THE UNIVERSITY of TENNESSEE KNOXVILLE

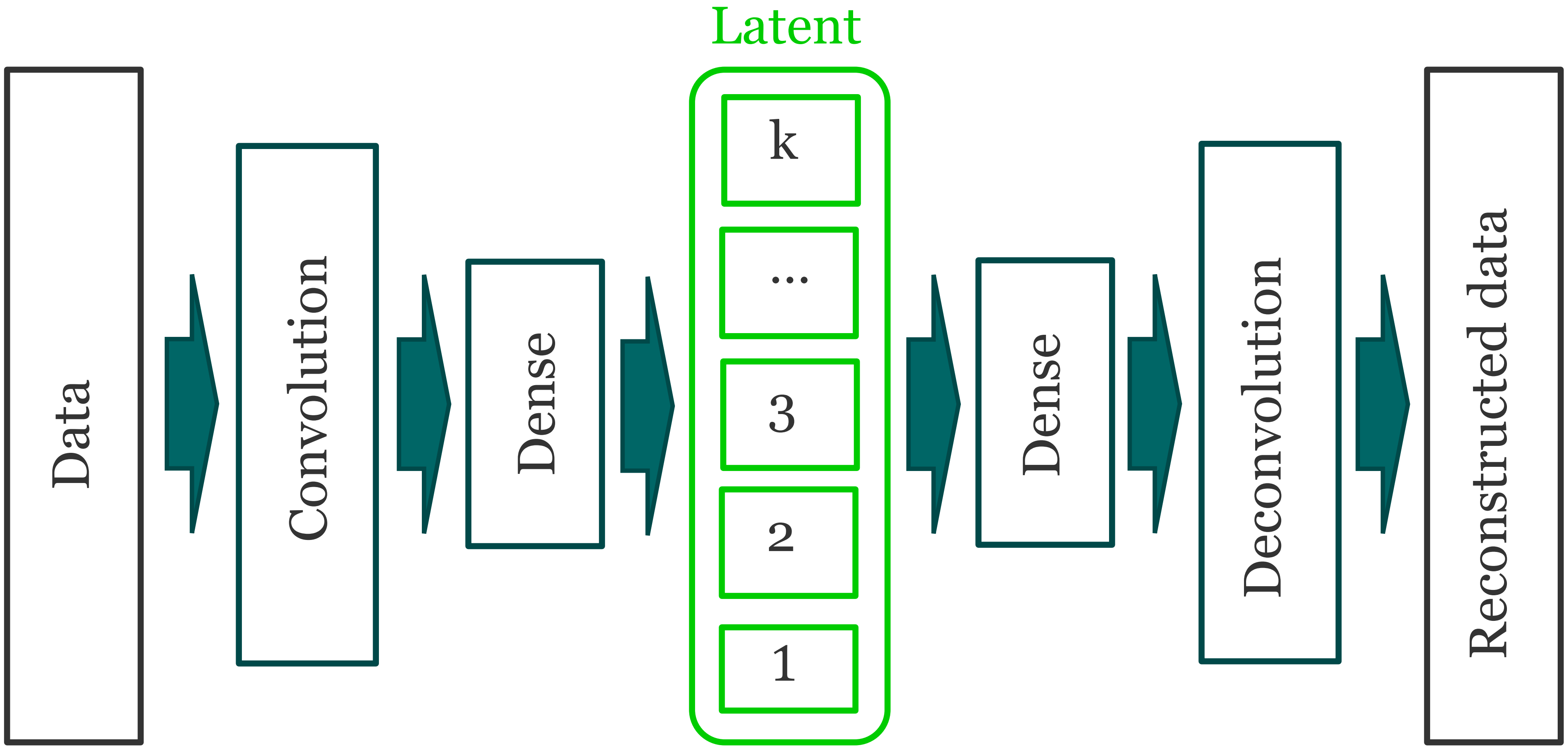
Day 4: Variational Autoencoders for imaging, spectra, and structure-property relationships

Sergei V. Kalinin

University of Tennessee, Knoxville and
Pacific Northwest National Laboratory

Pacific Northwest NATIONAL LABORATORY

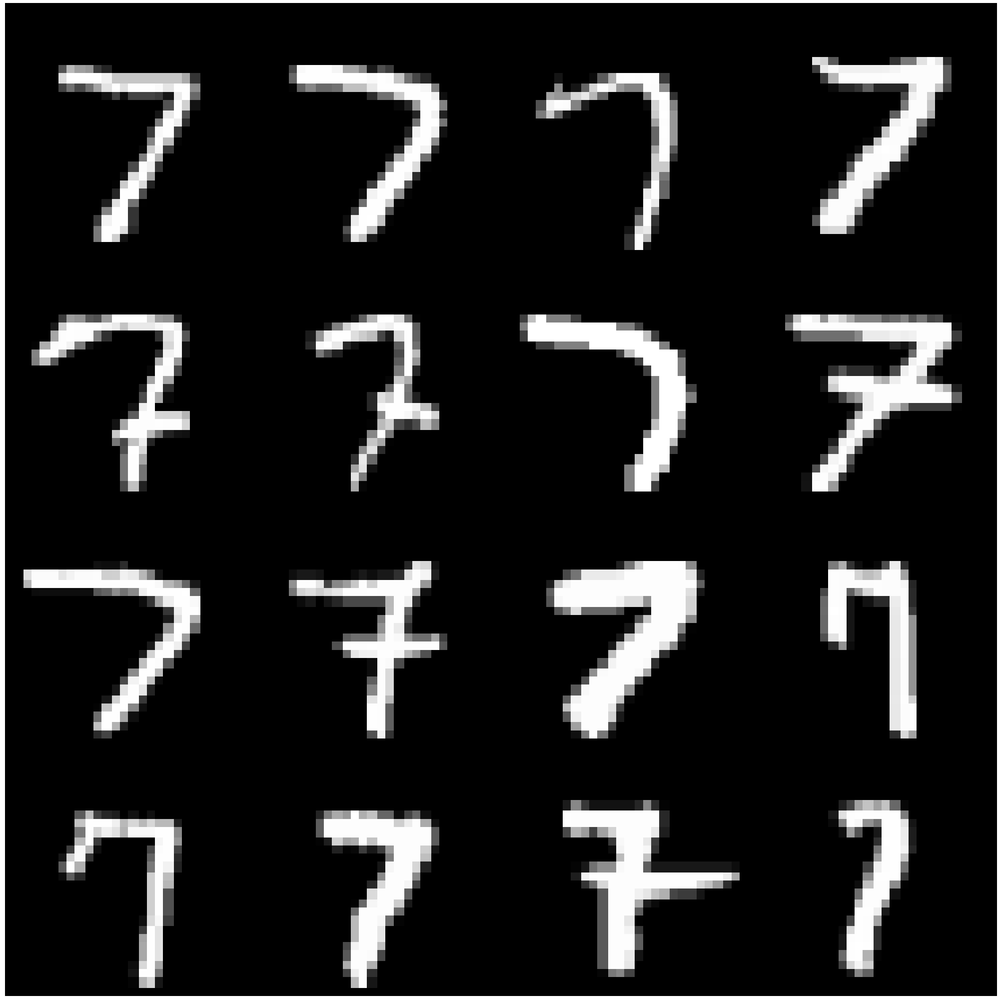
Autoencoders



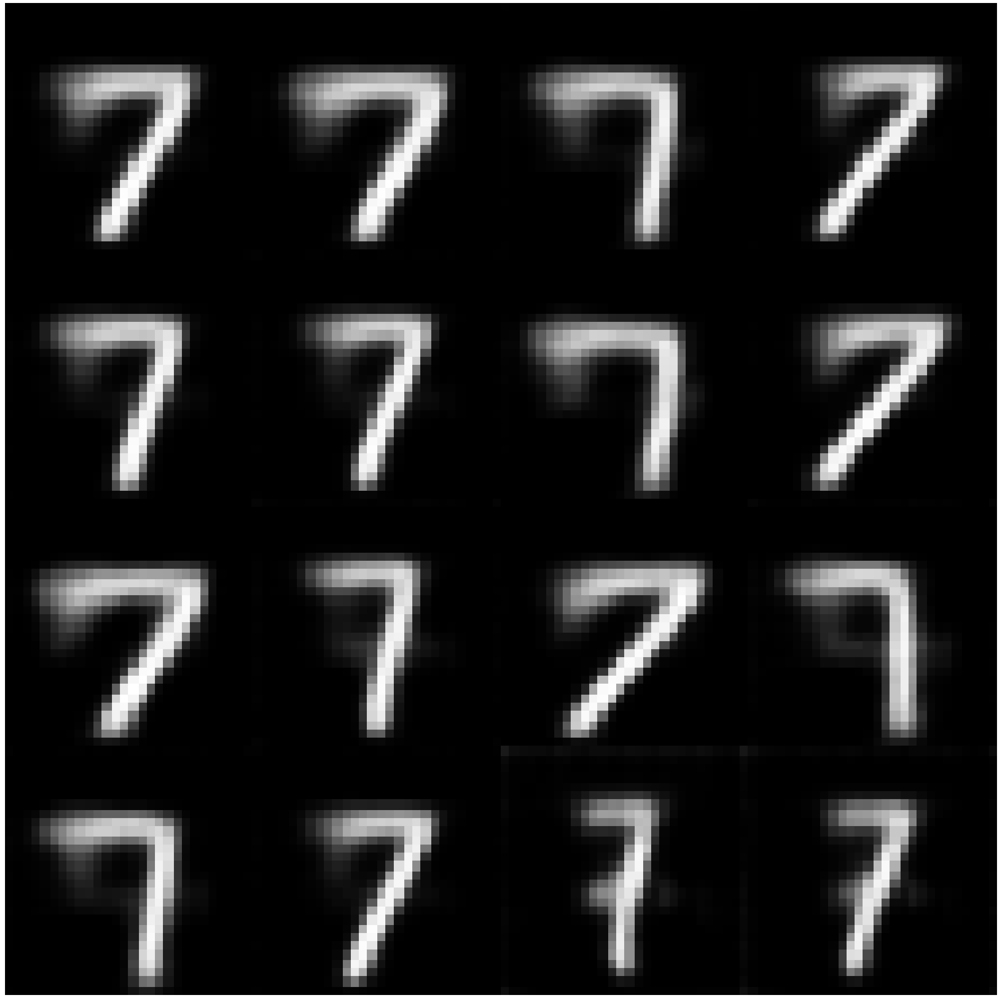
Loss: reconstruction loss

The AE reconstructs data

Input data



Decoded data



Why are AE important?

Geoffrey Hinton

FOLLOW

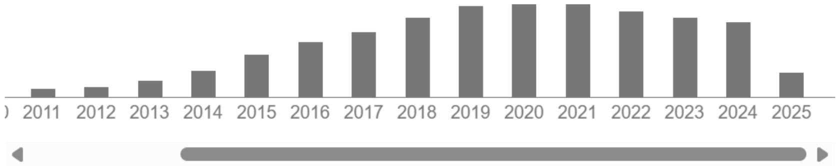
Emeritus Prof. Computer Science, University of Toronto
Verified email at cs.toronto.edu - Homepage

machine learning   psychology   artificial intelligence   cognitive science   computer science

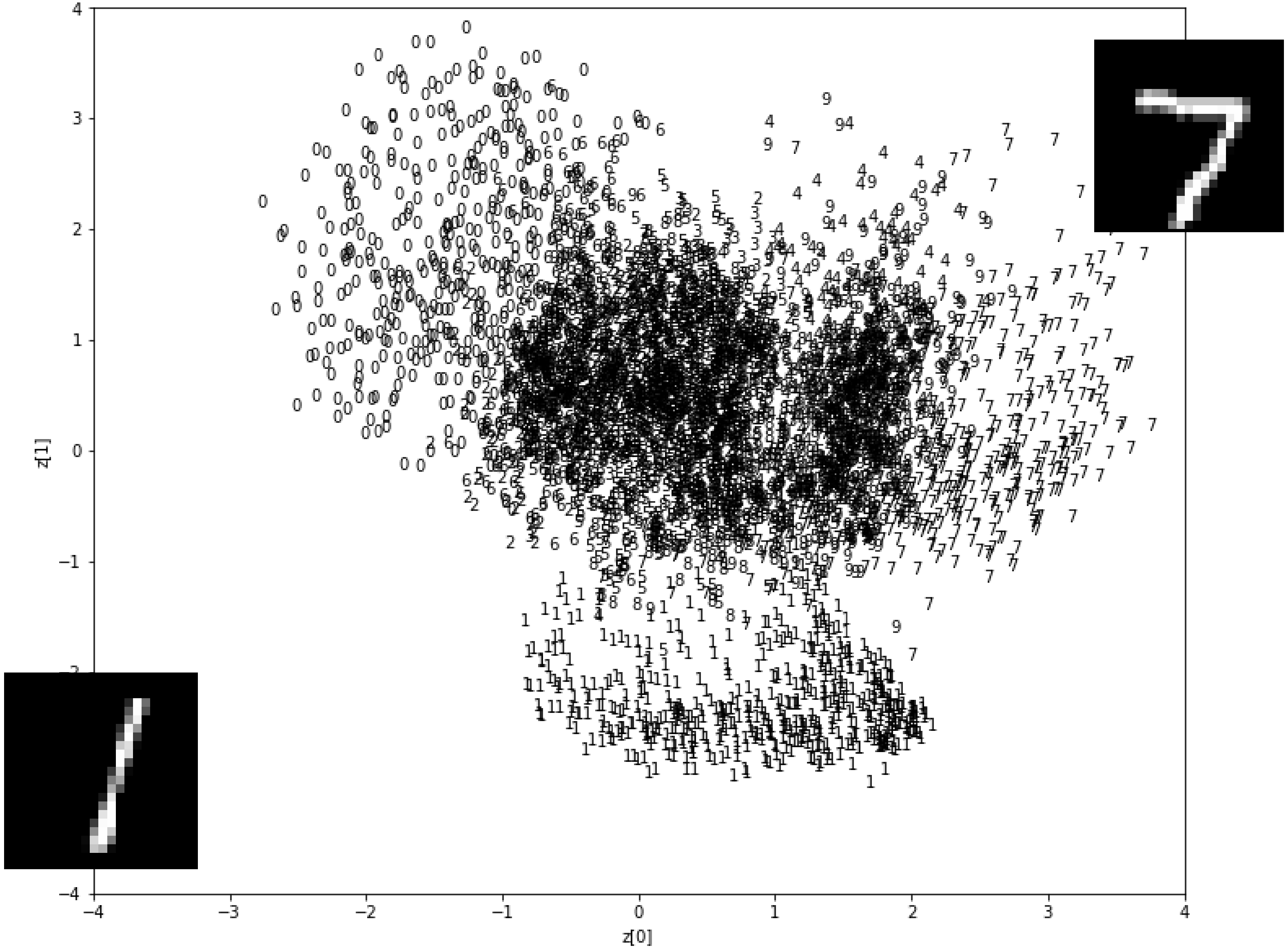
TITLE	CITED BY	YEAR
Imagenet classification with deep convolutional neural networks A Krizhevsky, I Sutskever, GE Hinton Advances in neural information processing systems 25	175478 *	2012
Deep learning Y LeCun, Y Bengio, G Hinton Nature 521 (7553), 436-44	94697	2015
Learning internal representations by error-propagation DE Rumelhart, GE Hinton, RJ Williams Parallel Distributed Processing: Explorations in the Microstructure of …	59850 *	1986
Visualizing data using t-SNE L van der Maaten, G Hinton Journal of Machine Learning Research 9 (Nov), 2579-2605	56540	2008
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	56535	2014
Learning representations by back-propagating errors DE Rumelhart, GE Hinton, RJ Williams Nature 323 (6088), 533-536	42715	1986
Learning multiple layers of features from tiny images A Krizhevsky, G Hinton	36460	2009
Rectified linear units improve restricted boltzmann machines V Nair, GE Hinton Proceedings of the 27th international conference on machine learning (ICML …	27590	2010
Reducing the dimensionality of data with neural networks GE Hinton, RR Salakhutdinov Science 313 (5786), 504-507	24999	2006
Distilling the knowledge in a neural network G Hinton, O Vinyals, J Dean arXiv preprint arXiv:1503.02531	24675	2015
A simple framework for contrastive learning of visual representations T Chen, S Kornblith, M Norouzi, G Hinton International conference on machine learning, 1597-1607	23542	2020

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 24999

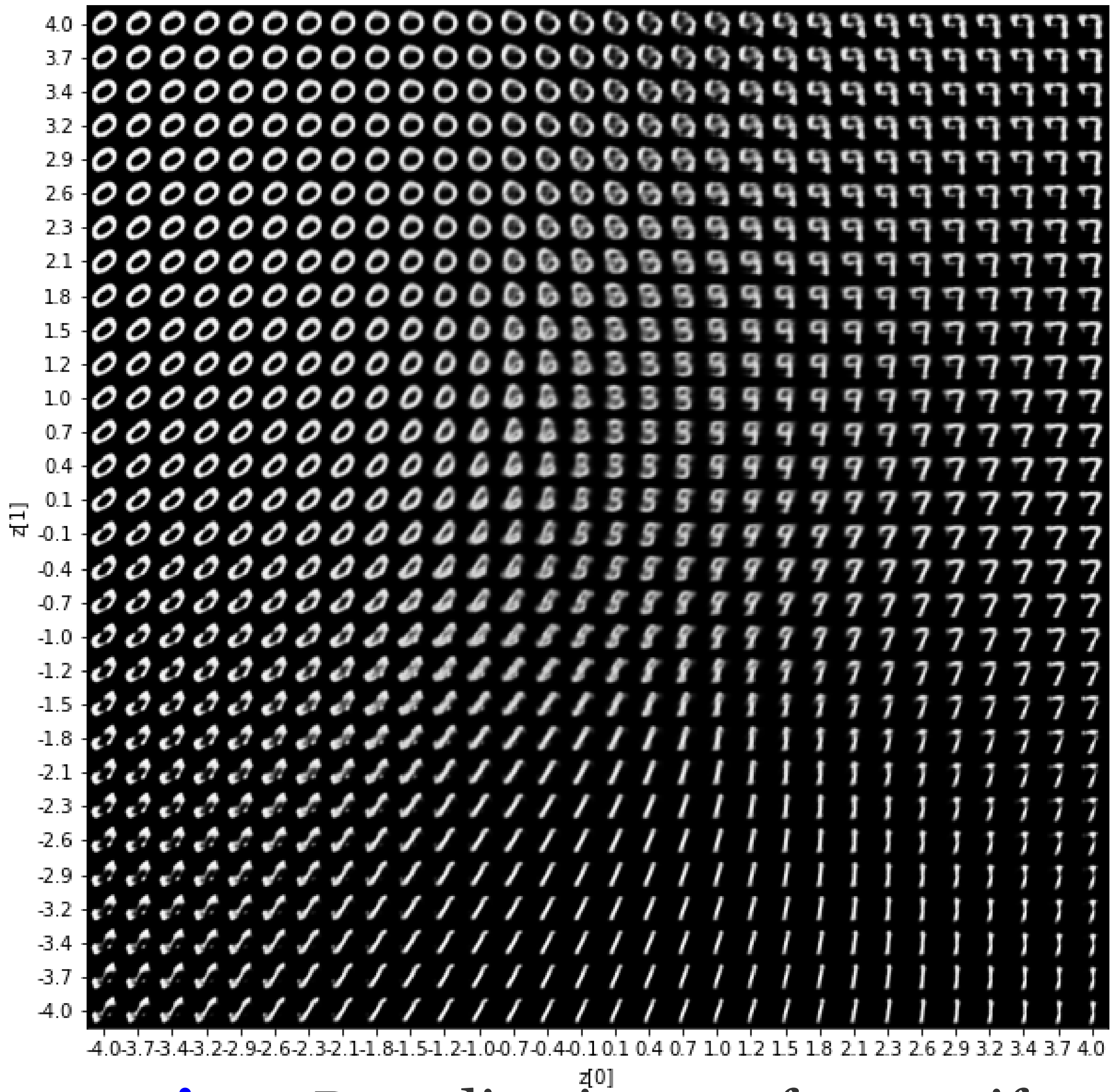


Encoding: Image → Latent Space



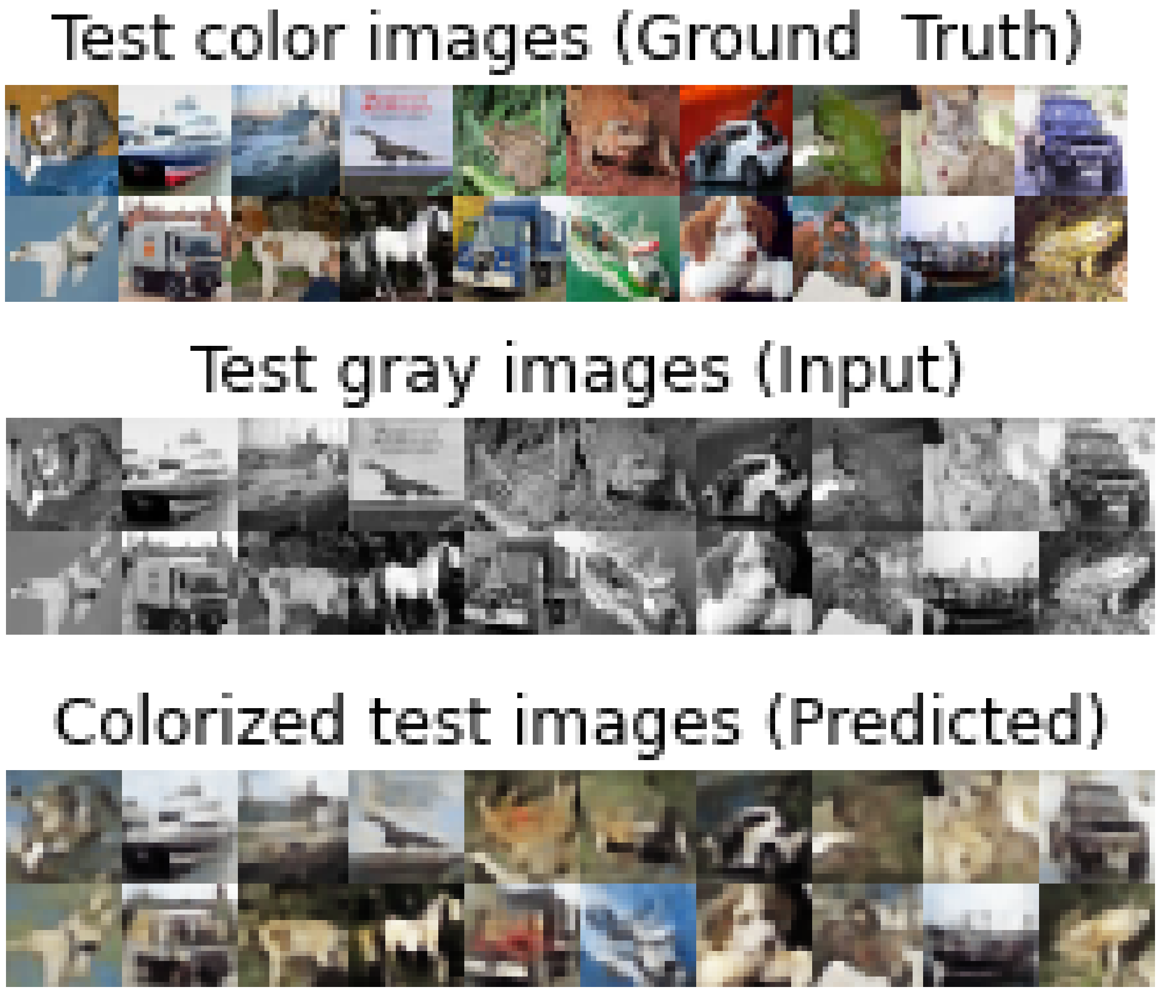
Latent distribution: Encoding the data via low dimensional vector

Decoding: Latent Space → Image



Latent representation: Decoding images from uniform grid in latent space

Image Reconstruction



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

Image Denoising

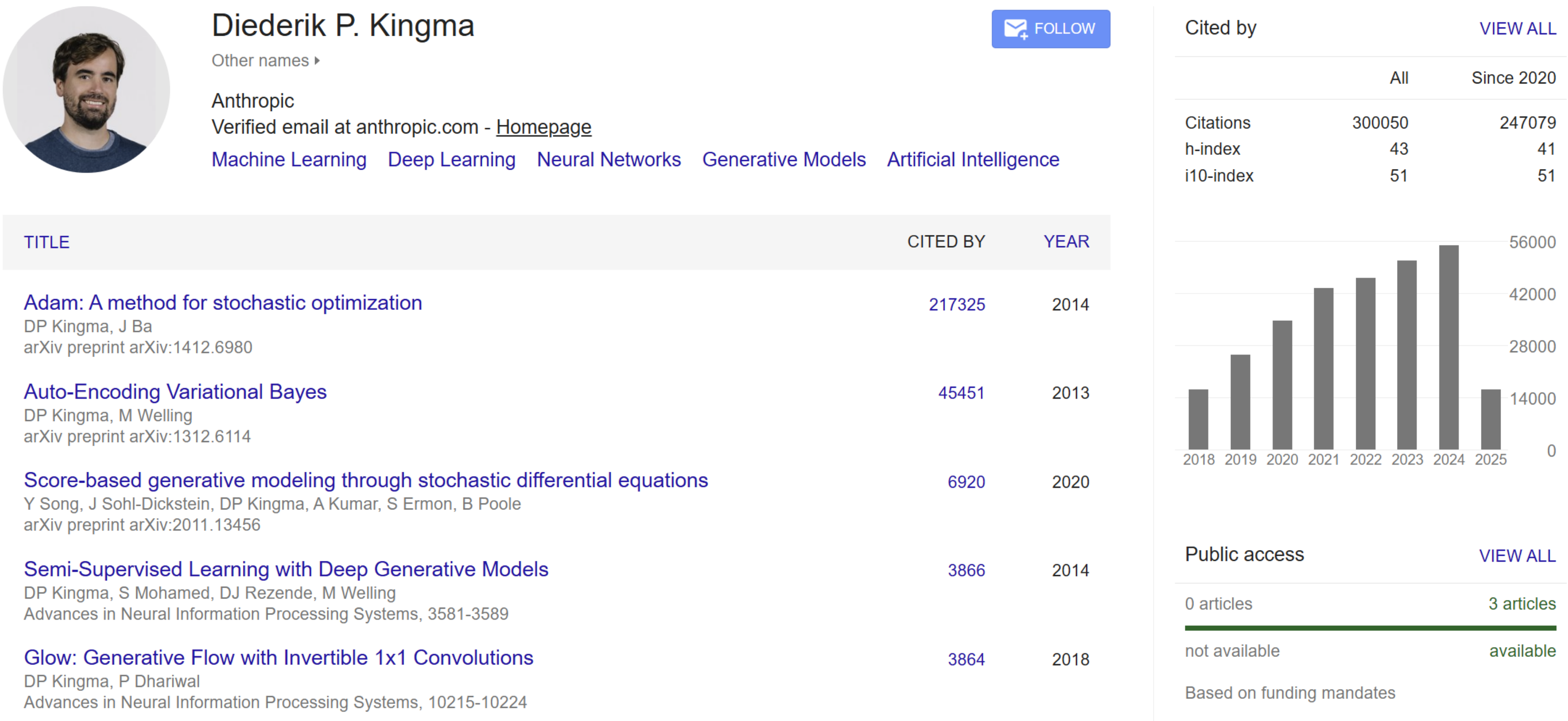
Ground truth

Noisy input

Reconstruction

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

Variational Autoencoders



Diederik P. Kingma

FOLLOW

Other names

Anthropic

Verified email at anthropic.com - Homepage

Machine Learning Deep Learning Neural Networks Generative Models Artificial Intelligence

TITLE	CITED BY	YEAR
Adam: A method for stochastic optimization DP Kingma, J Ba arXiv preprint arXiv:1412.6980	217325	2014
Auto-Encoding Variational Bayes DP Kingma, M Welling arXiv preprint arXiv:1312.6114	45451	2013
Score-based generative modeling through stochastic differential equations Y Song, J Sohl-Dickstein, DP Kingma, A Kumar, S Ermon, B Poole arXiv preprint arXiv:2011.13456	6920	2020
Semi-Supervised Learning with Deep Generative Models DP Kingma, S Mohamed, DJ Rezende, M Welling Advances in Neural Information Processing Systems, 3581-3589	3866	2014
Glow: Generative Flow with Invertible 1x1 Convolutions DP Kingma, P Dhariwal Advances in Neural Information Processing Systems, 10215-10224	3864	2018

Cited by VIEW ALL

	All	Since 2020
Citations	300050	247079
h-index	43	41
i10-index	51	51

Public access VIEW ALL

0 articles 3 articles

not available available

Based on funding mandates

- 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

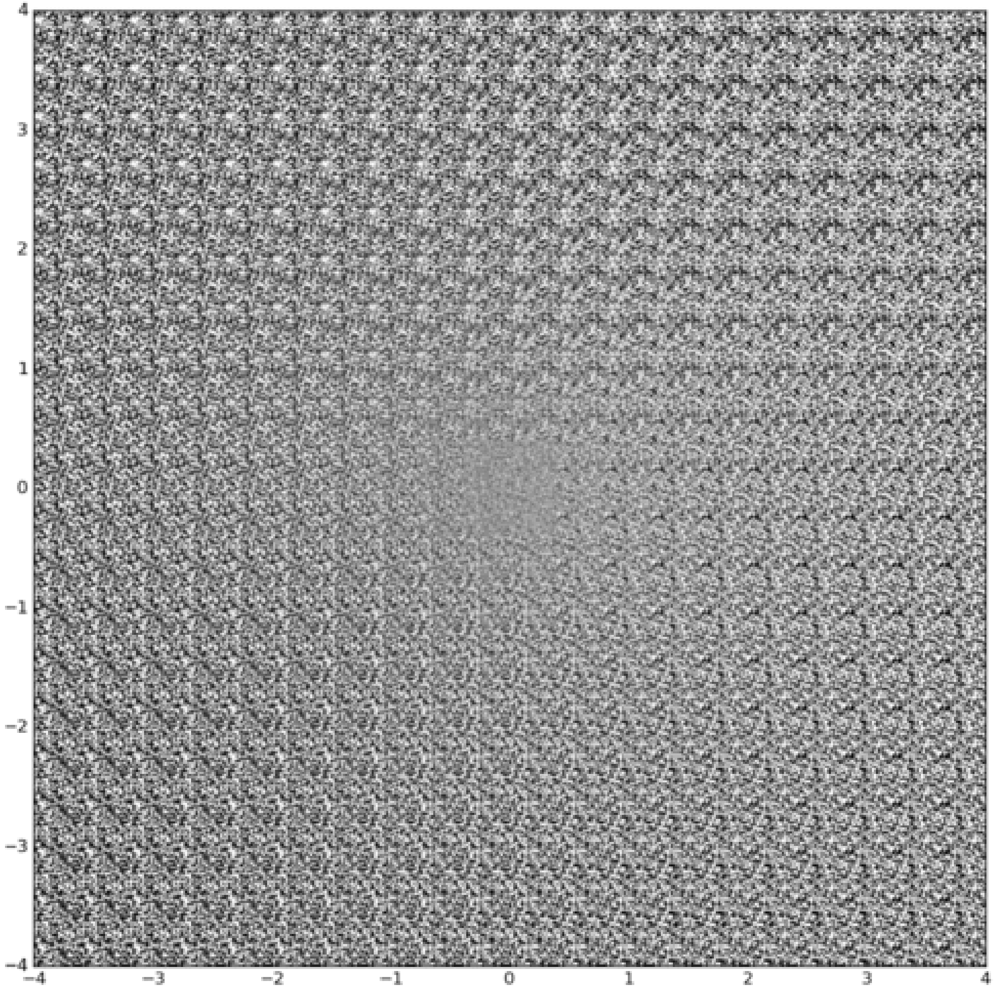
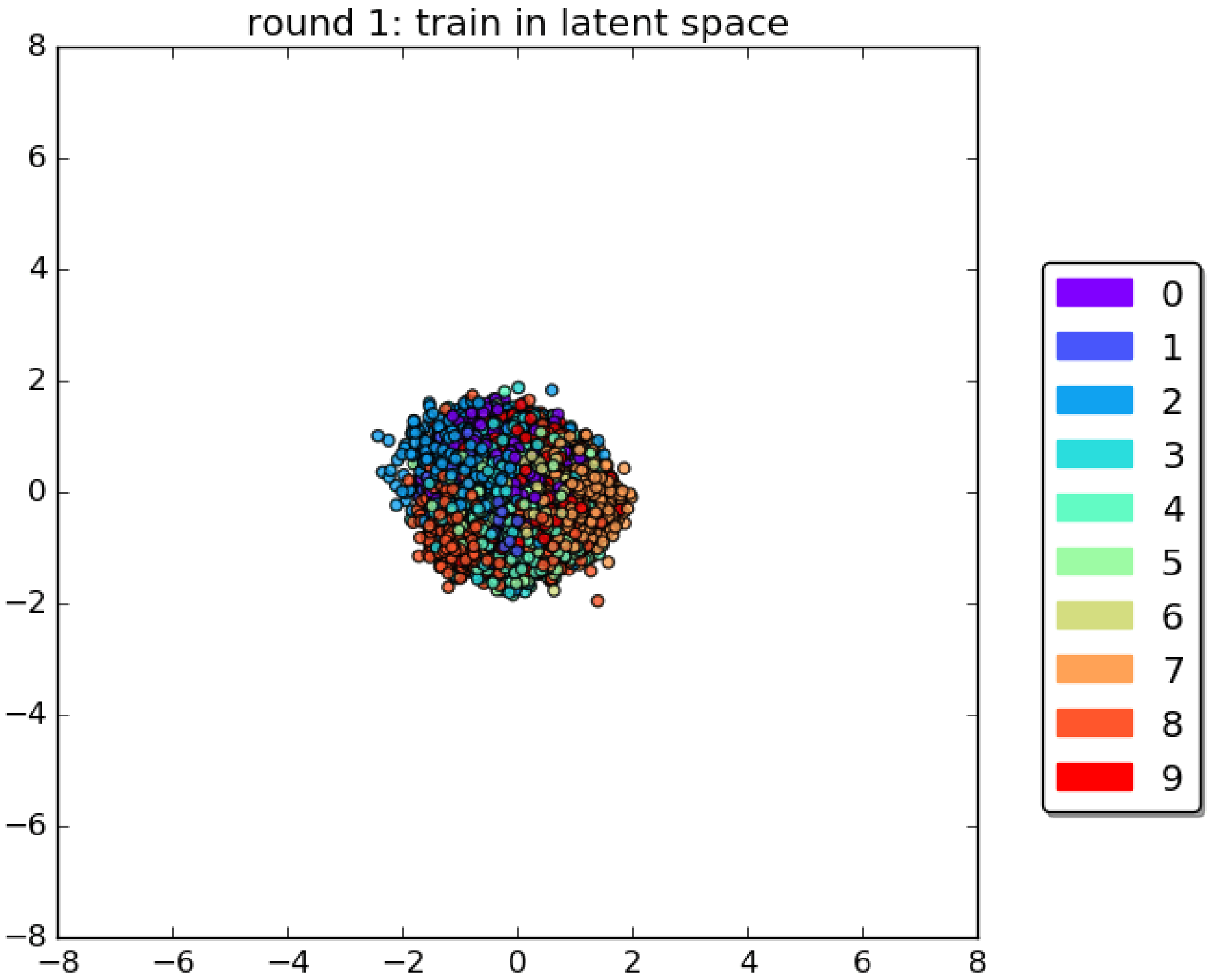
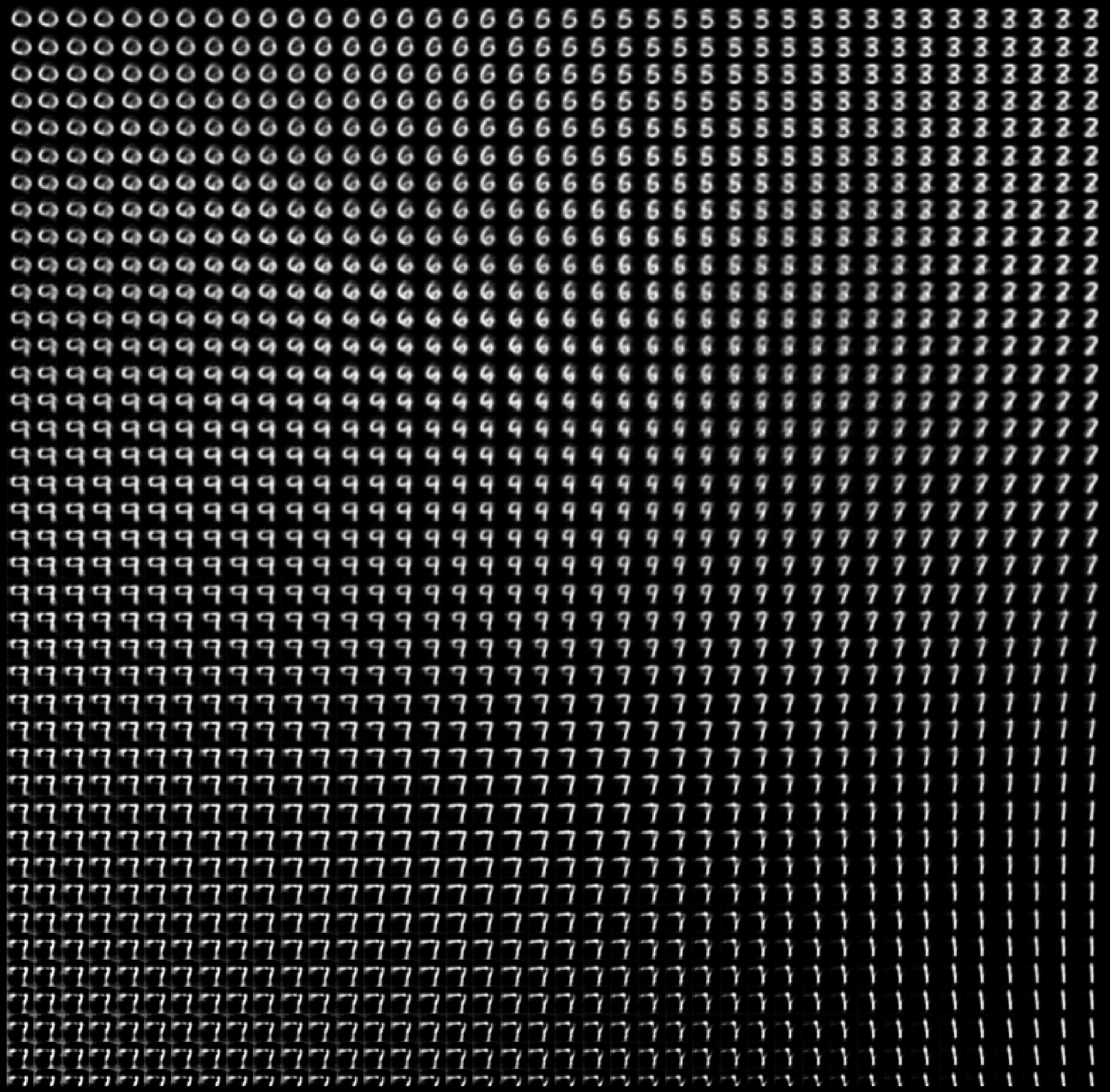


