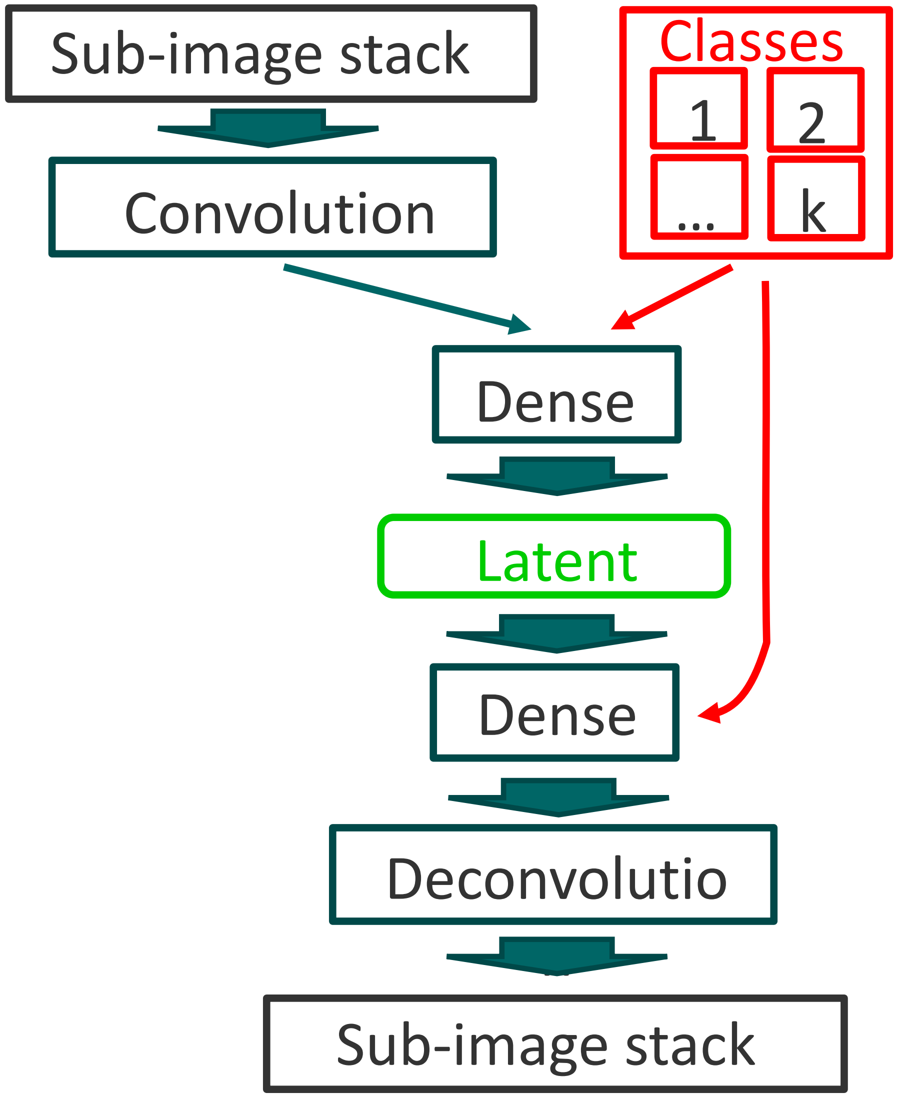
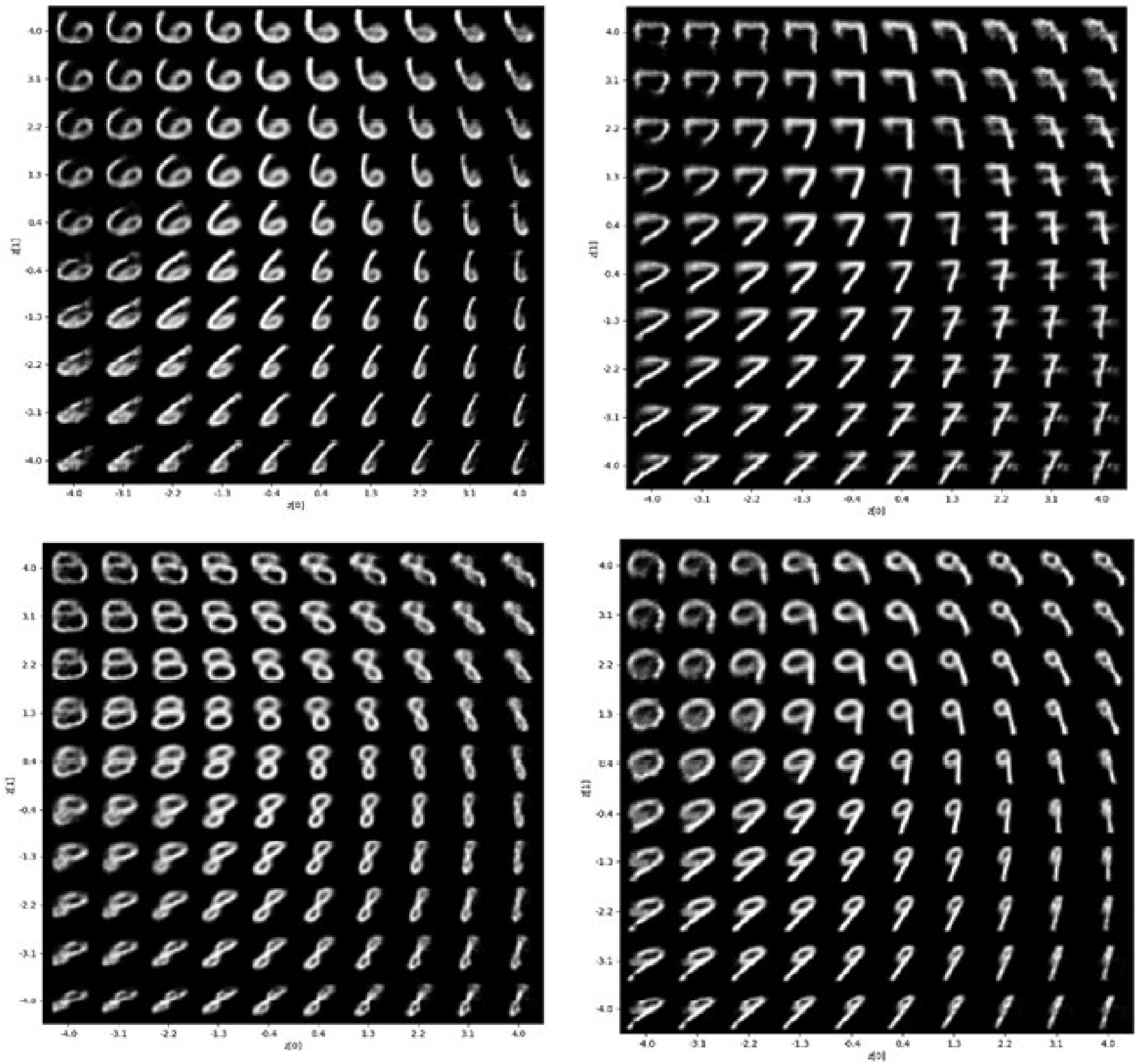
THE UNIVERSITY of TENNESSEE KNOXVILLE

Lecture 31: Invariant Variational Autoencoders

Sergei V. Kalinin

Conditional VAE

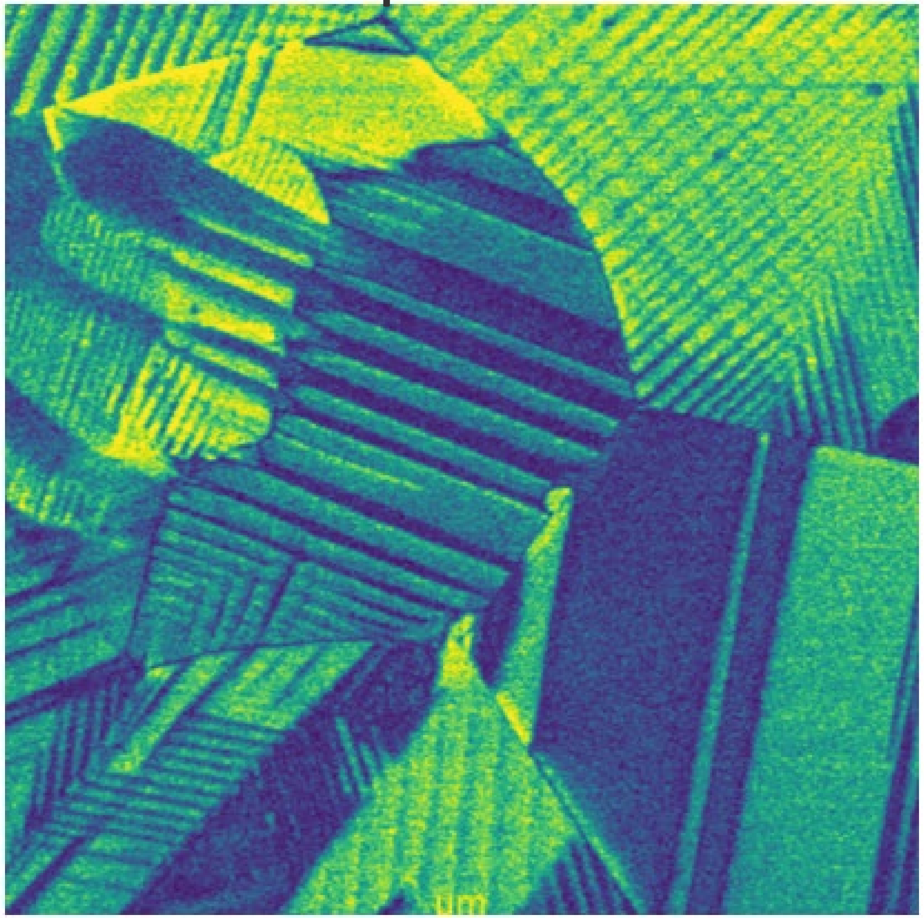




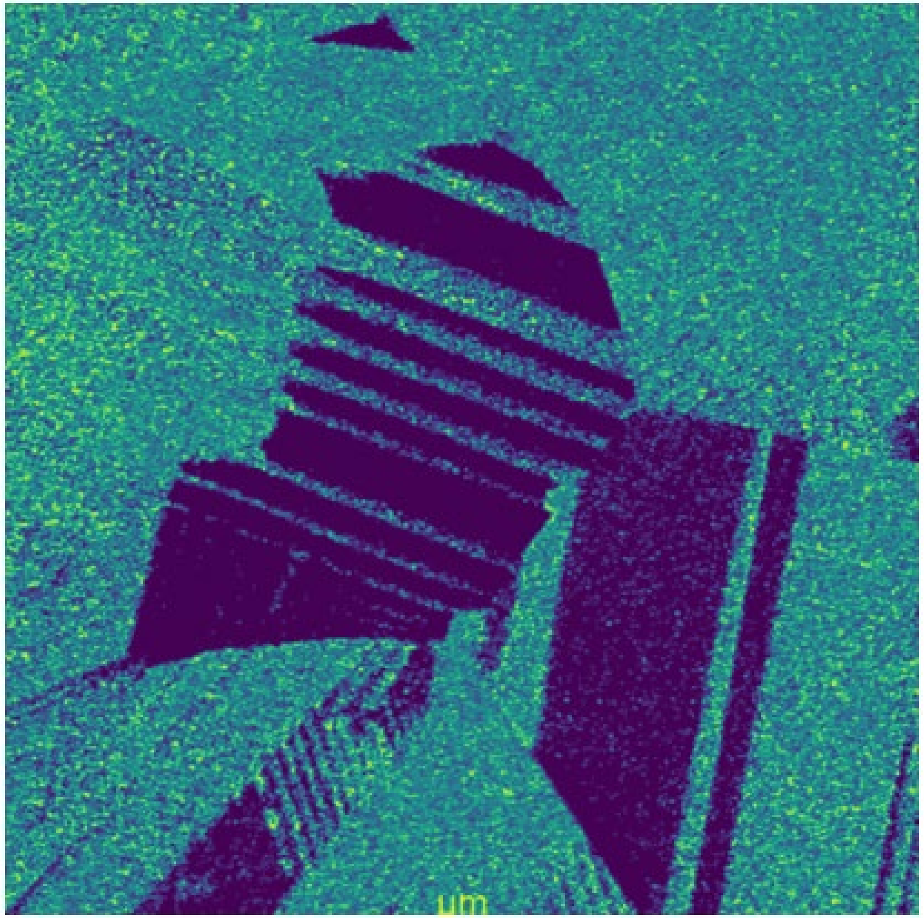
Note the trends in the latent representation for each digit: disentanglement of the representations

