_c(t)\left|\left(\overset{\text{Structure}}{D_{KL}(q(z|x)\|p(z))} + \overset{\text{Rotation}}{D_{KL}(q(\gamma|x)\|p(\gamma))}\right) - C_z\right| \quad \text{Continuous}$$

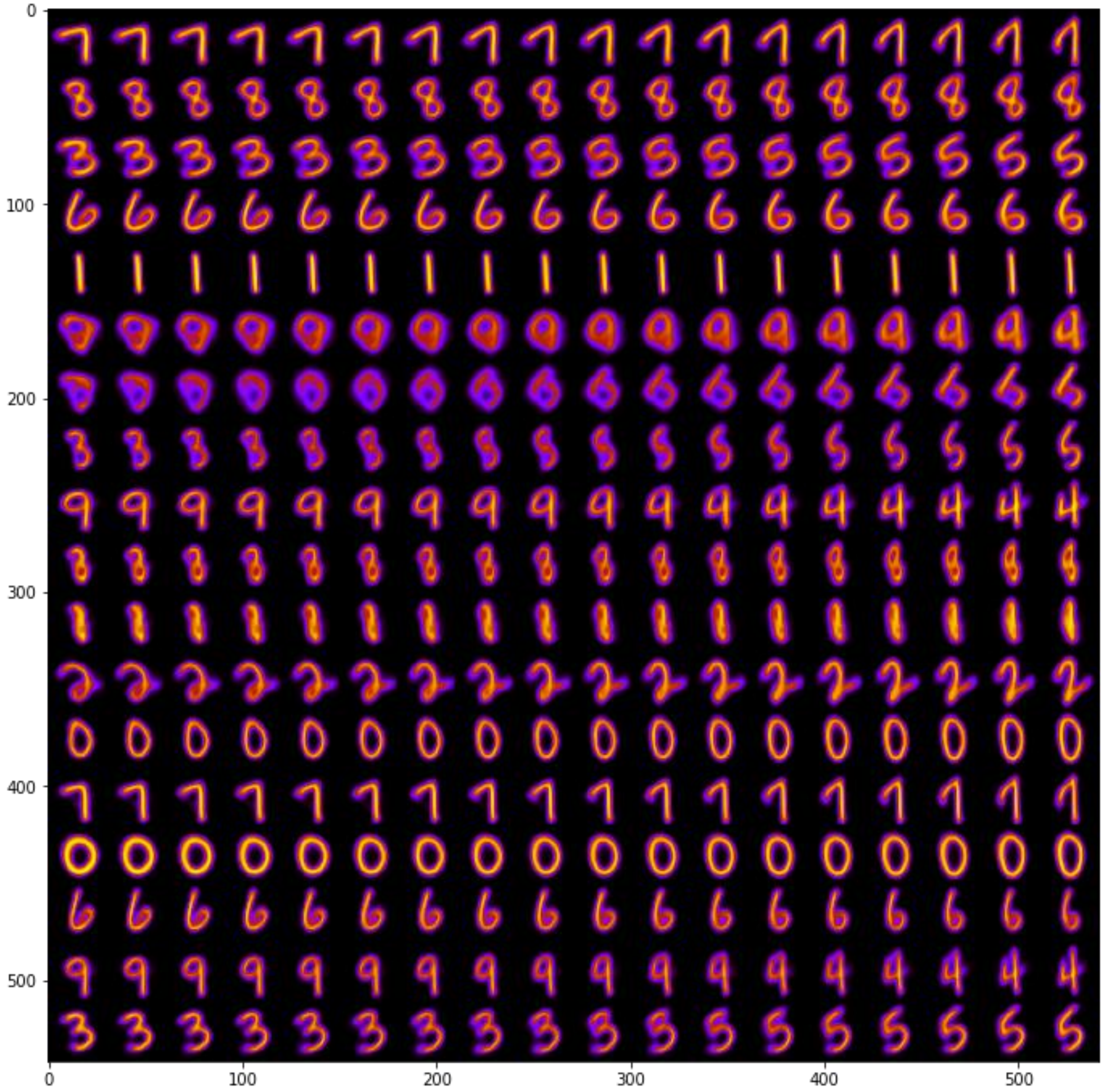
$$-\beta_d(t)\left|D_{KL}\left(q(\alpha|x)\|p(\alpha)\right) - C_\alpha\right| \quad \text{Discrete}$$

+ physics-based “loss” ?

