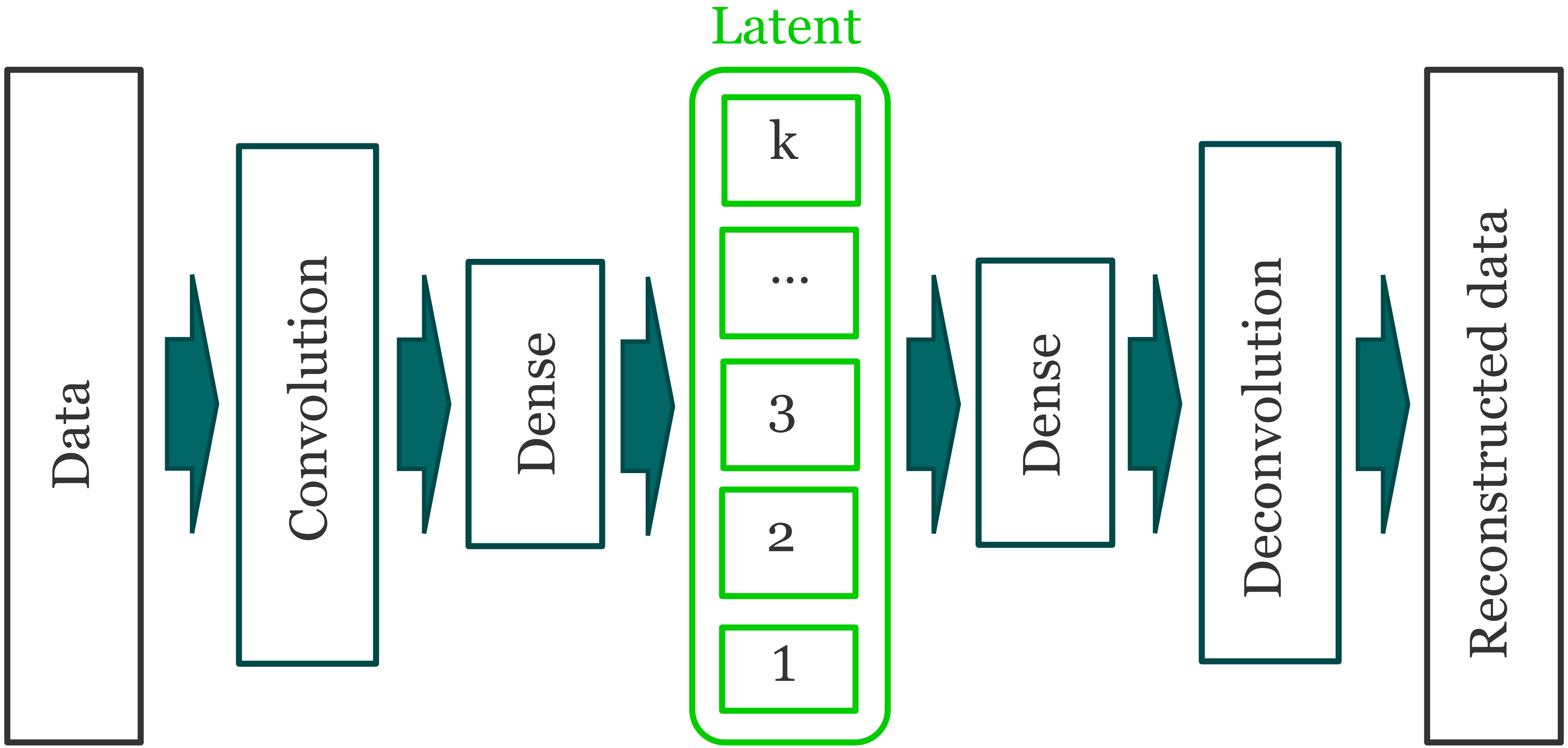
THE UNIVERSITY of TENNESSEE KNOXVILLE

Lecture 33: Simple and Variational Autoencoders

Instructor: Sergei V. Kalinin

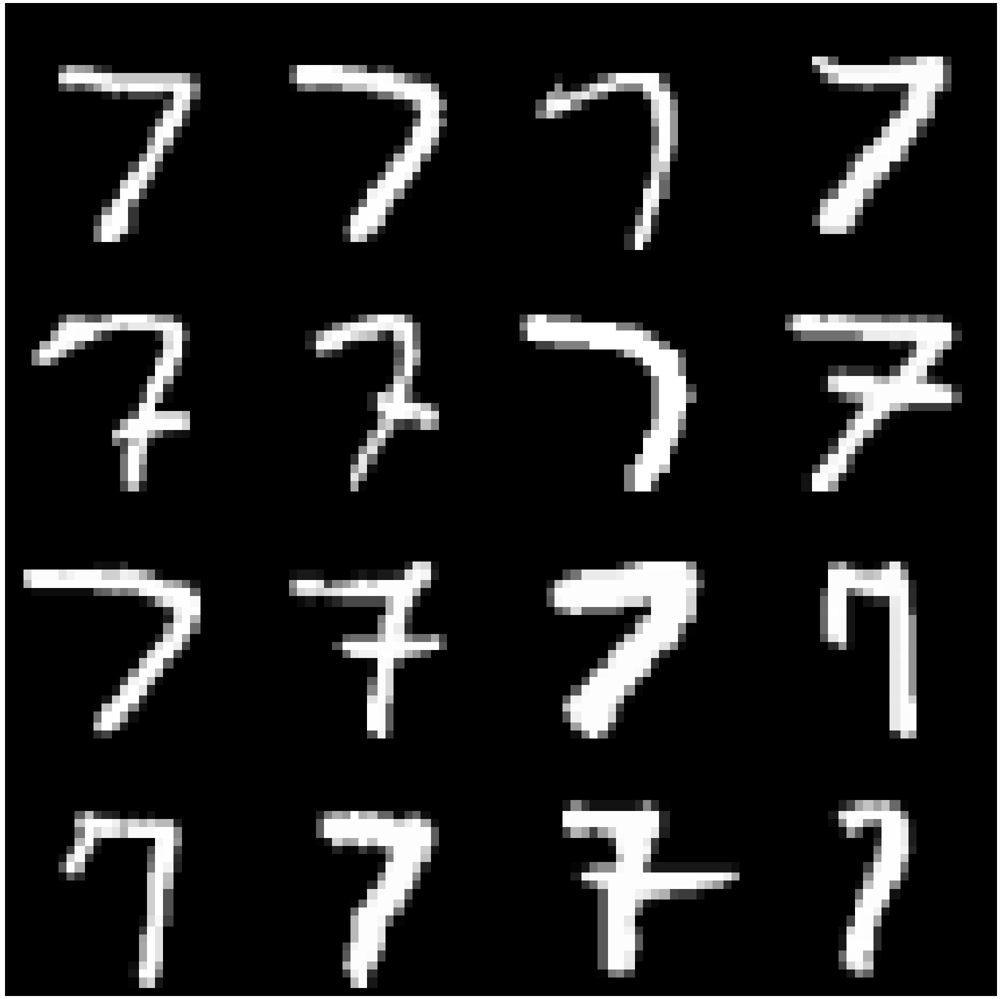
Autoencoders



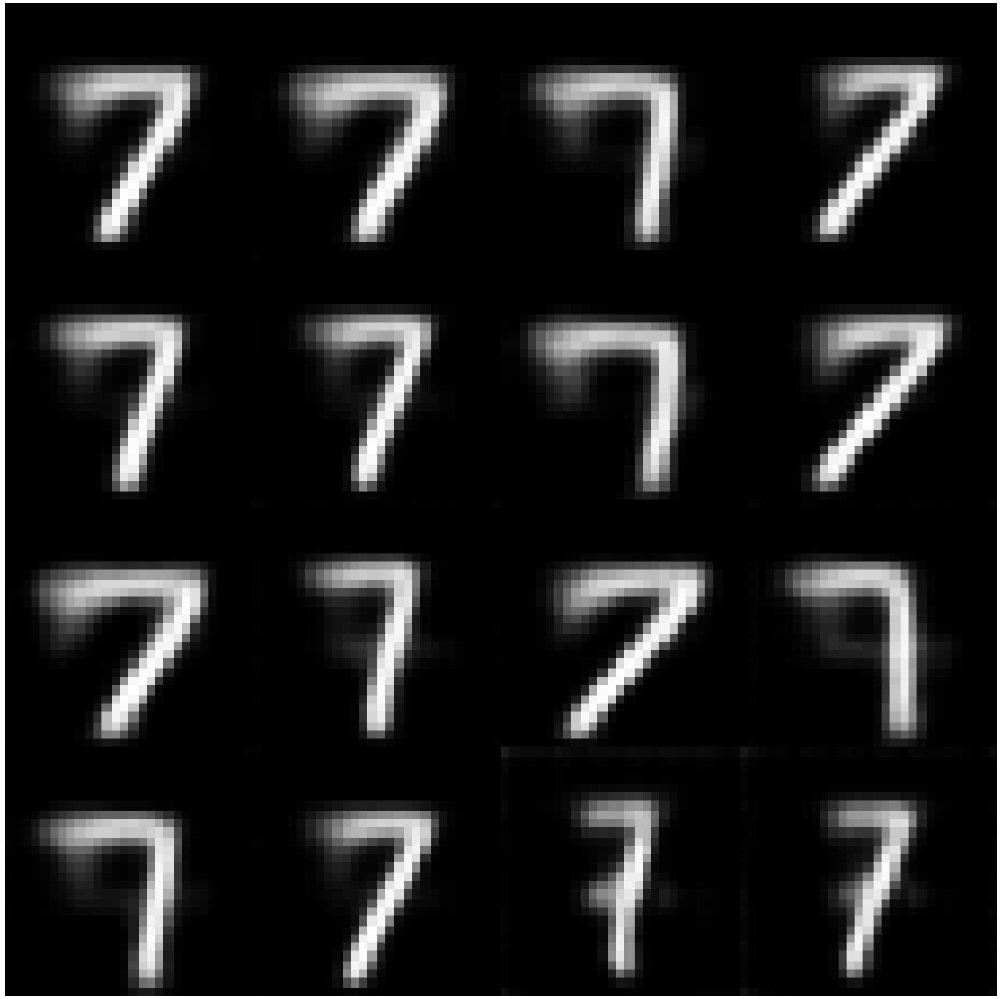
Loss: reconstruction loss

The AE reconstructs data

Input data



Decoded data



Why are AE important?