Ferroelectric domain and domain walls

Amplitude

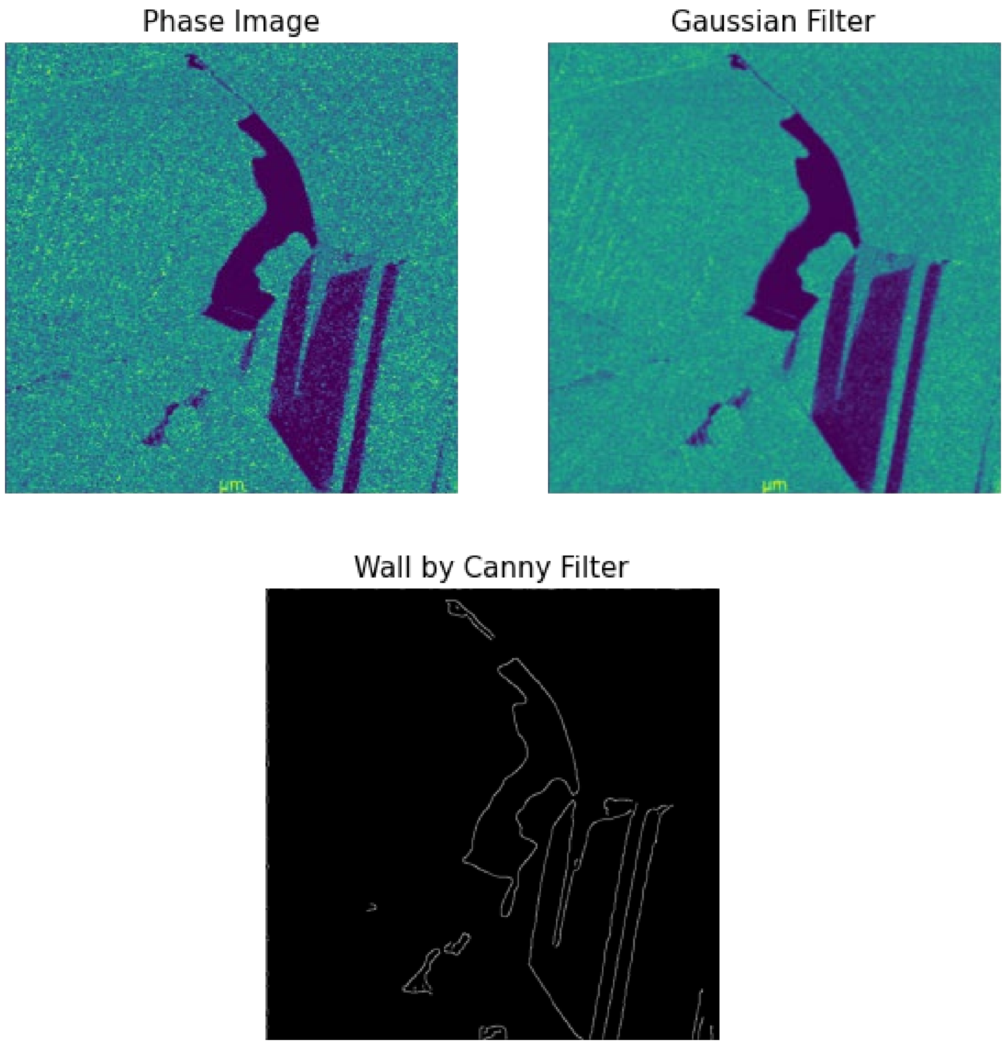


Phase

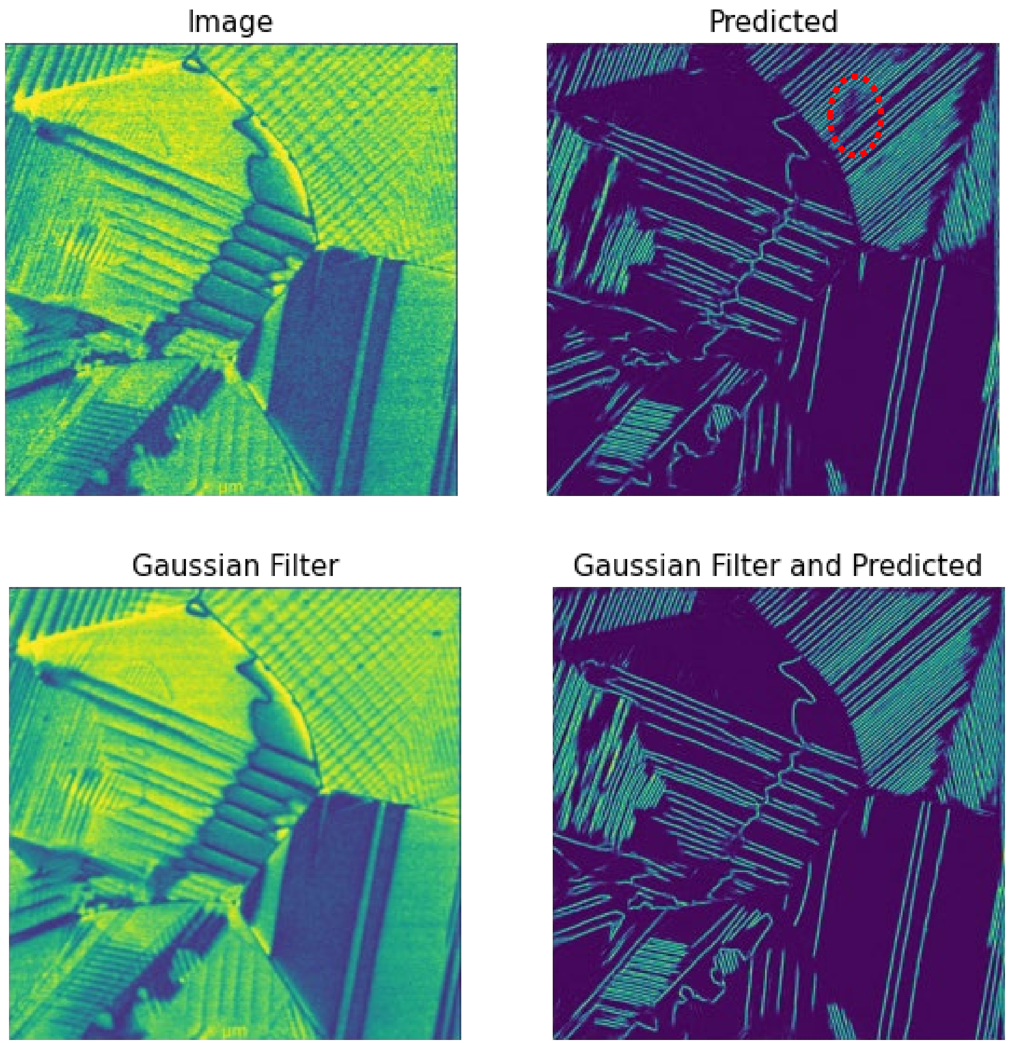


Detecting domain walls

Canny filter

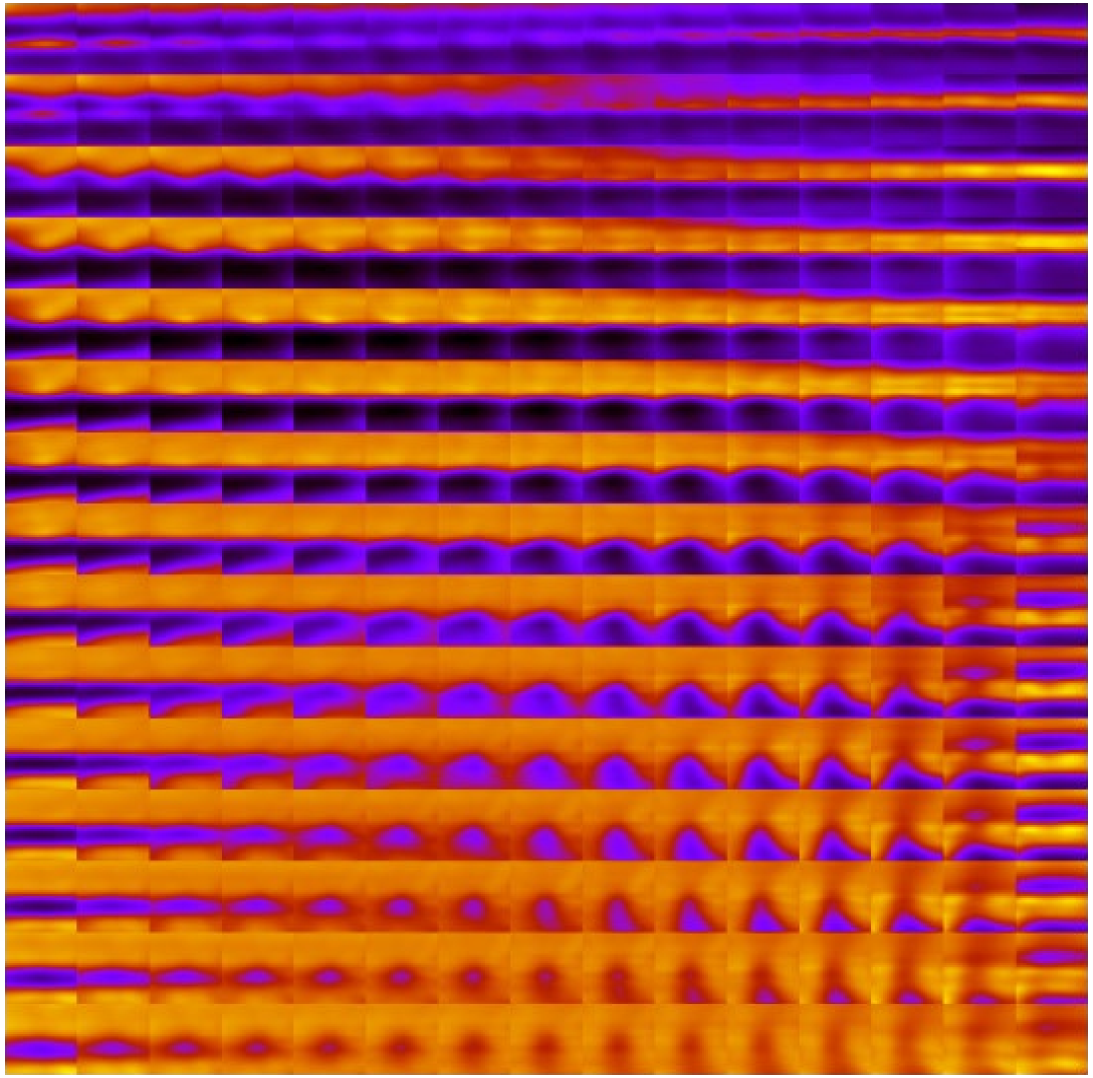


DCNN Prediction

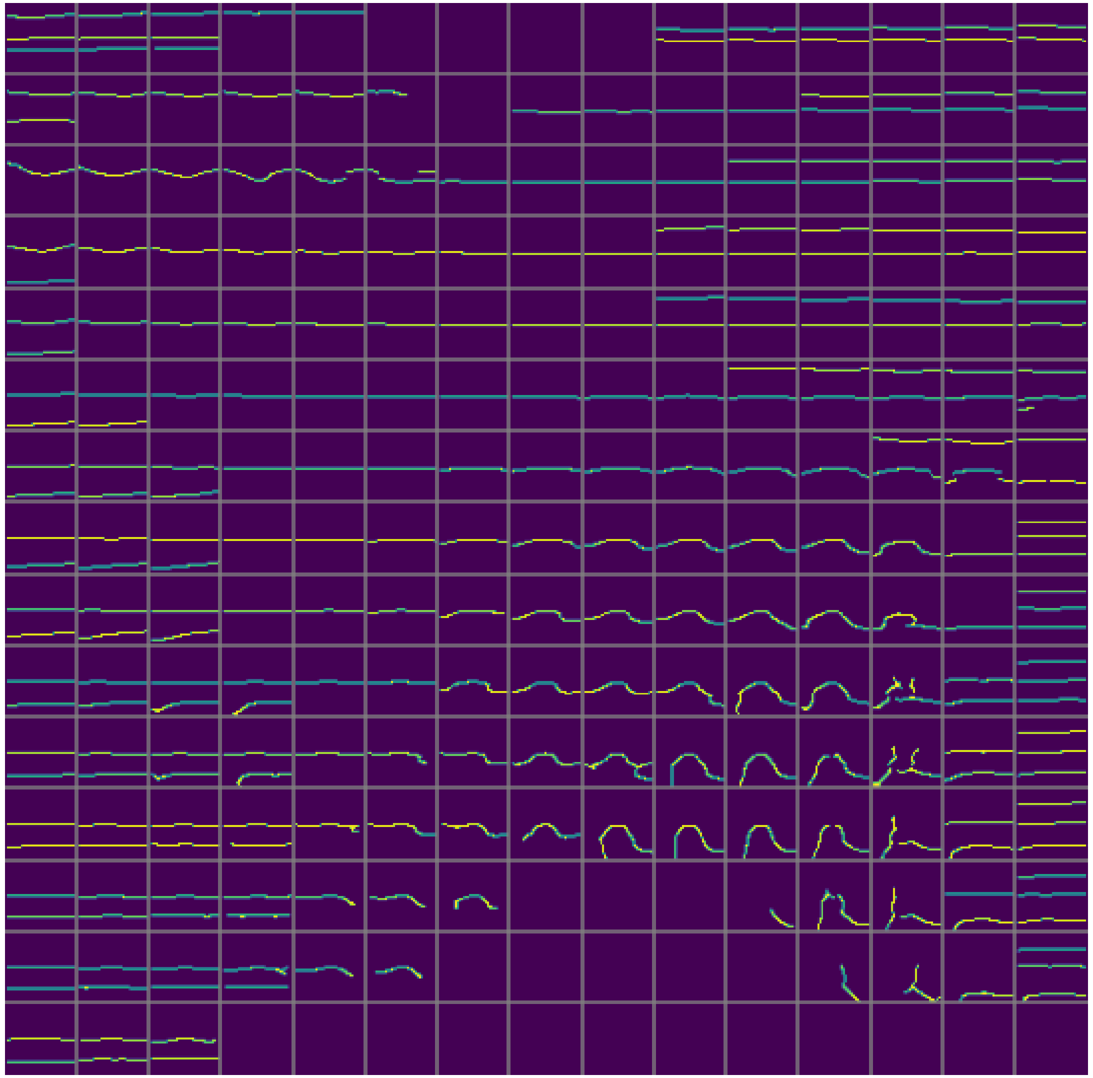


rVAE analysis

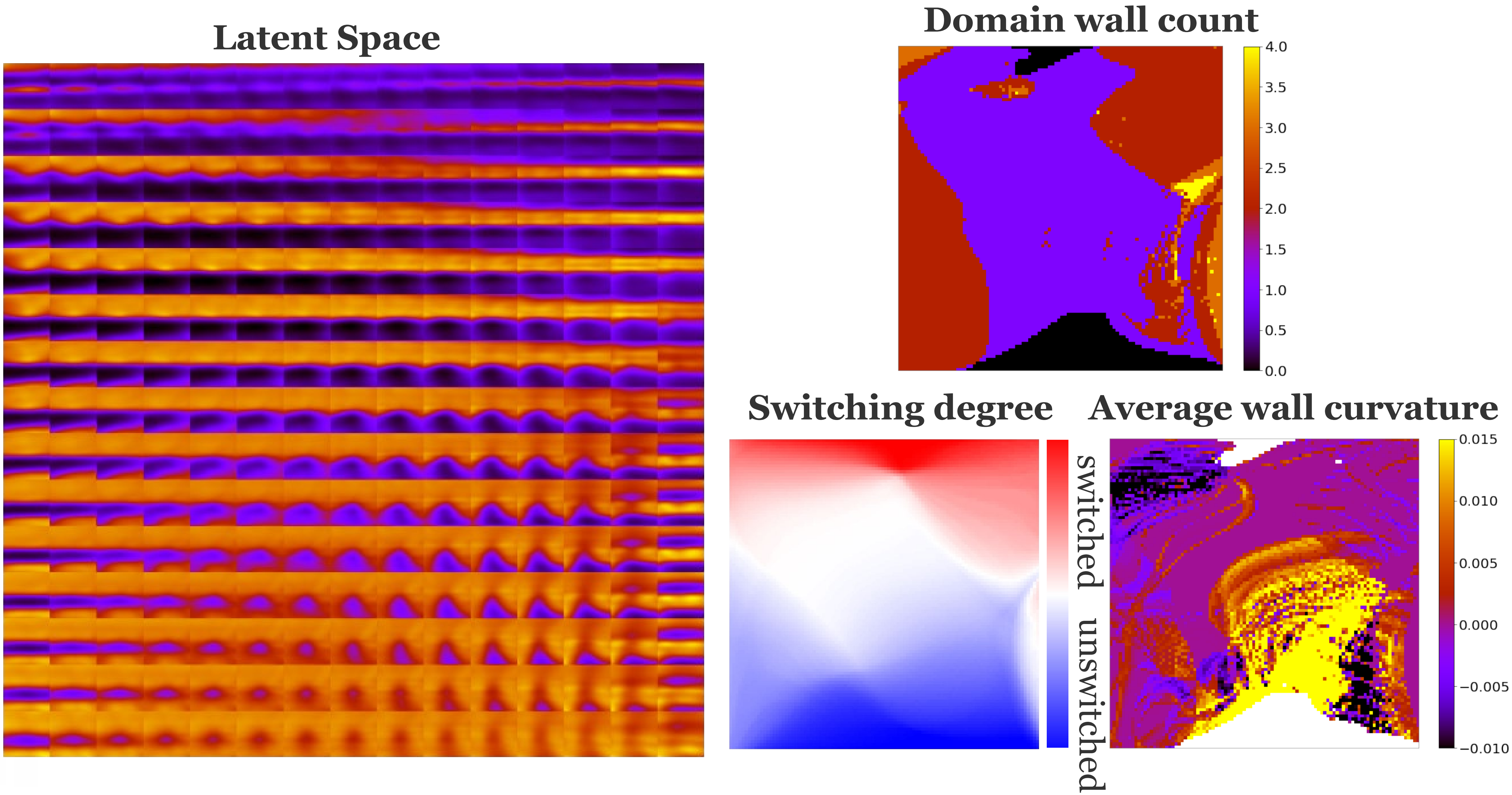
Latent Space



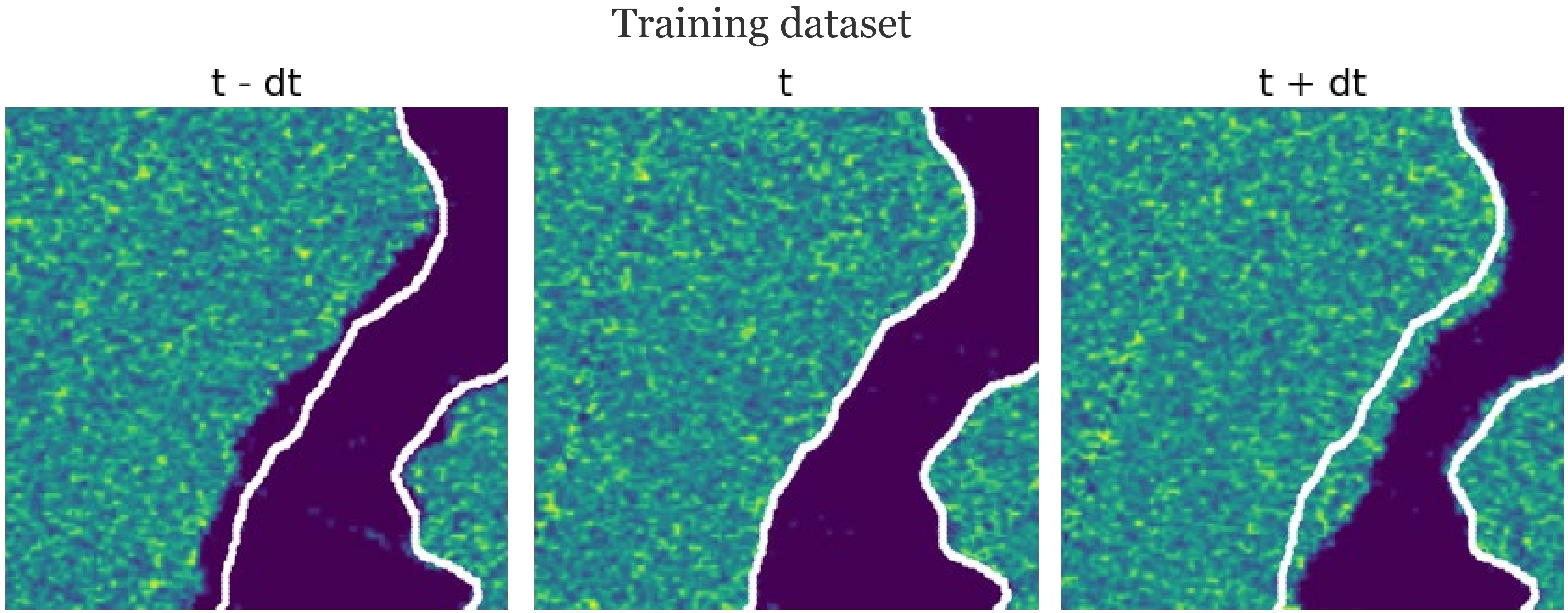
Domain Walls



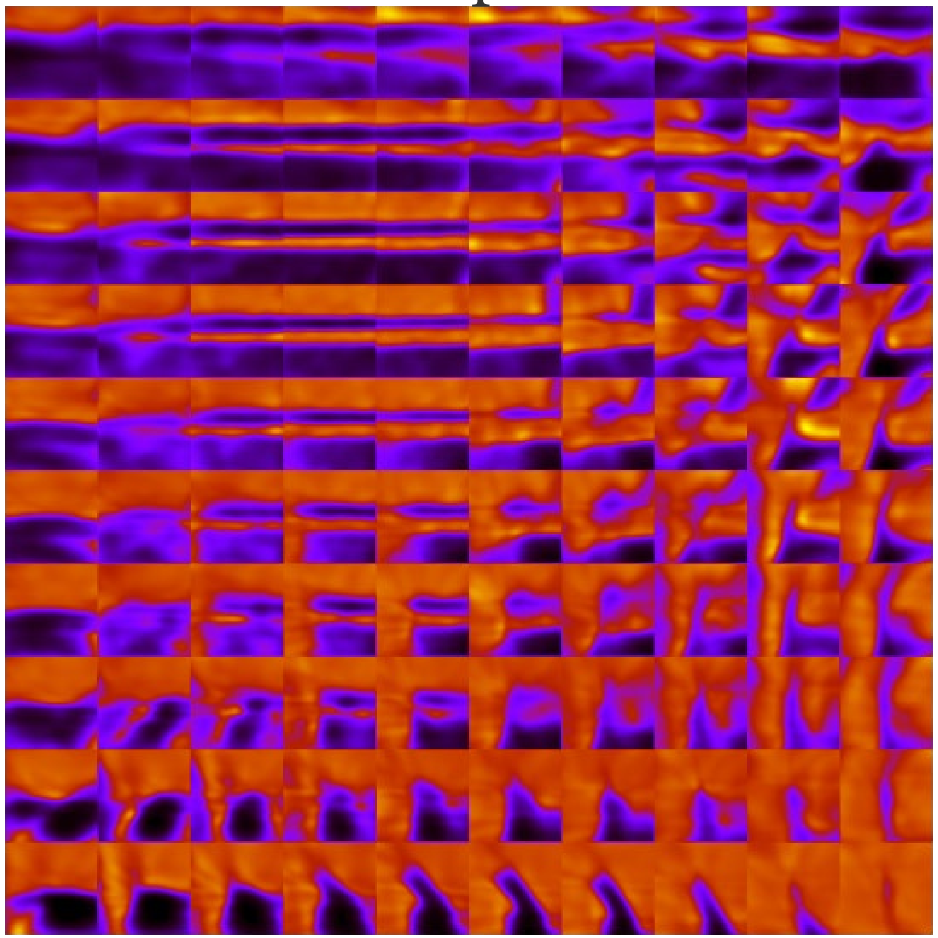
rVAE latent space

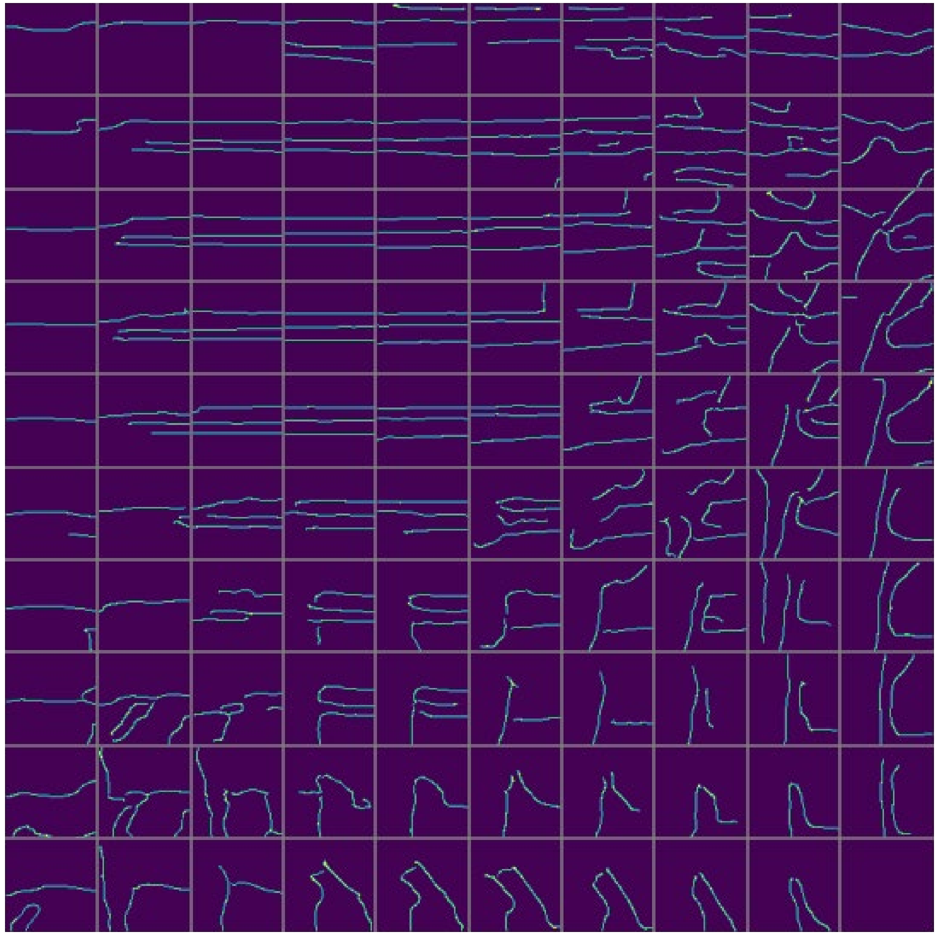


rVAE with time delay



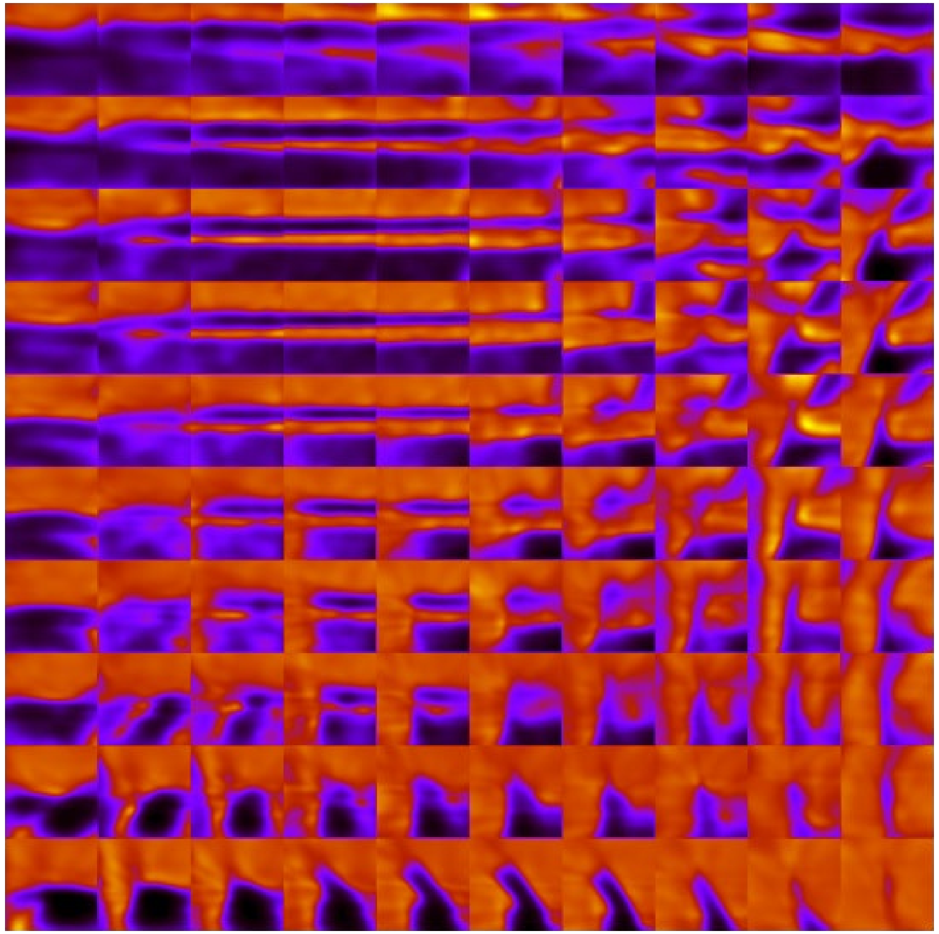
rVAE with time delay



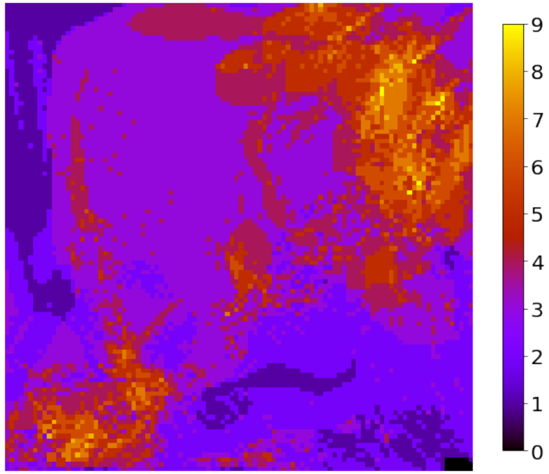


rVAE with time delay

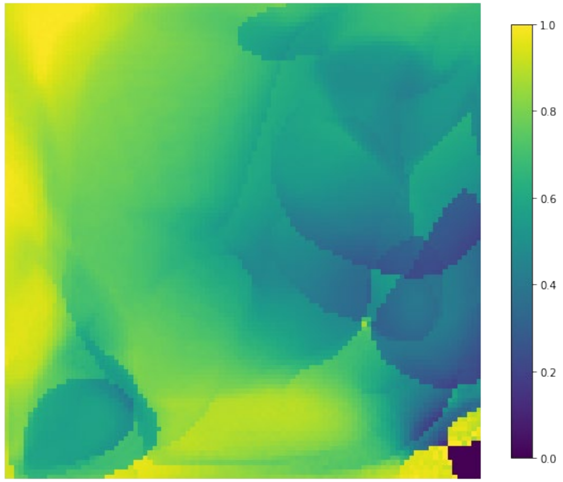
Latent space



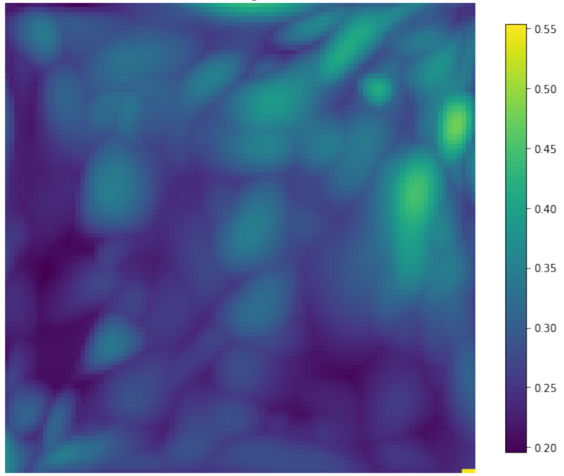
Wall count



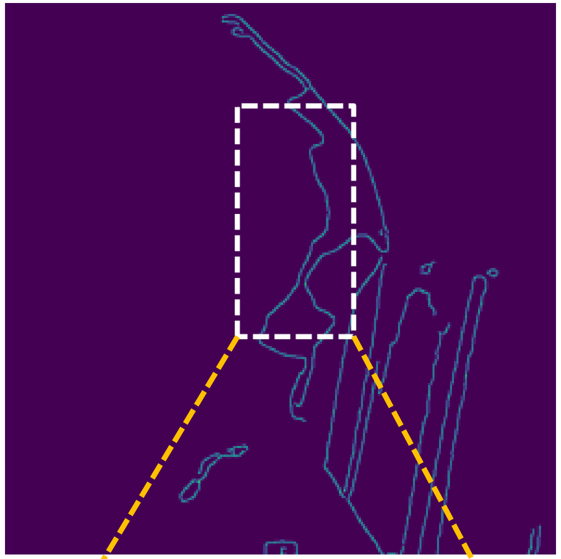
Domain convex

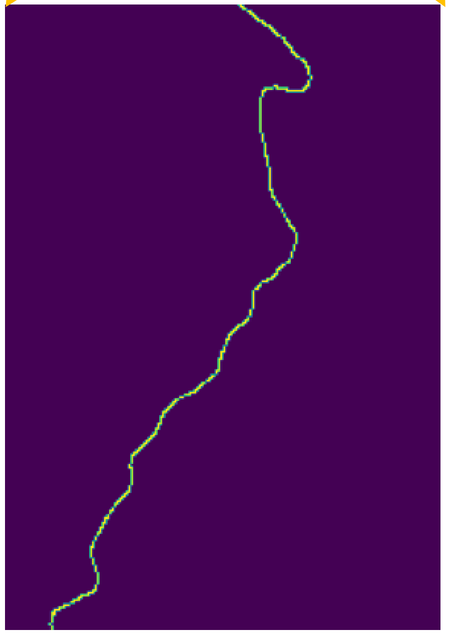


Switch significance

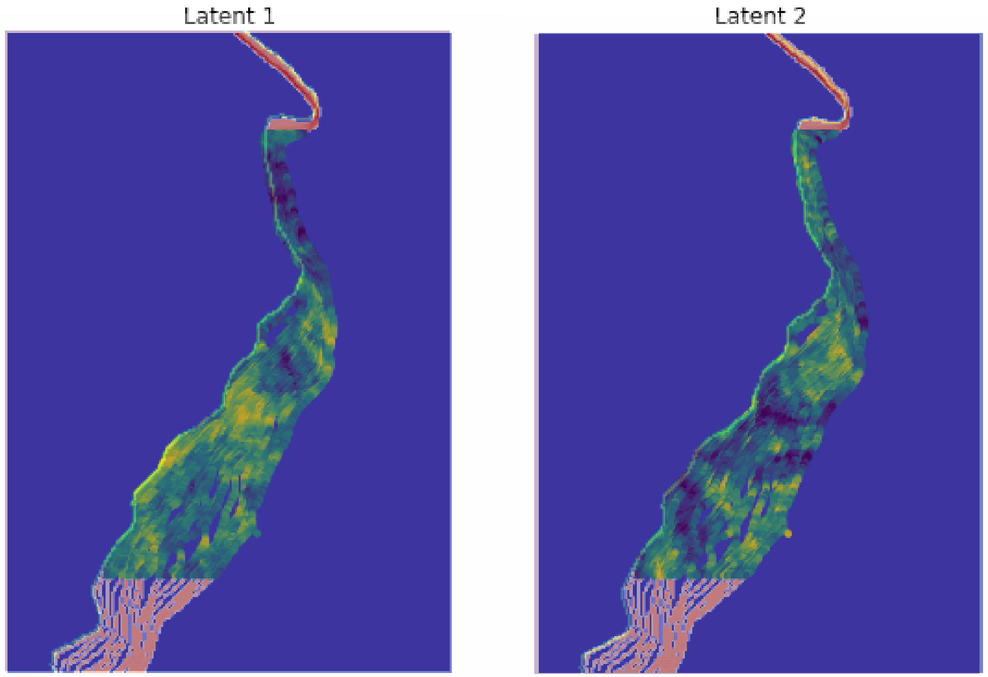


Domain wall evolution

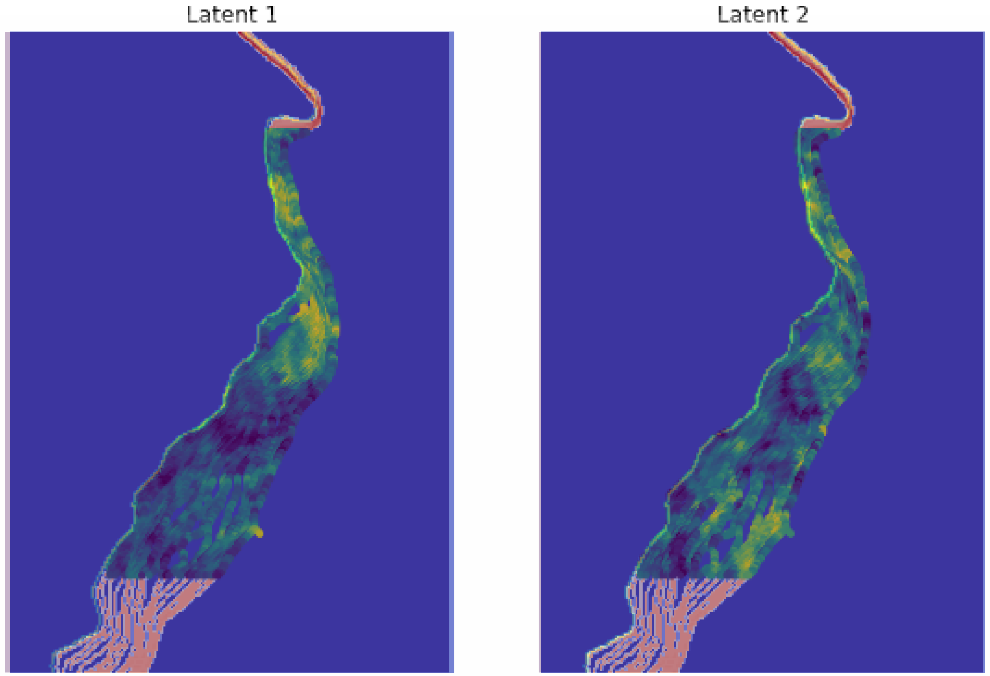




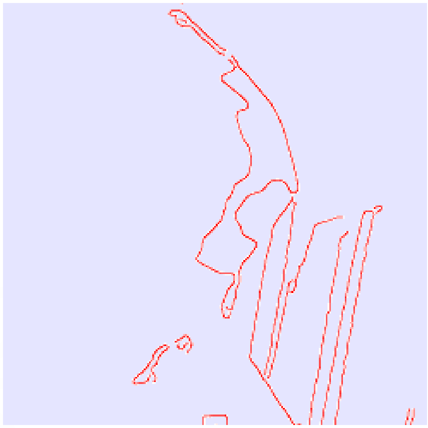
Forward:
t vs t+1

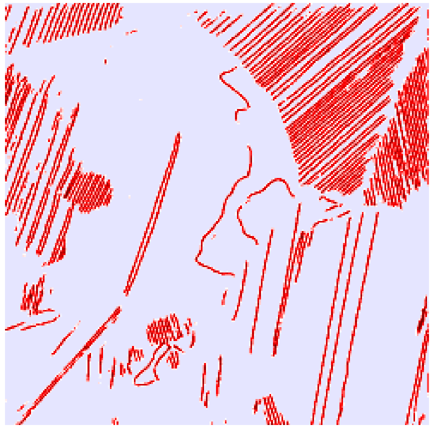


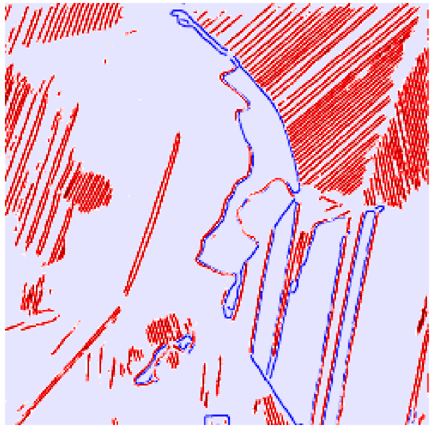
Reverse:
t vs t+1

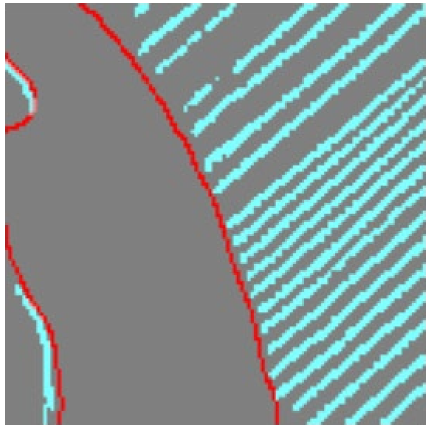