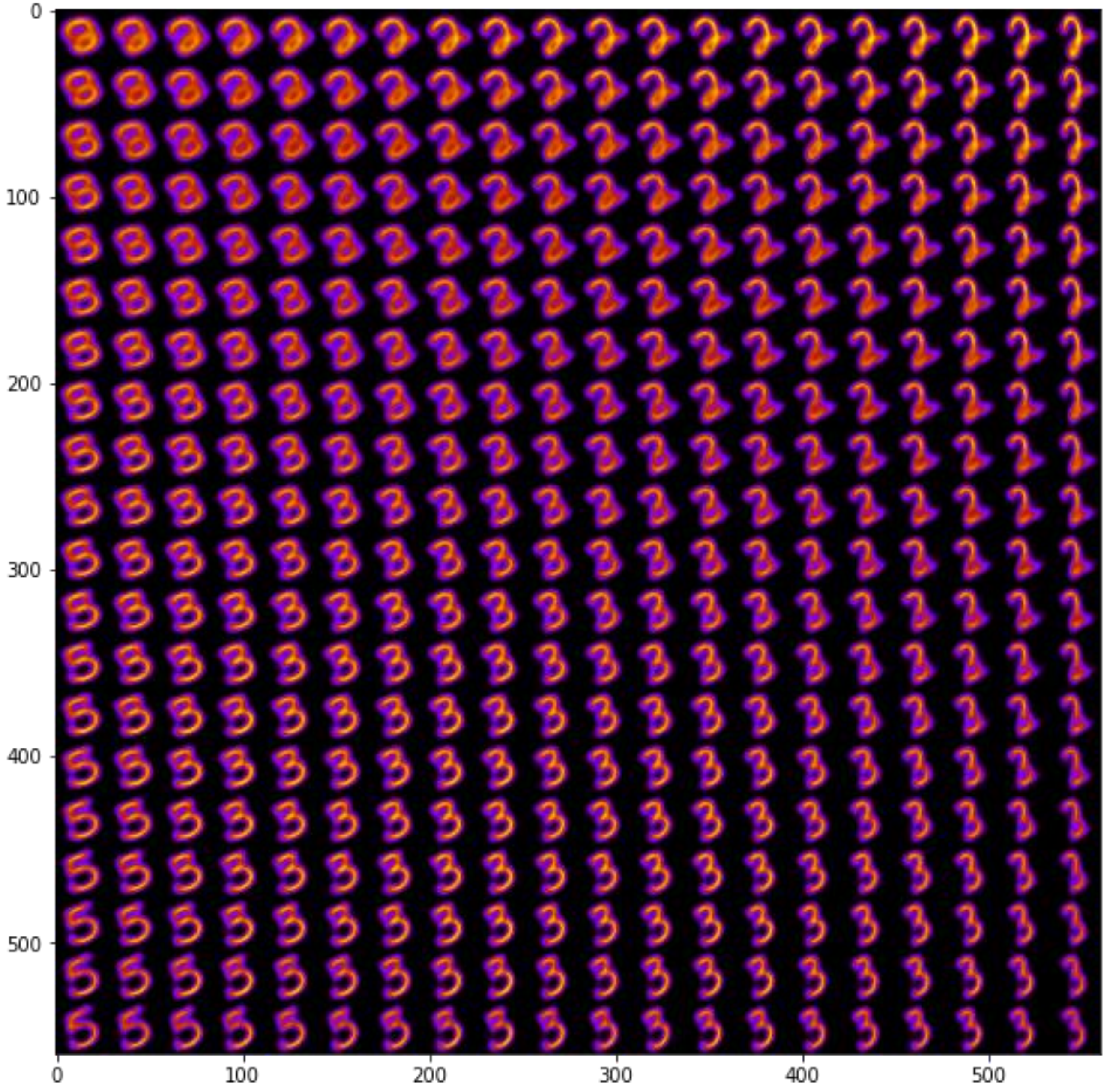
jVAE on MNIST





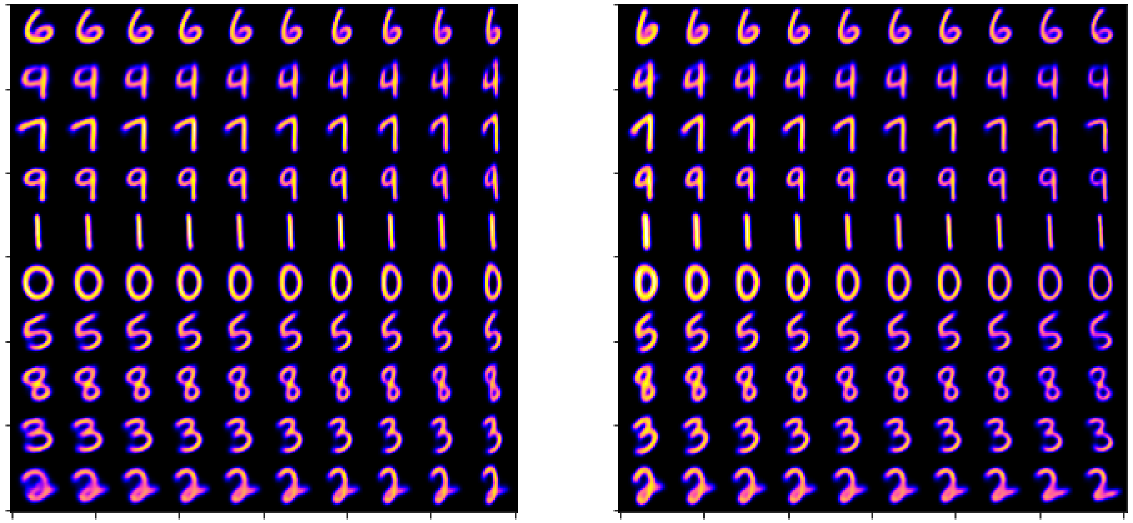


Latent representations

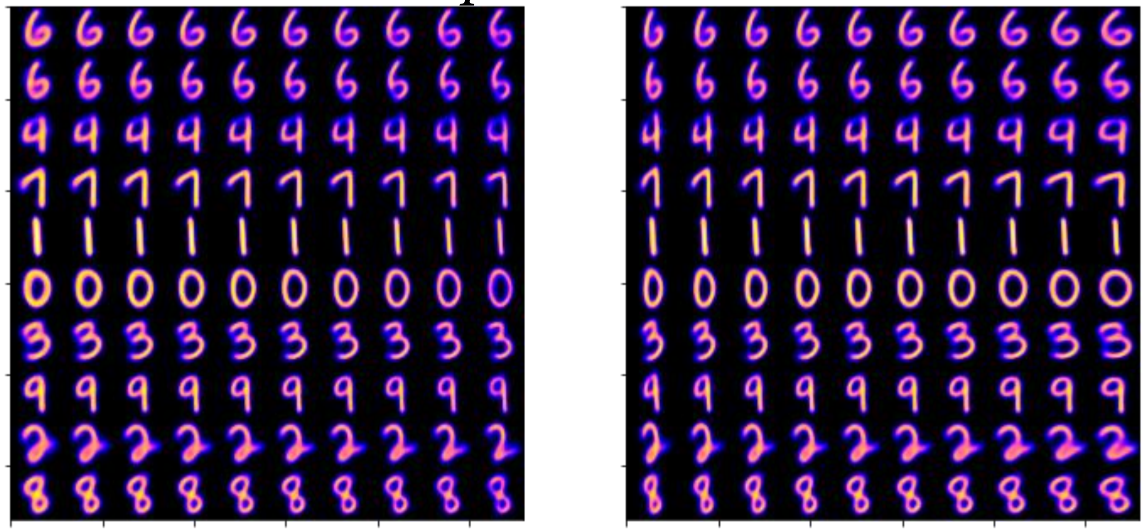


