THE UNIVERSITY of TENNESSEE KNOXVILLE

Variational Autoencoders

Sergei V. Kalinin

The VAE Story

- What are (Variational) autoencoders?
- Key notions:
  - Encoding and decoding
  - Latent distribution
  - Latent representations
  - Disentanglement of the representations
- Why invariances: rotational, translational, and shear
- Other colors of VAEs:
  - Semi-supervised
  - Conditional
  - Joint
- VAEs for real-world examples
- From VAEs to encoder-decoders (VED)
- Further opportunities:
  - Physics constraints
  - Representation learning
- Active learning: DKL

Autoencoders: Encoding

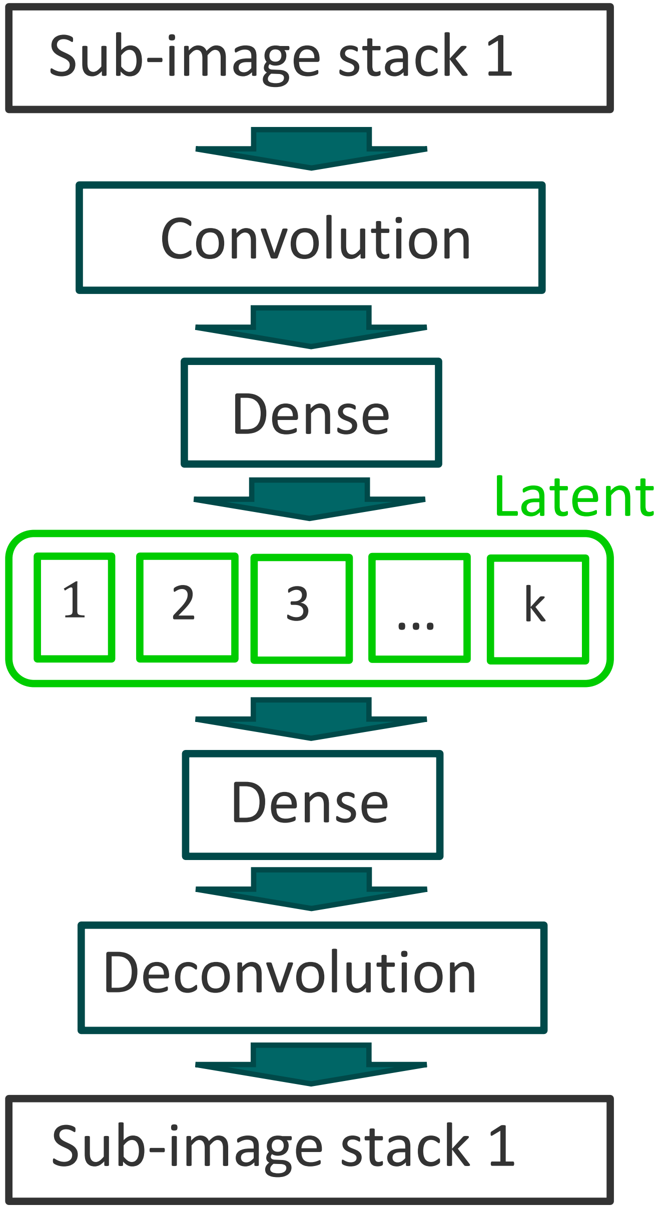
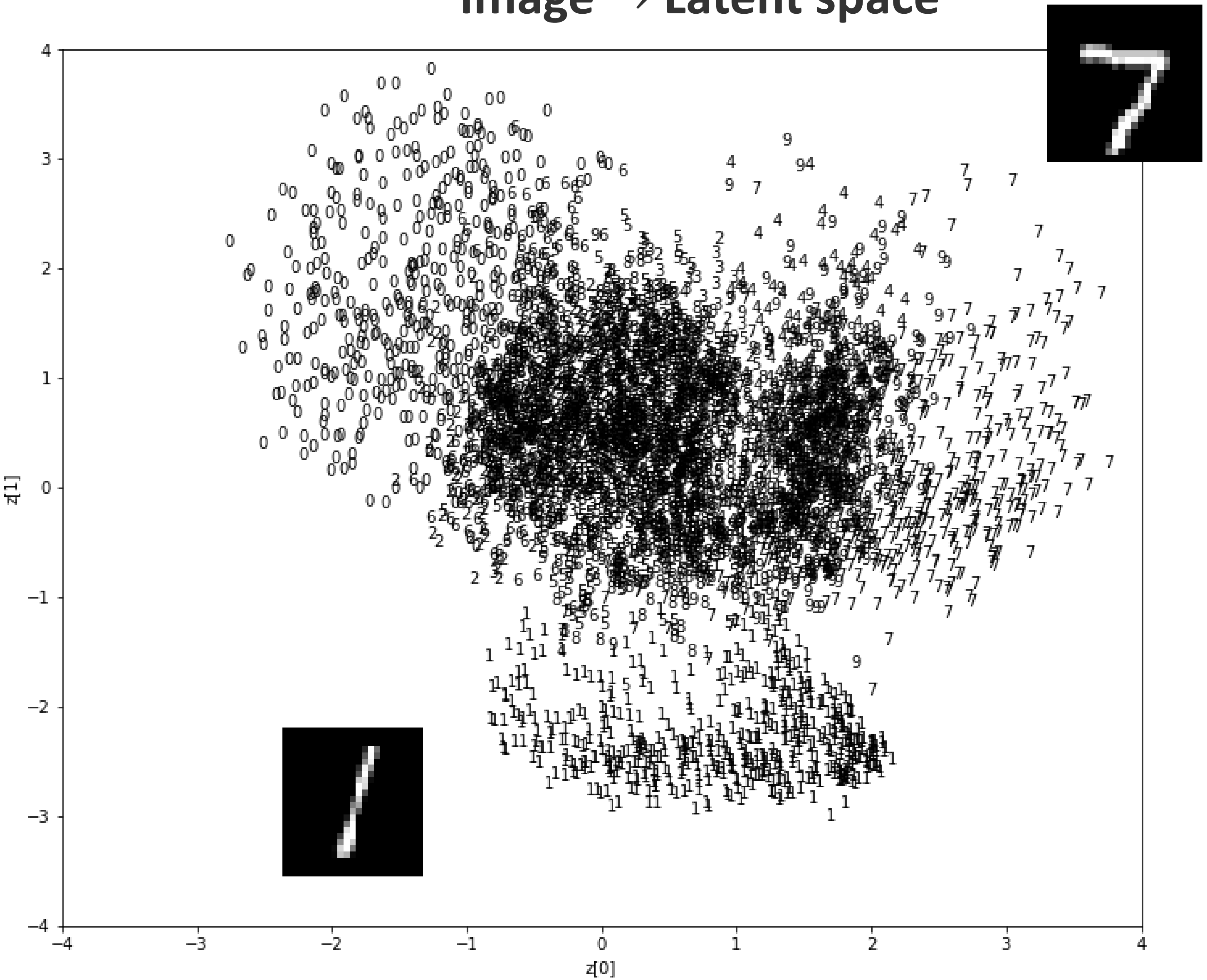
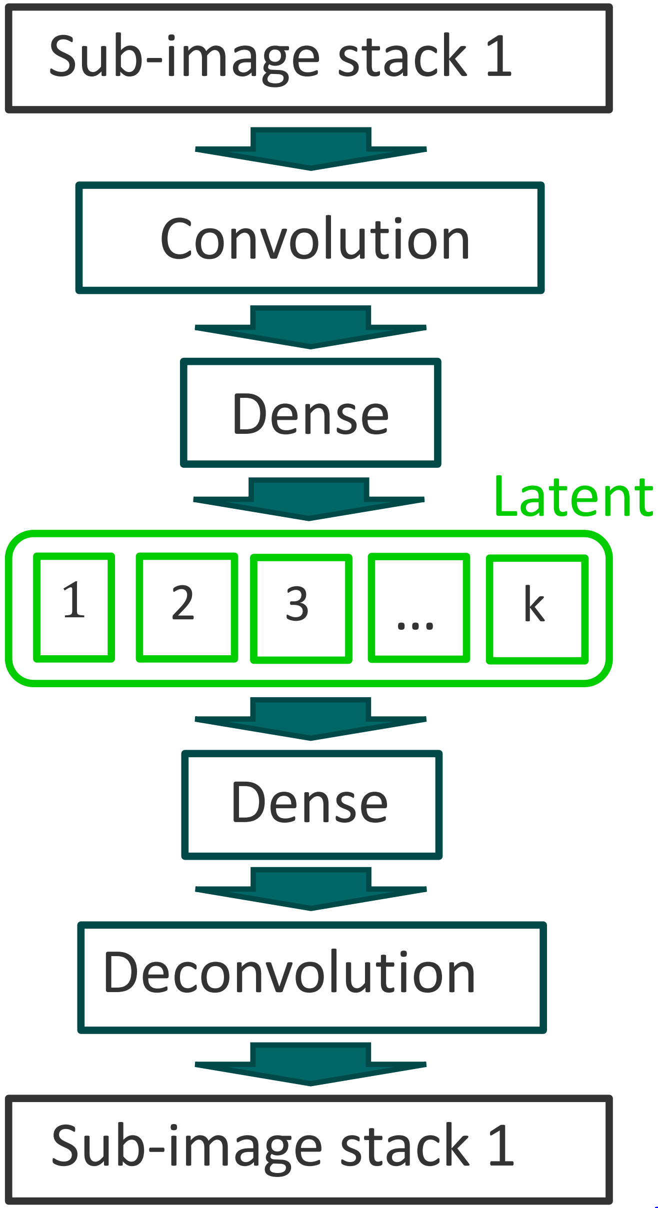


Image → Latent space

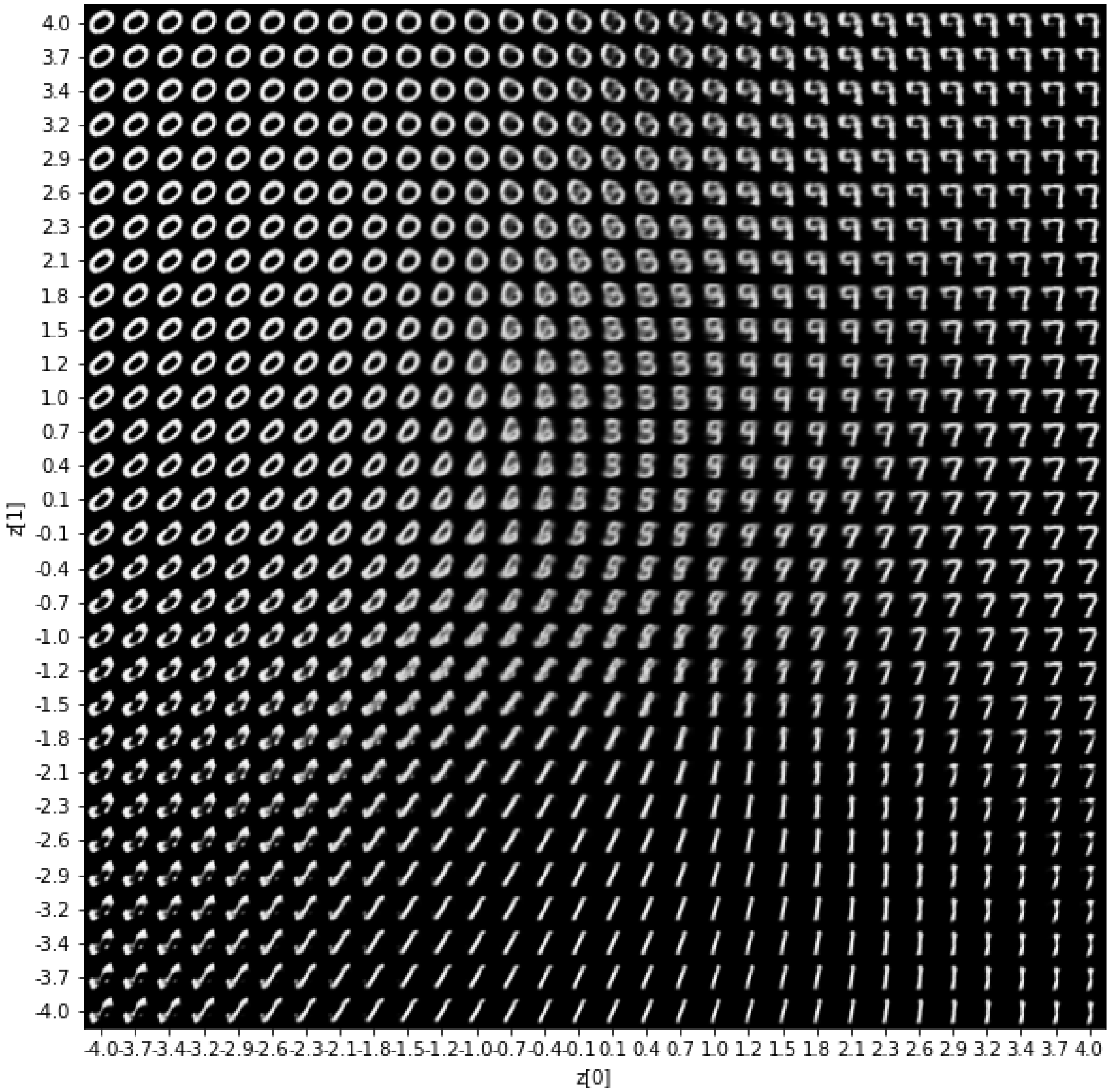


Latent distribution: Encoding the data via low dimensional vector

Autoencoders: Decoding



Latent space → Image



Latent representation: Decoding images from uniform grid in latent space

AE and Variational AE (VAE)

Geoffrey Hinton

FOLLOW

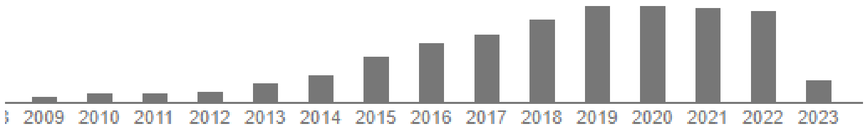
Emeritus Prof. Comp Sci, U.Toronto & Engineering Fellow, Google
Verified email at cs.toronto.edu - Homepage

machine learning psychology artificial intelligence cognitive science computer science

TITLE	CITED BY	YEAR
Imagenet classification with deep convolutional neural networks A Krizhevsky, I Sutskever, GE Hinton Communications of the ACM 60 (6), 84-90	130318	2017
Deep learning Y LeCun, Y Bengio, G Hinton Nature 521 (7553), 436-44	62790	2015
Dropout: a simple way to prevent neural networks from overfitting N Srivastava, G Hinton, A Krizhevsky, I Sutskever, R Salakhutdinov The journal of machine learning research 15 (1), 1929-1958	42078	2014
Visualizing data using t-SNE L van der Maaten, G Hinton Journal of Machine Learning Research 9 (Nov), 2579-2605	35035	2008
Learning representations by back-propagating errors DE Rumelhart, GE Hinton, RJ Williams Nature 323 (6088), 533-536	32239	1986
Learning internal representations by error-propagation DE Rumelhart, GE Hinton, RJ Williams Parallel Distributed Processing: Explorations in the Microstructure of ...	30711	1986
Schemata and sequential thought processes in PDP models. D Rumelhart, P Smolenksy, J McClelland, G Hinton Parallel distributed processing: Explorations in the microstructure of ...	28073 *	1986
Learning multiple layers of features from tiny images A Krizhevsky, G Hinton	21876	2009
Rectified linear units improve restricted boltzmann machines V Nair, GE Hinton Proceedings of the 27th international conference on machine learning (ICML ...	21050	2010
Reducing the dimensionality of data with neural networks GE Hinton, RR Salakhutdinov Science 313 (5786), 504-507	19930	2006

Reducing the dimensionality of data with neural networks

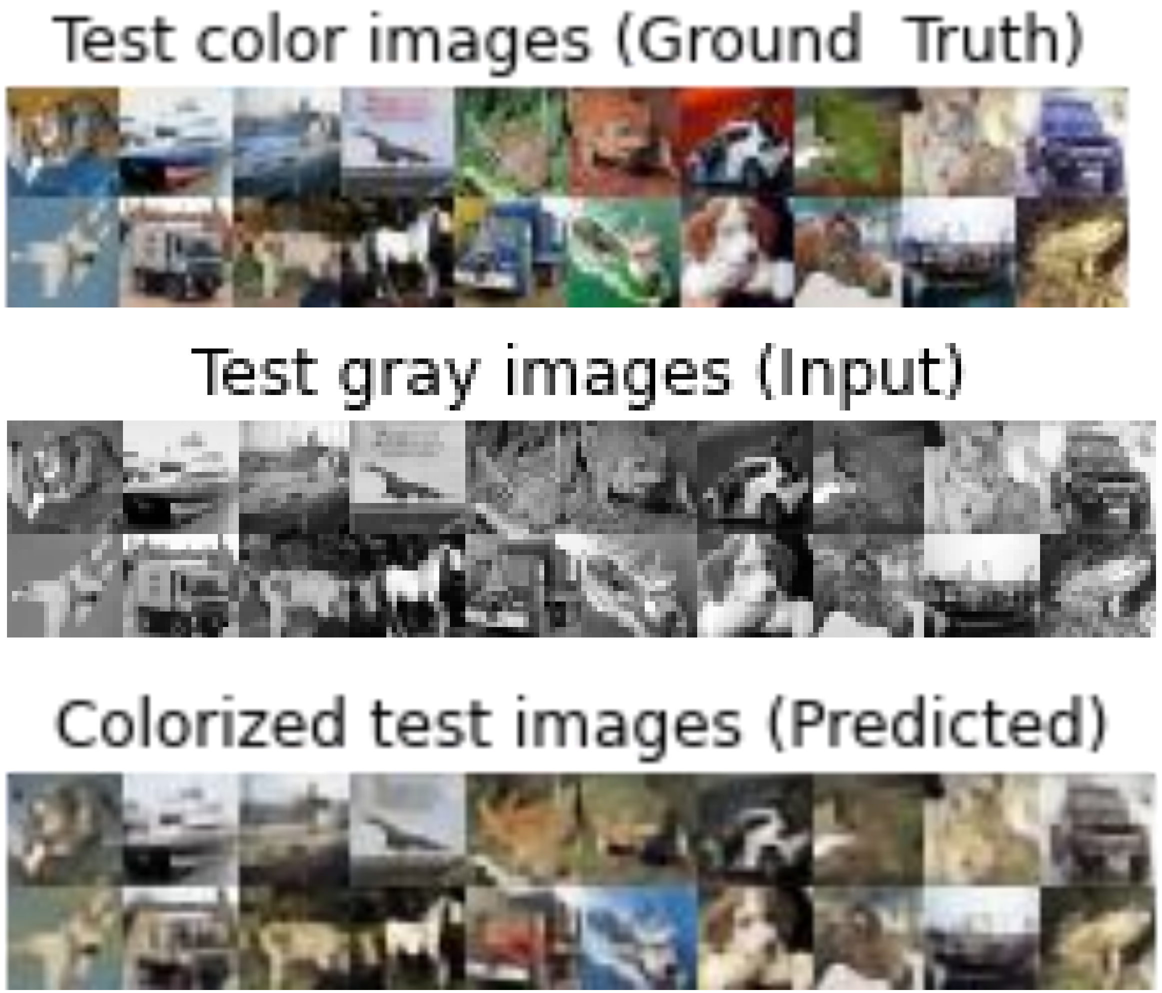
Authors	Geoffrey E Hinton, Ruslan R Salakhutdinov
Publication date	2006/7/28
Journal	Science
Volume	313
Issue	5786
Pages	504-507
Publisher	American Association for the Advancement of Science
Description	High-dimensional data can be converted to low-dimensional codes by training a multilayer neural network with a small central layer to reconstruct high-dimensional input vectors. Gradient descent can be used for fine-tuning the weights in such “autoencoder” networks, but this works well only if the initial weights are close to a good solution. We describe an effective way of initializing the weights that allows deep autoencoder networks to learn low-dimensional codes that work much better than principal components analysis as a tool to reduce the dimensionality of data.
Total citations	Cited by 19930



Denoising

- **Training:** pairs of the high-noise and low-noise images
- **Application:** new high noise images (from the same distribution)
- **Concern:** has to be from the same distribution

Restoration



- **Training:** pairs of the grayscale and color images
- **Application:** new grayscale images (from the same distribution)
- **Concern:** has to be from the same distribution

AE and Variational AE (VAE)

Diederik P. Kingma

Other names ▸

FOLLOW

Research Scientist, Google Brain
Verified email at google.com - Homepage

Machine Learning Deep Learning Neural Networks Generative Models Variational Inference

TITLE	CITED BY	YEAR
Adam: A Method for Stochastic Optimization DP Kingma, J Ba Proceedings of the 3rd International Conference on Learning Representations ...	141306	2014
Auto-Encoding Variational Bayes DP Kingma, M Welling arXiv preprint arXiv:1312.6114	26540	2013
Semi-Supervised Learning with Deep Generative Models DP Kingma, S Mohamed, DJ Rezende, M Welling Advances in Neural Information Processing Systems, 3581-3589	2946	2014

- Variational Autoencoder (VAE): uses “reparameterization trick” to sample from the latent space
- Can be used for same tasks as AE
- Have a much better-behaved latent space: disentanglement of the representations

VAE Training

Latent manifold -> Image space

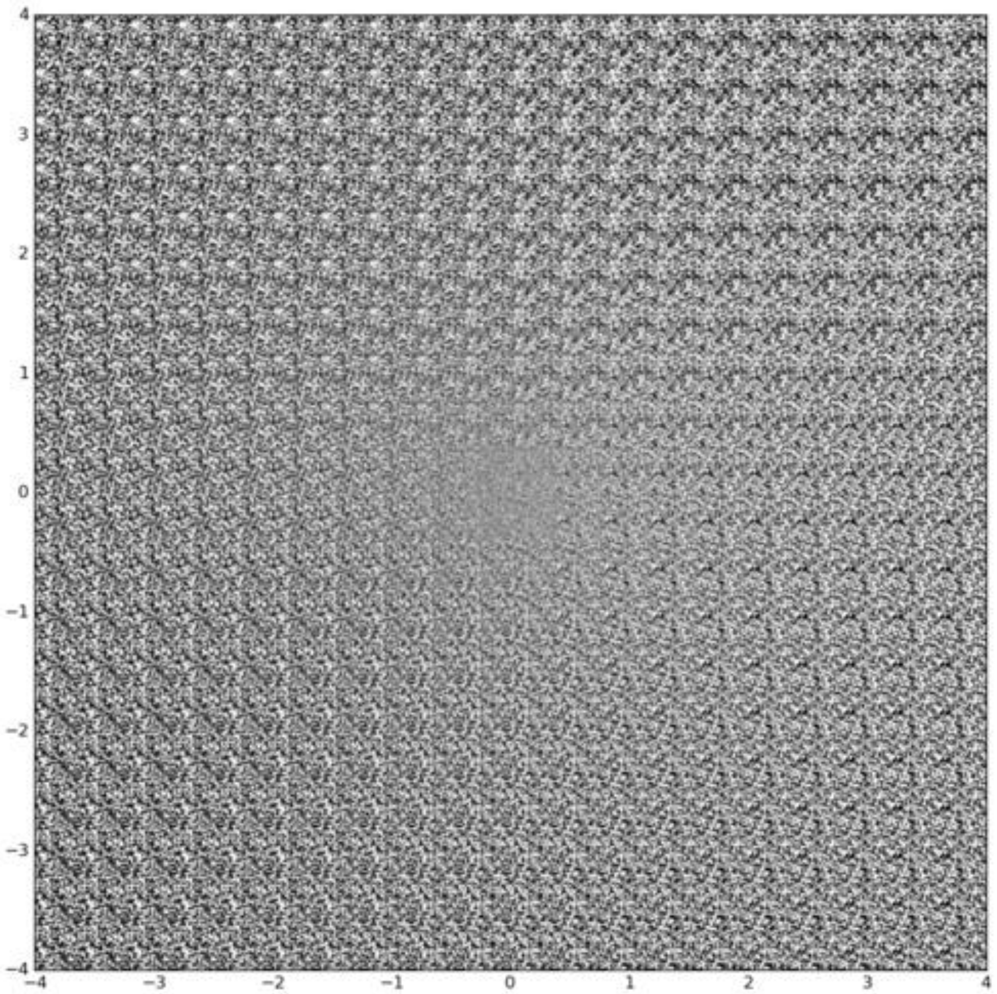
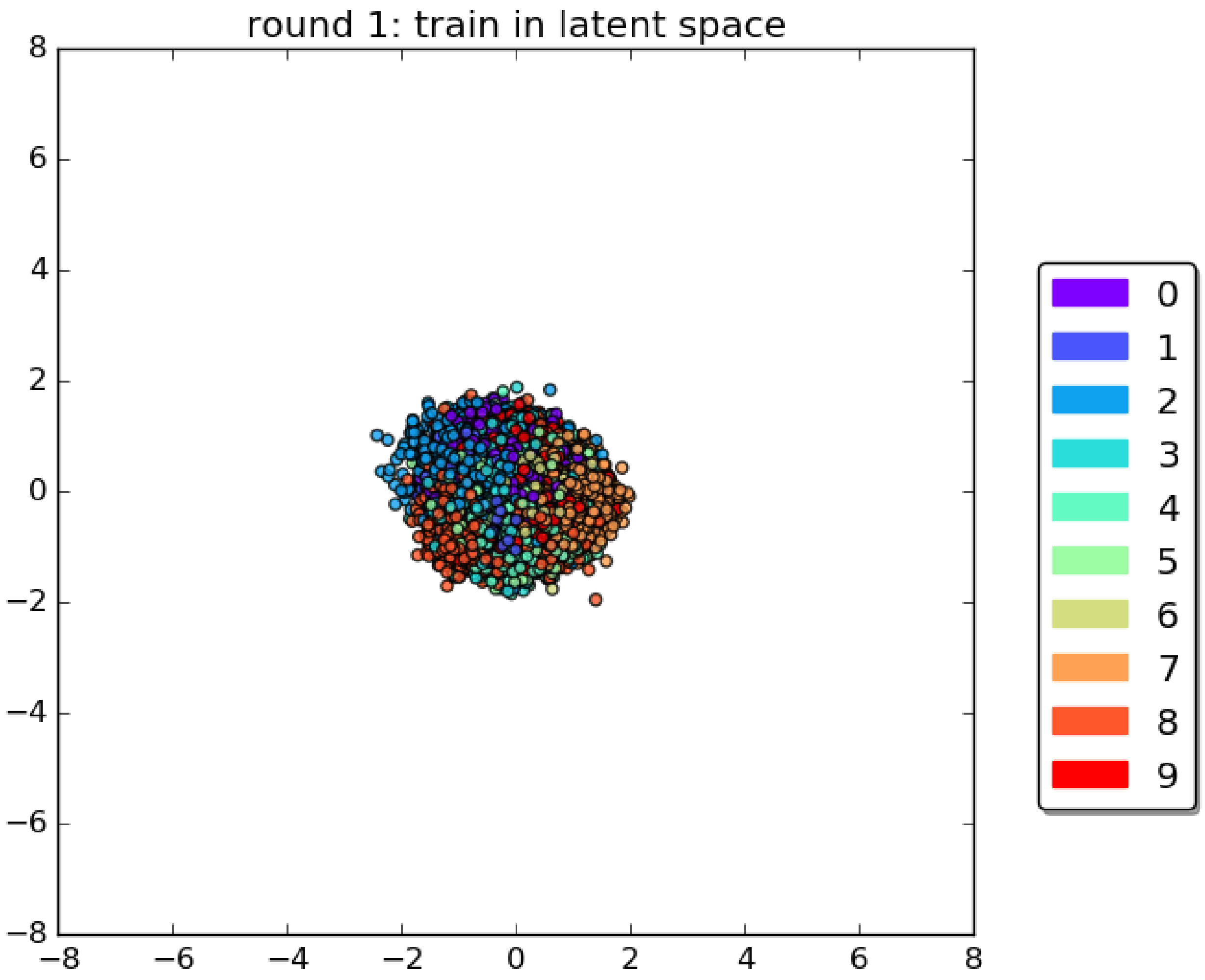
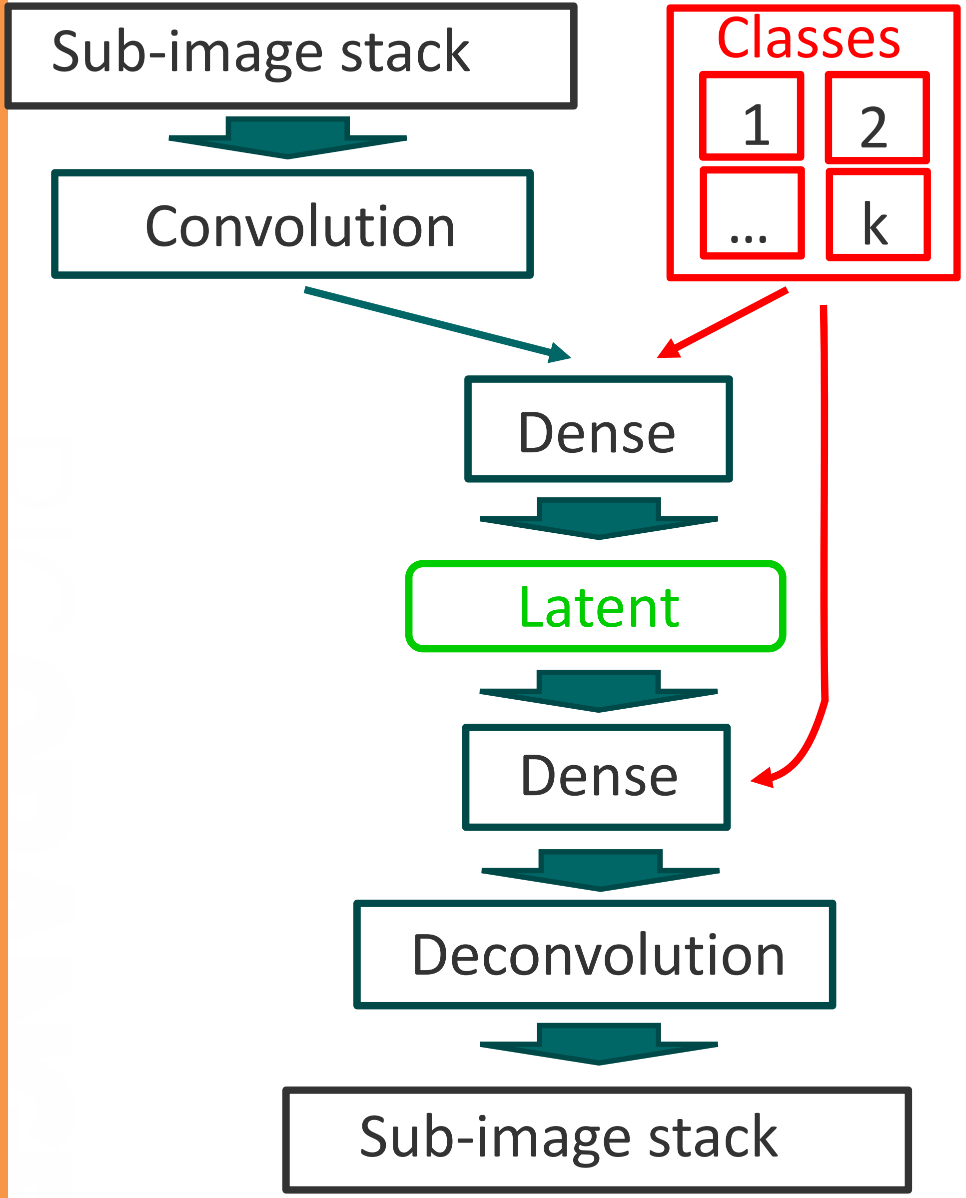
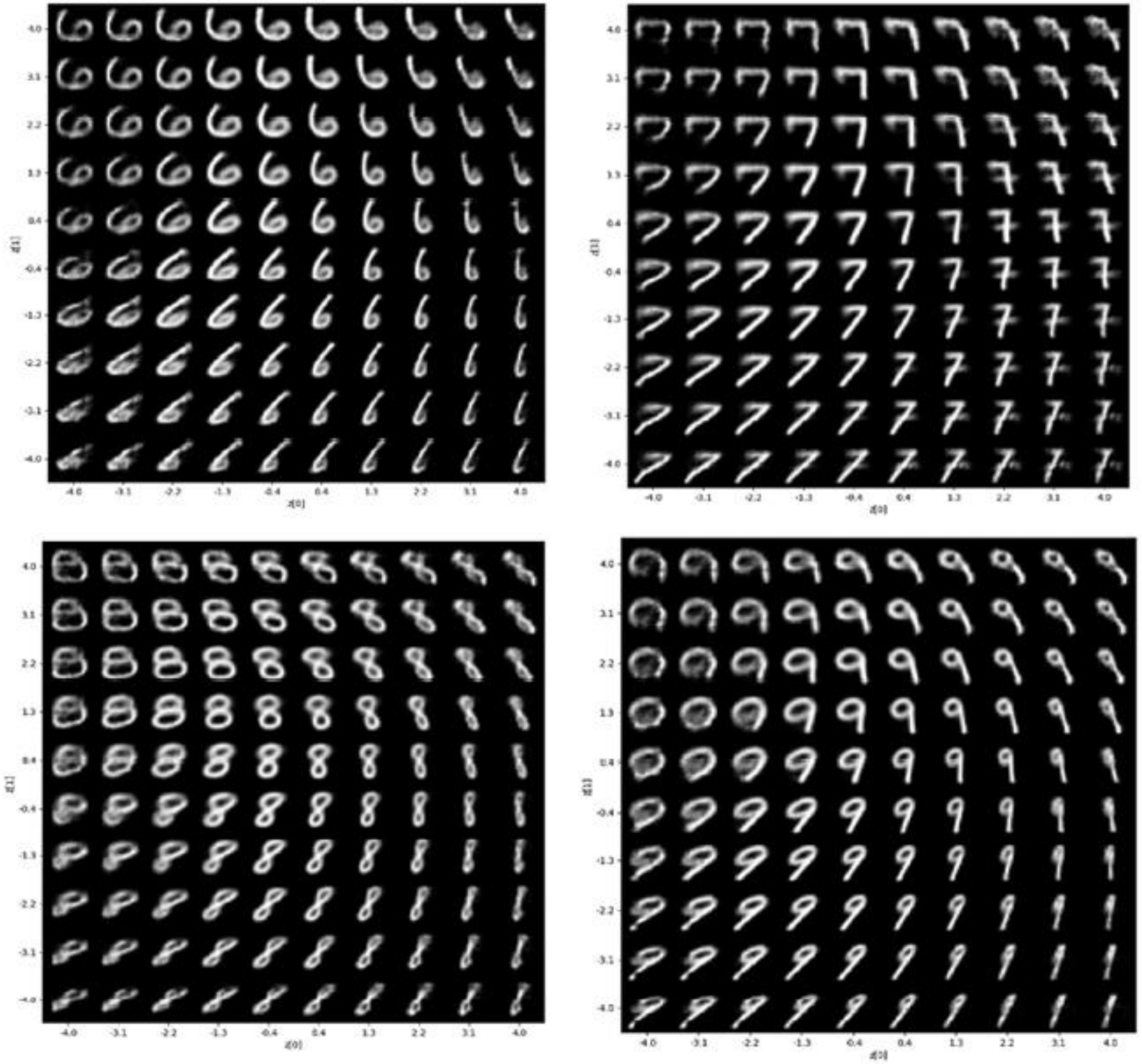