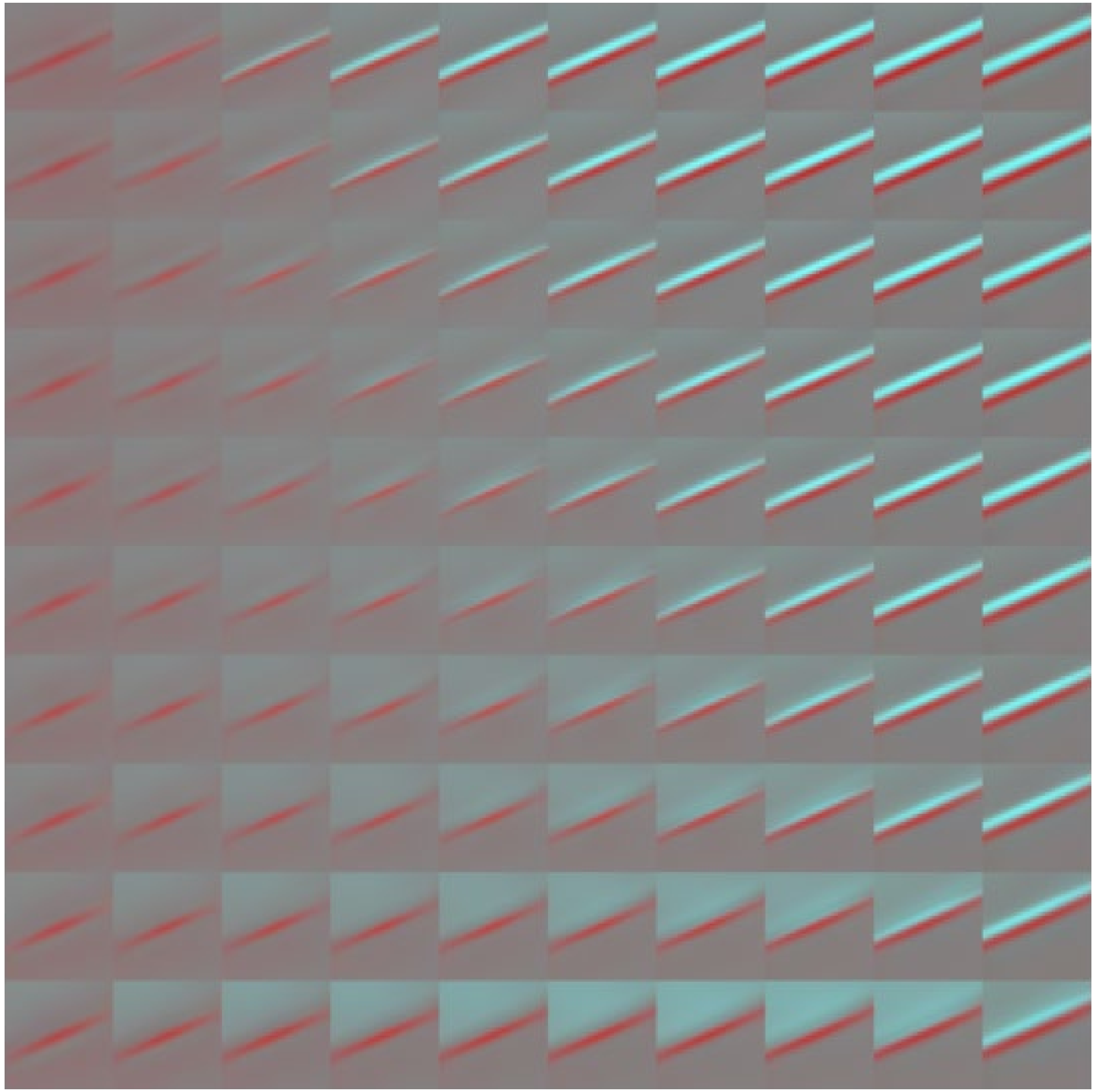
Multilayer rVAE



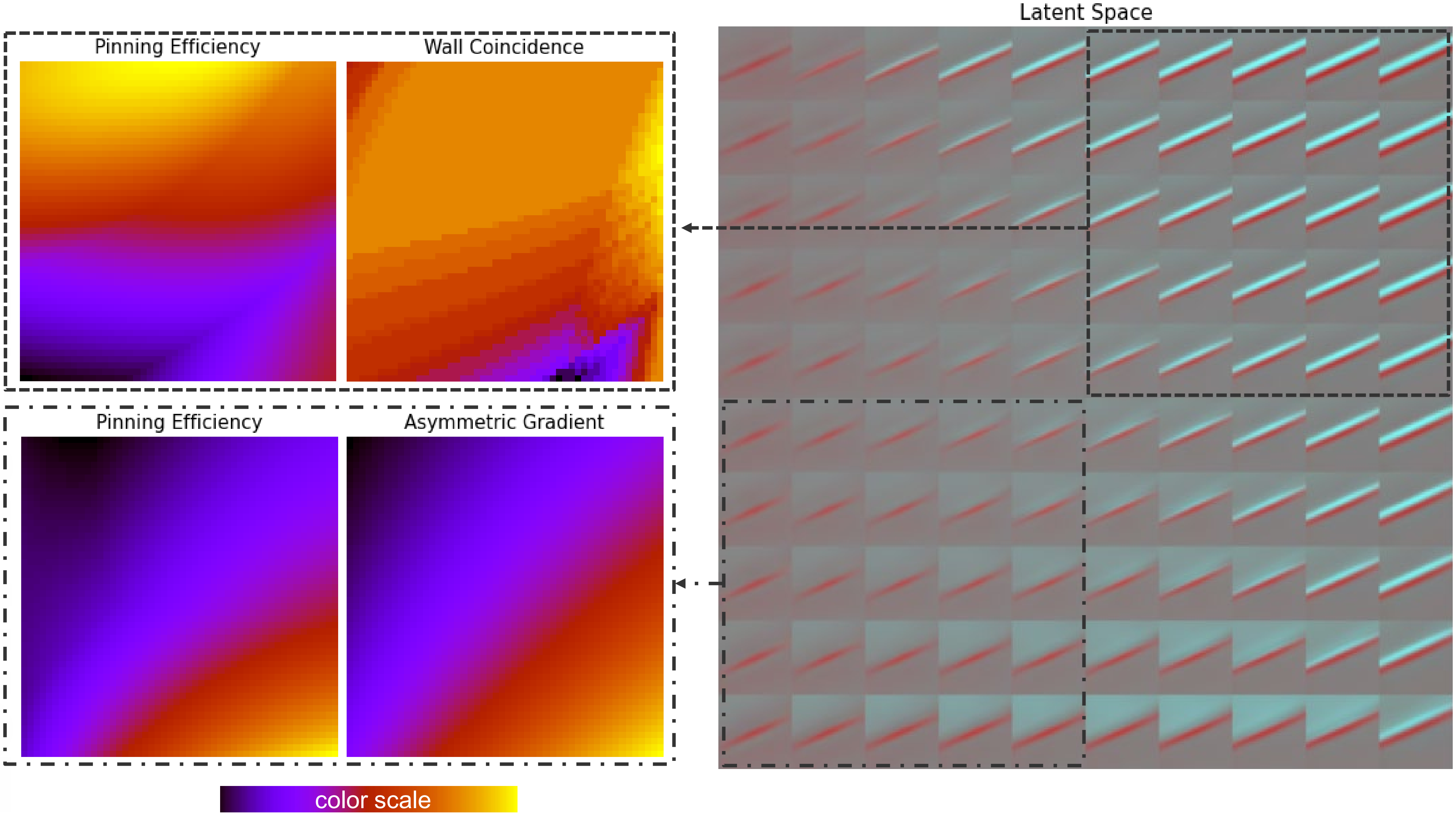








Pinning mechanism



What if we have multiple classes?

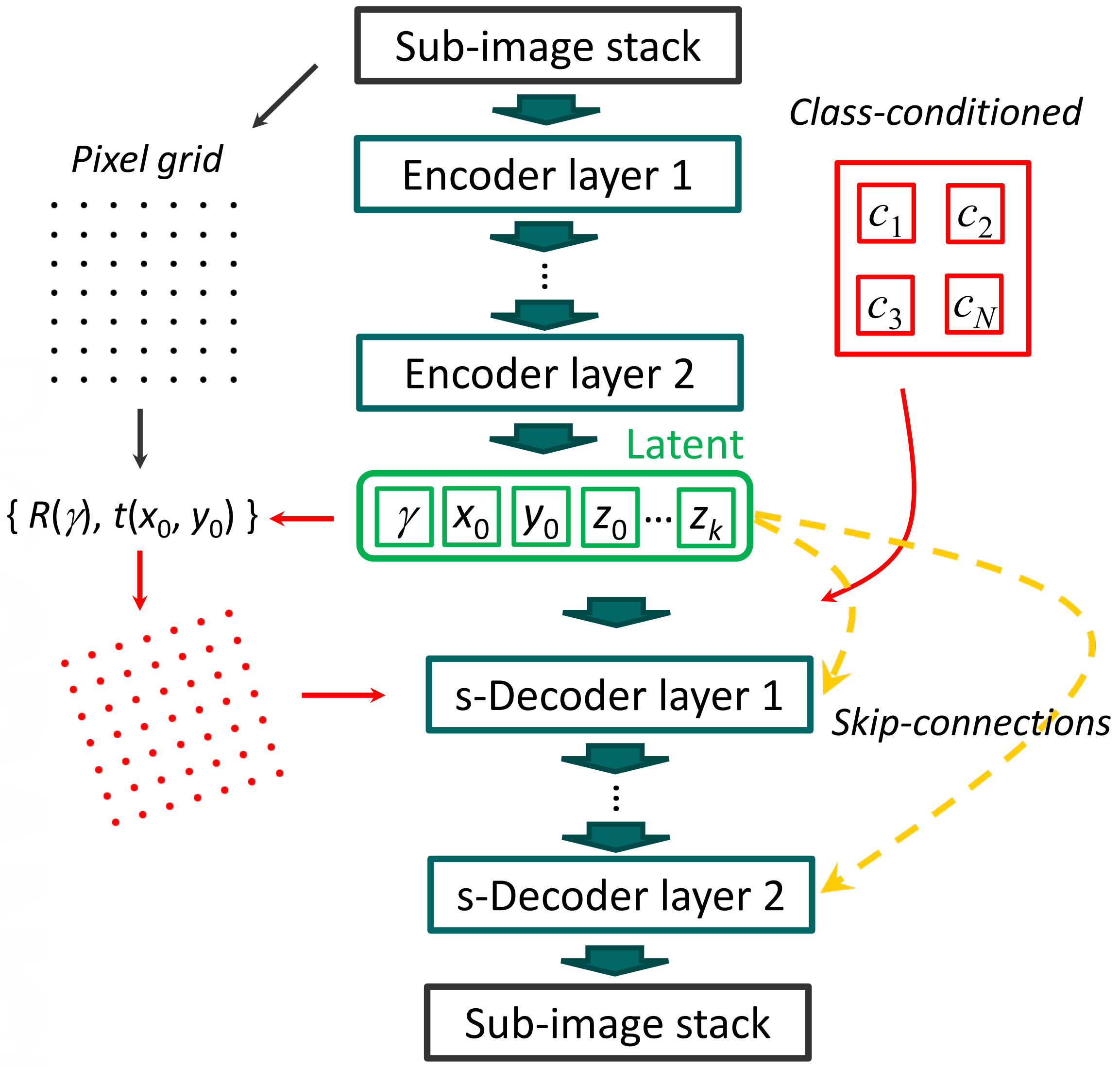
1. Classes are known: conditional (discrete) VAE

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

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

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

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

+ D_{KL}(physics-based “priors”) ?

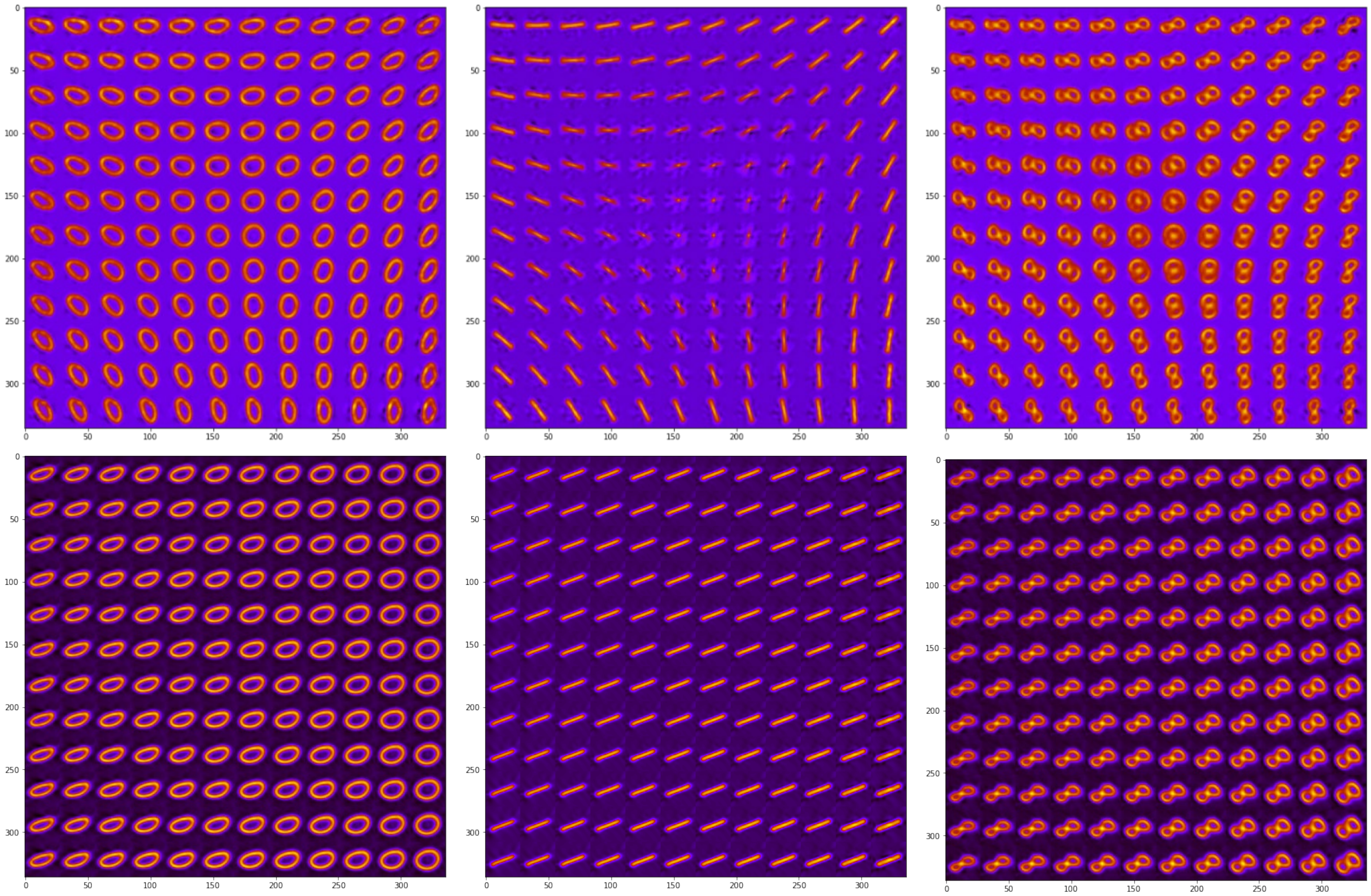
+ D (physics) ?

Bepler *et al.*, NeurIPS 2019

MNIST cVAE

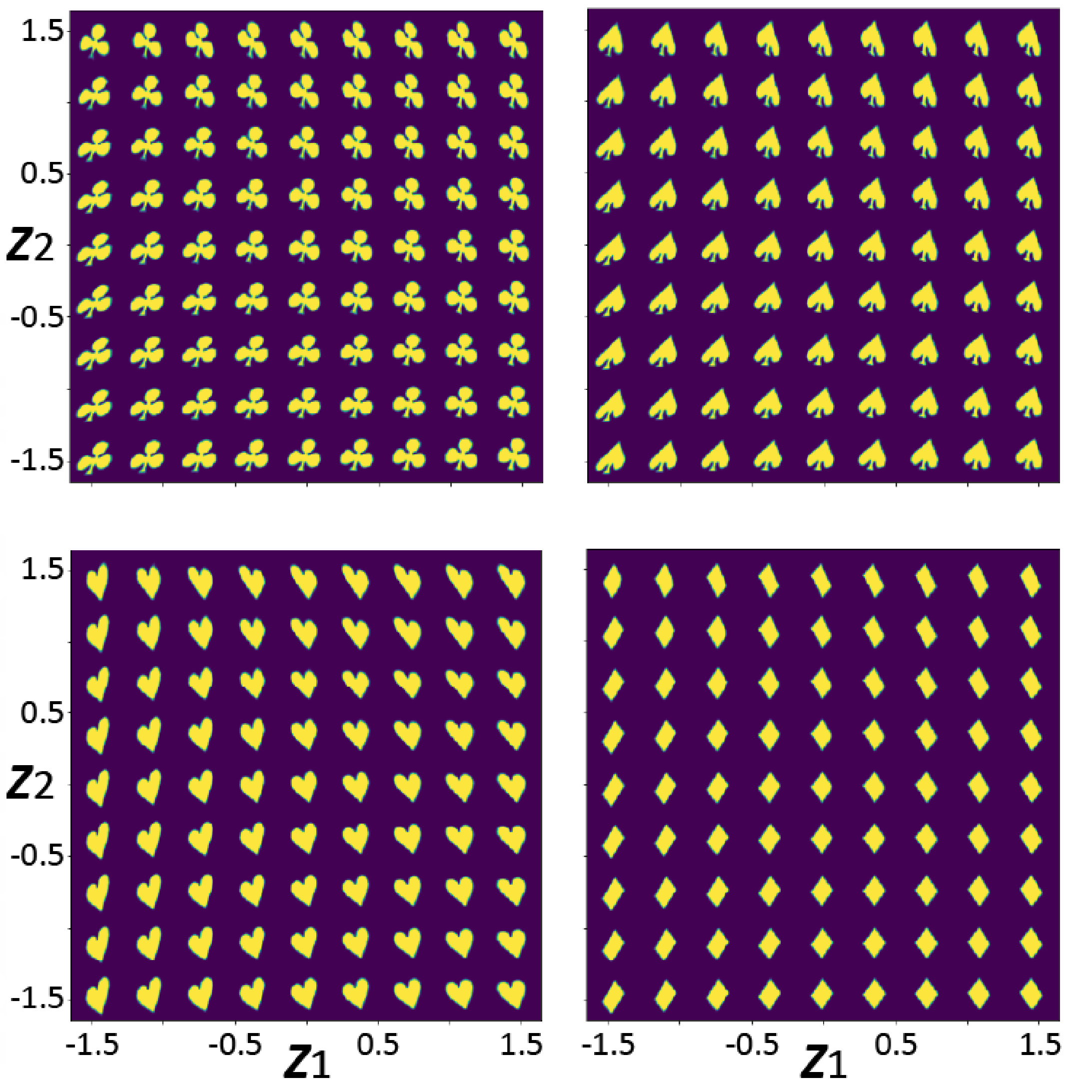
No rotations

With rotations

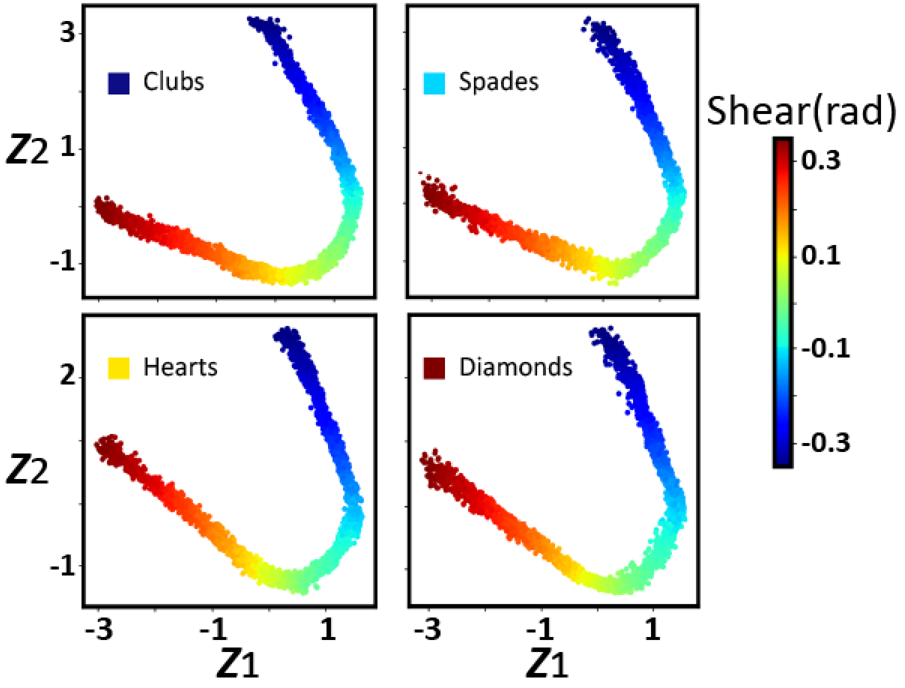


cVAE on cards data set

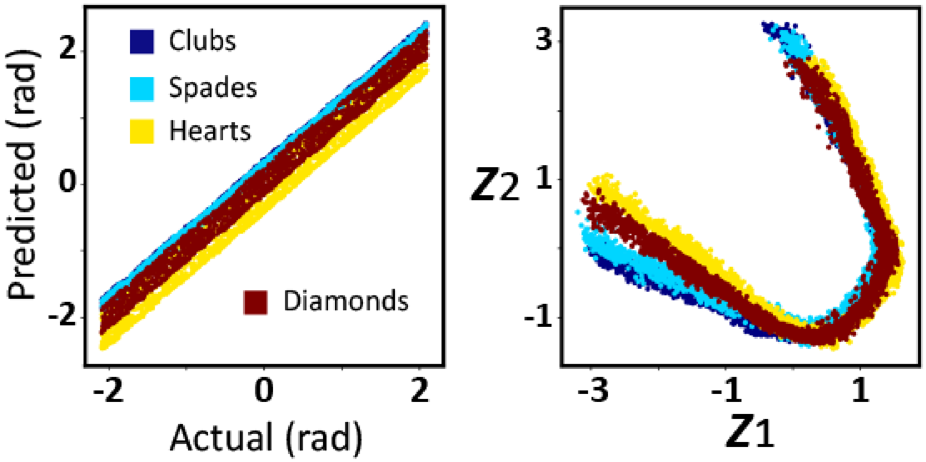
One independent latent space per class



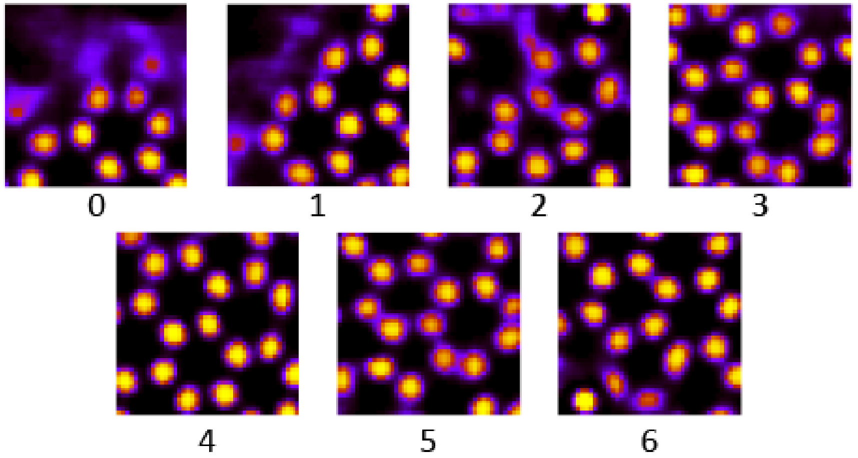
Shear distribution in latent space



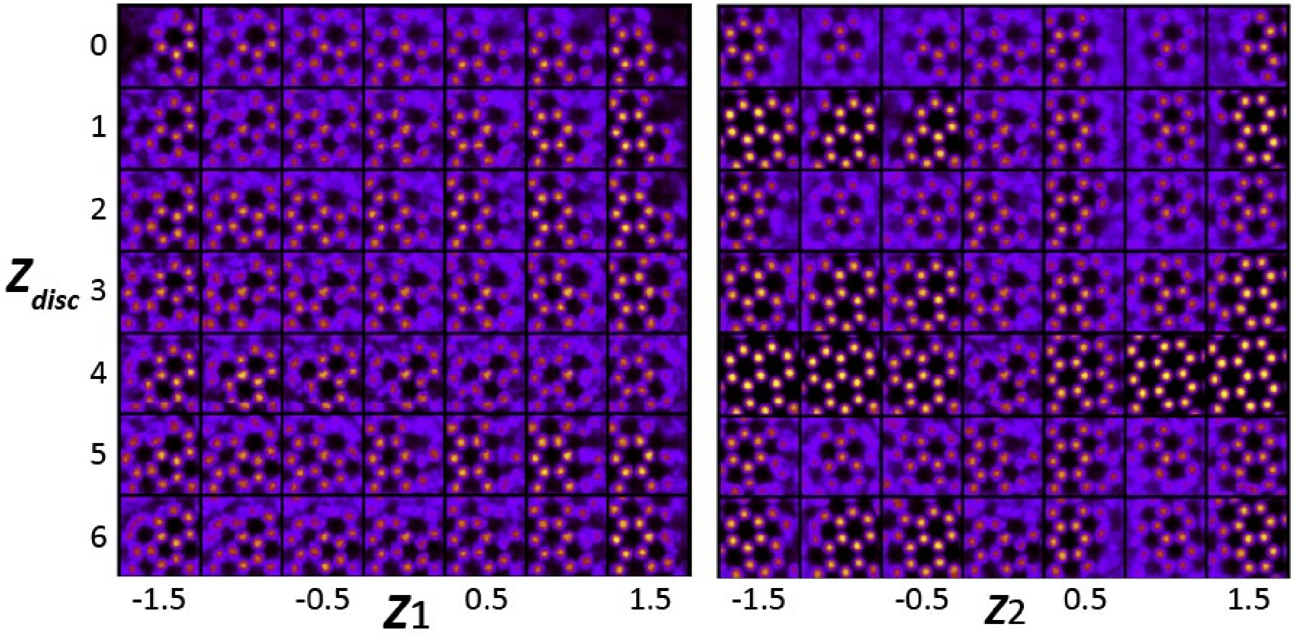
All four plotted jointly



cVAE on Graphene



Traversal manifolds



Autonomous sub-image labeling

0: Fewer than 2 rings

1: two rings of size [6, 6]

2: three rings of size [5, 6]

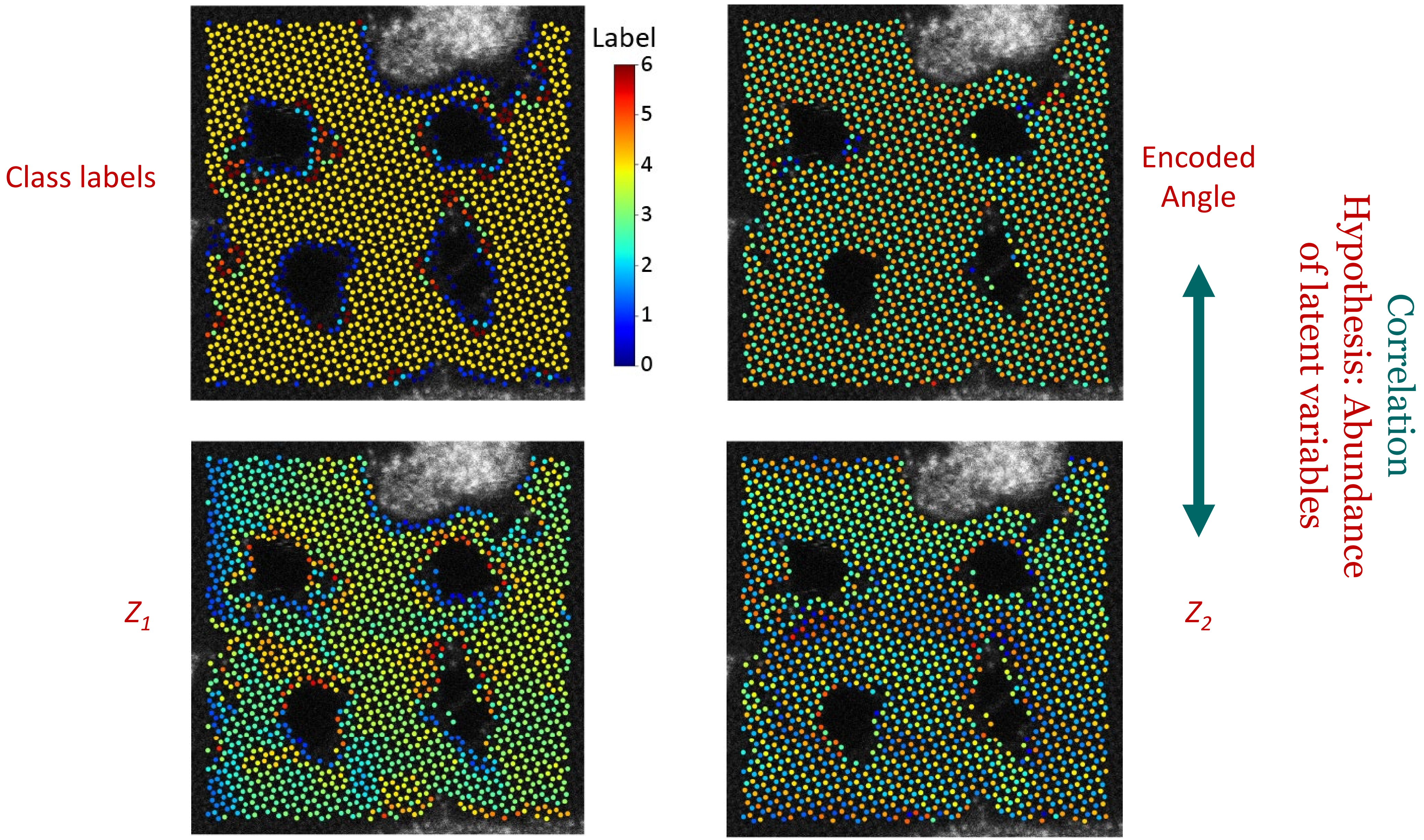
3: three rings of size [6, 6, 7]