Geoffrey Hinton

FOLLOW

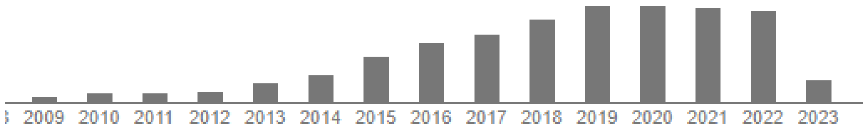
Emeritus Prof. Comp Sci, U.Toronto & Engineering Fellow, Google
Verified email at cs.toronto.edu - Homepage

machine learning   psychology   artificial intelligence   cognitive science   computer science

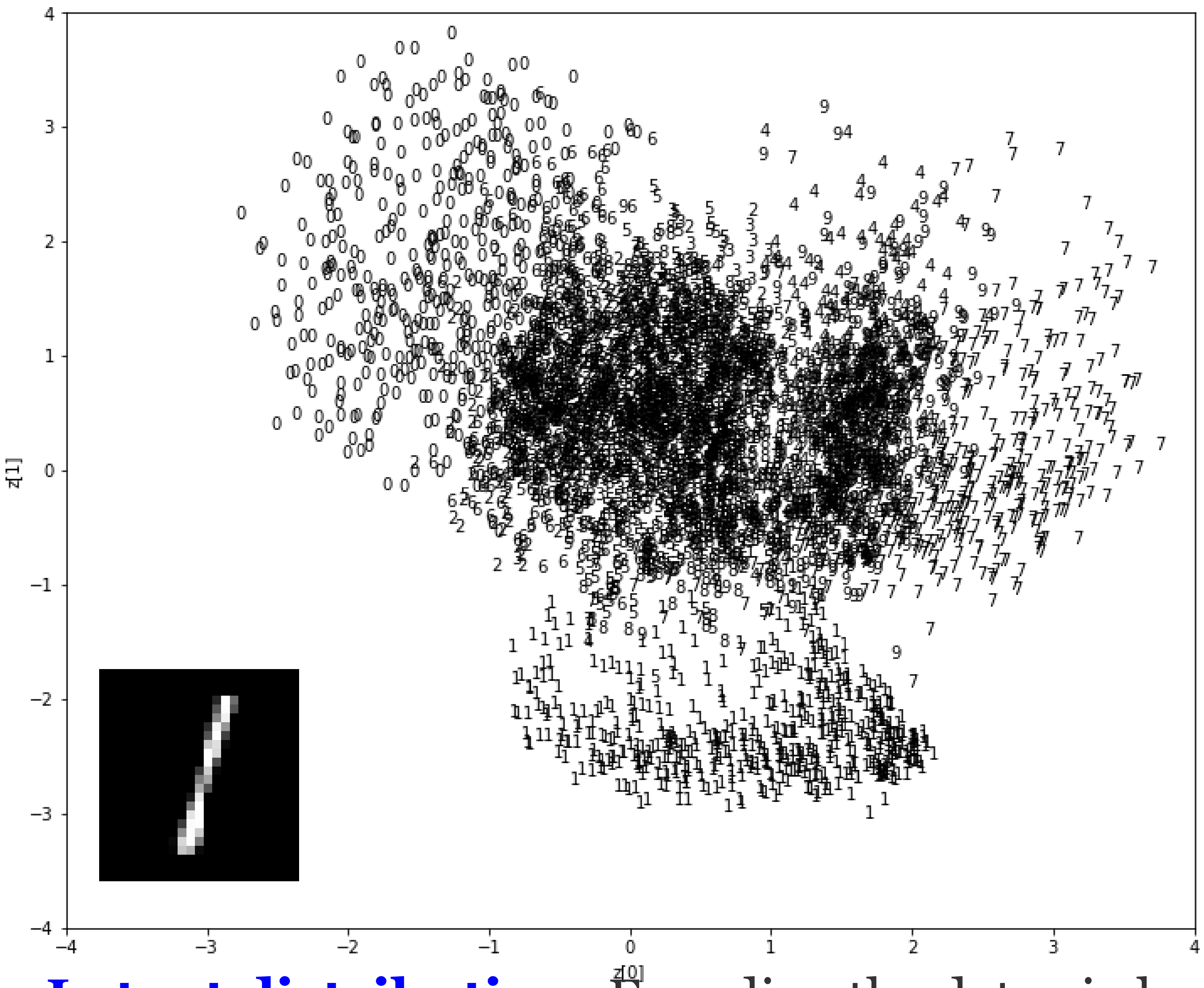
TITLE	CITED BY	YEAR
Imagenet classification with deep convolutional neural networks A Krizhevsky, I Sutskever, GE Hinton Communications of the ACM 60 (6), 84-90	130318	2017
Deep learning Y LeCun, Y Bengio, G Hinton Nature 521 (7553), 436-44	62790	2015
Dropout: a simple way to prevent neural networks 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 Smolensky, 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

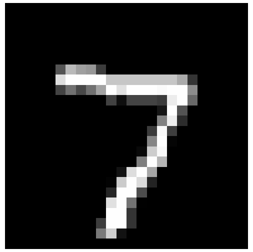
Reducing the dimensionality of data with neural networks

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



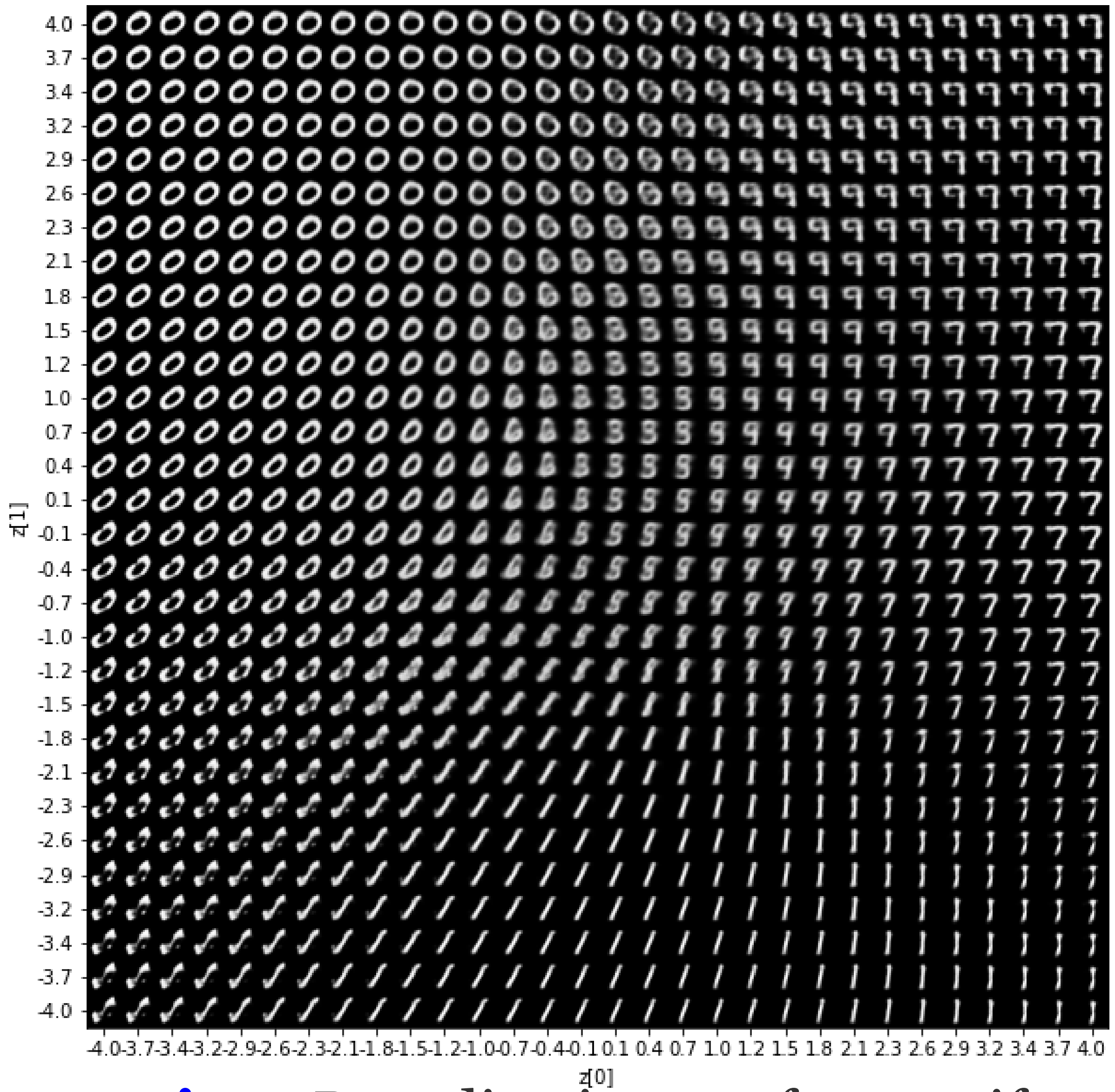
Encoding: Image → Latent Space





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

Test color images (Ground Truth)