Image space -> Latent space



Conditional VAE



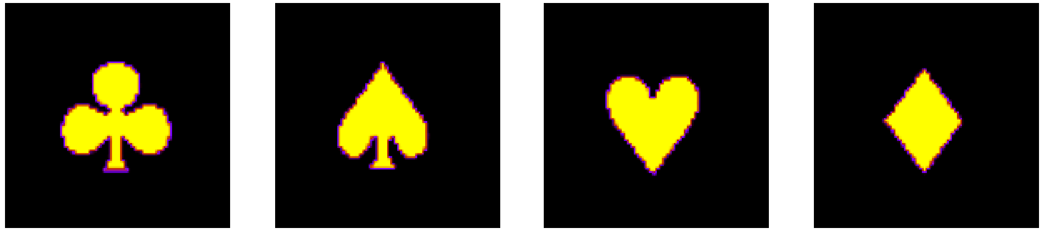


Note the trends in the latent representation for each digit: disentanglement of the representations

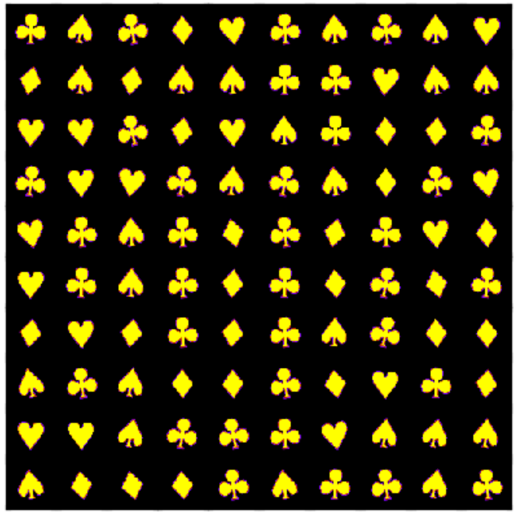
(R)VAE on Cards

Introduce the **cards** data set:

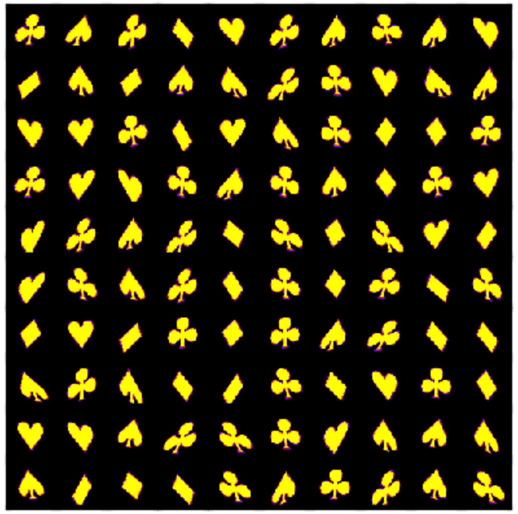
- Classical 4 hands (diamonds, clubs, pikes, hearts)
- Interesting similarities (pikes and hearts)
- And invariances on affine transforms (e.g. diamonds)



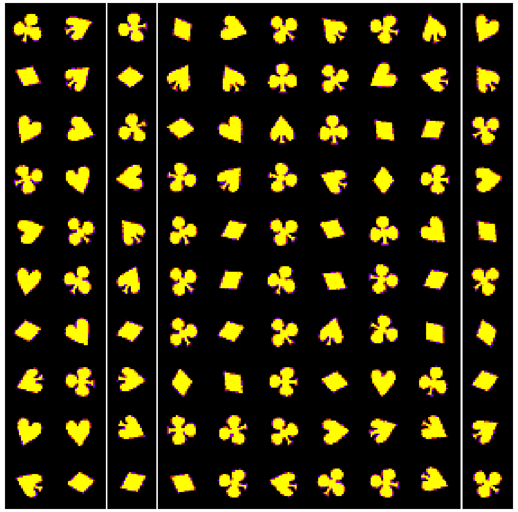
Cards 1: Low R (12 deg) and low S (1 deg)



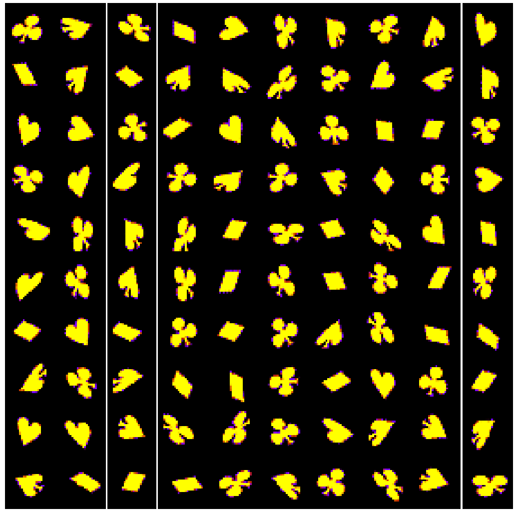
Cards 2: Low R (12 deg) and high S (20 deg)



Cards 3: High R (120 deg) and Low S (1 deg)



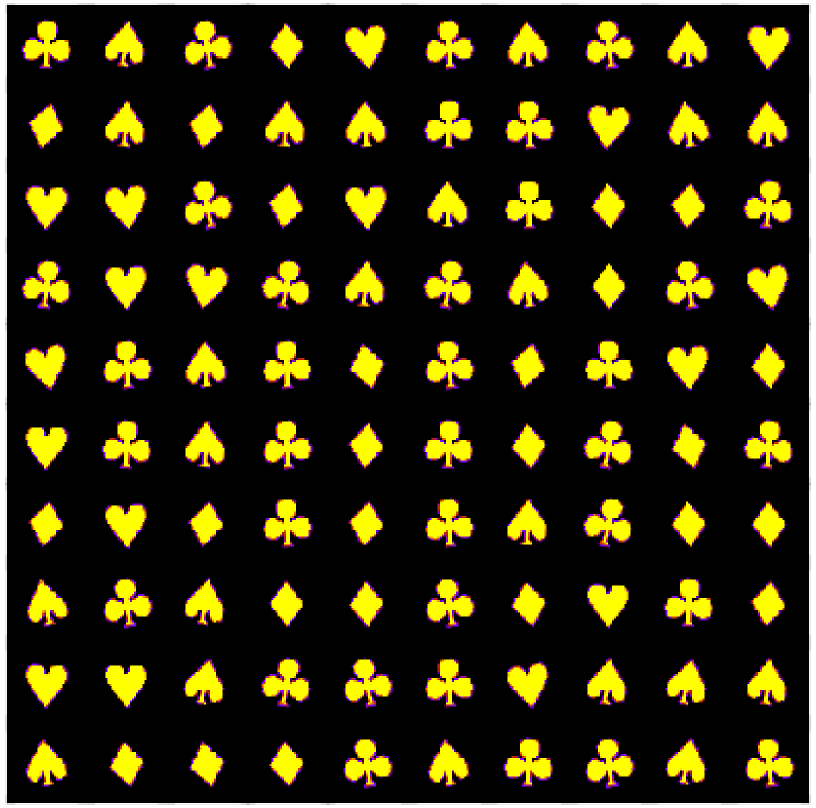
Cards 4: High R (120 deg) and high S (20 deg)



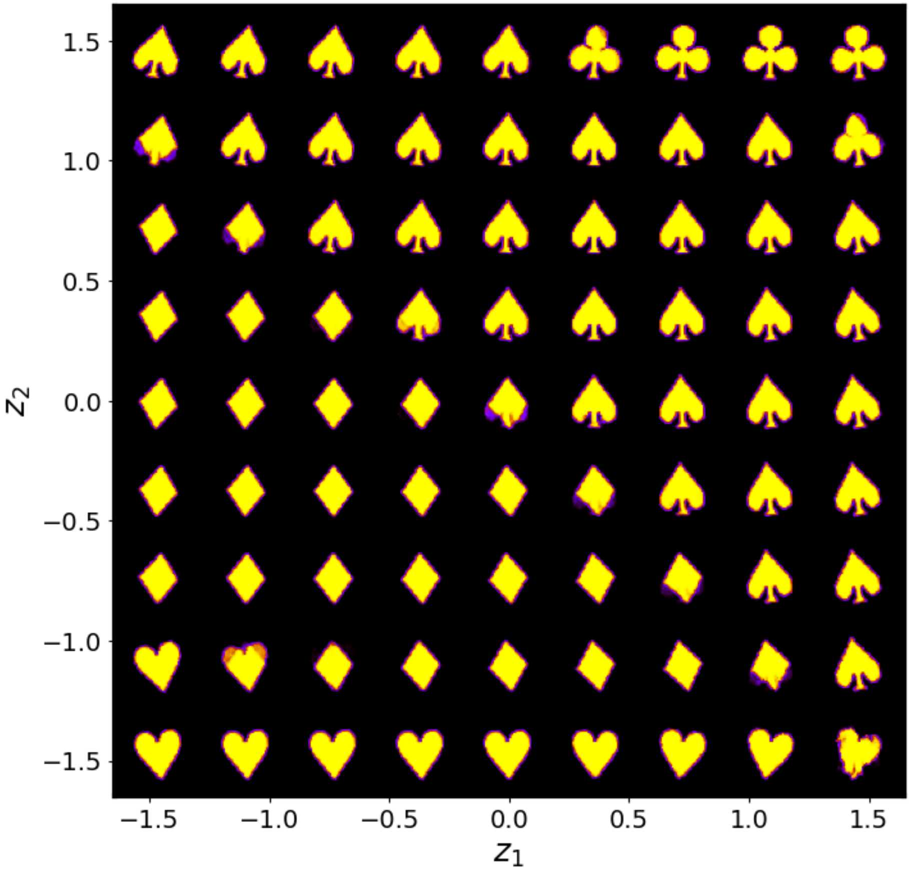
- Shear, rotations, and translations are **known** factors of variability (or traits) in data
- Can VAE disentangle representations and discover these factors of variability

VAE on Cards

Example of data

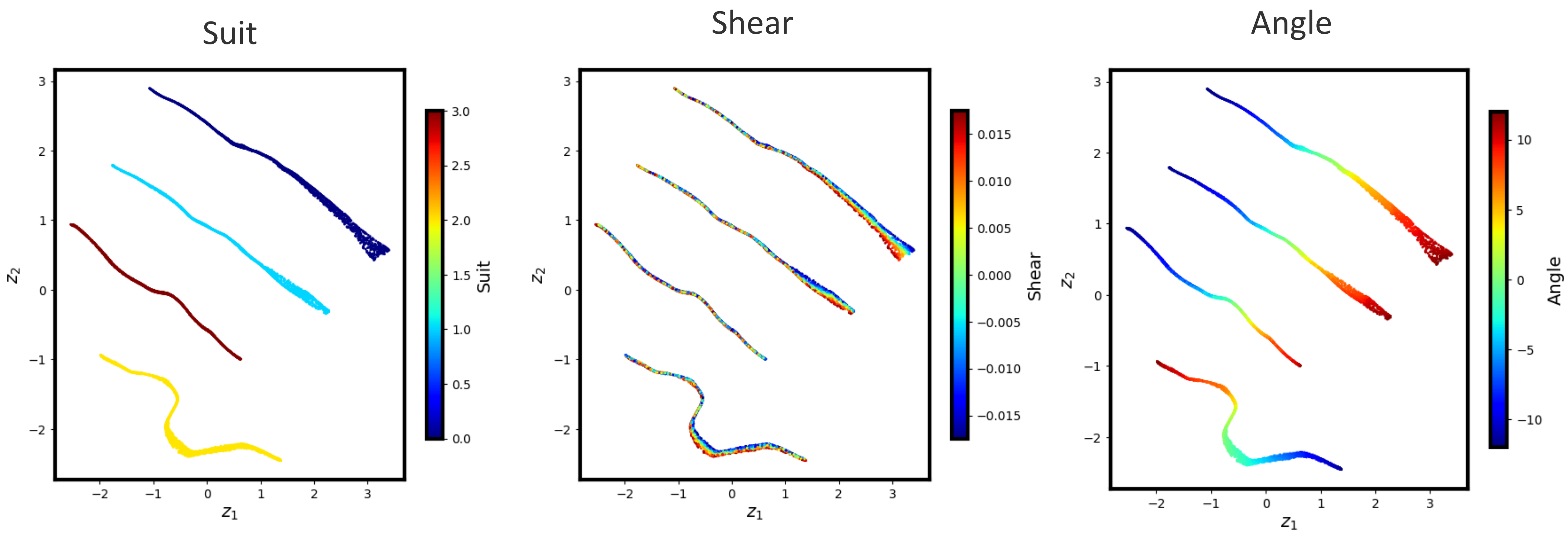


Latent representation



Cards 1: Low rotation (12 deg) and low shear (1 deg)

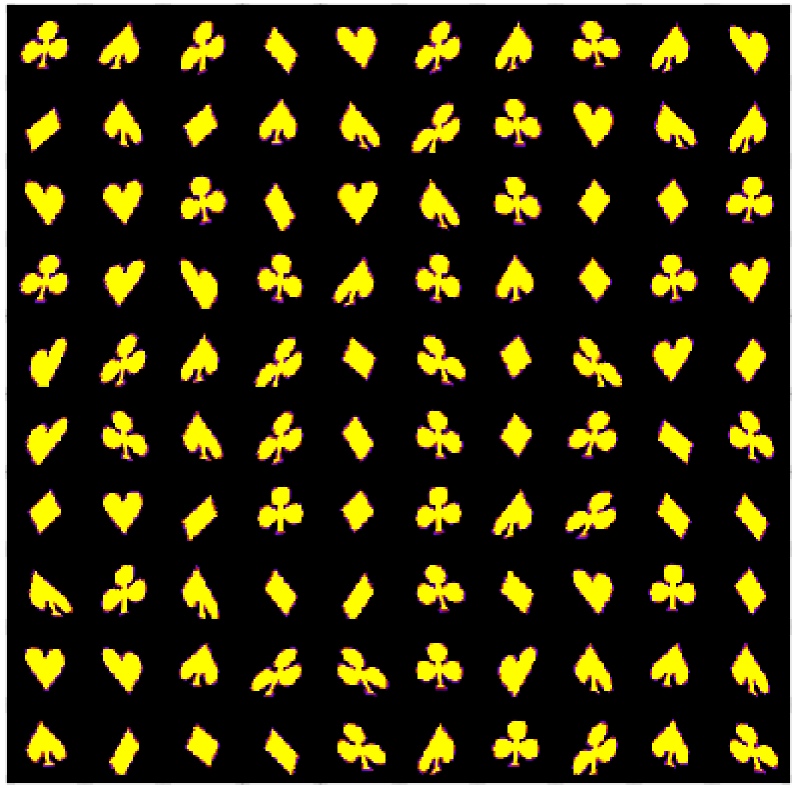
VAE on Cards



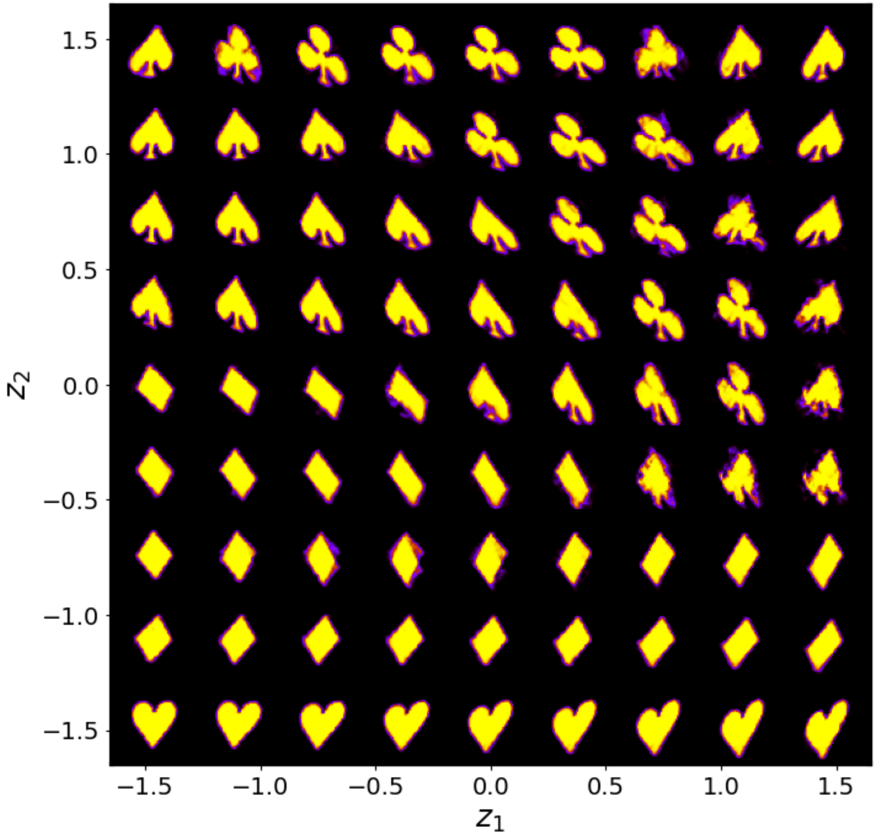
Cards 1: Low rotation (12 deg) and low shear (1 deg)

VAE on Cards

Example of data

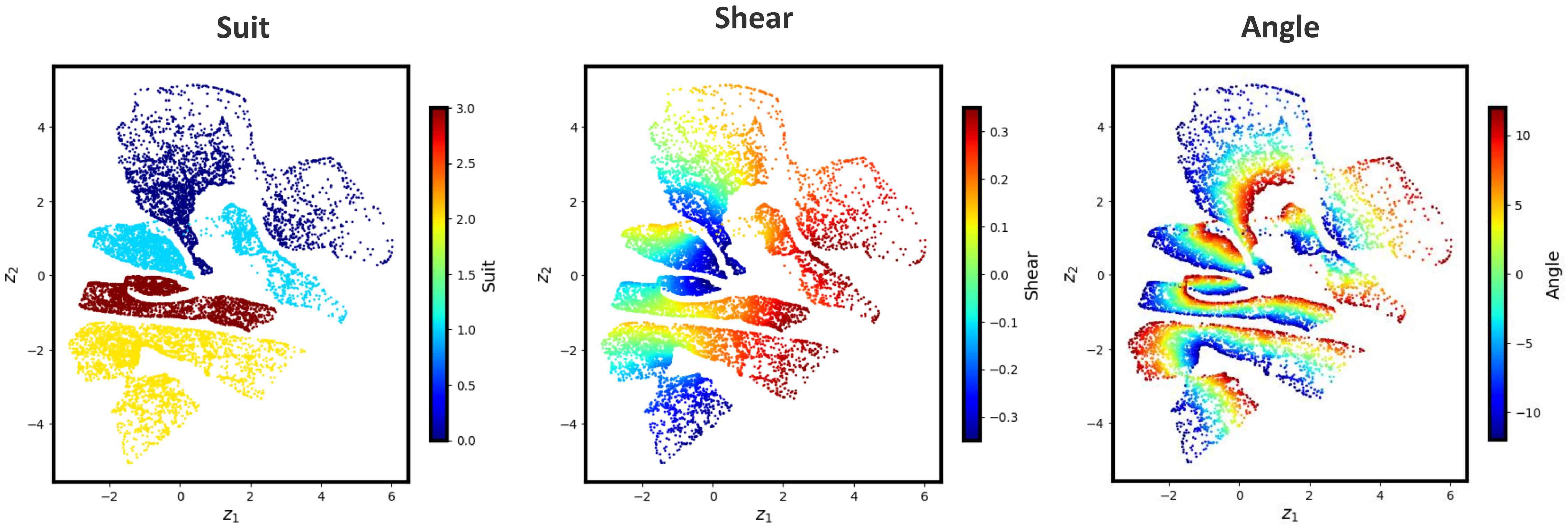


Latent representation



Cards 2: Low rotation (12 deg) and high shear (20 deg)

VAE on Cards

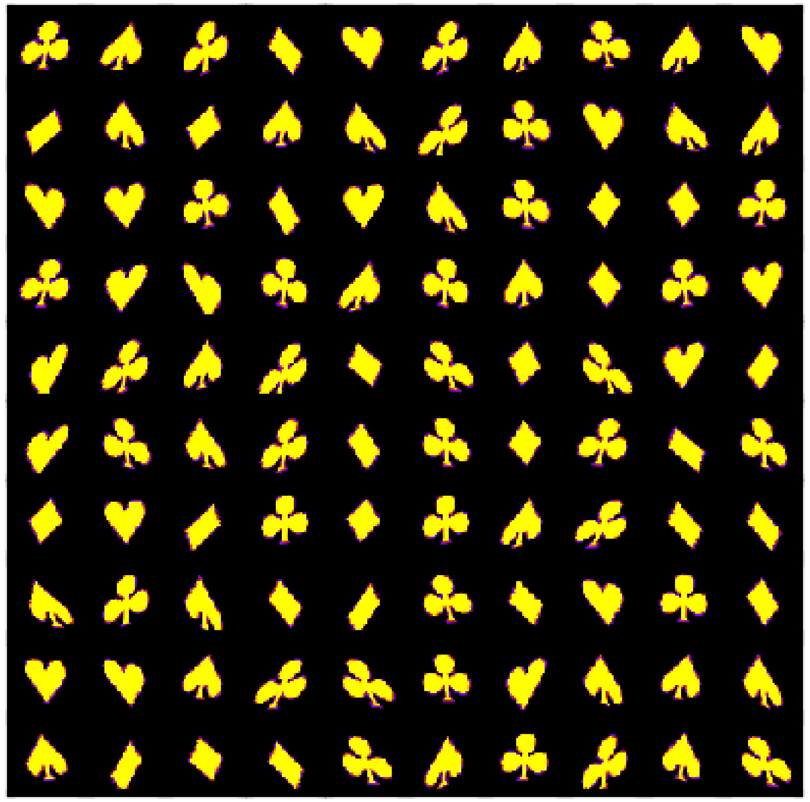


Cards 2: Low rotation (12 deg) and high shear (20 deg)

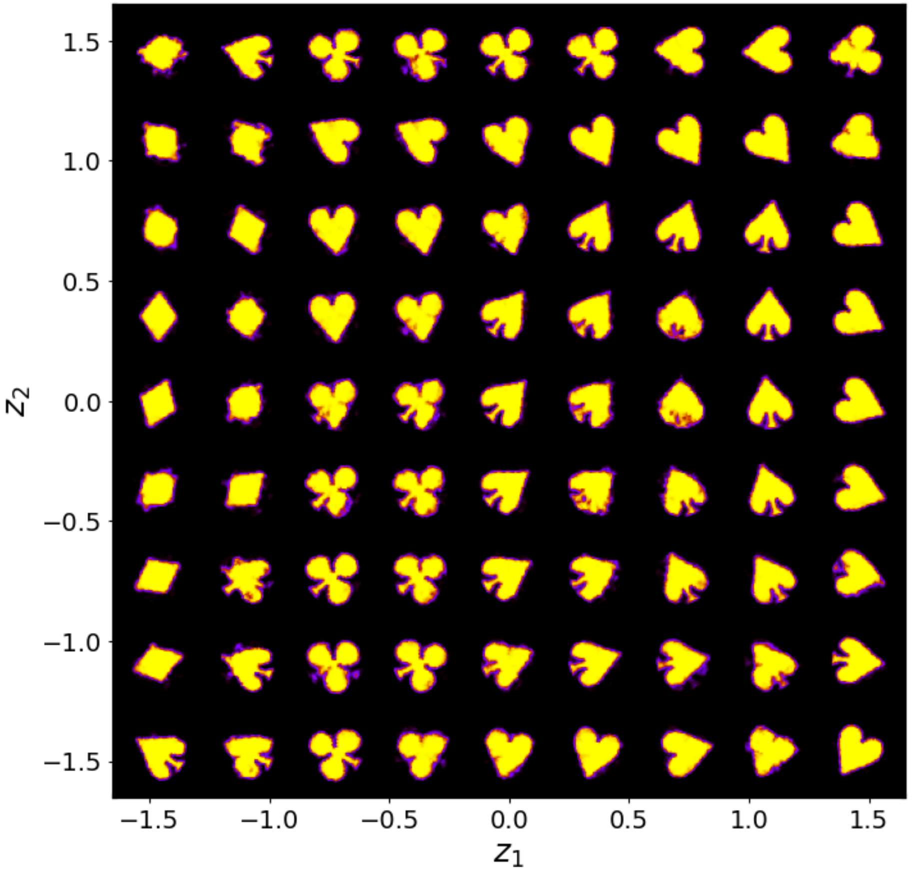
VAE on Cards

THE UNIVERSITY of TENNESSEE KNOXVILLE

Example of data

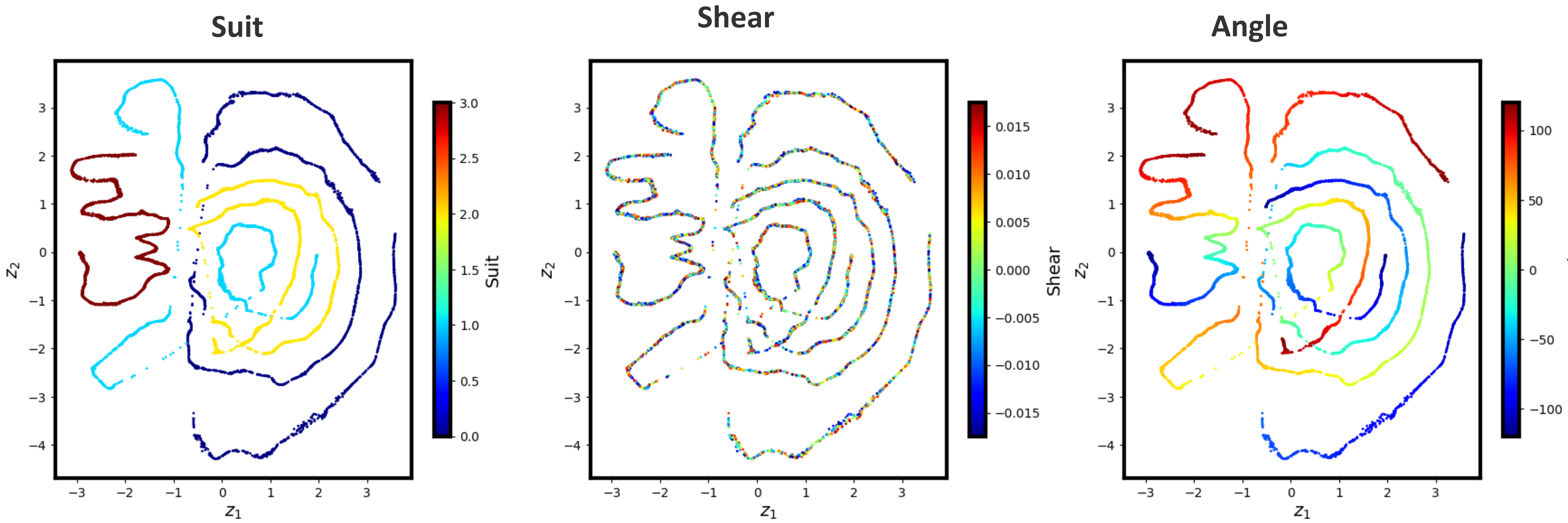


Latent representation



Cards 3: High rotation (120 deg) and low shear (1 deg)

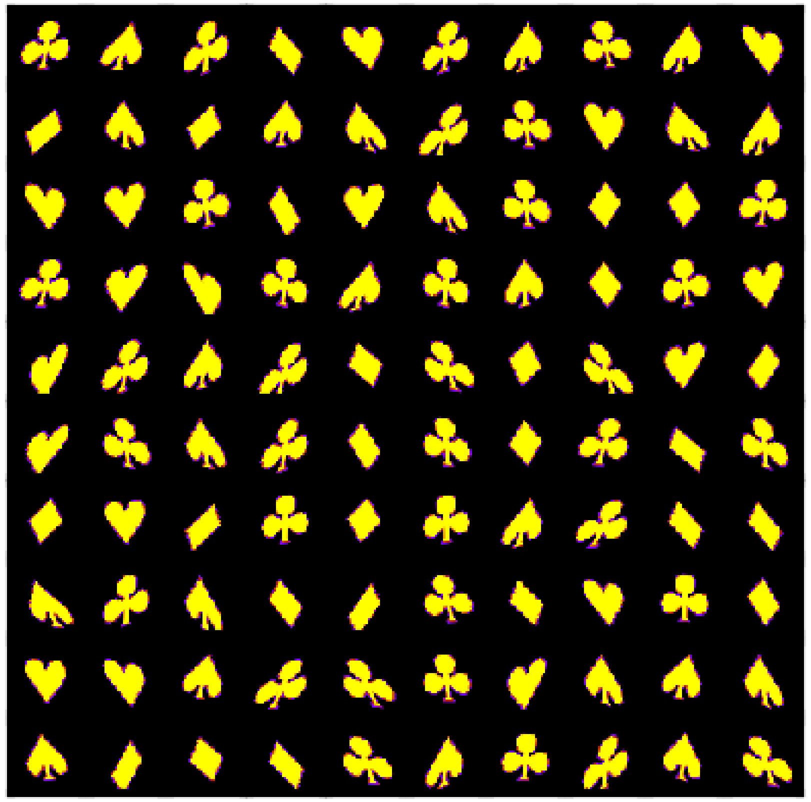
VAE on Cards



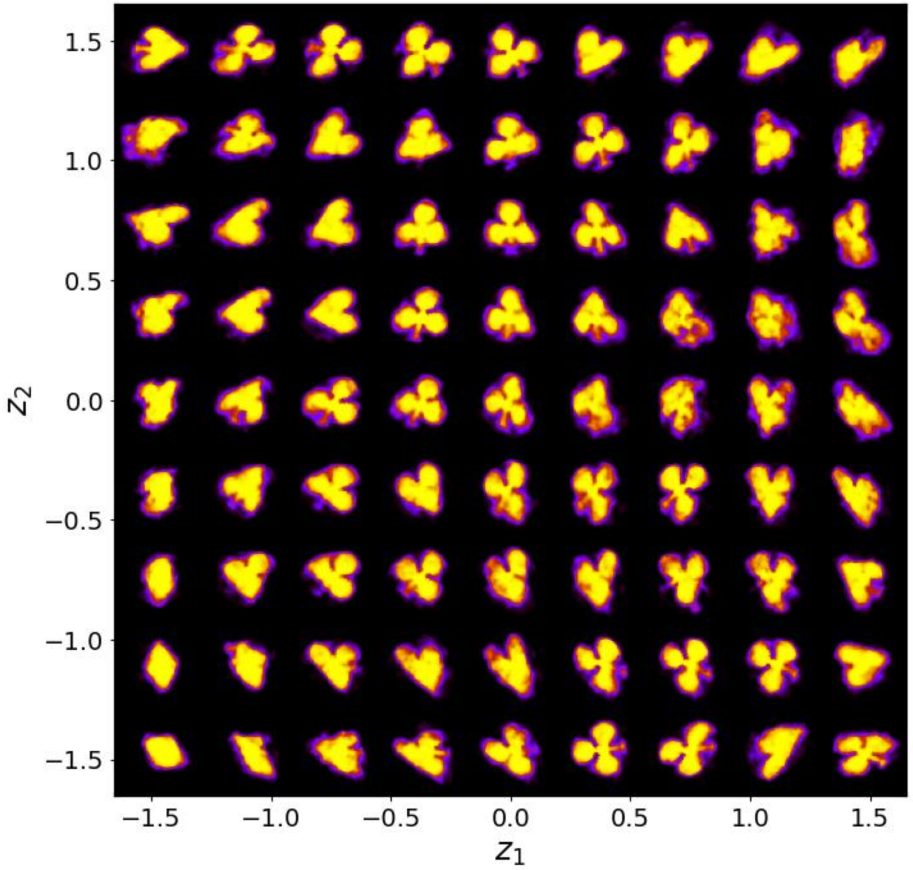
Cards 3: High rotation (120 deg) and low shear (1 deg)