4: three rings of size [6, 6, 6]

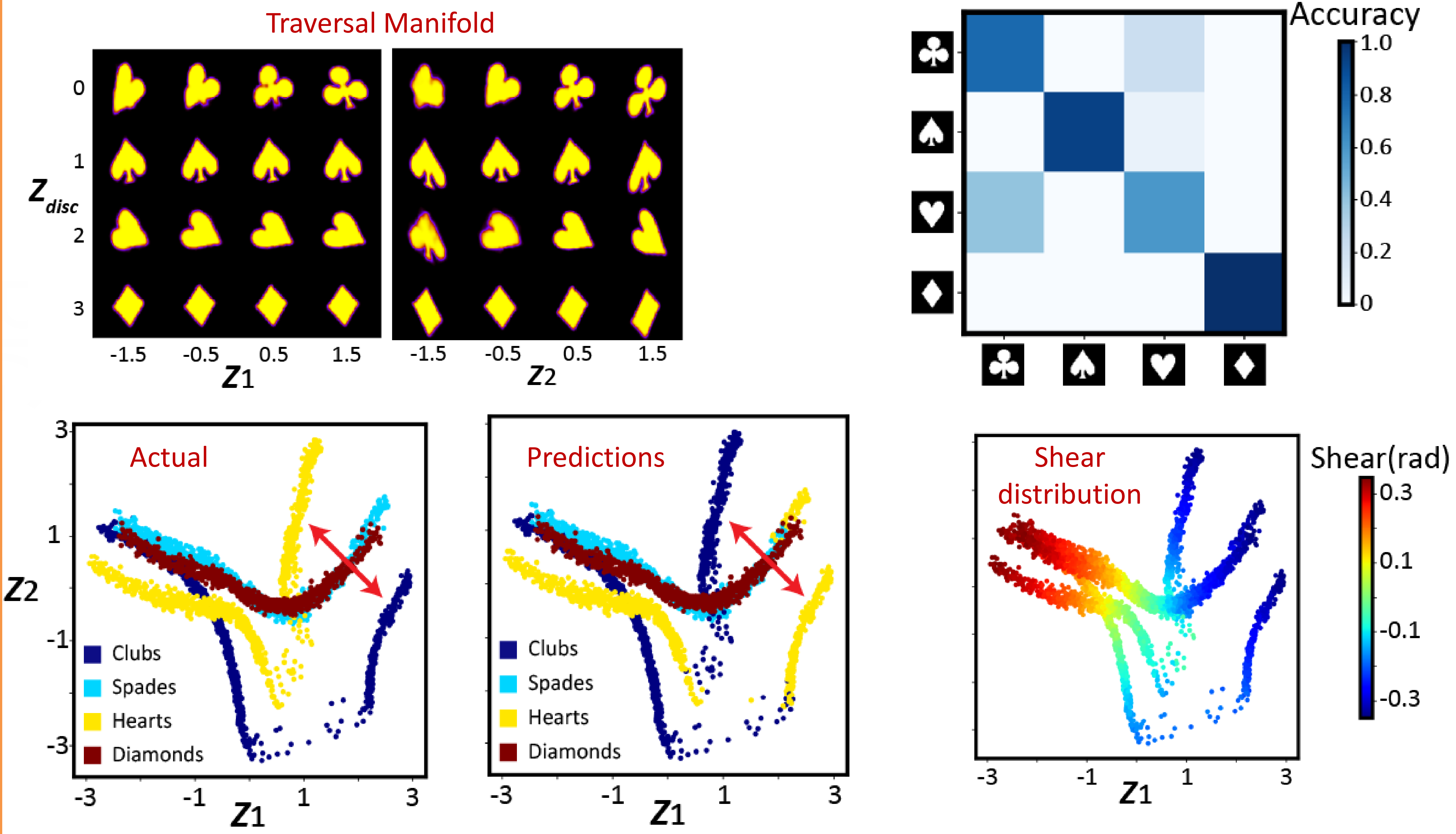
5: three rings of size [5, 6, 7]

6: three rings of size [5, 6, 6]

cVAE on graphene

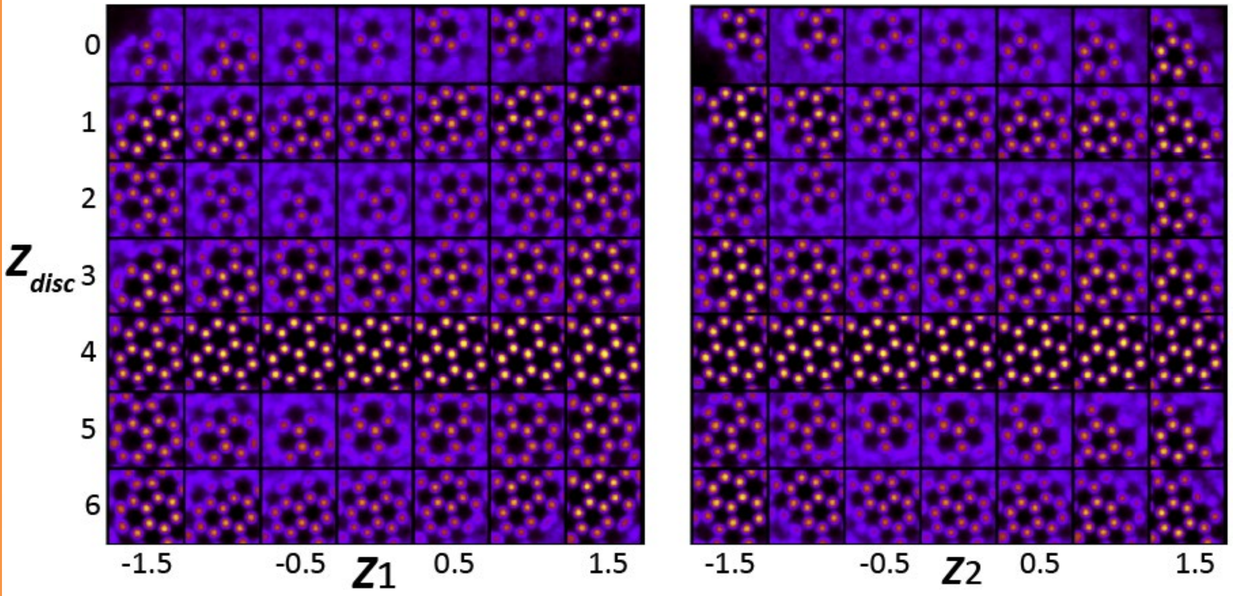


Semi-supervised VAE

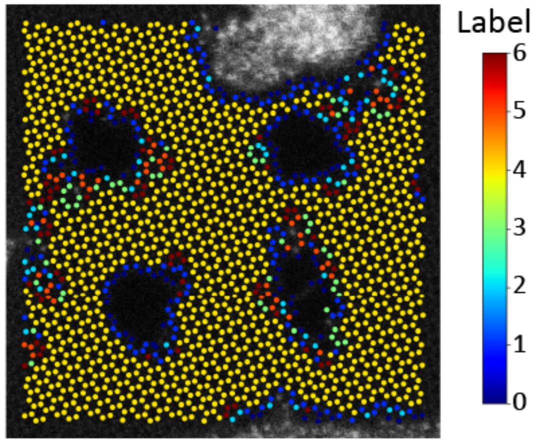


Semi-supervised VAE (graphene)

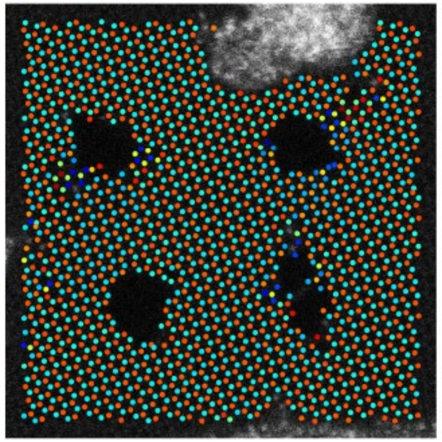
Traverse manifolds



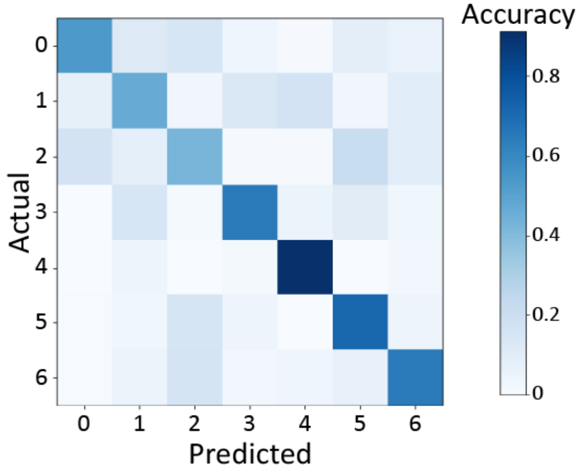
Class labels



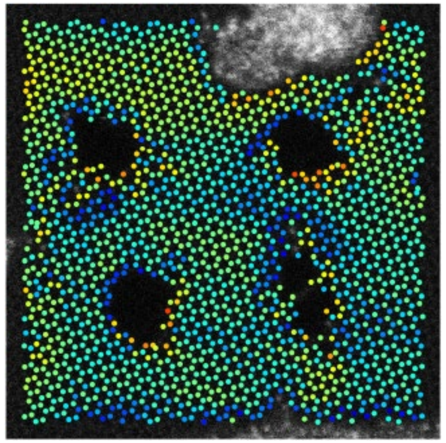
Angle Encoding



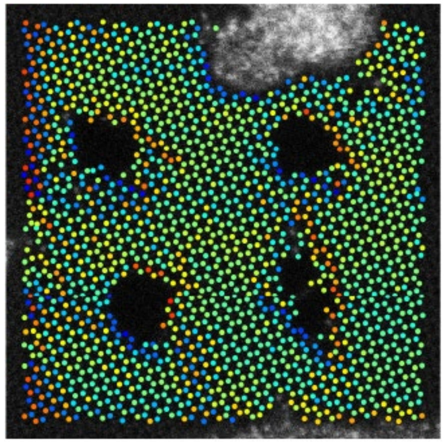
Confusion matrix



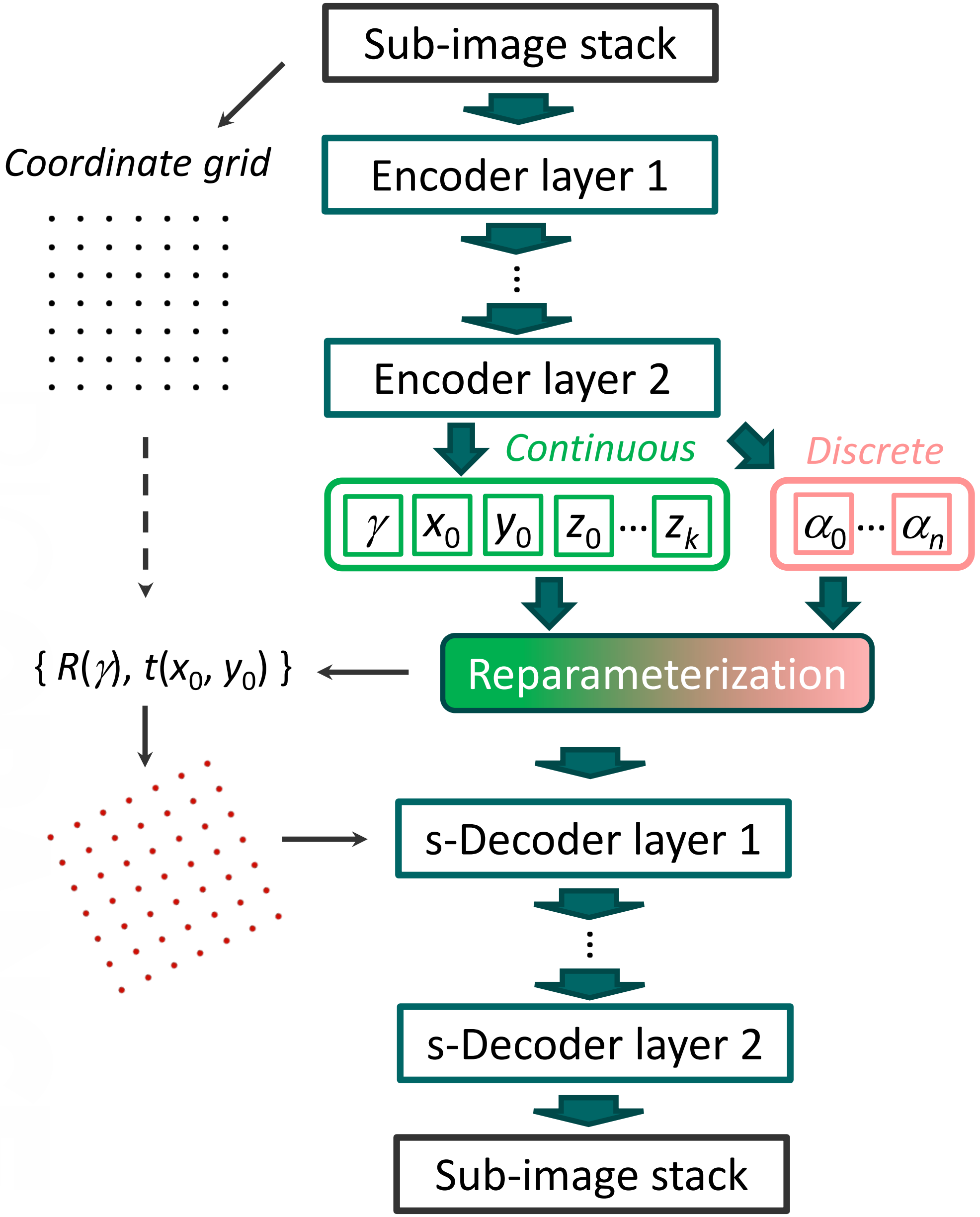
Z_1



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

$$ELBO =$$

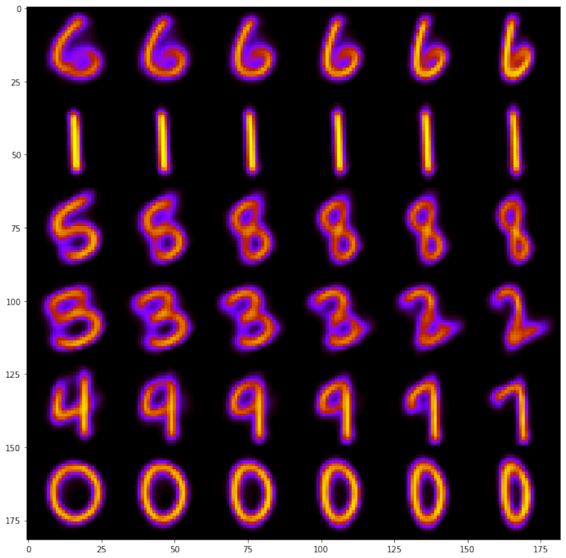
$$-\,Reconstruction\ Loss$$

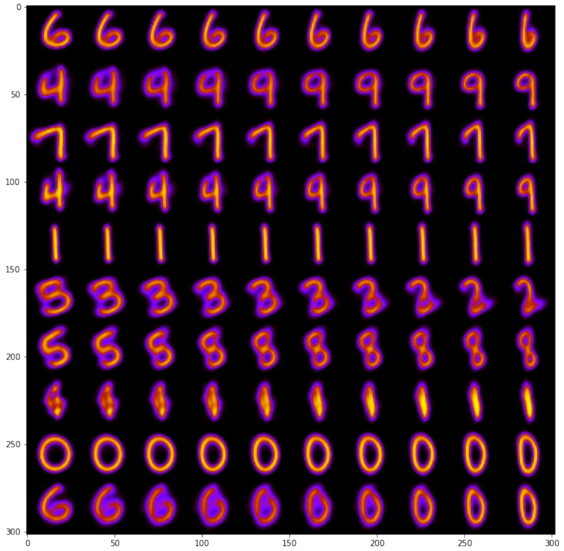
$$-\beta_c(t)\left|\left(\overset{\text{Structure}}{D_{KL}(q(z|x)\|p(z))} + \overset{\text{Rotation}}{D_{KL}(q(\gamma|x)\|p(\gamma))}\right) - C_z\right| \quad \text{Continuous}$$

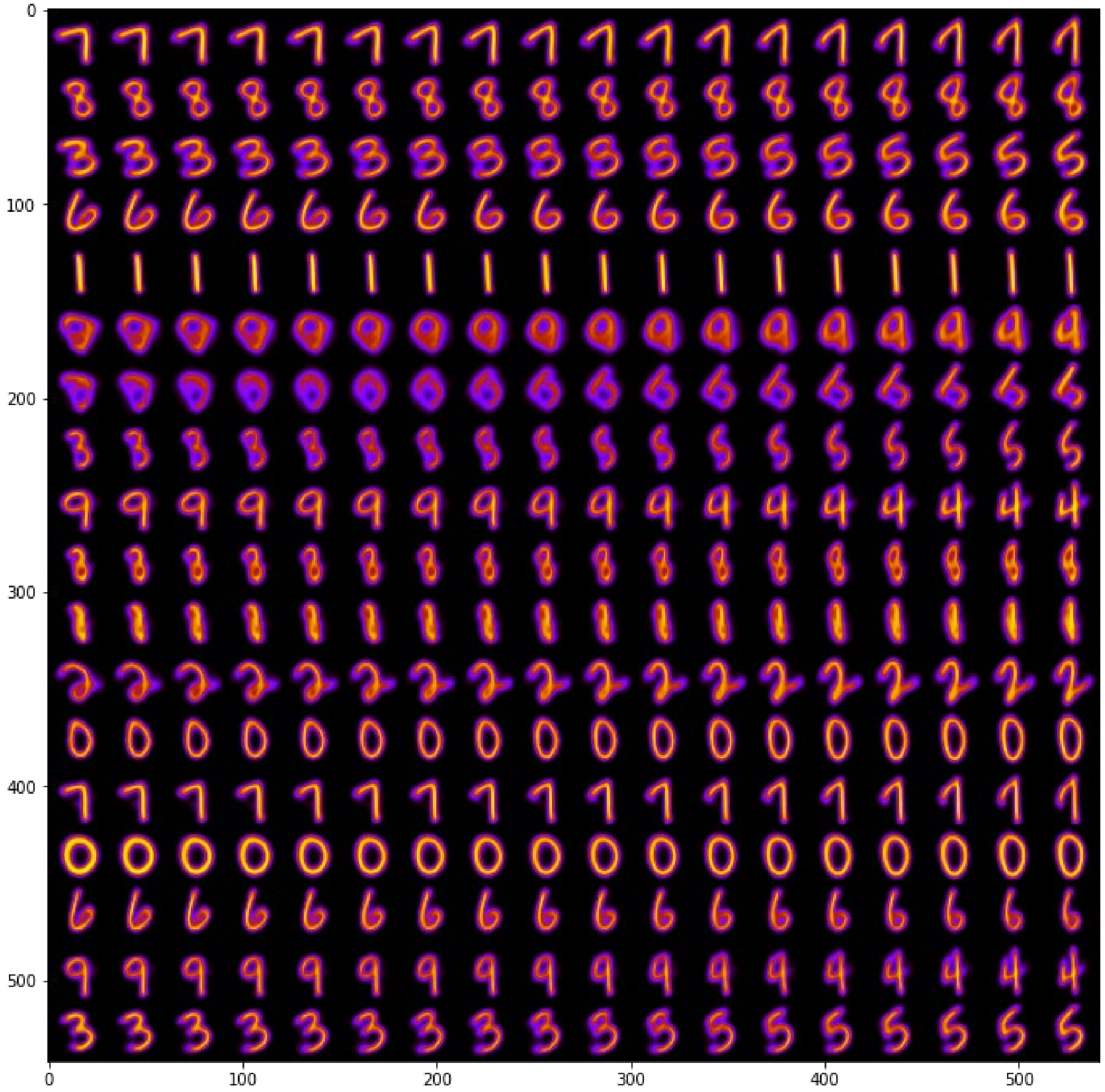
$$-\beta_d(t)\left|D_{KL}\left(q(\alpha|x)\|p(\alpha)\right) - C_\alpha\right| \quad \text{Discrete}$$

+ physics-based “loss” ?