Test gray images (Input)

Image Reconstruction

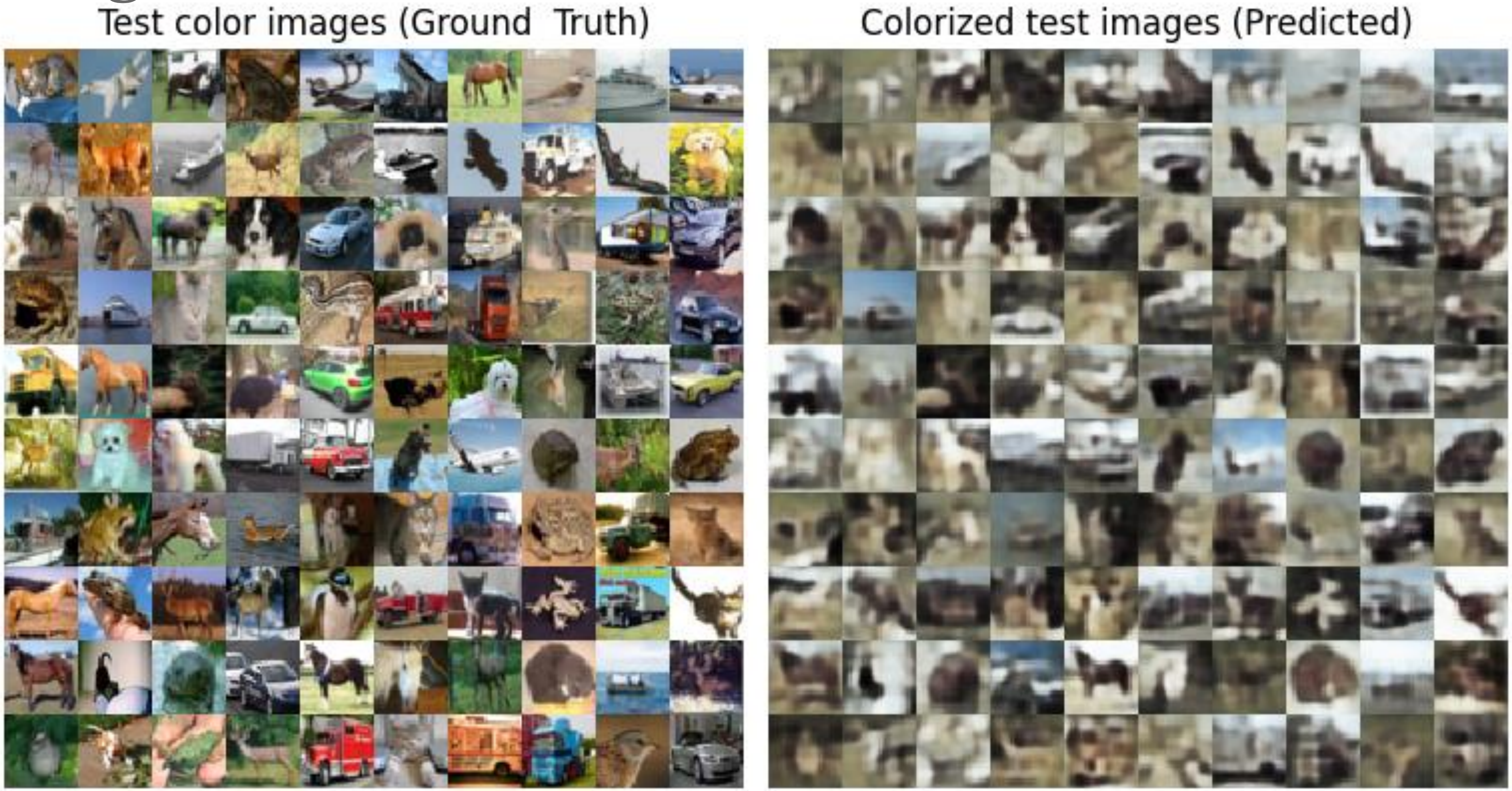
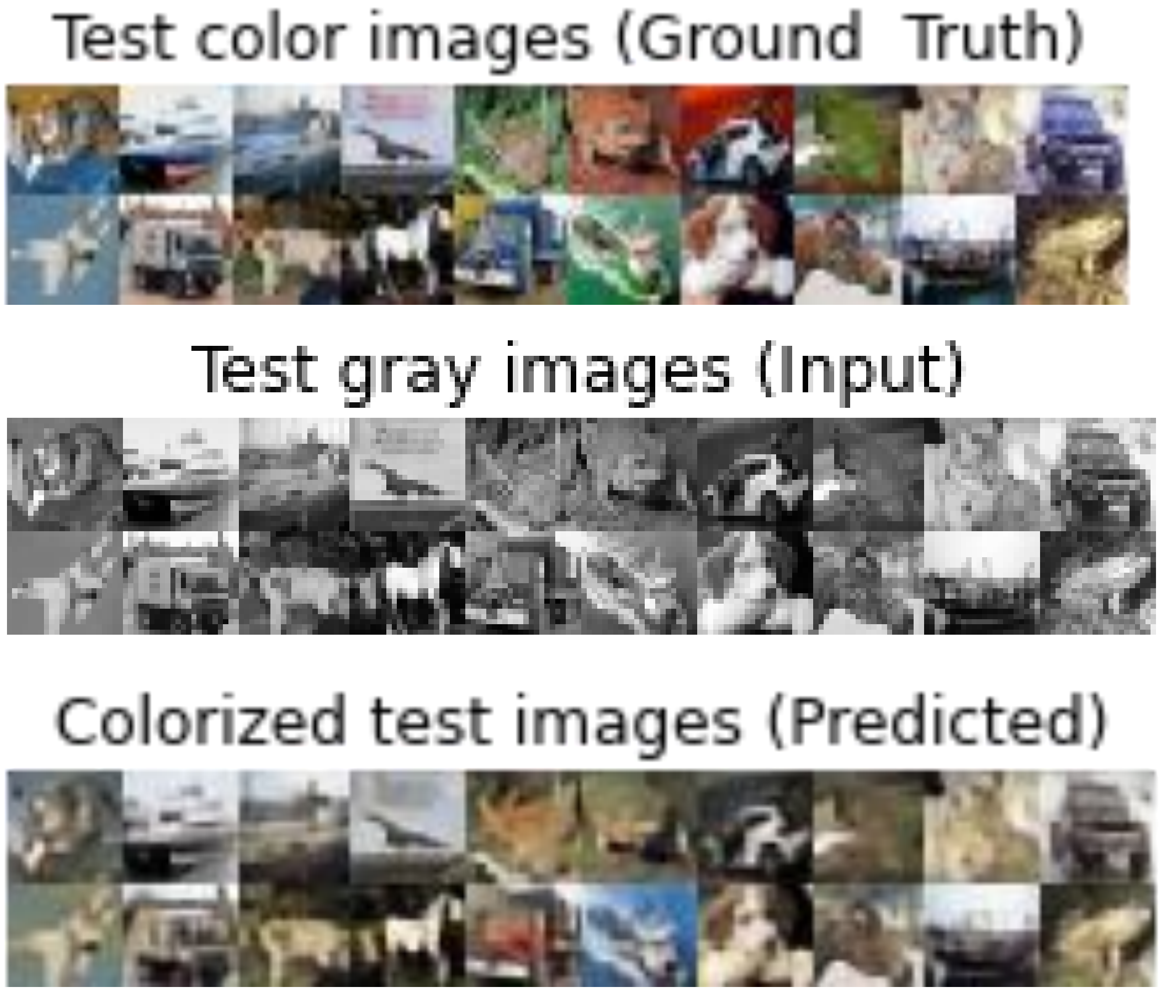


Image Reconstruction



- **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)
- **Concern:** has to be from the same distribution