VAE on Cards

Example of data



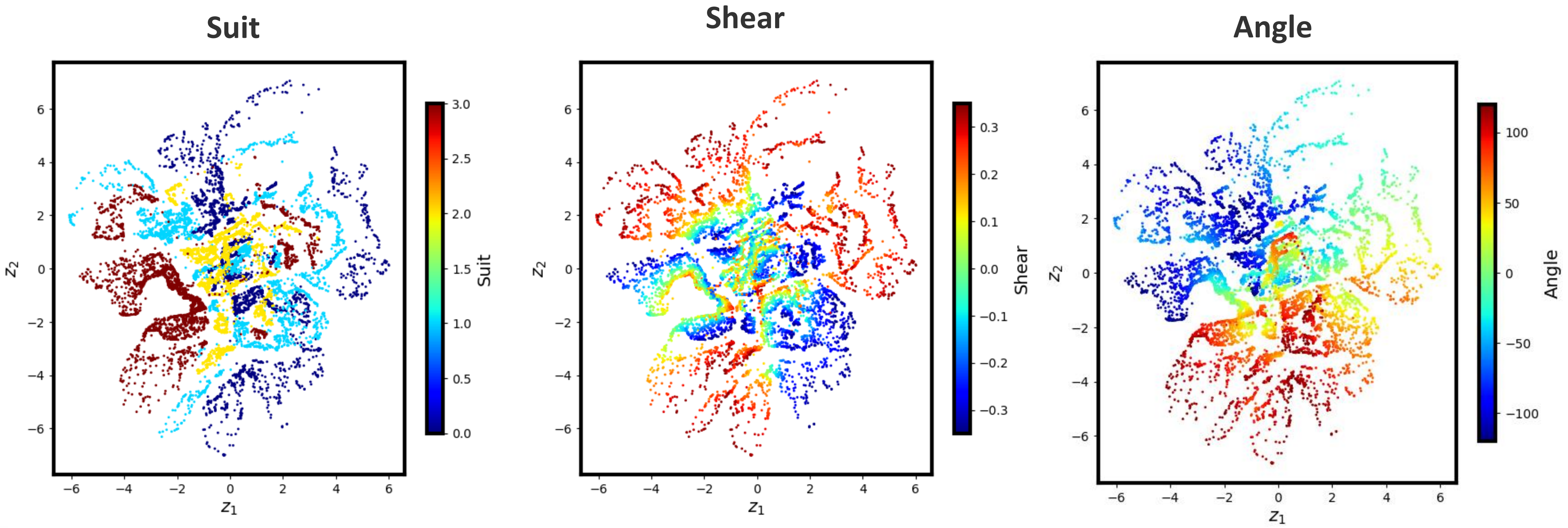
Latent representation



Cards 4: High rotation (120 deg) and high shear (20 deg)

VAE on Cards

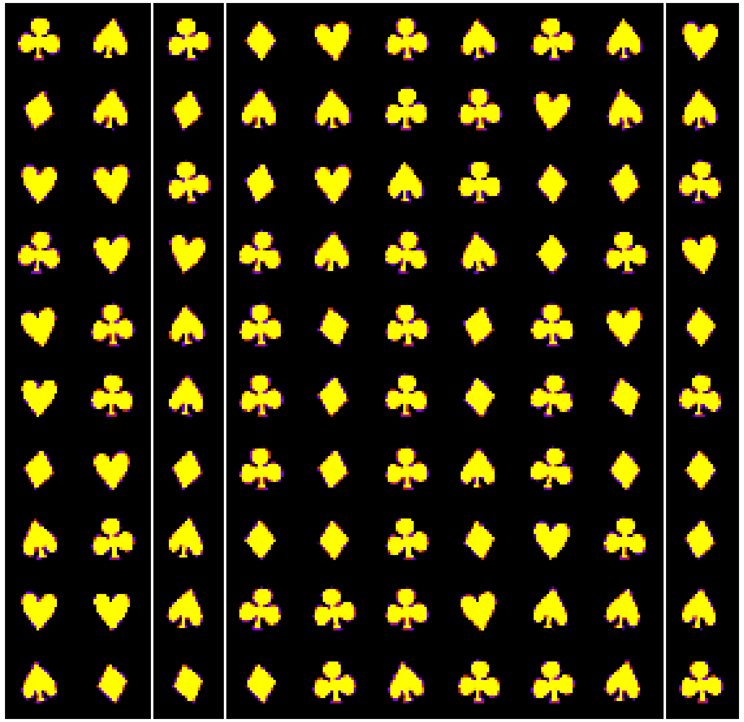
THE UNIVERSITY of TENNESSEE KNOXVILLE



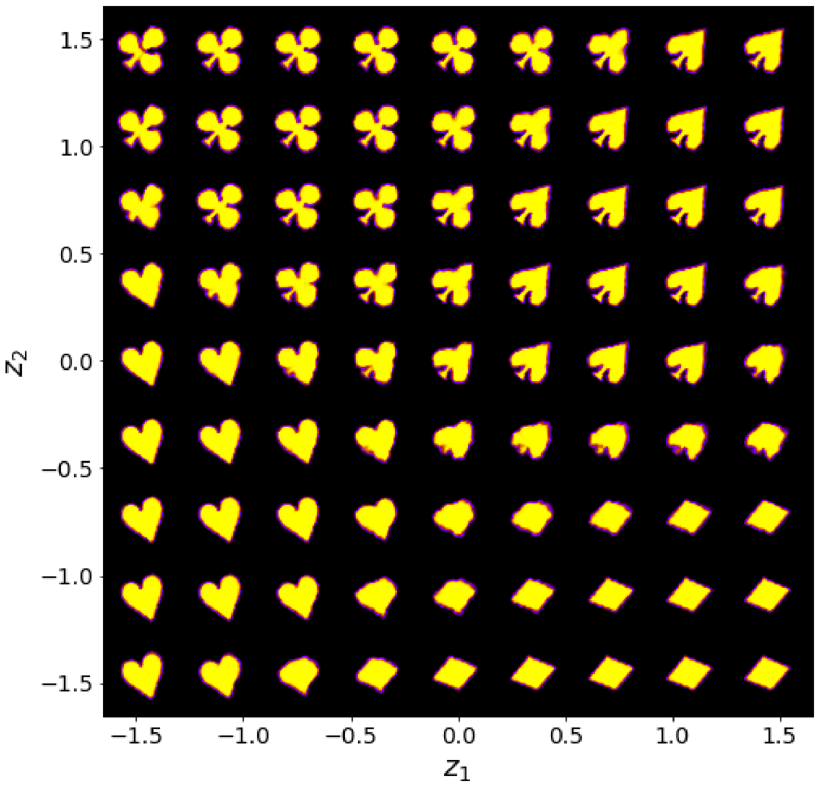
Cards 4: High rotation (120 deg) and high shear (20 deg)

rVAE on Cards

Example of data

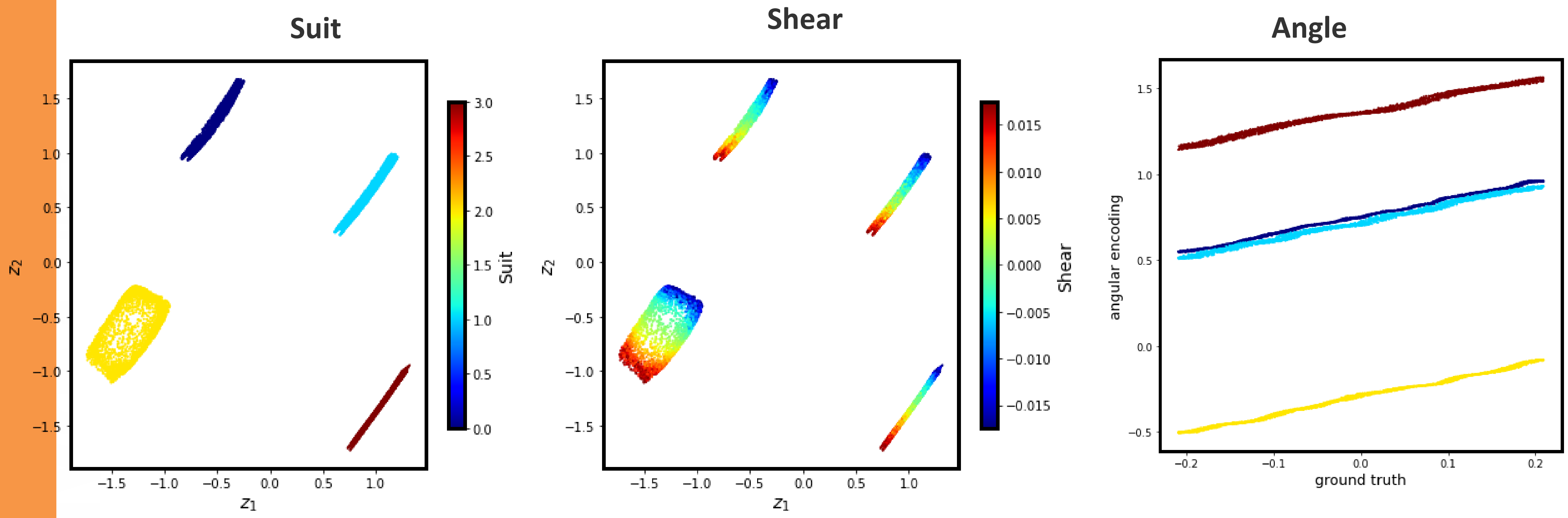


Latent representation



Cards 1: Low rotation (12 deg) and low shear (1 deg)

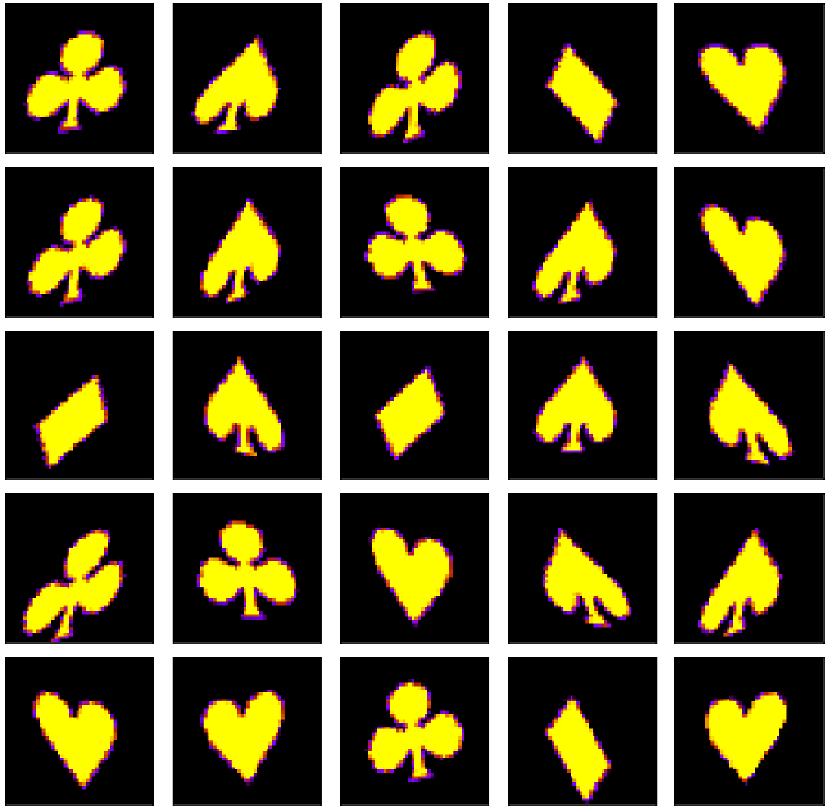
rVAE on Cards



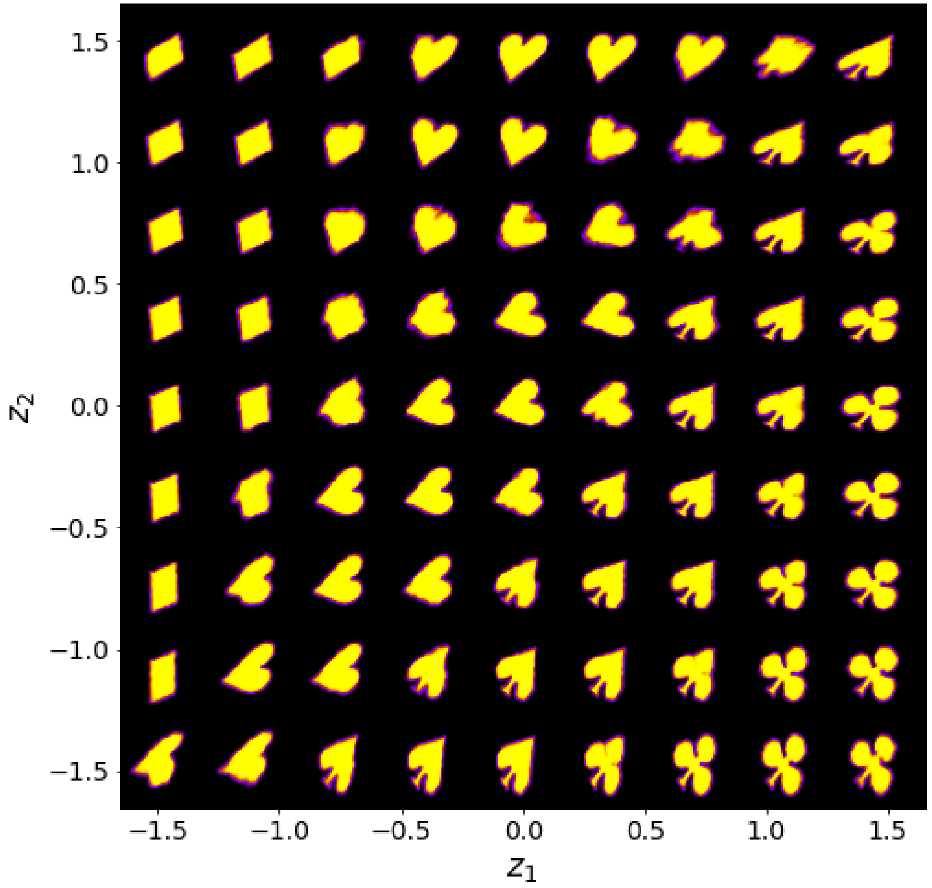
Cards 1: Low rotation (12 deg) and low shear (1 deg)

rVAE on Cards

Example of data

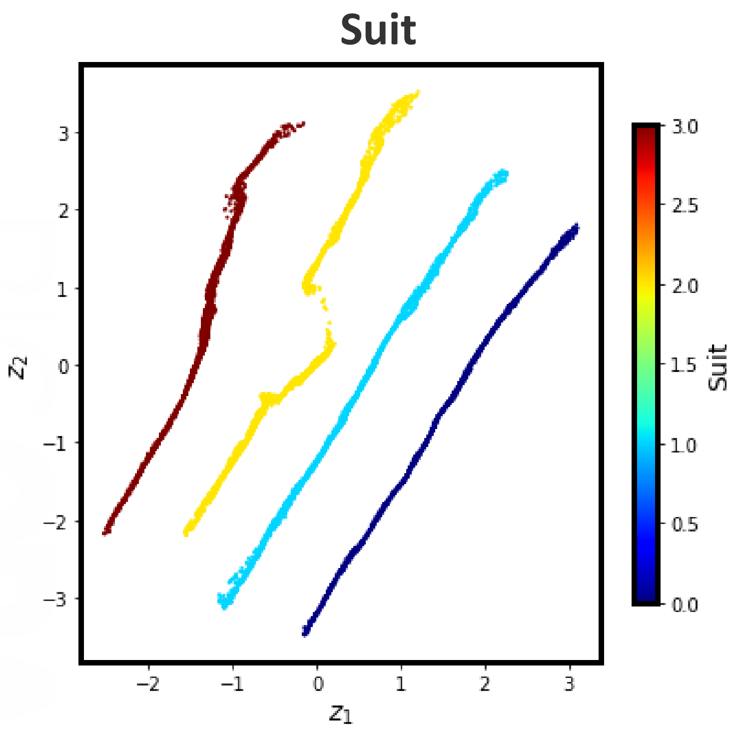


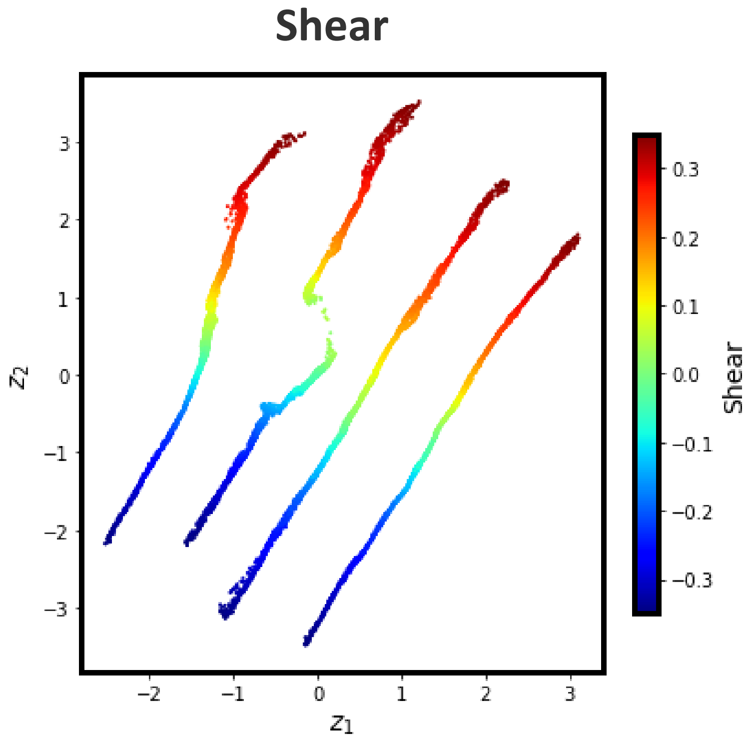
Latent representation

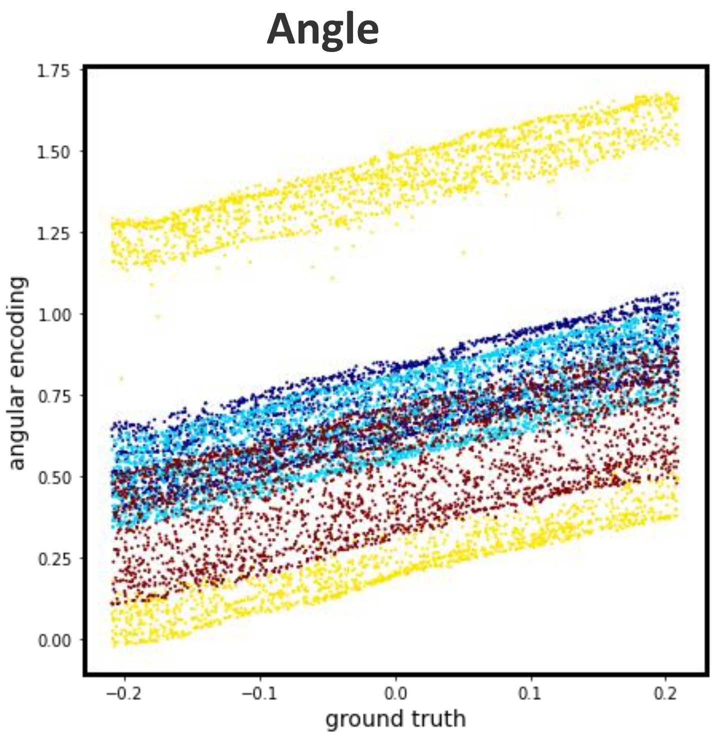


Cards 2: Low rotation (12 deg) and high shear (20 deg)

rVAE on Cards



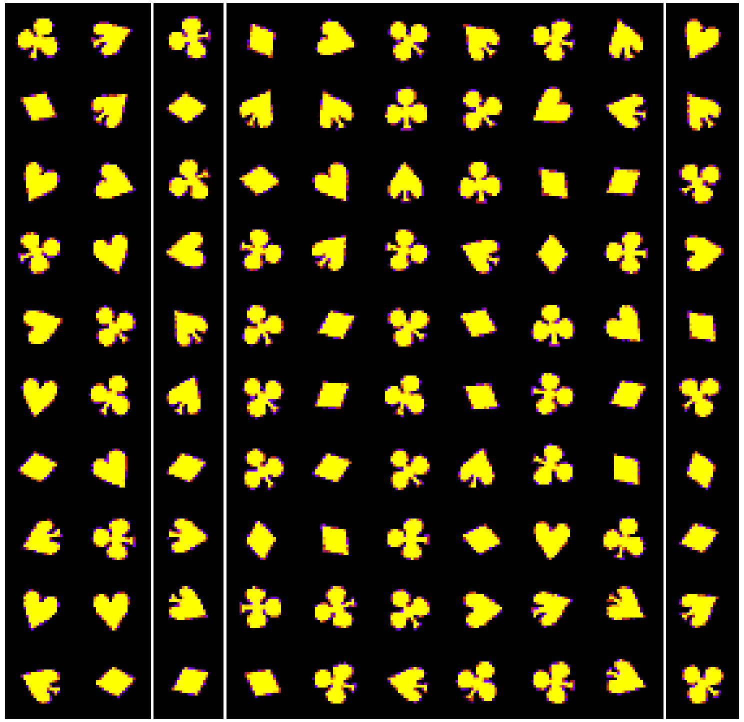




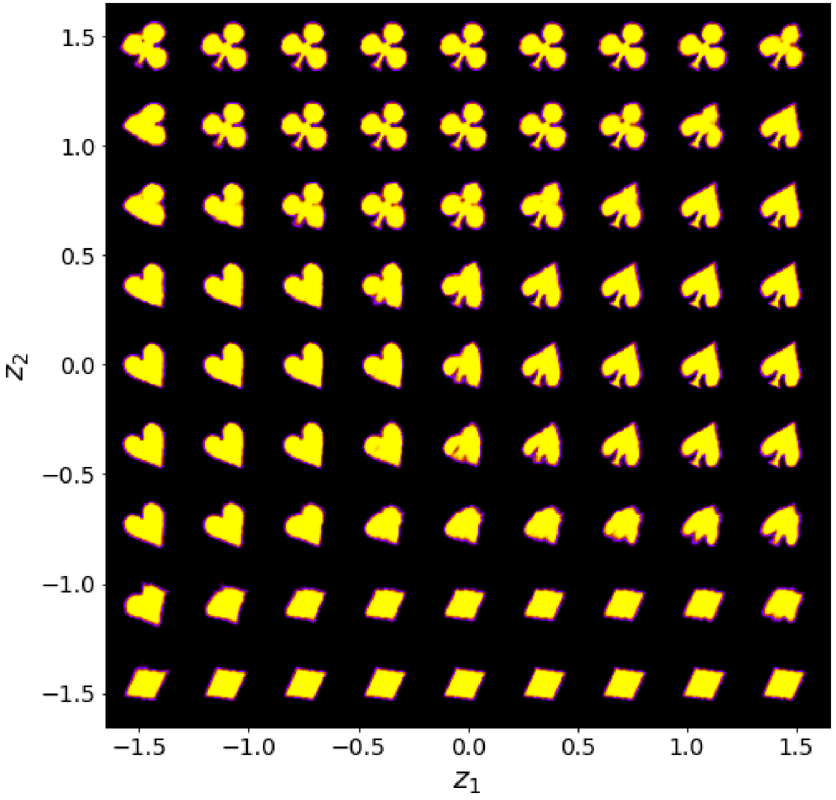
Cards 2: Low rotation (12 deg) and high shear (20 deg)

rVAE on Cards

Example of data

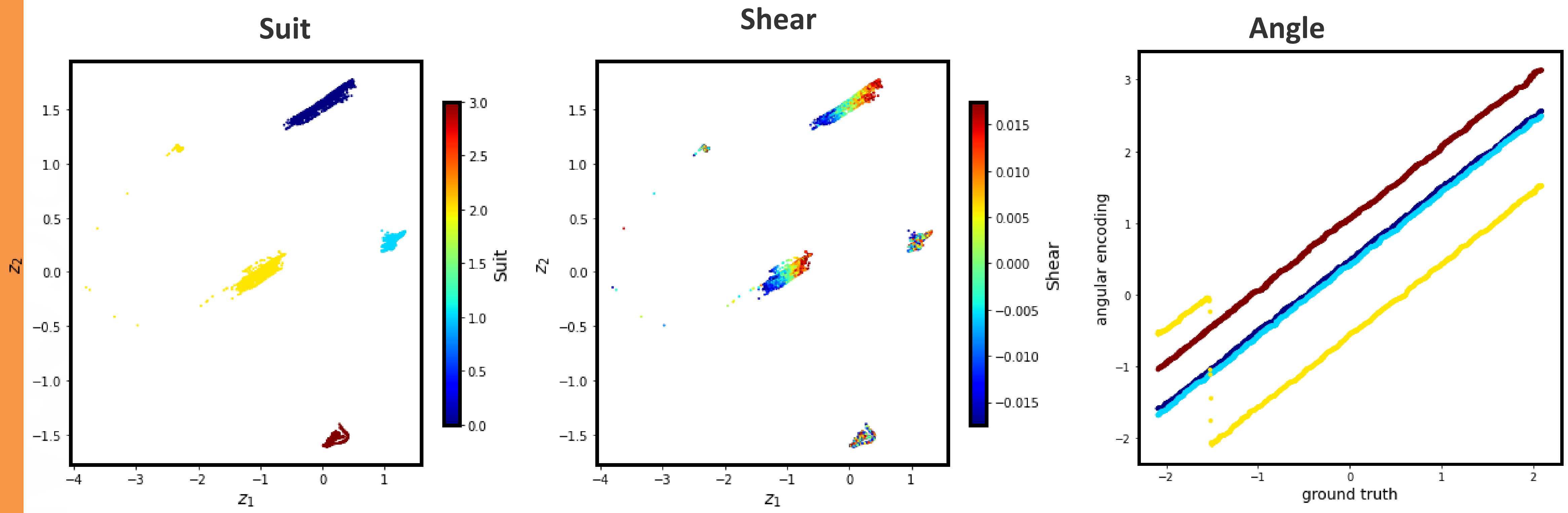


Latent representation



Cards 3: High rotation (120 deg) and low shear (1 deg)

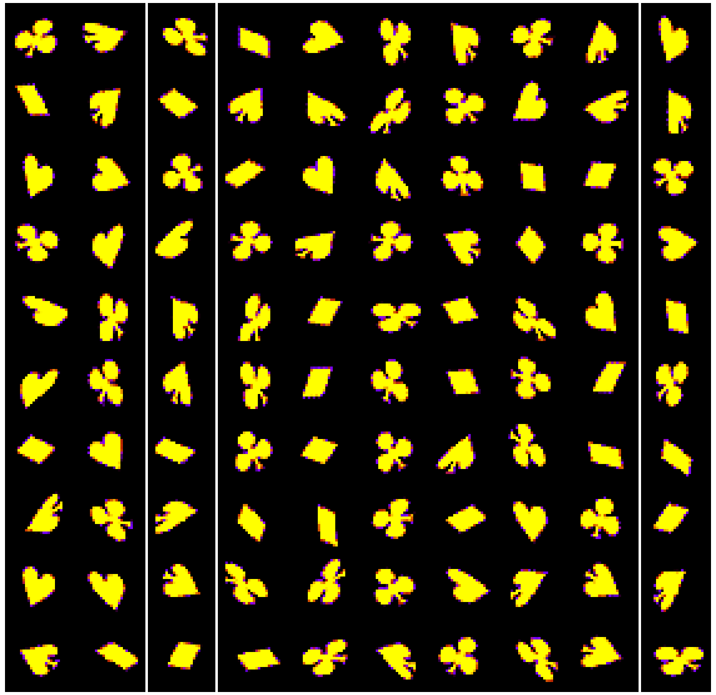
rVAE on Cards



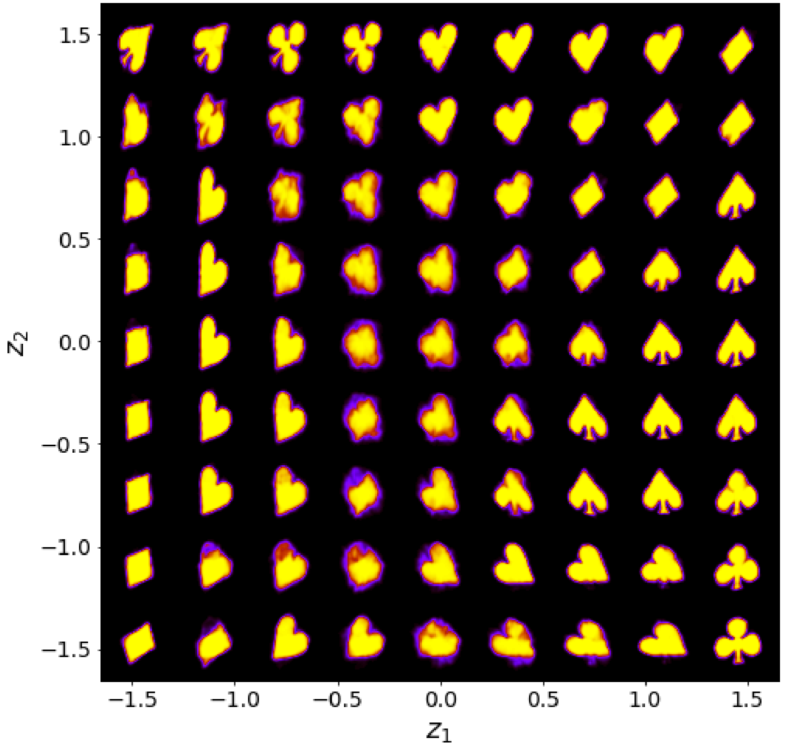
Cards 3: High rotation (120 deg) and low shear (1 deg)

rVAE on Cards

Example of data

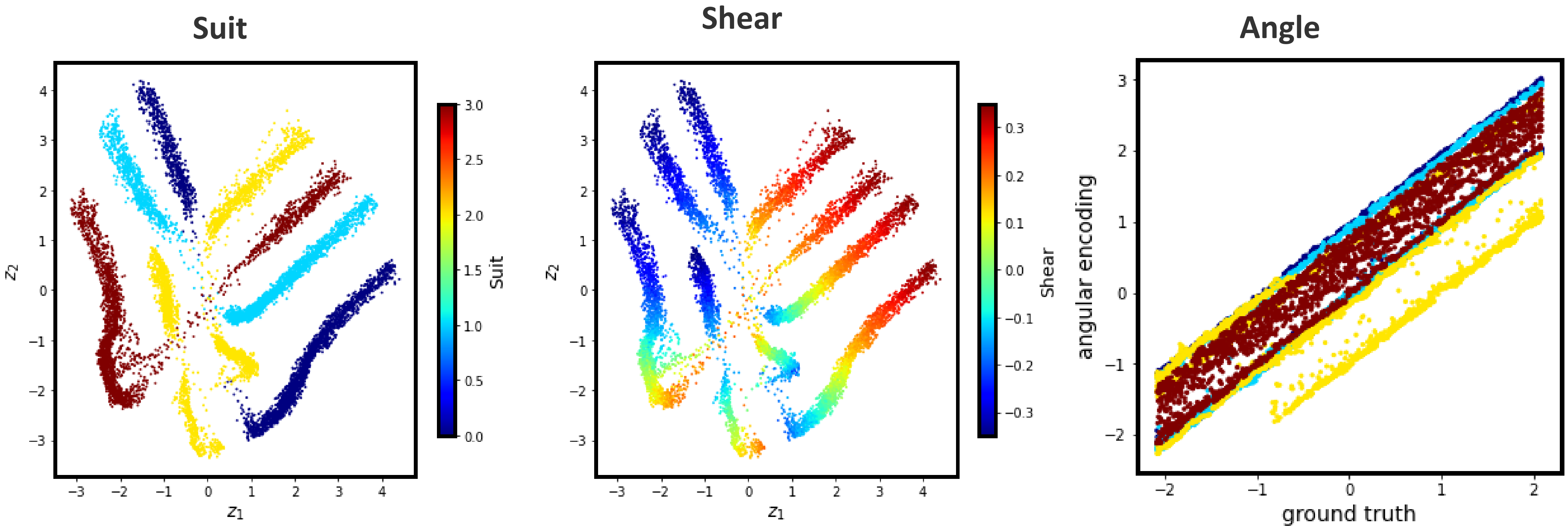


Latent representation



Cards 4: High rotation (120 deg) and high shear (20 deg)

rVAE on Cards



Cards 4: High rotation (120 deg) and high shear (20 deg)

What can (unsupervised) classification give us

- Our research deals with complex data sets containing 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 approaches to reduce dimensionality and explore similarities in these data sets.