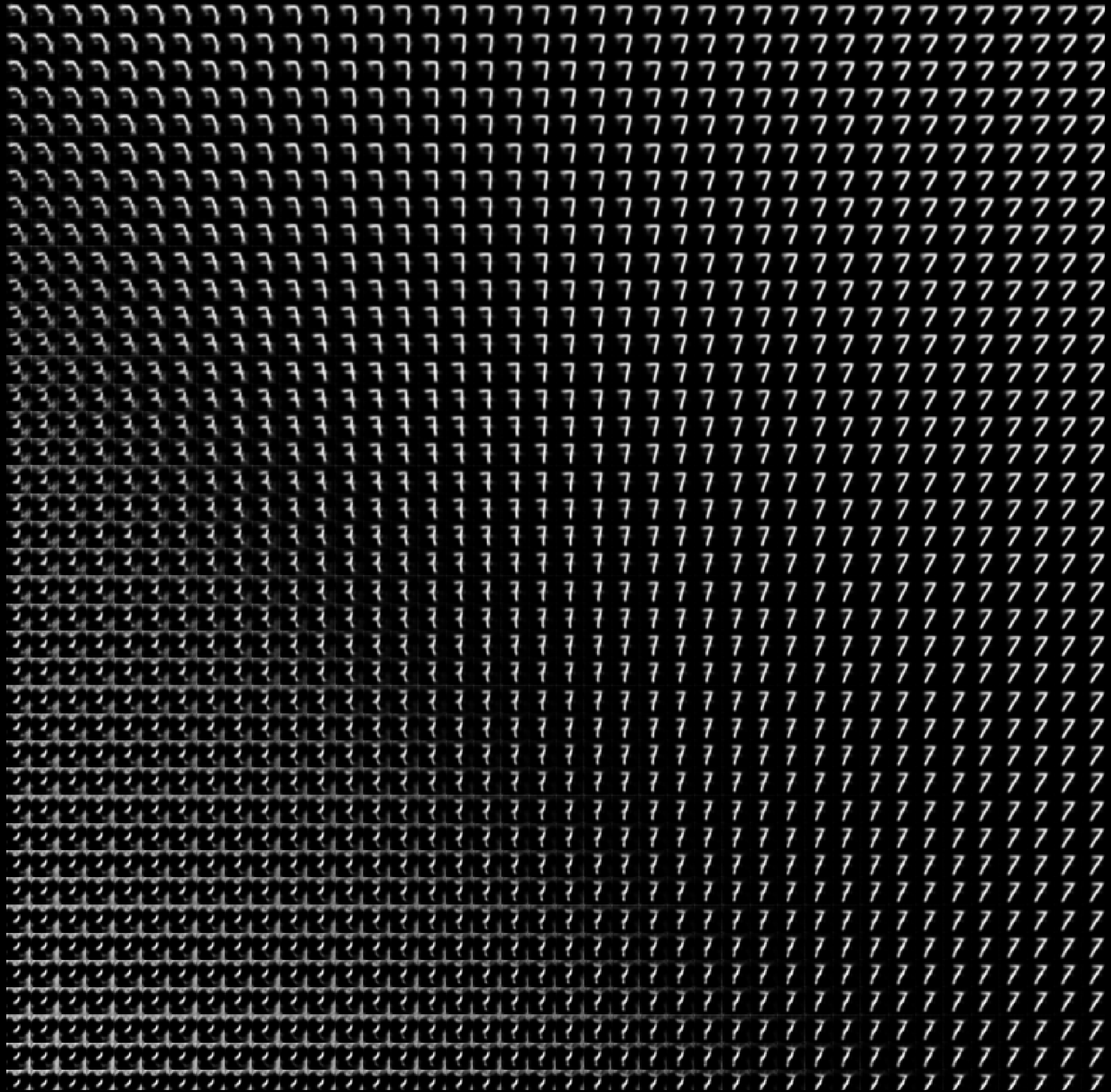
Image space → Latent space



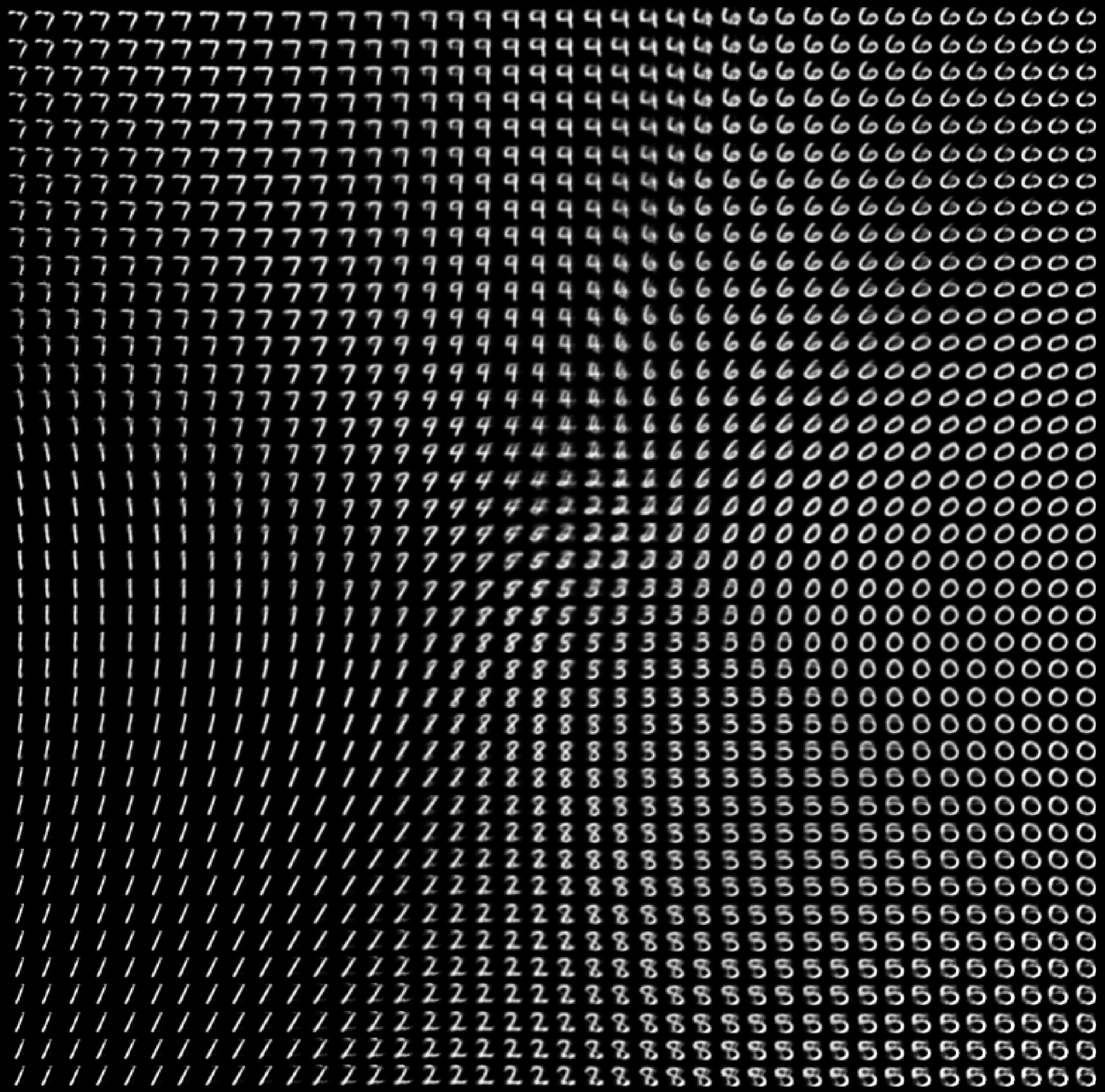
Autoencoder latent representation



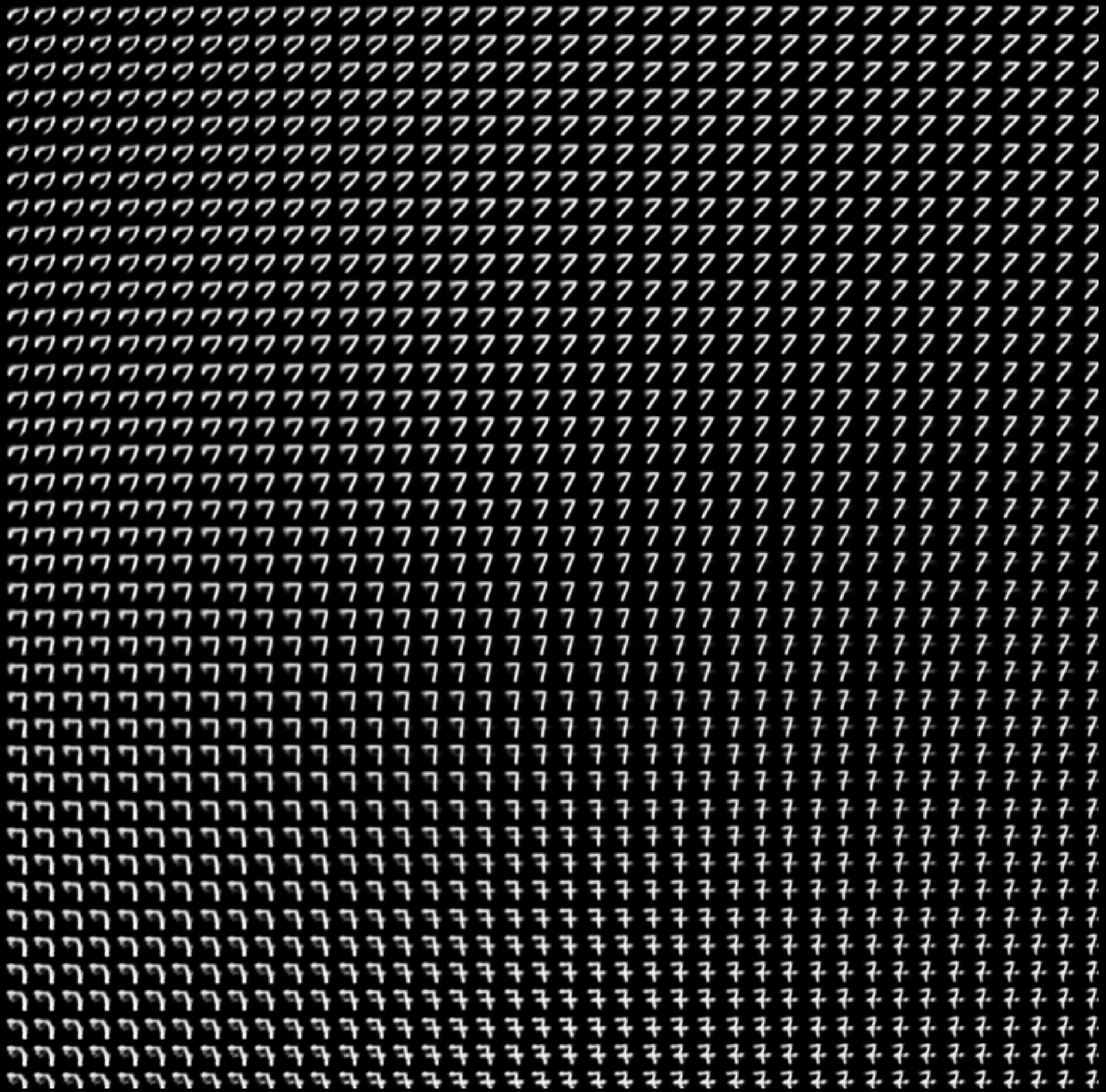
Autoencoder latent representation (digit 7)



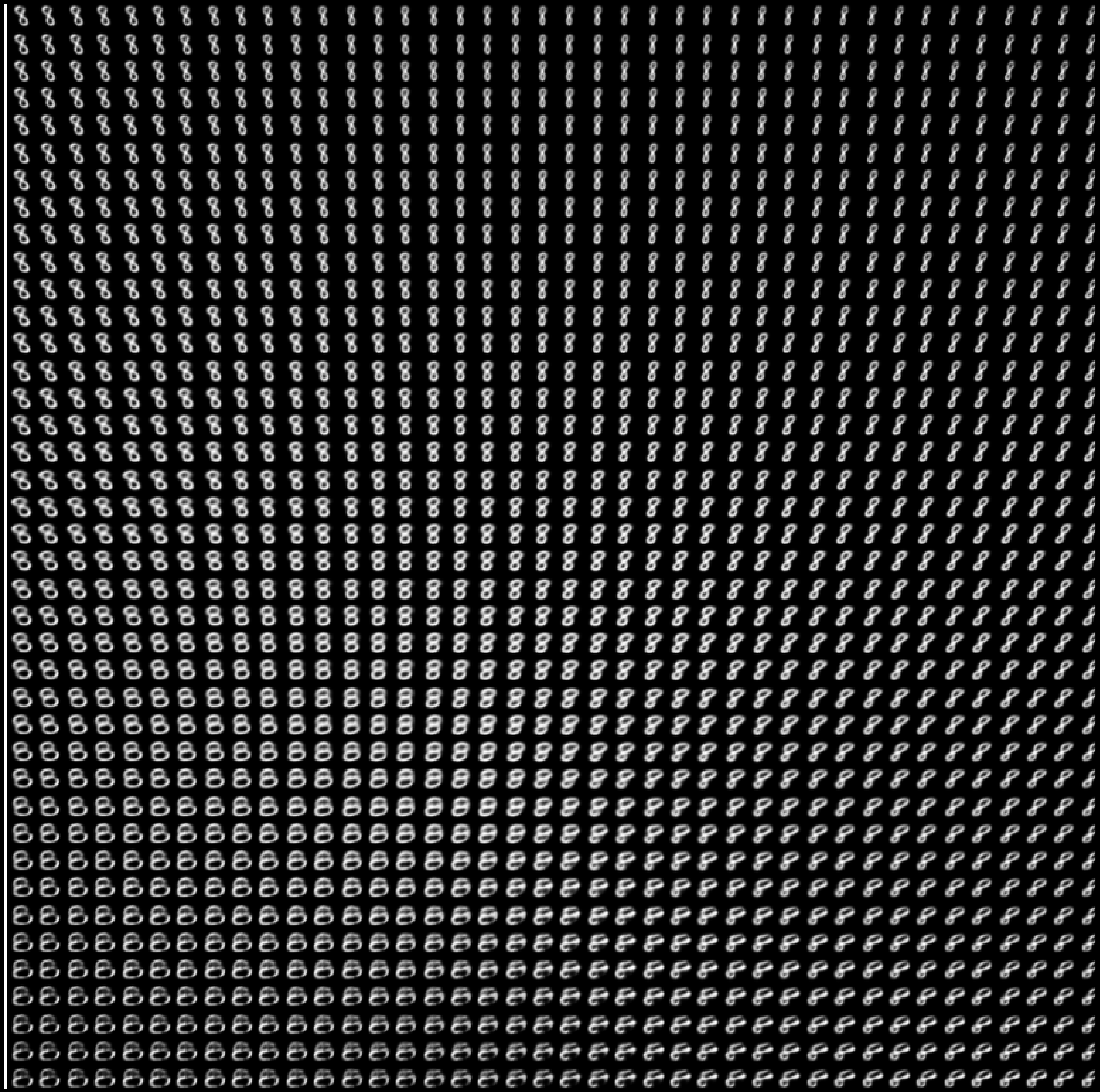
VAE latent representation



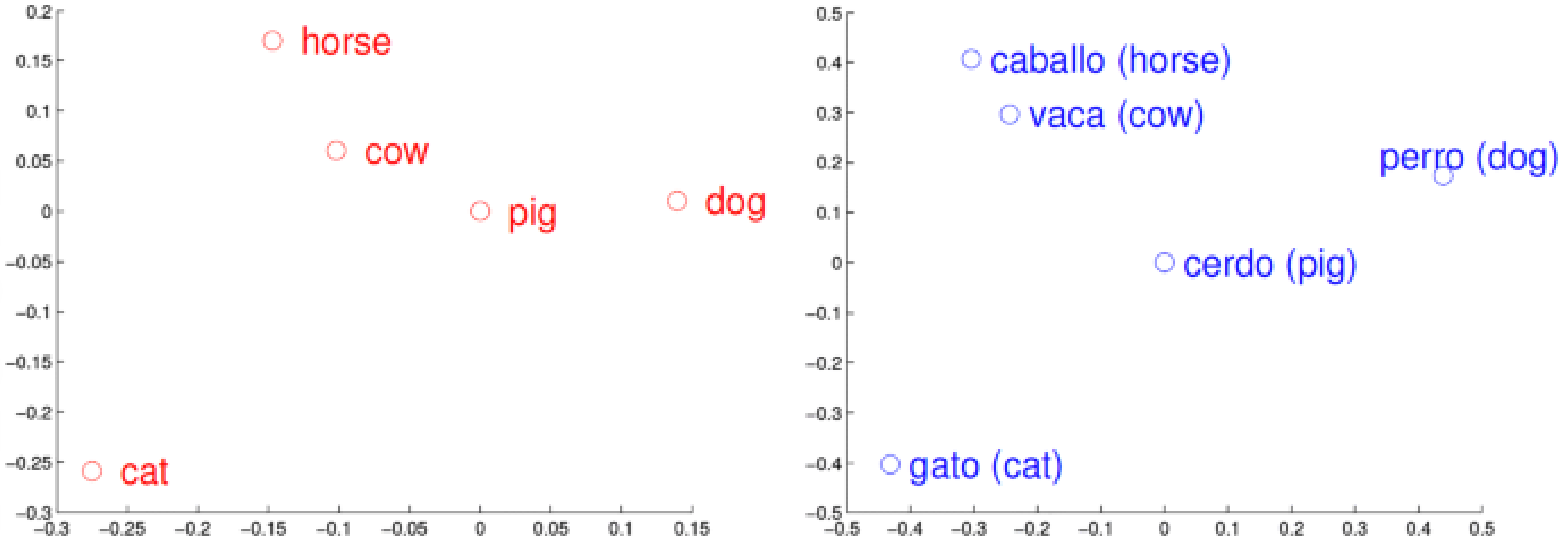
VAE latent representation (digit 7)



VAE latent representation (digit 8)

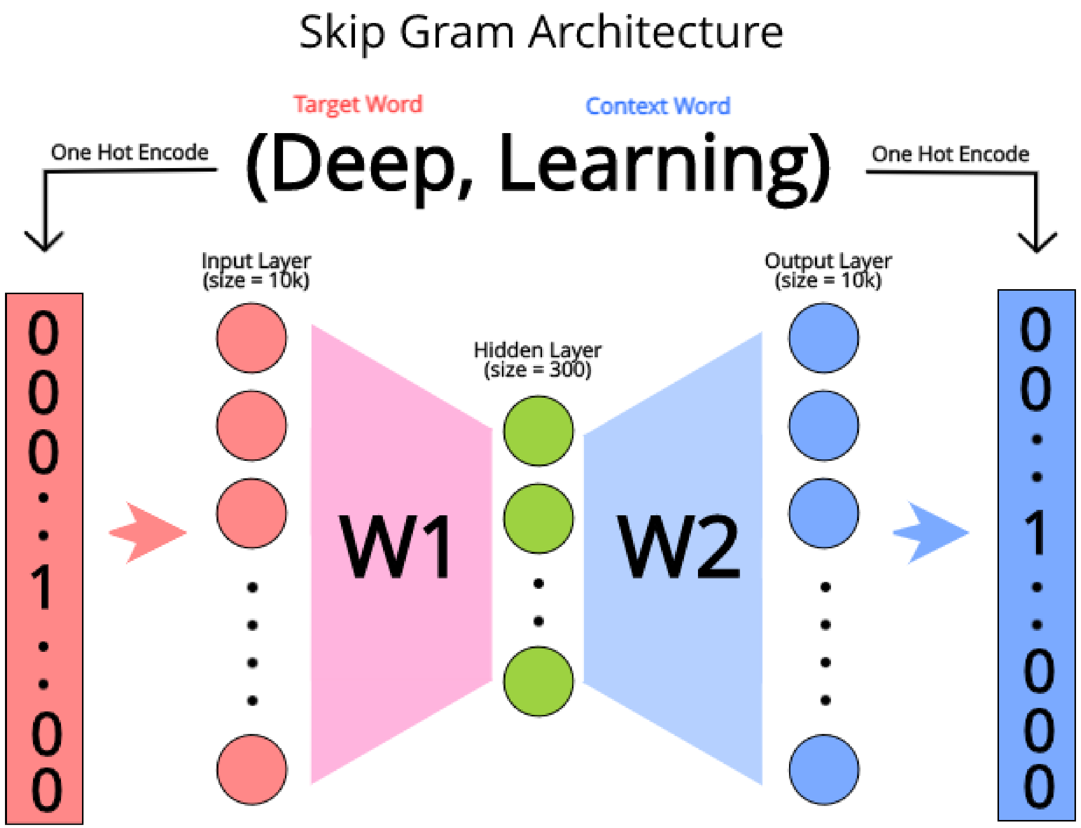


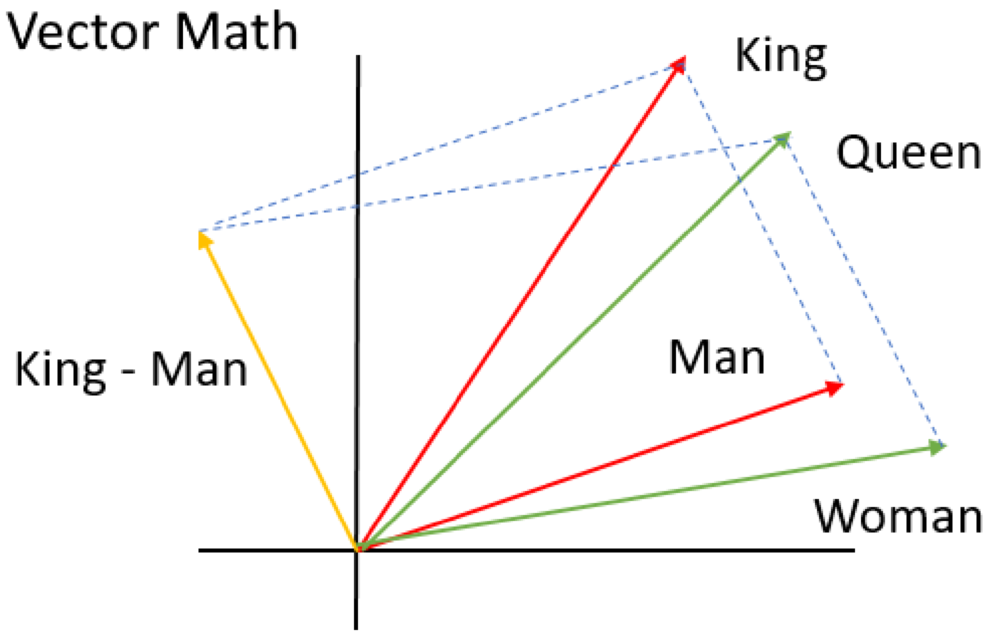
Word Embeddings



https://medium.com/geekculture/word-embeddings-in-ai-10a9e430cb59

Word Vectors



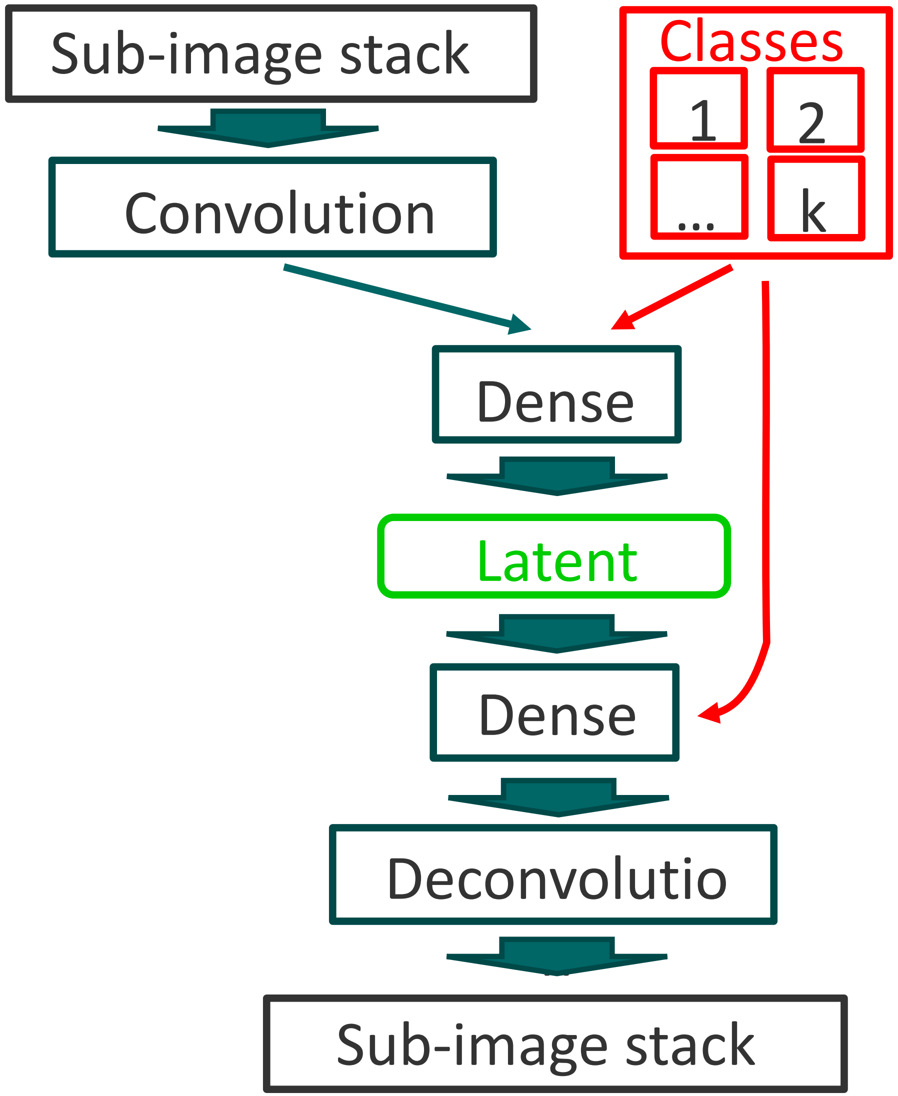


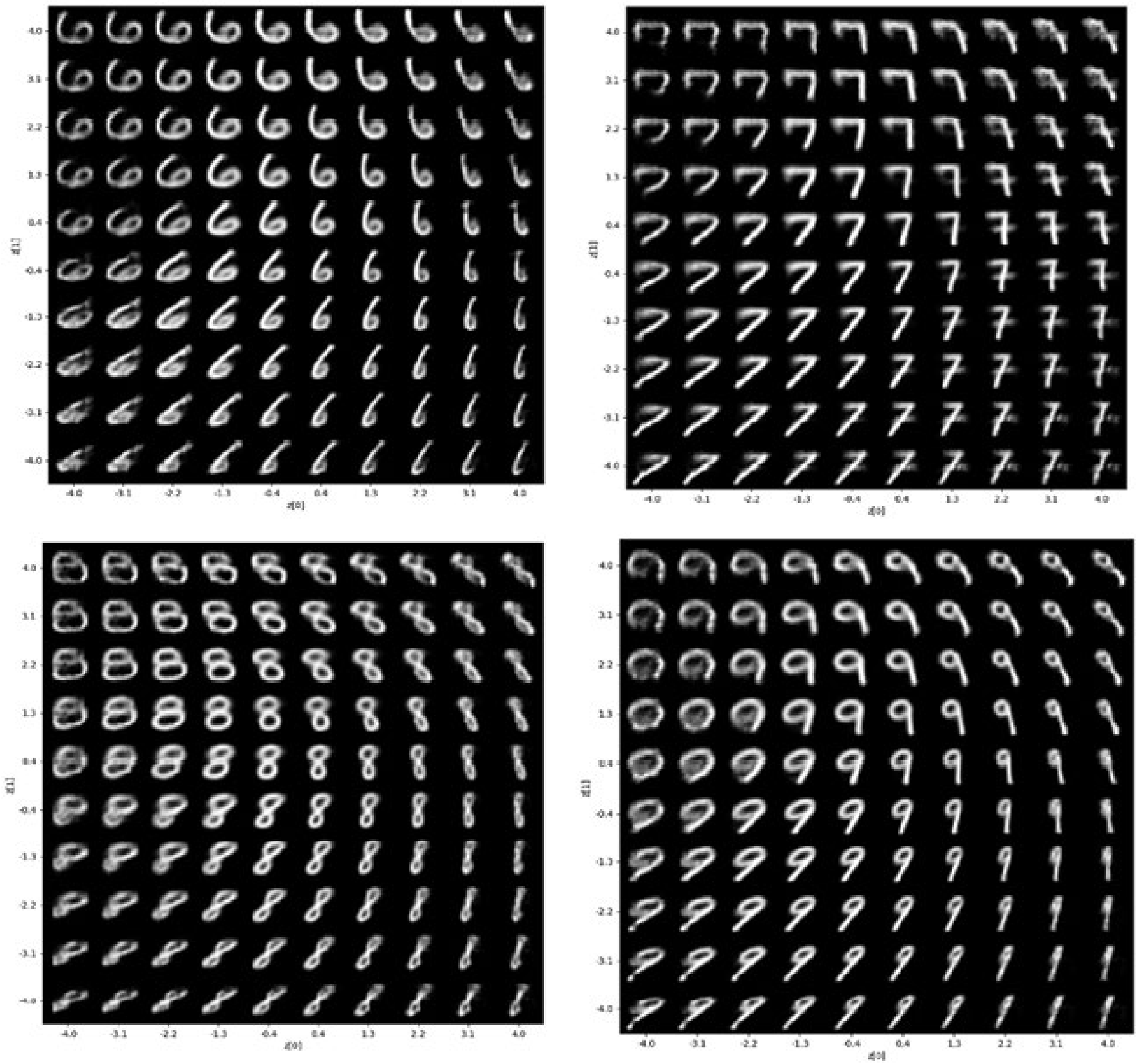
https://medium.com/analytics-vidhya/word-embeddings-in-nlp-word2vec-glove-fasttext-24d4d4286a73

Changing Attributes

https://rtoledo.me/post/2021-05-31-edit-face-attributes-using-vae/edit-face-attributes-using-vae/

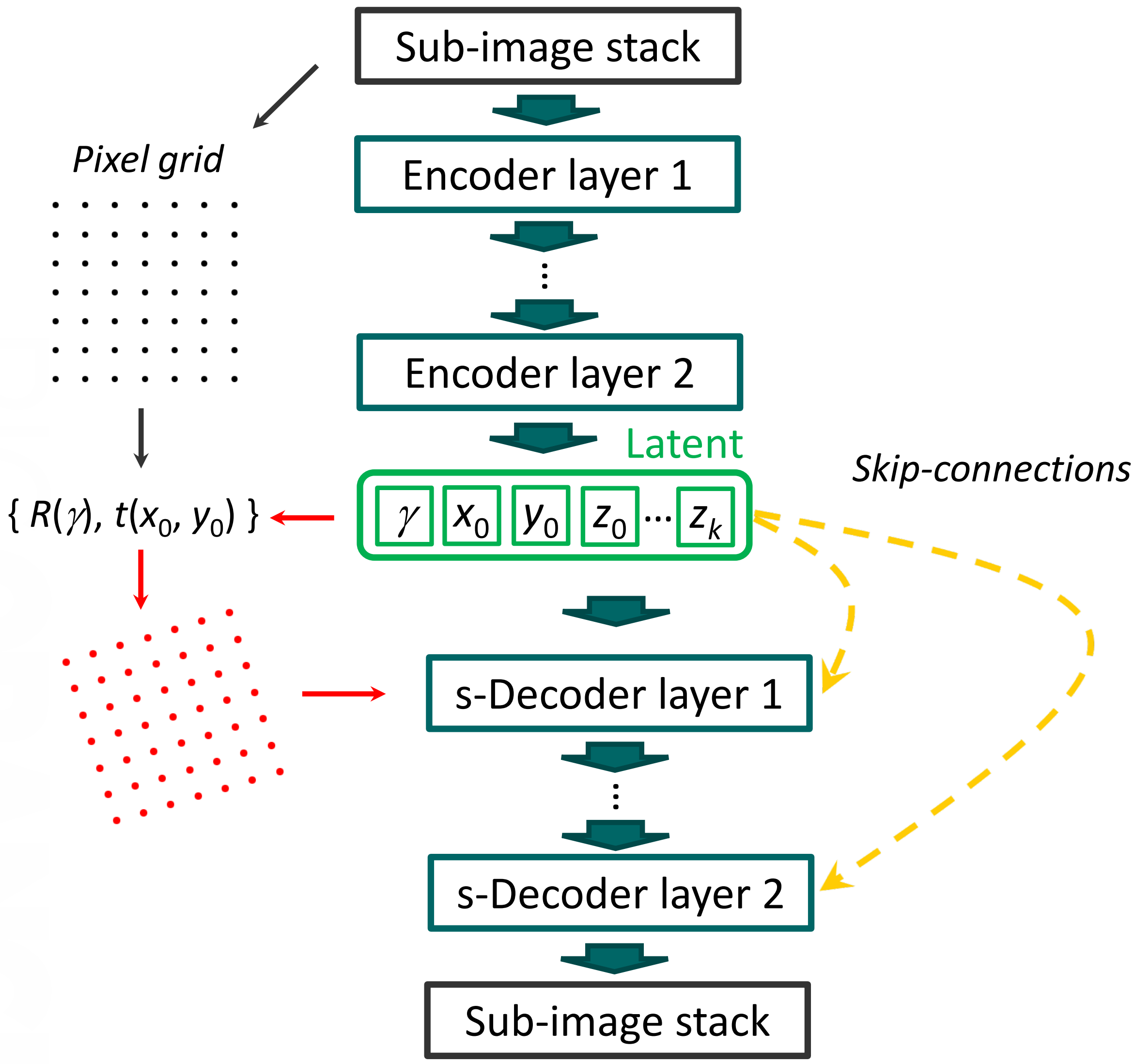
Conditional VAE





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

Invariant 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

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

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

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

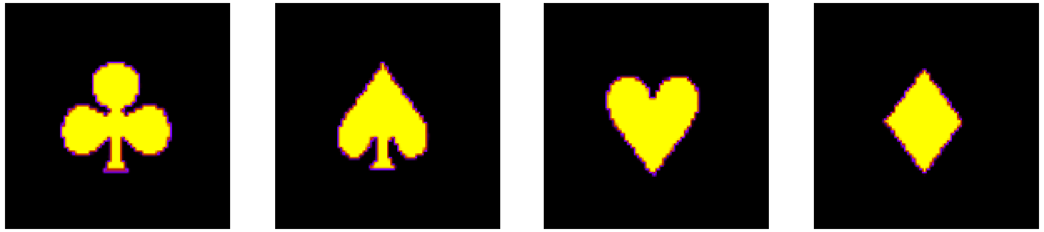
+ D_{KL}(physics-based “priors”) ?