Variational Autoencoders

Diederik P. Kingma

FOLLOW

Other names

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

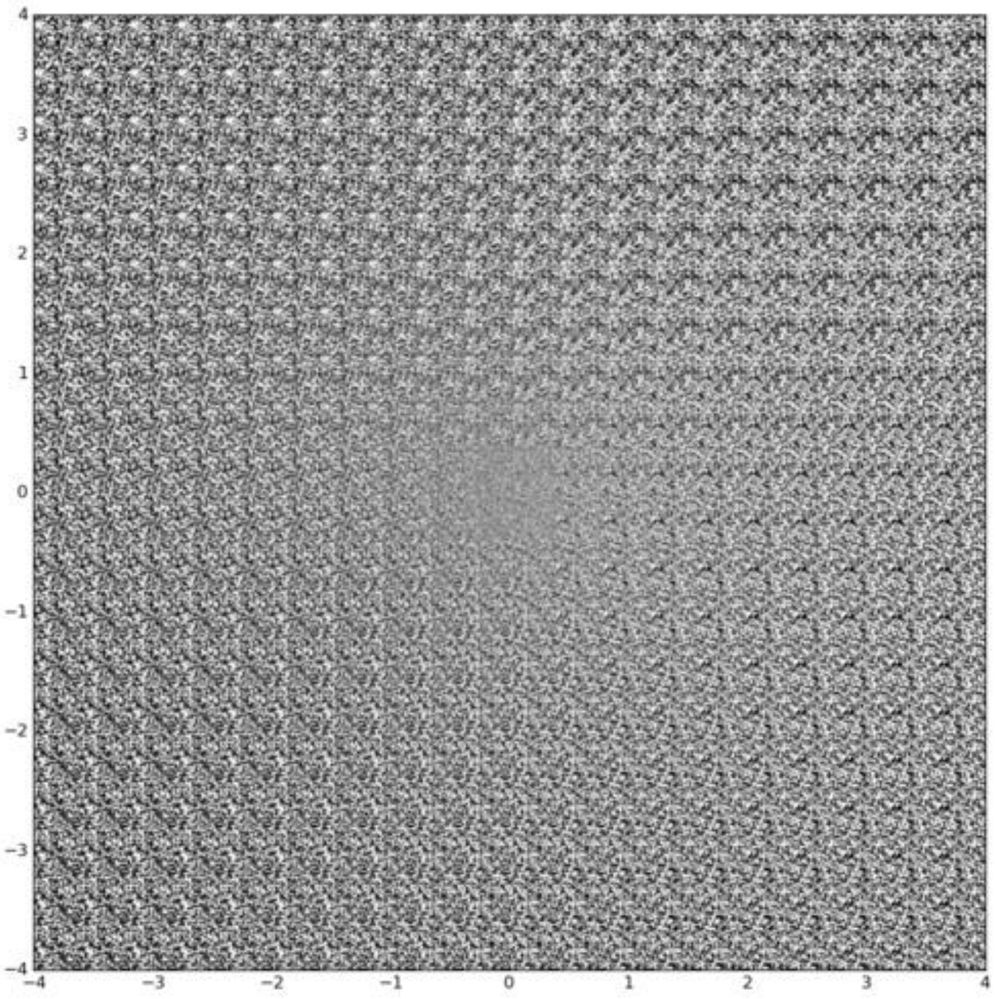
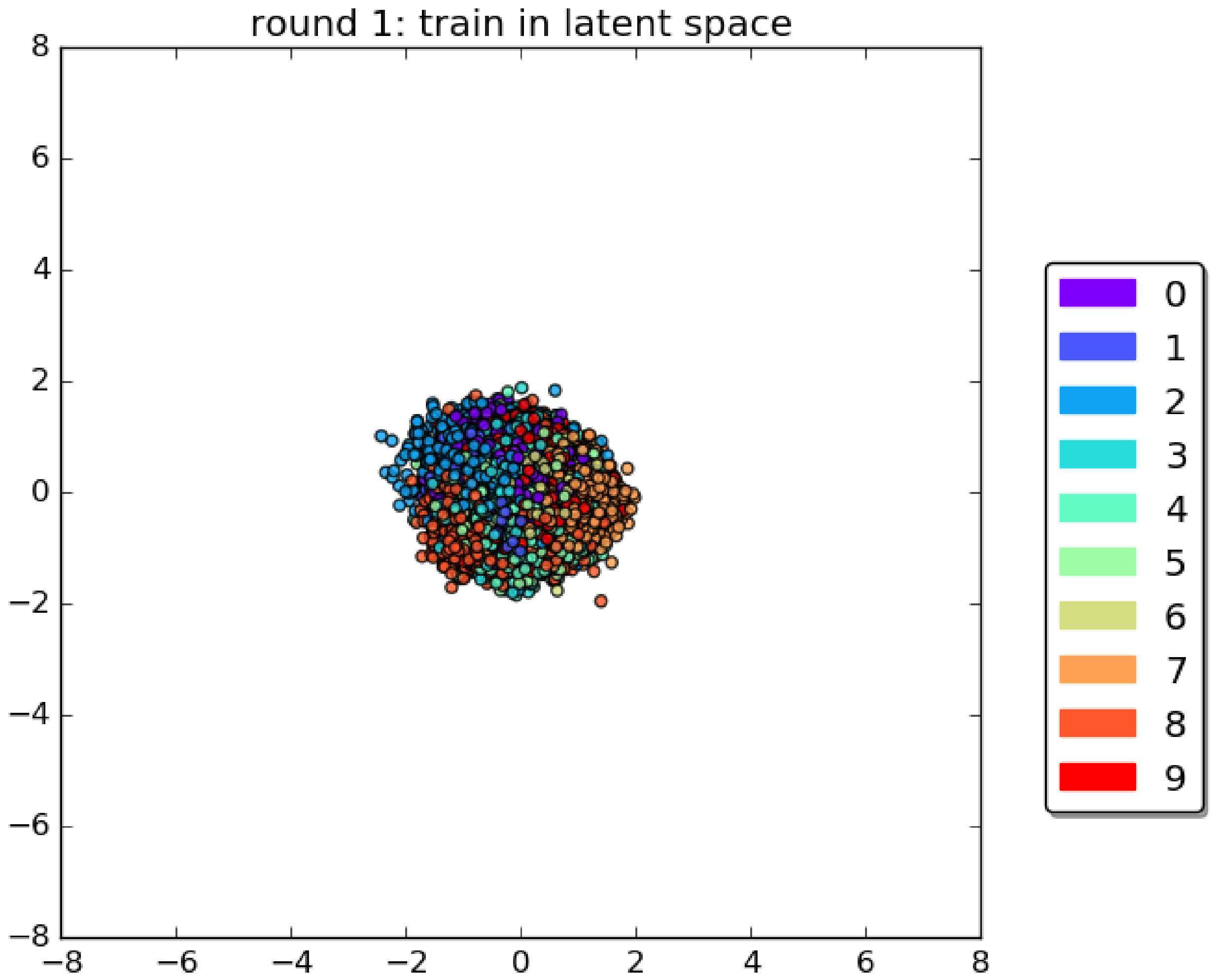
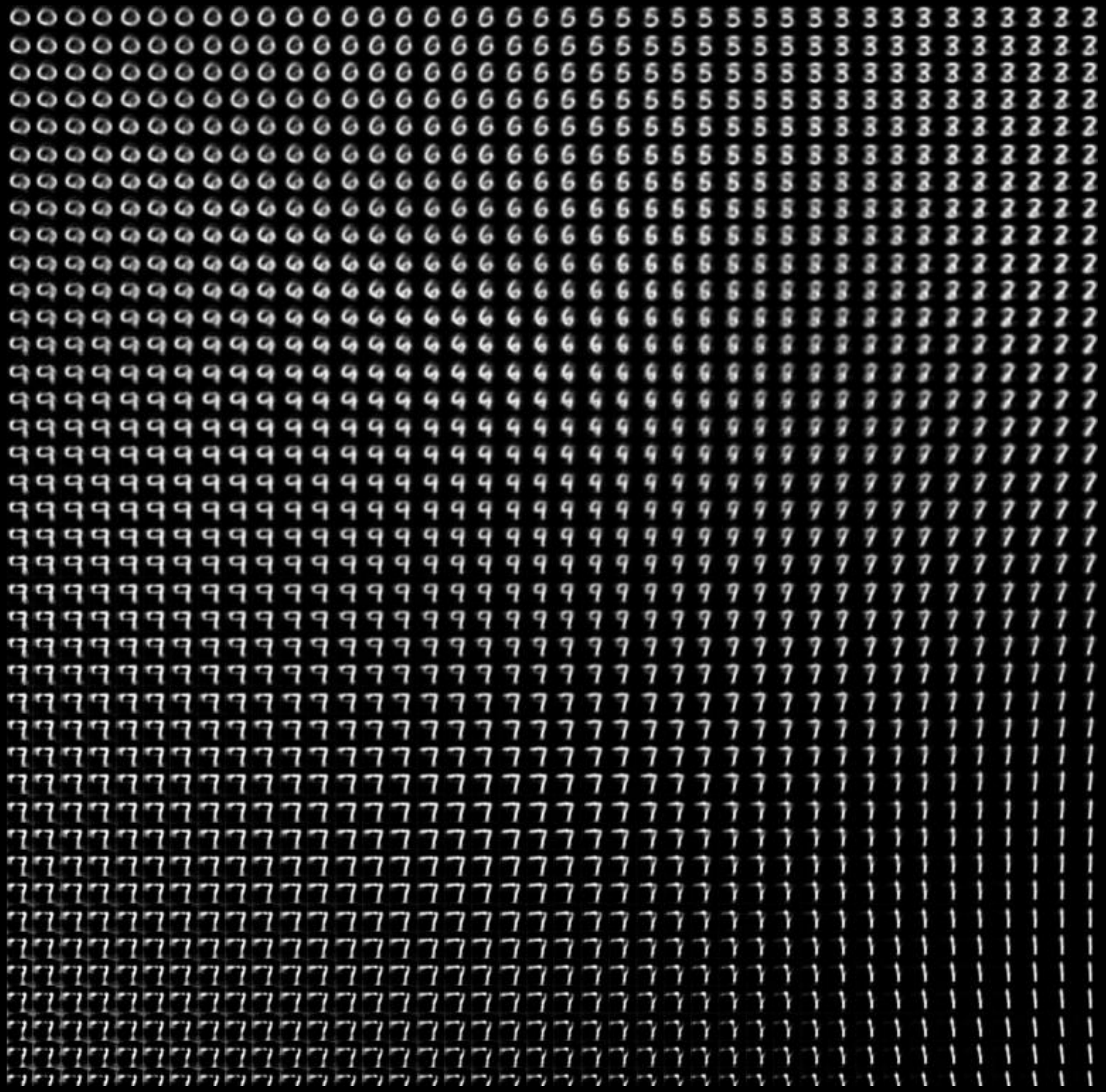