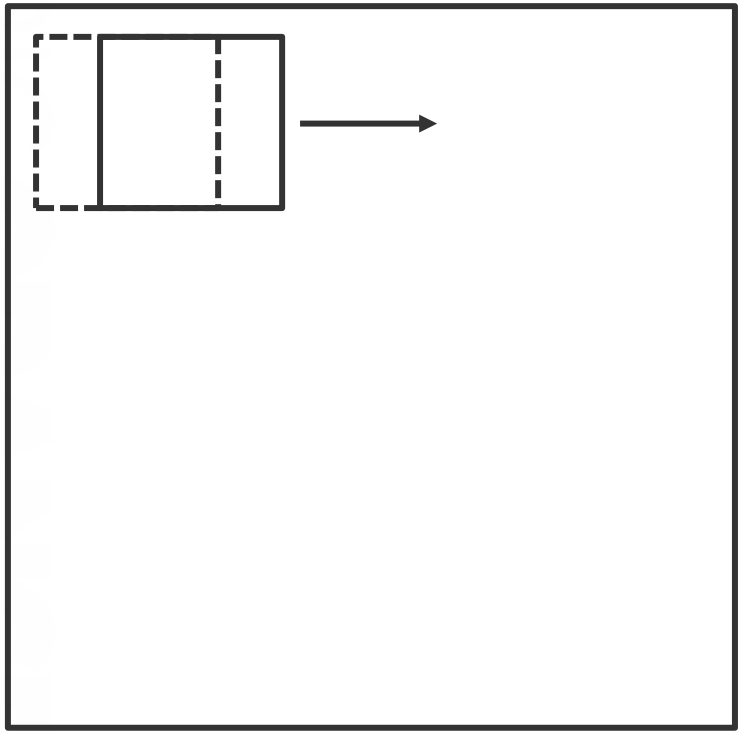
- When working with such data sets, two things matter: descriptors and ML method
- In analysis of EELS or CITS data, very often our descriptor is just the spectrum at each pixel. Typical analysis will be either linear or non-linear dimensionality reduction or clustering:
  - Linear dimensionality reduction: 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.

- What about images?

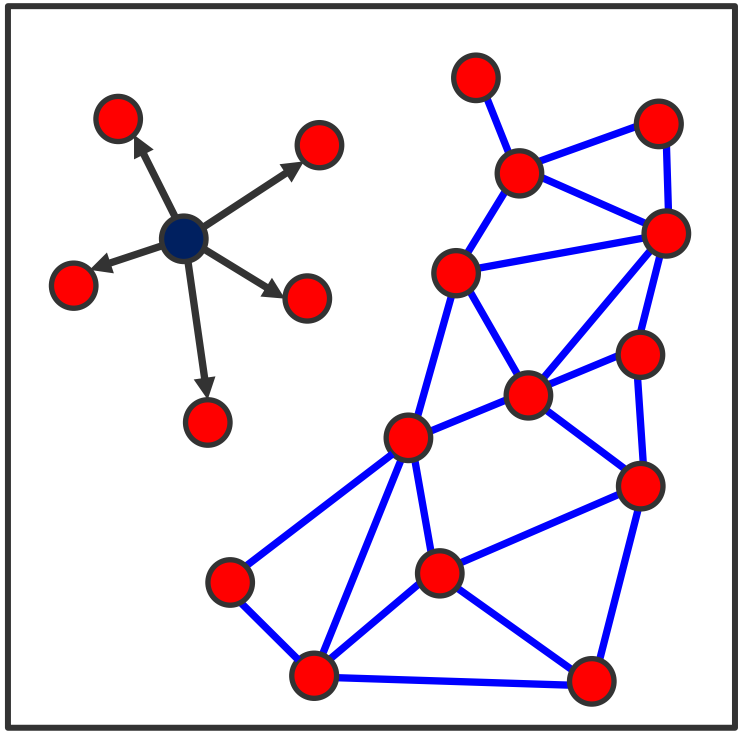
Describing the building blocks

- 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

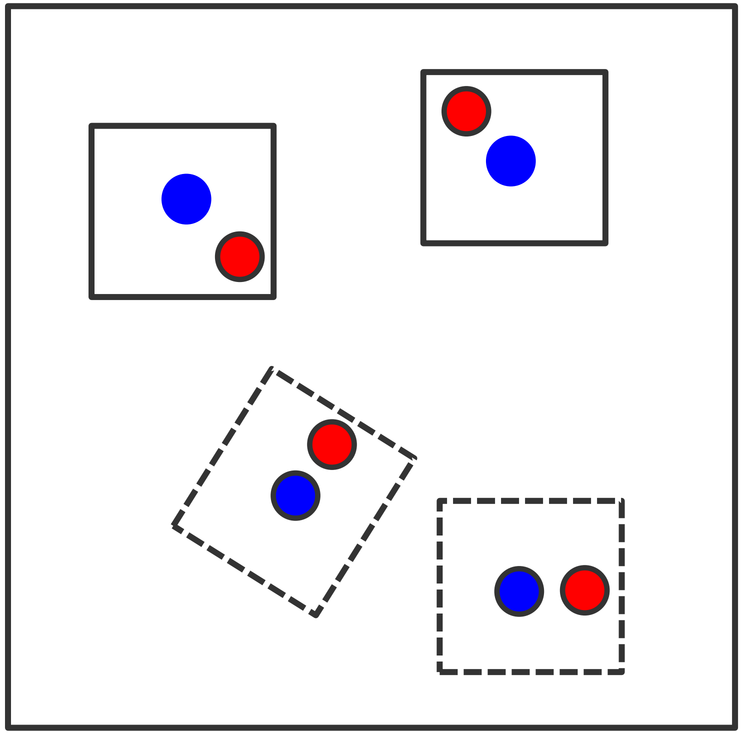
Continuous translational symmetry



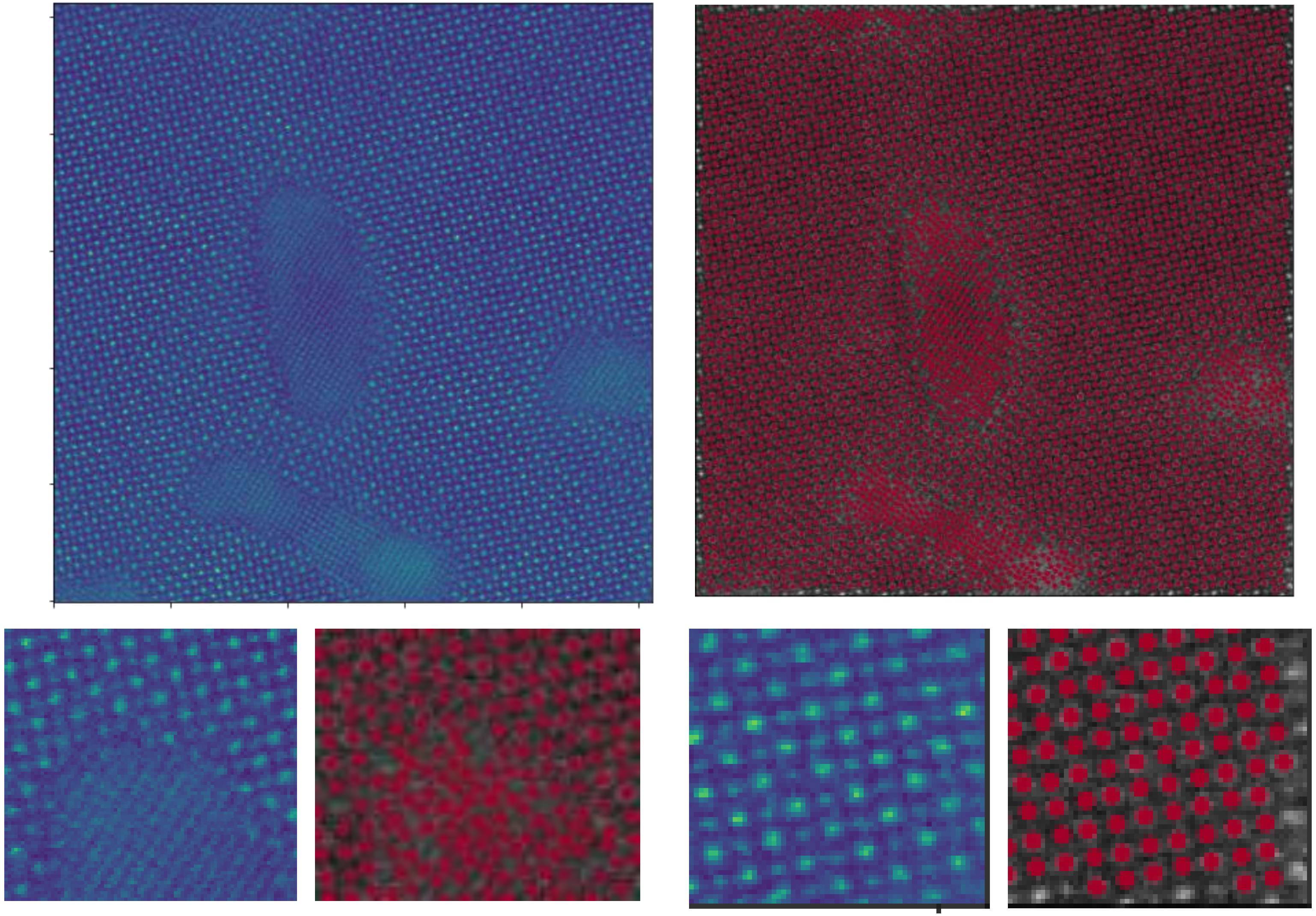
Atom based descriptions



Localized sub-images

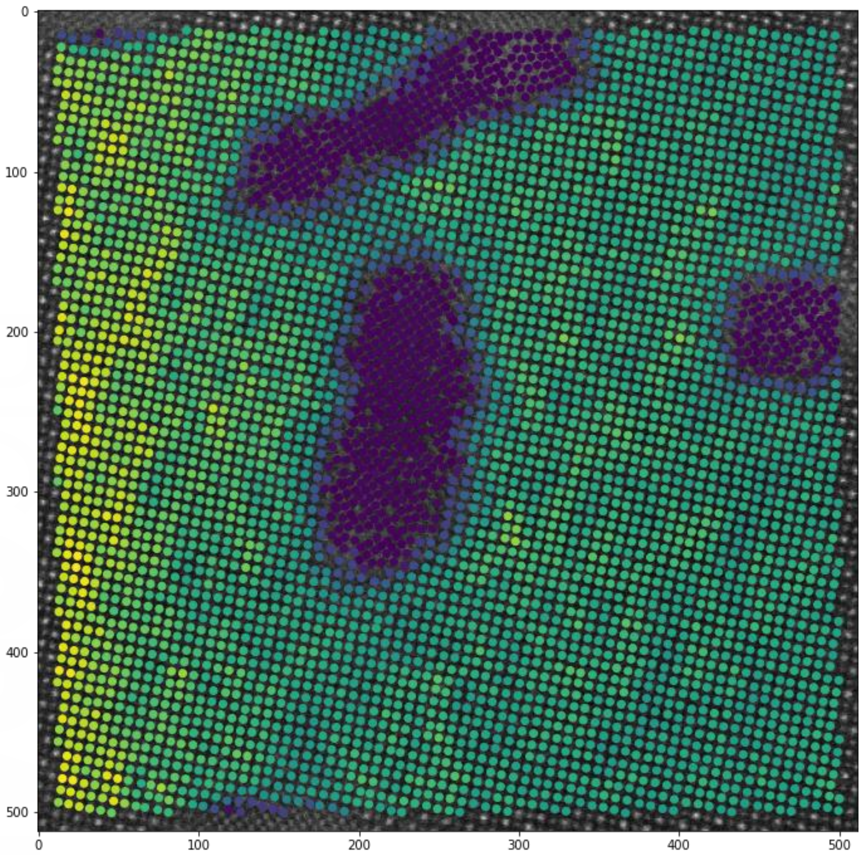


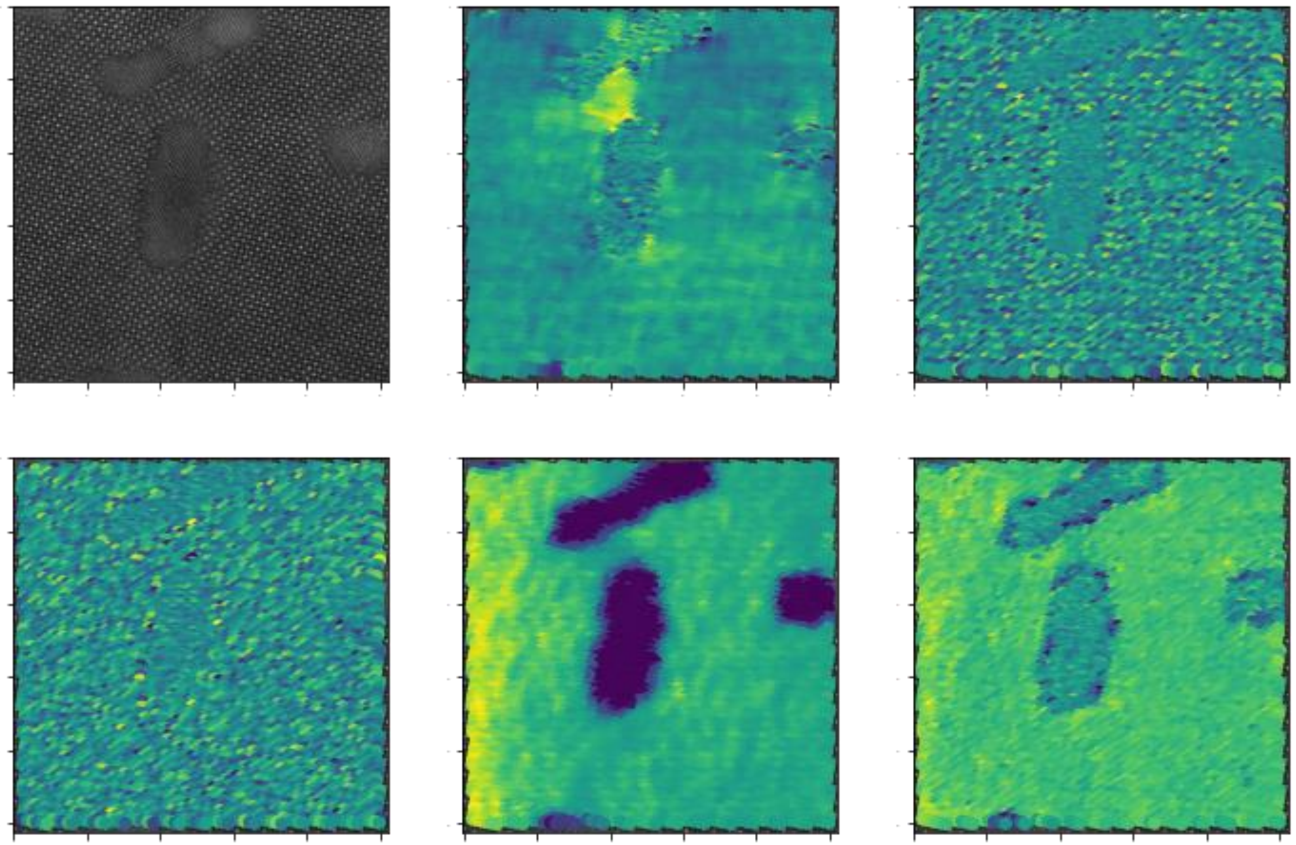
Let's put it all together!



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 came from

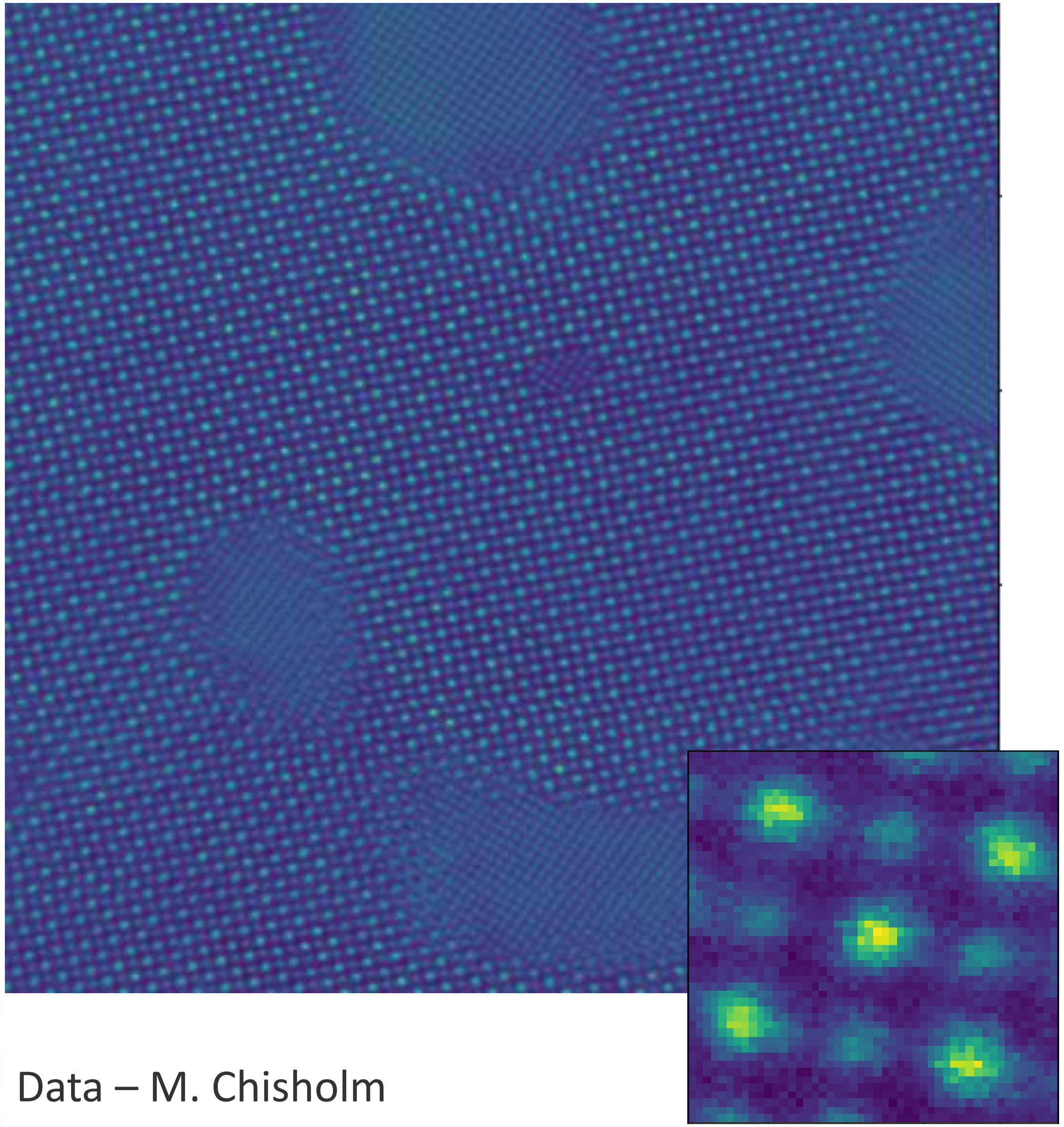
Step 3: rVAE

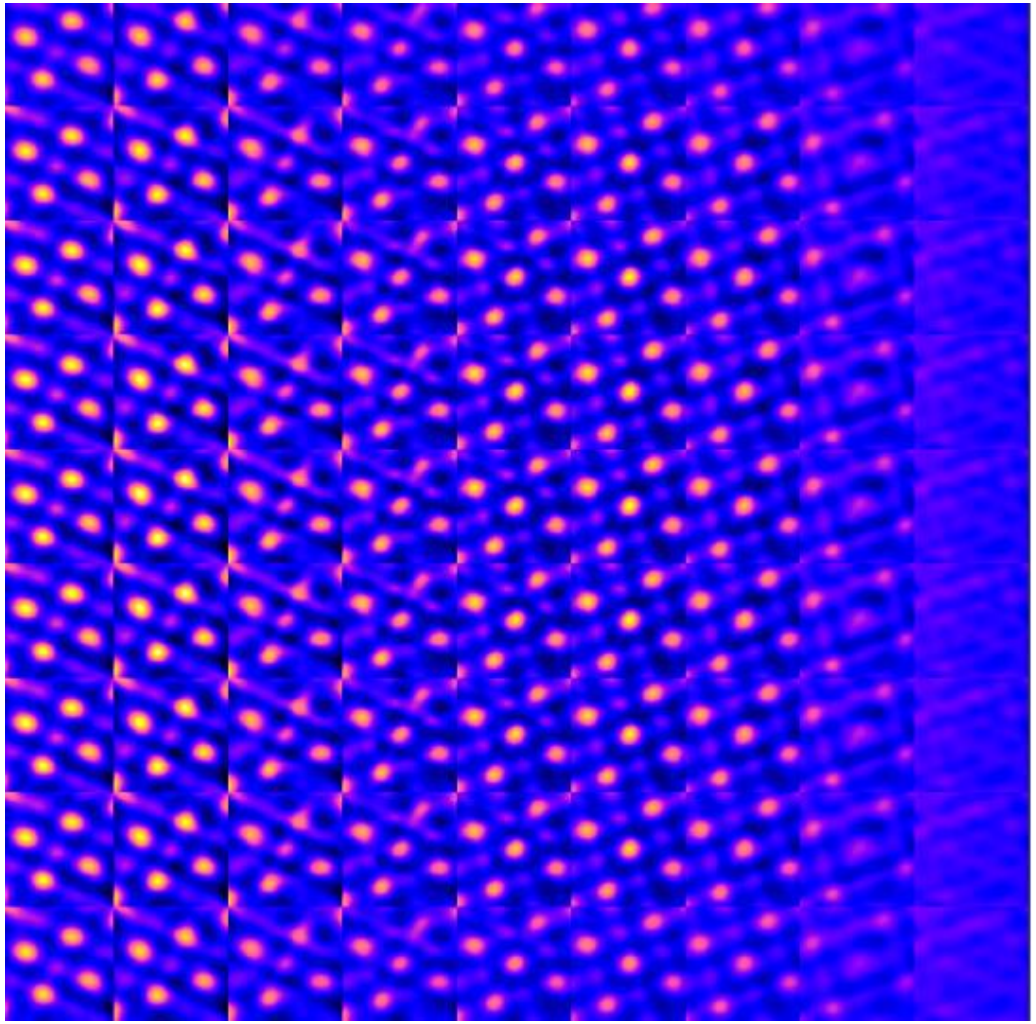




Output: Latent variable corresponding to local structure of each atomic site. Can be visualized on top of the original atomically resolved image, or as 2D maps (but – not rectangular array!)

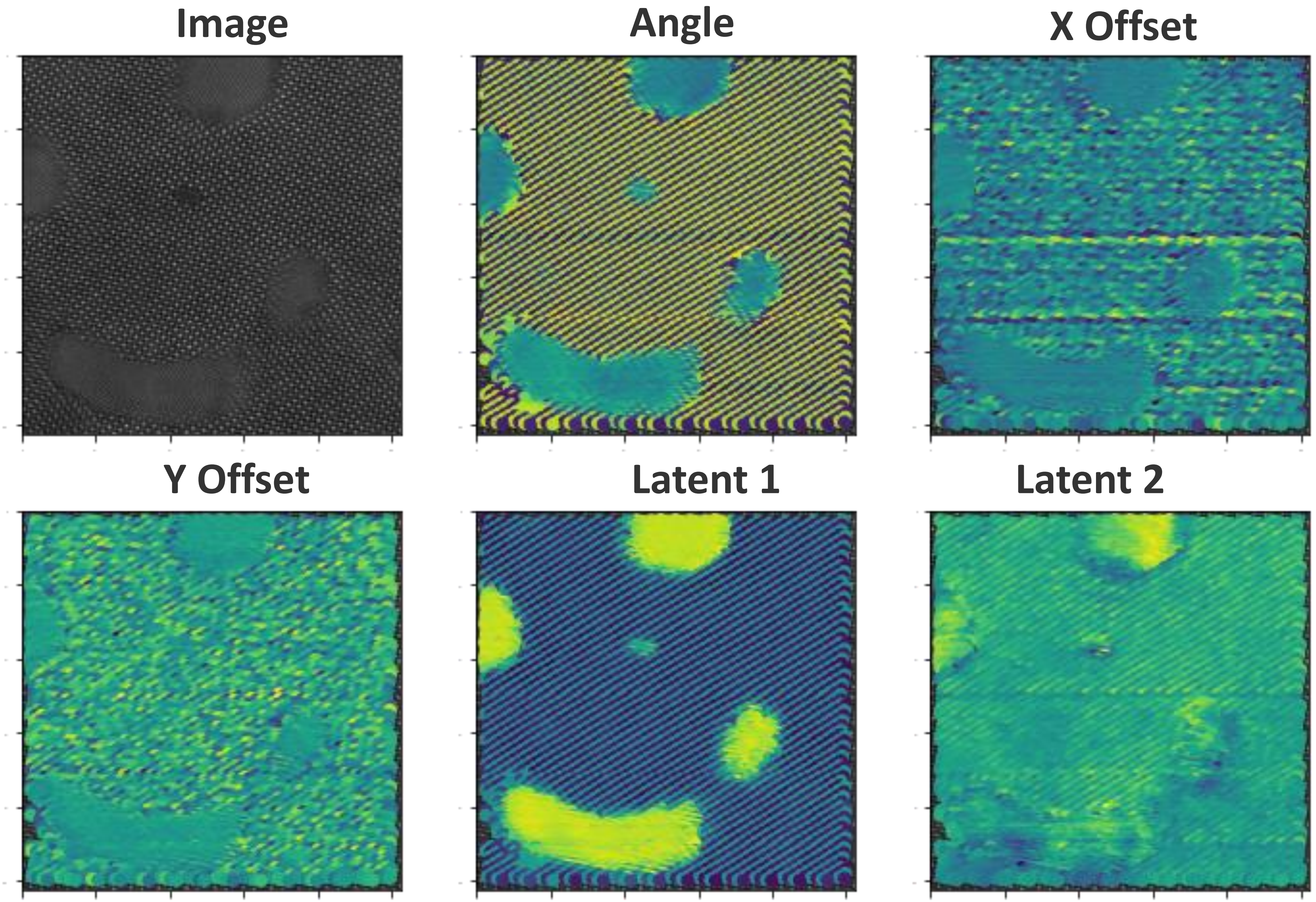
Analysis of the NiO-LSMO



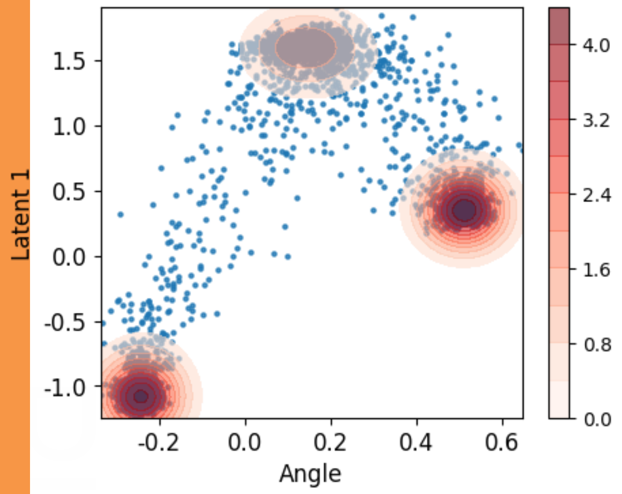


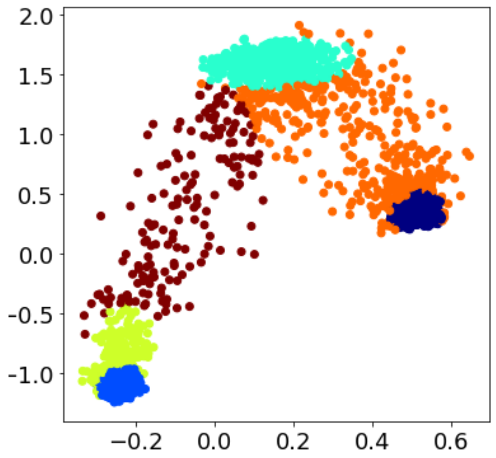
Data – M. Chisholm

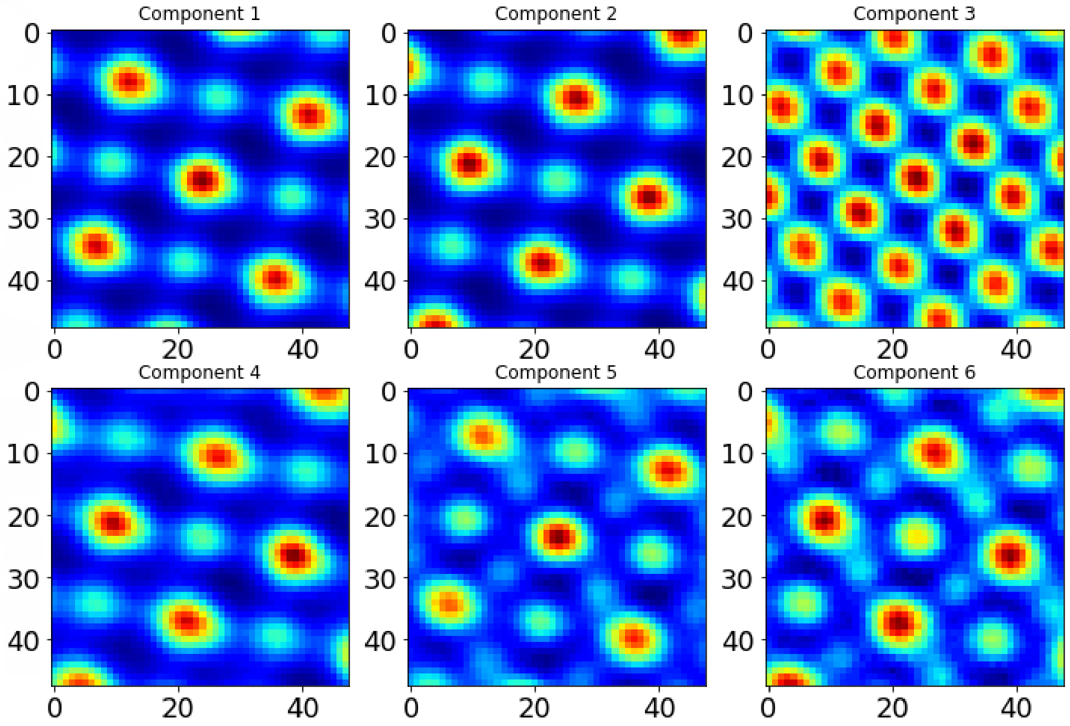
Let's look at latent space



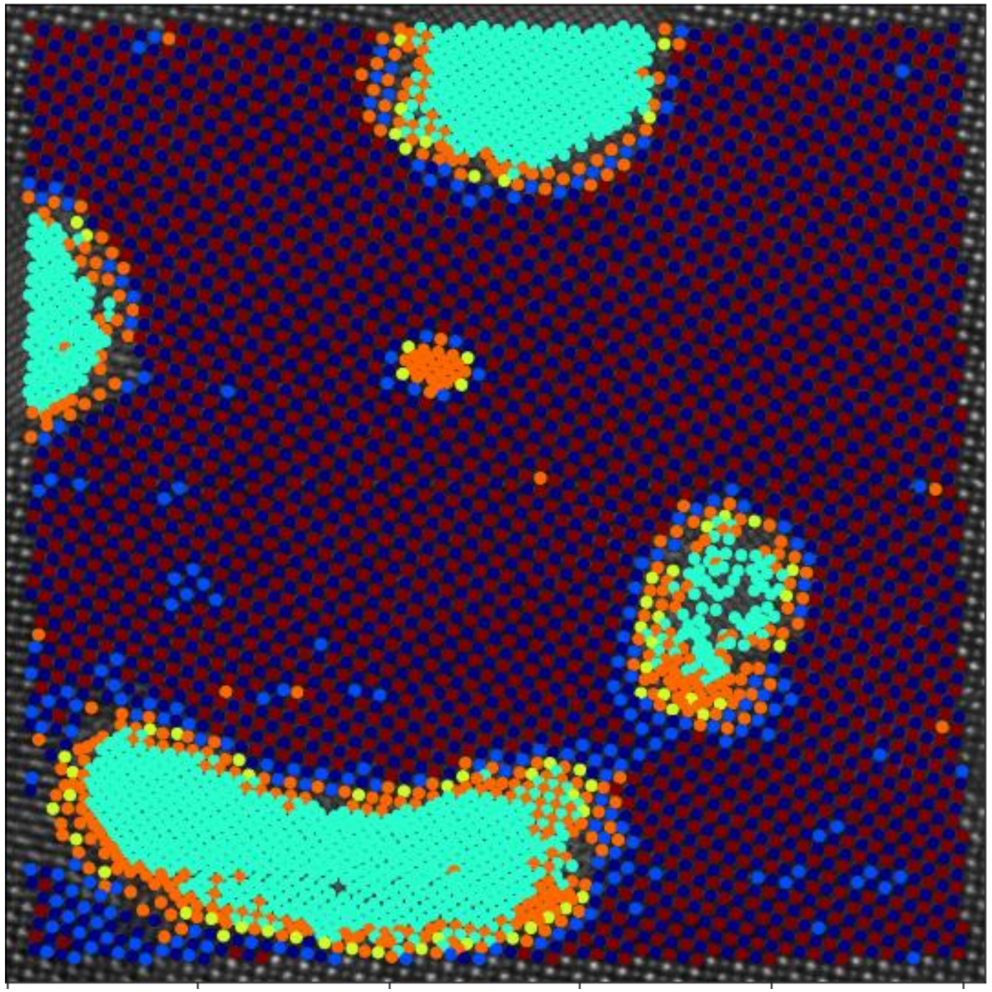
Exploring latent distributions







Labeled image

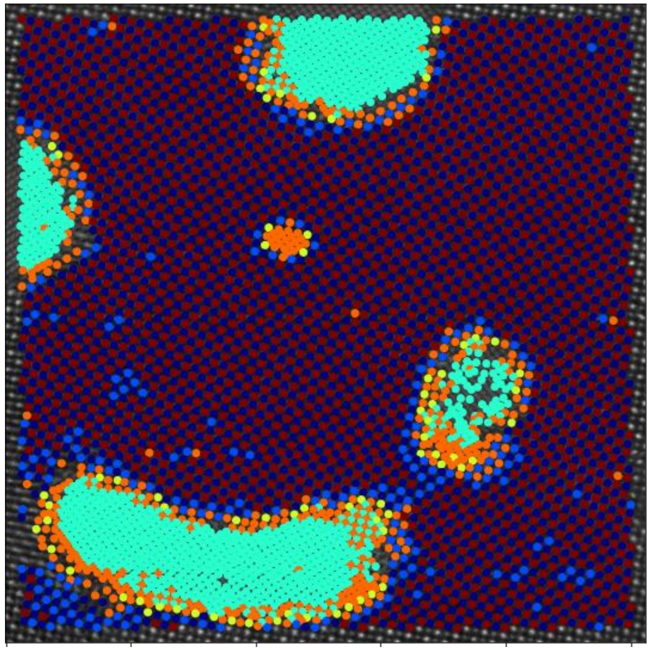


- Classes and variability are mixed in a single latent space
- Disentangling of representation