+ D (physics) ?

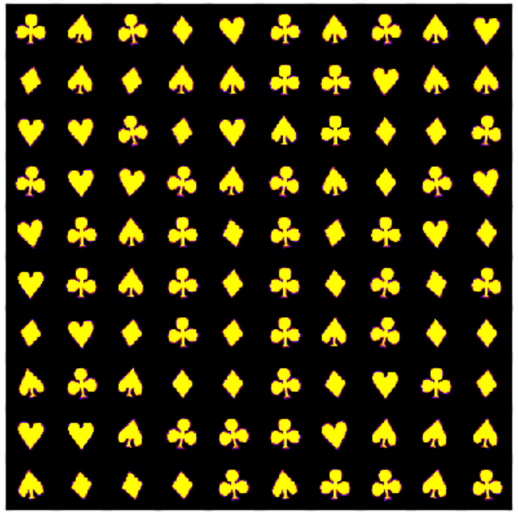
Bepler *et al.*, NeurIPS 2019

(R)VAE on Cards

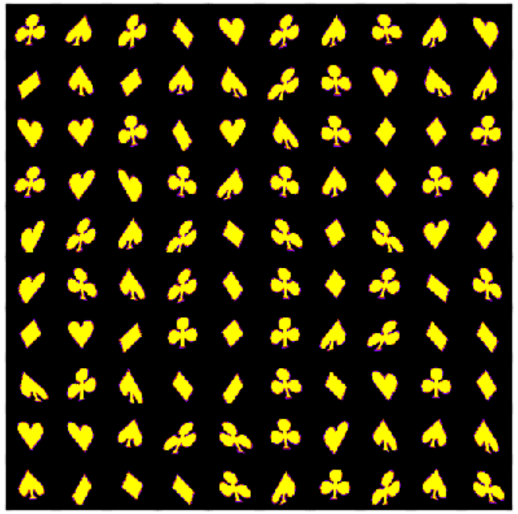
Introduce the **cards** data set:
- Classical 4 hands (diamonds, clubs, pikes, hearts)
- Interesting similarities (pikes and hearts)
- And invariances on affine transforms (e.g. diamonds)



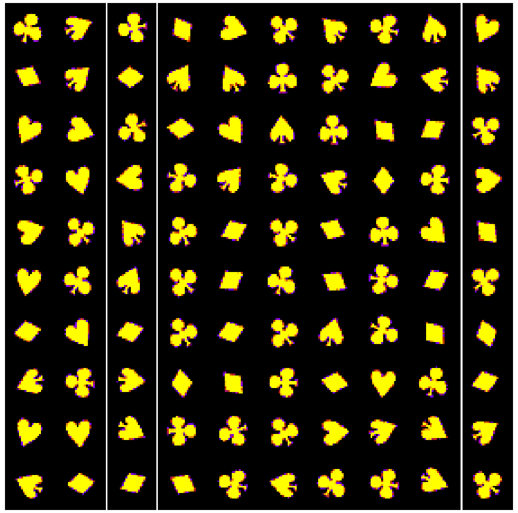
Cards 1: Low R (12 deg) and low S (1 deg)



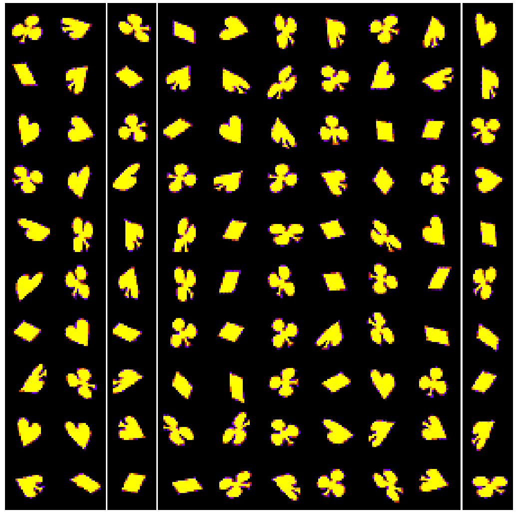
Cards 2: Low R (12 deg) and high S (20 deg)



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



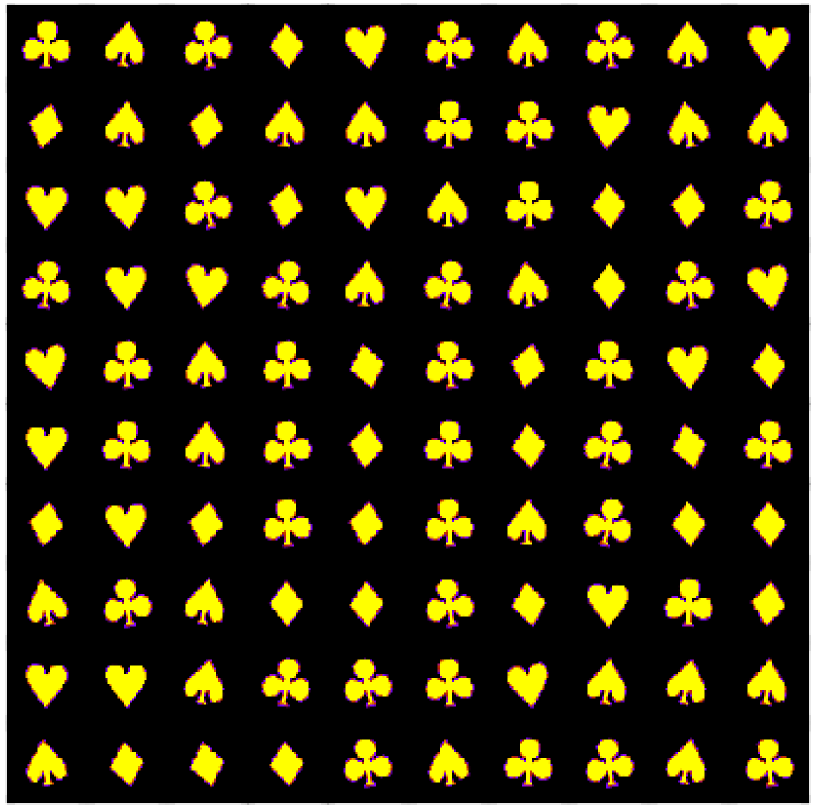
Cards 4: High R (120 deg) and high S (20 deg)



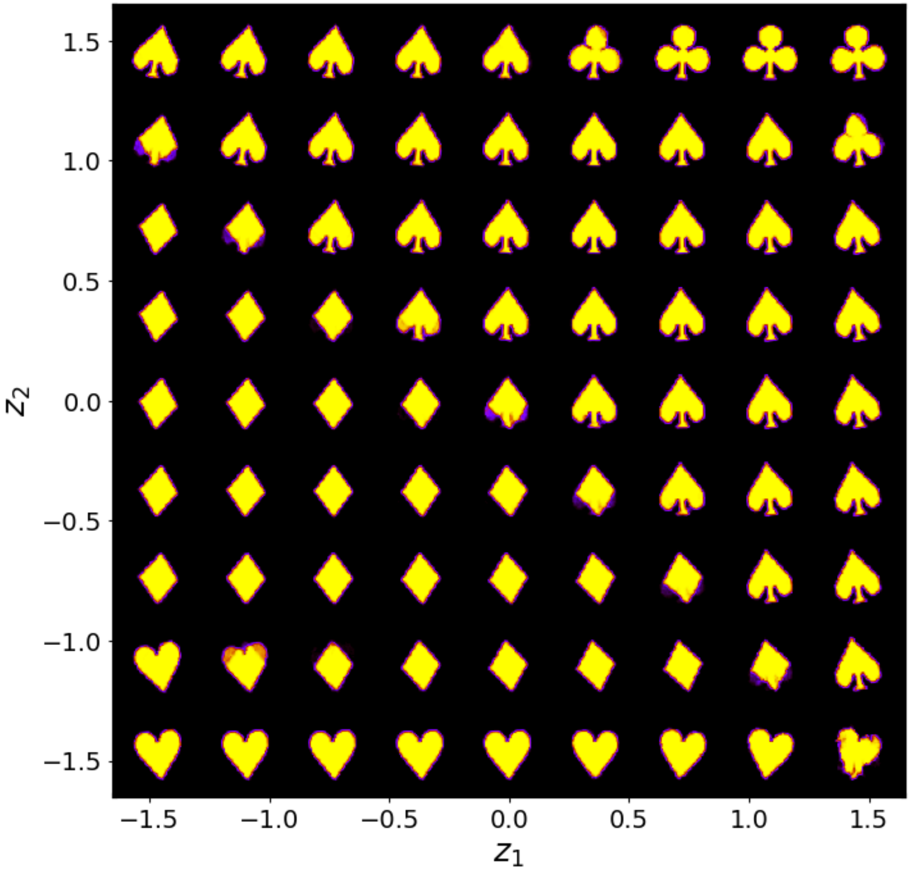
- Shear, rotations, and translations are **known** factors of variability (or traits) in data
- Can VAE disentangle representations and **discover** these factors of variability

VAE on Cards

Example of data

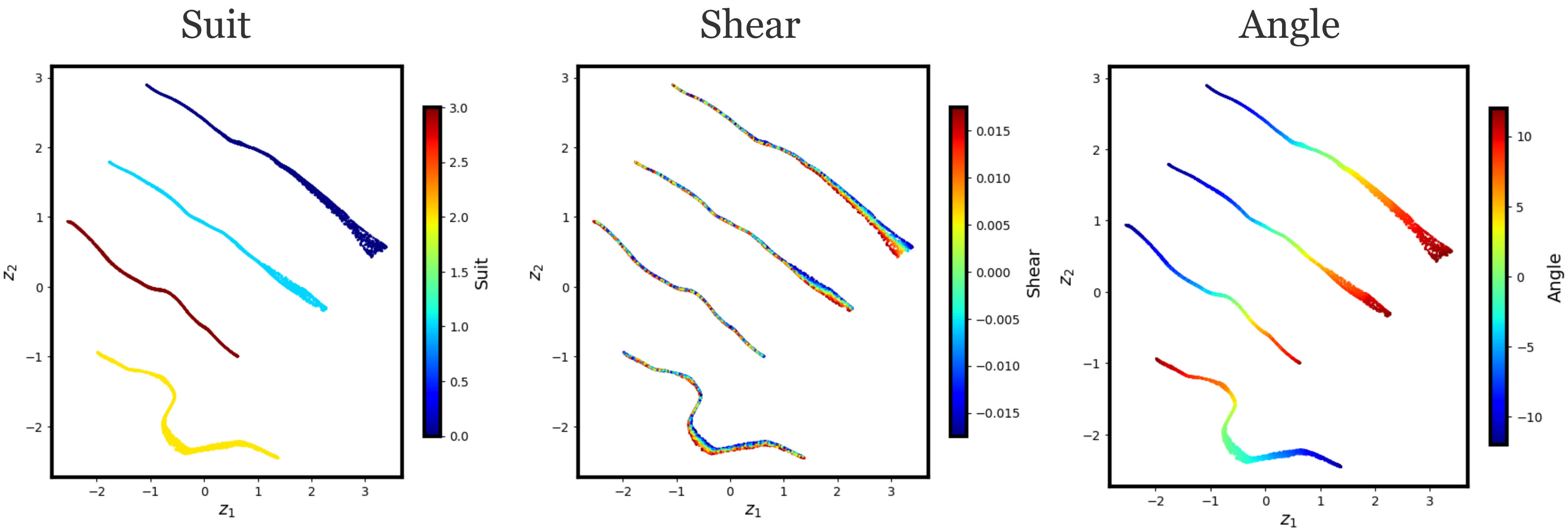


Latent representation



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

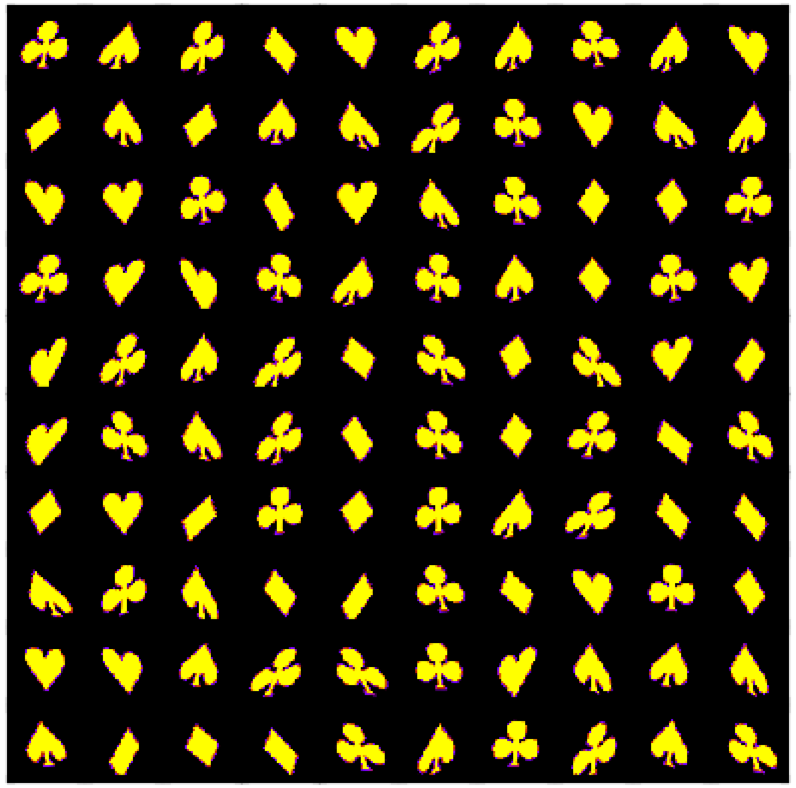
VAE on Cards



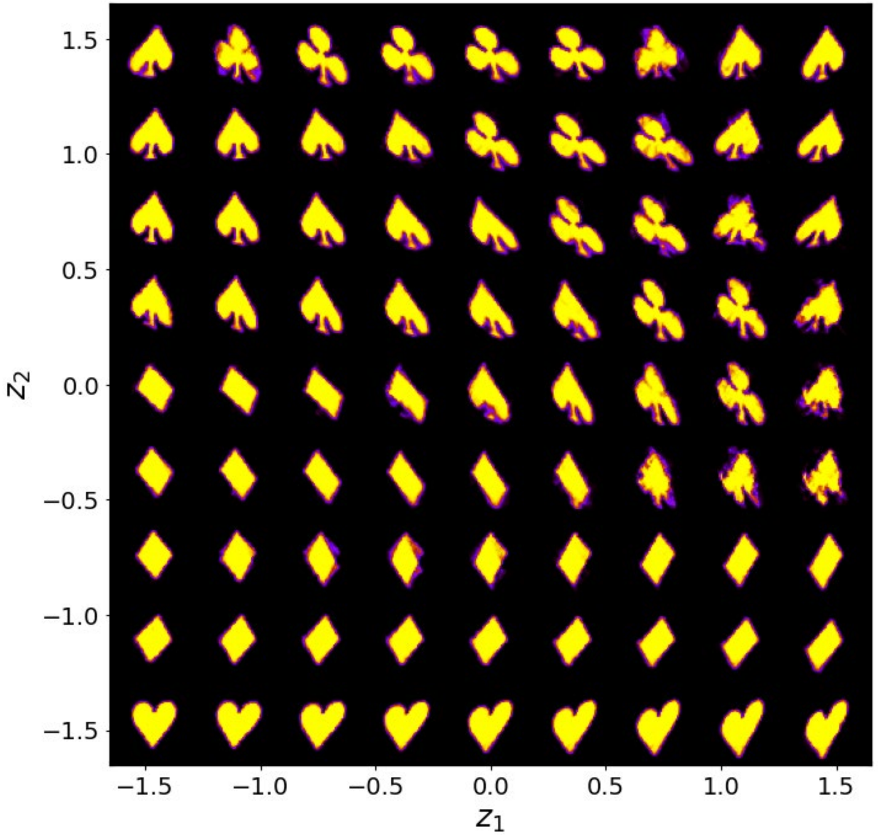
Cards 1: Low rotation (12 deg) and low shear (1 deg)

VAE on Cards

Example of data

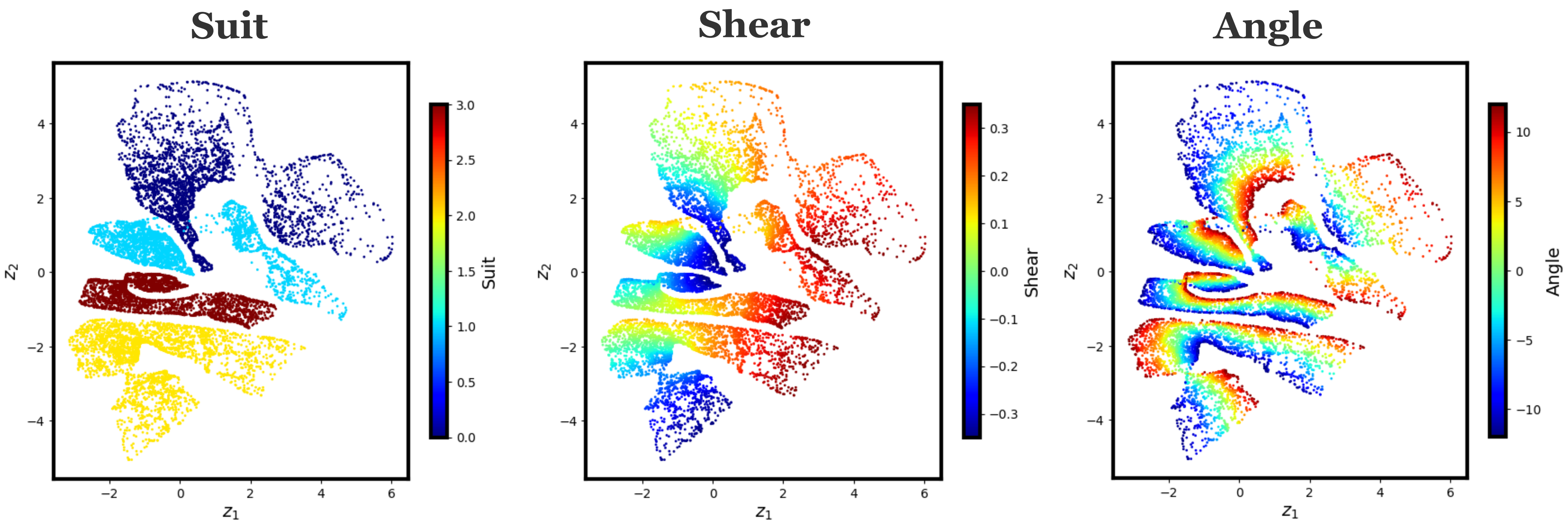


Latent representation



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

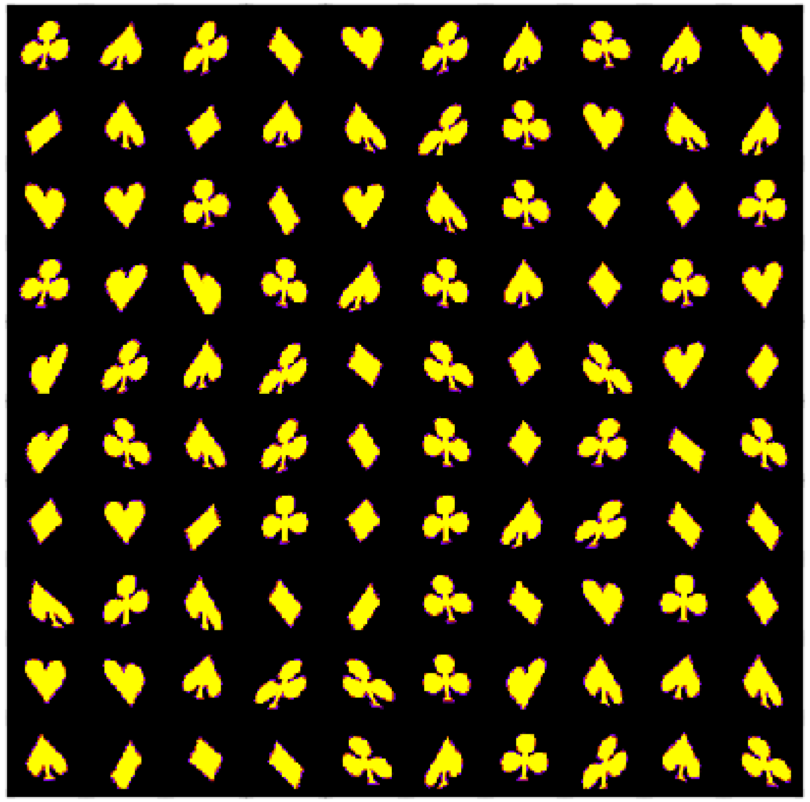
VAE on Cards



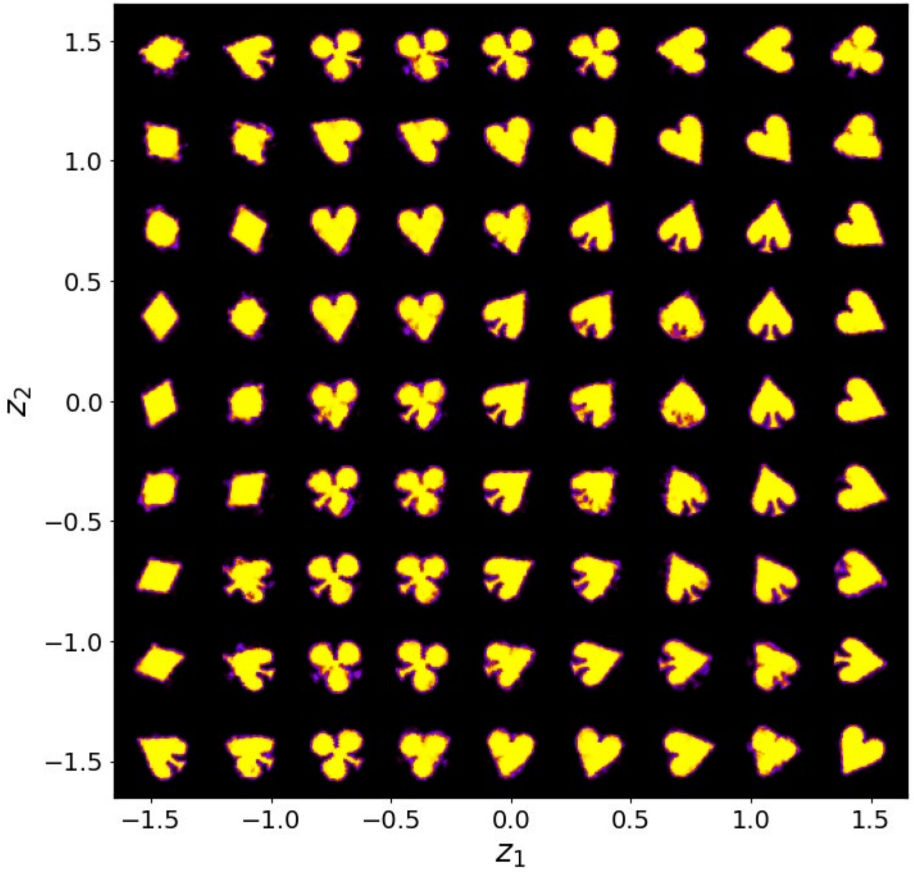
Cards 2: Low rotation (12 deg) and high shear (20 deg)

VAE on Cards

Example of data

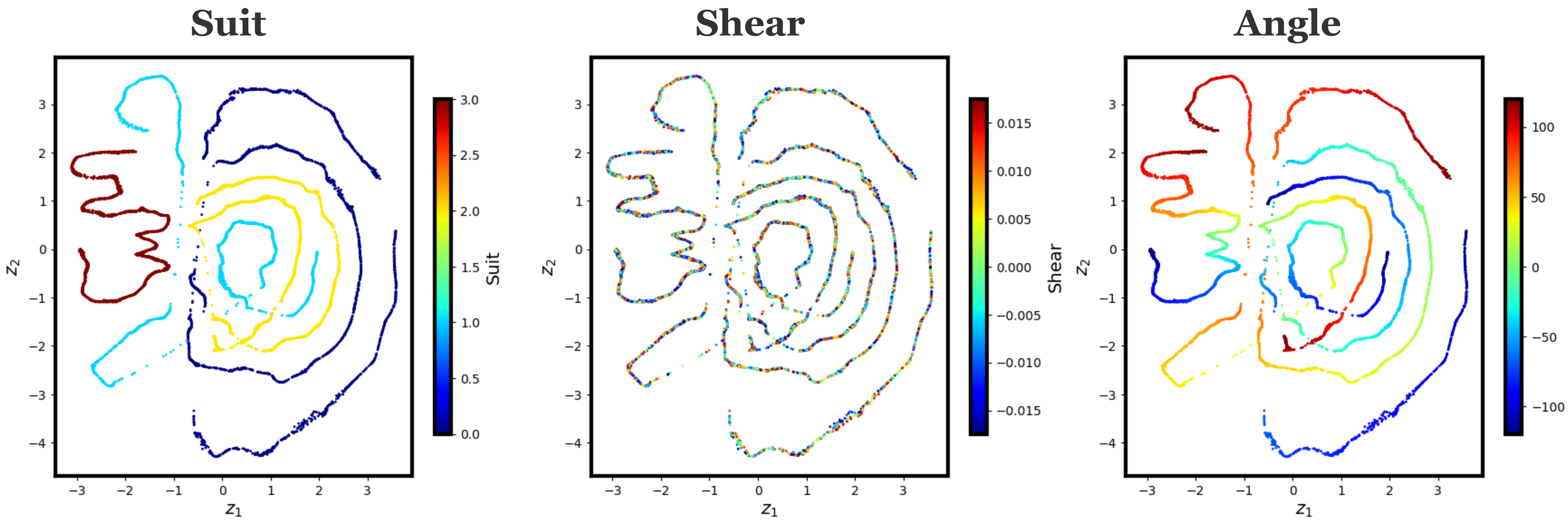


Latent representation



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

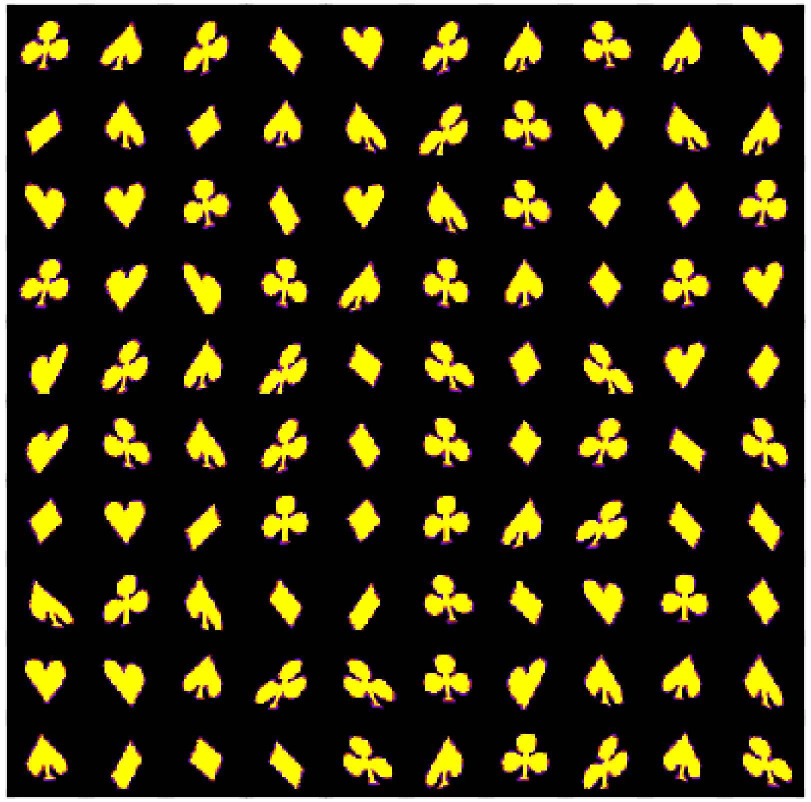
VAE on Cards



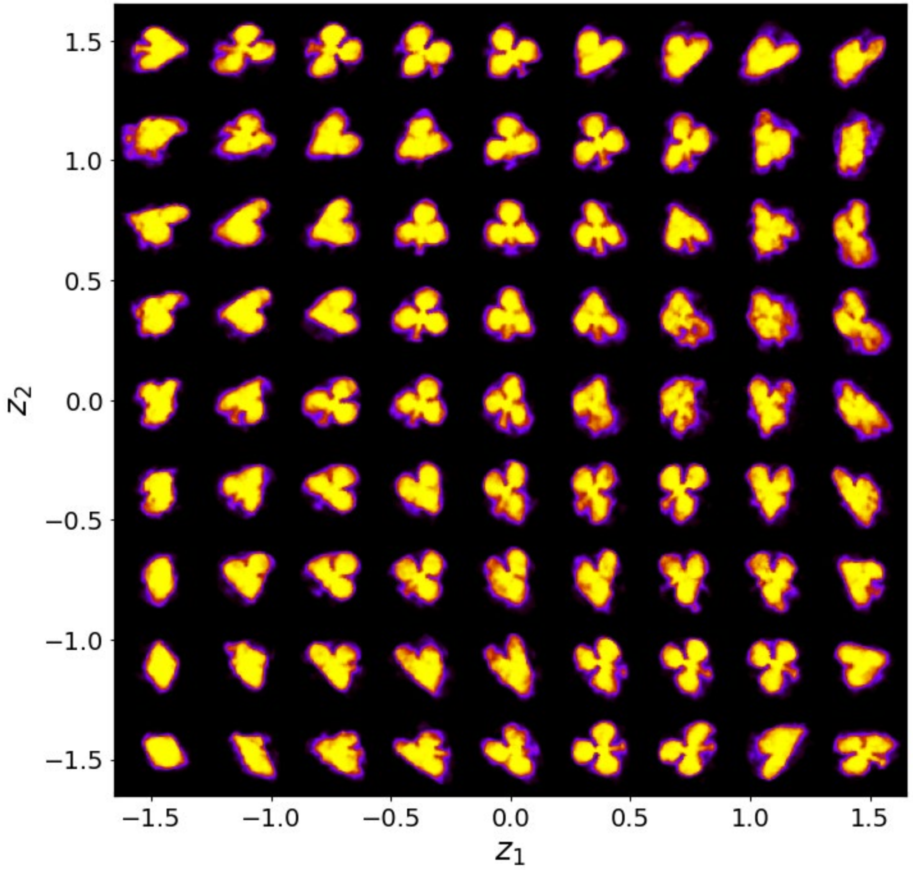
Cards 3: High rotation (120 deg) and low shear (1 deg)