Ensemble jVAE

Predictions from different ensemble models



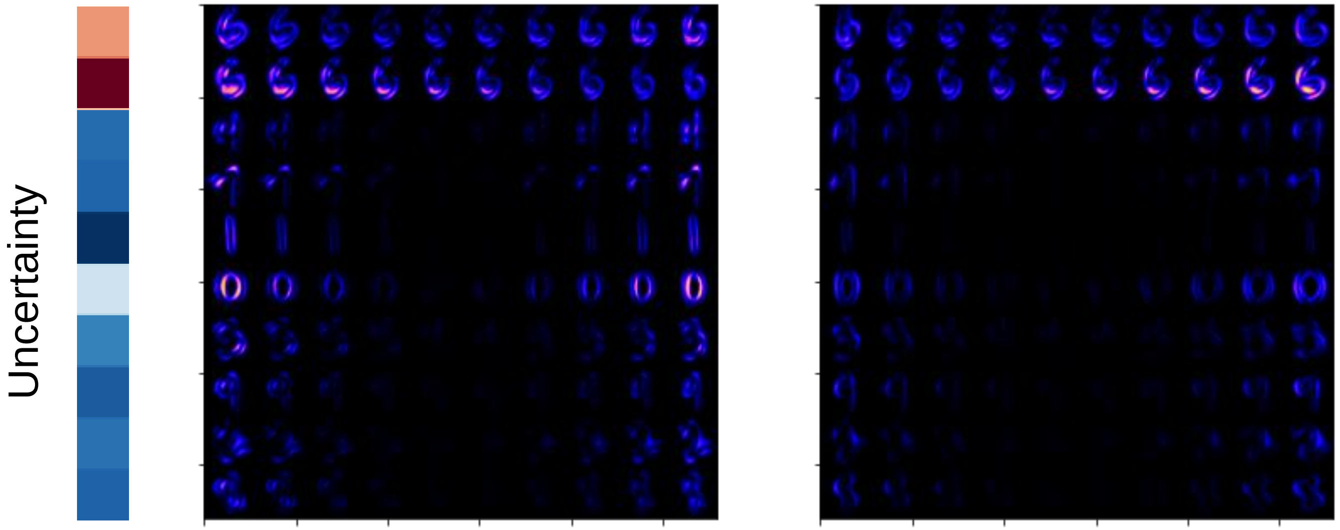
Mean prediction



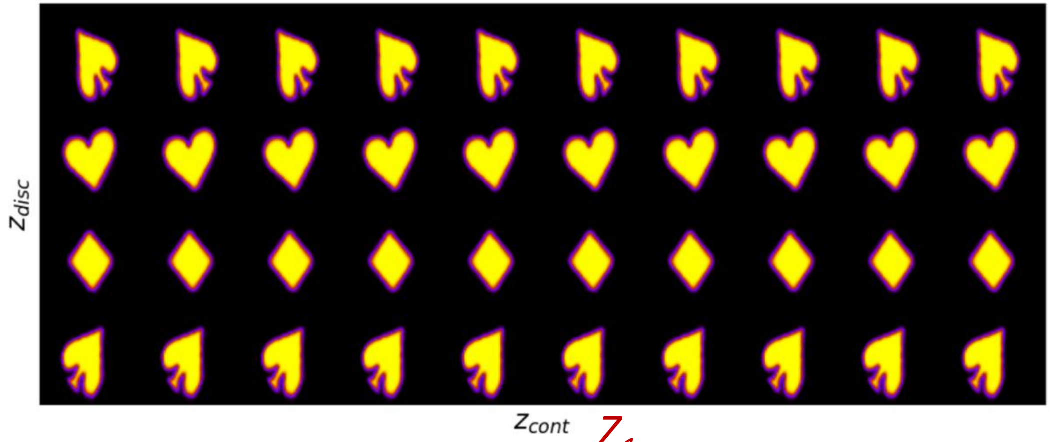
Baseline: 10 epochs
Ensemble models: 8

- The unstable classes show the largest “uncertainty”
- Indication of the quality of separation and/or a guide for selection of the number of classes