That’s where the jVAE has come from

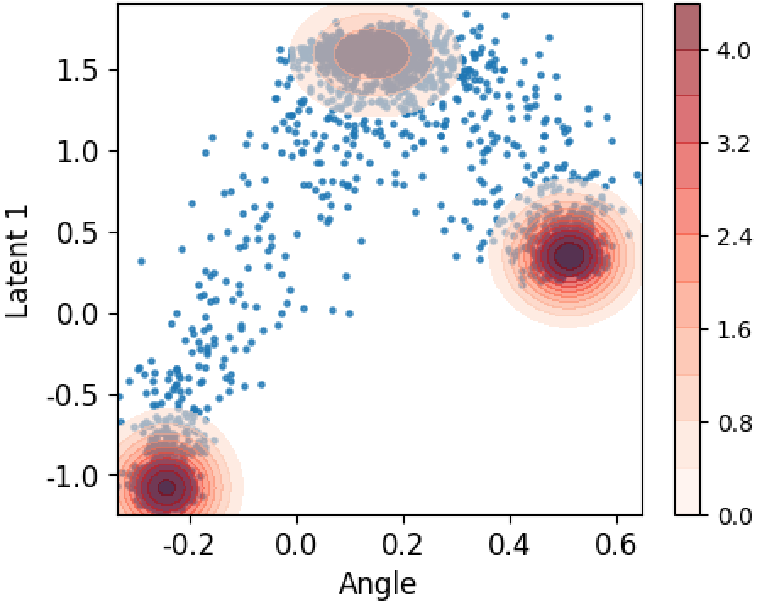
Currently, we have variants of invariant VAE that include:

- Convolutional or dense layers (reconfigurable via **kwargs)
- Rotational invariance
- With and without offsets (as latent variables)
- Multilayer inputs

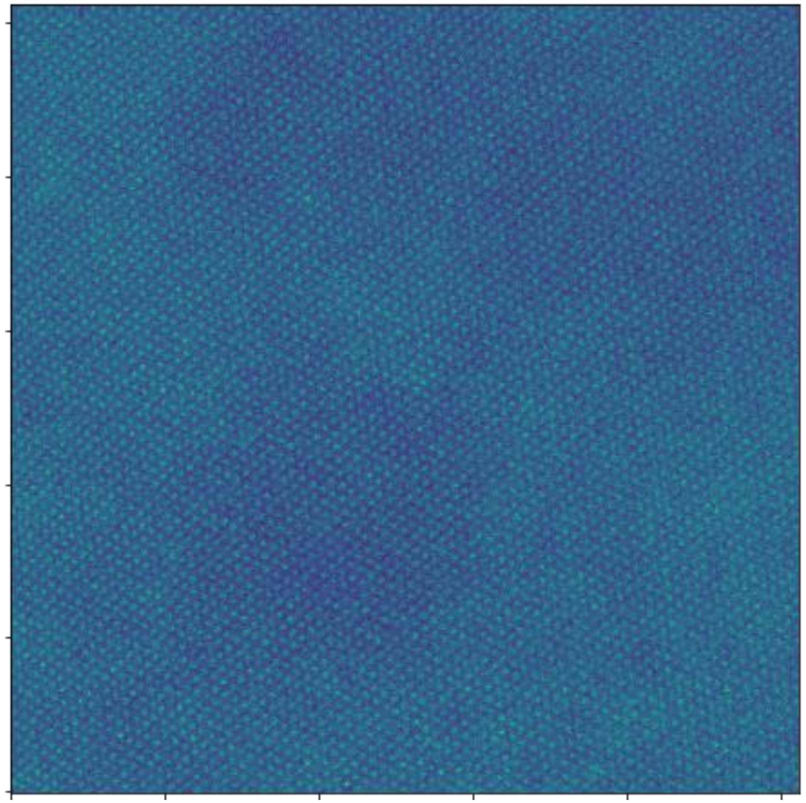


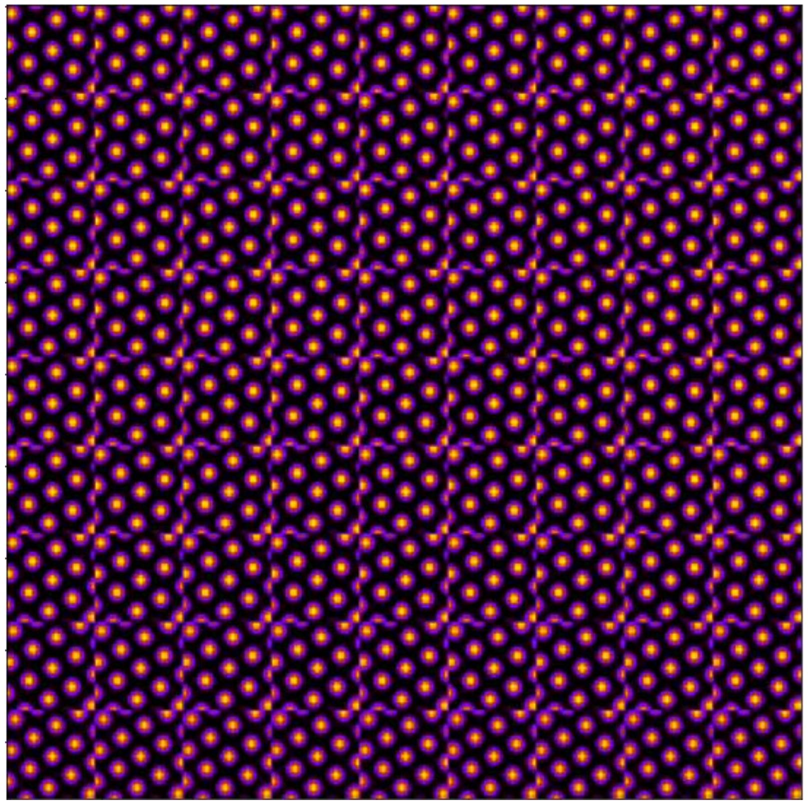
However, our rVAE collects everything in a single latent space. Realistically, very often we deal with system where we expect the presence of finite number of classes that may be known, partially known, or unknown, with certain continuous traits within classes.

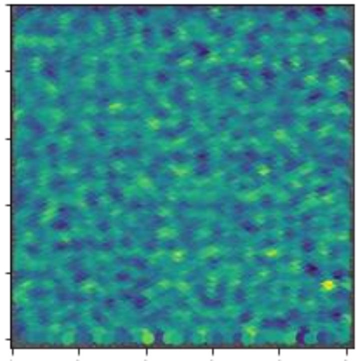
- **Graphene and MX2:** structural units (discrete) and strain states
- **Crystalline solids:** phases and ferroic variants, strain states
- **Plasmonic EELS:** particle spectra, off-particle spectra, edge states
- **CITS:** lattice and defects, strain states

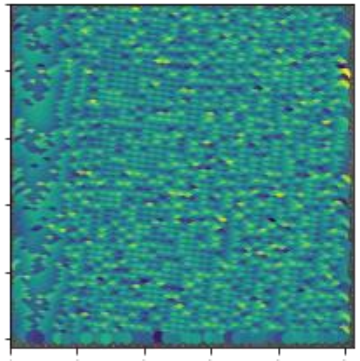


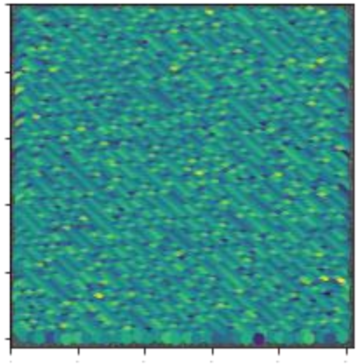
Out of curiosity: single crystal?

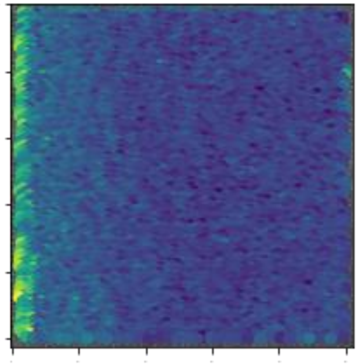


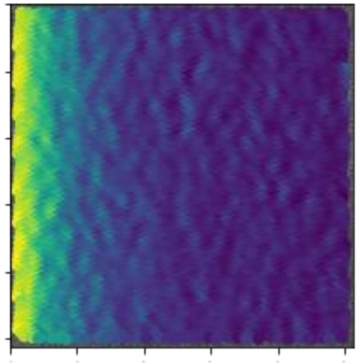






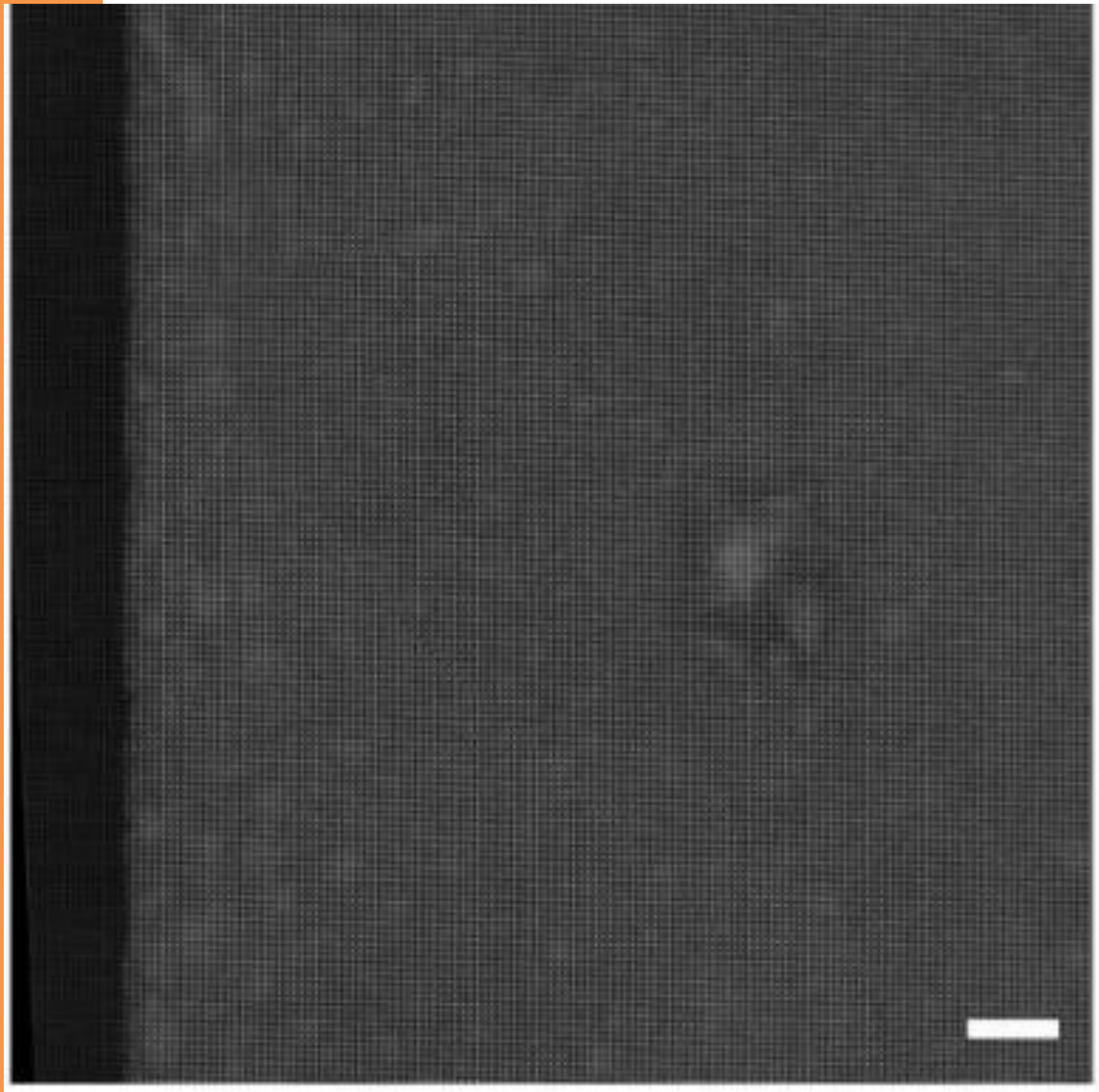




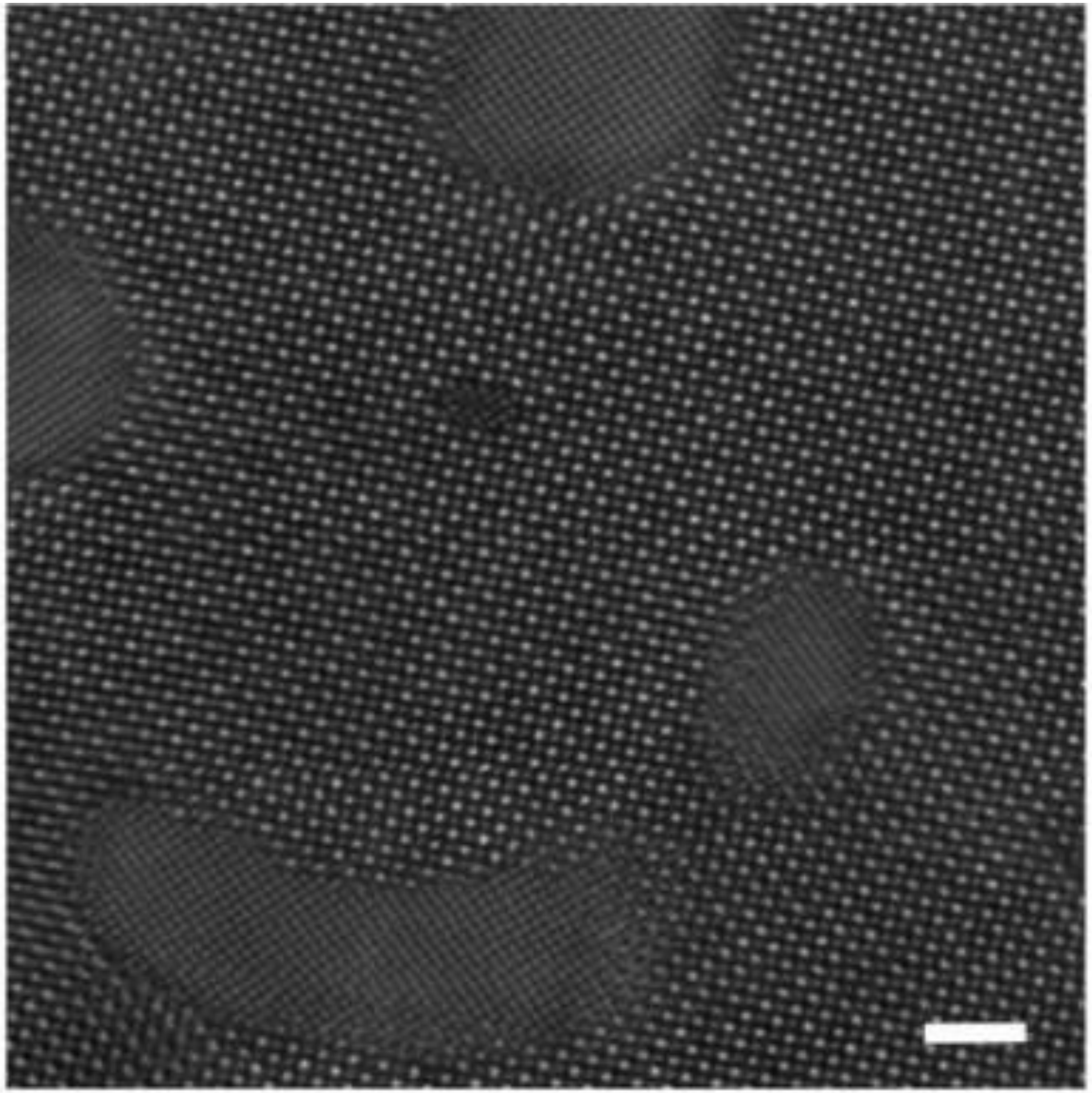


VAE without Atom Finding

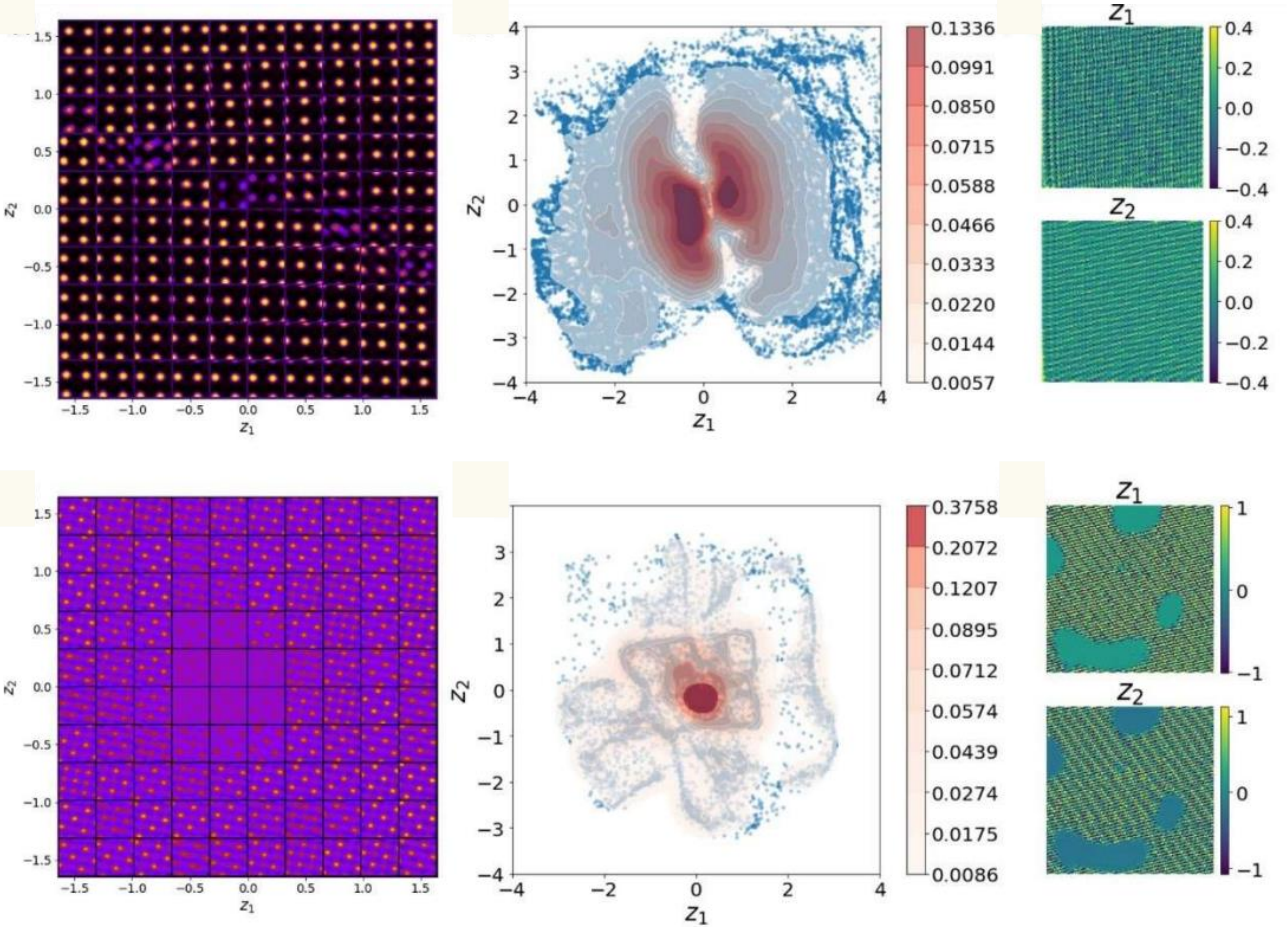
Ferroelectric $BiFeO_3$



$NiO – La_xSr_{1-x}MnO_3$

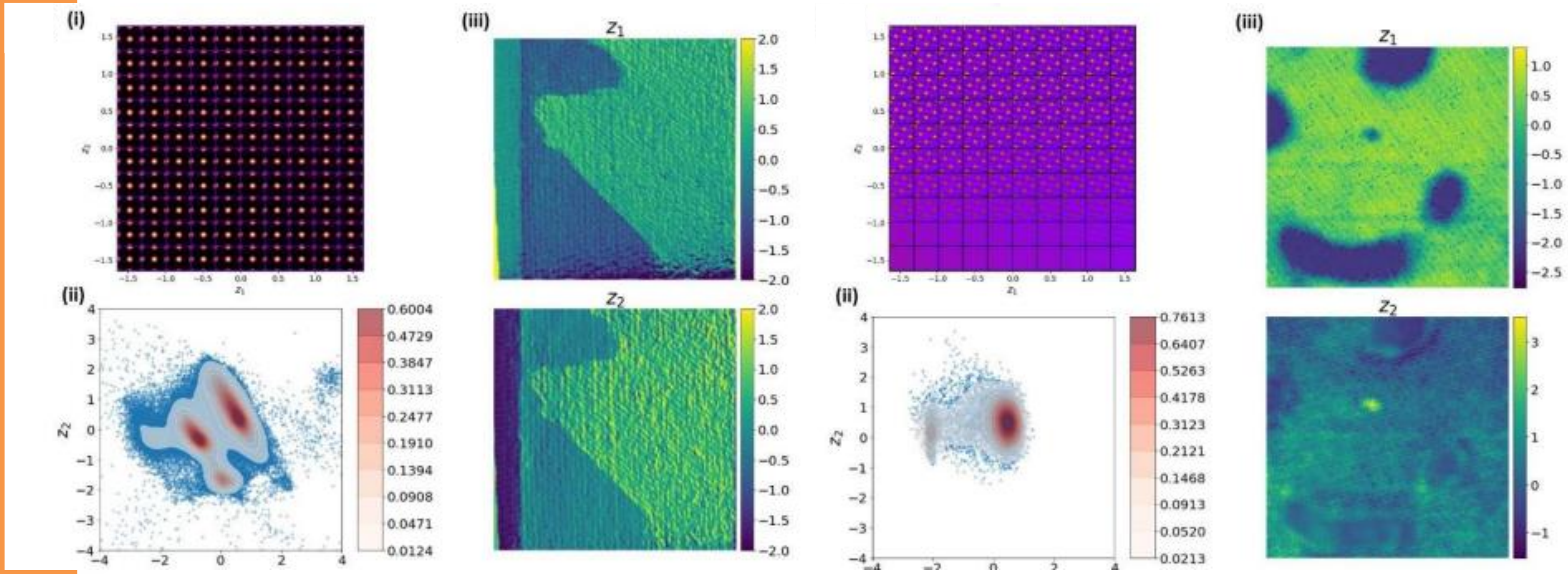


Simple VAE

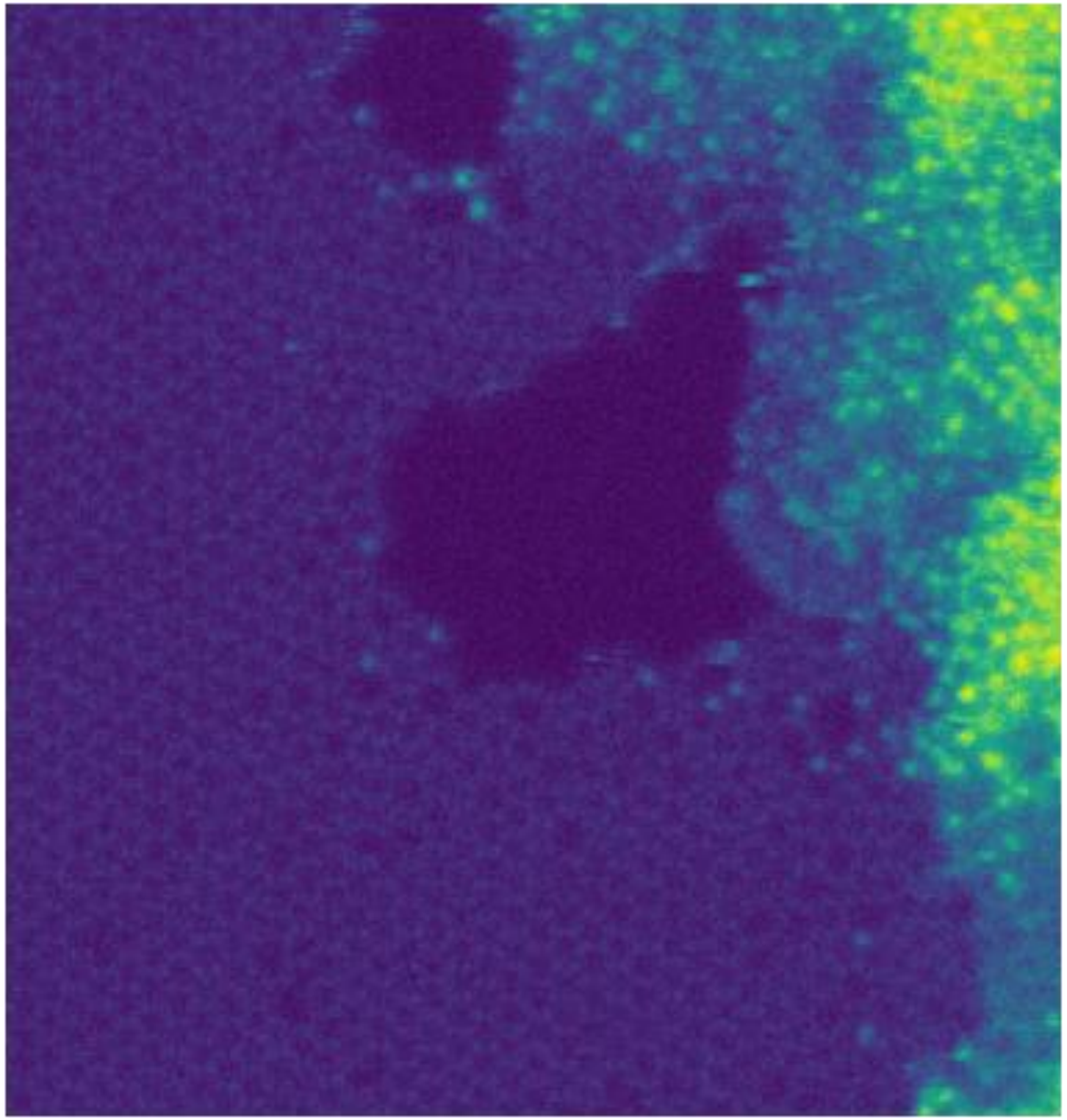


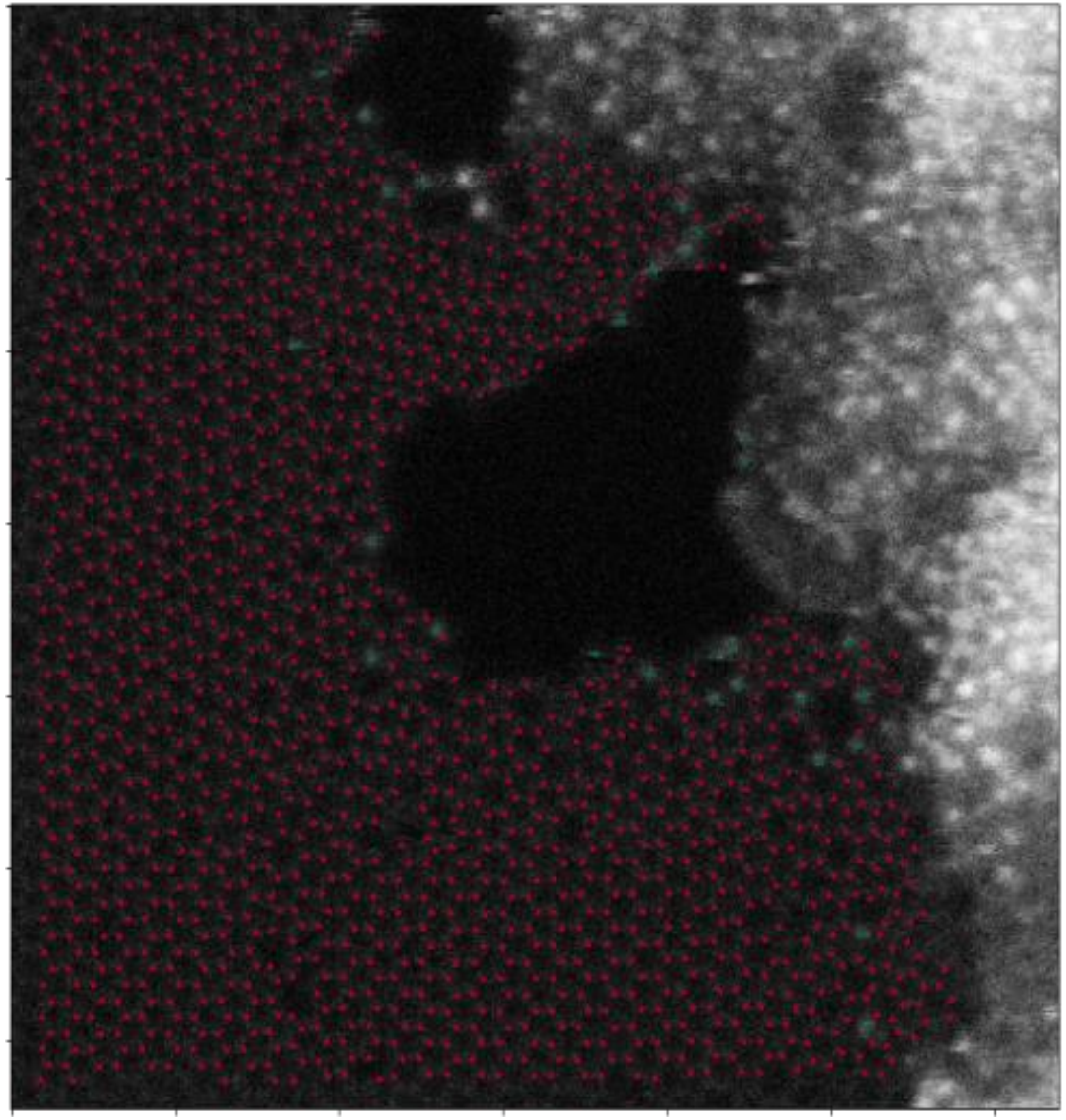
z_1
z_2
0.1336
0.0991
0.0850
0.0715
0.0588
0.0466
0.0333
0.0220
0.0144
0.0057
0.3758
0.2072
0.1207
0.0895
0.0712
0.0574
0.0439
0.0274
0.0175
0.0086