VAE on Cards

Example of data

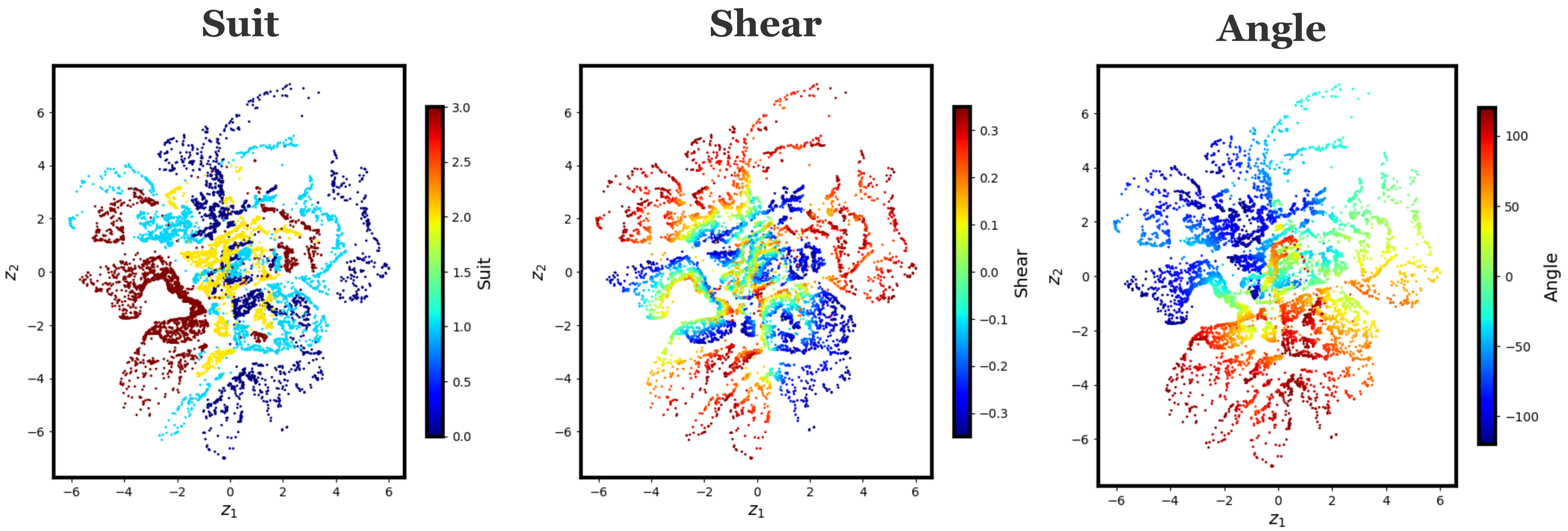


Latent representation



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

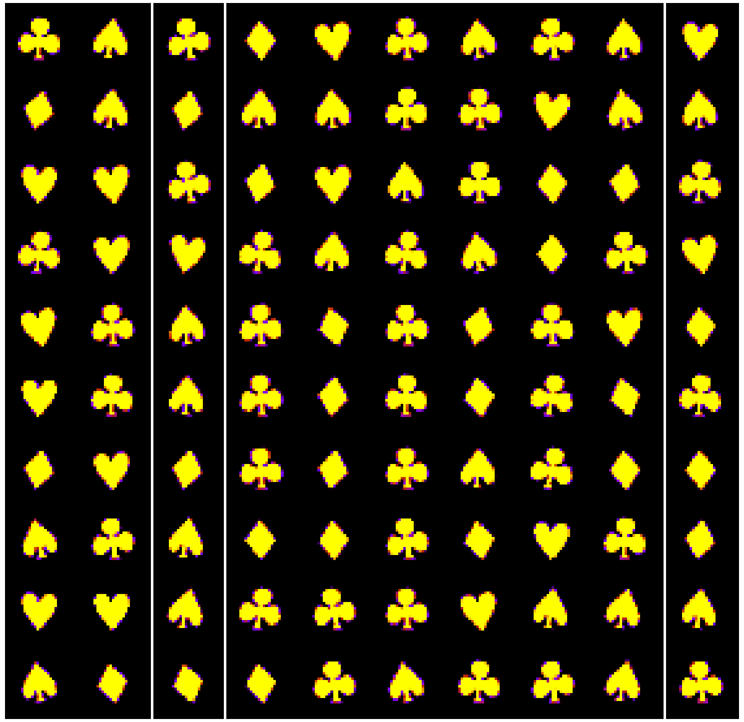
VAE on Cards



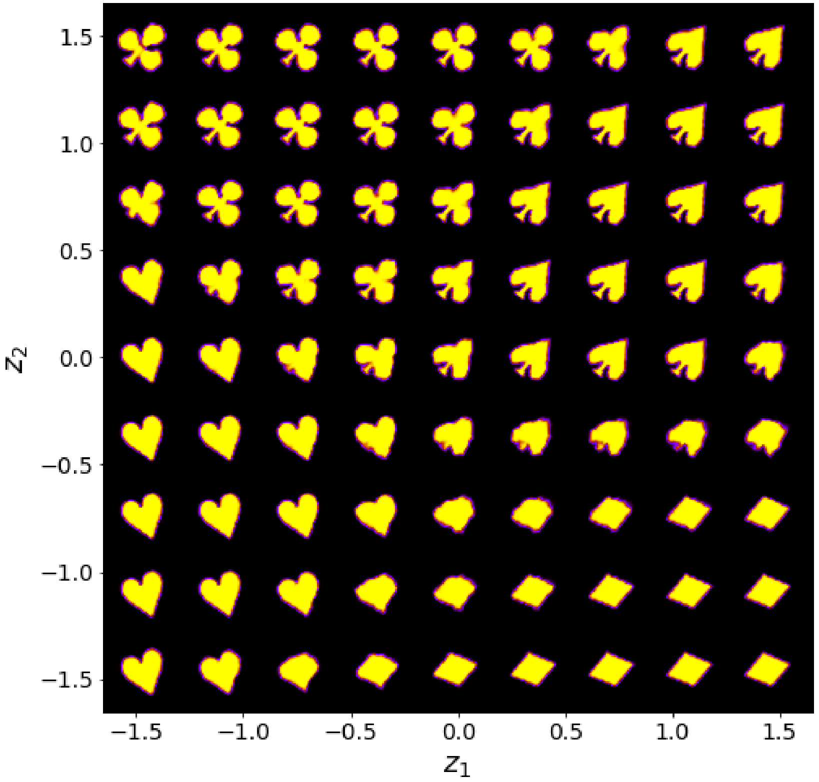
Cards 4: High rotation (120 deg) and high shear (20 deg)

rVAE on Cards

Example of data

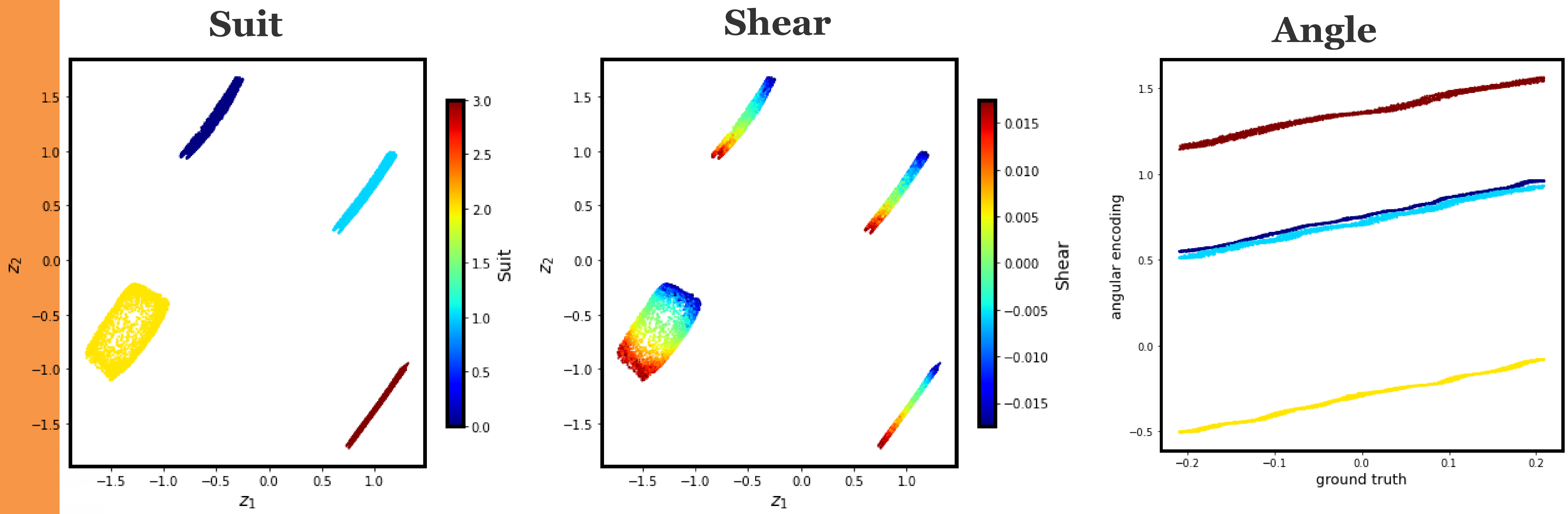


Latent representation



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

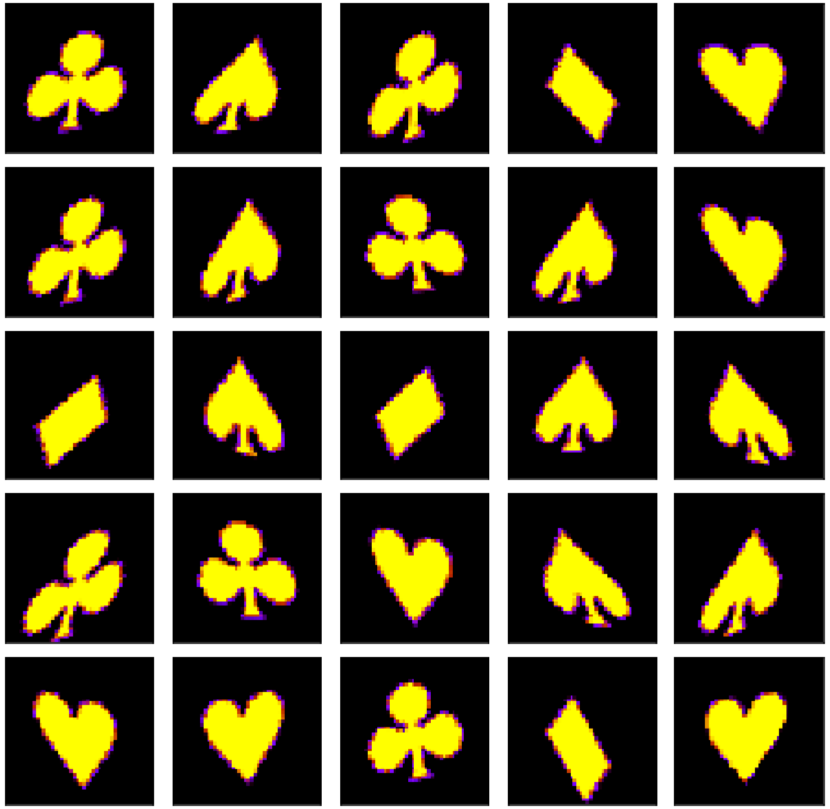
rVAE on Cards



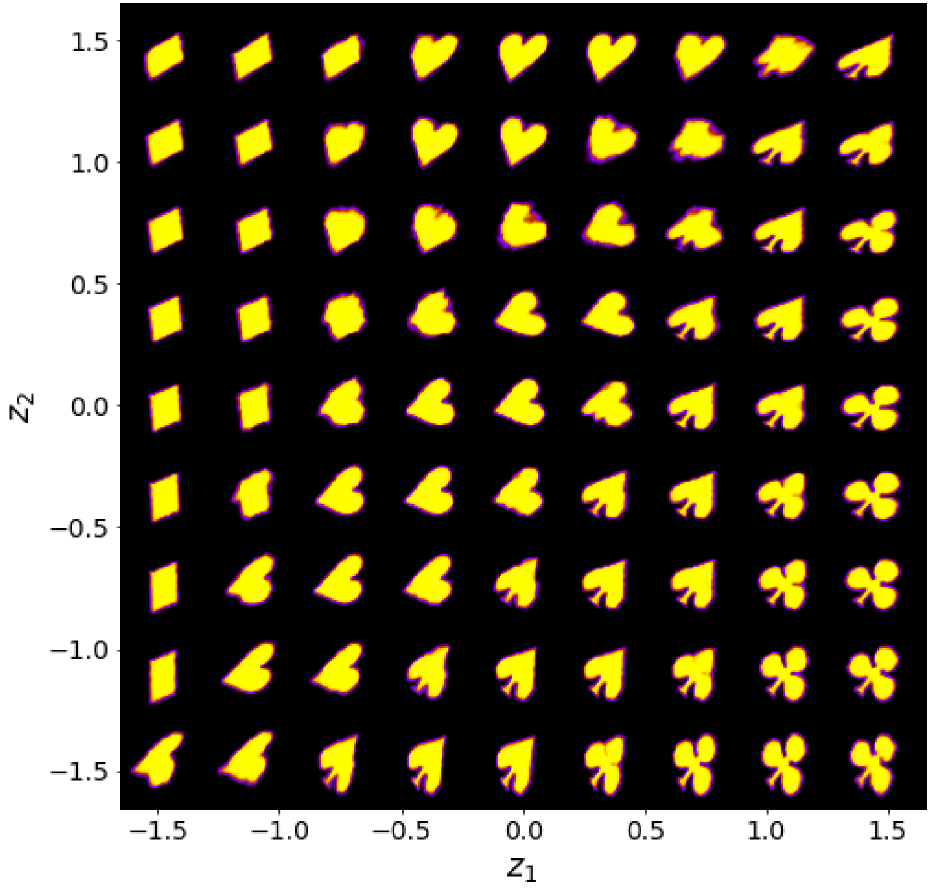
Cards 1: Low rotation (12 deg) and low shear (1 deg)

rVAE on Cards

Example of data

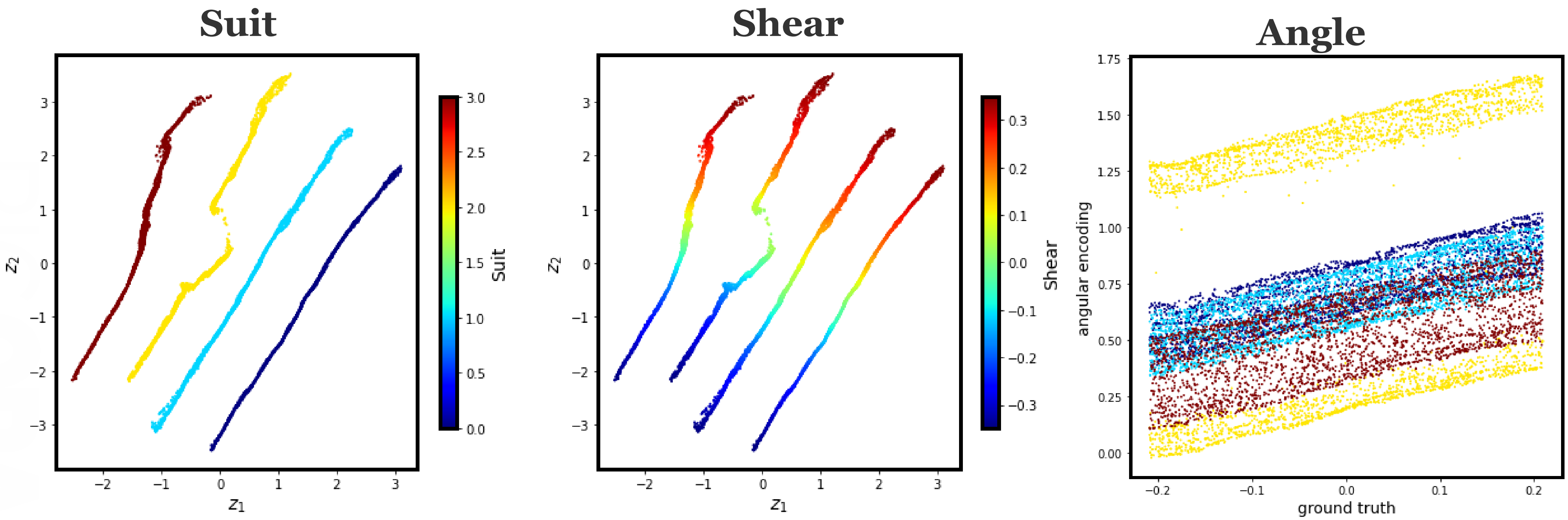


Latent representation



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

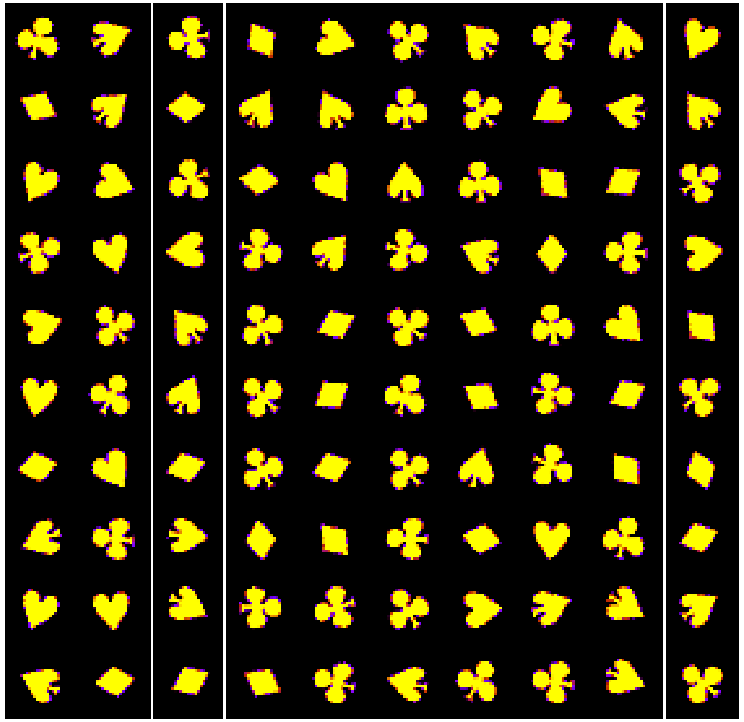
rVAE on Cards



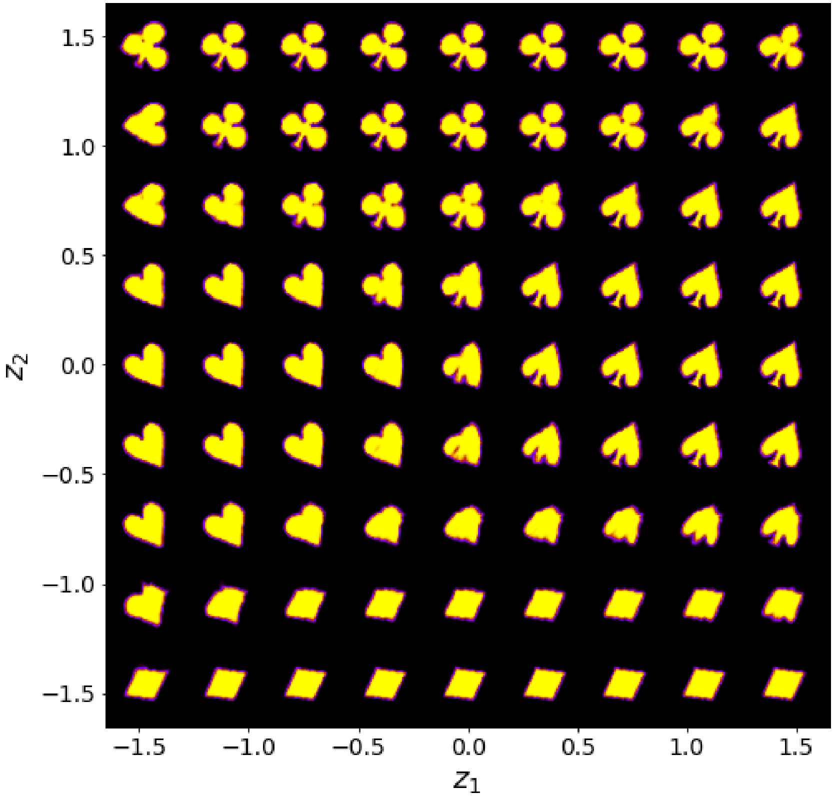
Cards 2: Low rotation (12 deg) and high shear (20 deg)

rVAE on Cards

Example of data



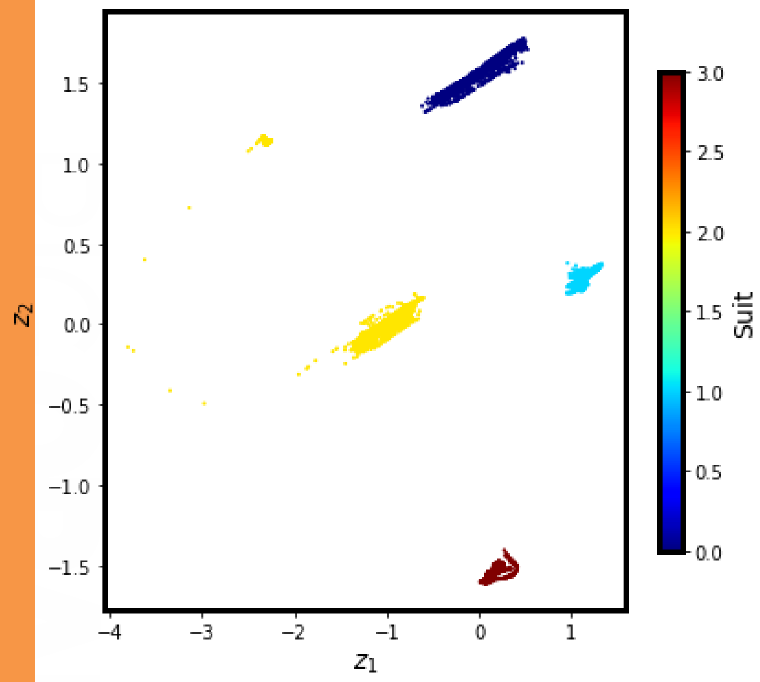
Latent representation



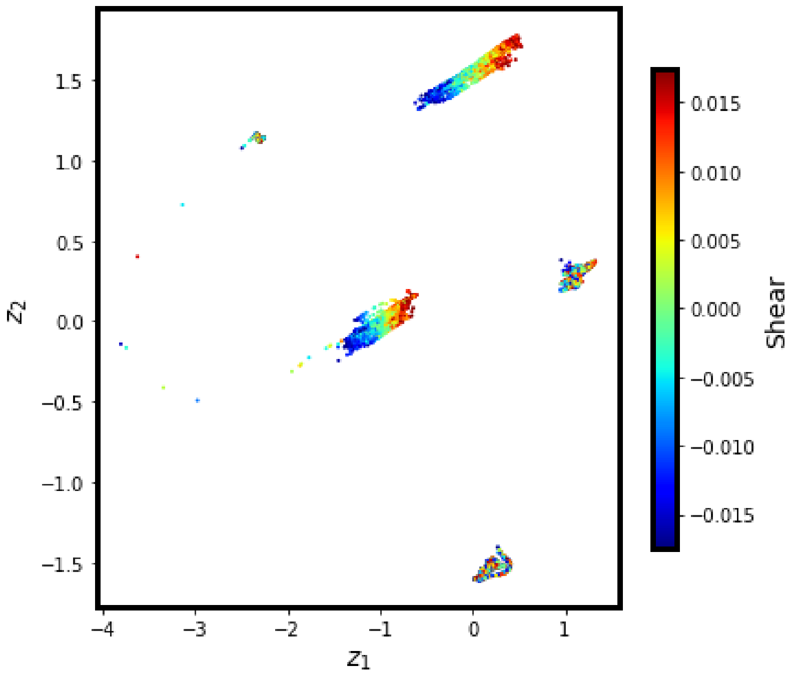
Cards 3: High rotation (120 deg) and low shear (1 deg)

rVAE on Cards

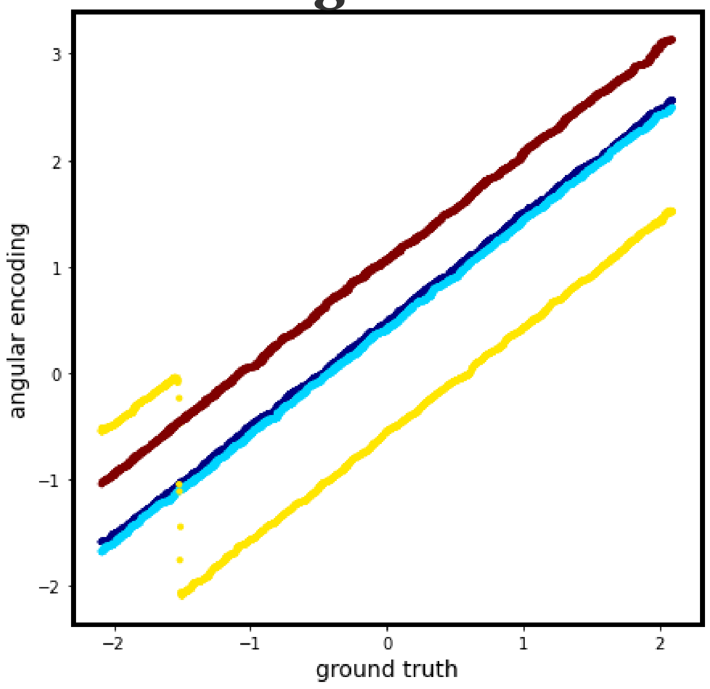
Suit



Shear



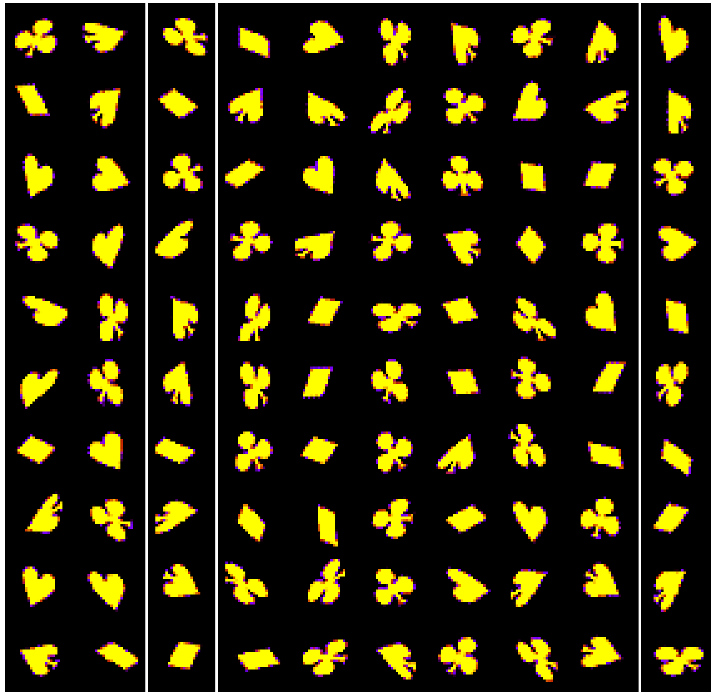
Angle



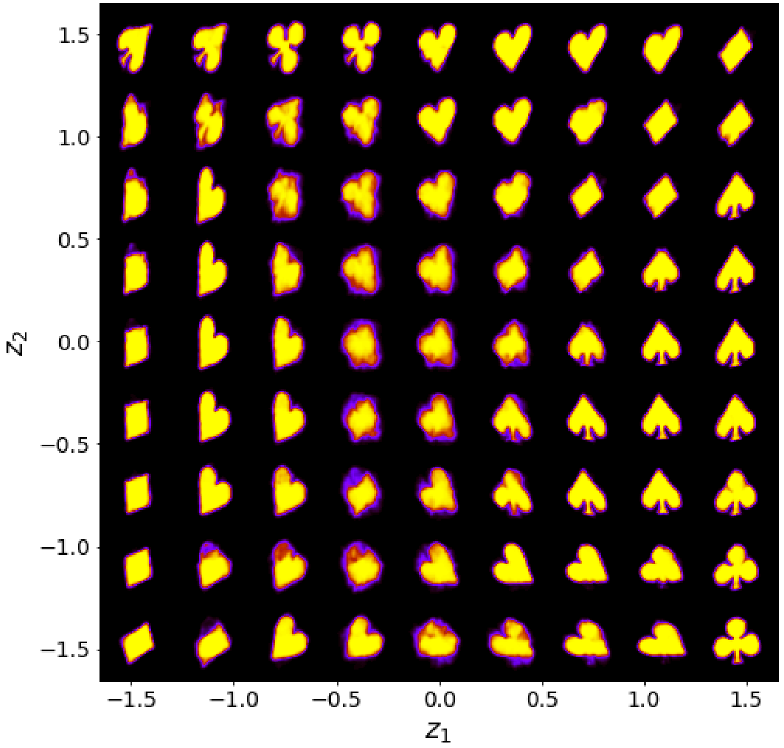
Cards 3: High rotation (120 deg) and low shear (1 deg)

rVAE on Cards

Example of data



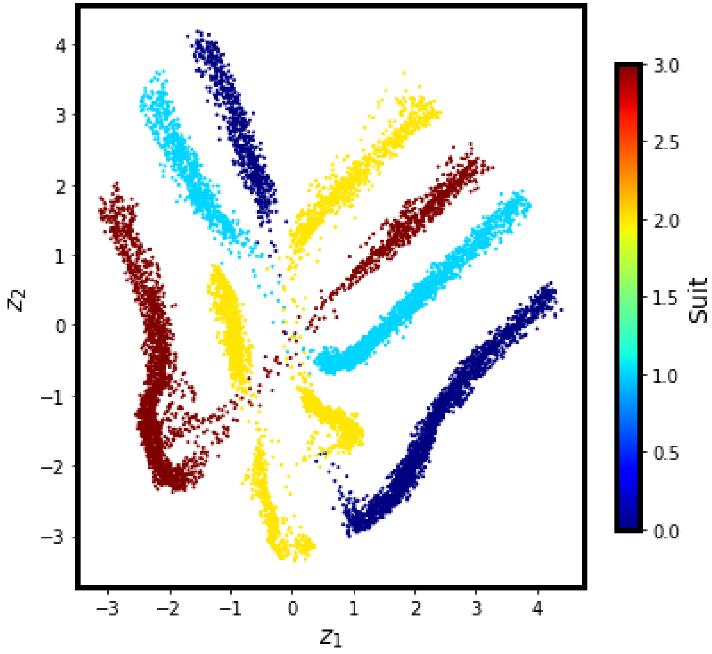
Latent representation



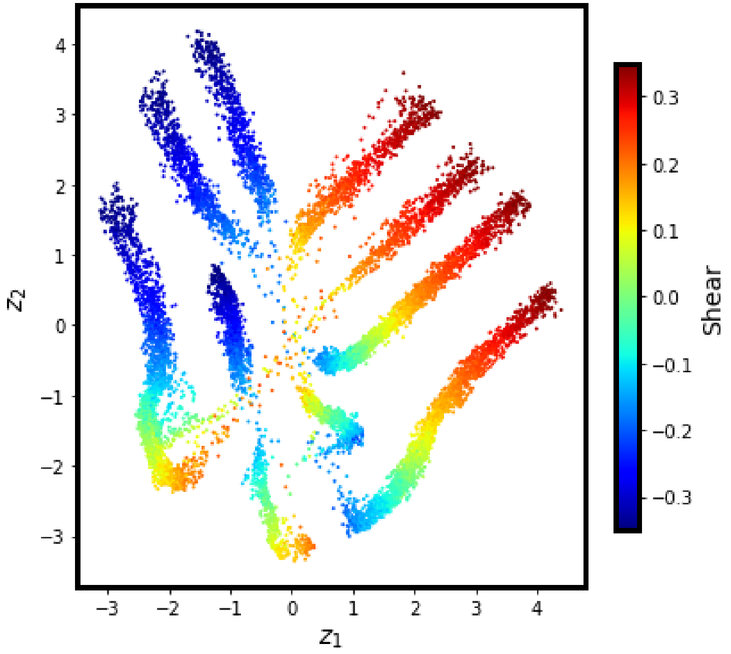
Cards 4: High rotation (120 deg) and high shear (20 deg)

rVAE on Cards

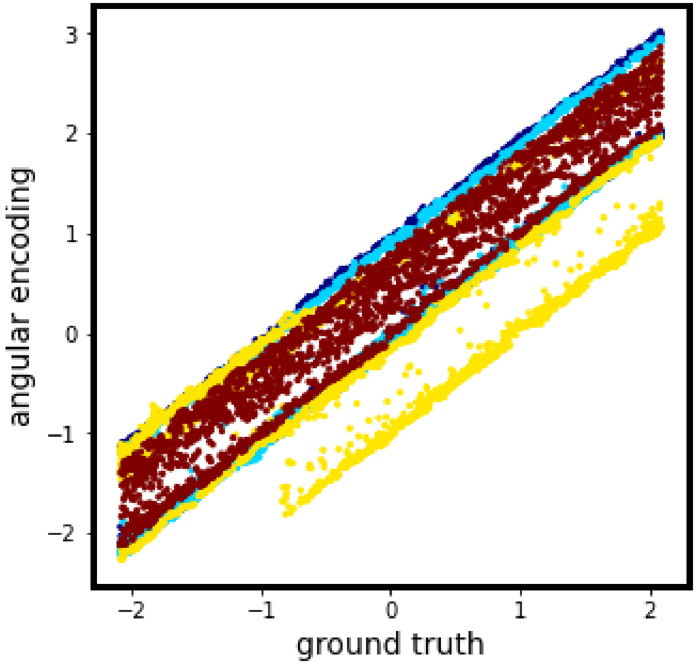
Suit



Shear



Angle

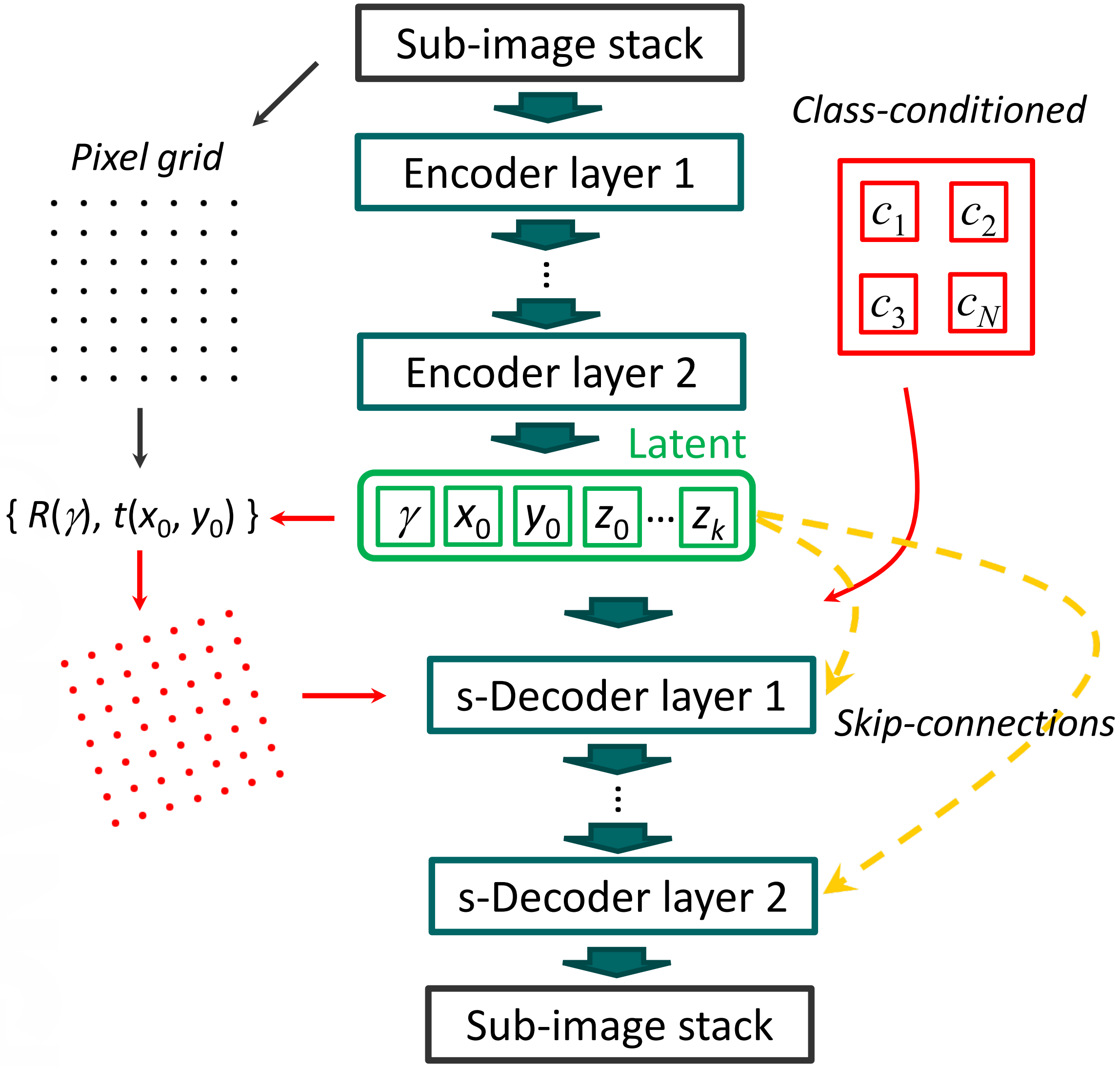


Cards 4: High rotation (120 deg) and high shear (20 deg)