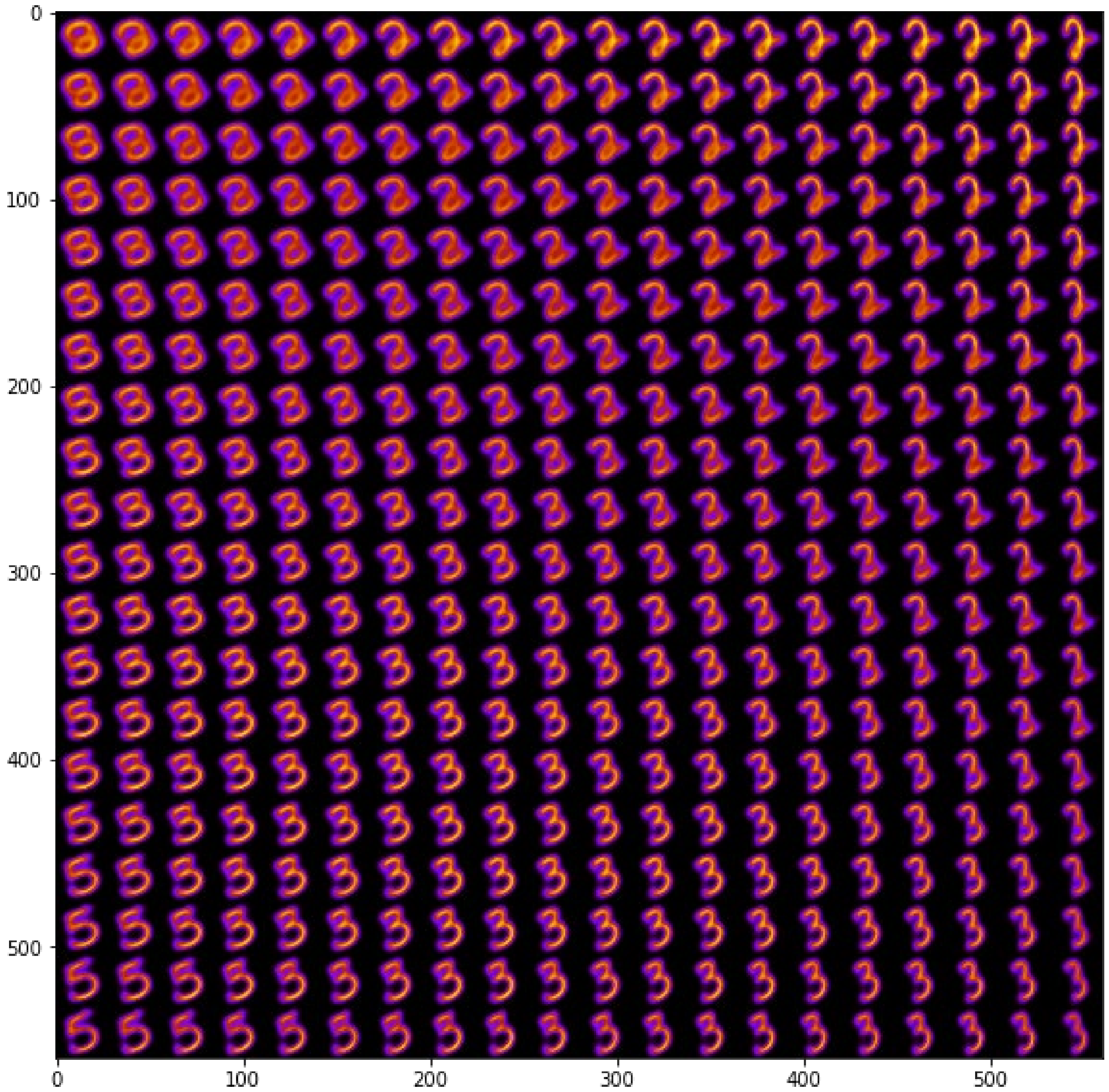
jVAE on MNIST





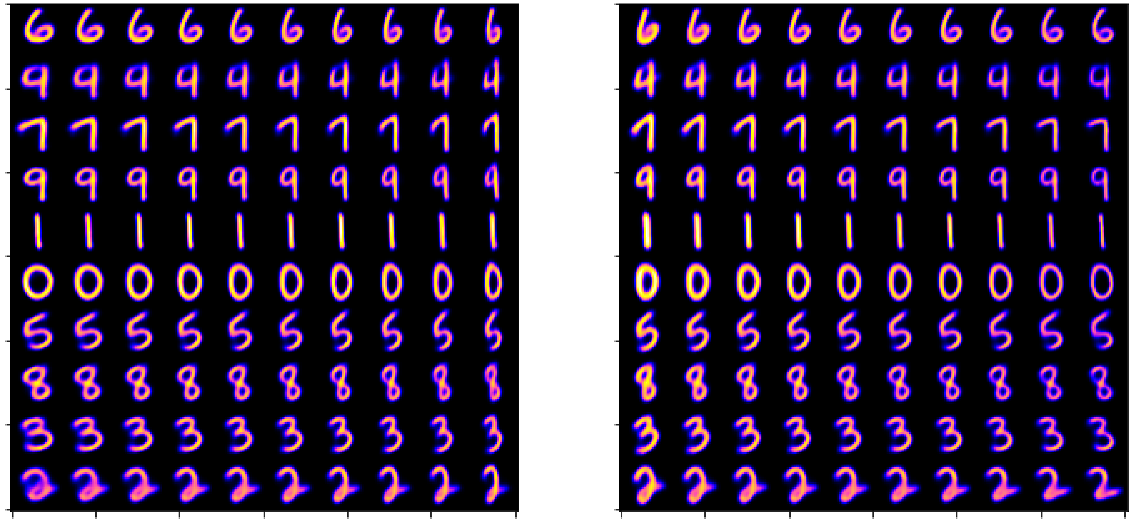


Latent representations

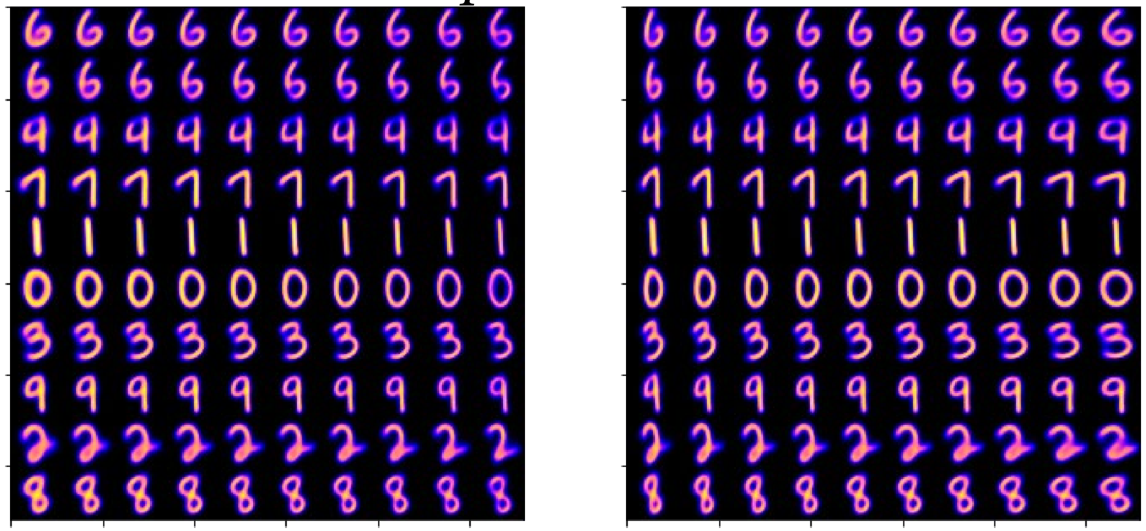


Ensemble jVAE

Predictions from different ensemble models



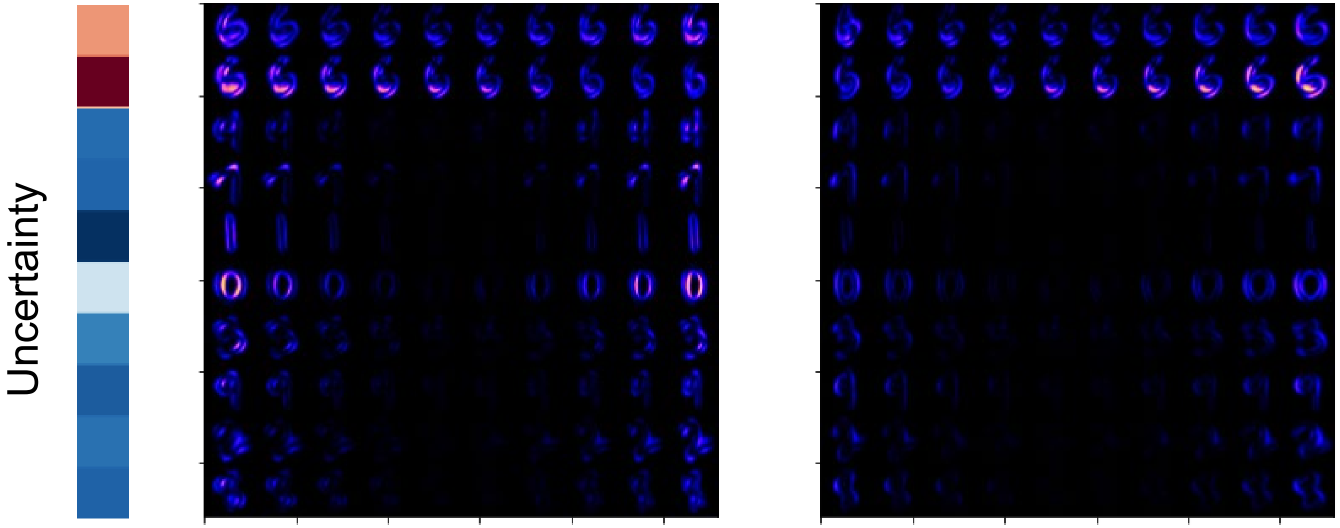
Mean prediction



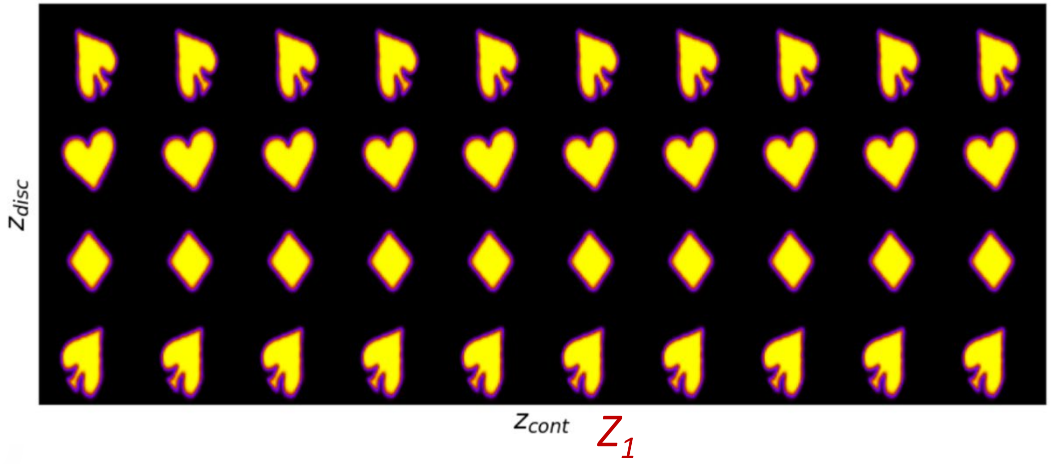
Baseline: 10 epochs
Ensemble models: 8

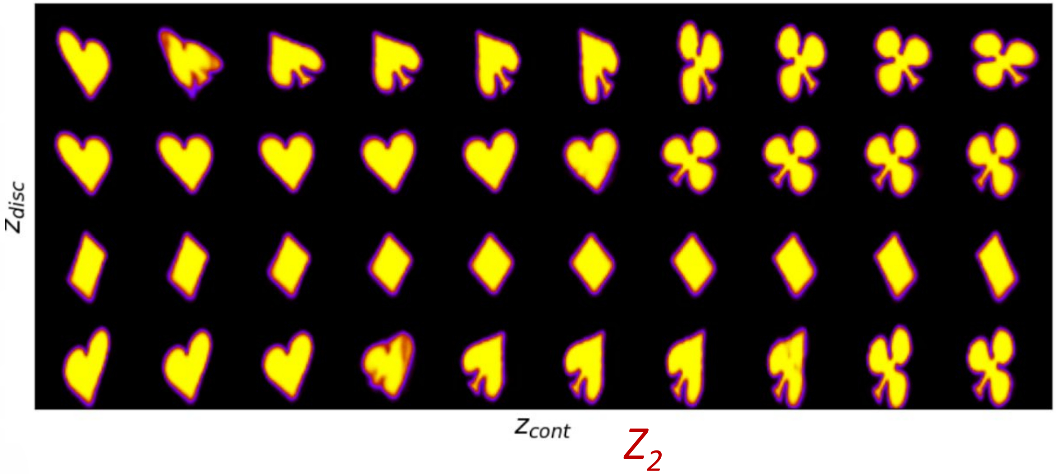
- The unstable classes show the largest "uncertainty"
- Indication of the quality of separation and/or a guide for selection of the number of classes

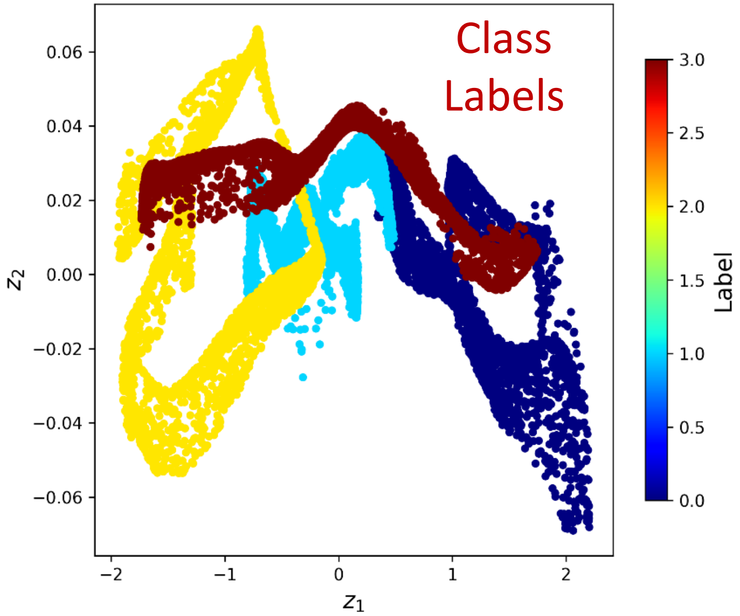
Dispersion in predictions ('uncertainty')

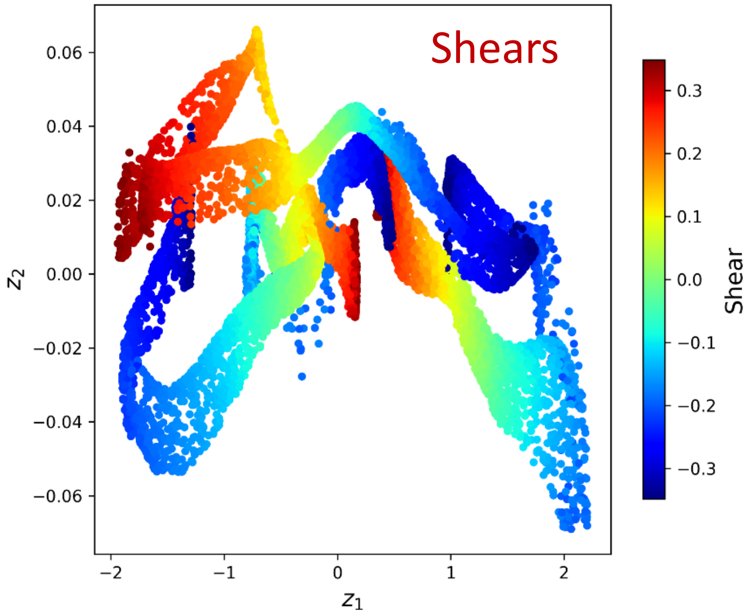


jVAE on cards

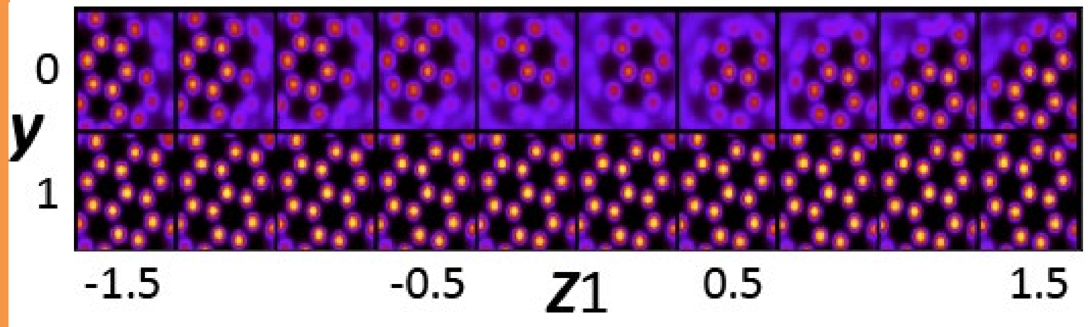


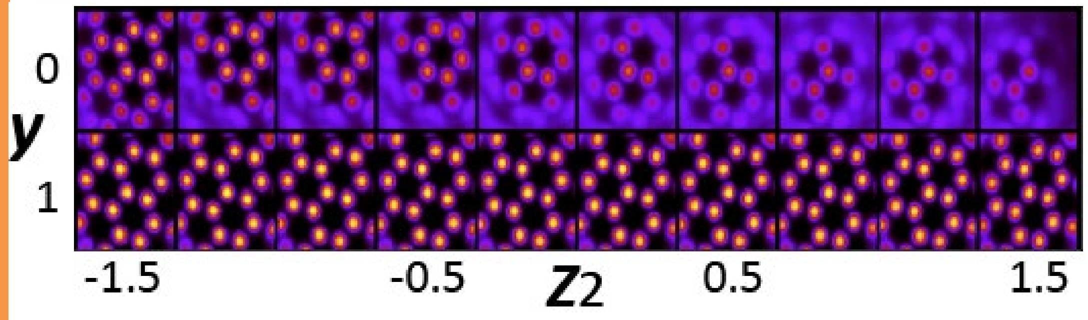




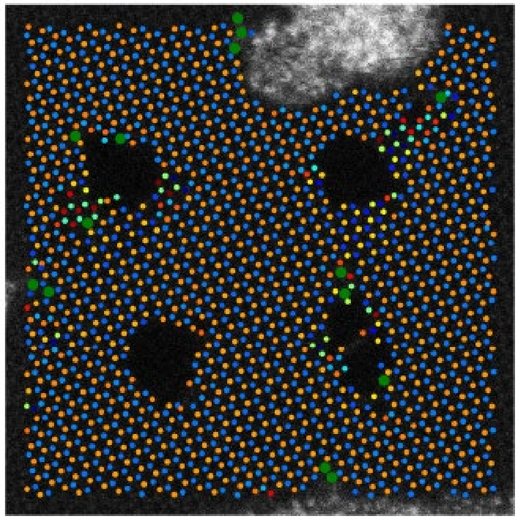


jVAE on graphene

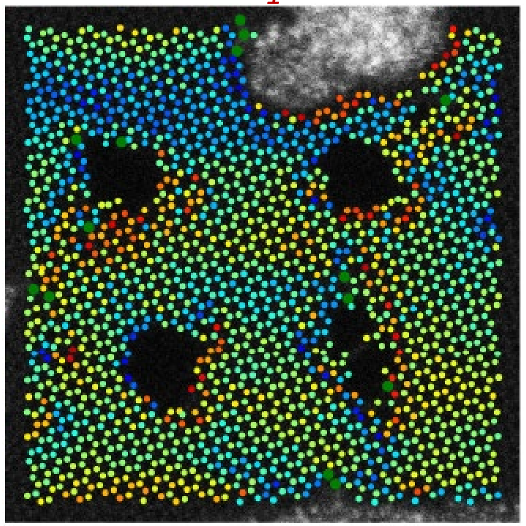




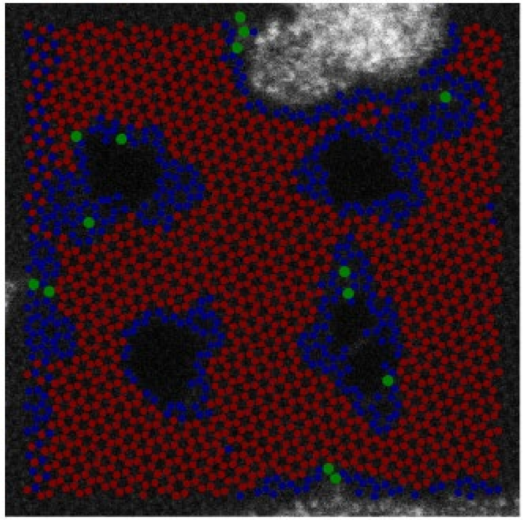
Angle encoding



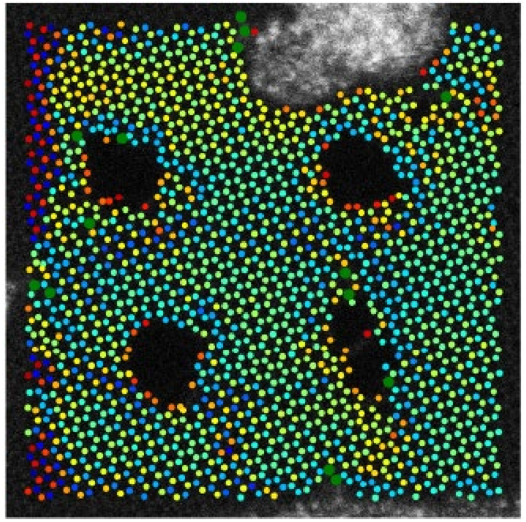
z_1



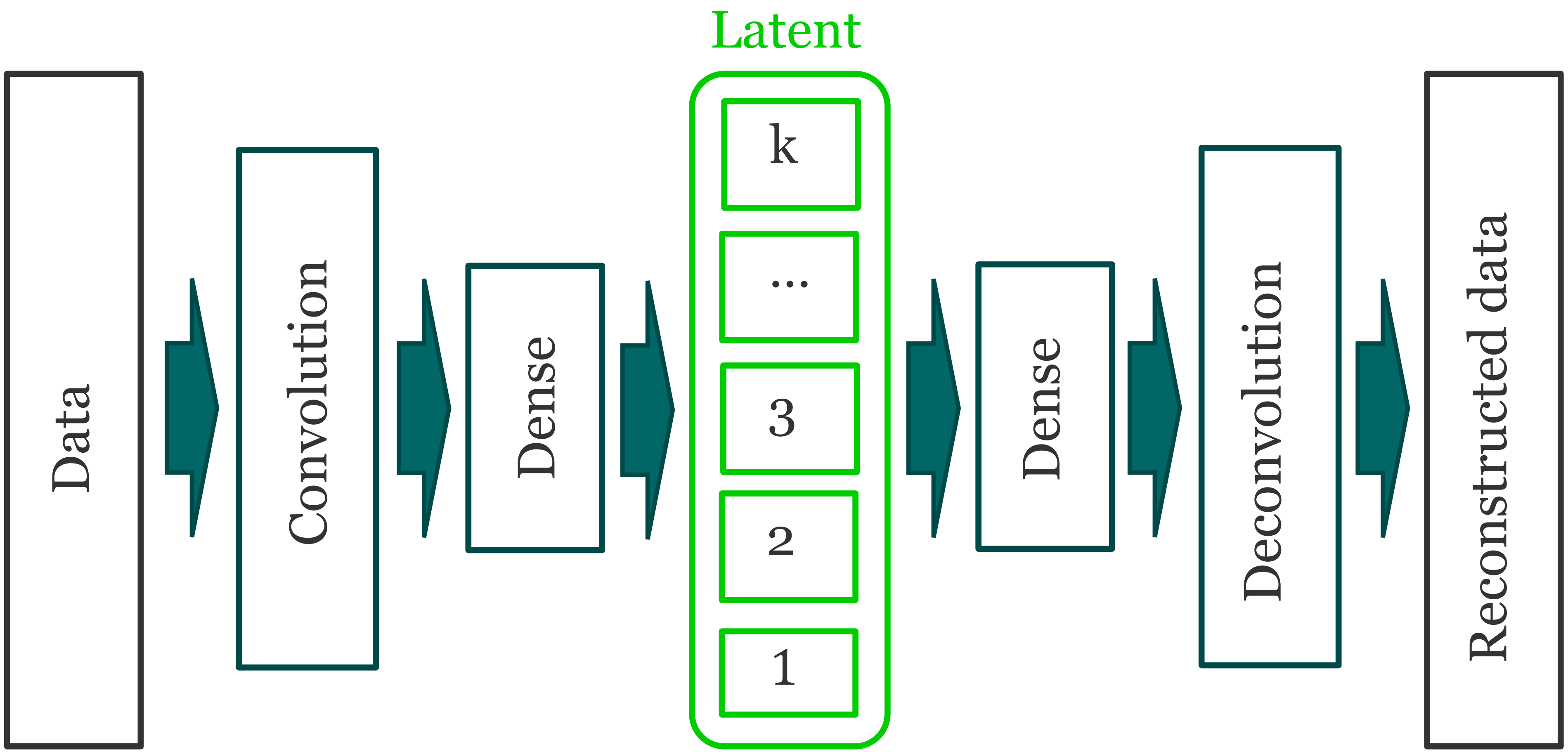
Class labels



z_2



Autoencoders



Loss: (some form of) reconstruction loss

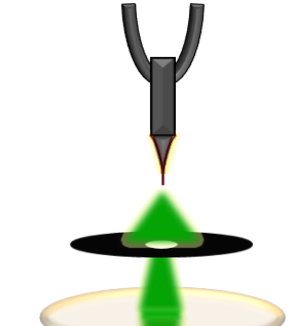
Do we have to encode and decode the same type of objects?

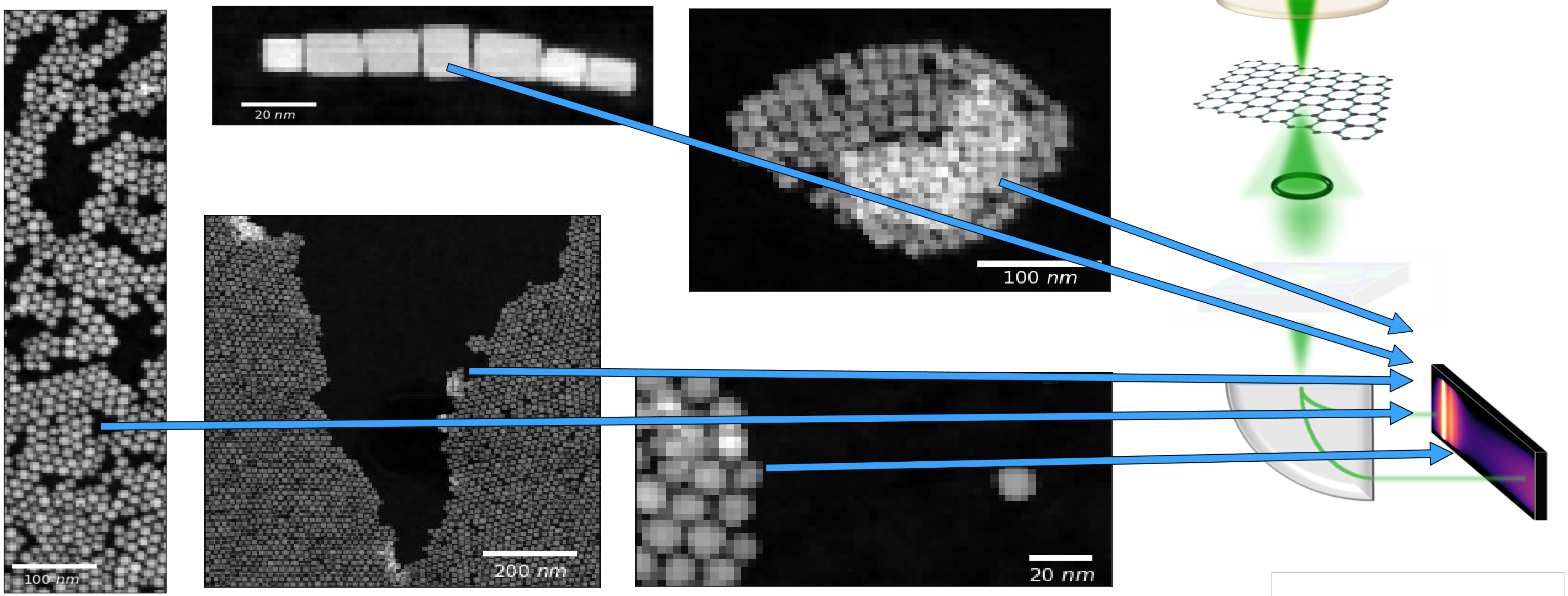
Examples of structure property relationships

- Molecular structure – optical spectroscopy
- Atomic configurations in catalysts – catalytic activity
- Protein sequence – geometry
- Photonic structure – optical adsorption
- Materials microstructure – dielectric/conductive properties
- Composite structure – electrochemical properties
- Antenna shape – emission characteristics
- … and many more

Practical Use Case: EELS

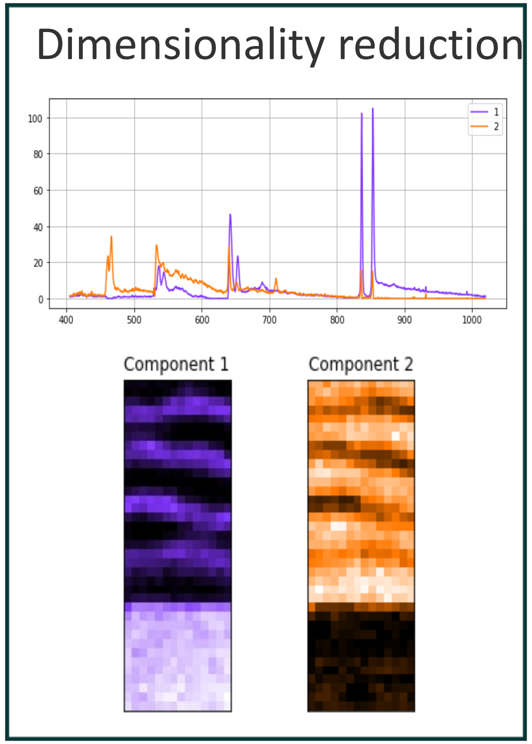
- Self-assembled monolayer of **metal oxide nanoparticles** (**F**,**Sn** co-doped indium oxide)
- **Sn** tunes the plasmon resonance by supply of additional e^- (F concentration fixed)
- Variety of geometric configurations also present