Shift VAE: Translational Invariance

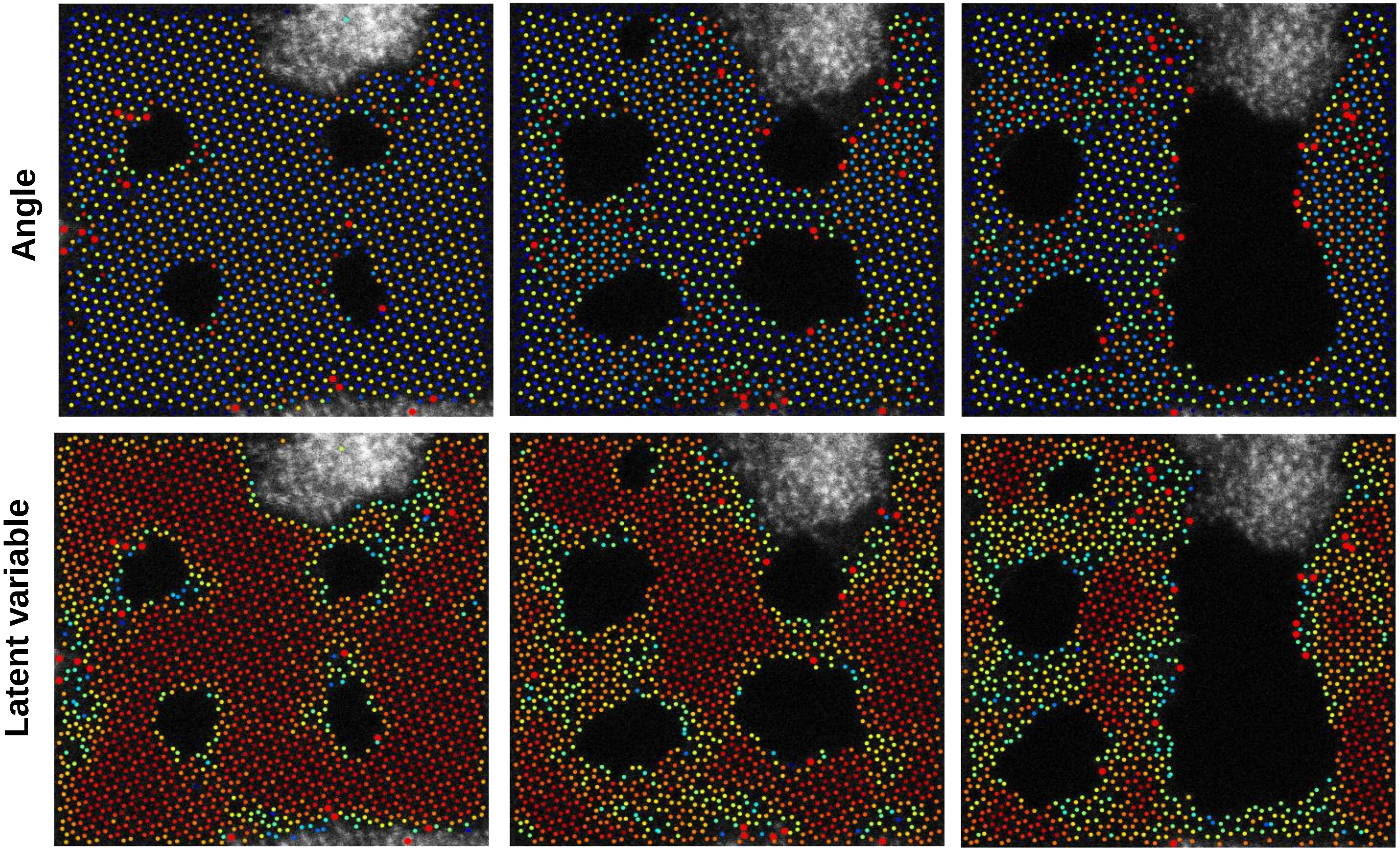


Off to chemically-disordered systems

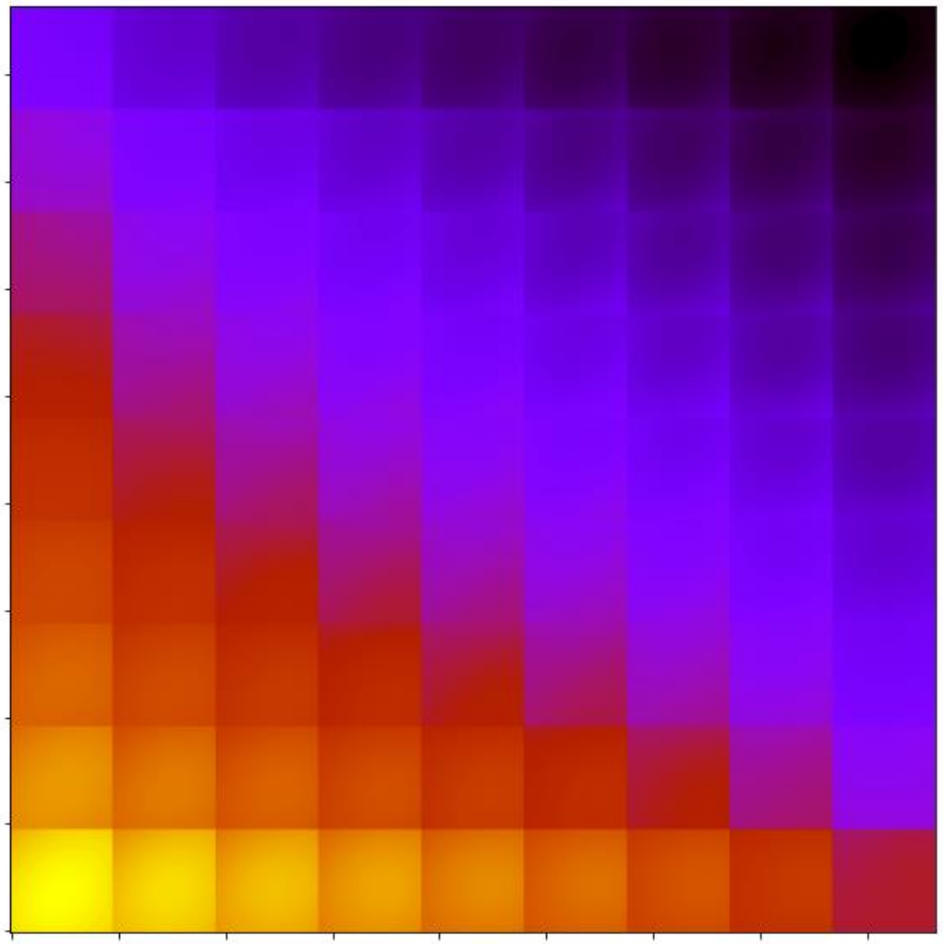


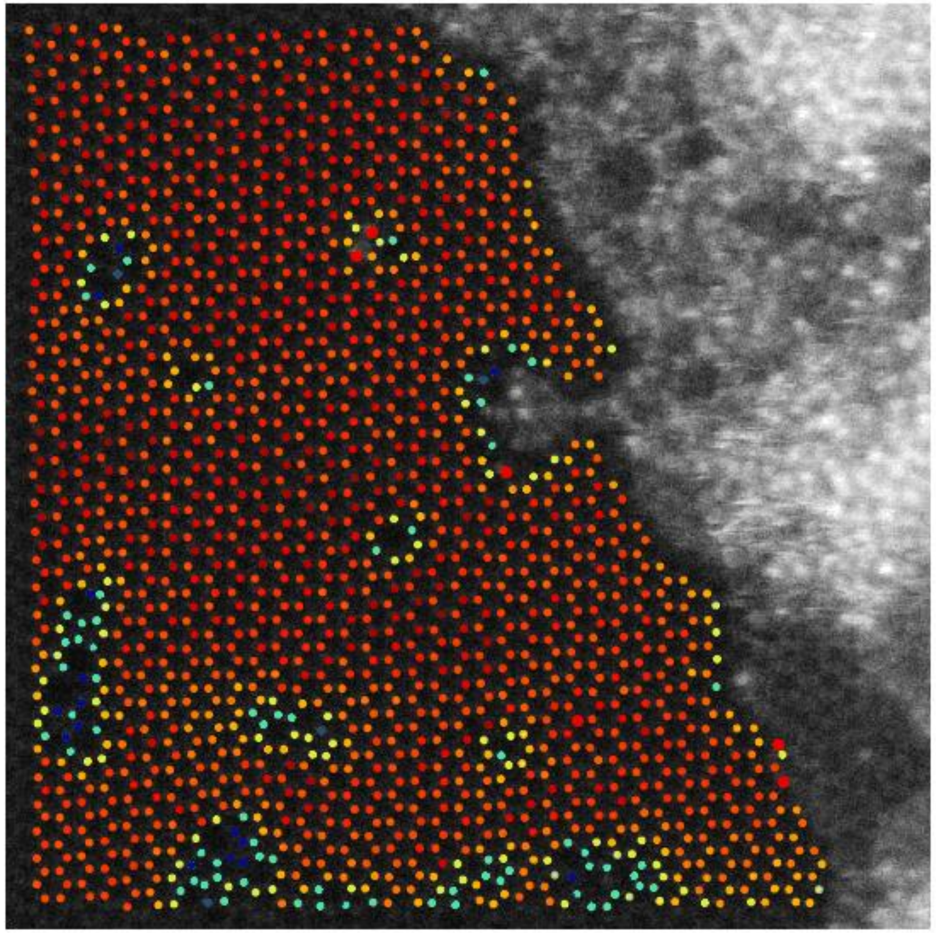


rVAE analysis at different time steps

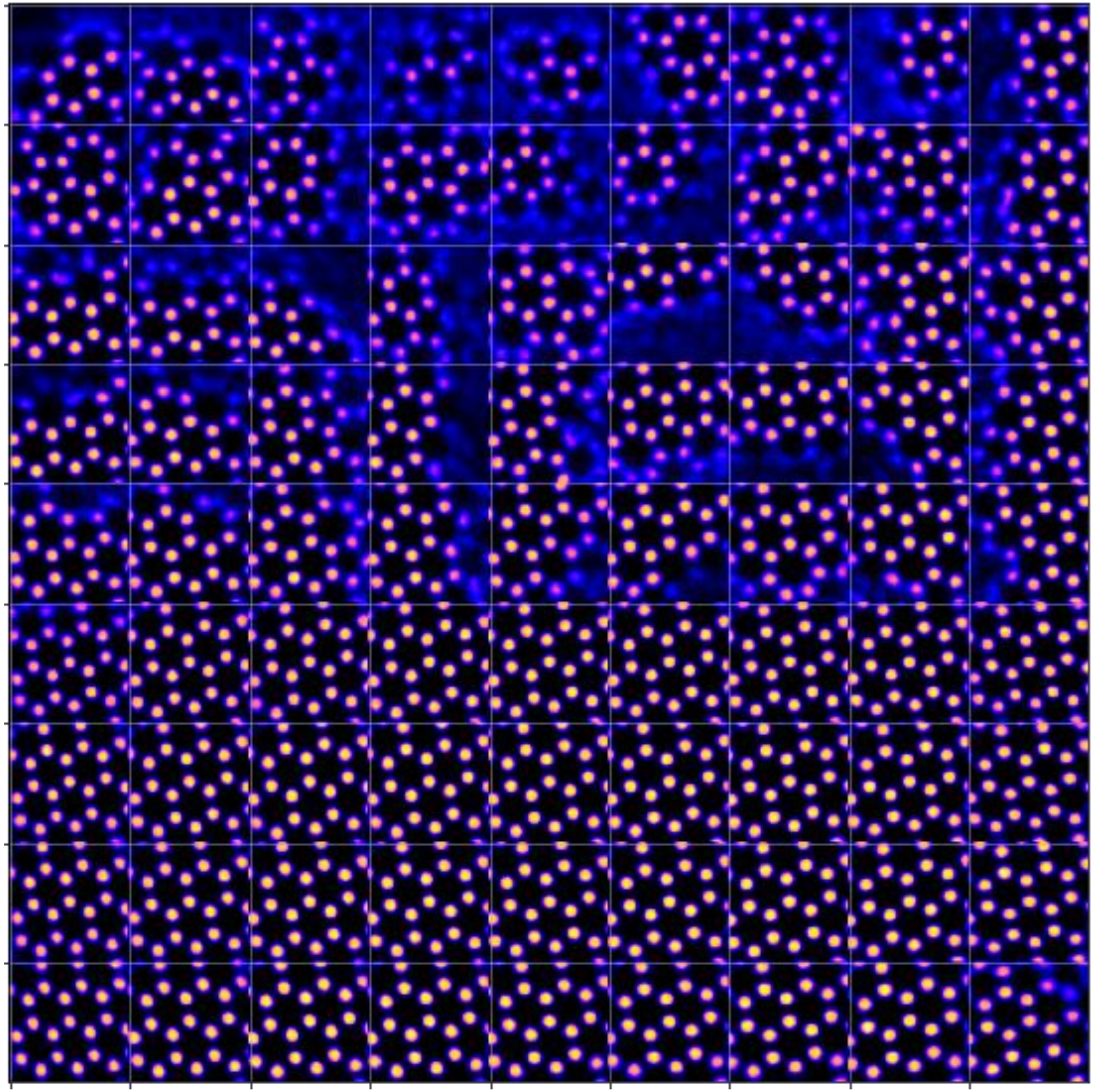


There is nothing as beautiful as training VAE

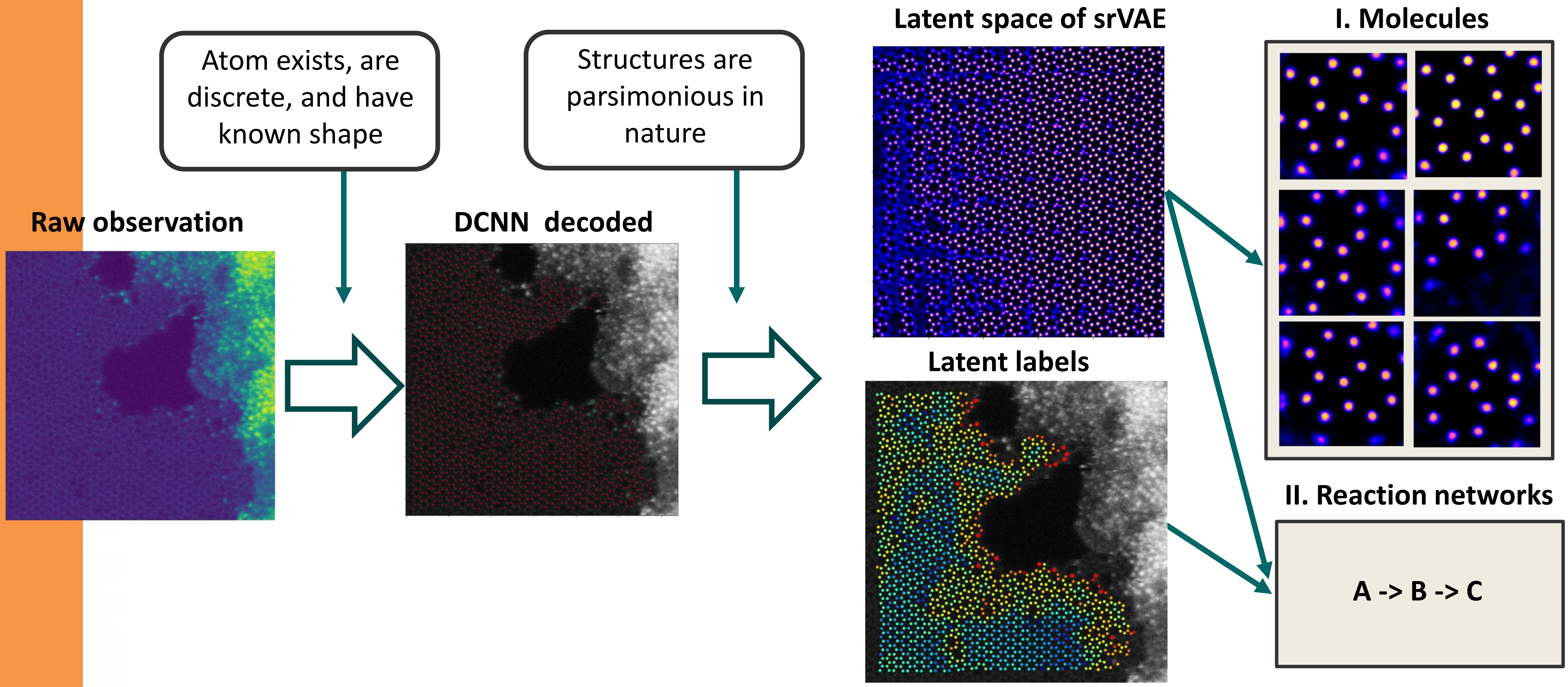




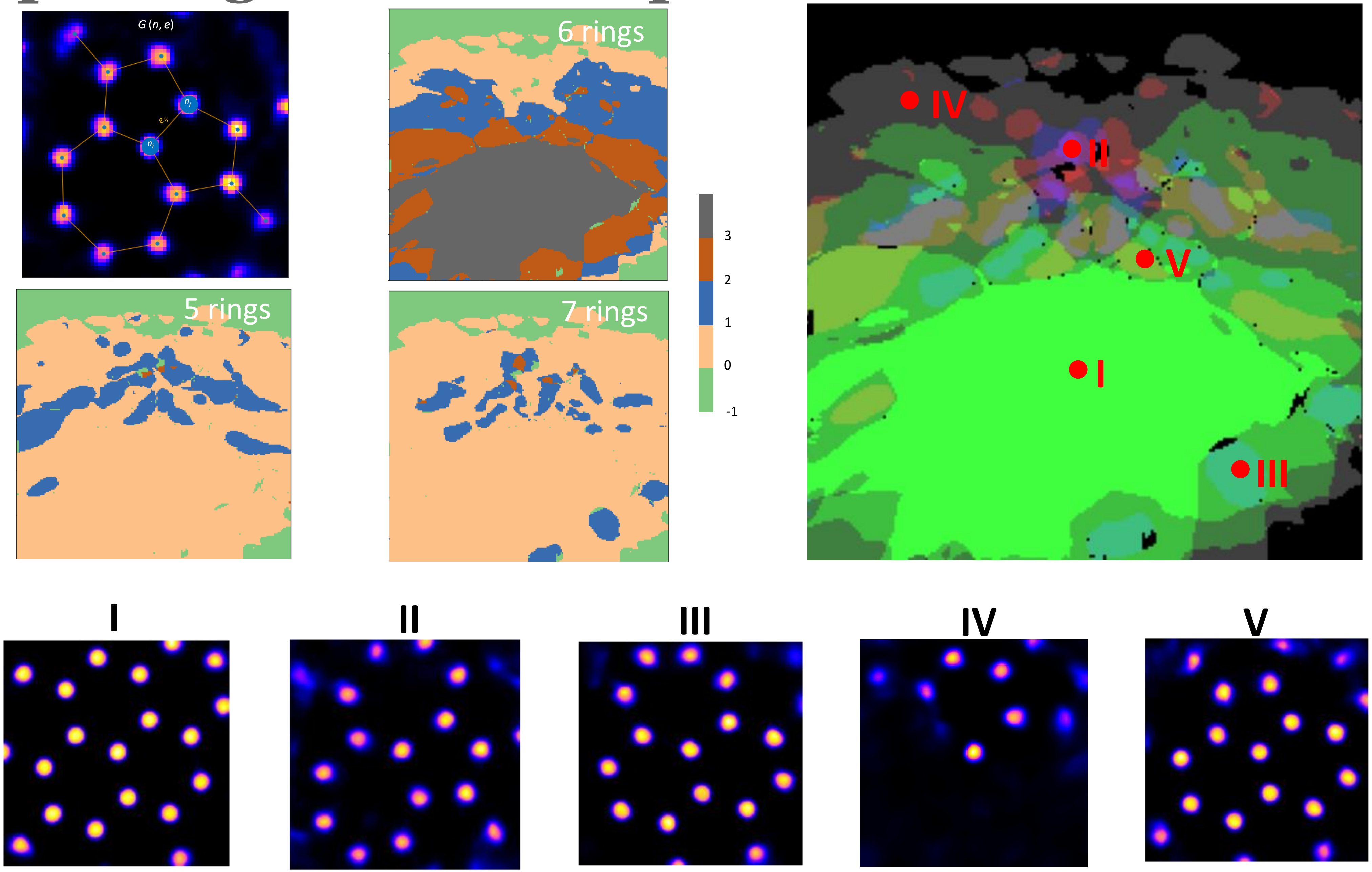
Next step: skip-rVAE



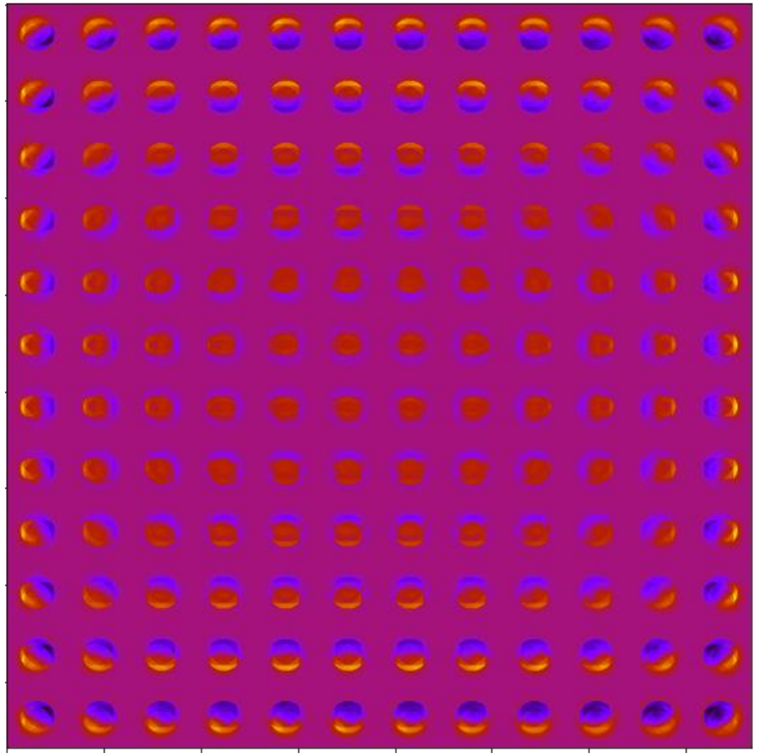
Unsupervised discovery of molecules



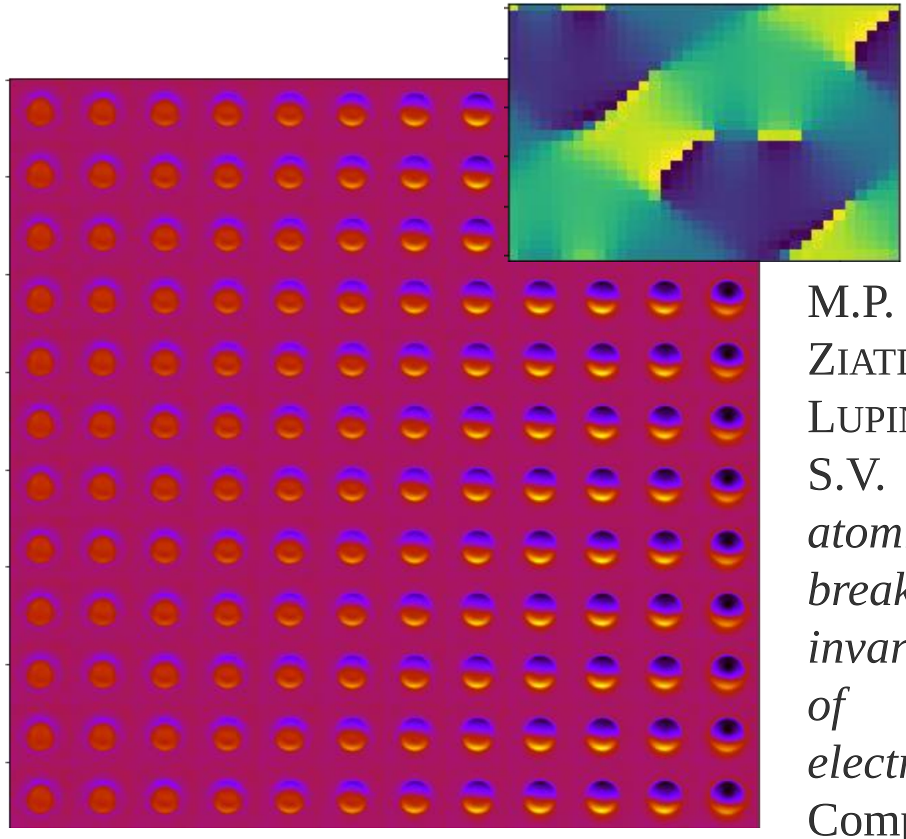
Exploring the latent space structure



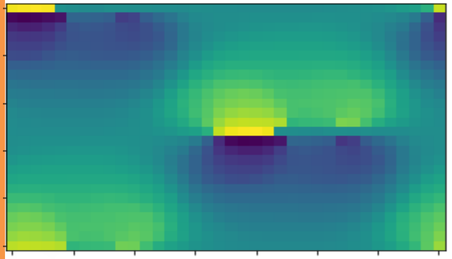
Simple VAE vs. rVAE

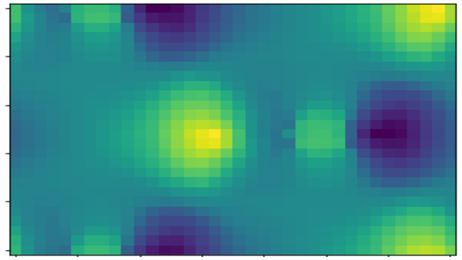


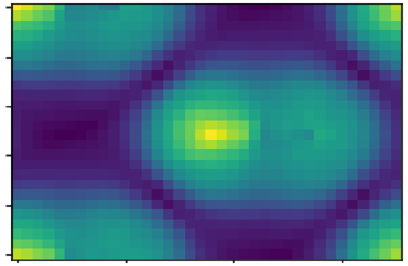
Angle

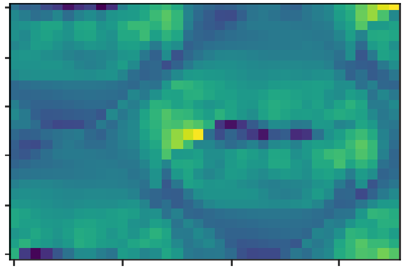


Latent 1 | Latent 2 | Latent 1 | Latent 2







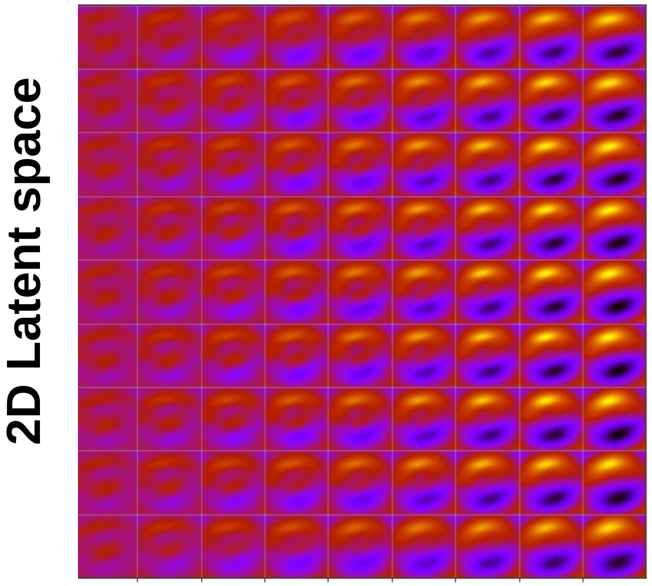


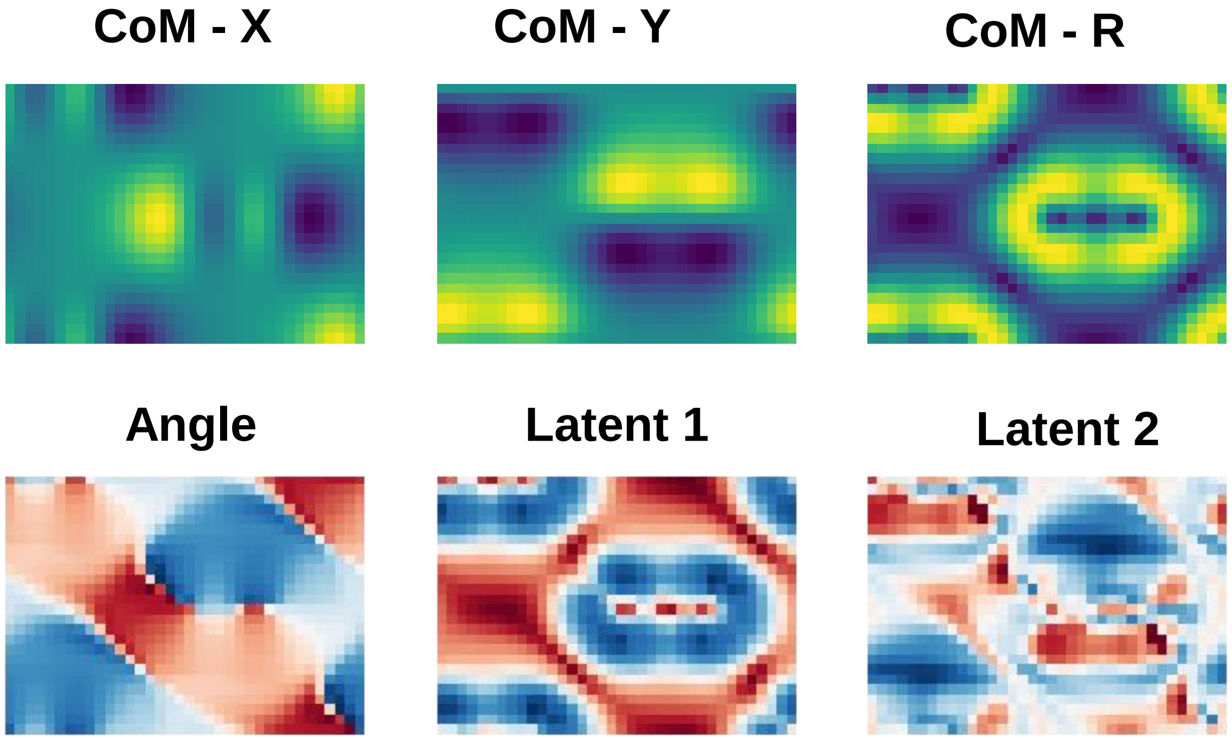
M.P. OXLEY, M. ZIATDINOV, O. DYCK, A.R. LUPINI, R. VASUDEVAN, and S.V. KALININ, *Probing atomic-scale symmetry breaking by rotationally invariant machine learning of multidimensional electron scattering*, npj Comp. Mat. **7**, 1 (2021) arXiv:2009.10758

Using simple VAE on the 4DSTEM data leads to suboptimal encoding

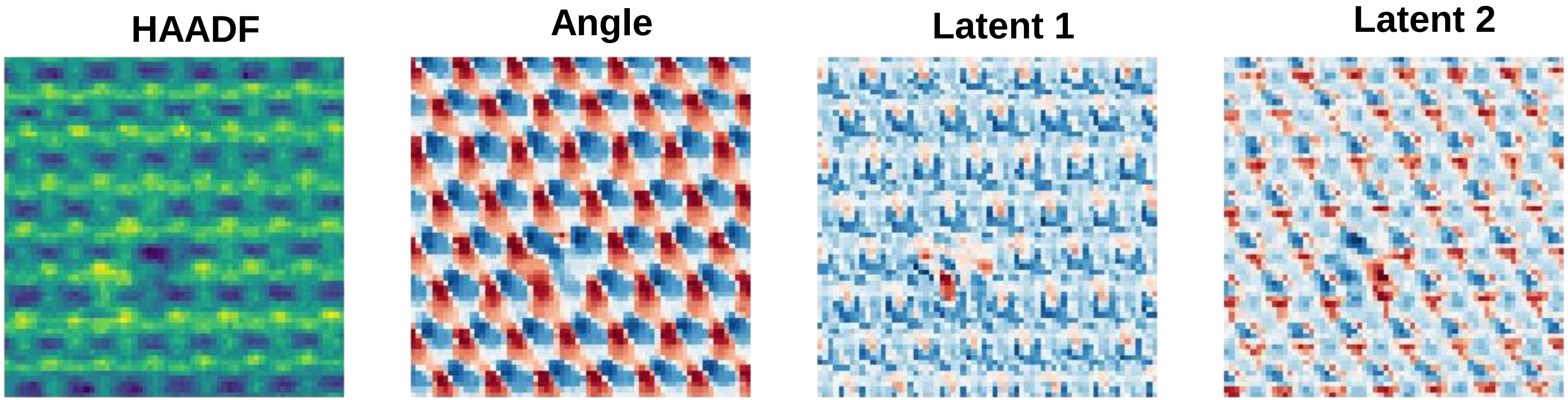
What about 4D STEM?