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

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

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

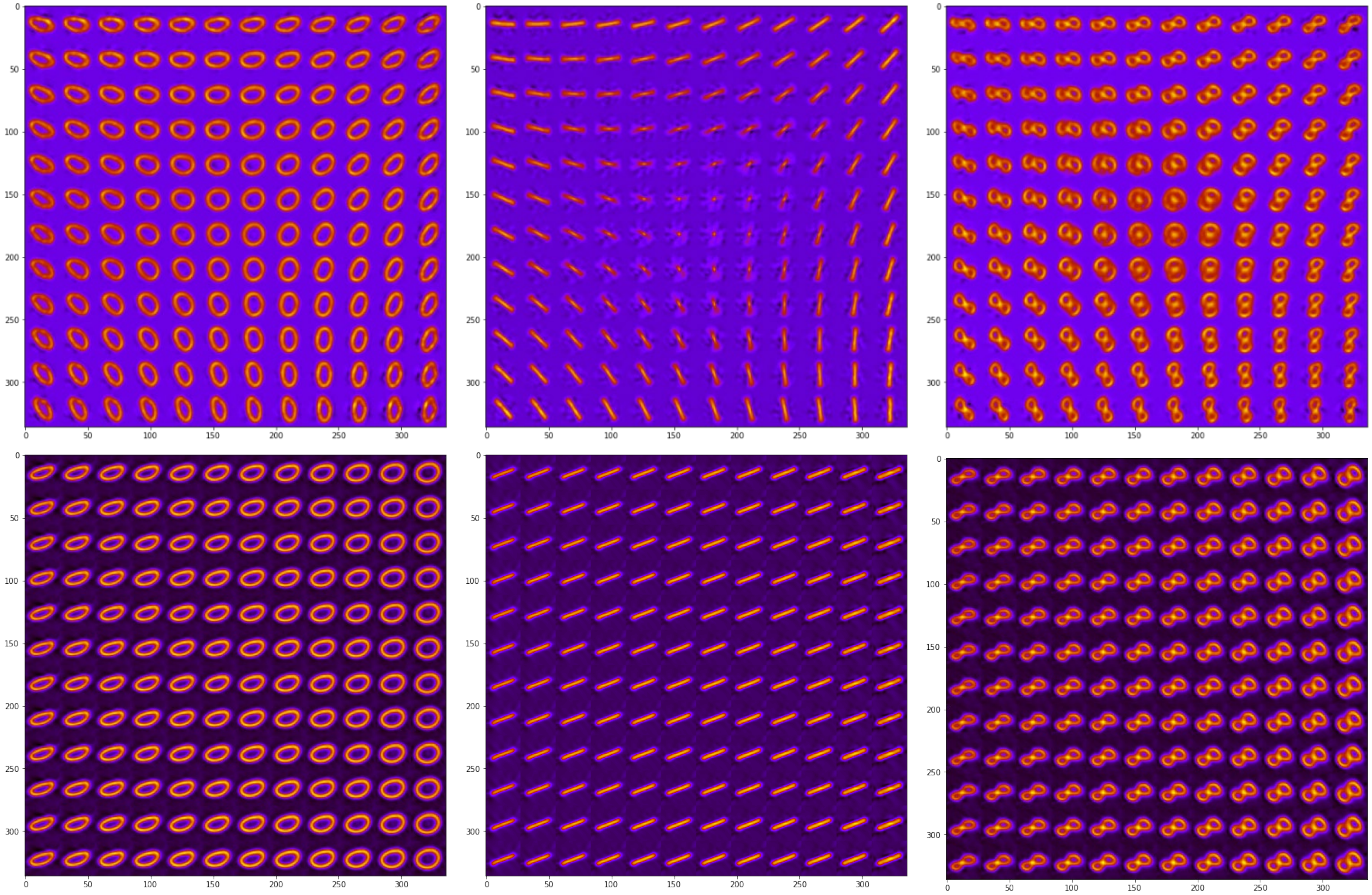
$$- D_{KL}\left(q(\Delta r|x)\|\mathcal{N}(0, s_{\Delta r}^2)\right) \quad \text{Translation}$$

+ D_{KL}(physics-based “priors”) ?

+ D (physics) ?

Bepler *et al.*, NeurIPS 2019

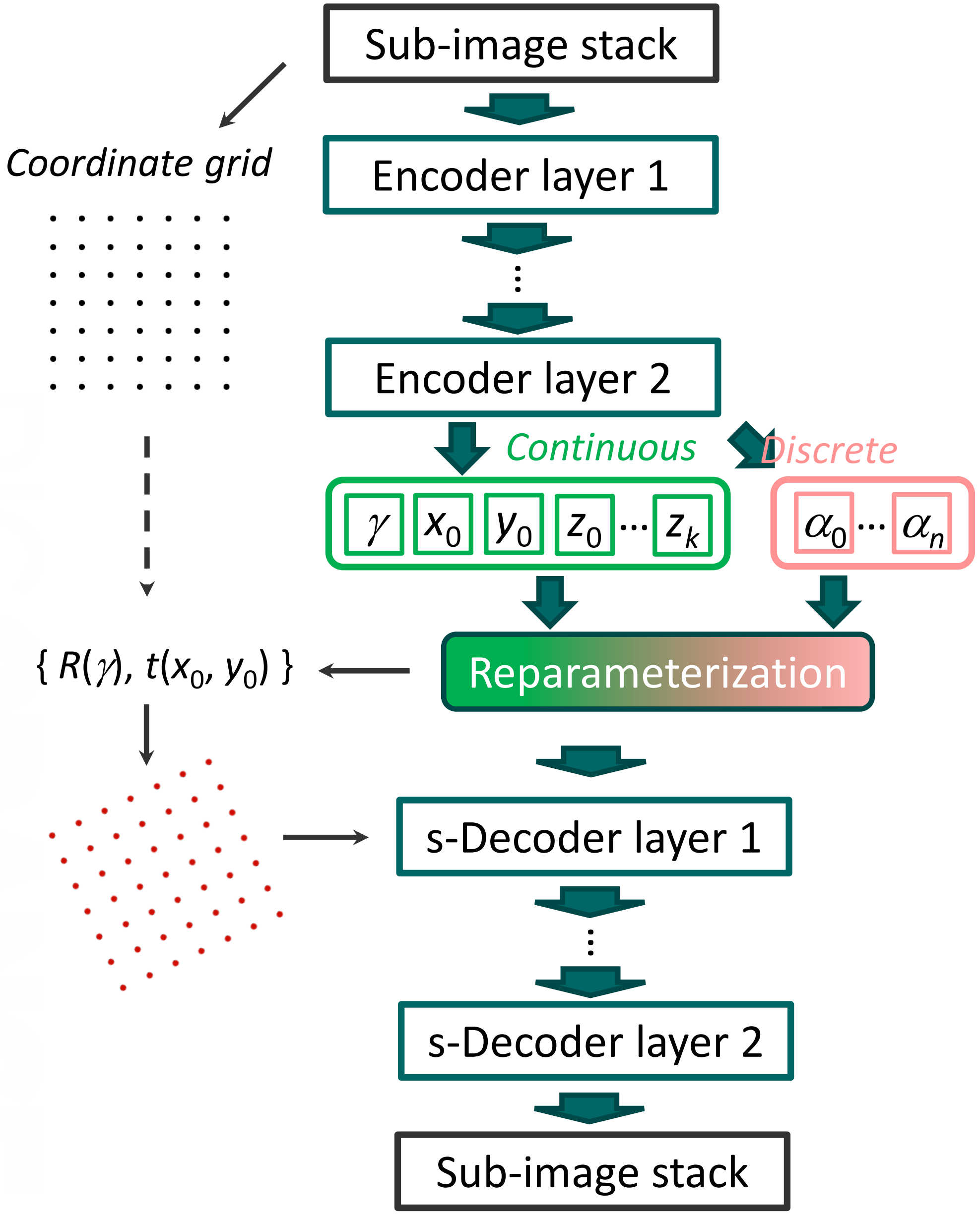
MNIST: cVAE



No rotations

With rotations

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

$$ELBO =$$

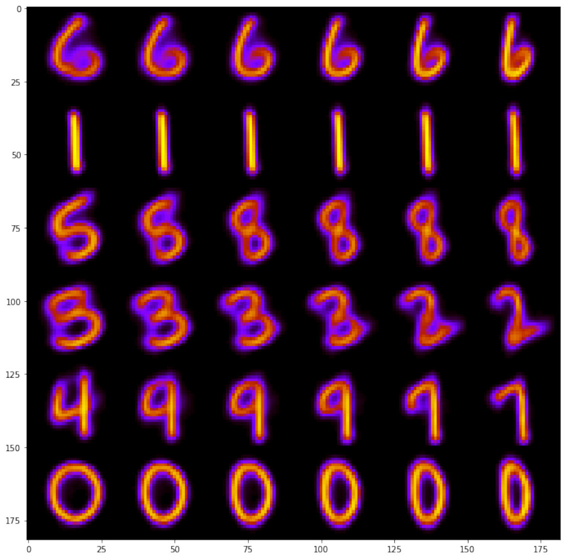
$$-\,Reconstruction\ Loss$$

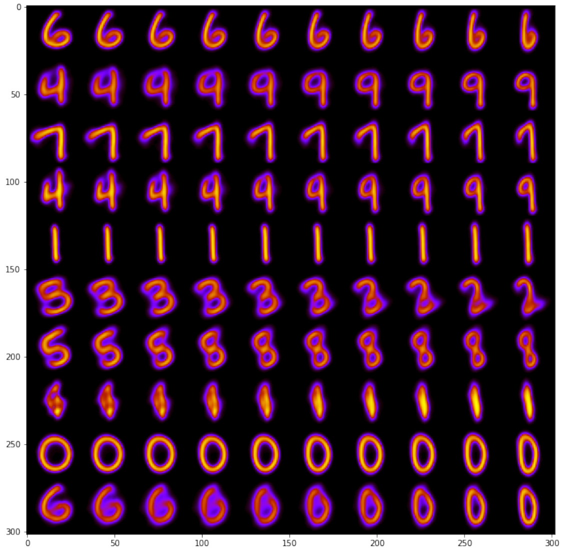
$$-\beta_c(t)\left|\left(\underbrace{D_{KL}(q(z|x)\|p(z))}_{\text{Structure}} + \underbrace{D_{KL}(q(\gamma|x)\|p(\gamma))}_{\text{Rotation}} - C_z\right|\right. \quad \text{Continuous}$$

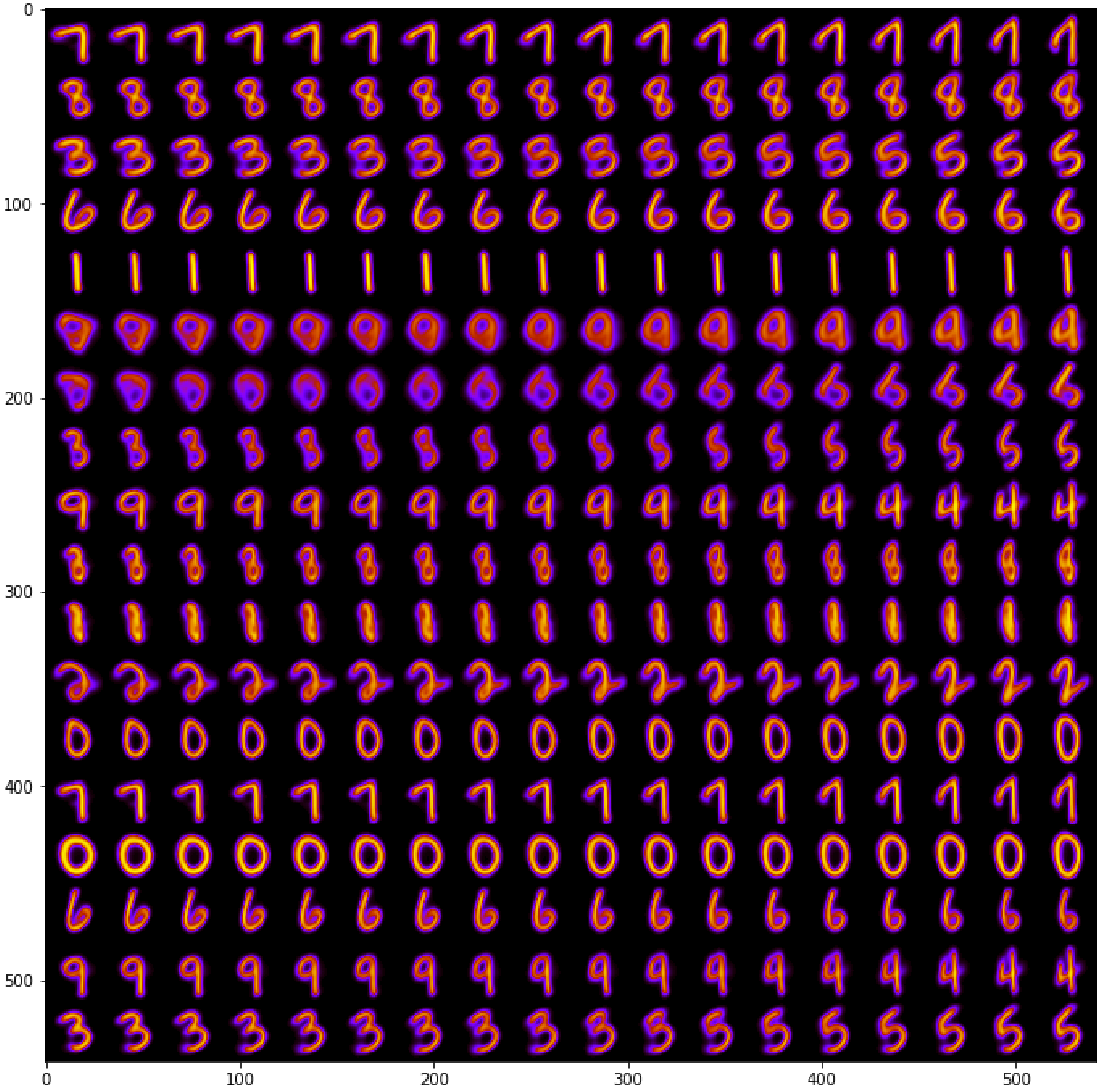
$$-\beta_d(t)\left|D_{KL}(q(\alpha|x)\|p(\alpha)) - C_\alpha\right| \quad \text{Discrete}$$

+ physics-based "loss" ?

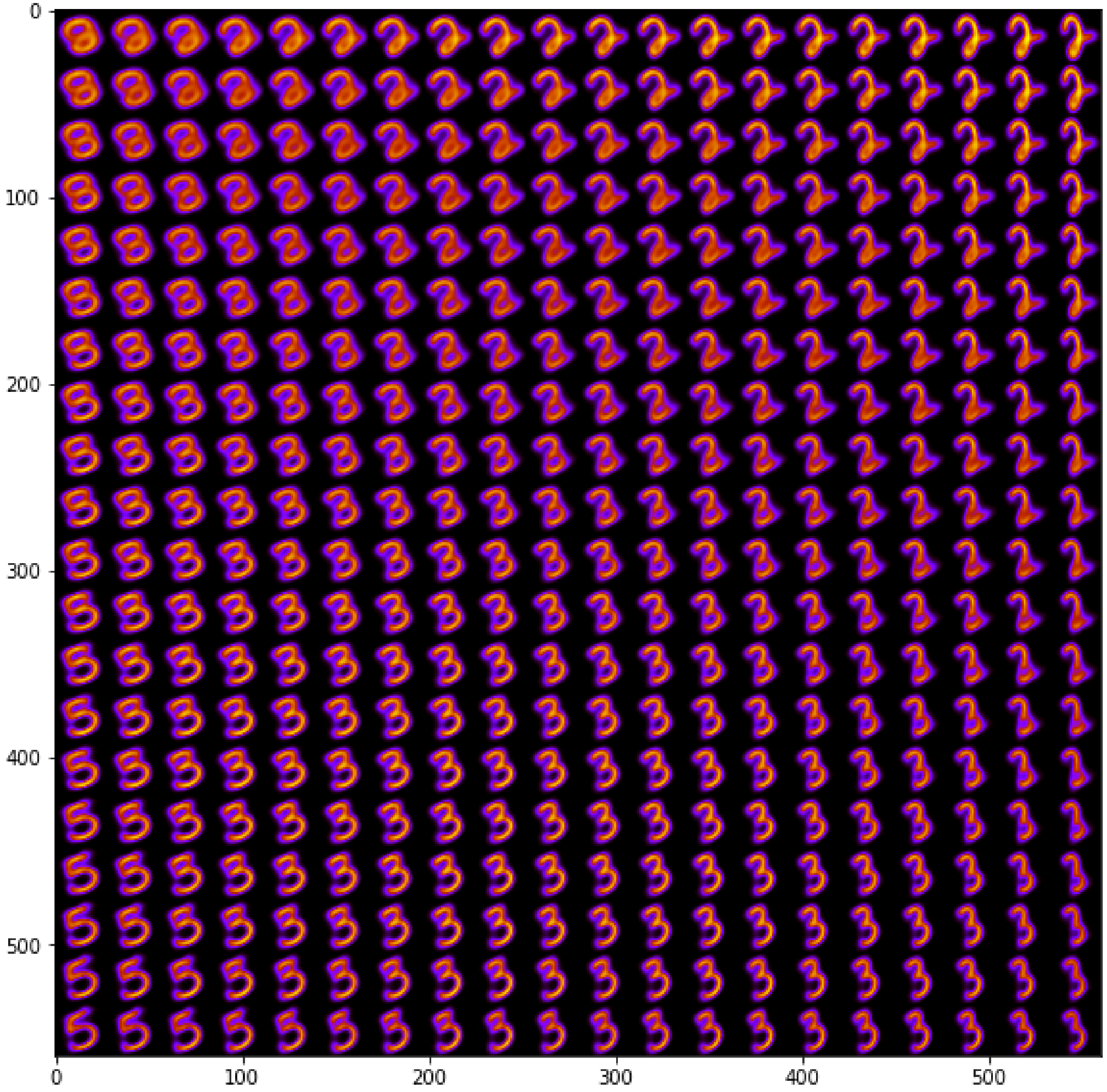
jVAE of MNIST





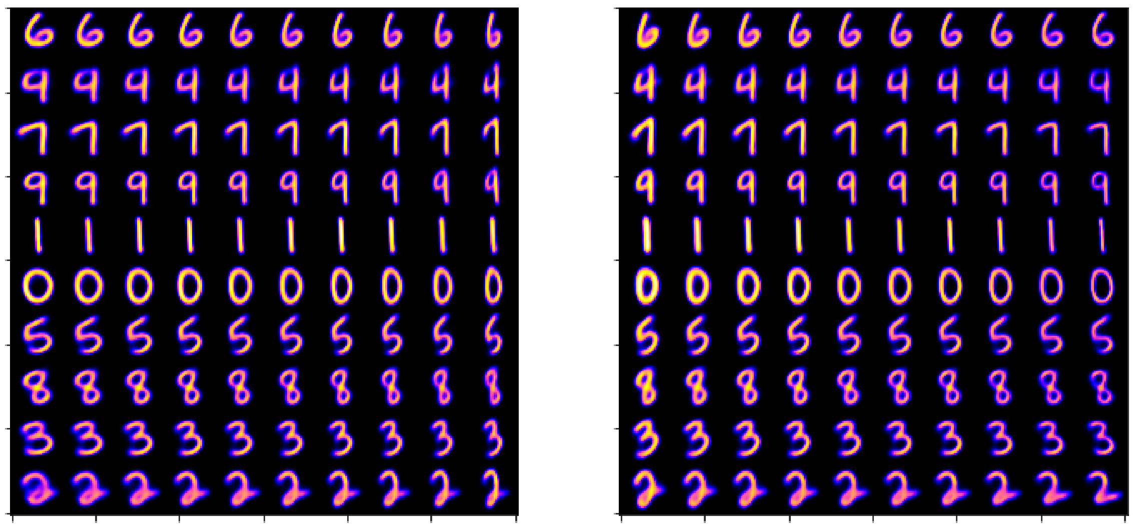


Latent representation

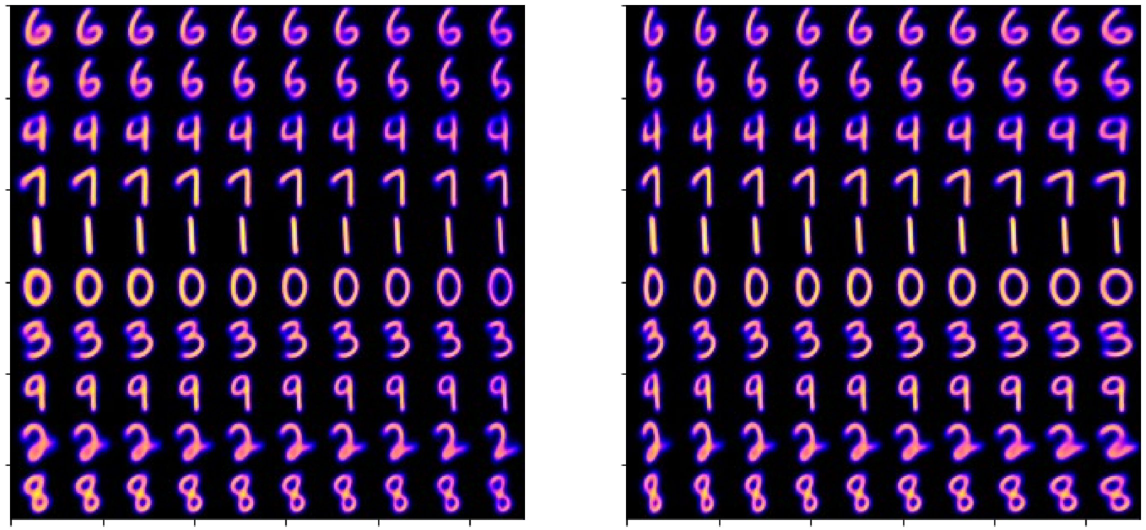


Ensemble jVAE

Predictions from different ensemble models



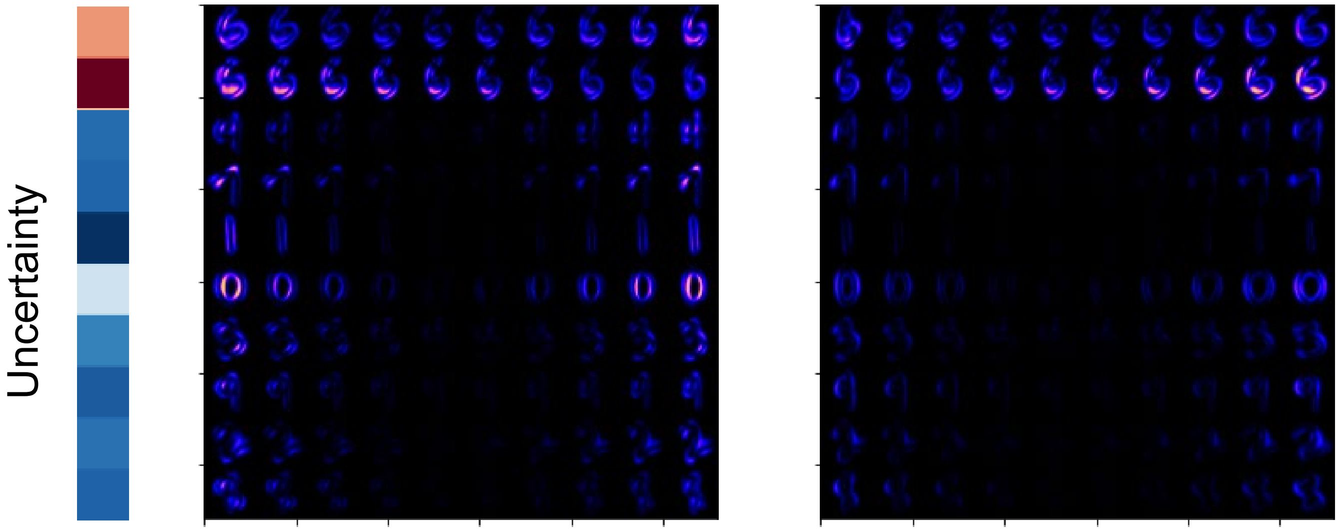
Mean prediction



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')



What can (unsupervised) classification give us

- Two things matter:
  - descriptors
  - 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?

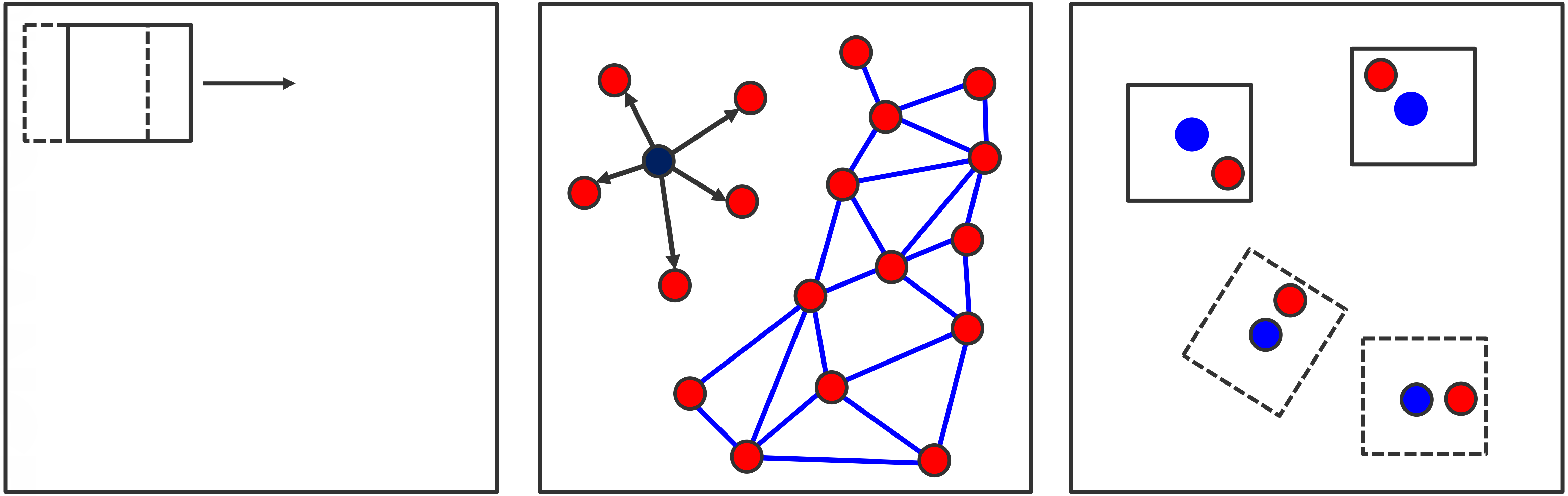
Describing the building blocks

- The classical physical descriptions (symmetry) 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

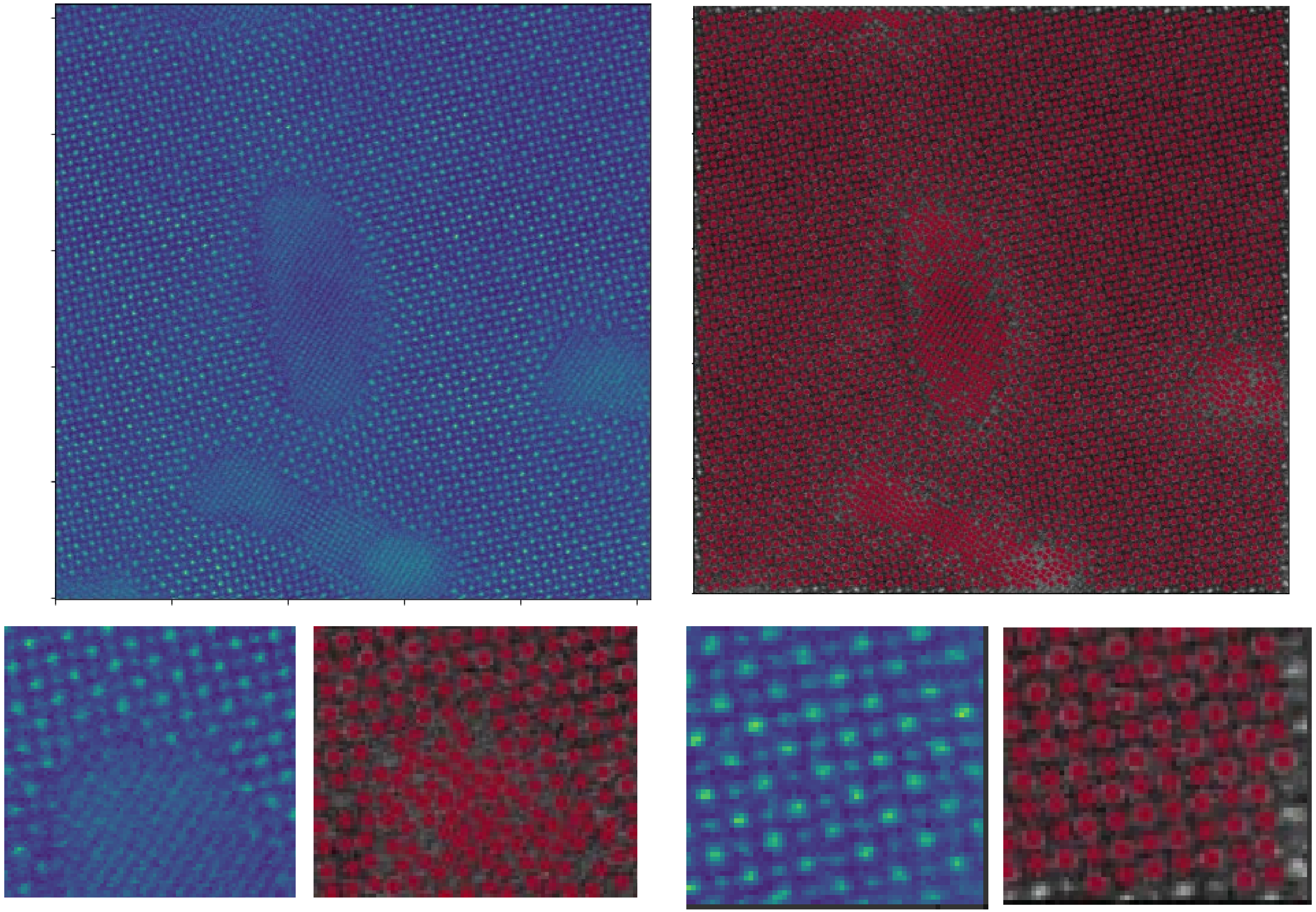
Continuous translational symmetry

Atom based descriptions

Localized sub-images

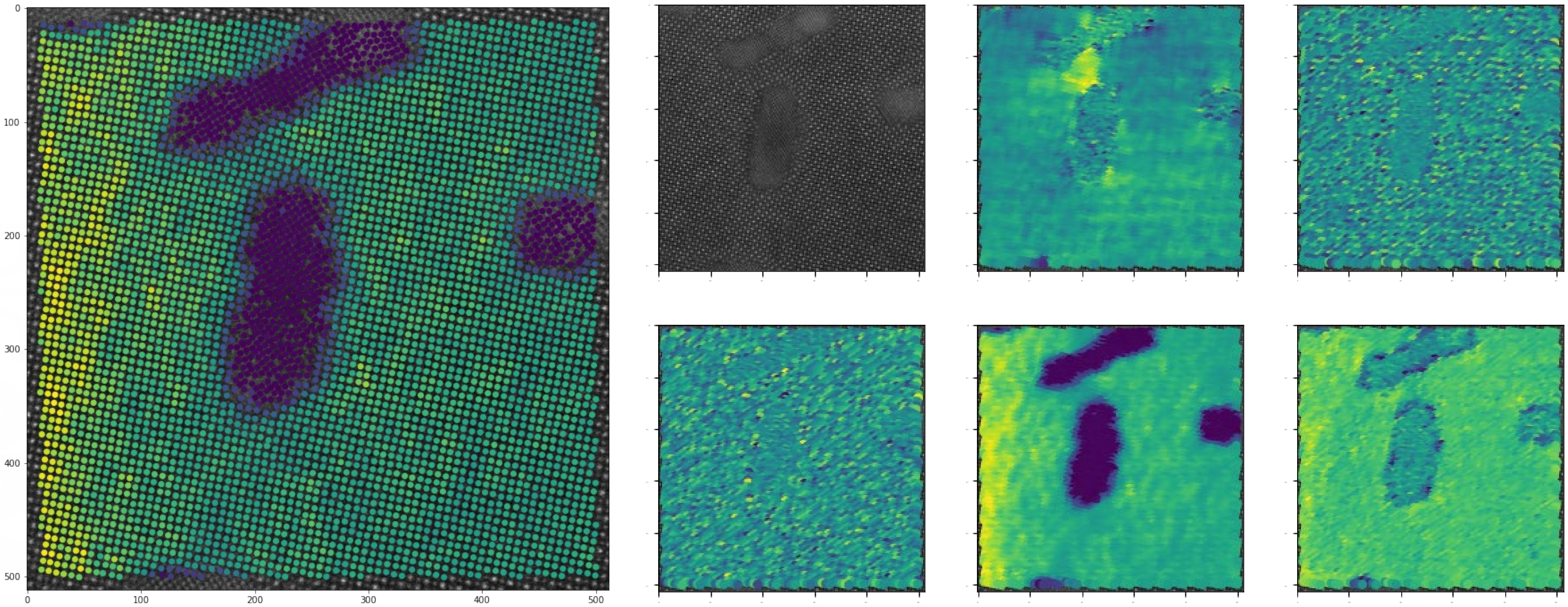


Let’s put it all together!