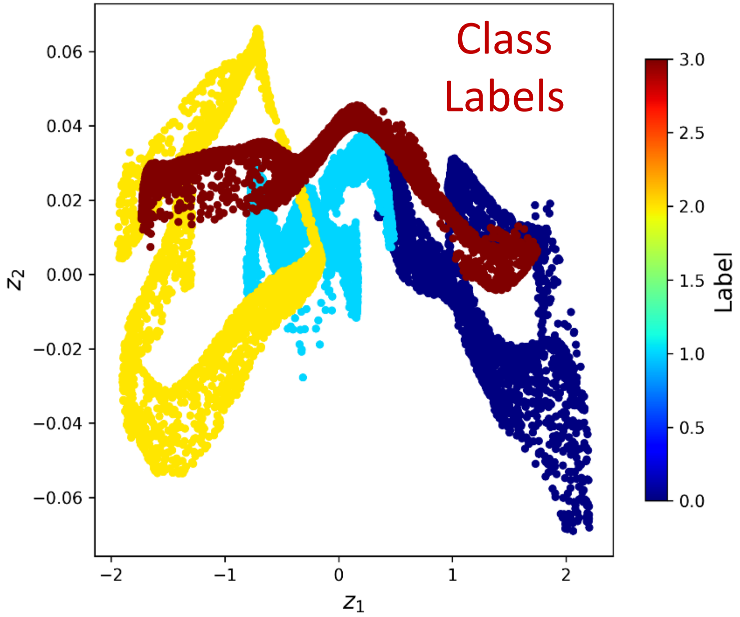
Dispersion in predictions (‘uncertainty’)

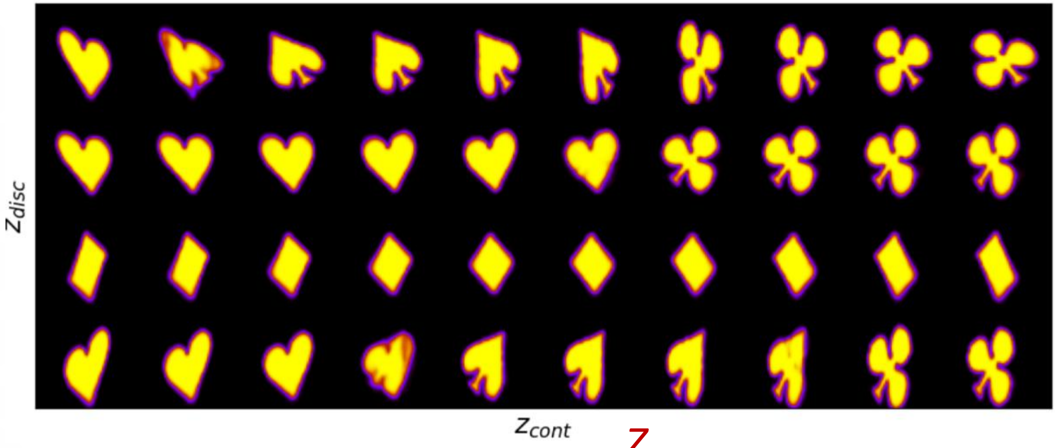


jVAE on cards

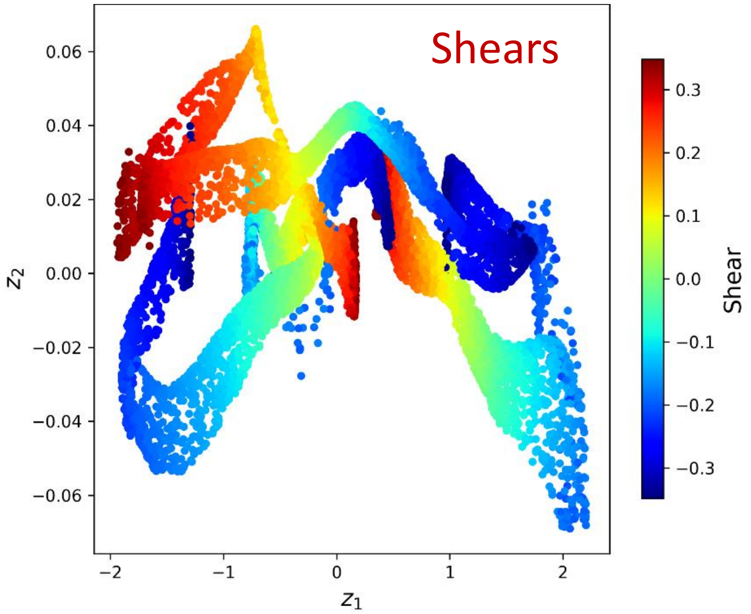


Z_1





Z_2



jVAE on graphene