Simulated 4D STEM for graphene



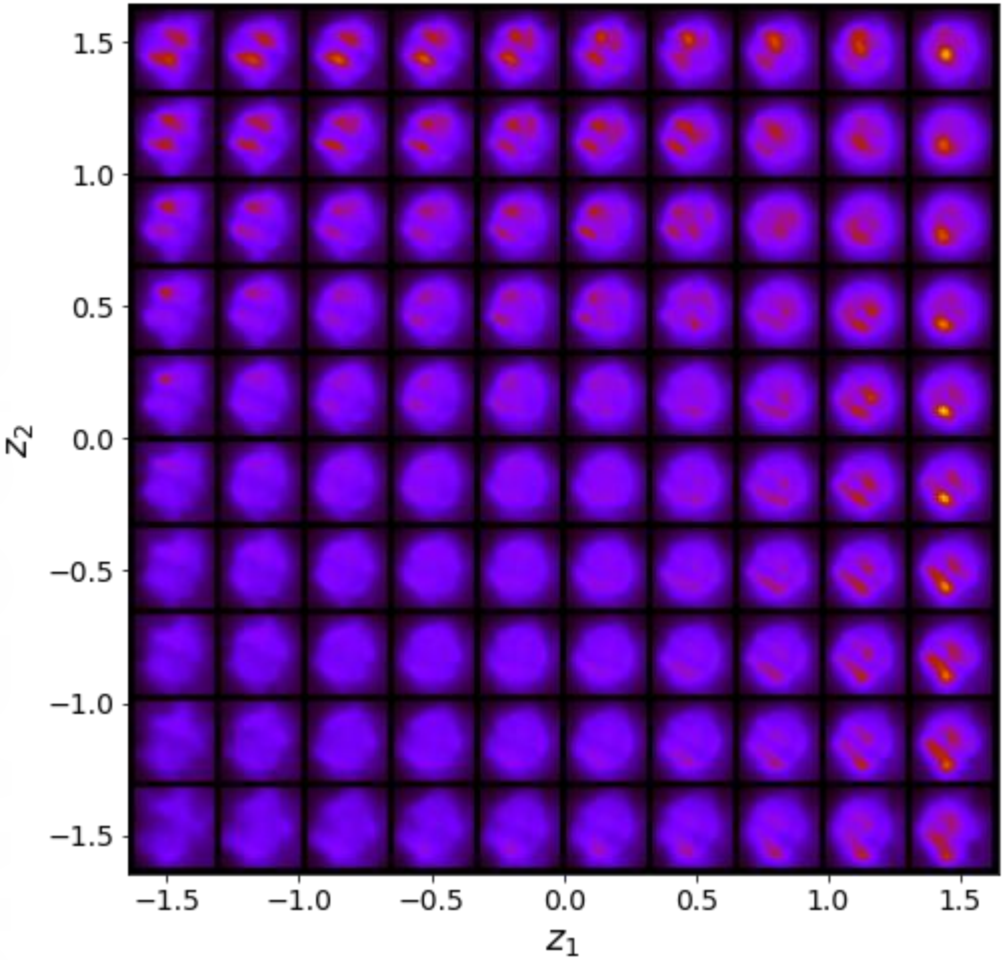


Experiment: vacancy in graphene

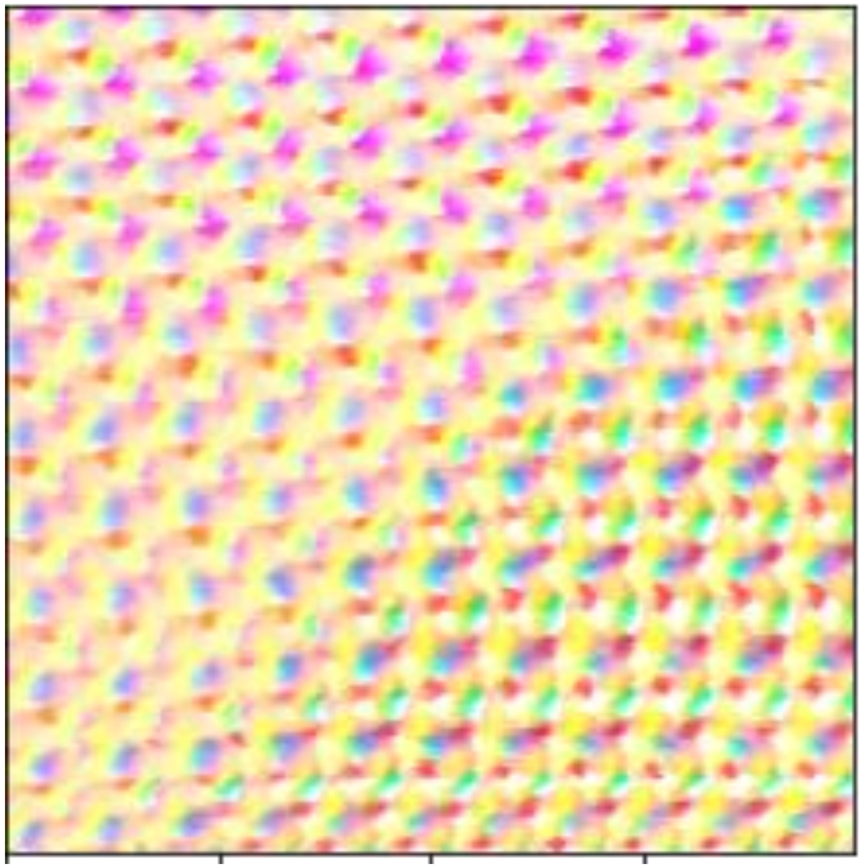


VAE on 4D STEM of Domain Wall in BiFeO3

Latent representations

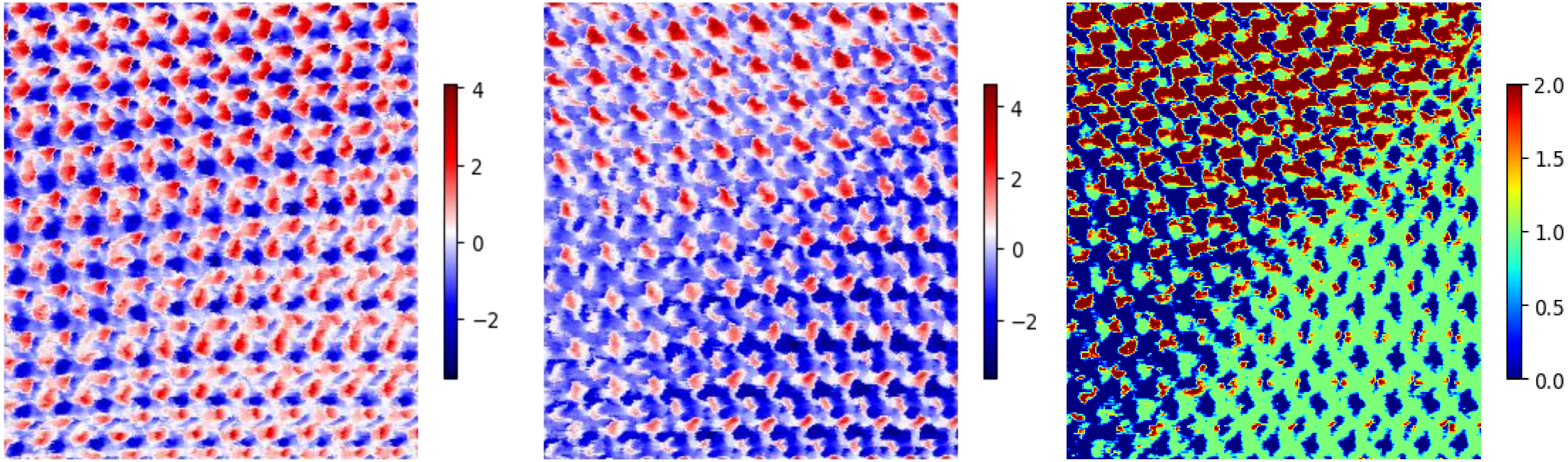


Latent image (3D VAE)



- Extract factors of variability in the 4D STEM data set
- Separate these latent factors from physical factors of variability such as rotation and shear

Joint VAE

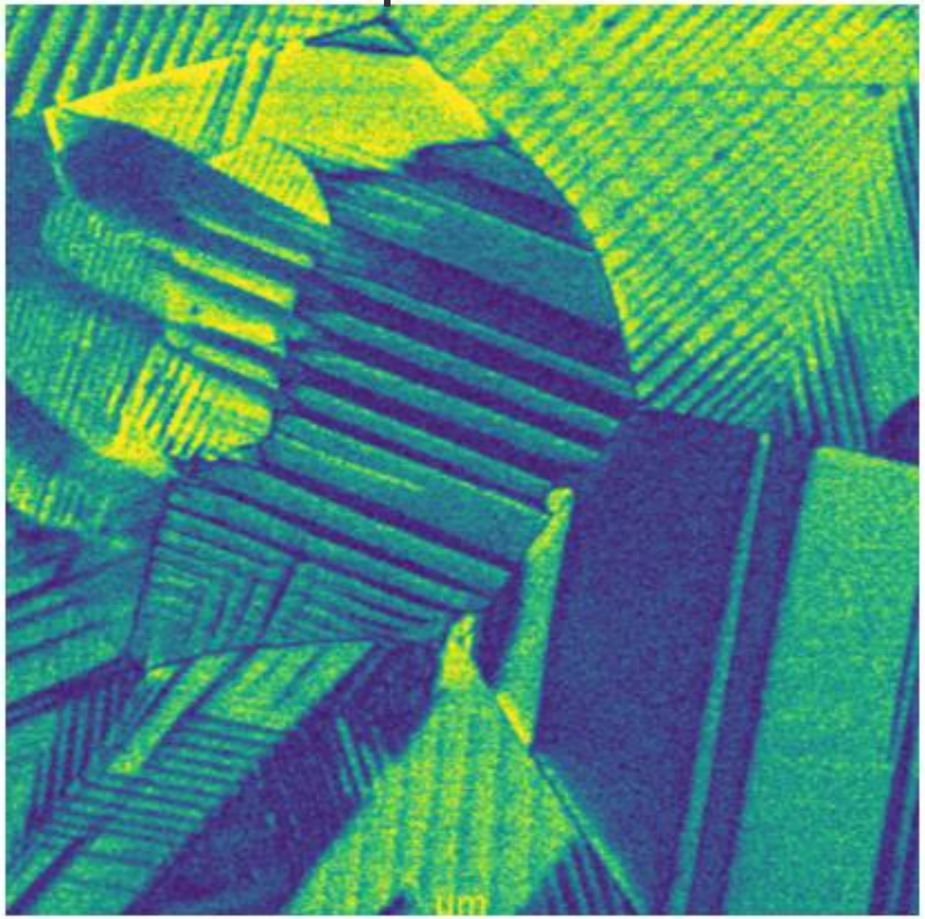


- Disentanglement of representations
- Separation in classes
- Extraction of invariances (rotation and shear)

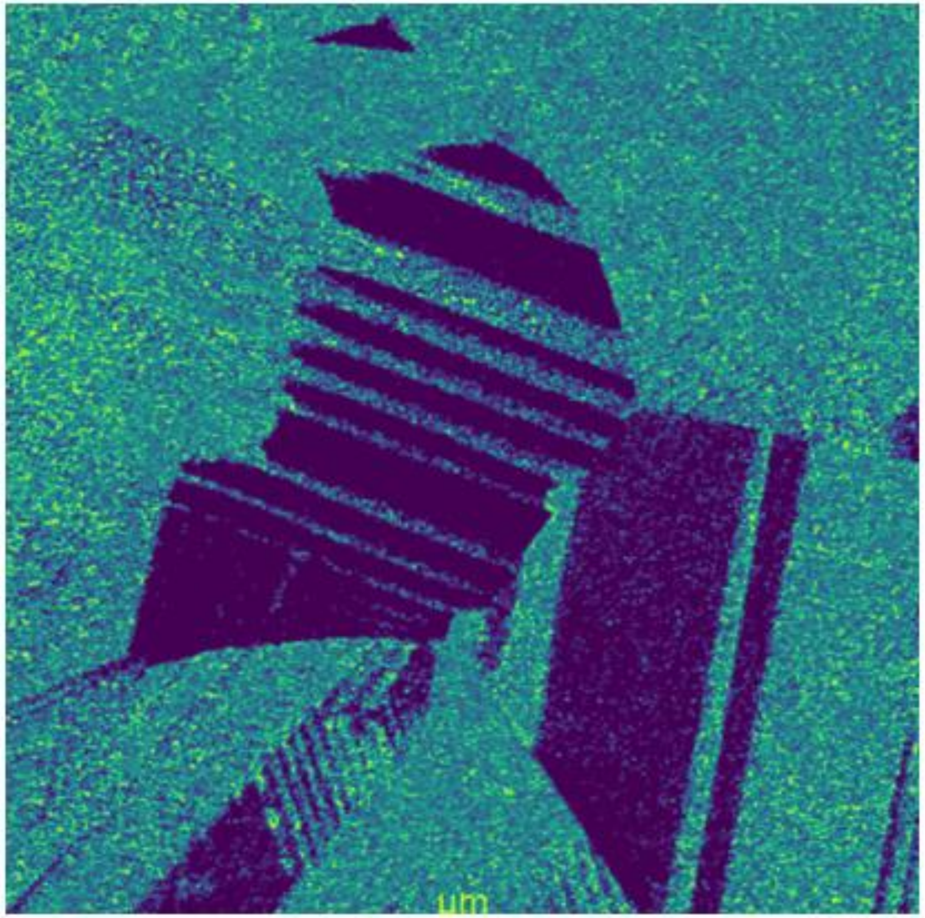
Now, **open source** software (part of PyroVED)