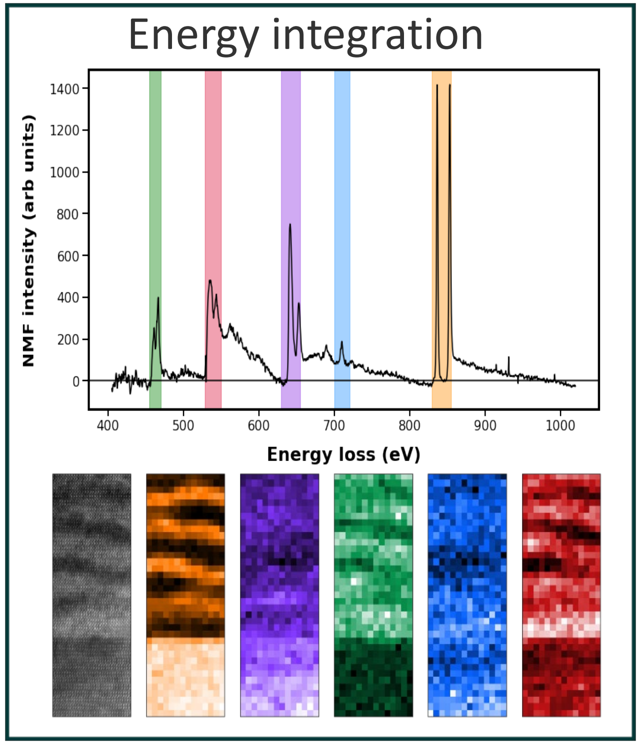


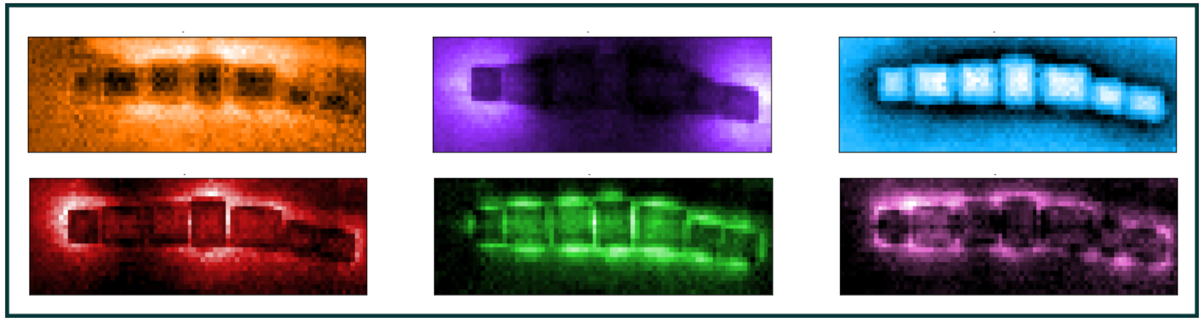


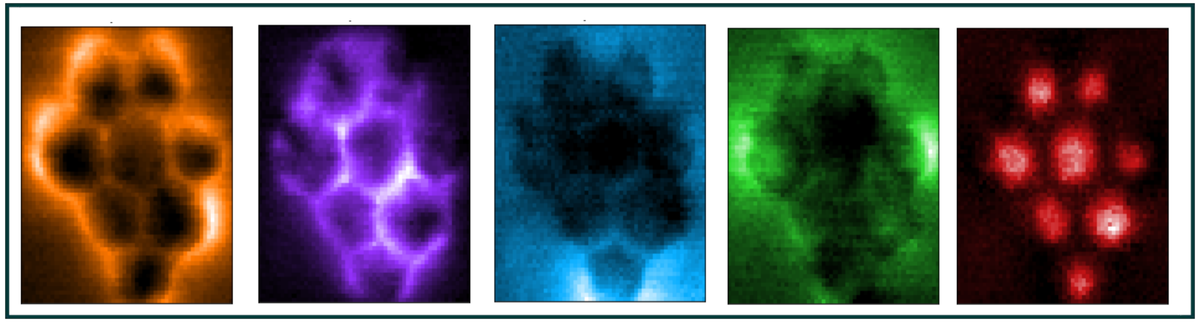
EELS

- Can consider to be another signal in the form of a **1D spectrum**
- Collect EEL spectrum in an (x,y) grid: EELS imaging
- To better visualize / understand these 3D signals, can integrate specific spectral bands, or dimensionally reduce (PCA, NMF) them
- Applications in plasmonics & nanophotonics

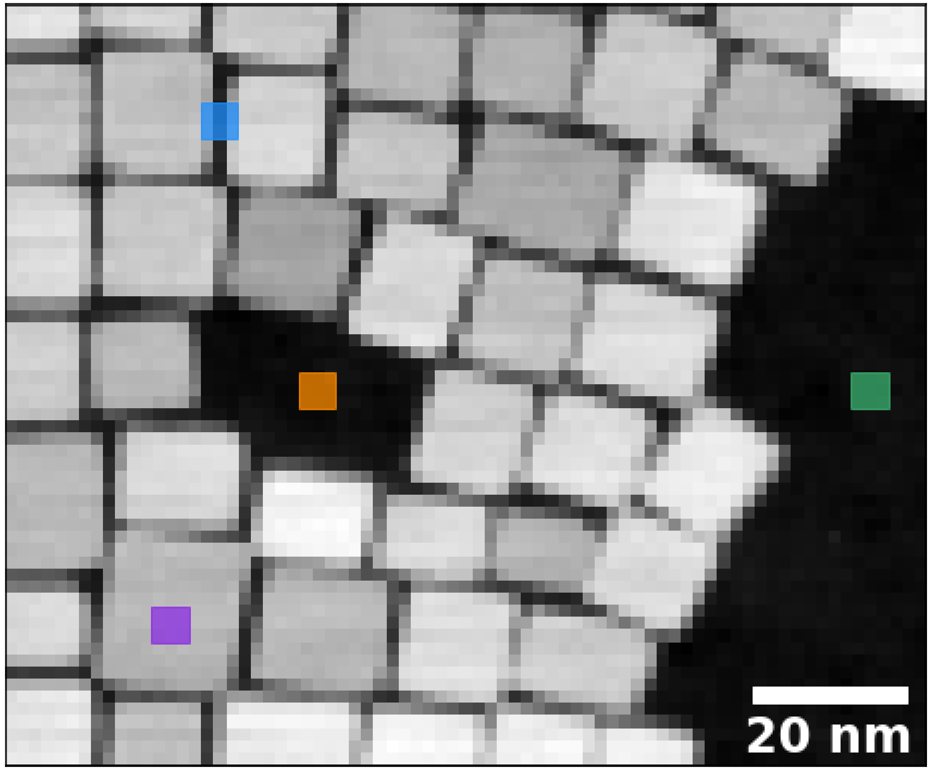


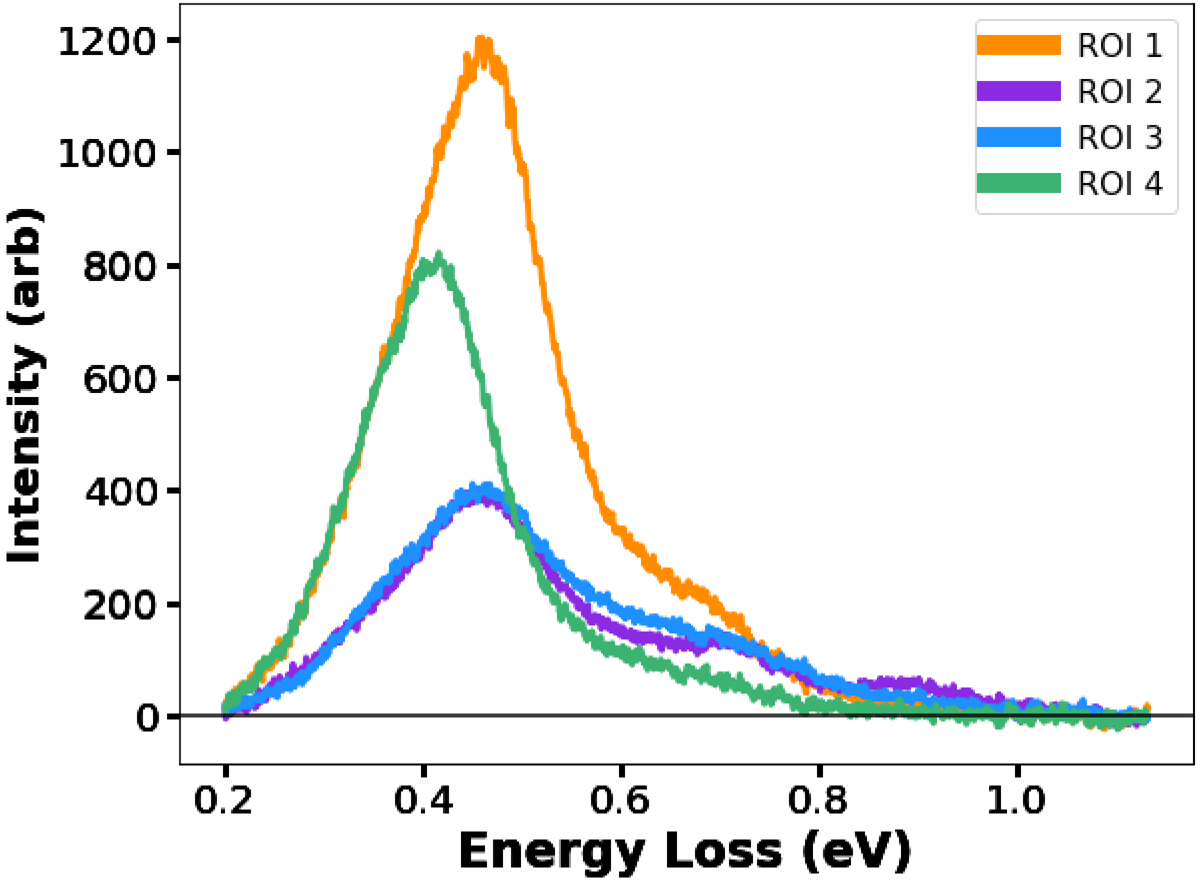






But what about the structure-property relations?

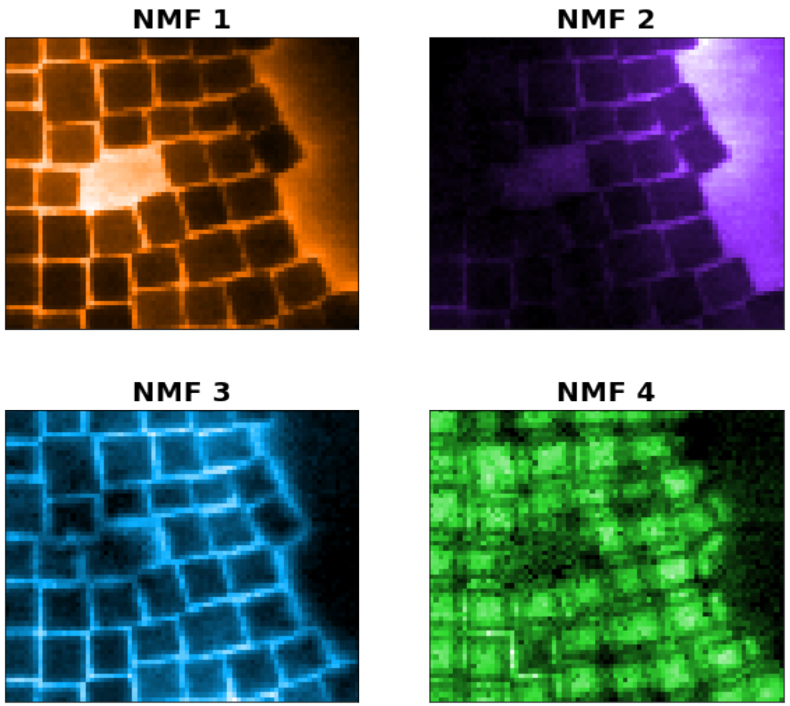


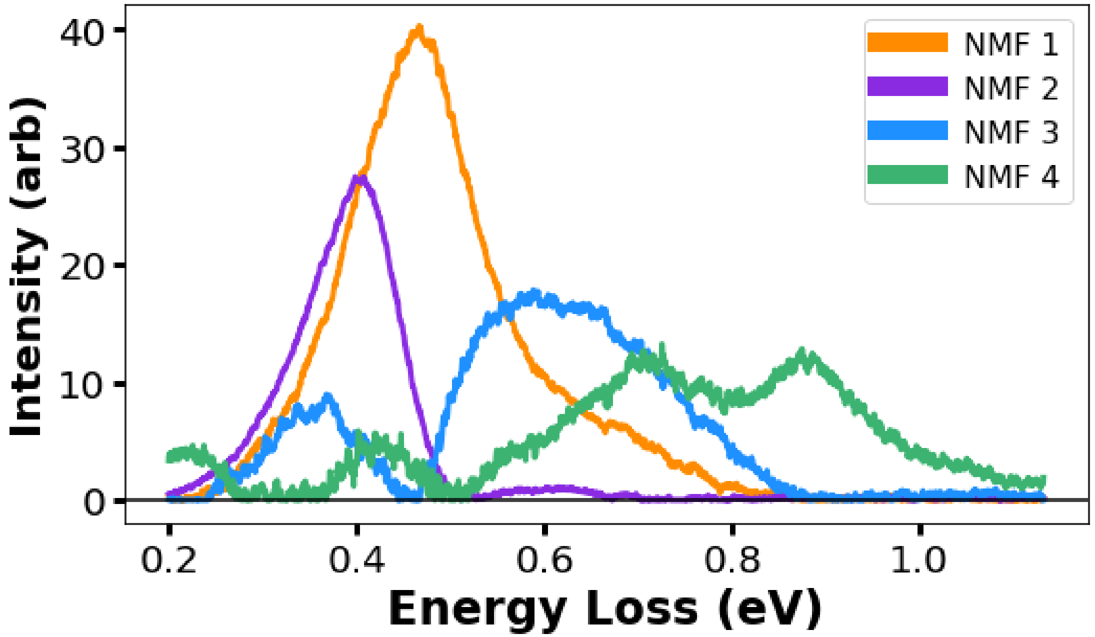


By inspection, we can note some characteristic aspects of spectra from specific types of geometries. However:

- How can we prove it and quantify this relationship?
- How universal is it for similar structures?
- Can we discover structures that will have the properties that we want?

Exploratory data analysis

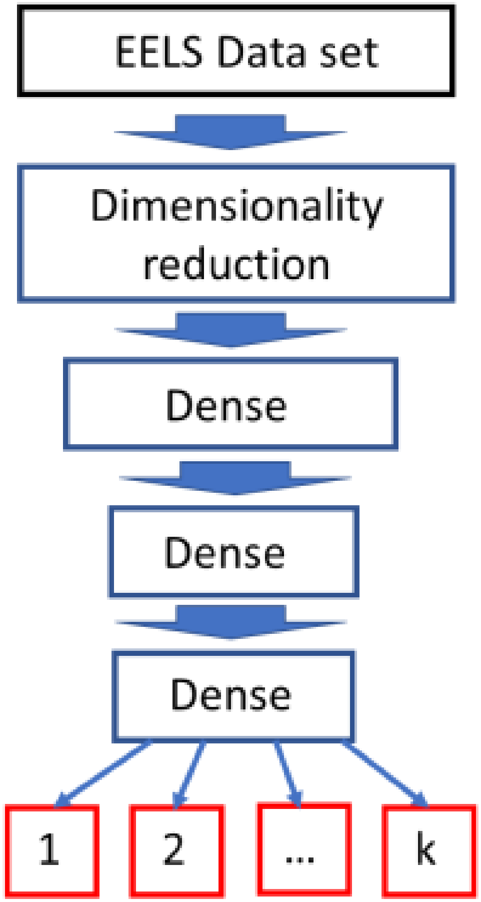


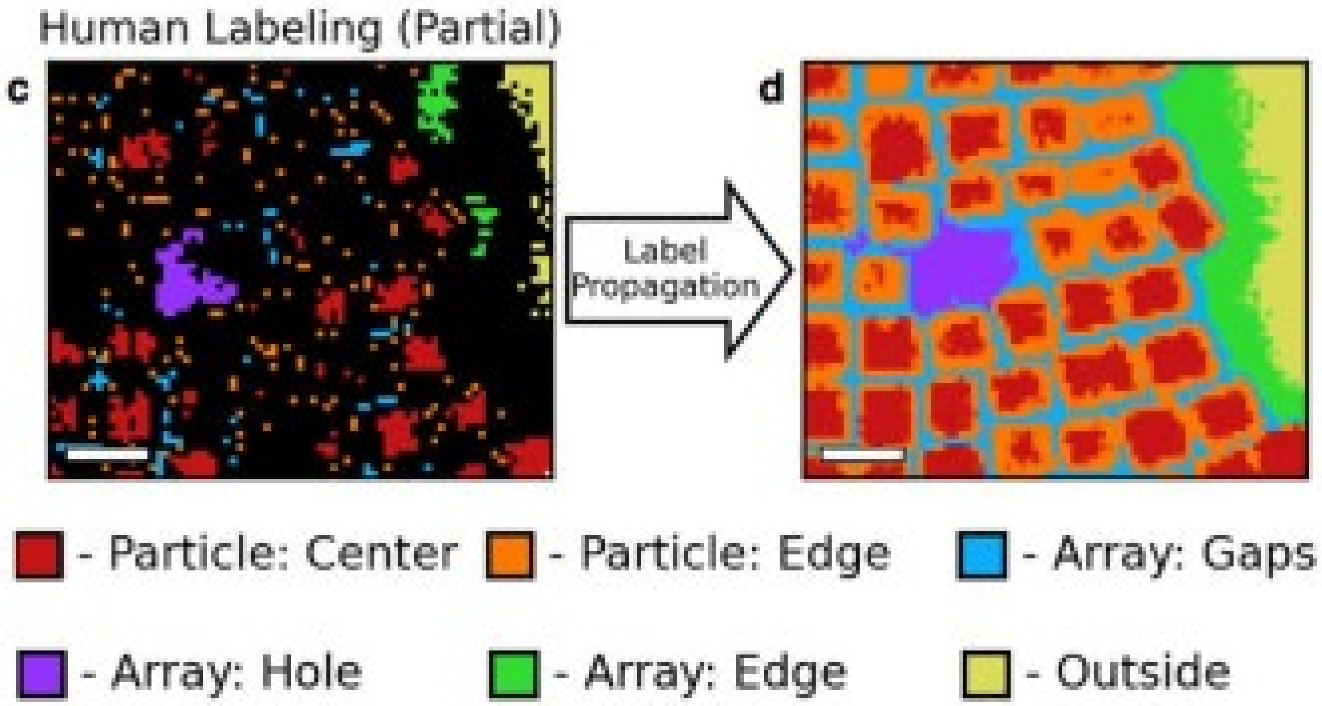


- Great way to explore system
- Visualization of hyperspectral data

- Multiple modes per pixel
- Non-physical extraction
- No relationship with geometry established

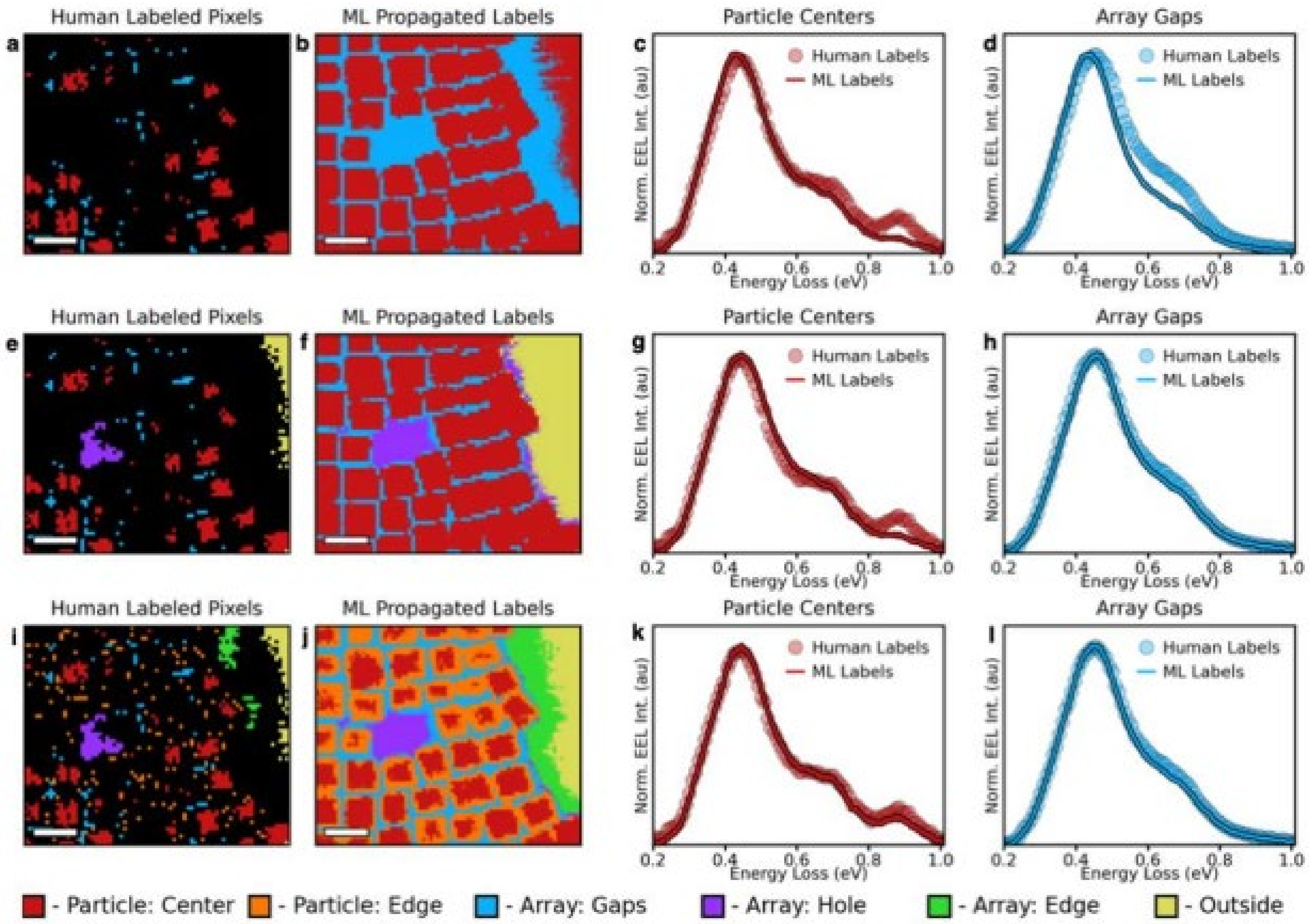
Strategy 1: Labelling





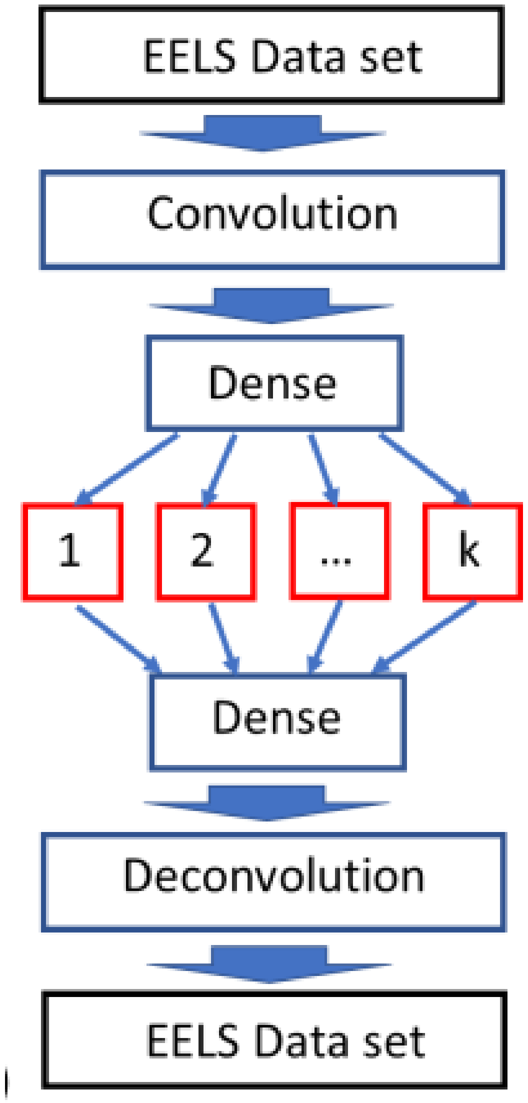
- Define attributes by semi-manual inspection
- Create training data set
- Train the classifier
- Apply the classifier to the remainder of the data set