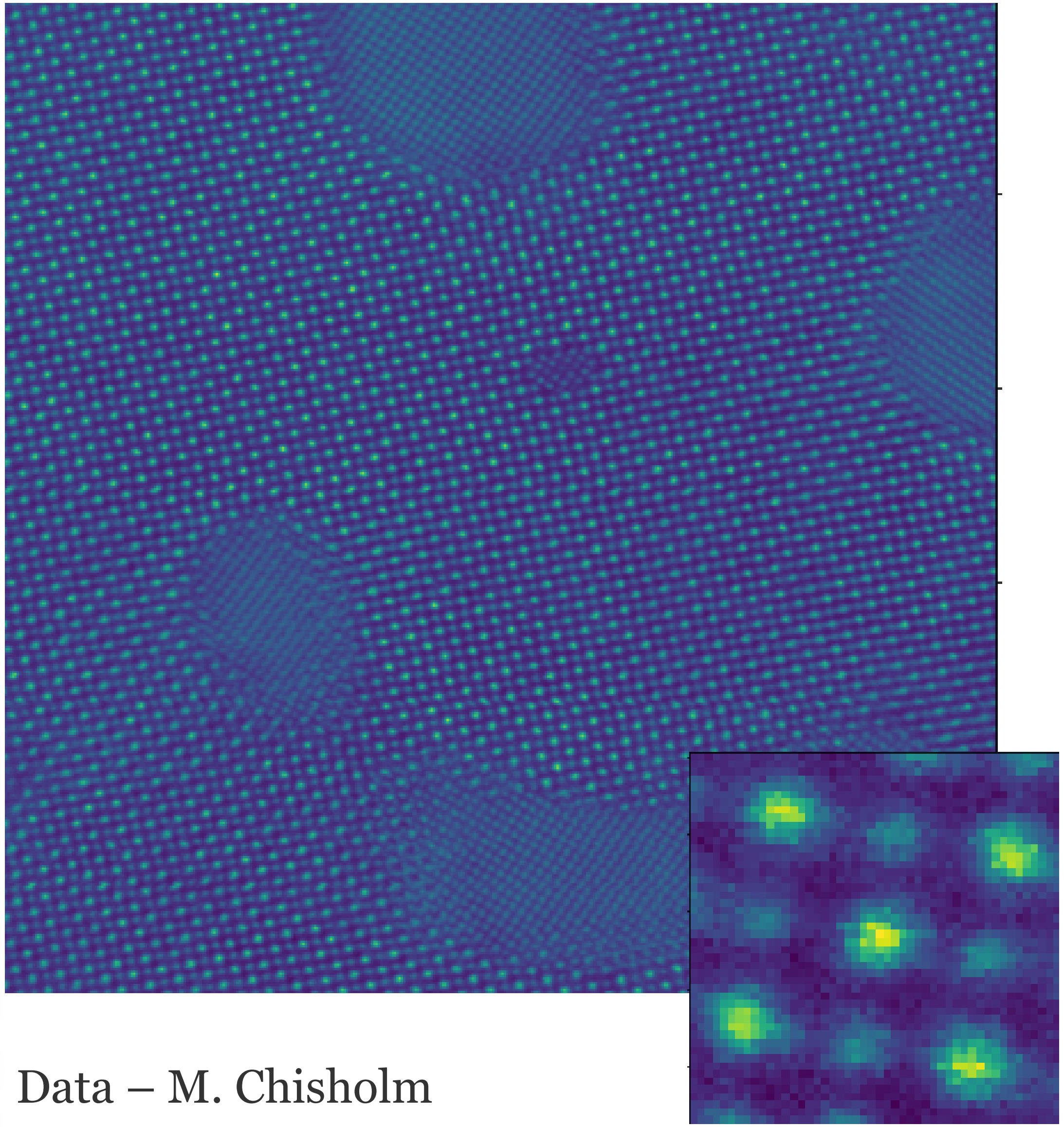
Step 1: Find all atoms (or all that you can) – use maximum finders, blob-log, or DCNNs
Step 2: Create descriptors – patches centered on atoms. Keep track on what part of image (or stack) it came from

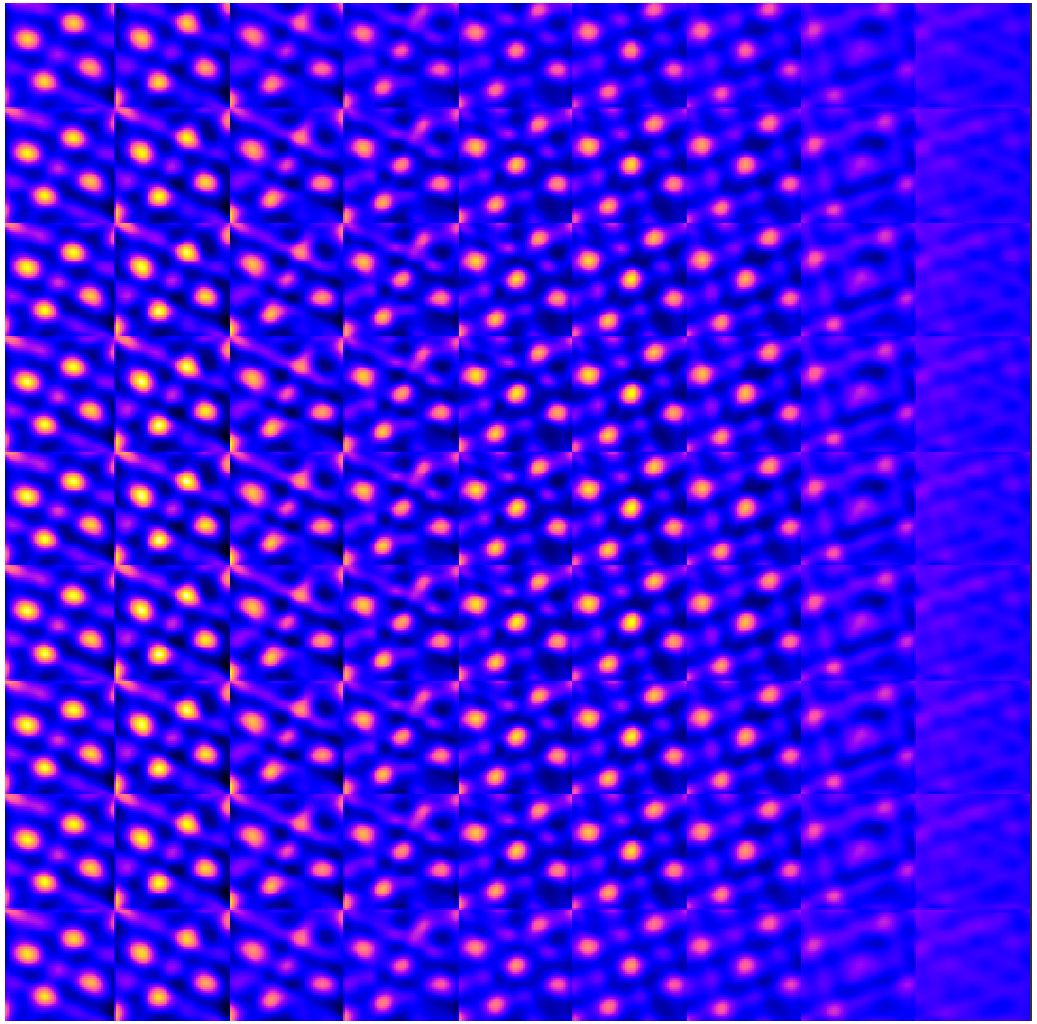
Step 3: rVAE



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!)

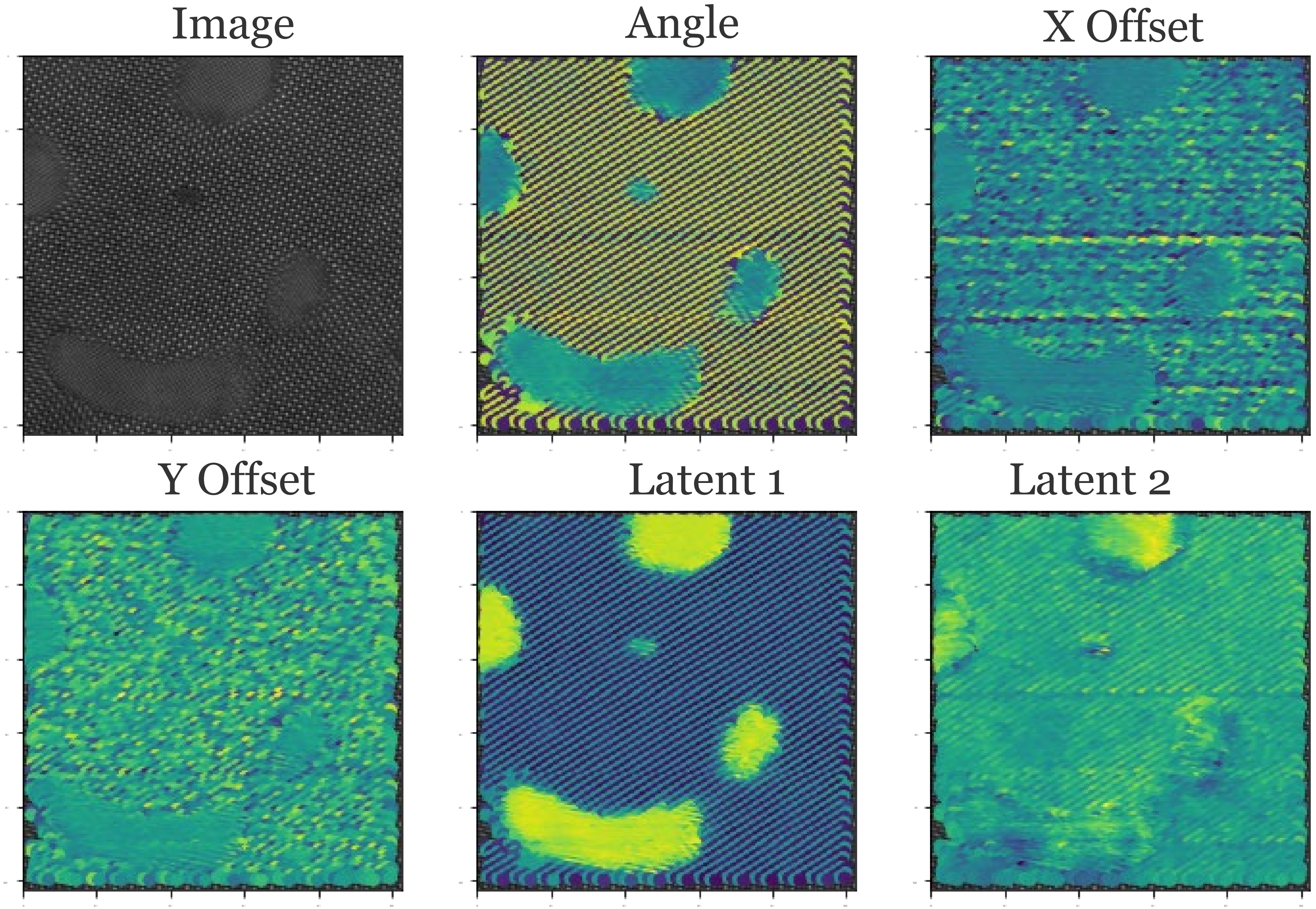
Analysis of the NiO-LSMO



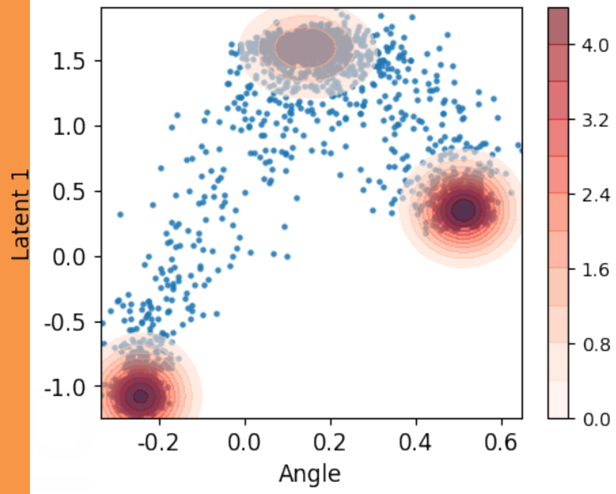


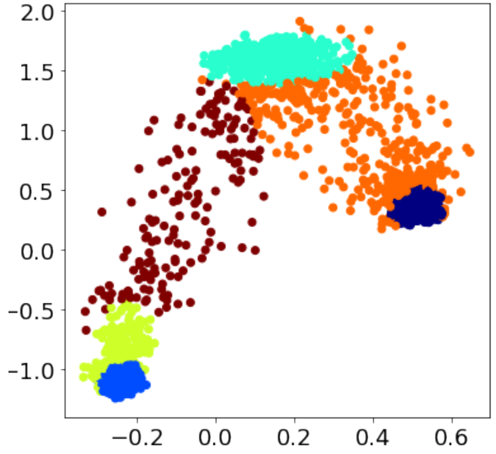
Data – M. Chisholm

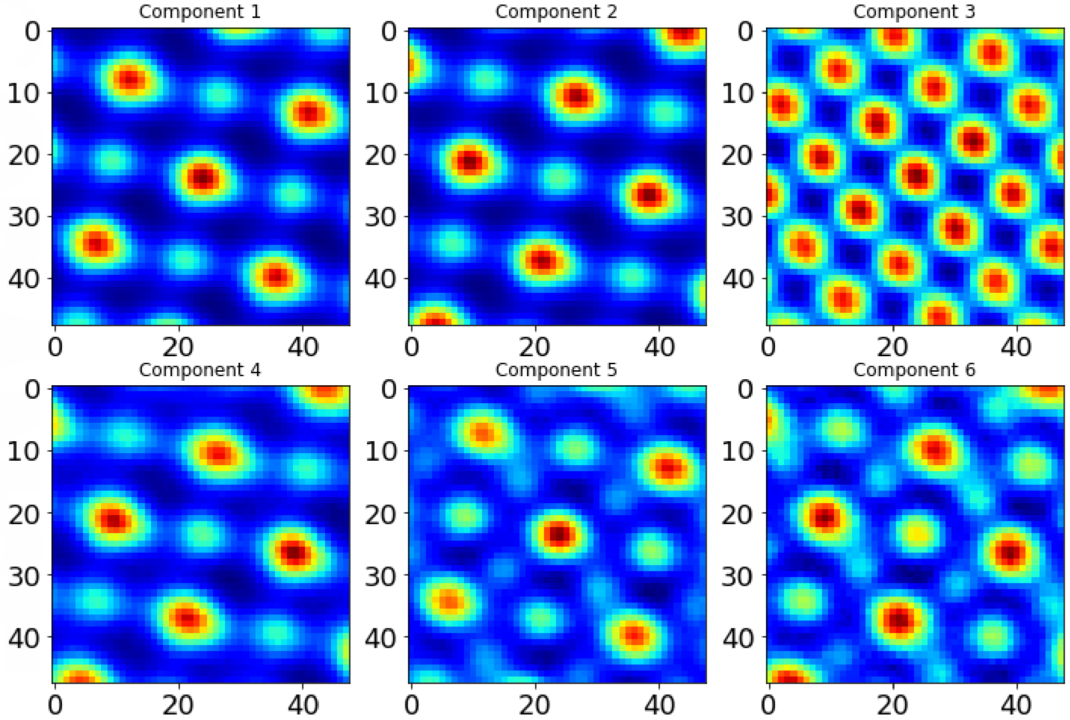
Let's look at latent space



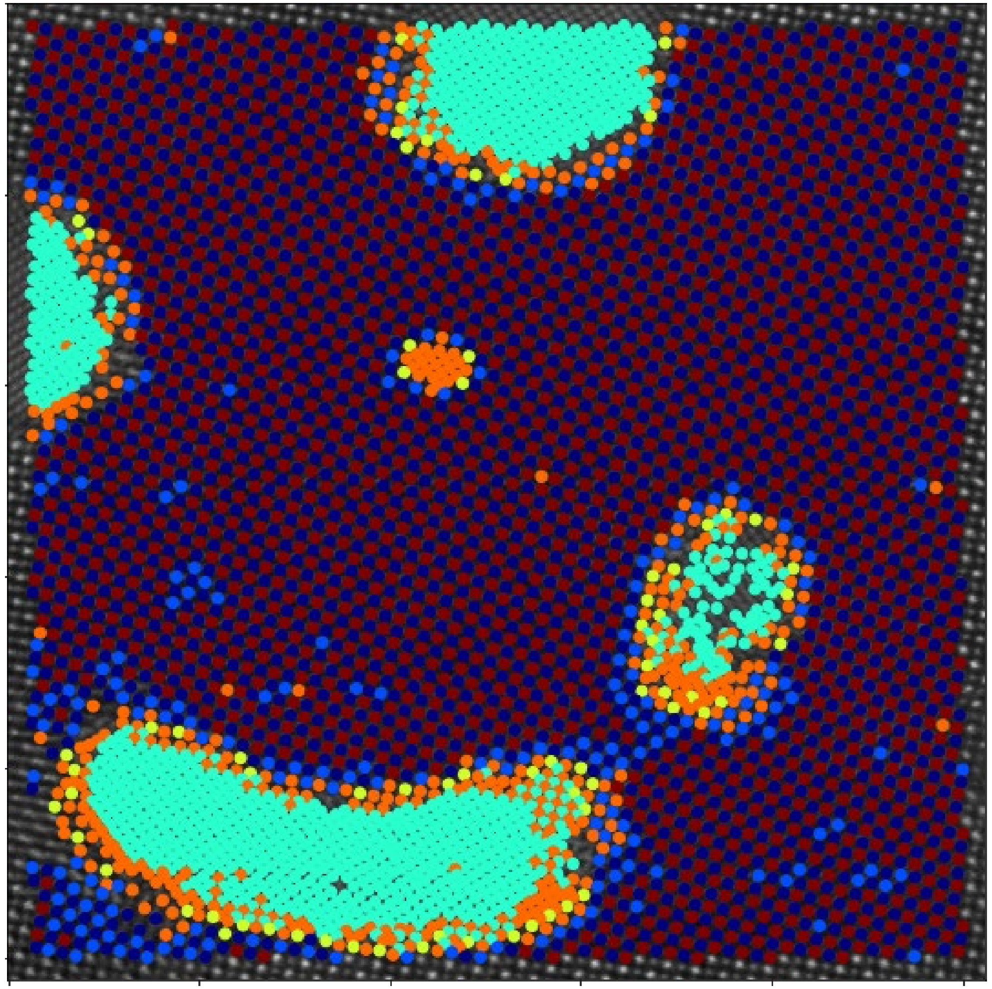
Exploring latent distributions







Labeled image

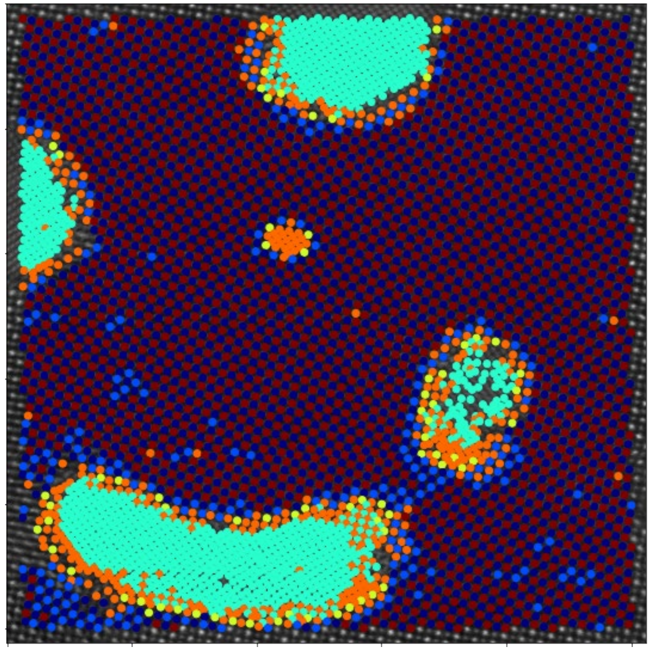


- Classes and variability are mixed in a single latent space
- Disentangling of representation

That's where the jVAE has come from

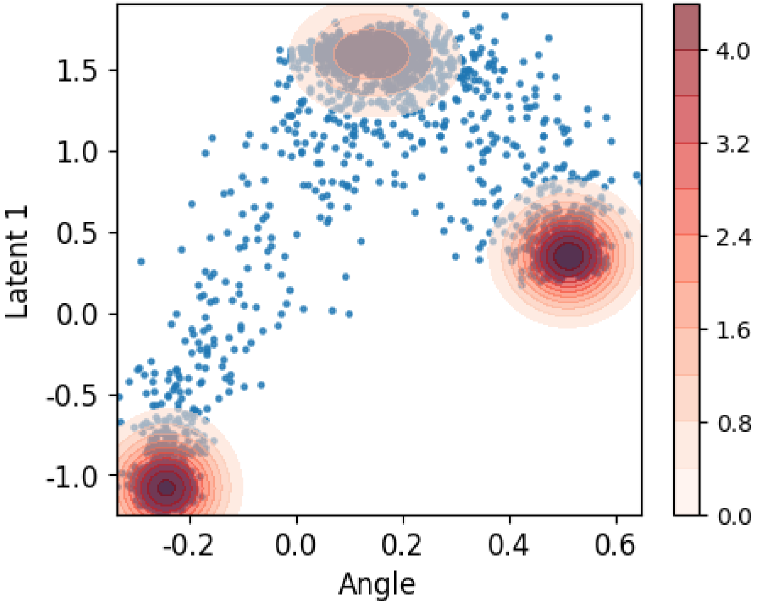
Currently, we have variants of invariant VAE that include:

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

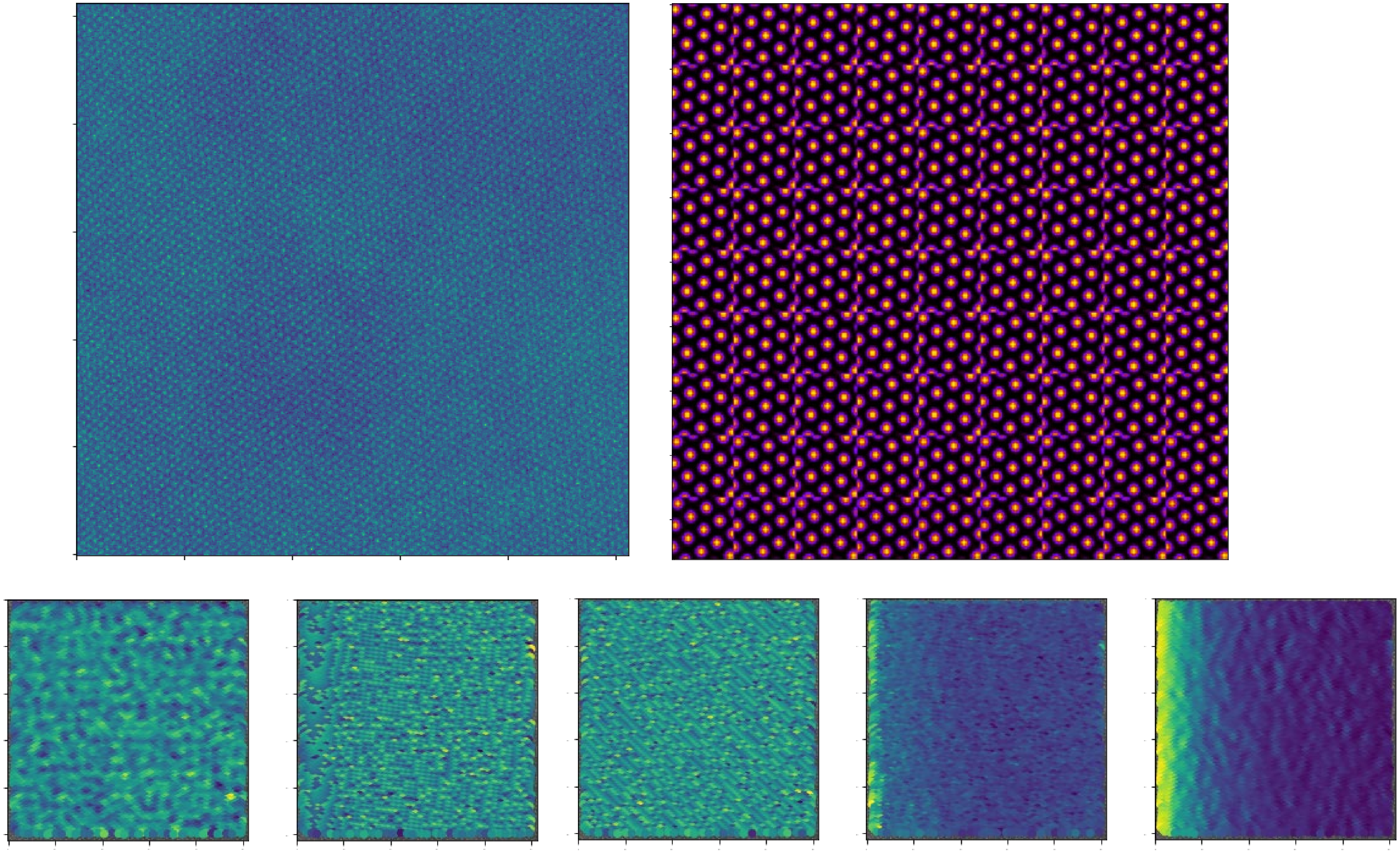


However, our rVAE collects everything in a single latent space. Realistically, very often we deal with system where we expect the presence of finite number of classes that may be known, partially known, or unknown, with certain continuous traits within classes.

- **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

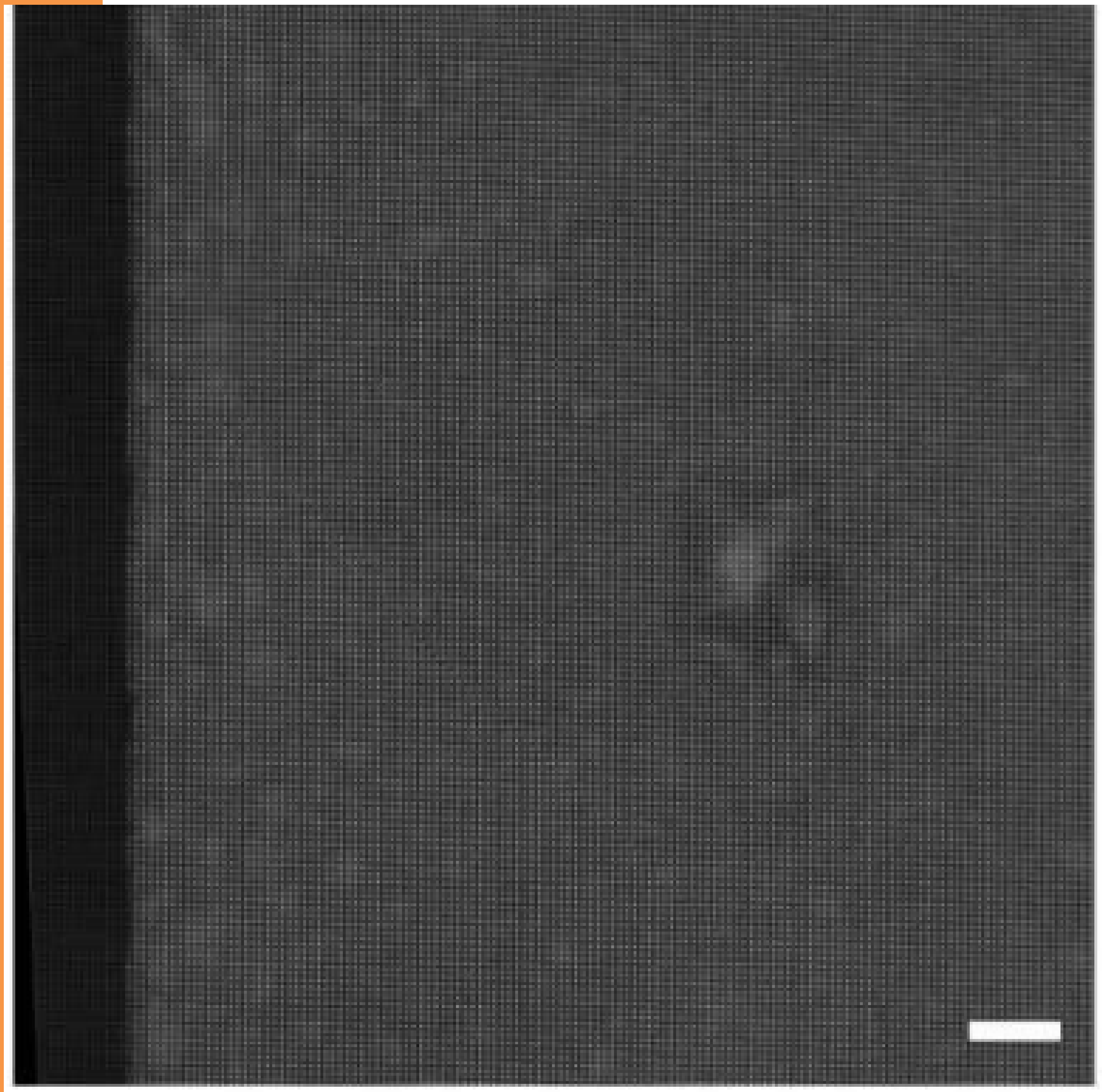


Out of curiosity: single crystal?