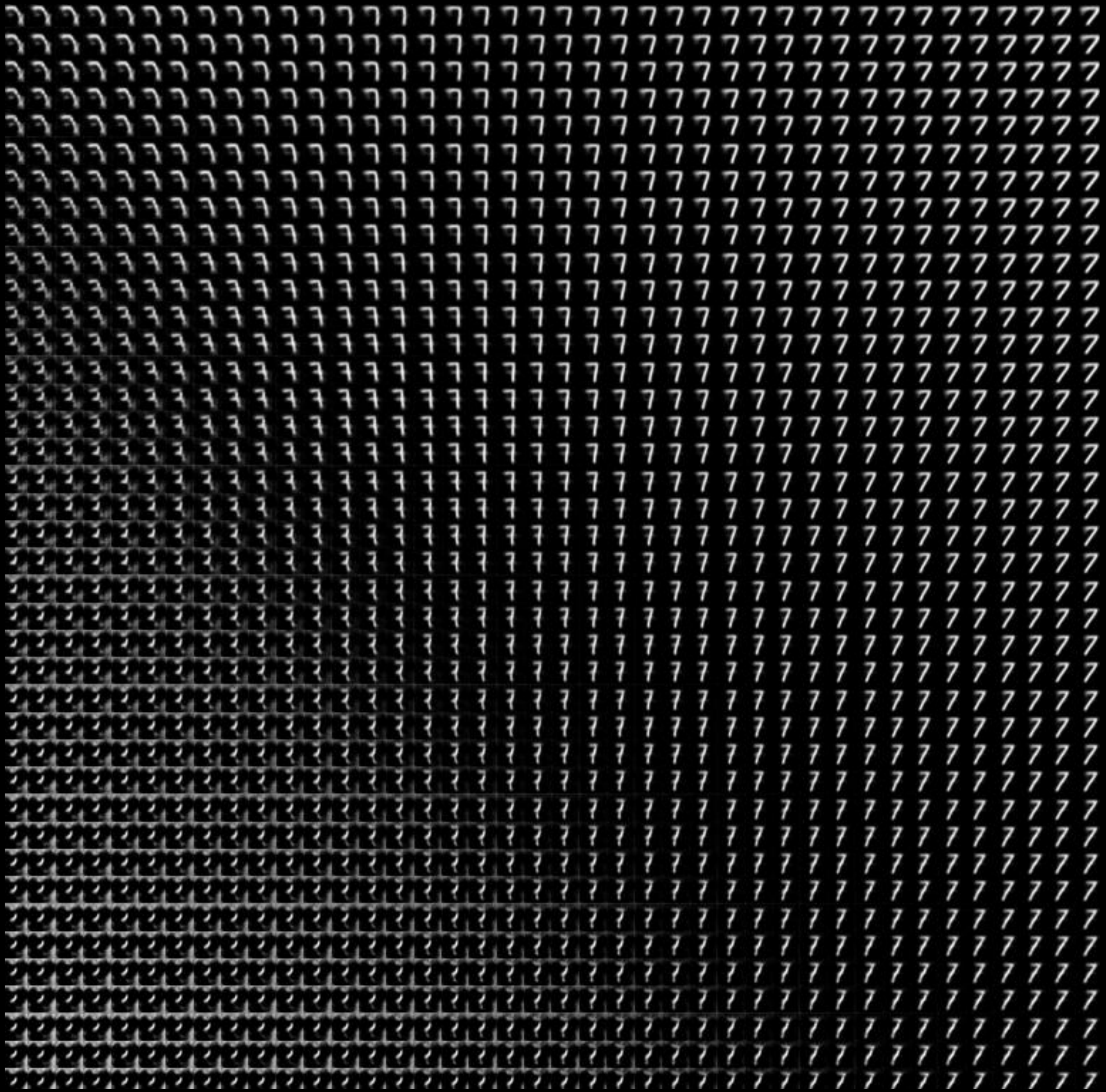
Image space → Latent space



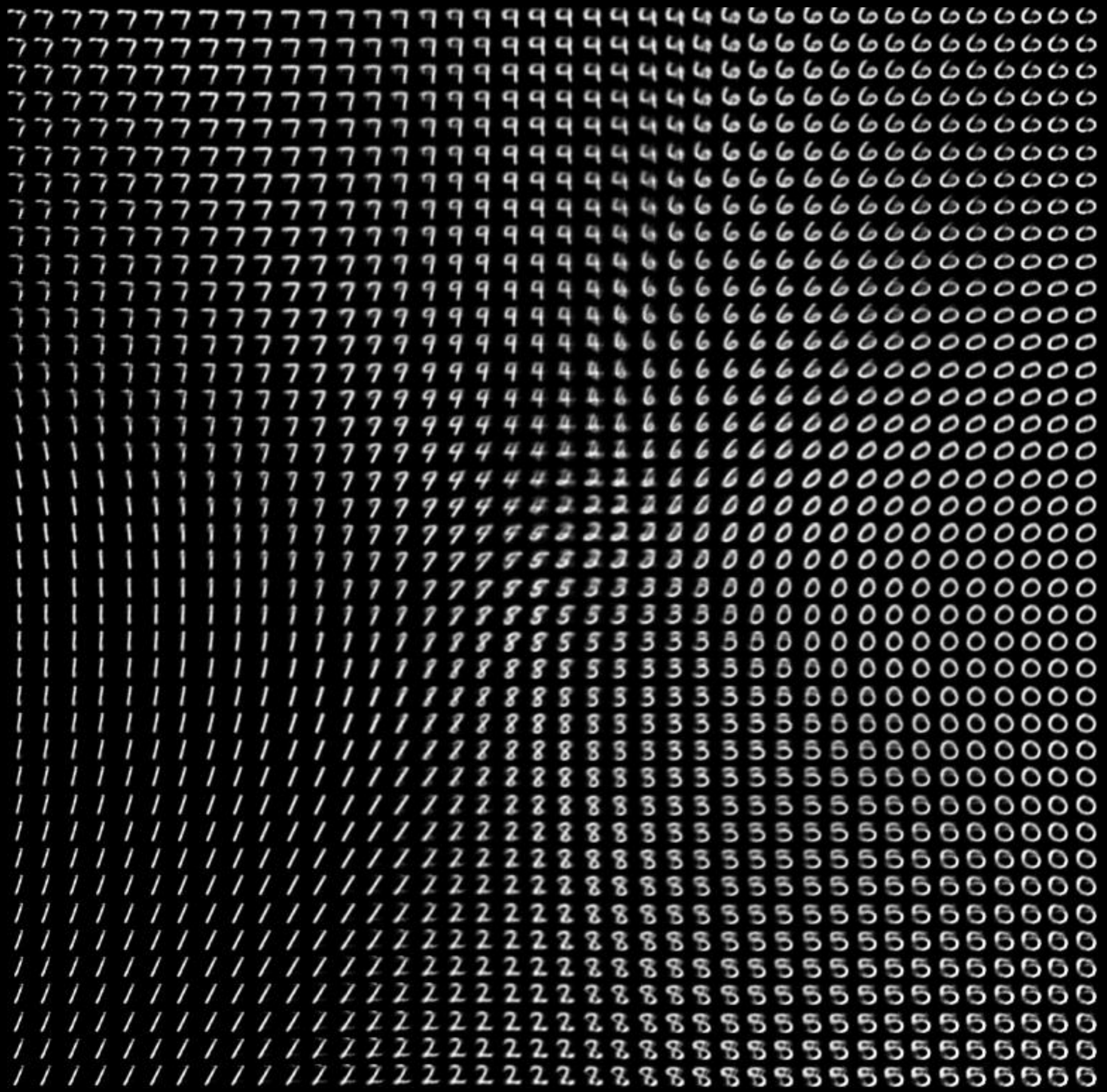
Autoencoder latent representation



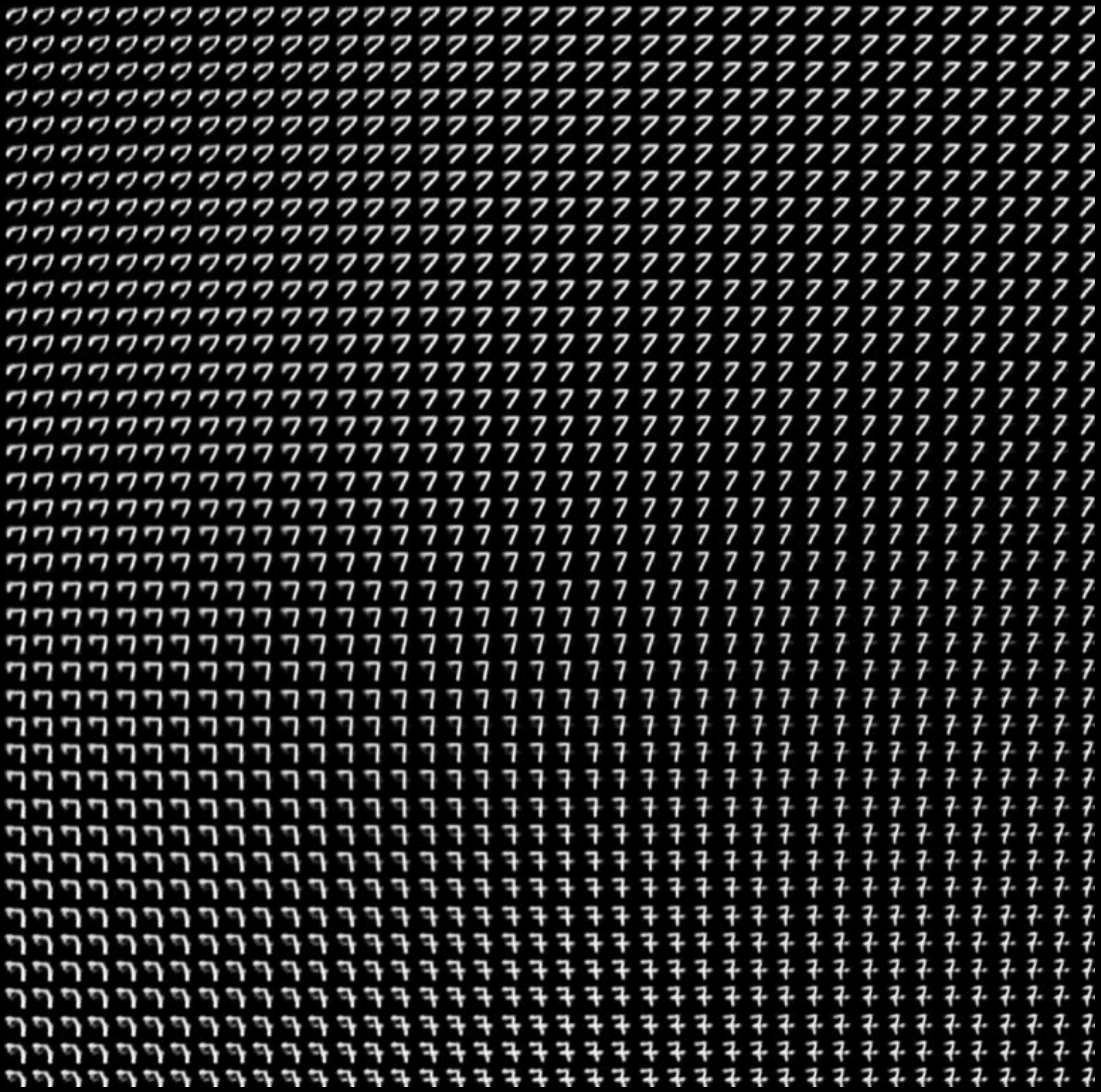