Ferroelectric domain and domain walls

Amplitude



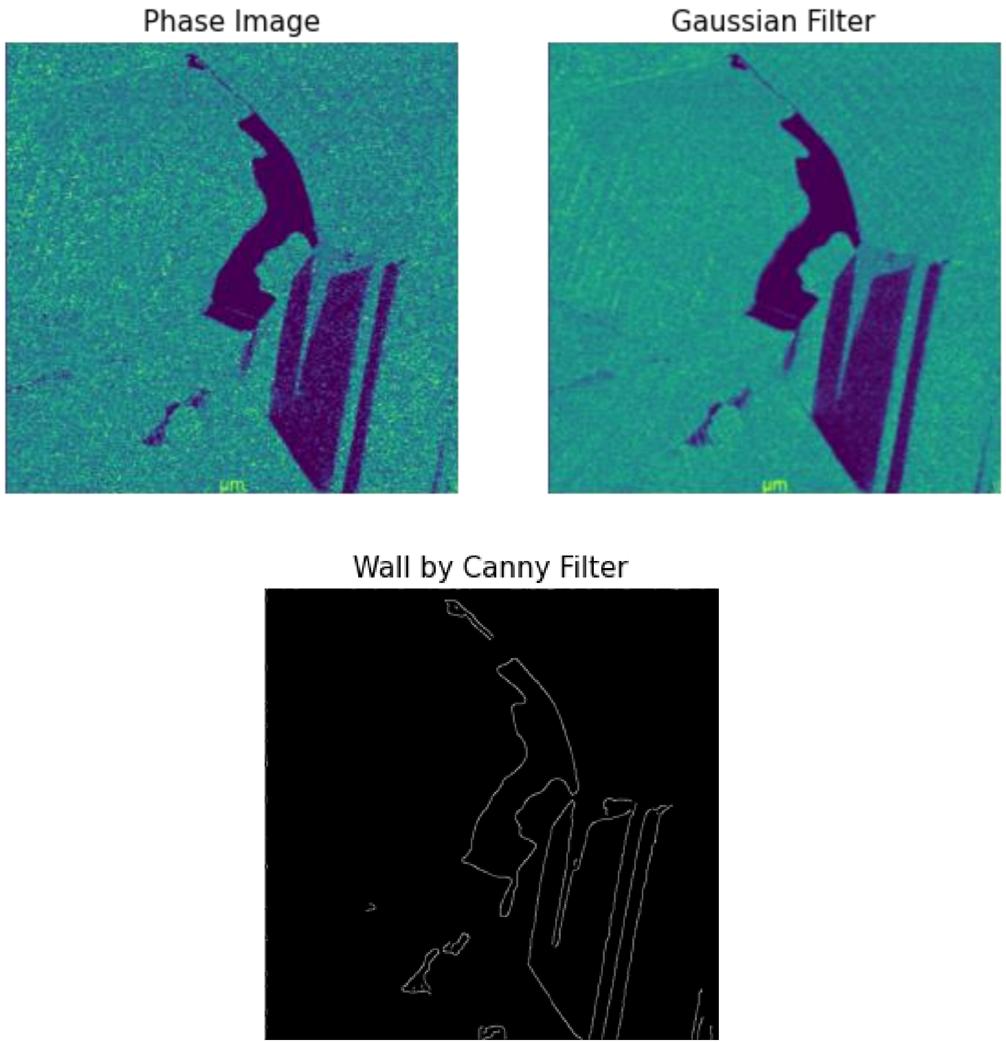
Phase



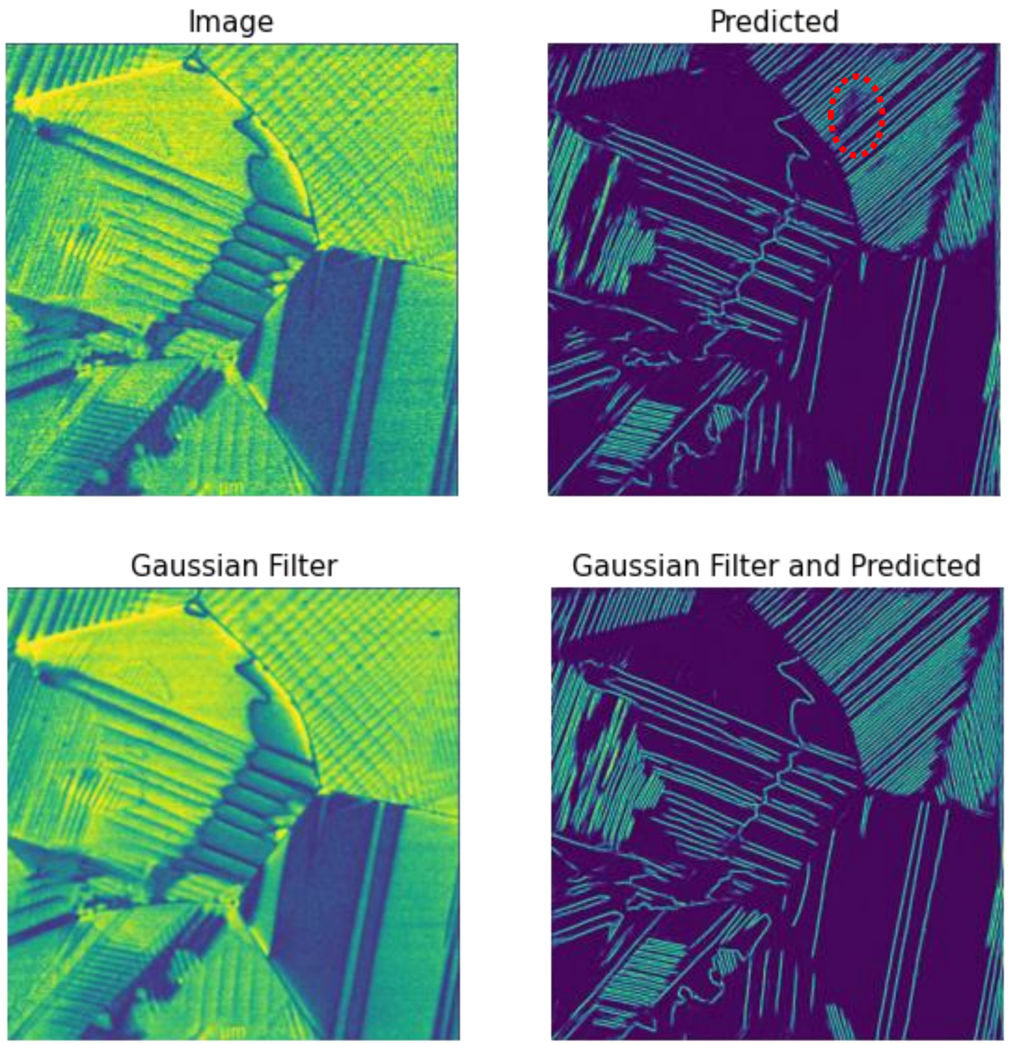
Detecting domain walls

THE UNIVERSITY of TENNESSEE KNOXVILLE

Canny filter

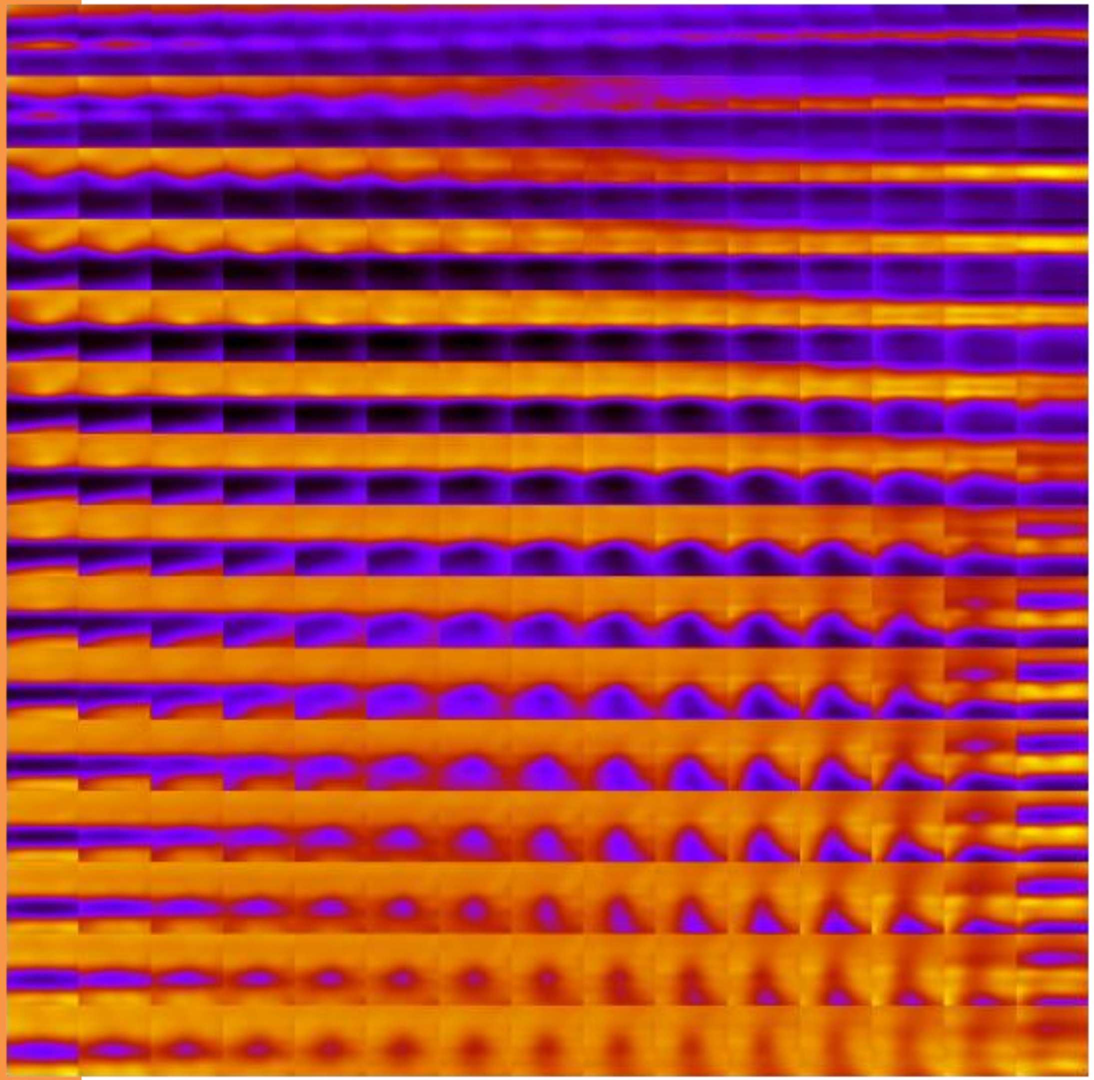


DCNN Prediction

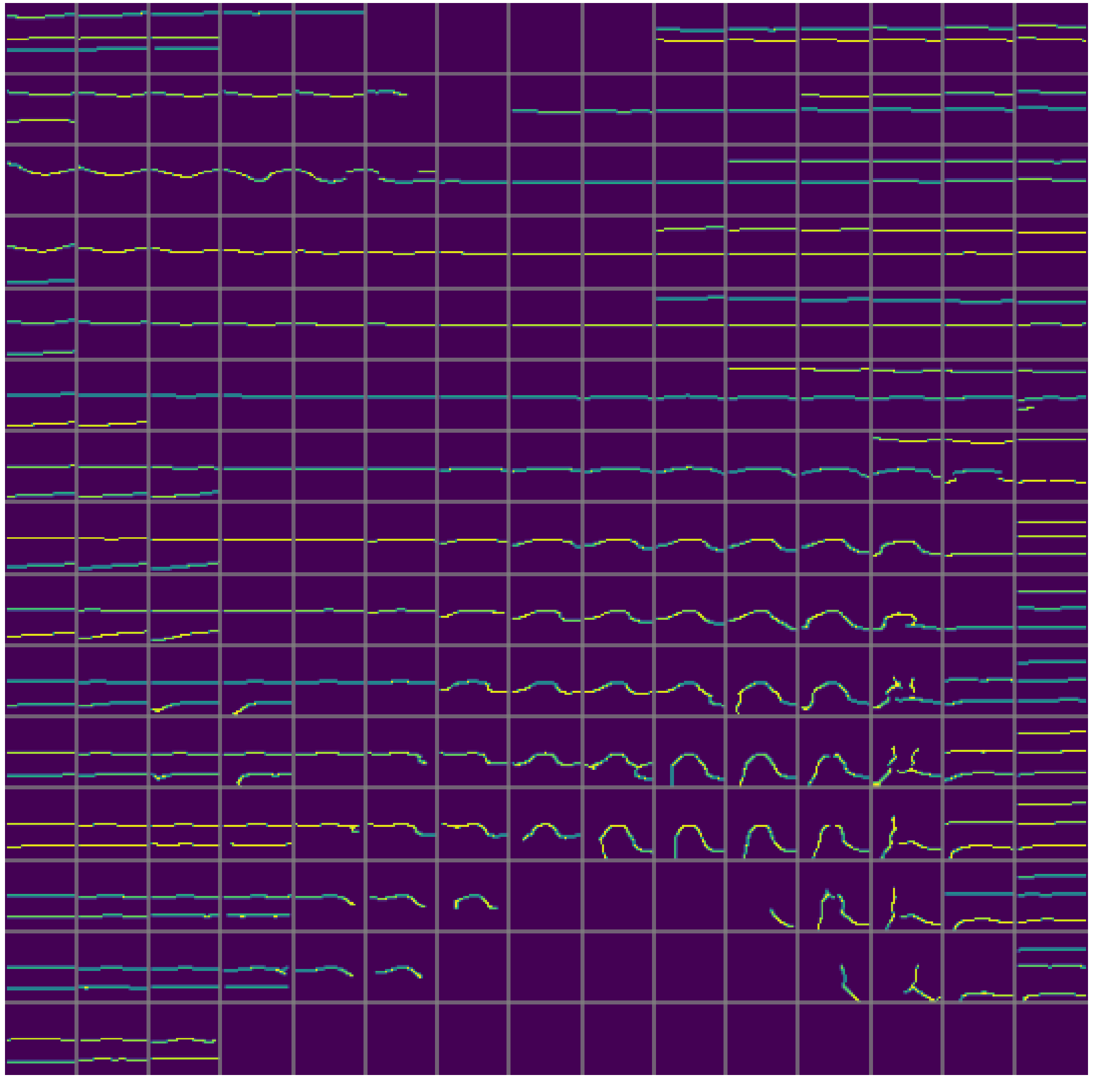


rVAE analysis

Latent Space

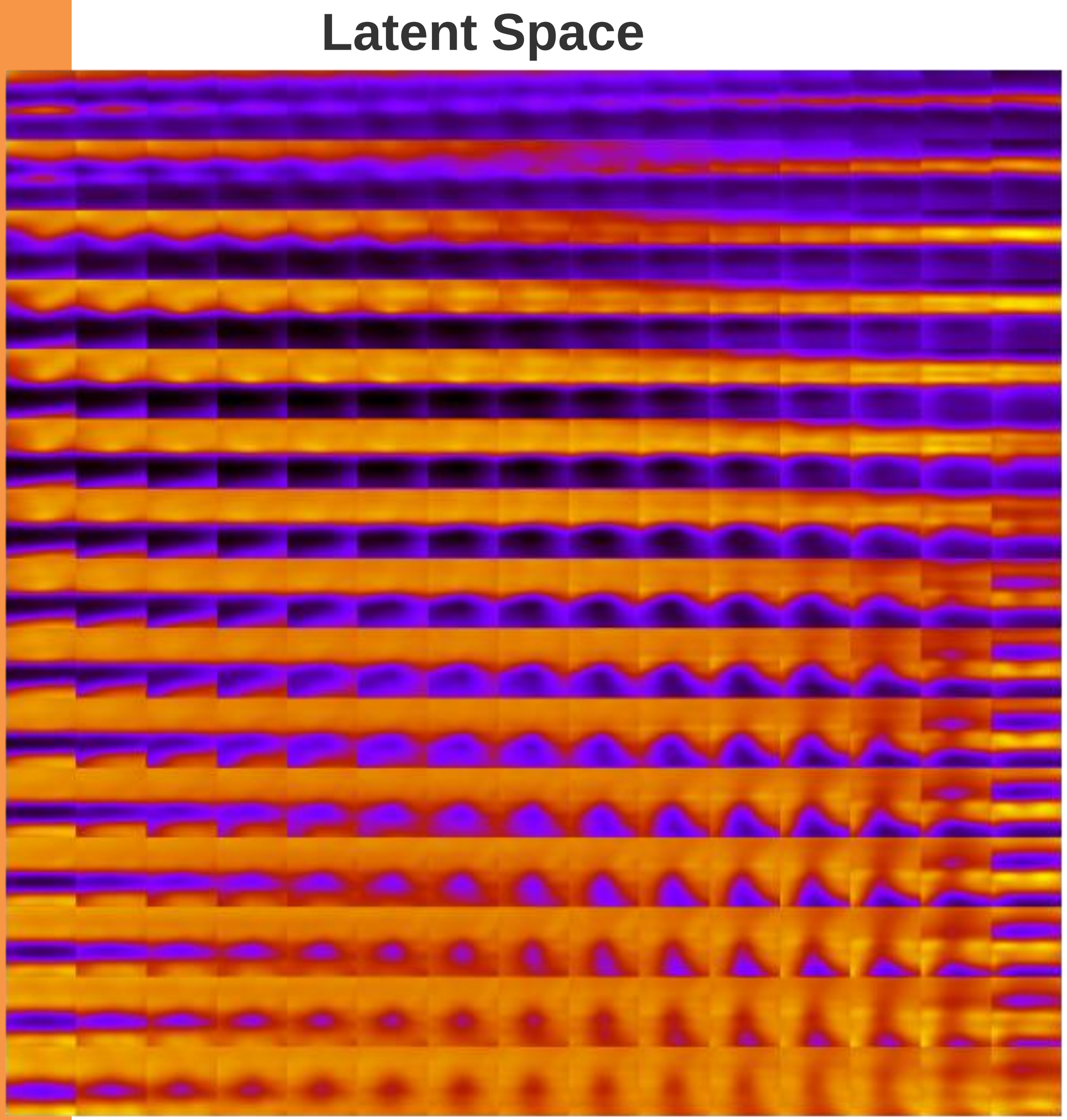


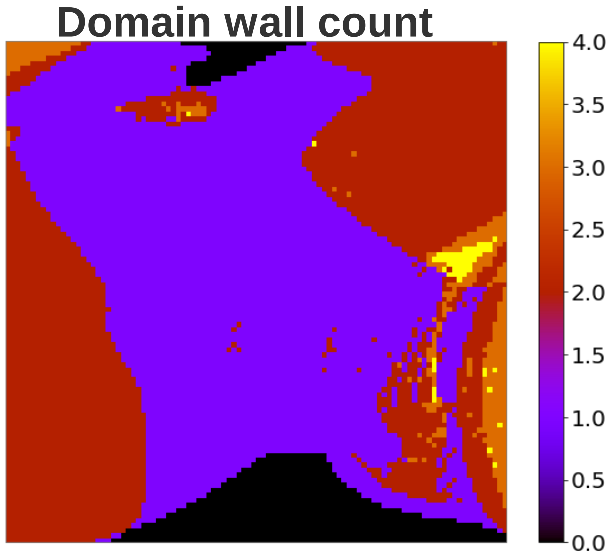
Domain Walls

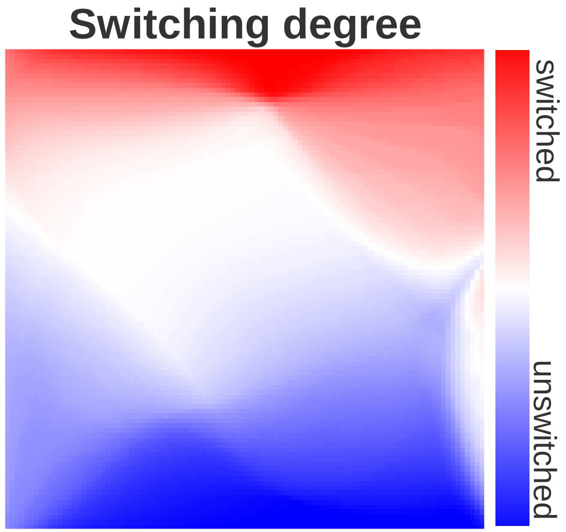


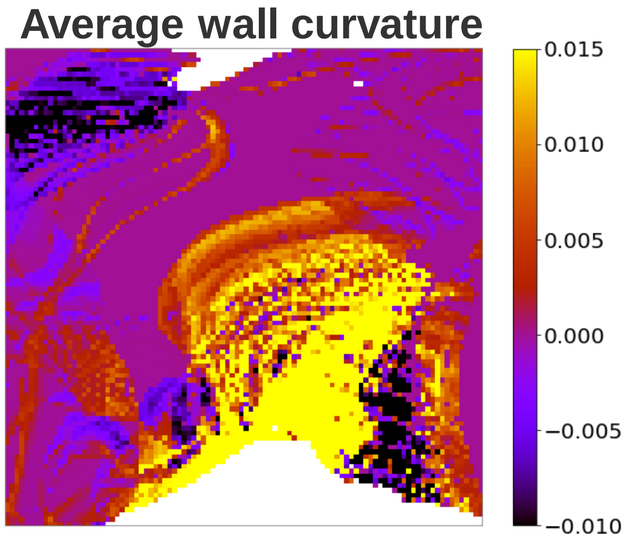
rVAE latent space

THE UNIVERSITY of TENNESSEE KNOXVILLE

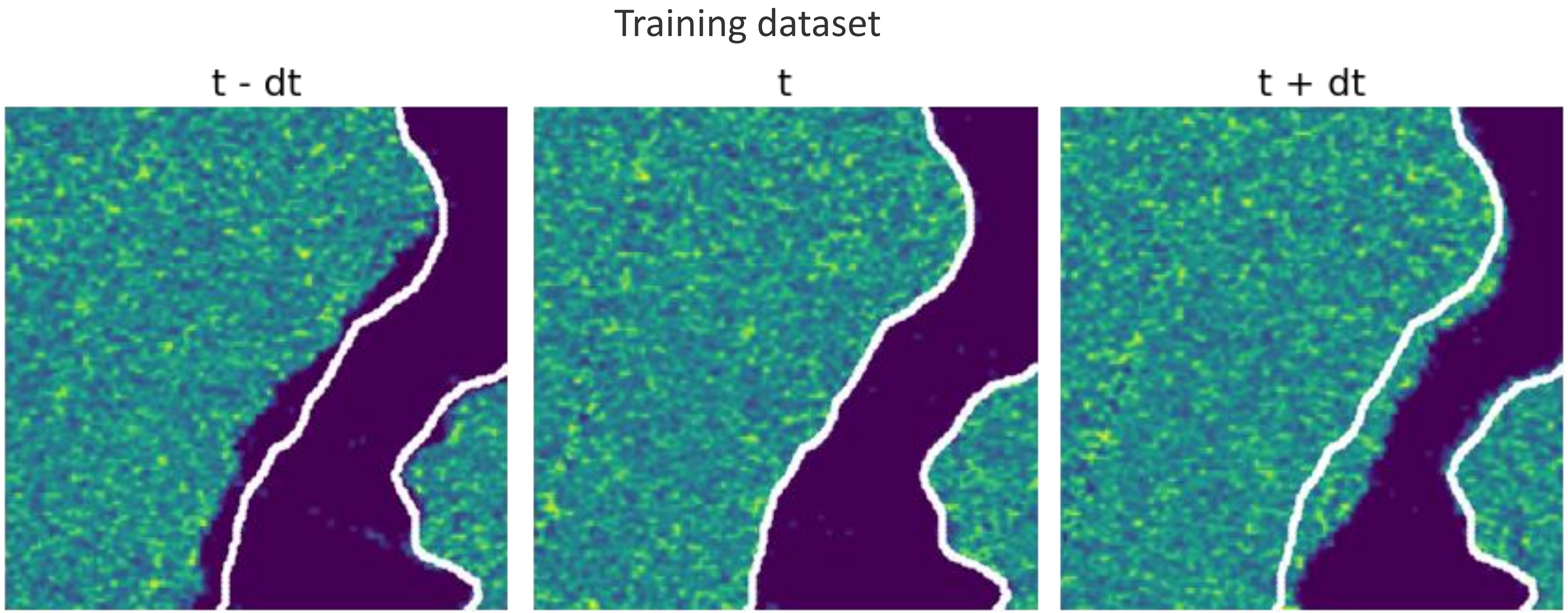








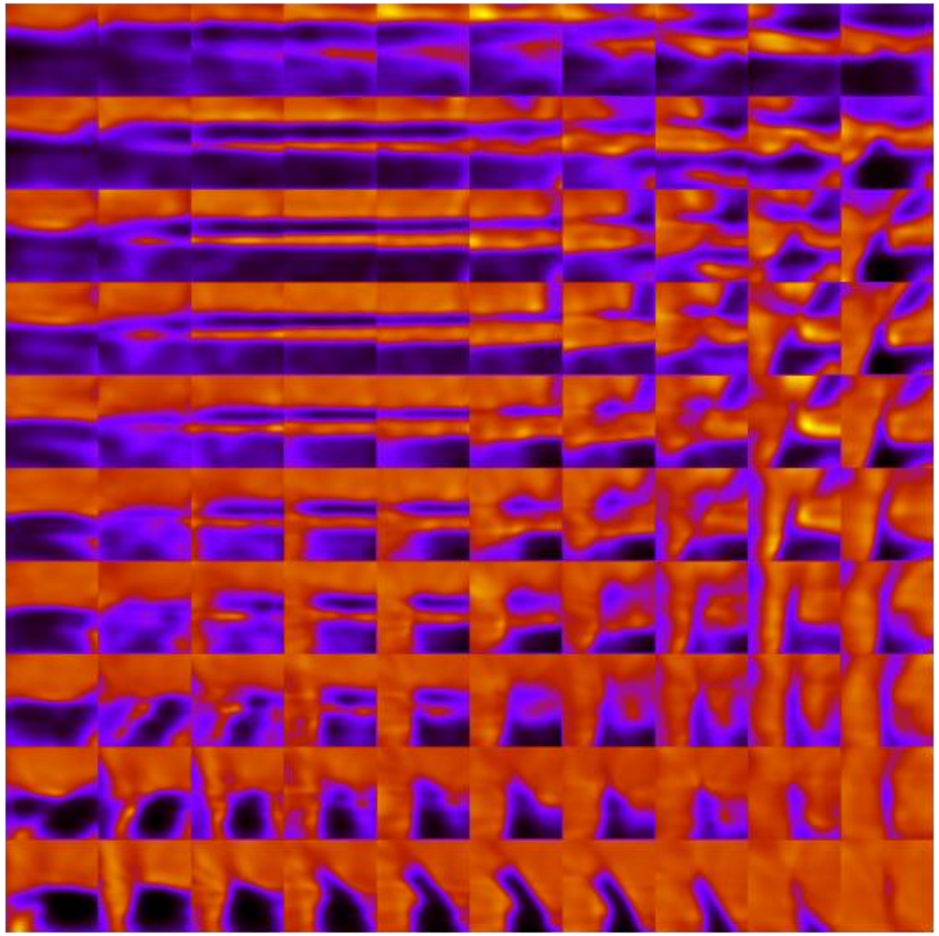
rVAE with time delay



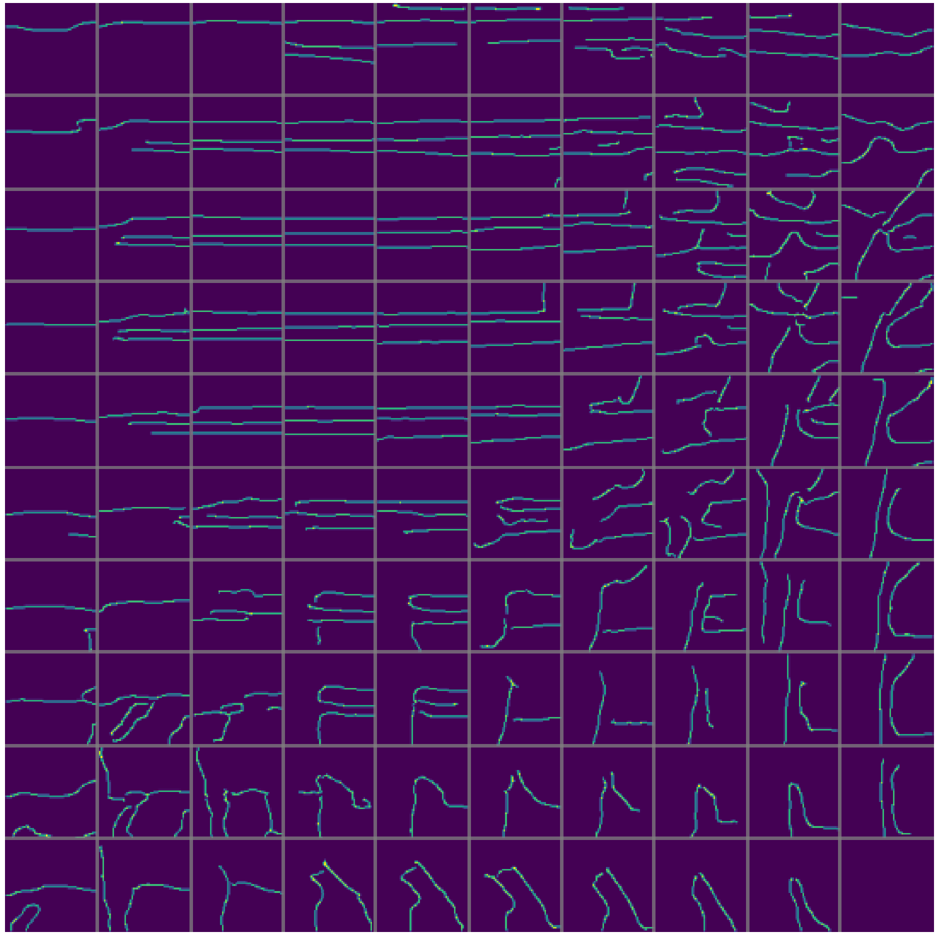
rVAE with time delay

THE UNIVERSITY of TENNESSEE KNOXVILLE

Latent space

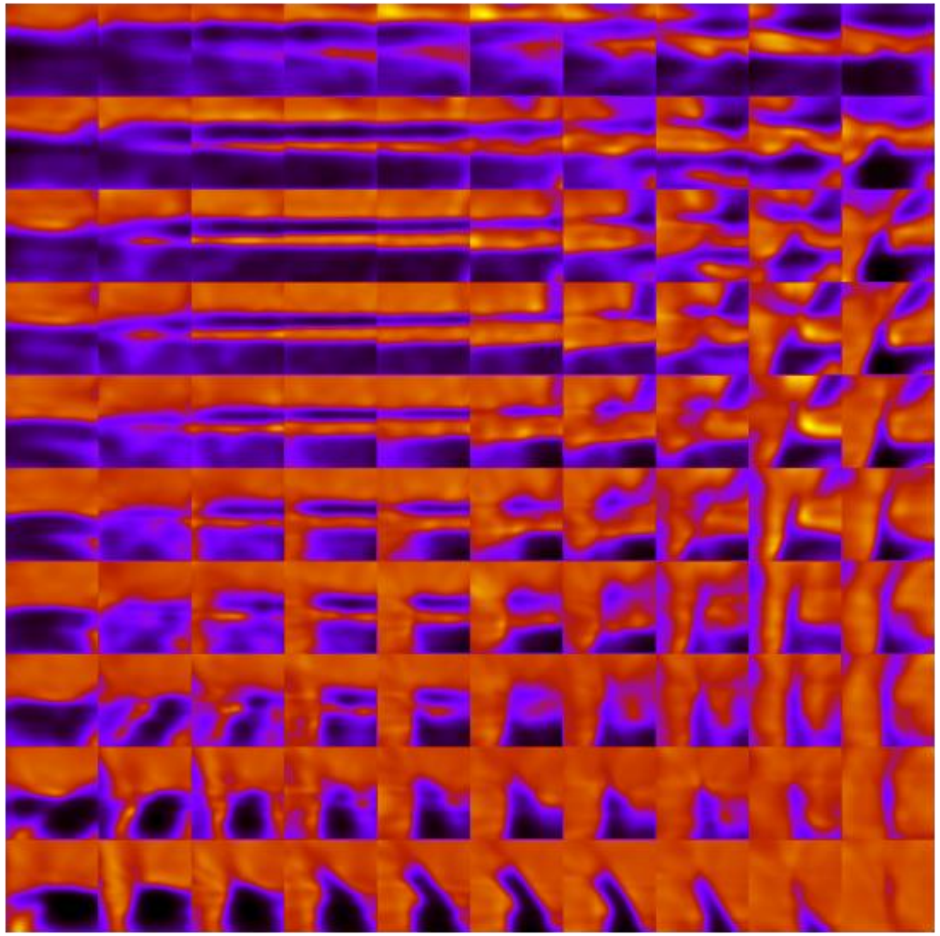


Domain wall

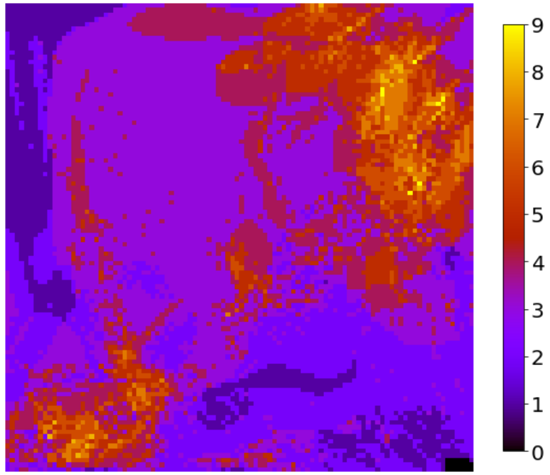


rVAE with time delay

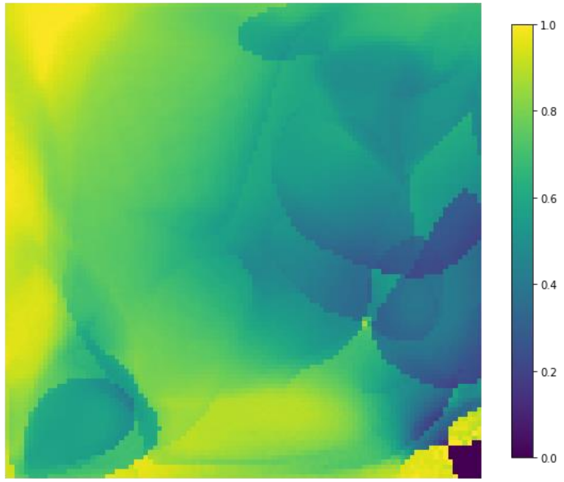
Latent space



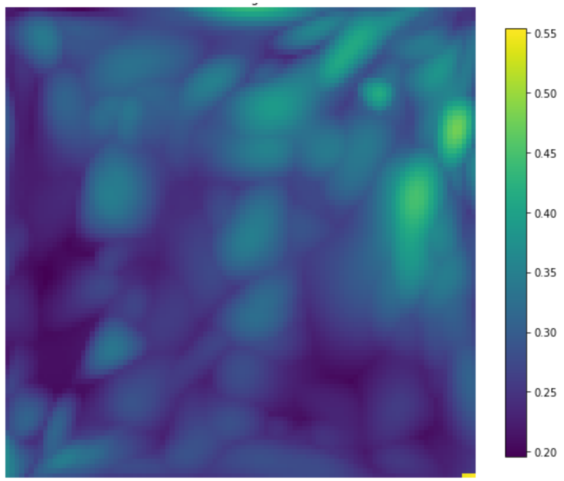
Wall count



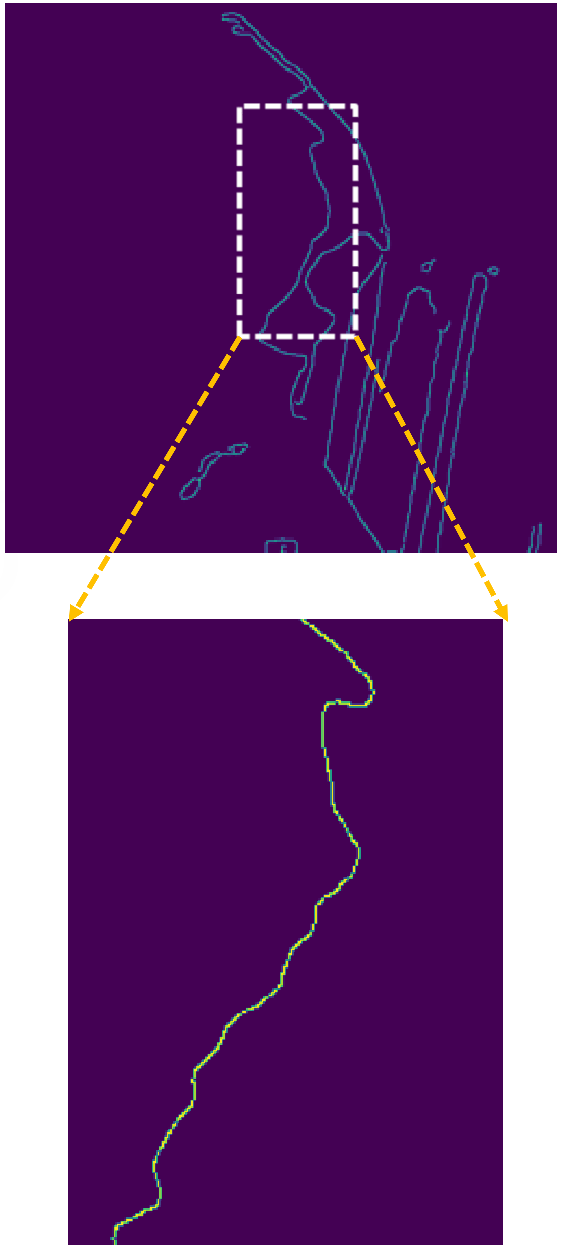
Domain convex

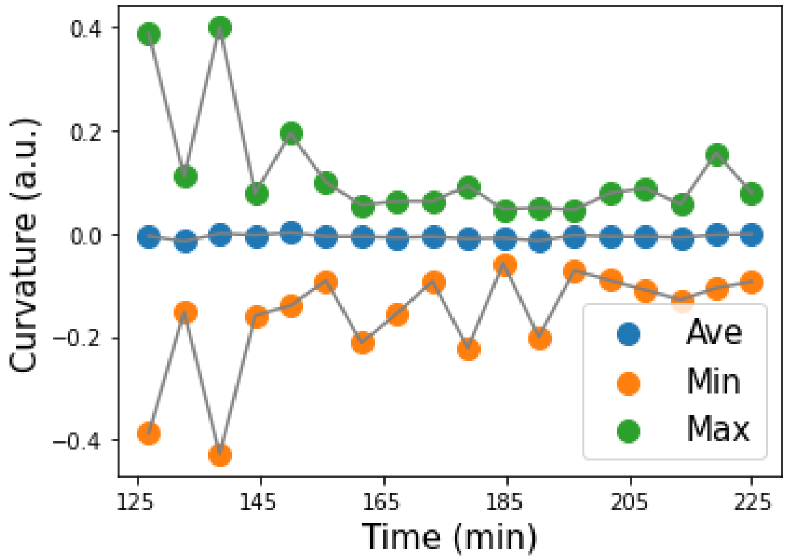


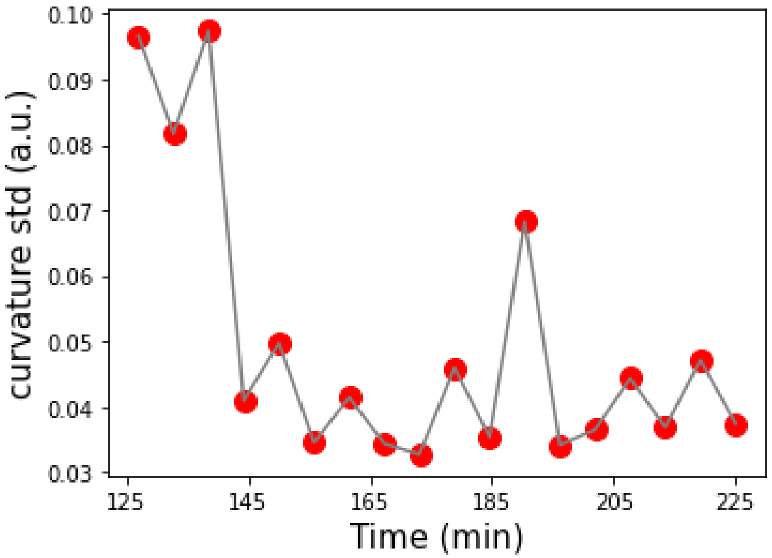
Switch significance



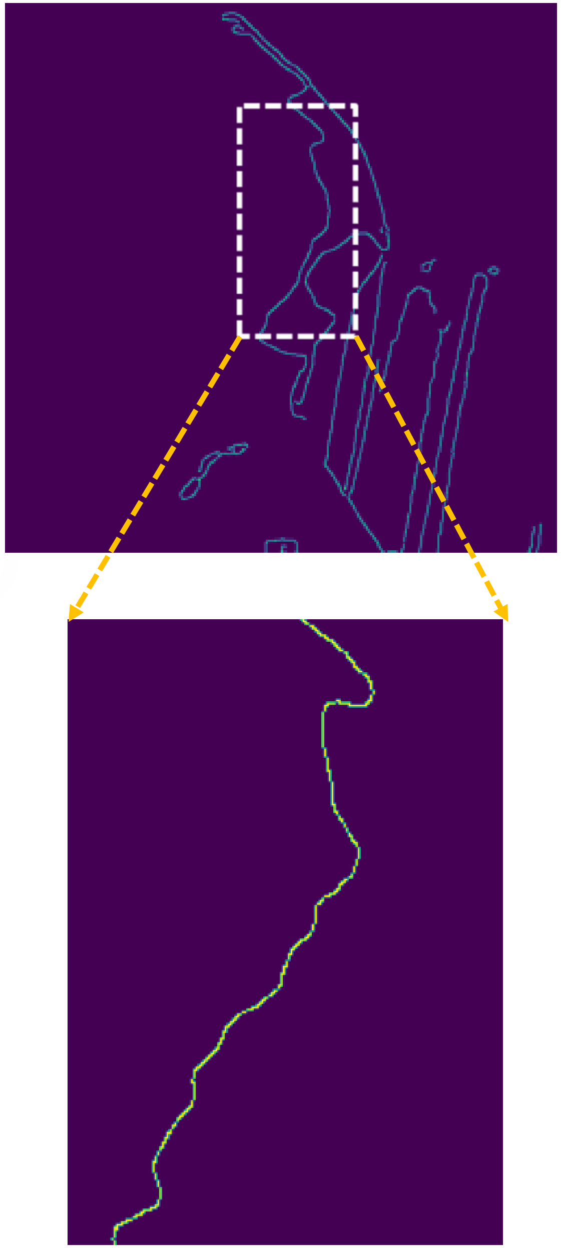
Domain wall evolution





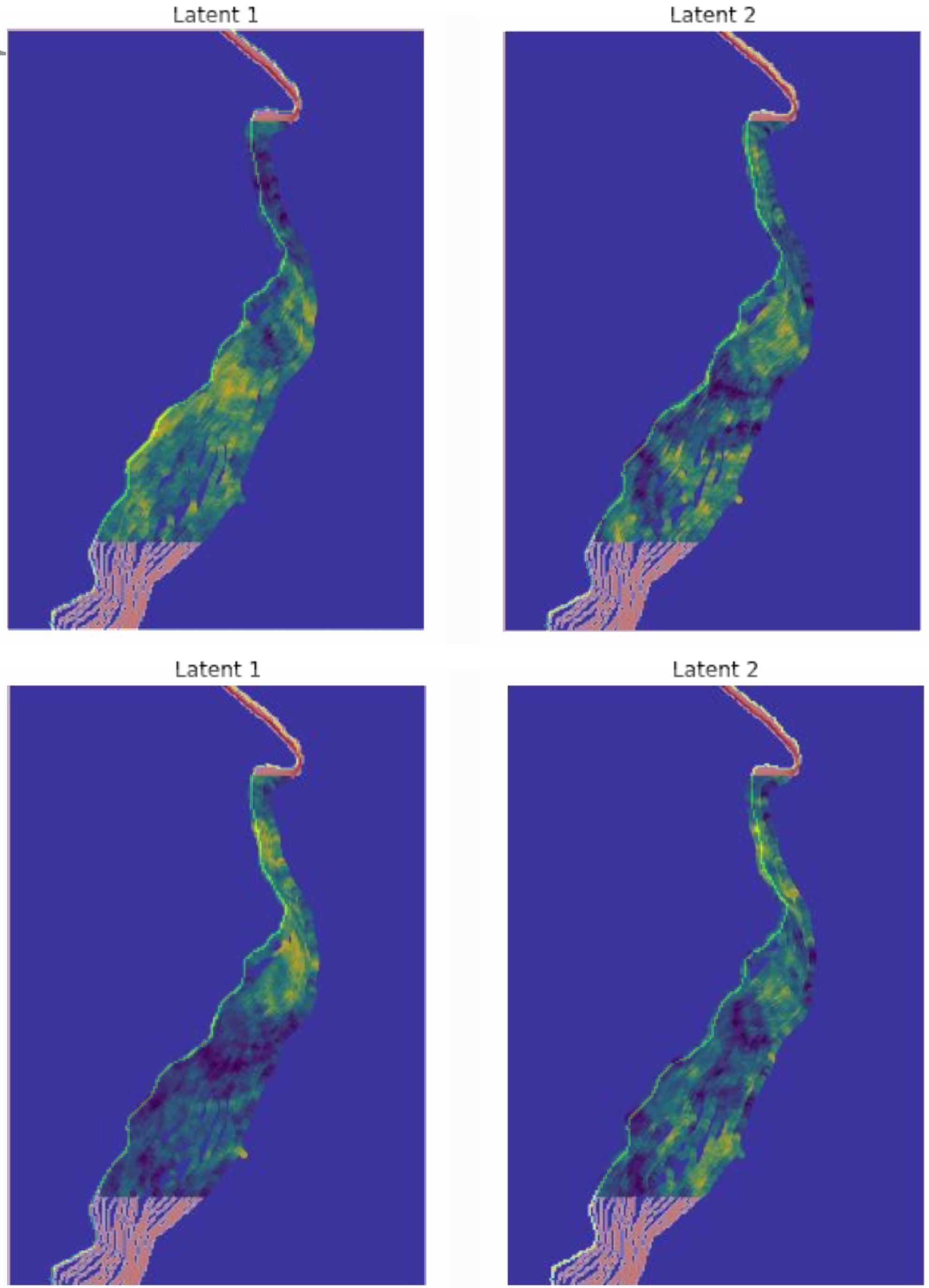


Domain wall evolution



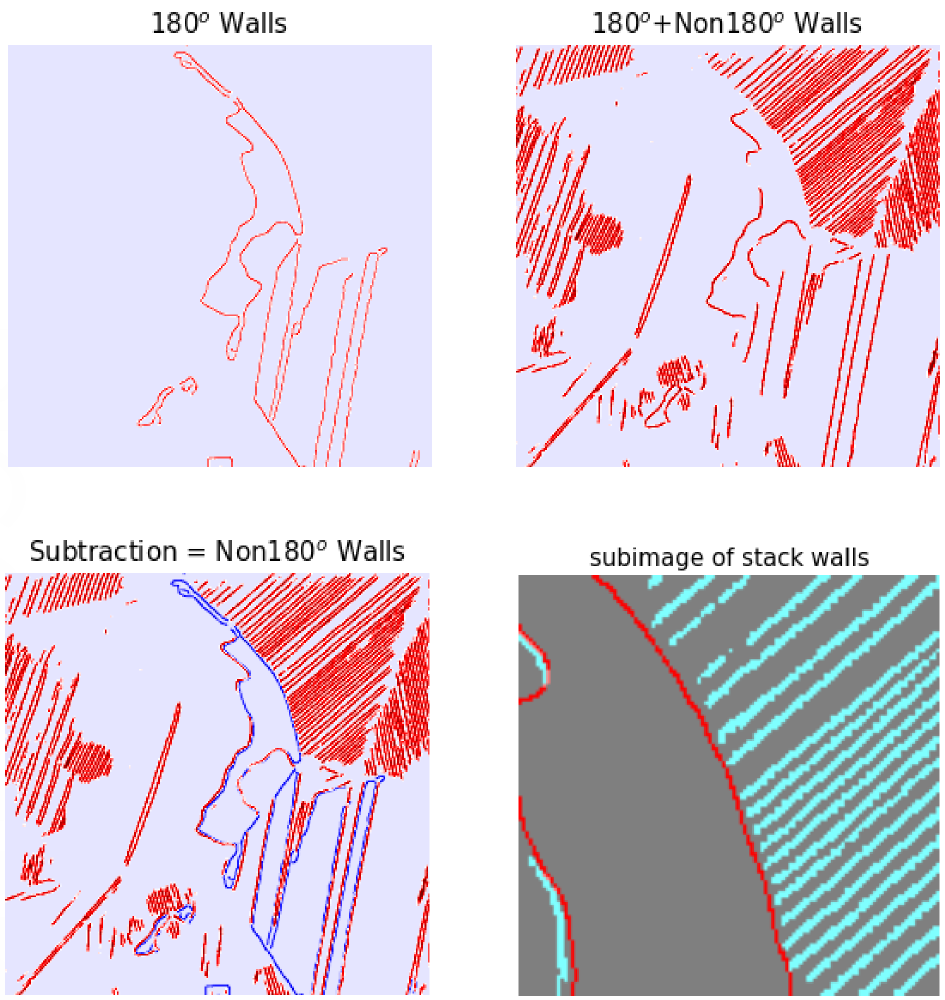
Forward:
t vs t+1

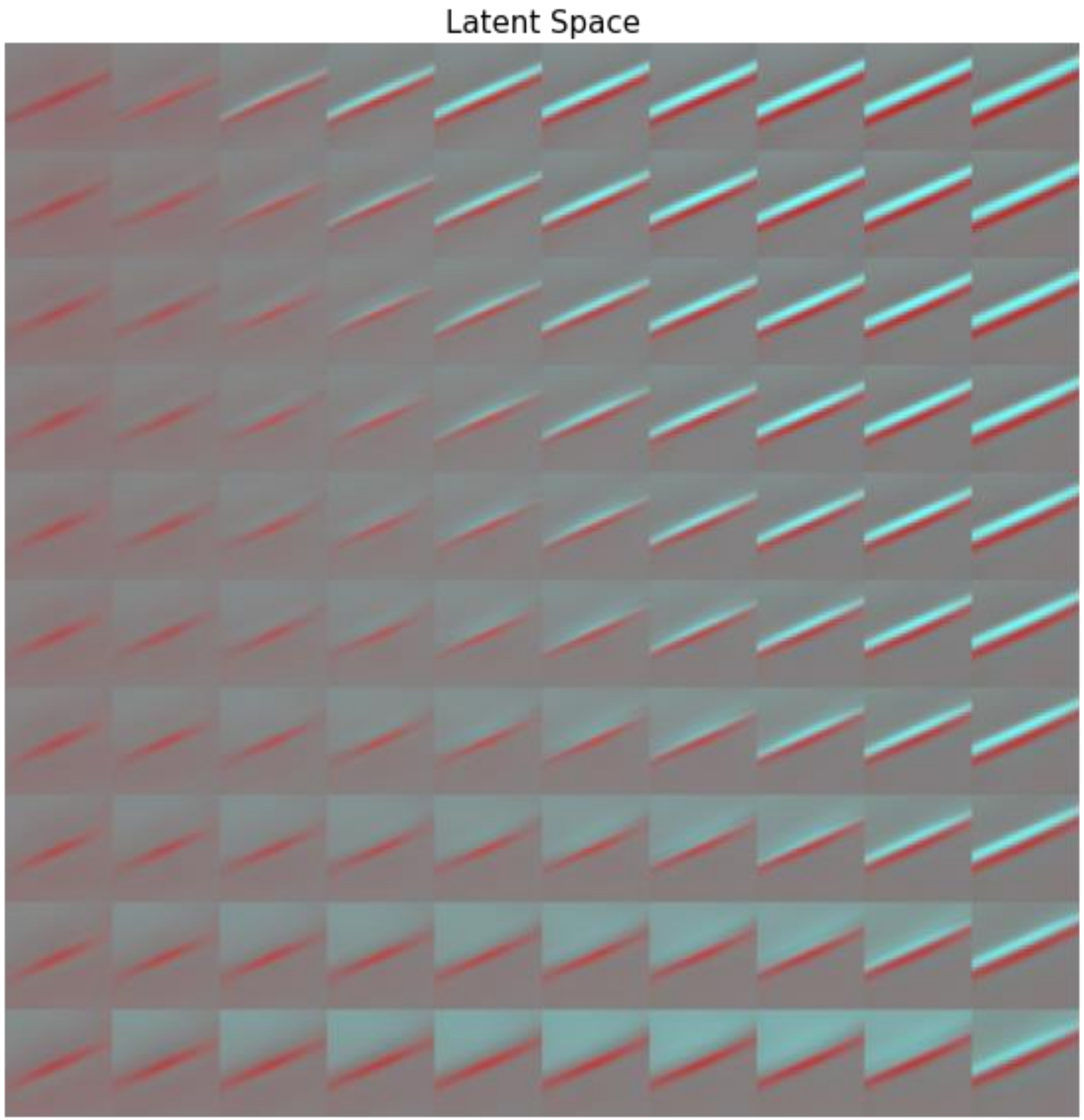
Reverse:
t vs t+1



Multilayer rVAE

THE UNIVERSITY of TENNESSEE KNOXVILLE





Pinning mechanism

THE UNIVERSITY of TENNESSEE KNOXVILLE