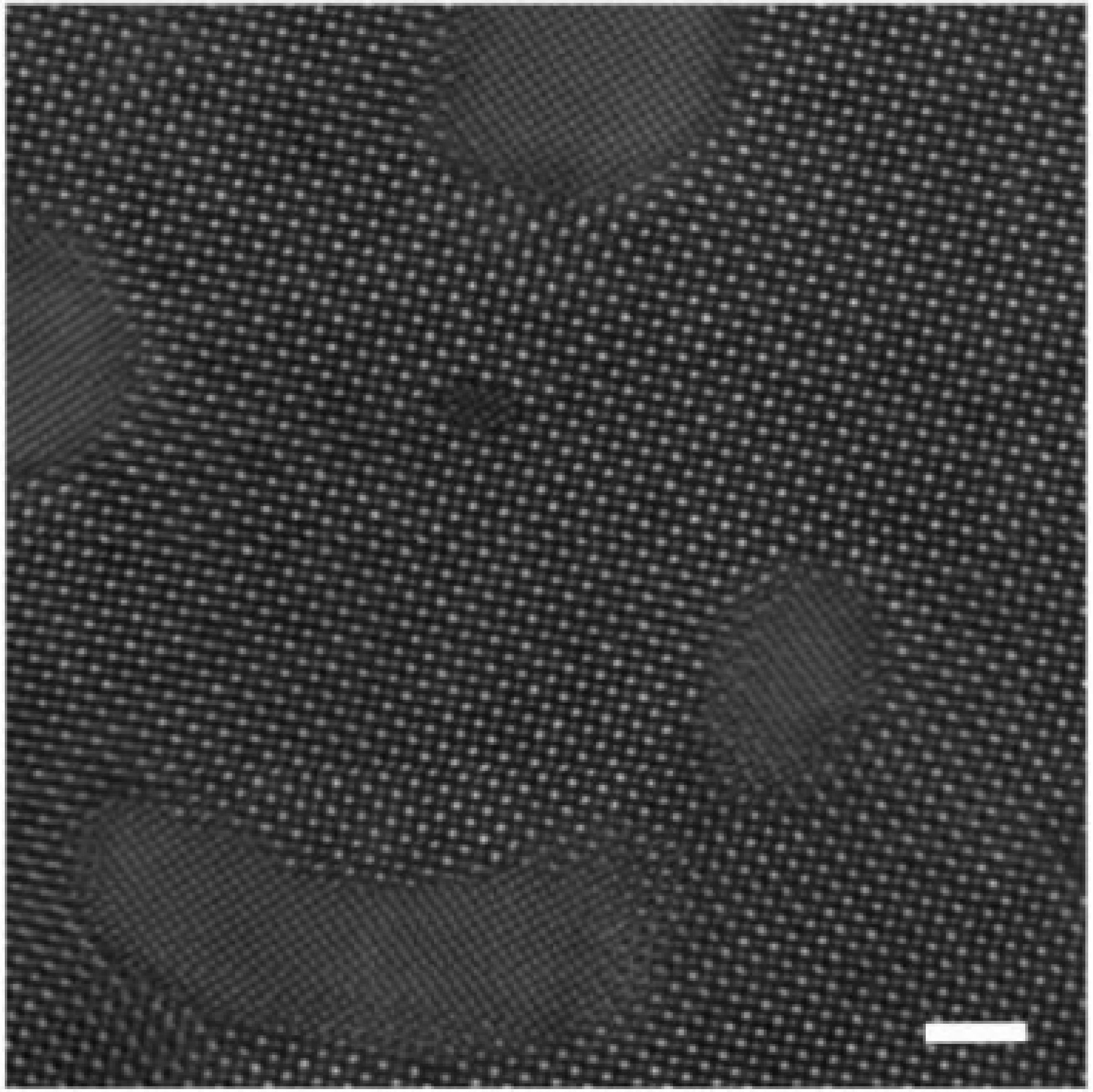


VAE without Atom Finding

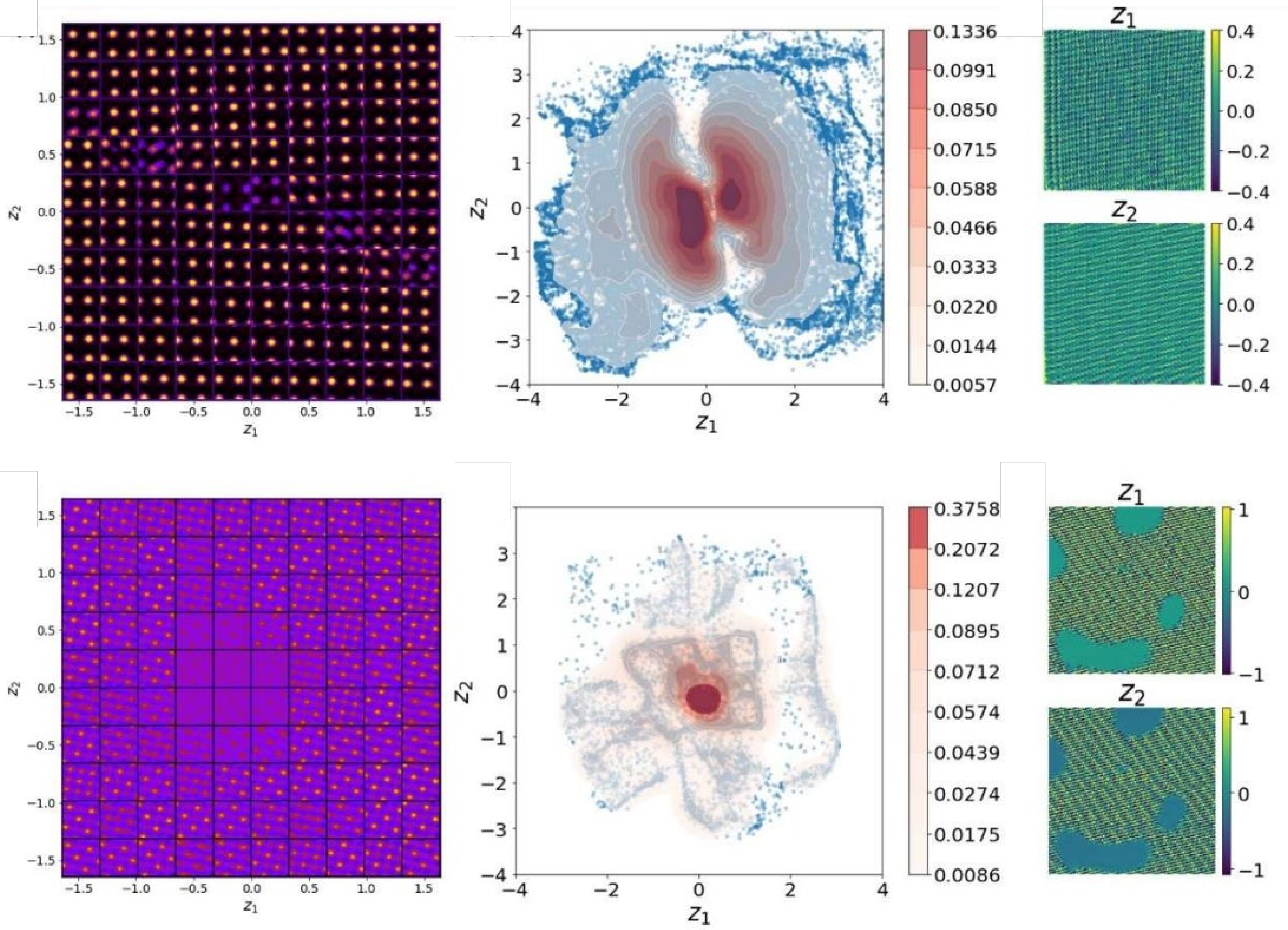
Ferroelectric $BiFeO_3$



NiO – $La_xSr_{1-x}MnO_3$

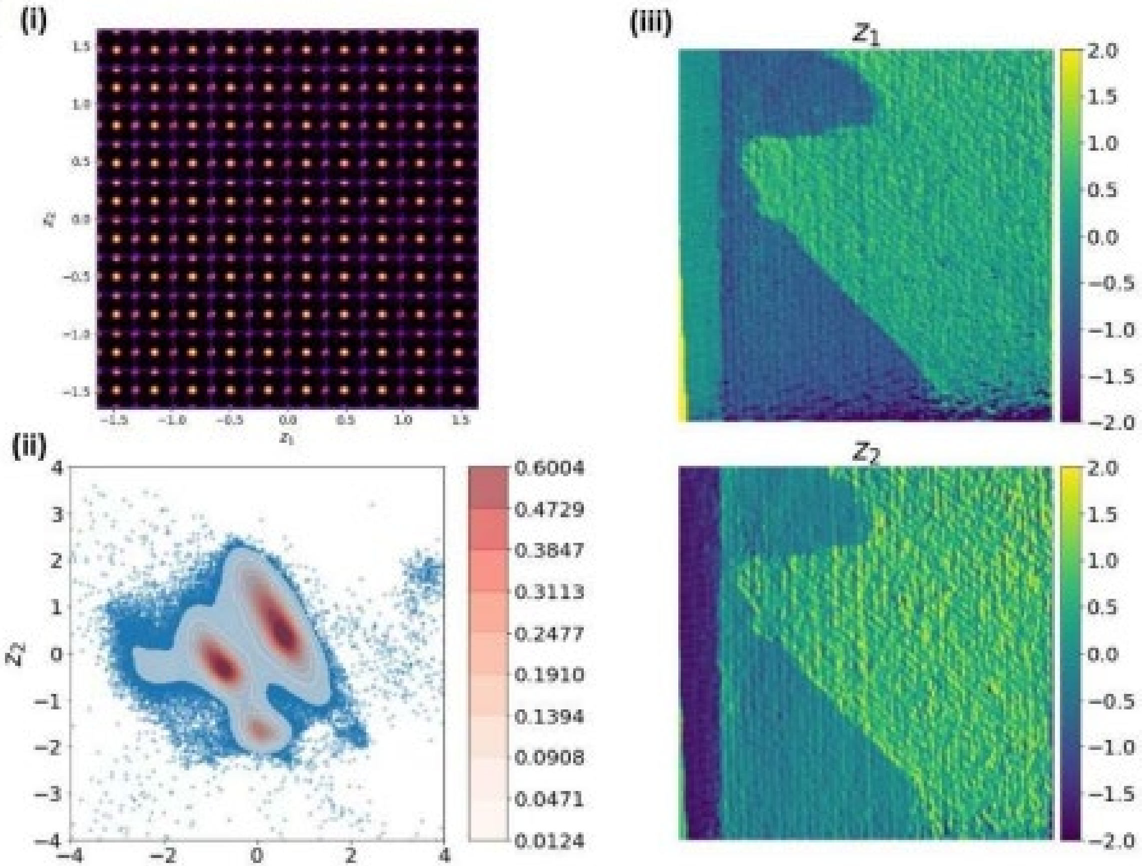


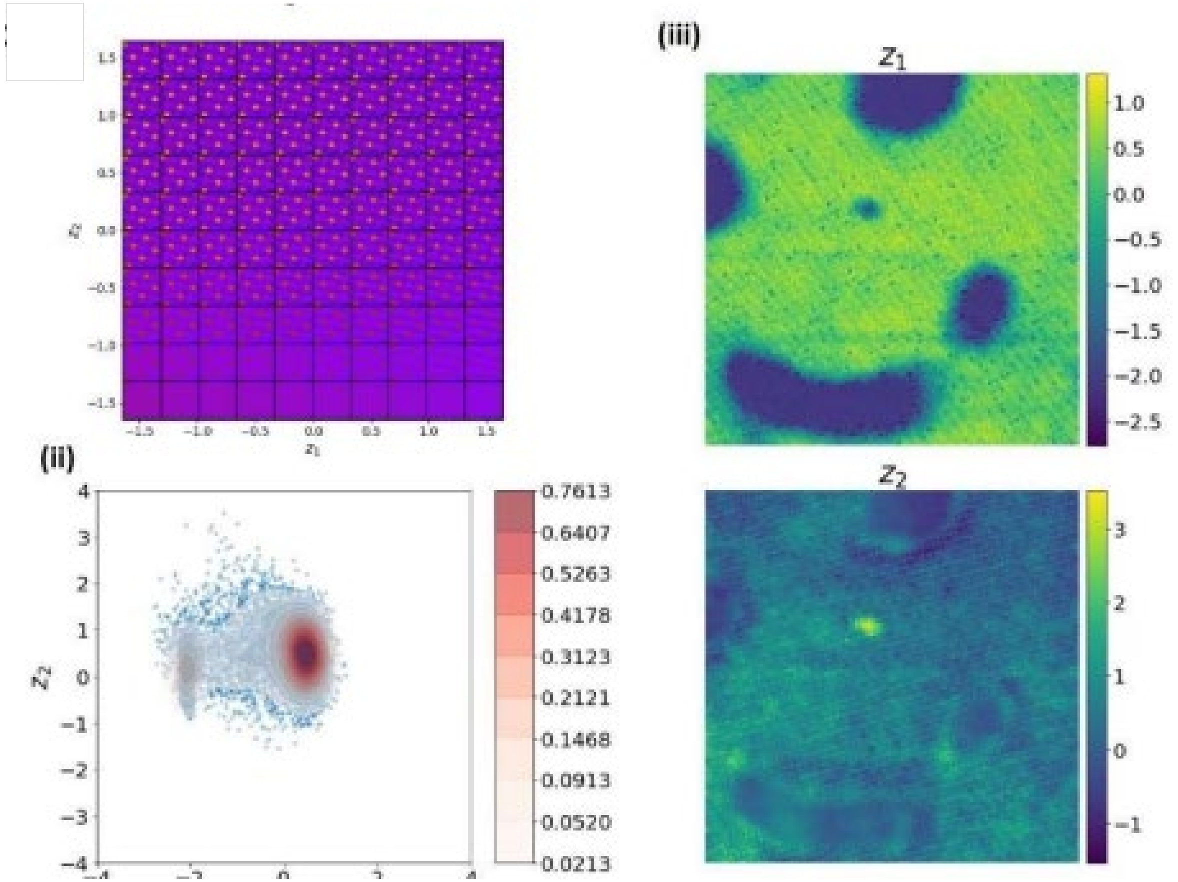
Simple VAE



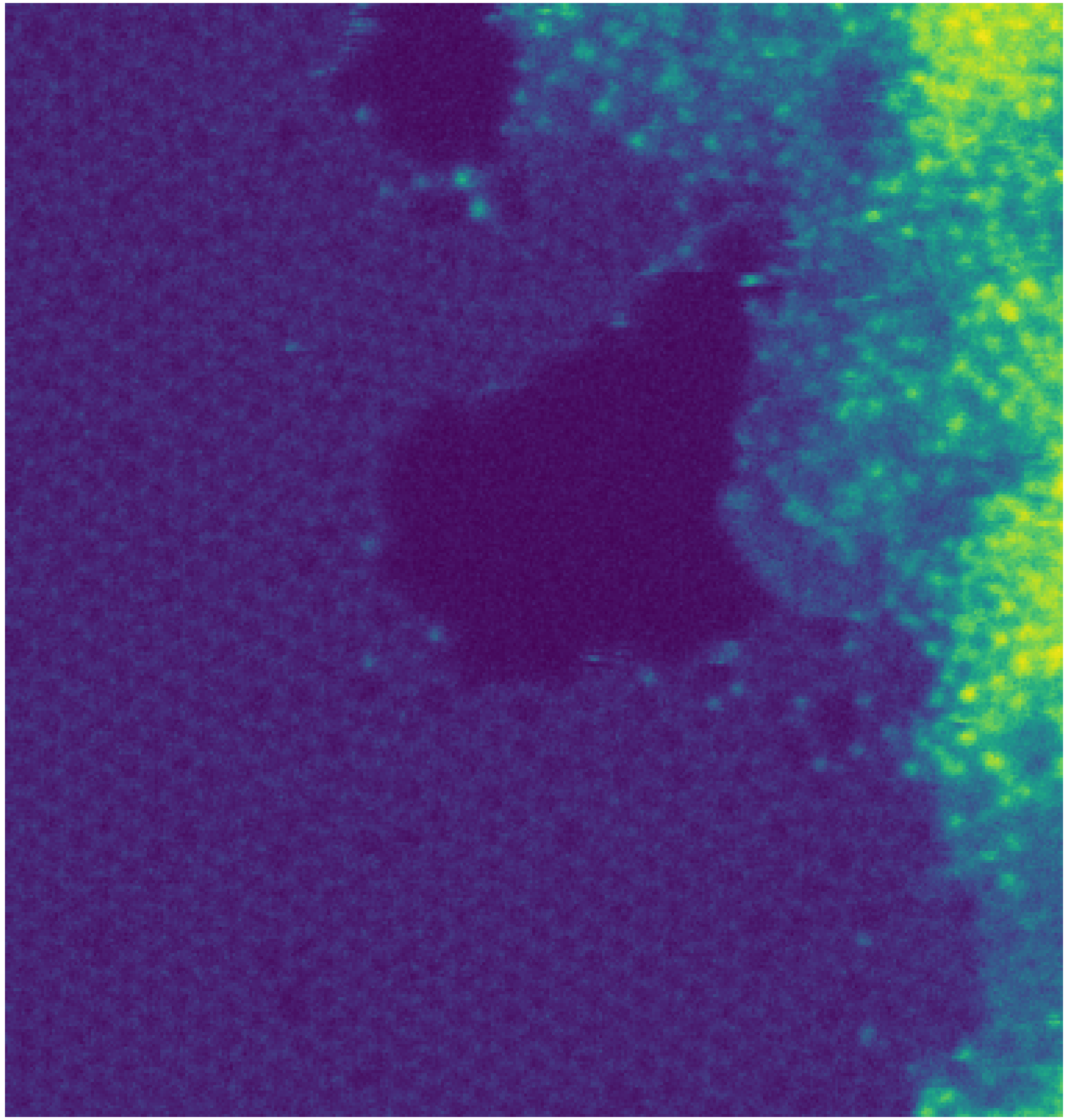
z_1
z_2
0.1336
0.0991
0.0850
0.0715
0.0588
0.0466
0.0333
0.0220
0.0144
0.0057
0.3758
0.2072
0.1207
0.0895
0.0712
0.0574
0.0439
0.0274
0.0175
0.0086