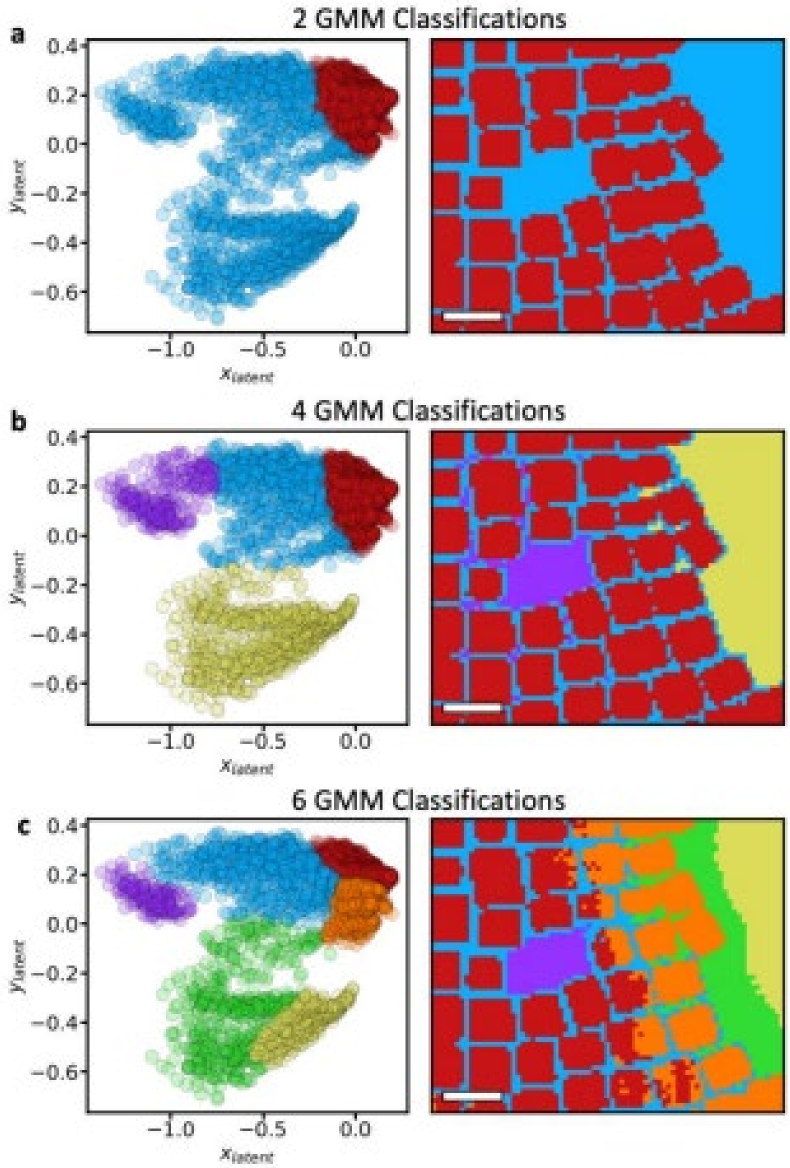
Strategy 1: Labelling



- Problem: human labels can be ambiguous

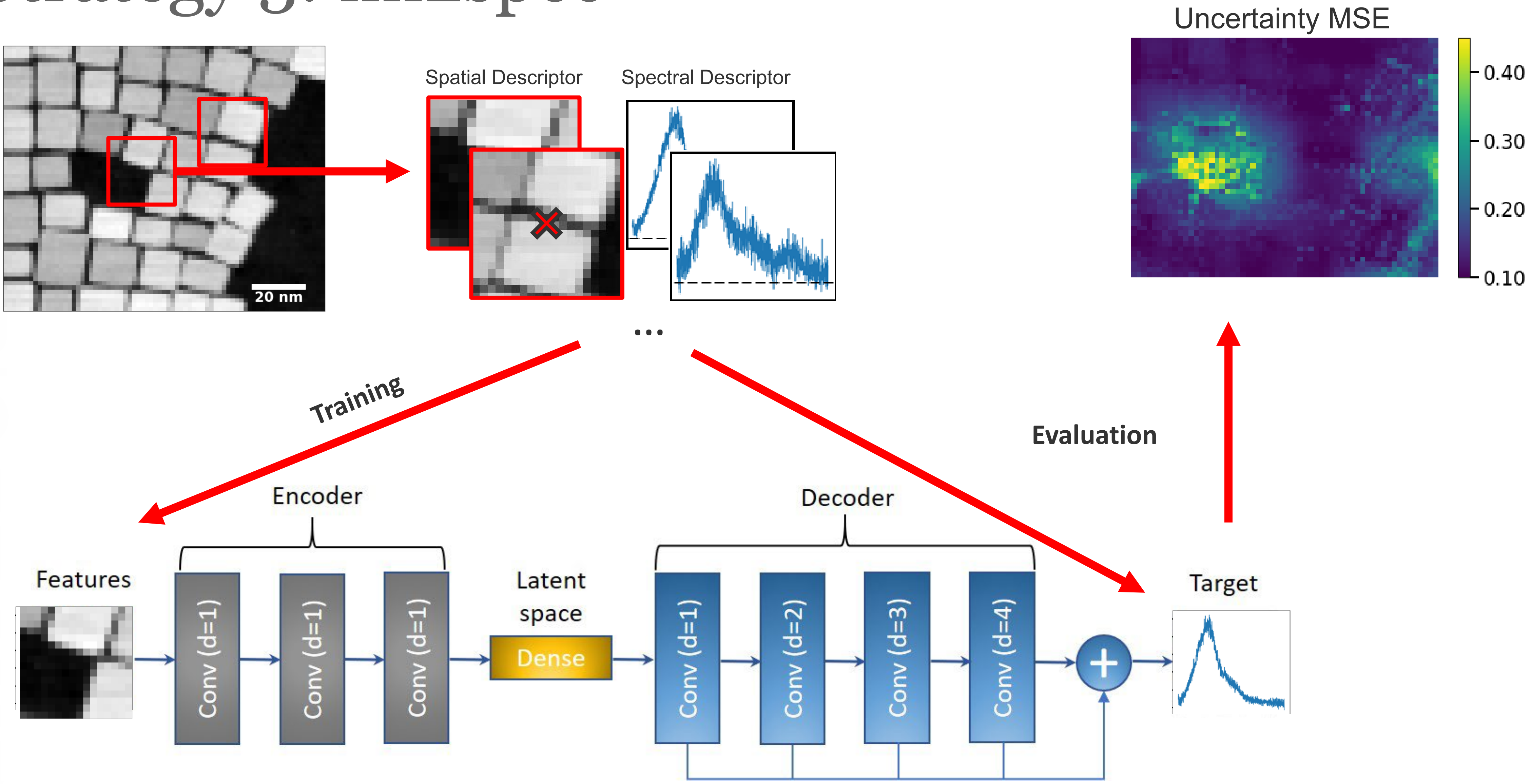
Strategy 2: AE Labels





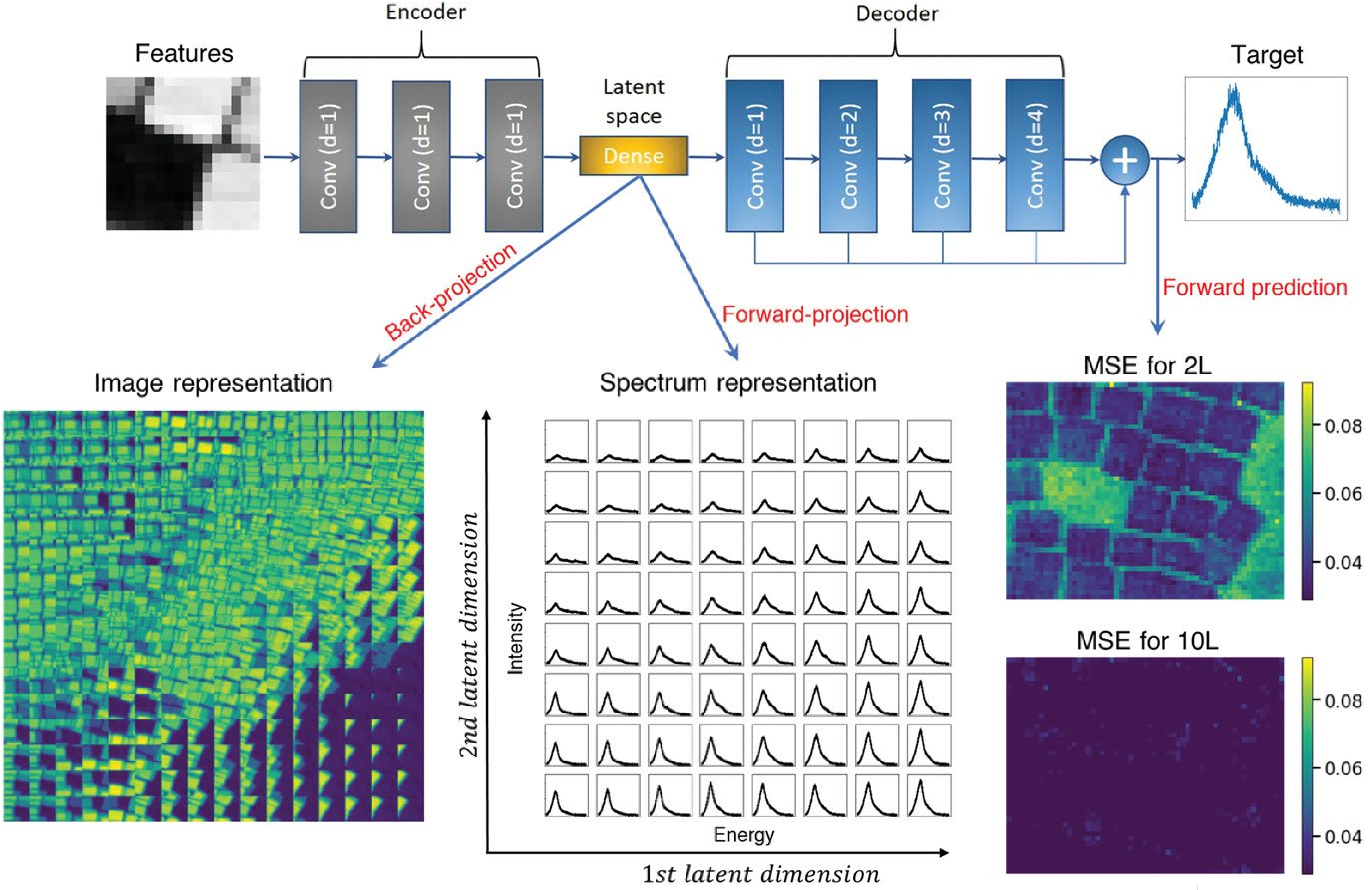
- Unsupervised learning on spectral data: we make conclusions based on maps

Strategy 3: im2spec

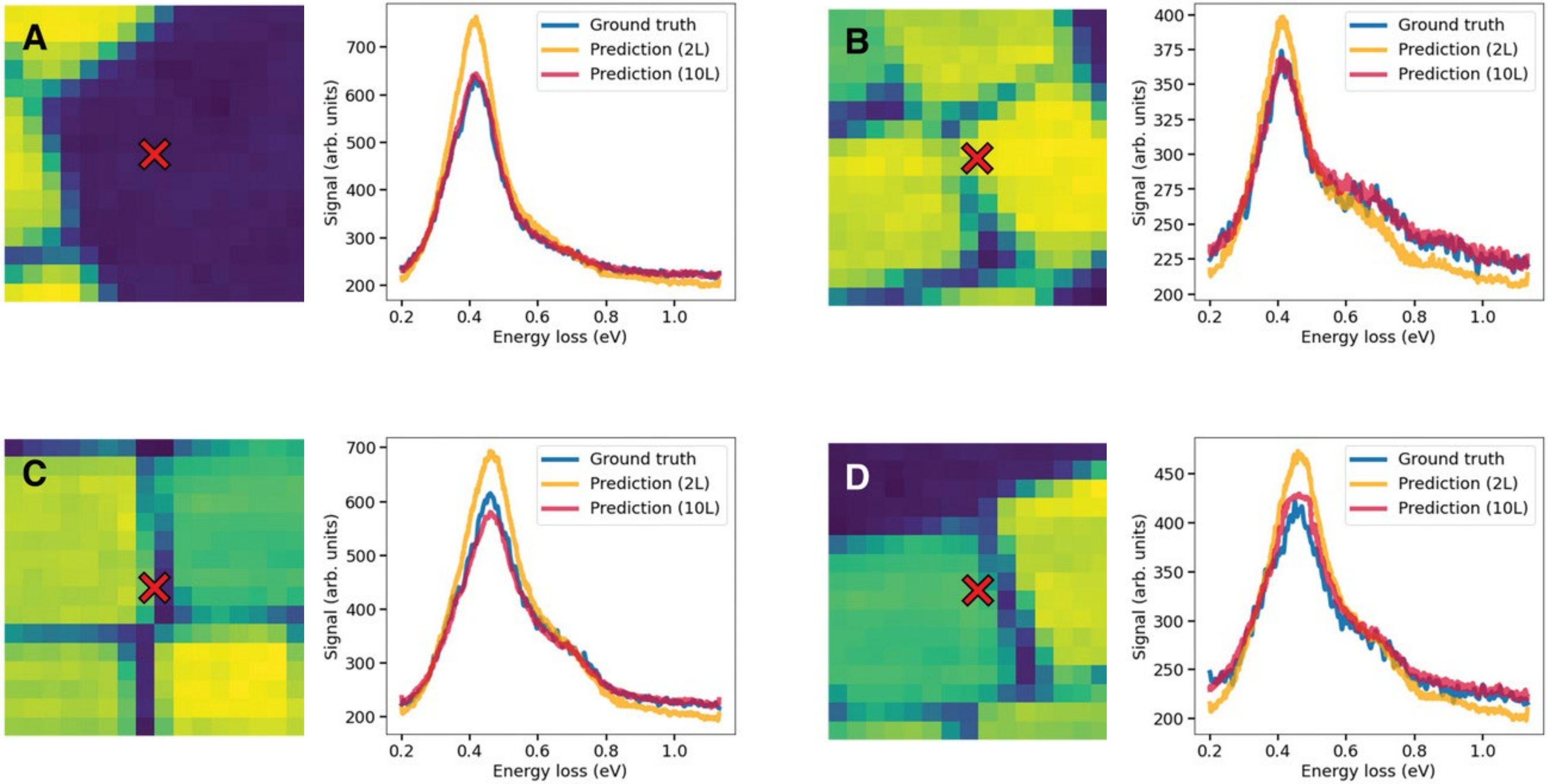


What if we create a network that encodes structure, and decodes spectra?

Im2spec: latent space visualization

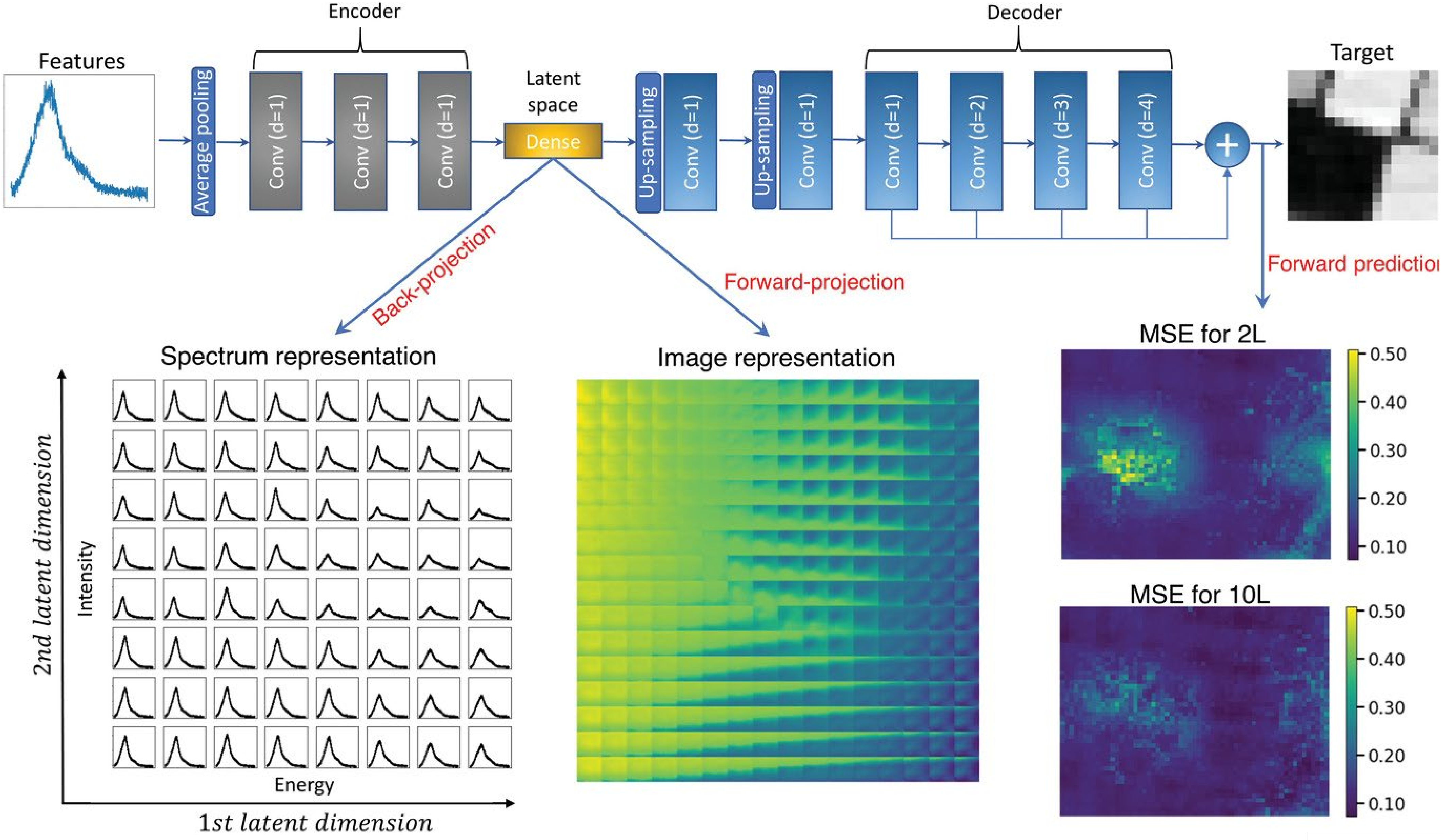


Im2spec prediction



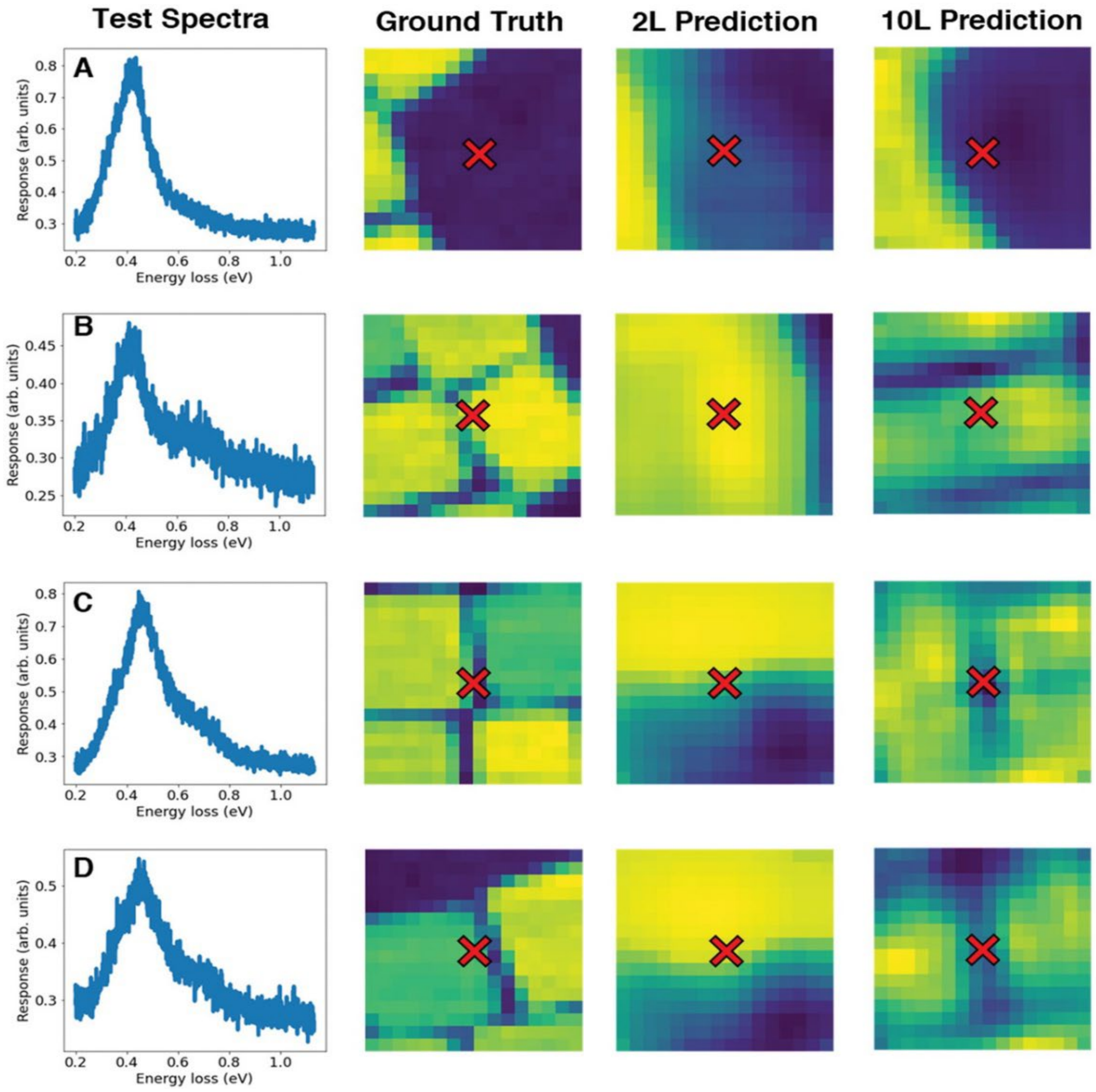
- After training, predict the spectral response of a geometric arrangement that the network has **never encountered**
- **Library** of geometric-plasmonic relationships
- Can be used for solution of inverse design in nanophotonics and other fields

Spec2im



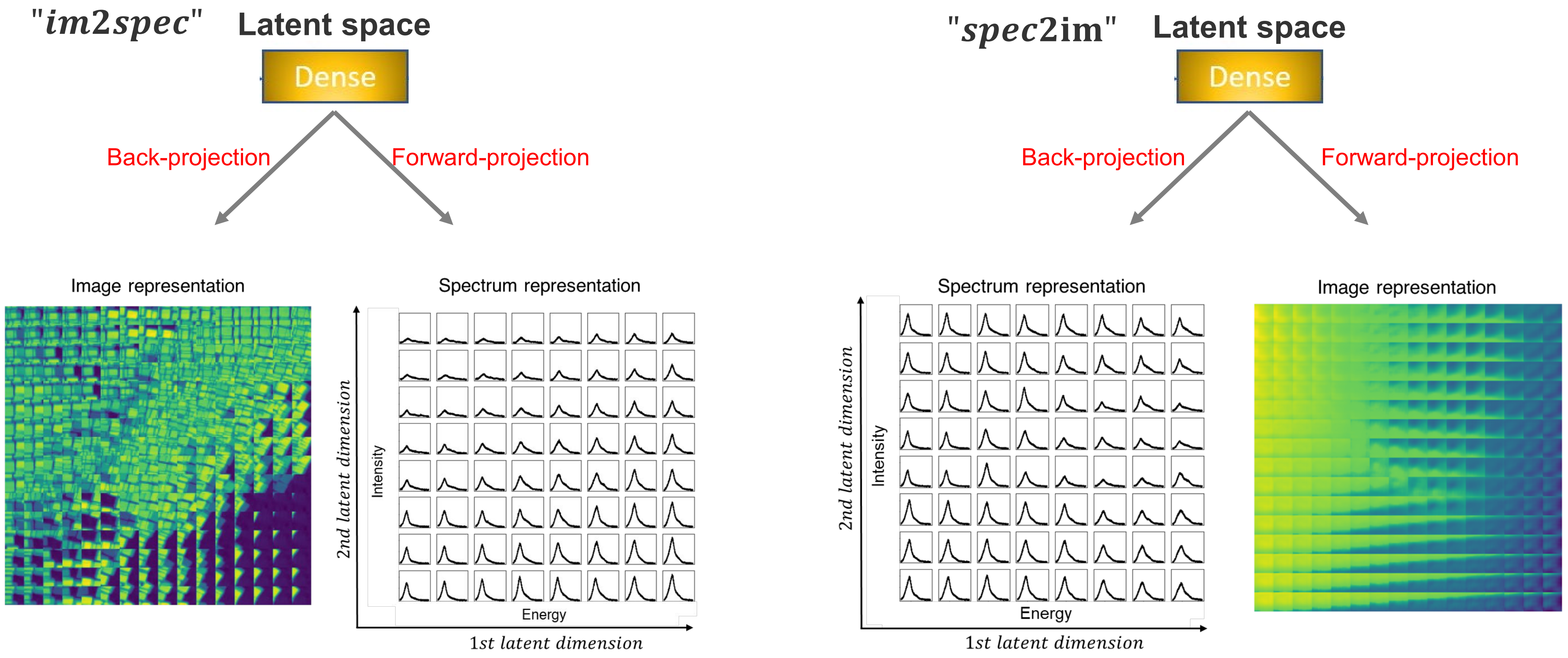
Encoder
Decoder
Features
Average pooling
Conv (d=1)
Conv (d=1)
Conv (d=1)
Latent space
Dense
Up-sampling
Conv (d=1)
Up-sampling
Conv (d=1)
Conv (d=1)
Conv (d=2)
Conv (d=3)
Conv (d=4)
Target
Back-projection
Forward-projection
Forward predictio
Spectrum representation
Image representation
MSE for 2L
MSE for 10L
0.50
0.40
0.30
0.20
0.10
2nd latent dimension
Intensity
Energy
1st latent dimension

Spec2im predictions



- After training, predict the spectral response of a geometric arrangement that the network has **never encountered**
- **2L** and **10L** refers to number of latent dimensions chosen

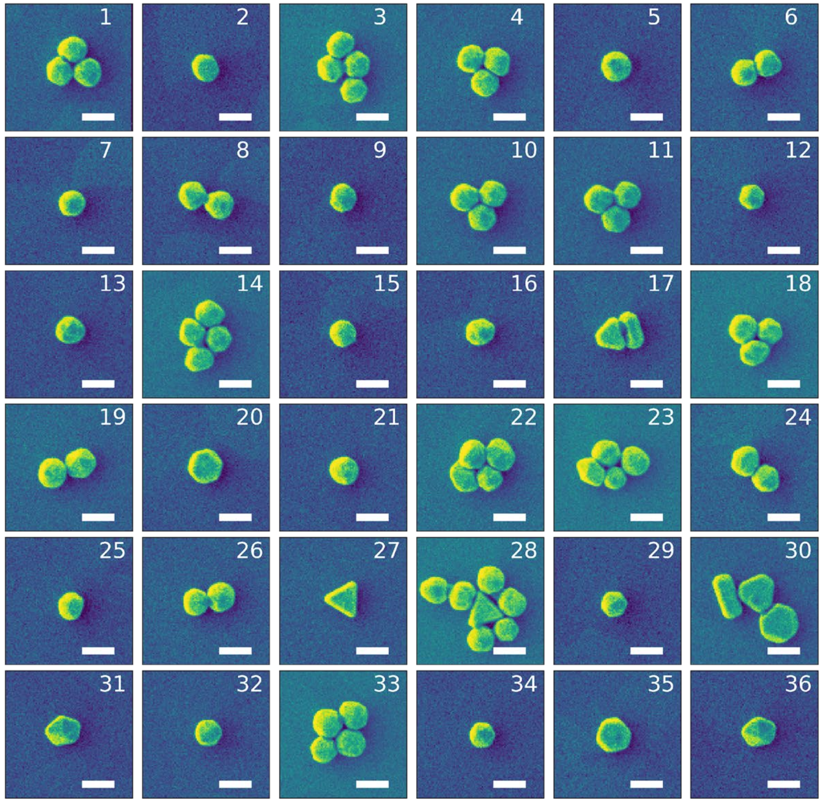
Spec2im and im2spec



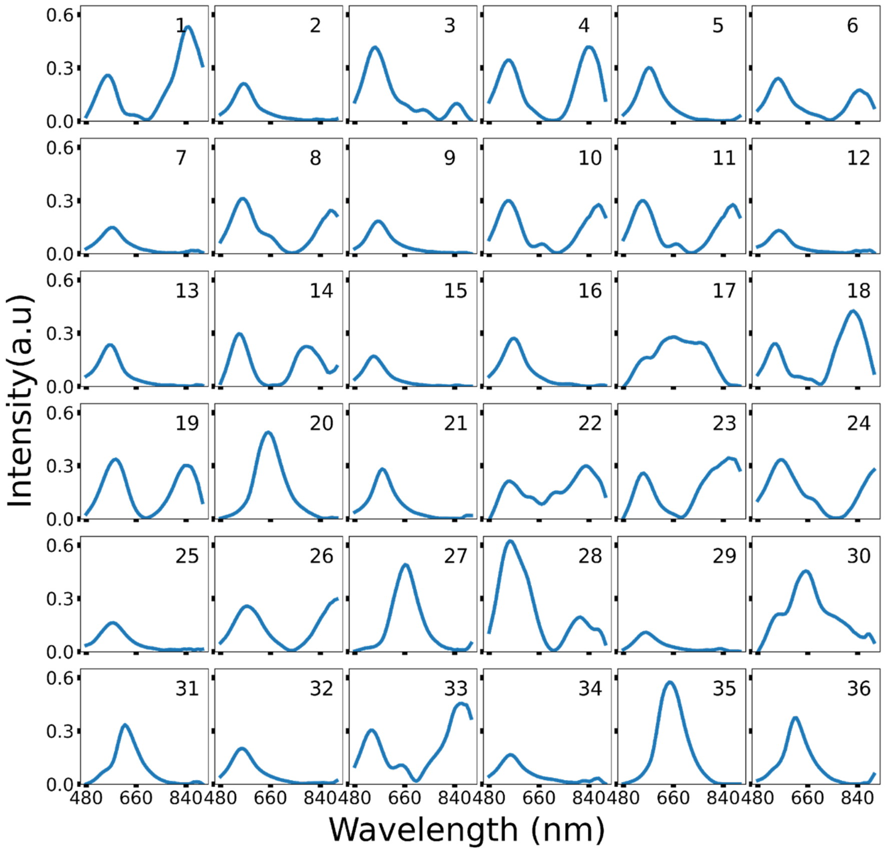
- Problem with im2spec and spec2im: they are generative only wrt. one transformation

Dual VAE: structure-property relationships

SEM images: "Structure Information"