Autoencoder latent representation (digit 7)



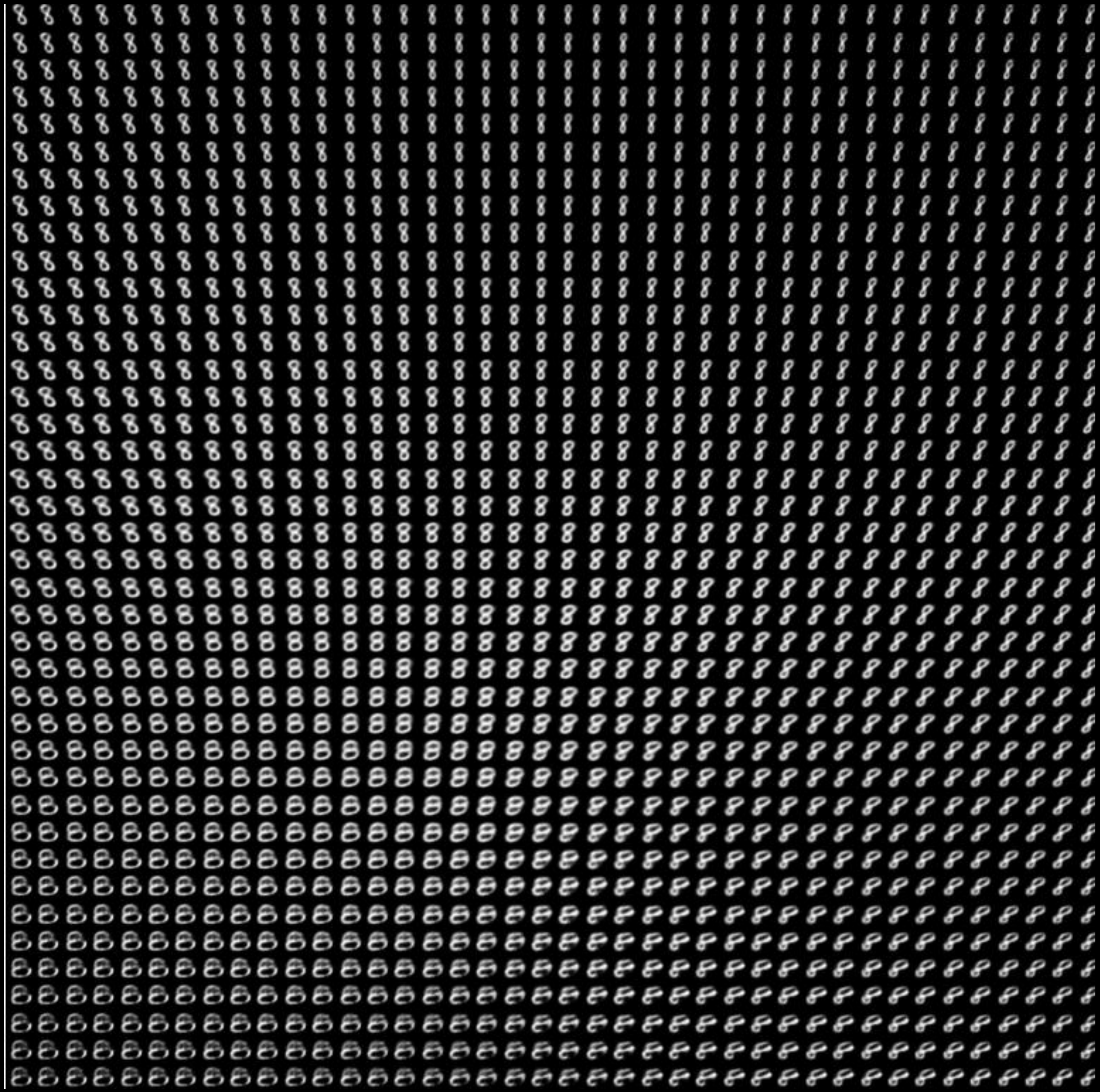
VAE latent representation



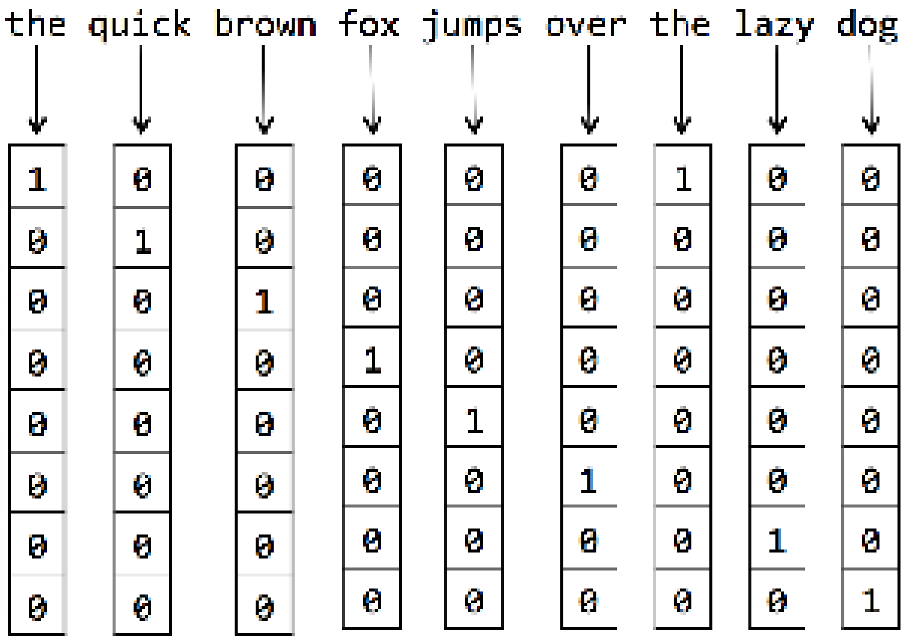
VAE latent representation (digit 7)