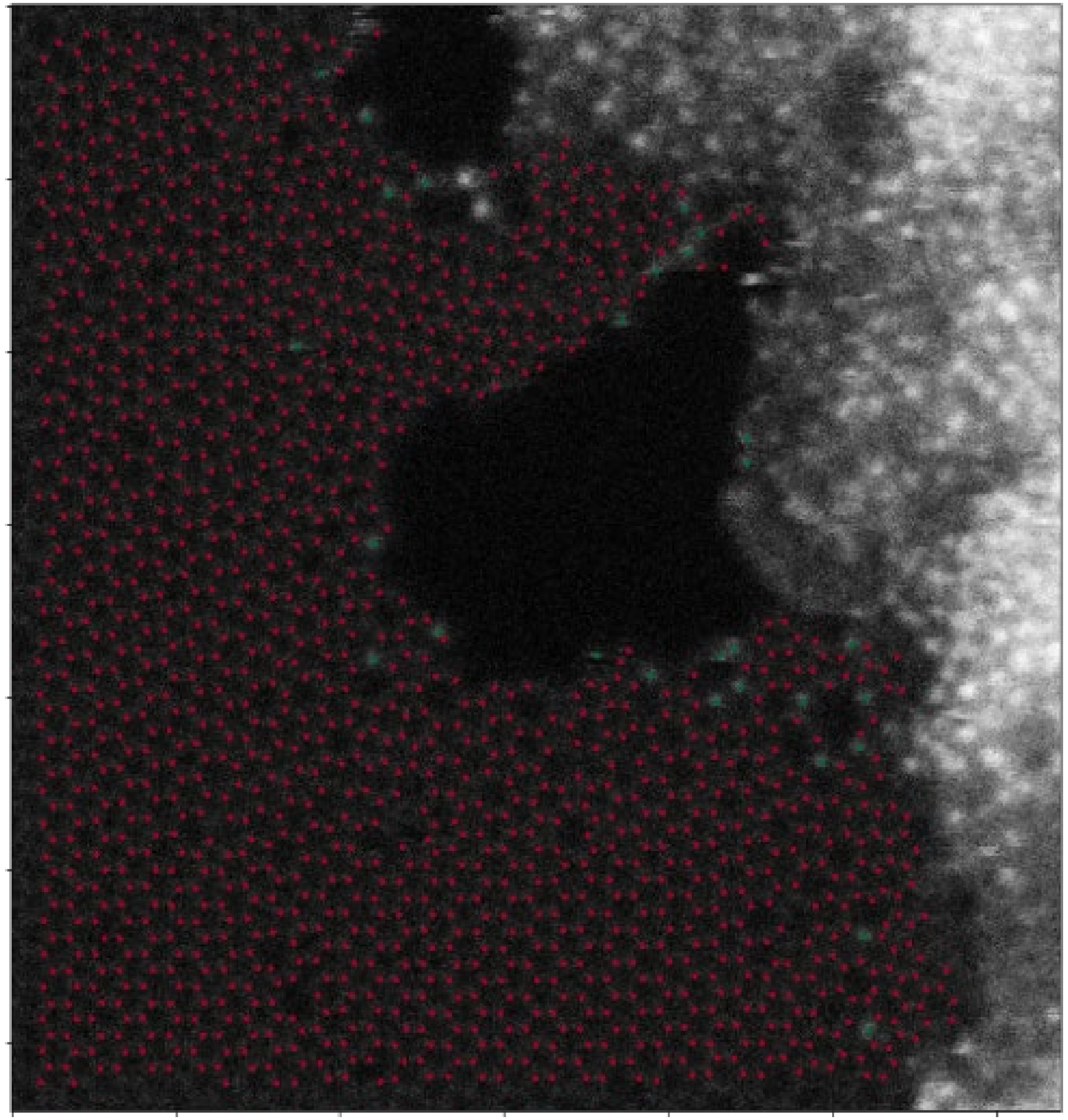
Shift VAE: Translational Invariance



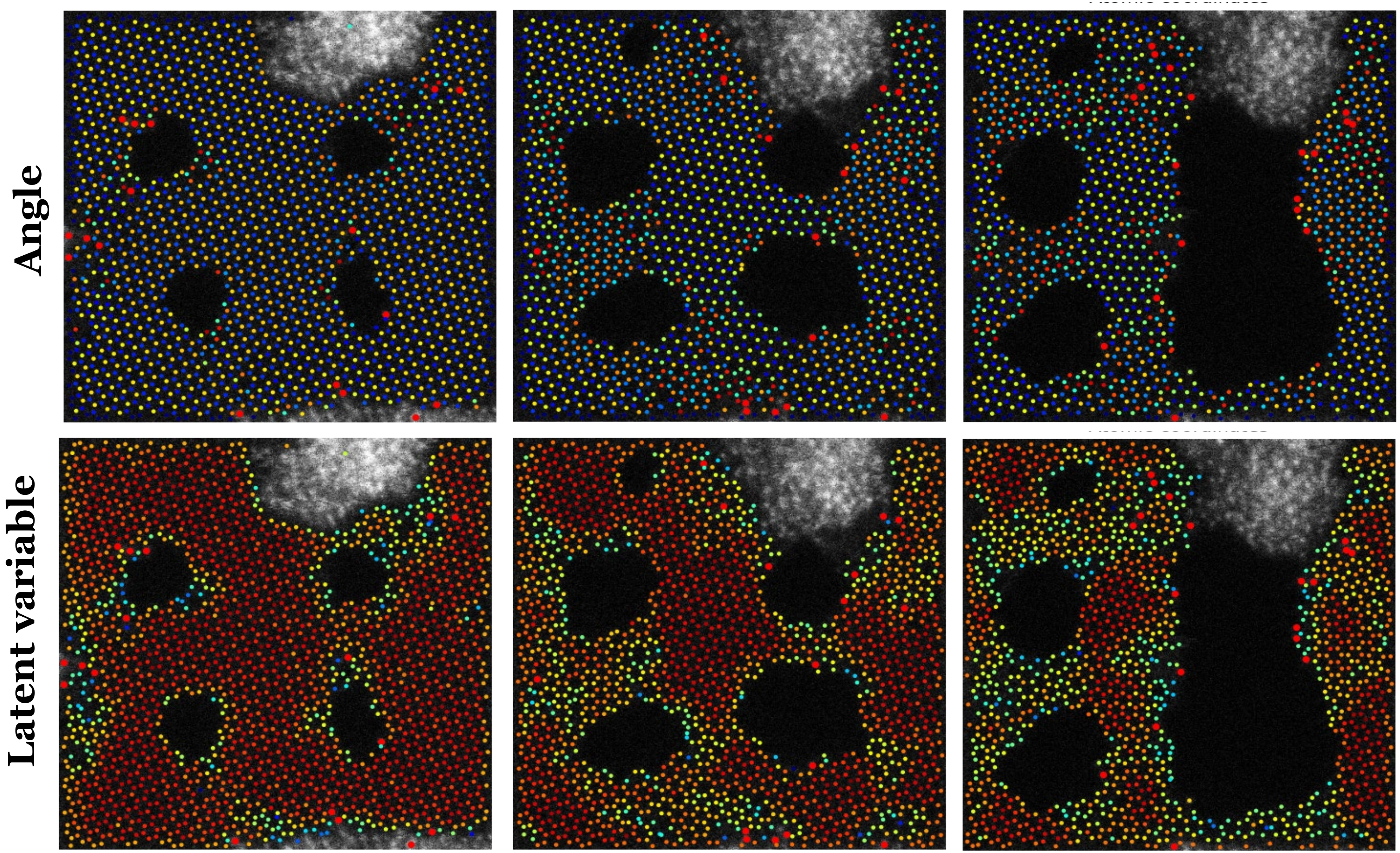


Off to chemically-disordered systems

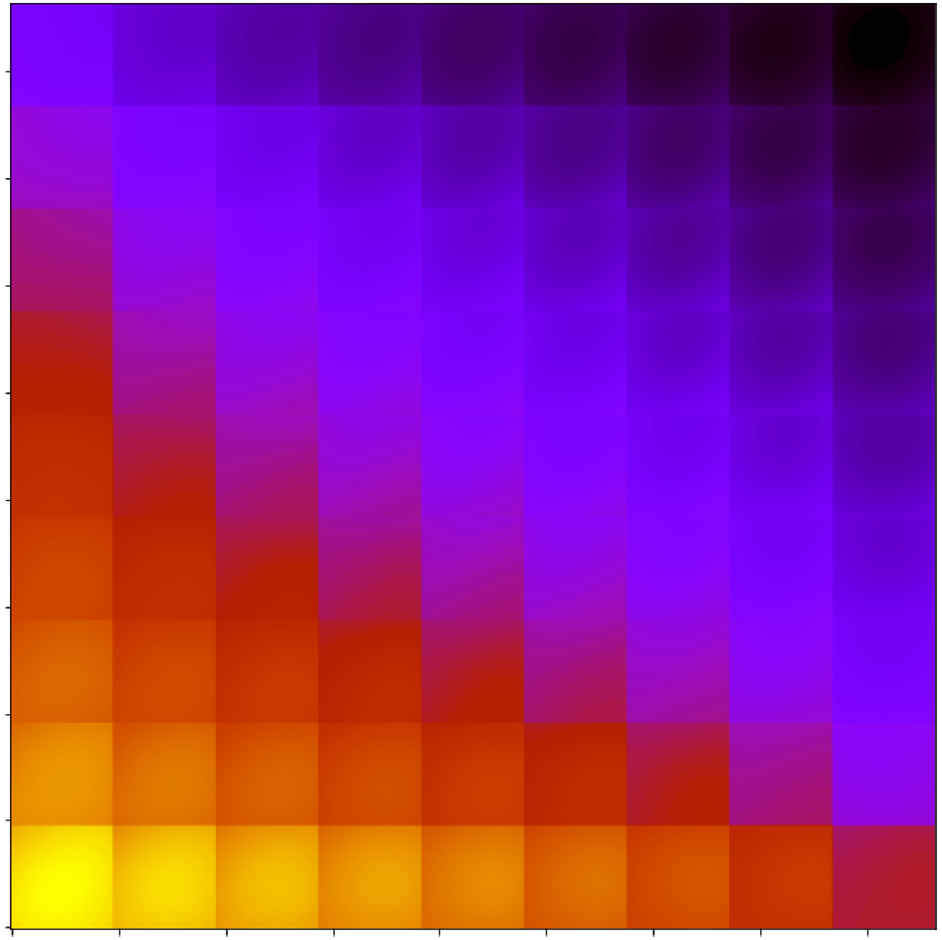


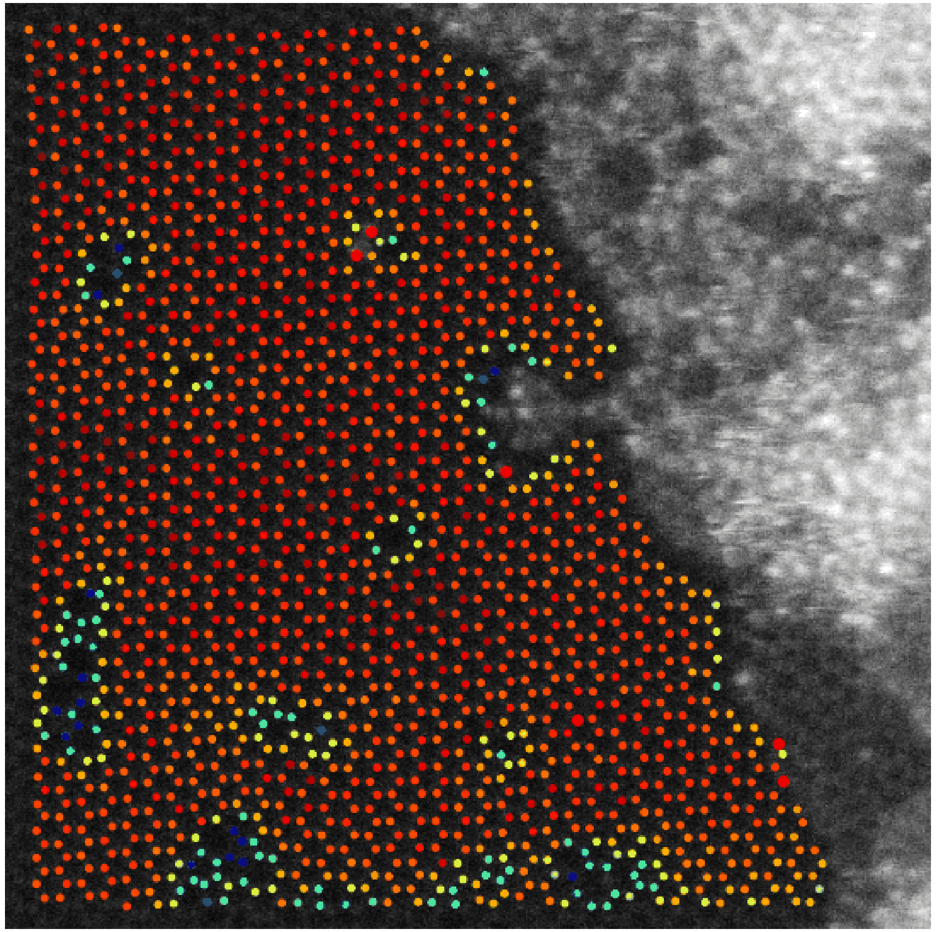


rVAE analysis at different time steps

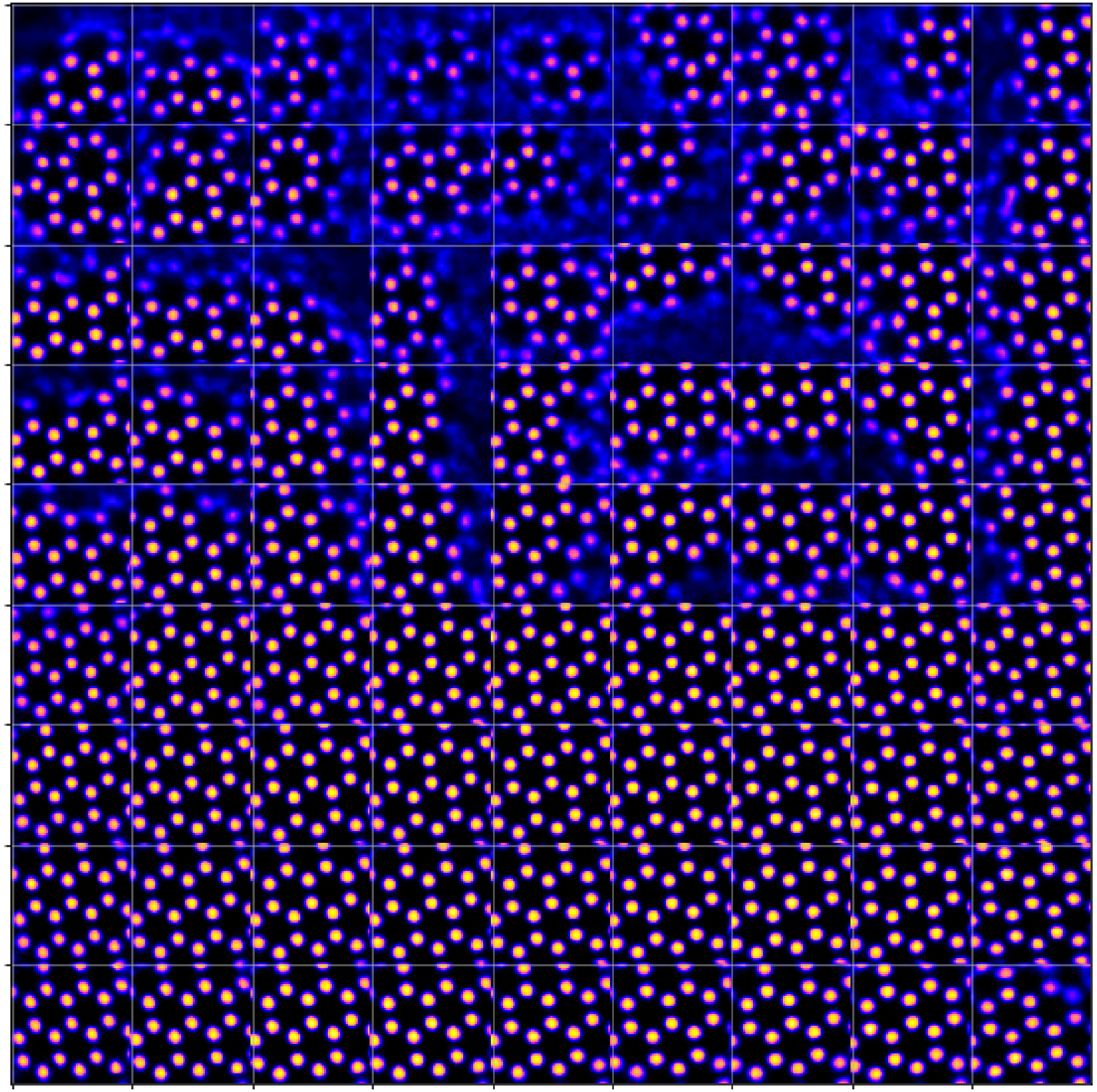


There is nothing as beautiful as training VAE

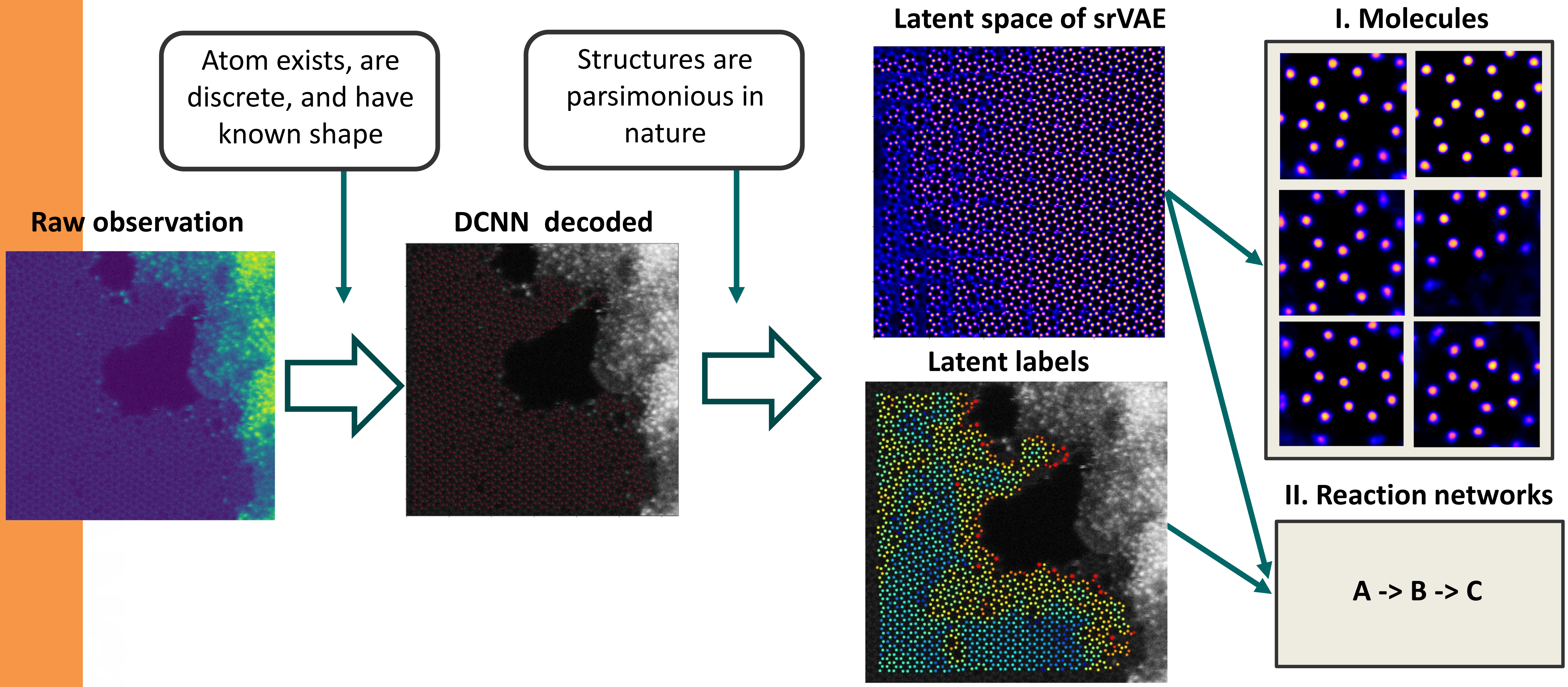




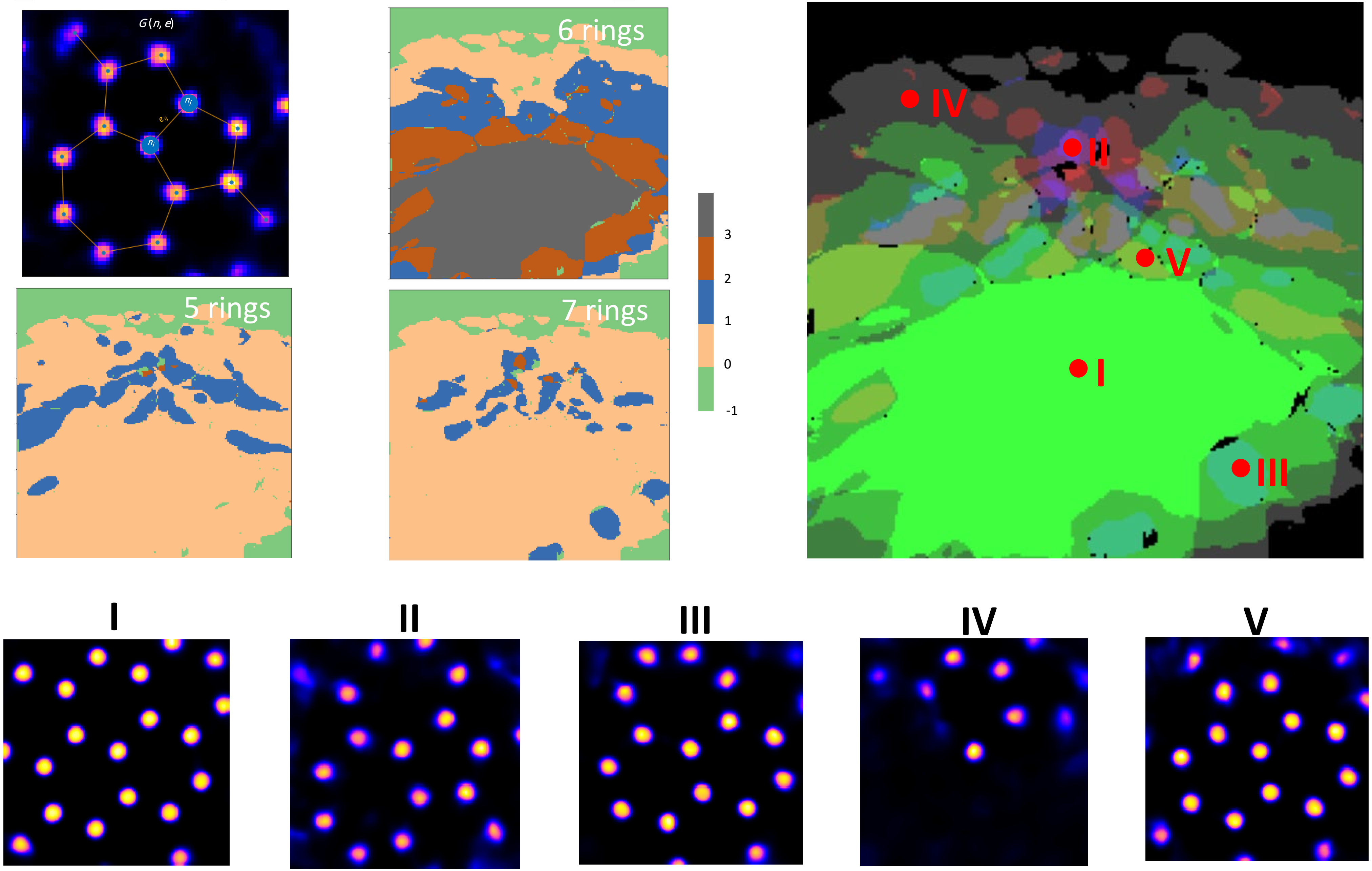
Next step: skip-rVAE



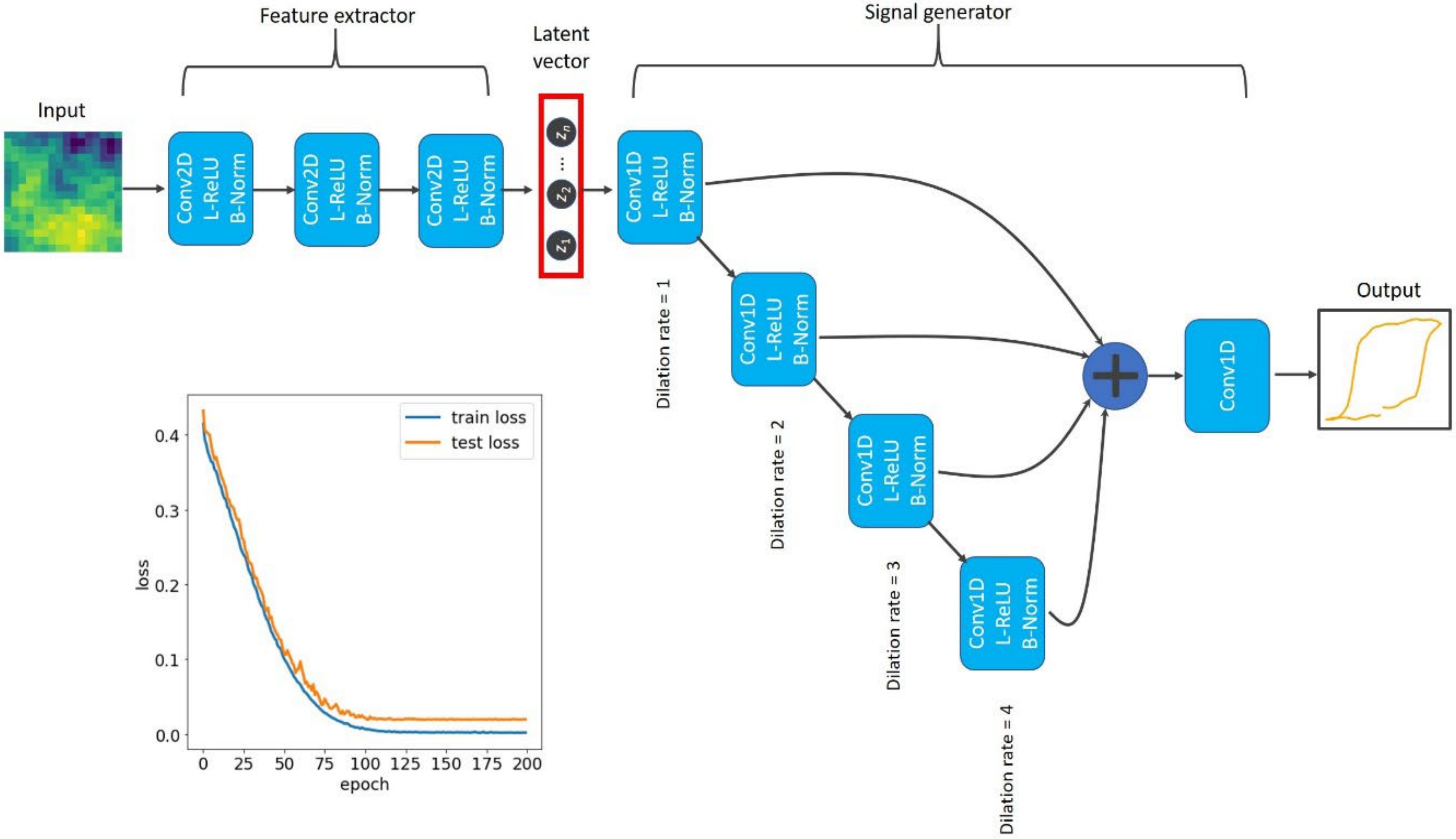
Unsupervised discovery of molecules



Exploring the latent space structure



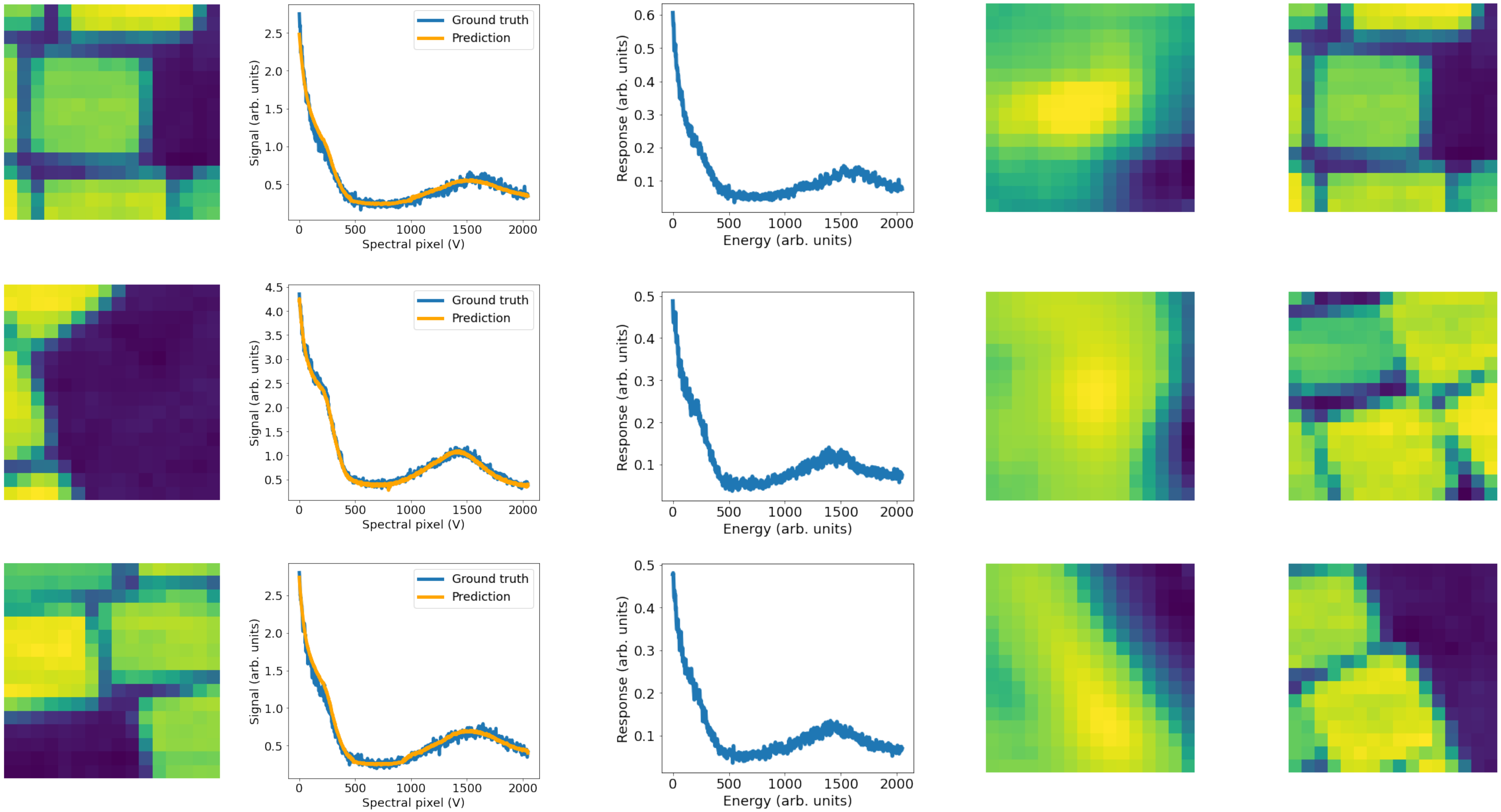
Encoders-Decoders



- Use encoder-decoder architecture to transform local structure to local spectra
- And spectra to images
- Predictive within the image

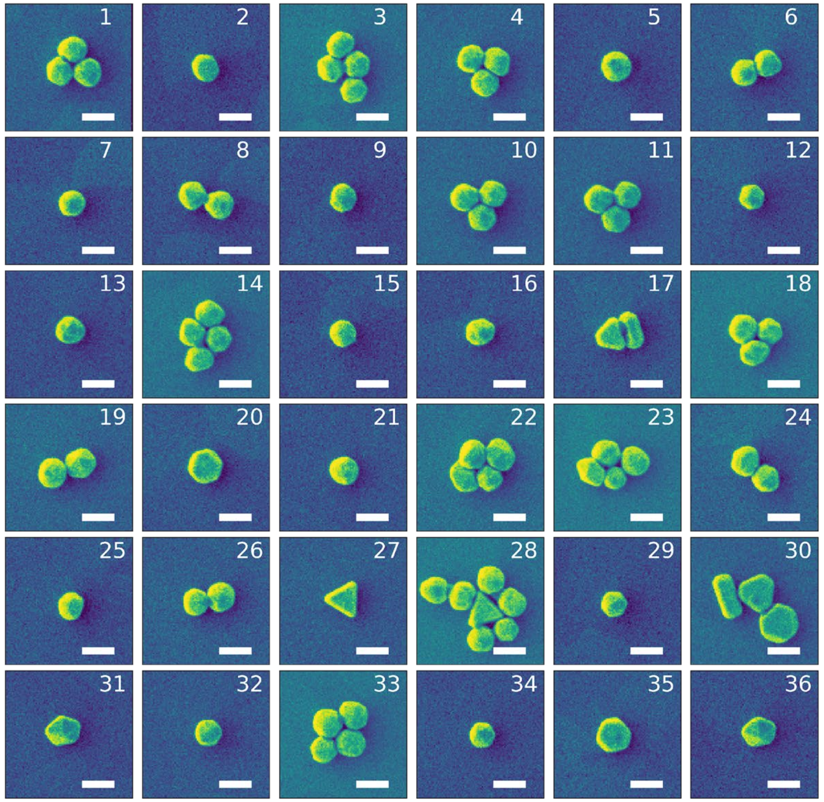
Encoders-Decoders

Sub-image | **Prediction vs. truth** | **Spectrum** | **Prediction** | **Ground truth**

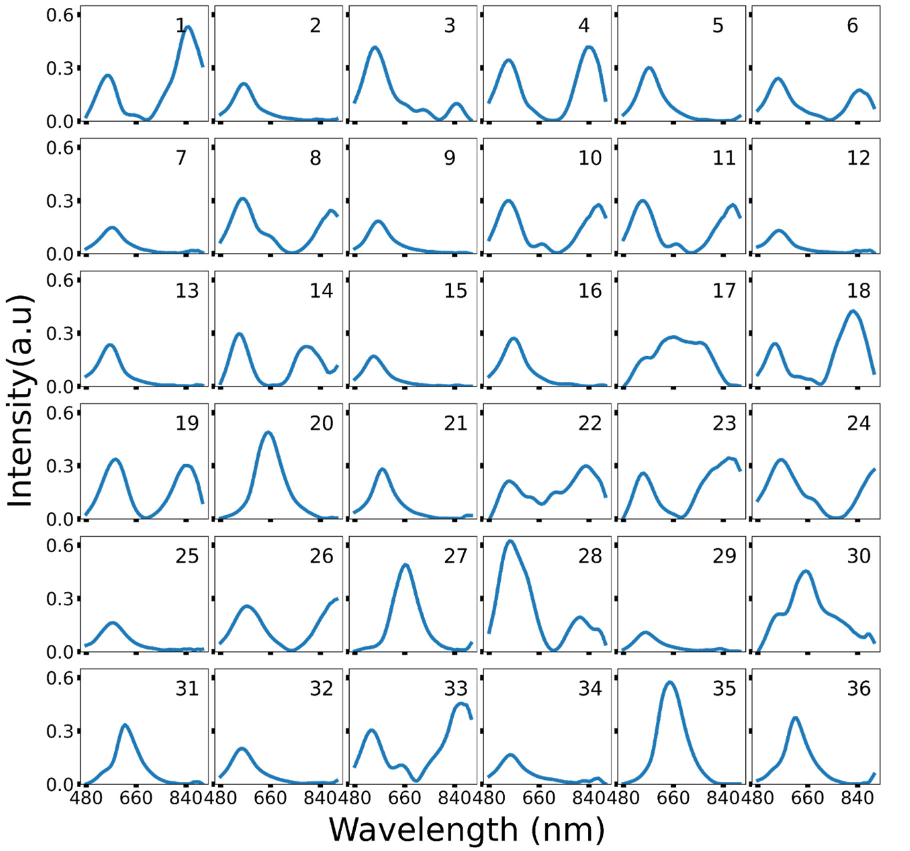


Dual VAE: structure-property relationships

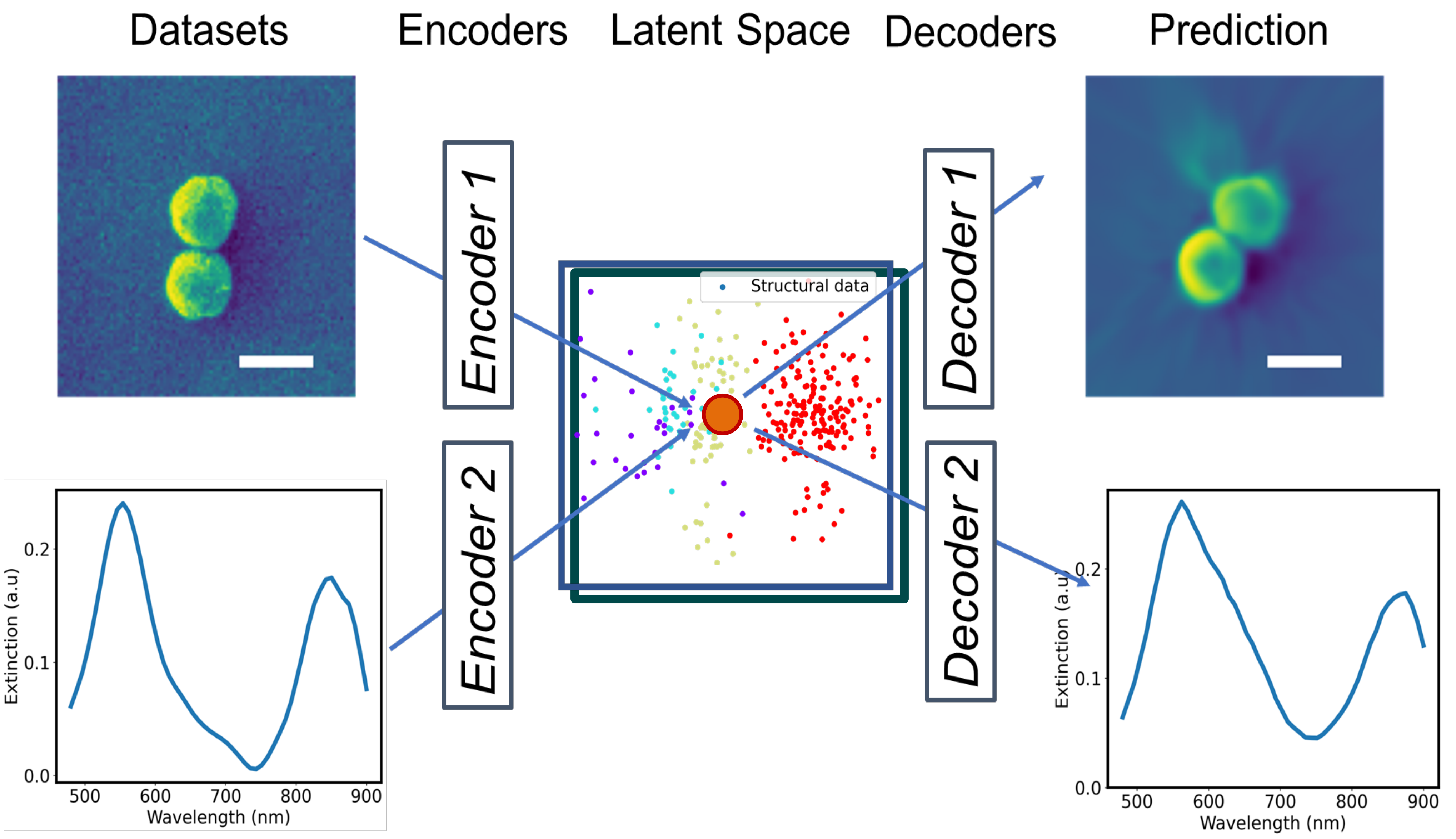
SEM images: “Structure Information”



Hyperspectral microscope: “Property Information”



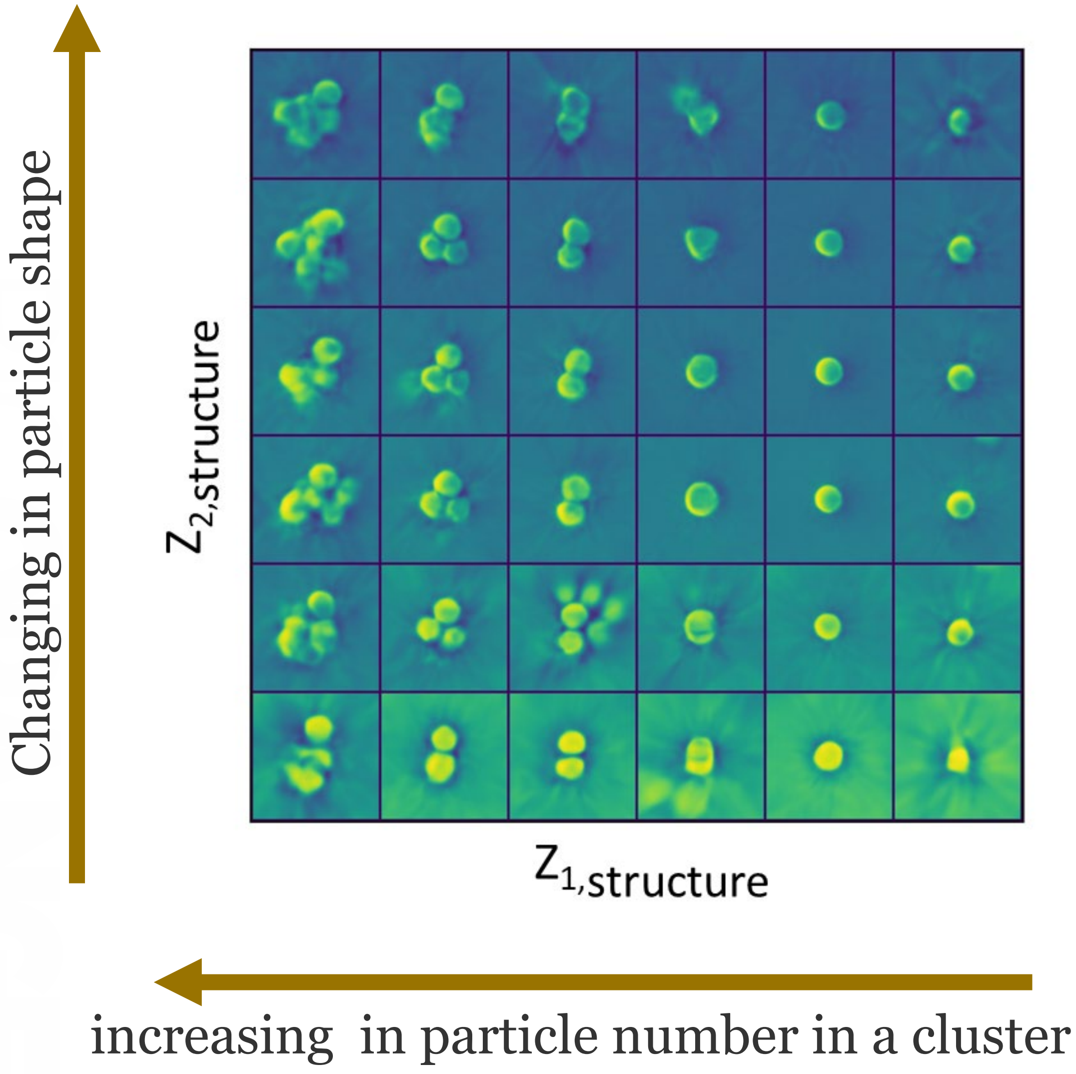
Dual VAE

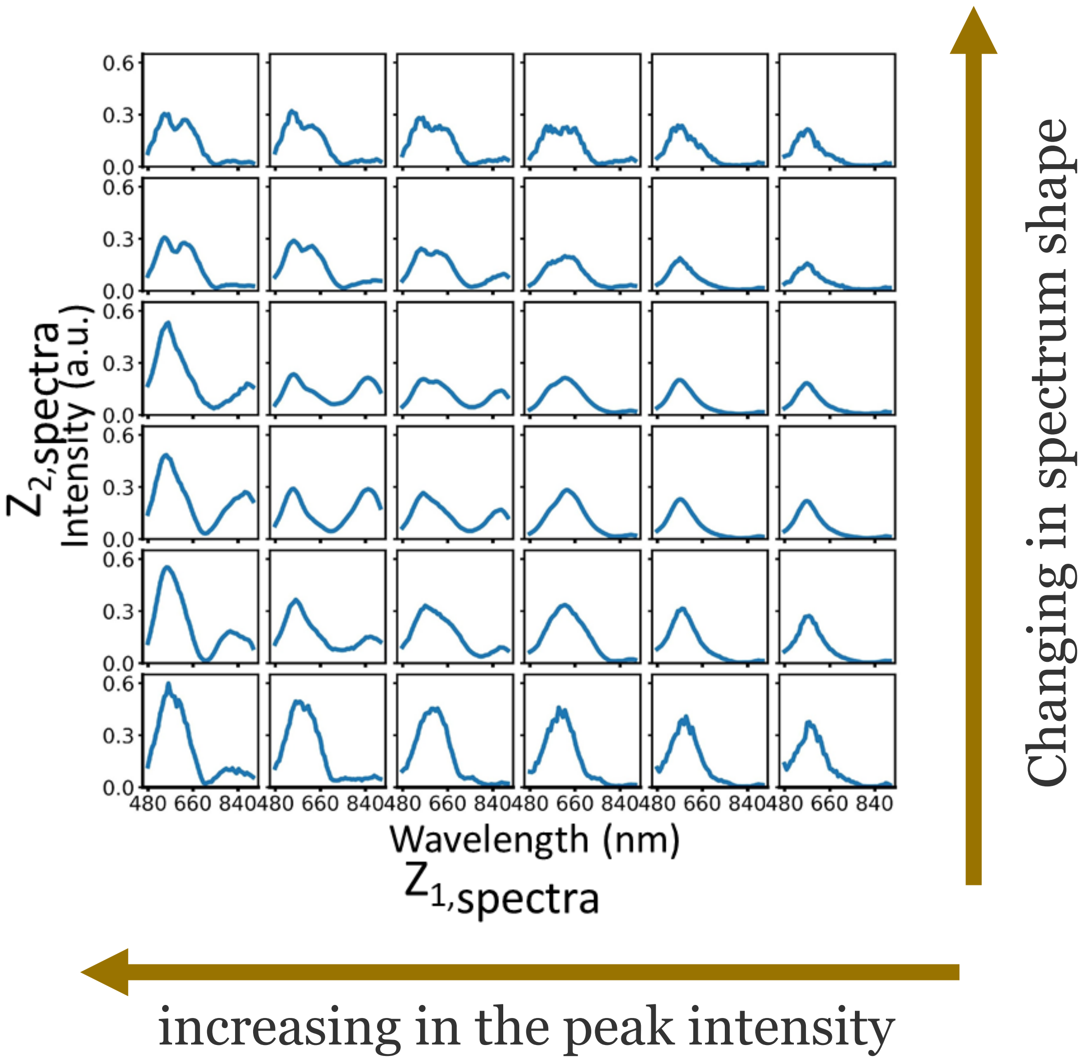


Datasets
Encoders
Latent Space
Decoders
Prediction
Encoder 1
Encoder 2
Decoder 1
Decoder 2
Structural data
Extinction (a.u)
Wavelength (nm)
0.0
0.1
0.2
500
600
700
800
900

Dual VAE: Latent Representations

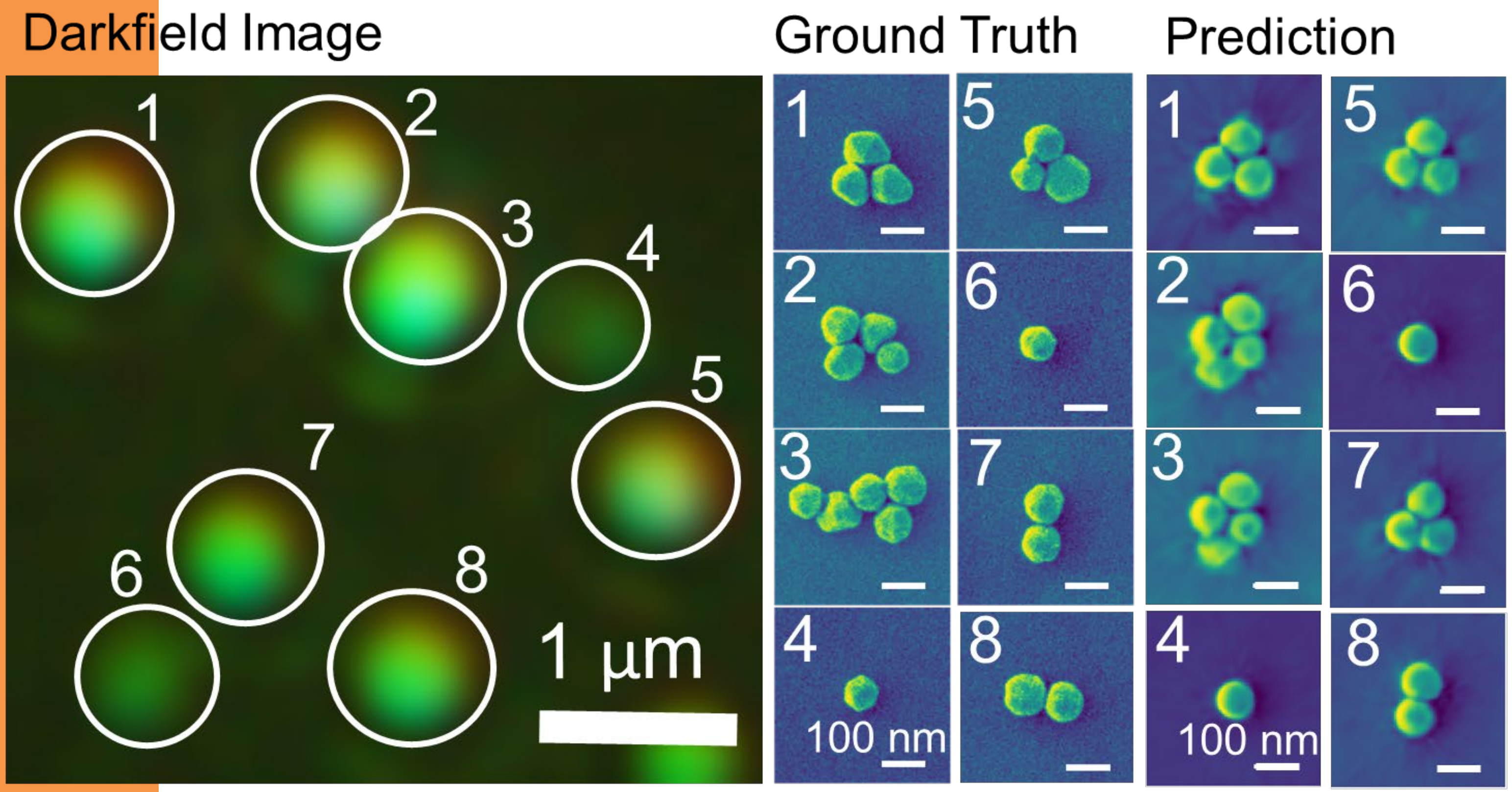
Manifold Representation



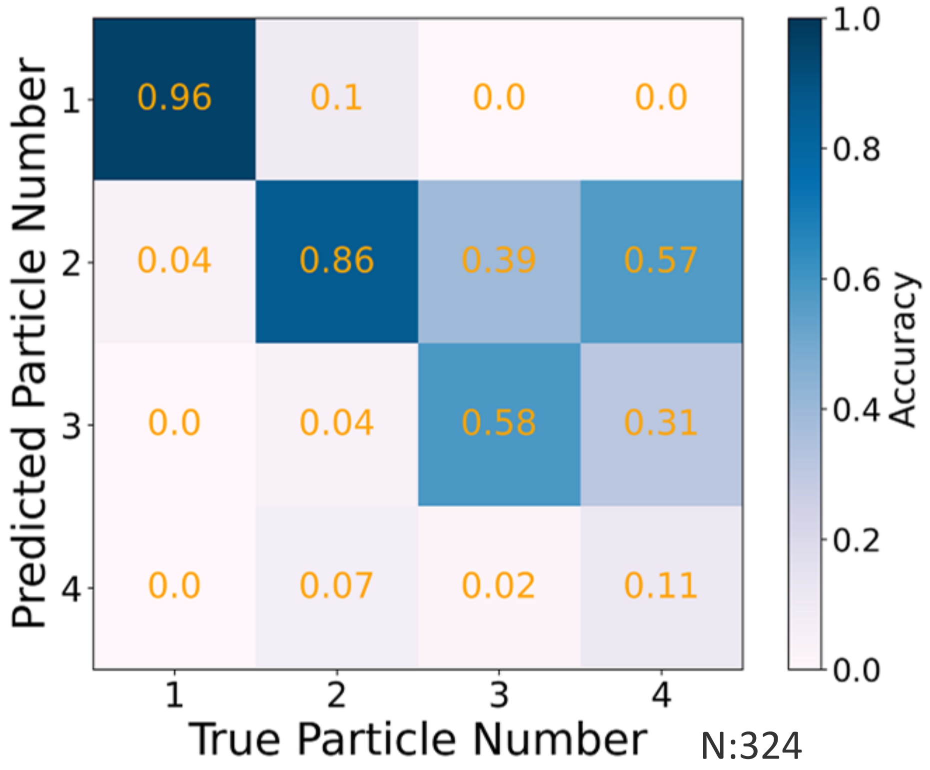


Dual VAE: Predictions

Example



Overall Particles



Article | Open access | Published: 14 August 2024

Physics and chemistry from parsimonious representations: image analysis via invariant variational autoencoders

Mani Valleti, Maxim Ziatdinov, Yongtao Liu & Sergei V. Kalinin

npj Computational Materials **10**, Article number: 183 (2024) | Cite this article

Abstract

Electron, optical, and scanning probe microscopy methods are generating ever increasing volume of image data containing information on atomic and mesoscale structures and functionalities. This necessitates the development of the machine learning methods for discovery of physical and chemical phenomena from the data, such as manifestations of symmetry breaking phenomena in electron and scanning tunneling microscopy images, or variability of the nanoparticles. Variational autoencoders (VAEs) are emerging as a powerful paradigm for the unsupervised data analysis, allowing to disentangle the factors of variability and discover optimal parsimonious representation. Here, we summarize recent developments

Variational Autoencoder Toolkit

Kamyar Barakati, Chris Nelson, Anna N. Morozovska, Maxim A. Ziatdinov, Eugene A. Eliseev, Xiaohang Zhang, Ichiro Takeuchi, Sergei V. Kalinin

April 17, 2025 · https://doi.org/10.69761/udpm2547

VAE Quickstart

Abstract

Variational Autoencoders (VAEs) provide a robust framework for extracting latent structures and revealing physical insights in high-dimensional microscopy data. Although initial clustering with Gaussian Mixture Models (GMM) in the original descriptor space reveals distinct domain configurations and substrate regions, we miss underlying details, such as domain walls associated with unit cell rotations, chemical and mis-tilt effects, and the distinction between positive and negative wall orientations. By combining invariant and conditional VAEs, we capture nuanced features by encoding intrinsic factors of variation, such as orientation and translation, while incorporating cation type. This approach effectively links structural characteristics with physical phenomena. This approach demonstrates how invariant and conditional VAEs can extract both data-driven and physically interpretable features, offering an adaptable toolkit for analyzing ferroic materials, multiphase systems, and similar complex datasets.

https://www.elementalmicroscopy.com/articles/EM000006