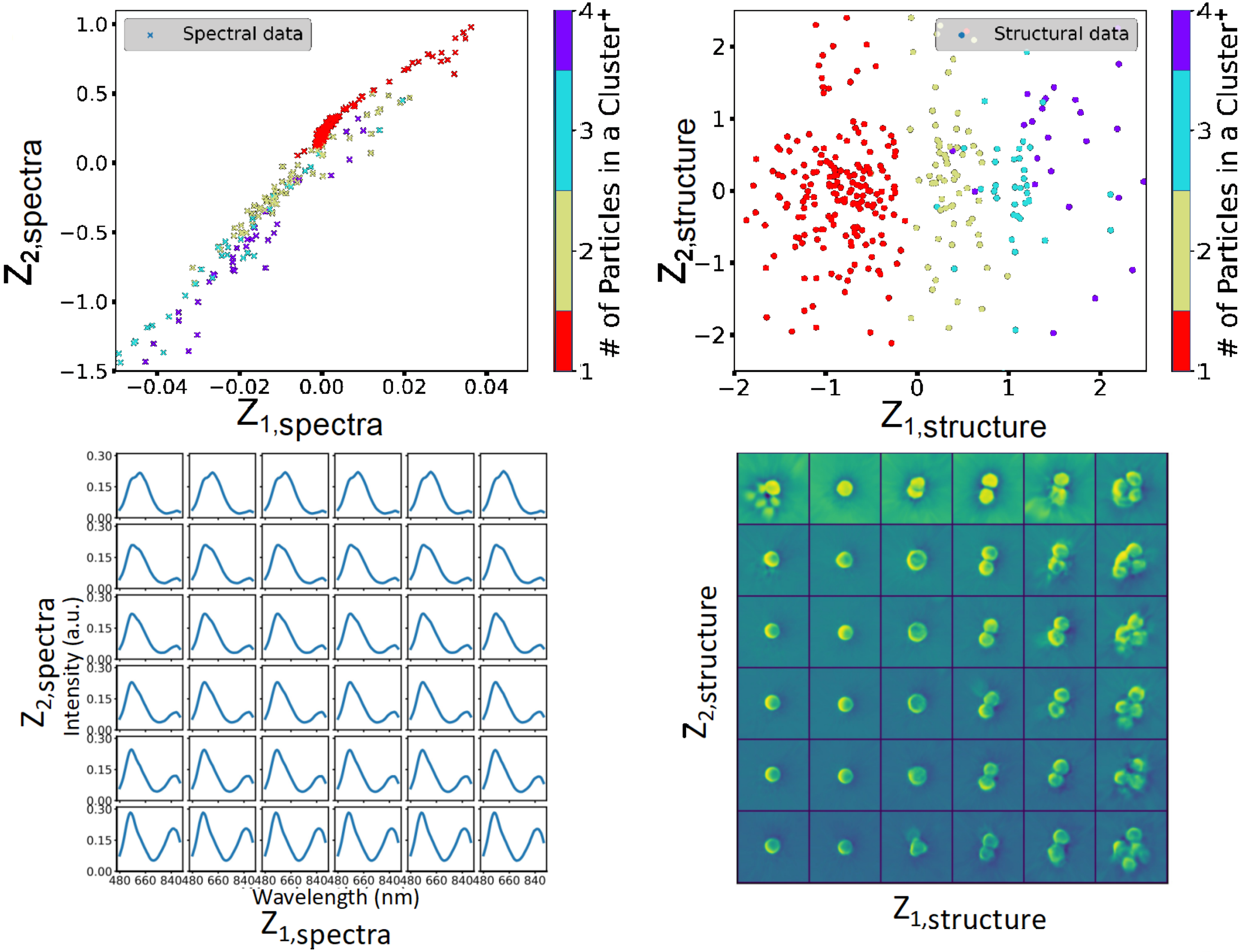


Hyperspectral microscope: "Property Information"

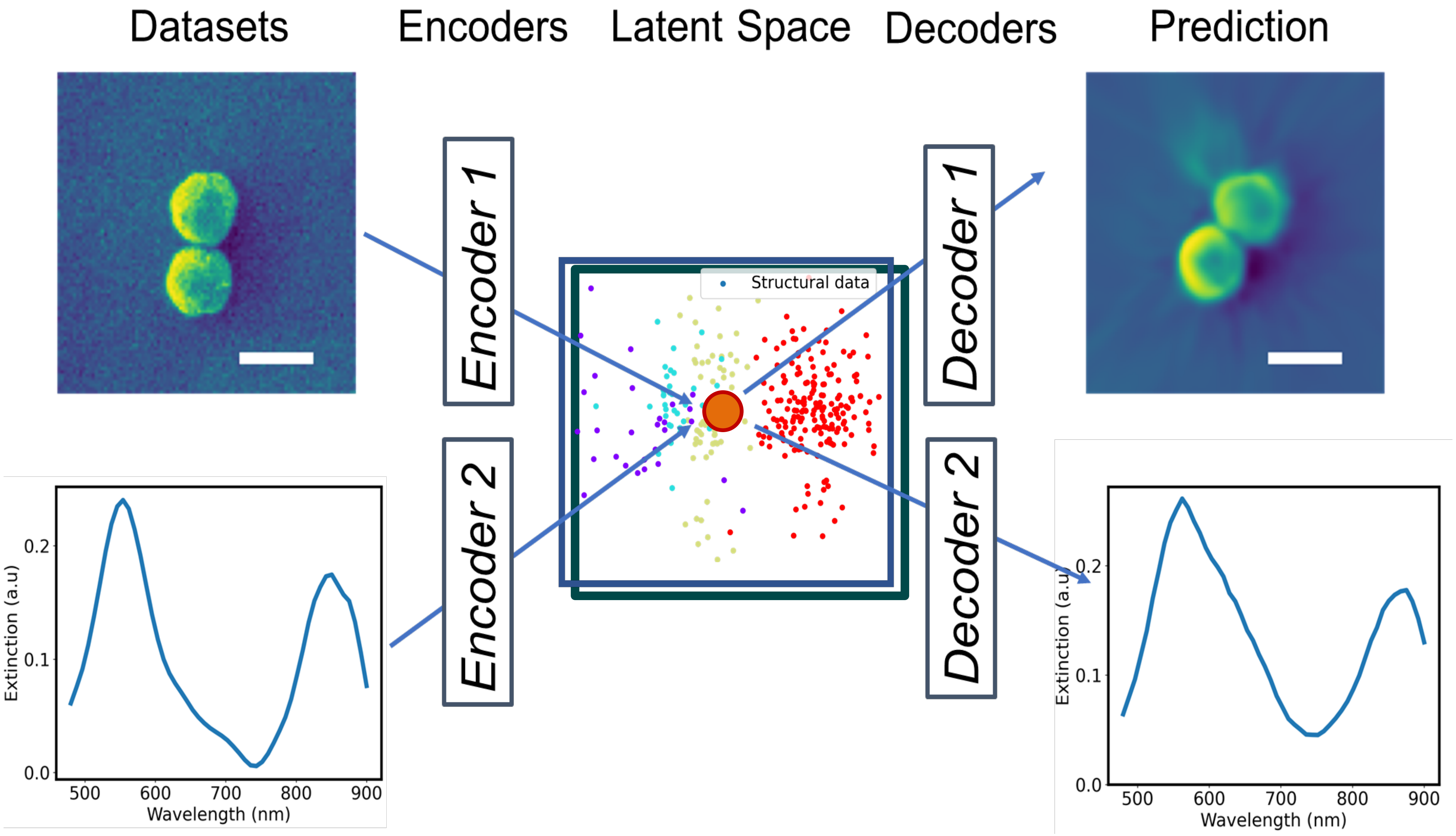


- Far field optical spectroscopy: images and spectra
- Here, we also have simple labels (number of clusters)

Separated VAE



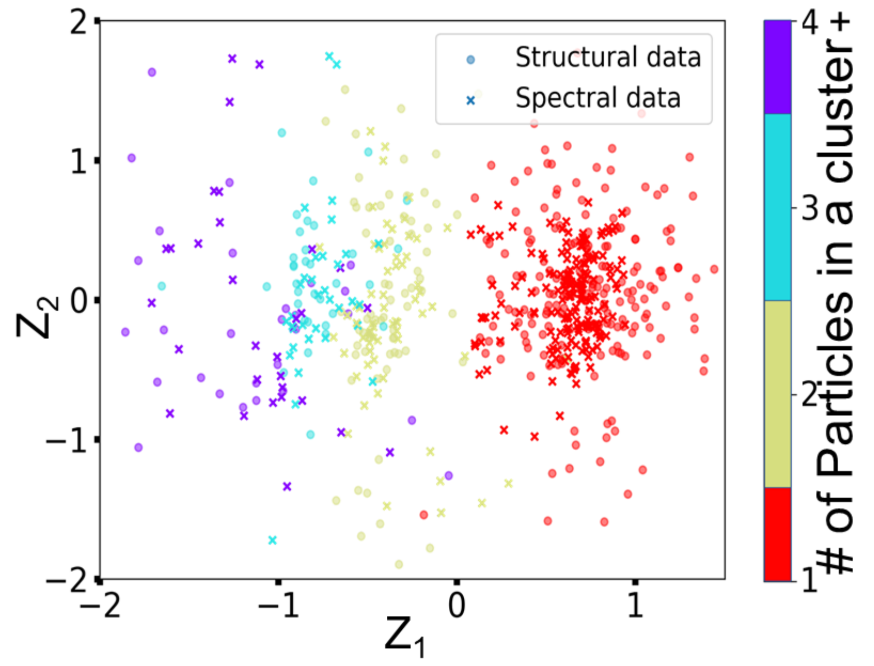
Dual VAE


Datasets
Encoders
Latent Space
Decoders
Prediction
Encoder 1
Encoder 2
Structural data
Decoder 1
Decoder 2
Extinction (a.u)
0.2
0.1
0.0
500
600
700
800
900
Wavelength (nm)

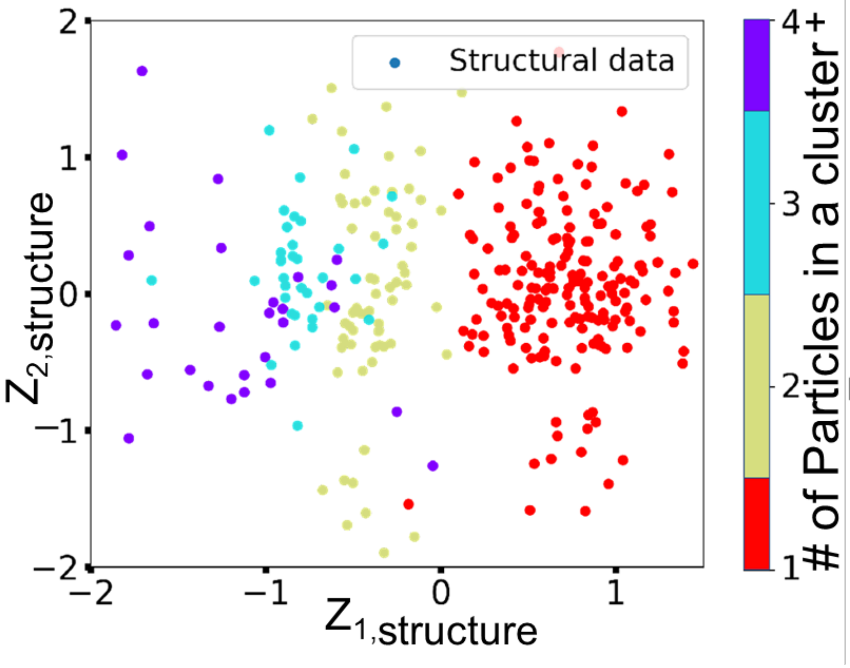
Dual VAE: Latent Distributions

Latent Space Representation

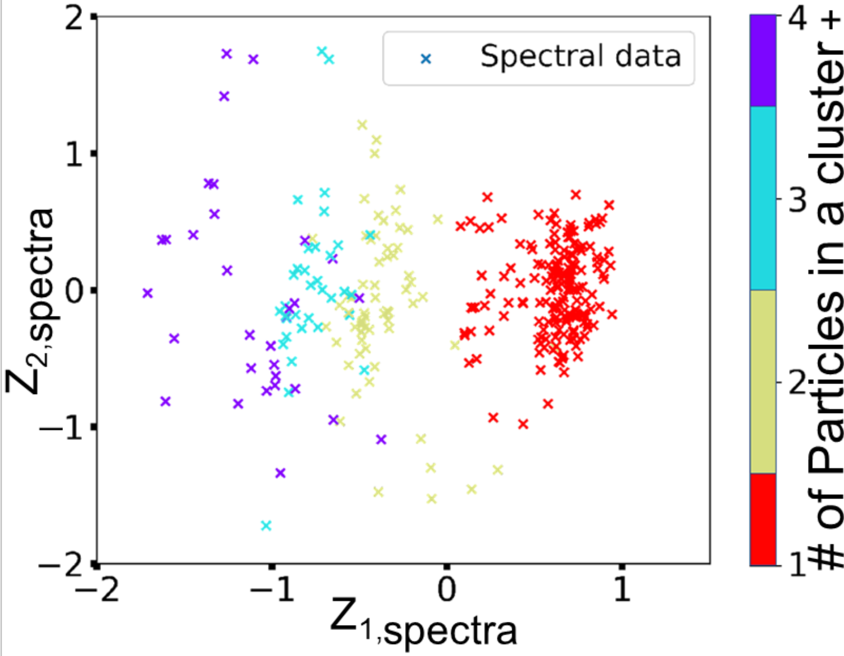
"Dual Latent Space"



"Structure Information"



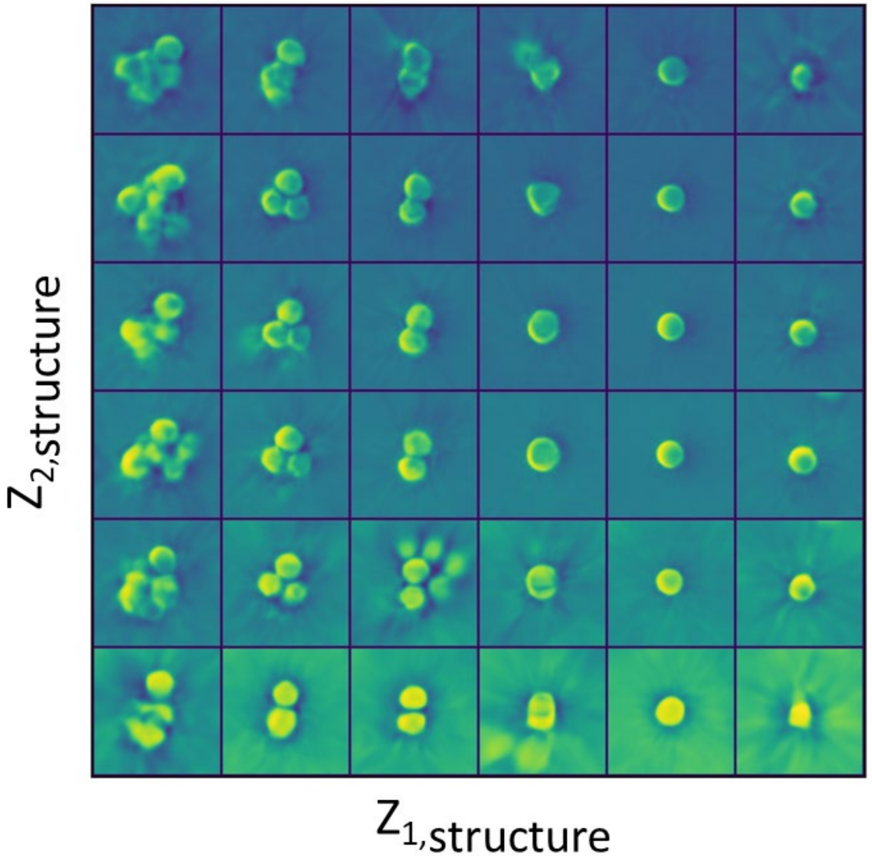
"Property Information"

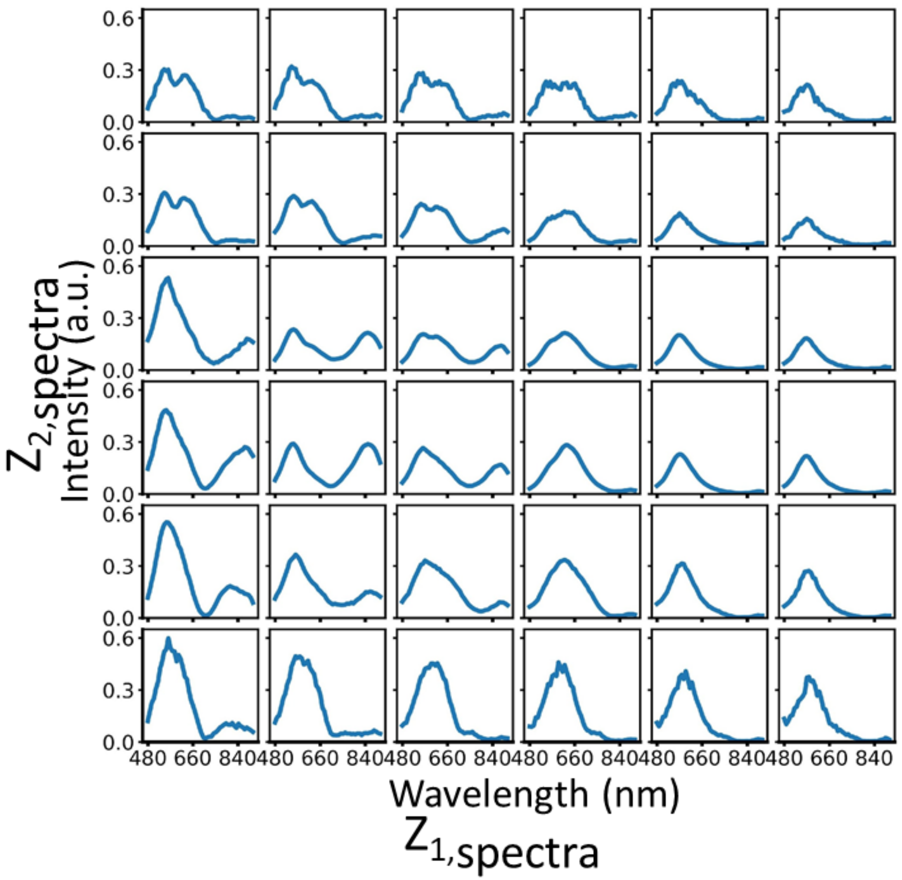


Dual VAE: Latent Representations

Manifold Representation

Changing in particle shape





Changing in spectrum shape

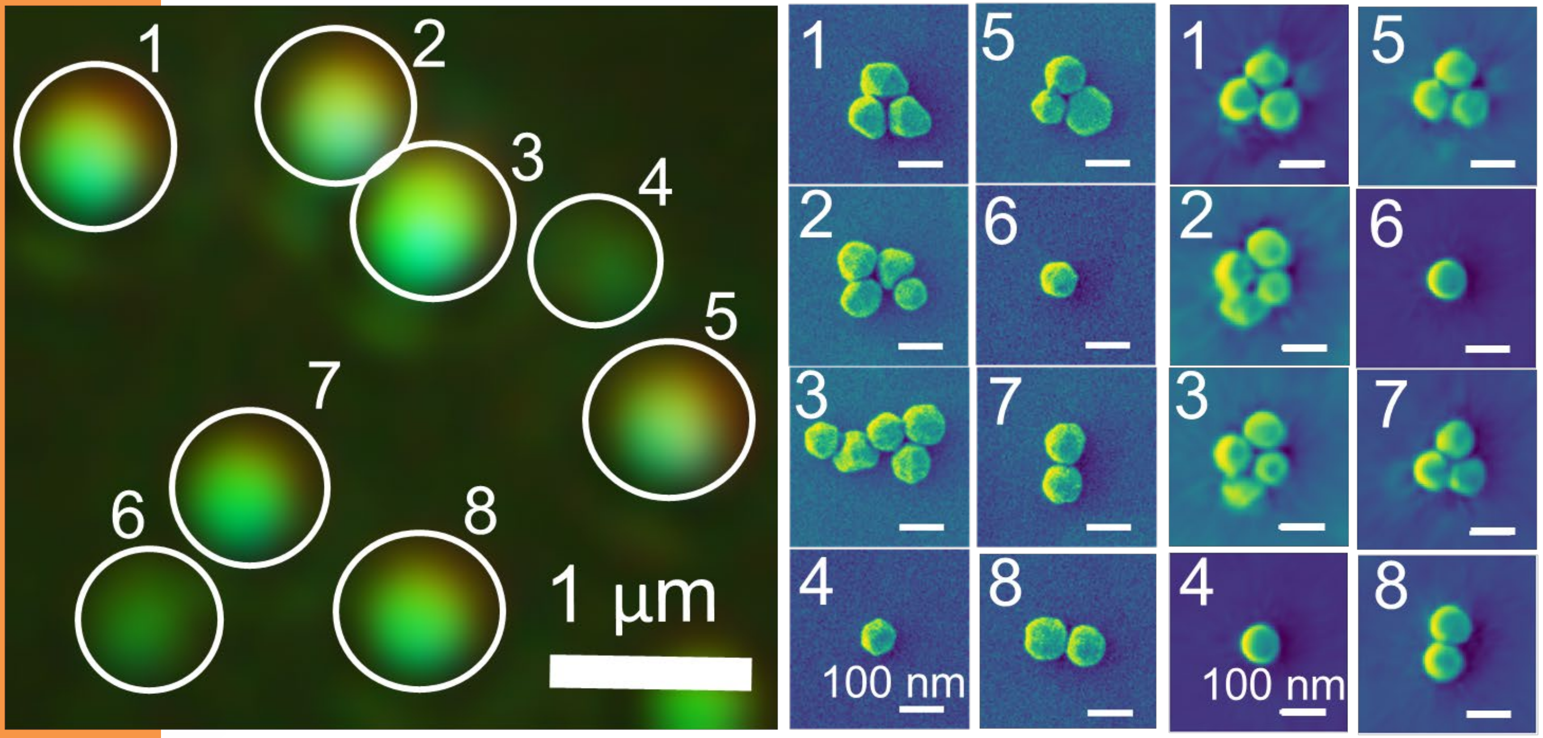
increasing in particle number in a cluster

increasing in the peak intensity

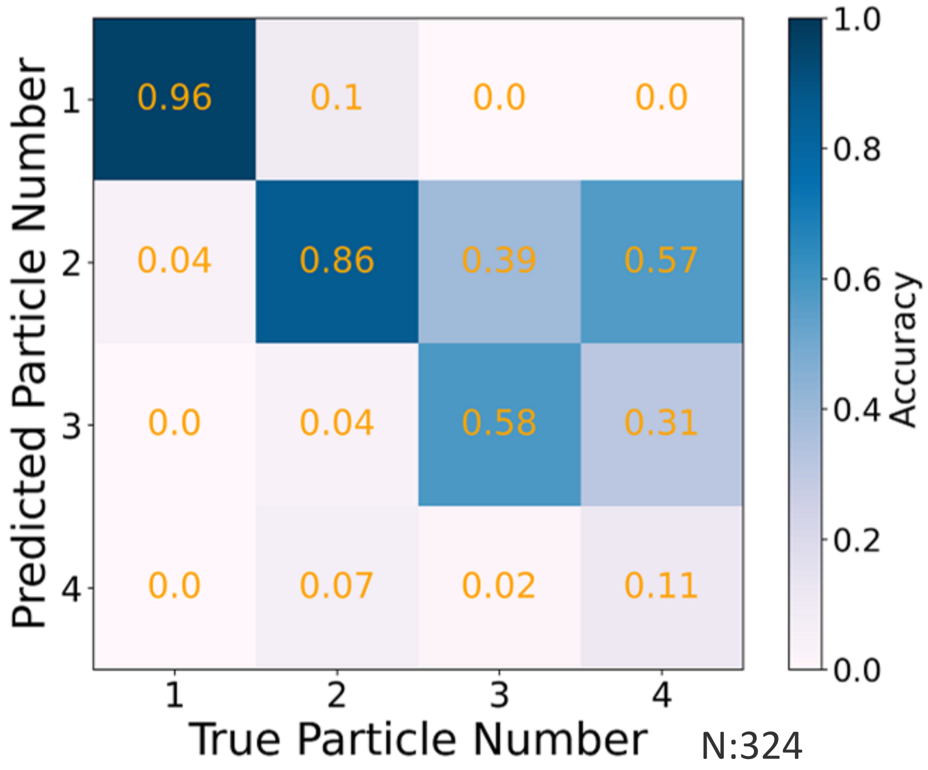
Dual VAE: Predictions

Example

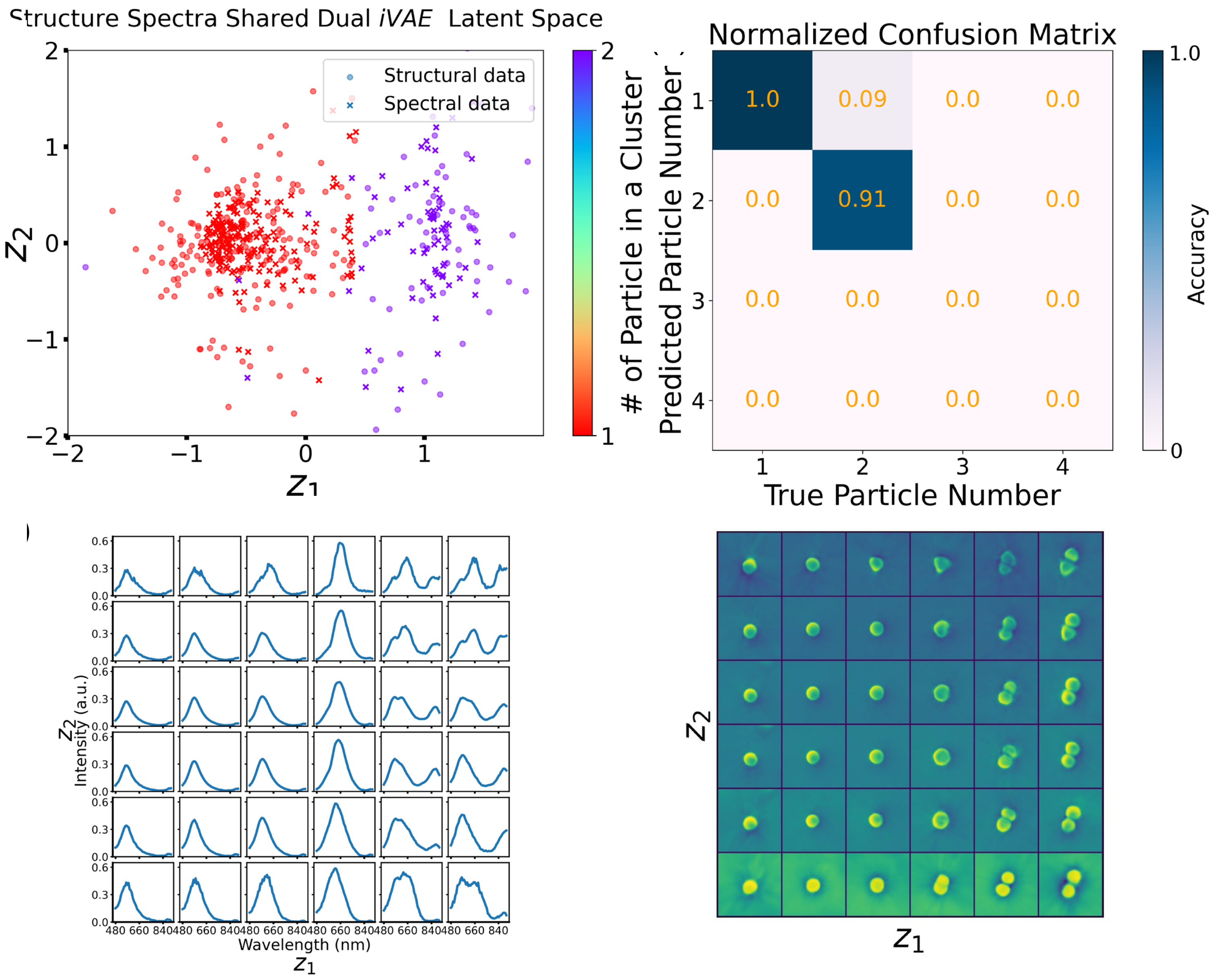
Darkfield Image Ground Truth Prediction



Overall Particles



Dual VAE: Predictions for only 1 and 2 mers



Questions to ask when using ML

Data science:

- What is the dimensionality of feature and target spaces
- How variable are the features in these spaces (VAE can help)
- How much data do I have?

Domain knowledge:

- How predictive do I expect this relationship to be
- What other factors matter from materials side
- Can measurements introduce additional factors of variance?

Experiment planning:

- How much data can I have?
- Is it a static learning problem, or am I interested in ML-assisted experiment?