VAE latent representation (digit 8)



Word Embeddings



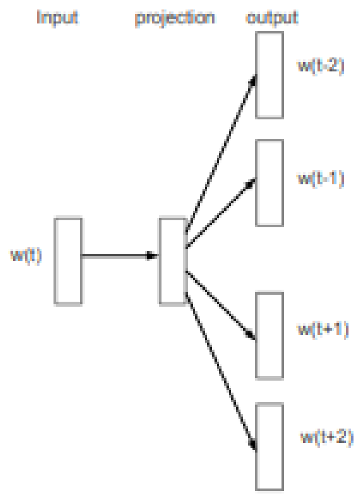
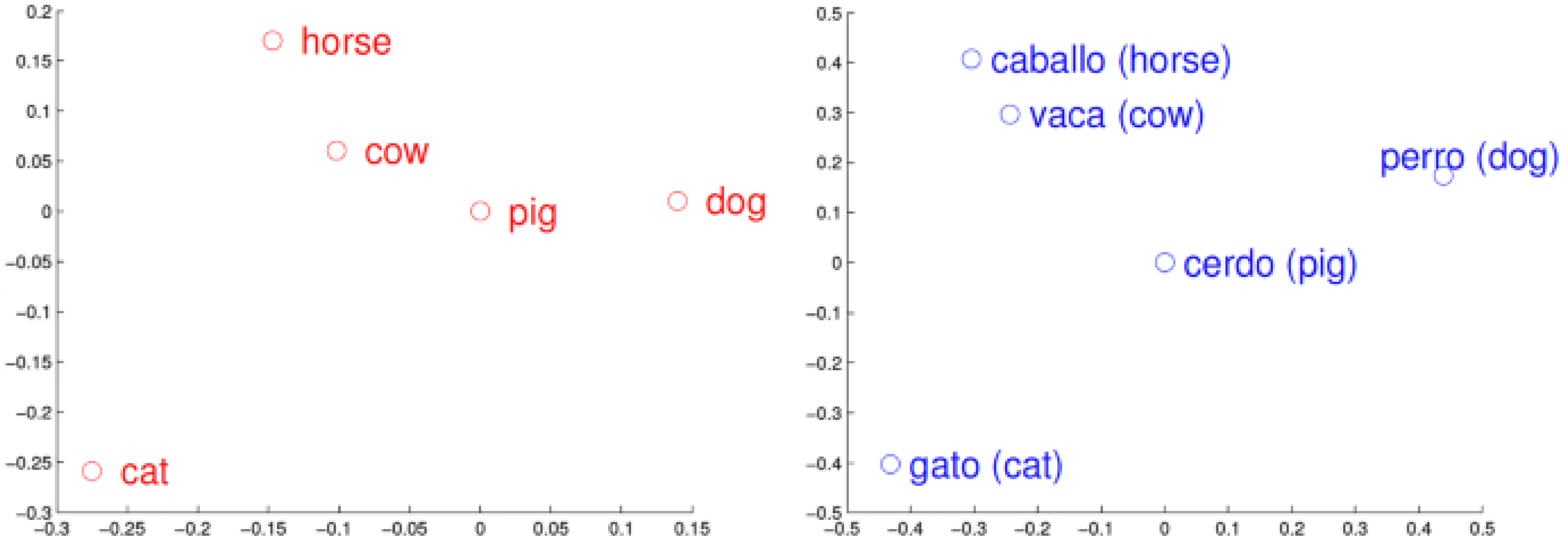


Figure 1: The Skip-gram model architecture. The training objective is to learn word vector representations that are good at predicting the nearby words.

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

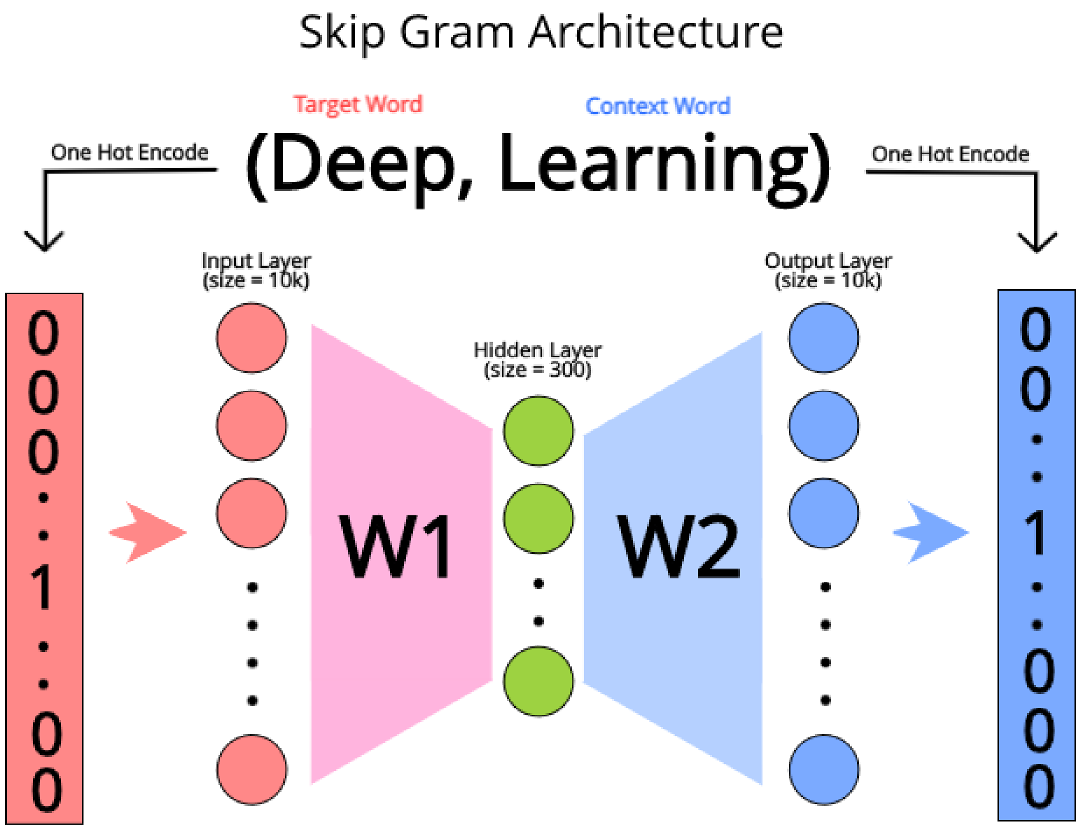
https://proceedings.neurips.cc/paper_files/paper/2013/file/9aa42b31882ec039965f3c4923ce901b-Paper.pdf

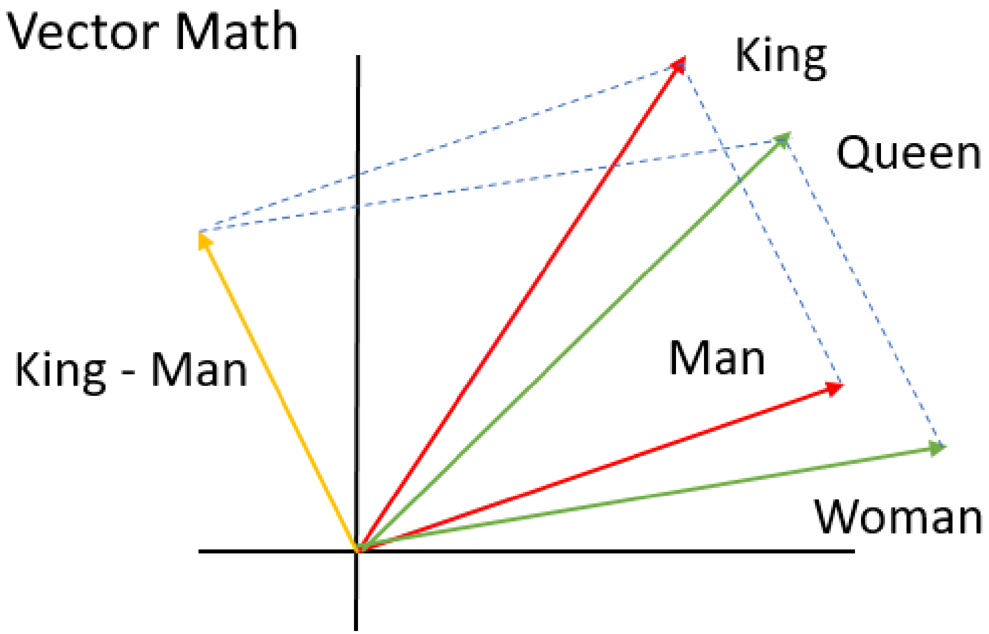
Word Embeddings



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

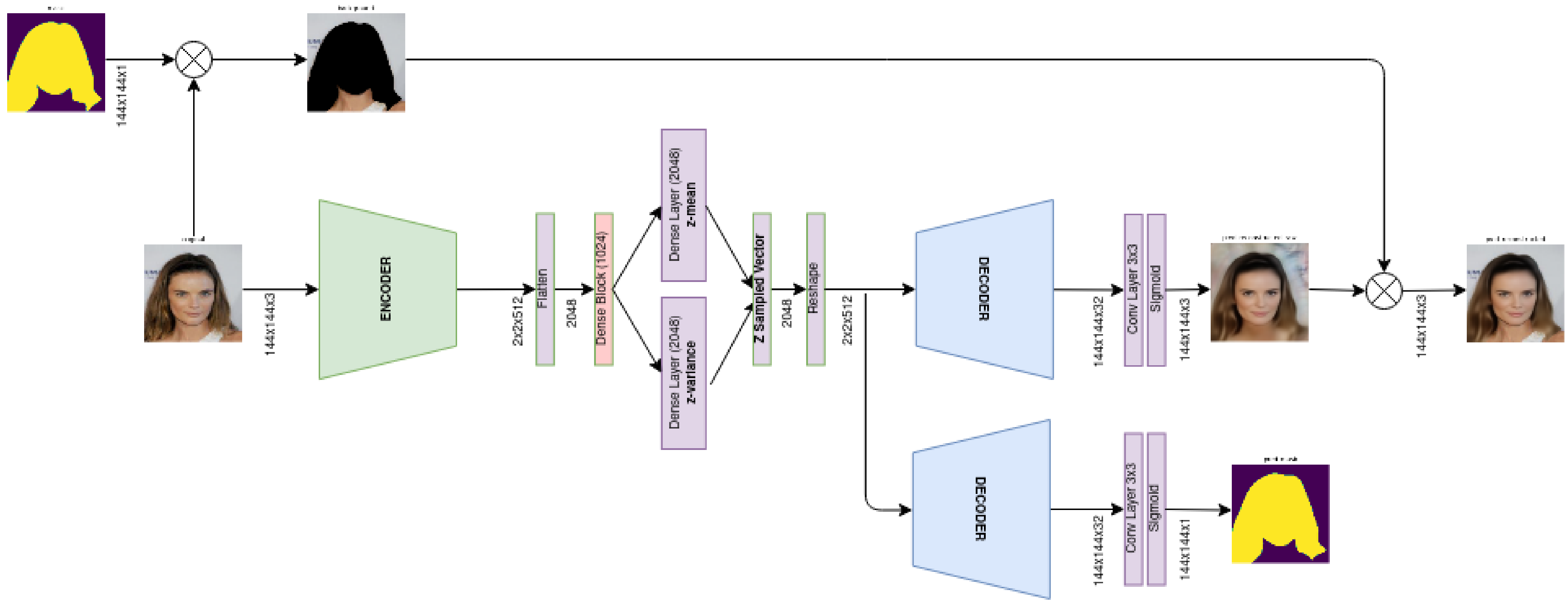
Word Vectors





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

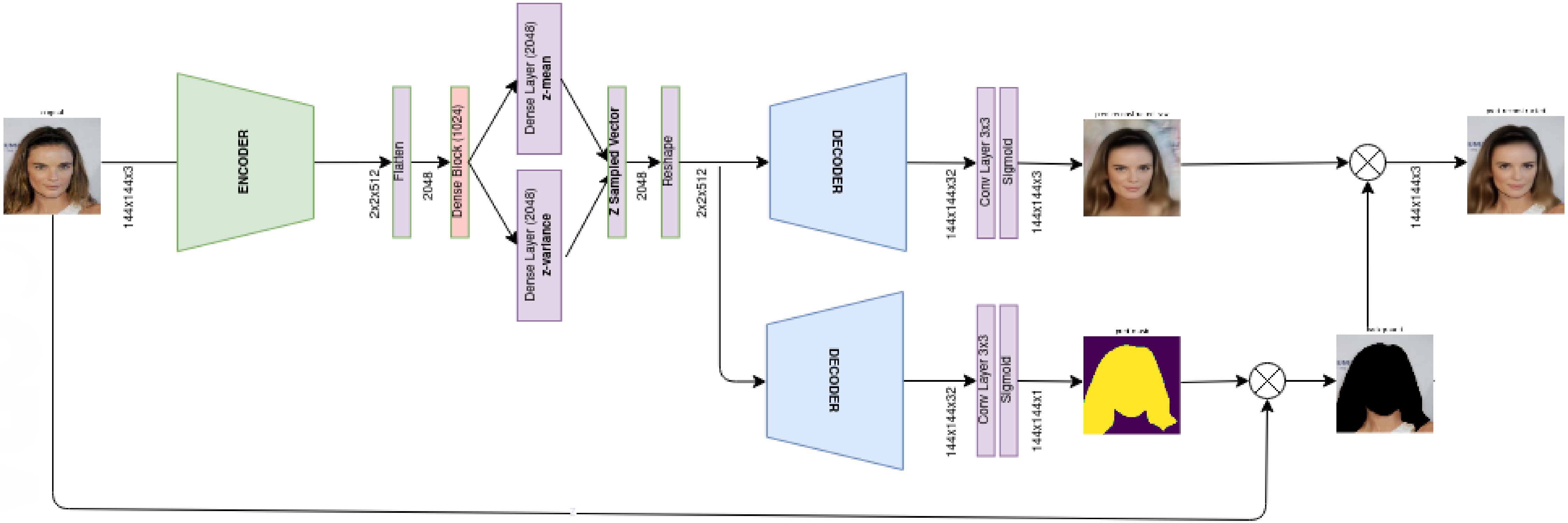
Making Deep Fakes with VAE



During training, the labels of the face masks are used too, which replaces the background of the reconstructed image such that the loss function is applied only over the face pixels.

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

Making Deep Fakes with VAE



In the prediction mode, the background replacement is done by the predicted mask itself, not requiring any extra input but a sample image.

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

Reconstruction

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

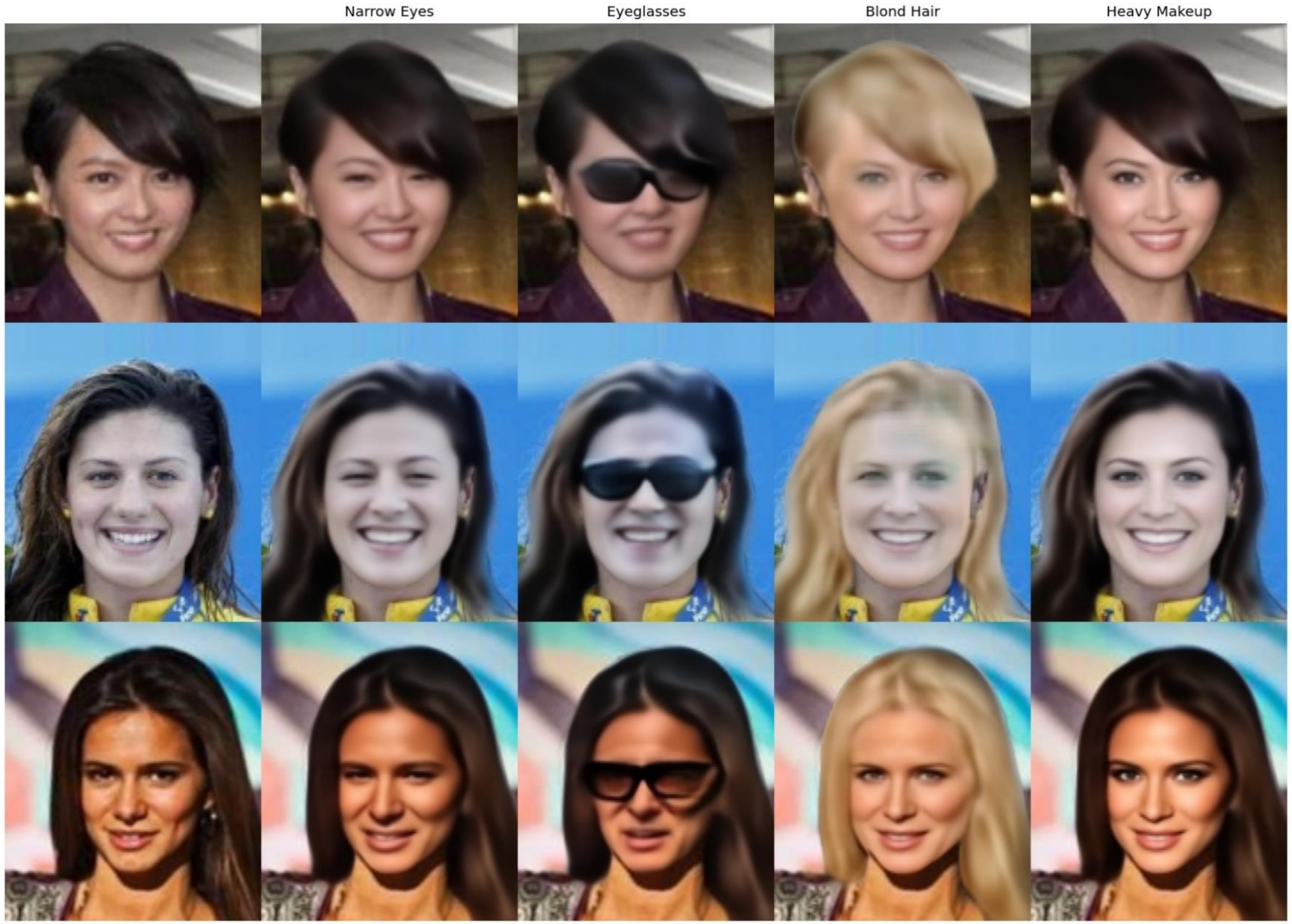
Changing Attributes

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

Changing Attributes

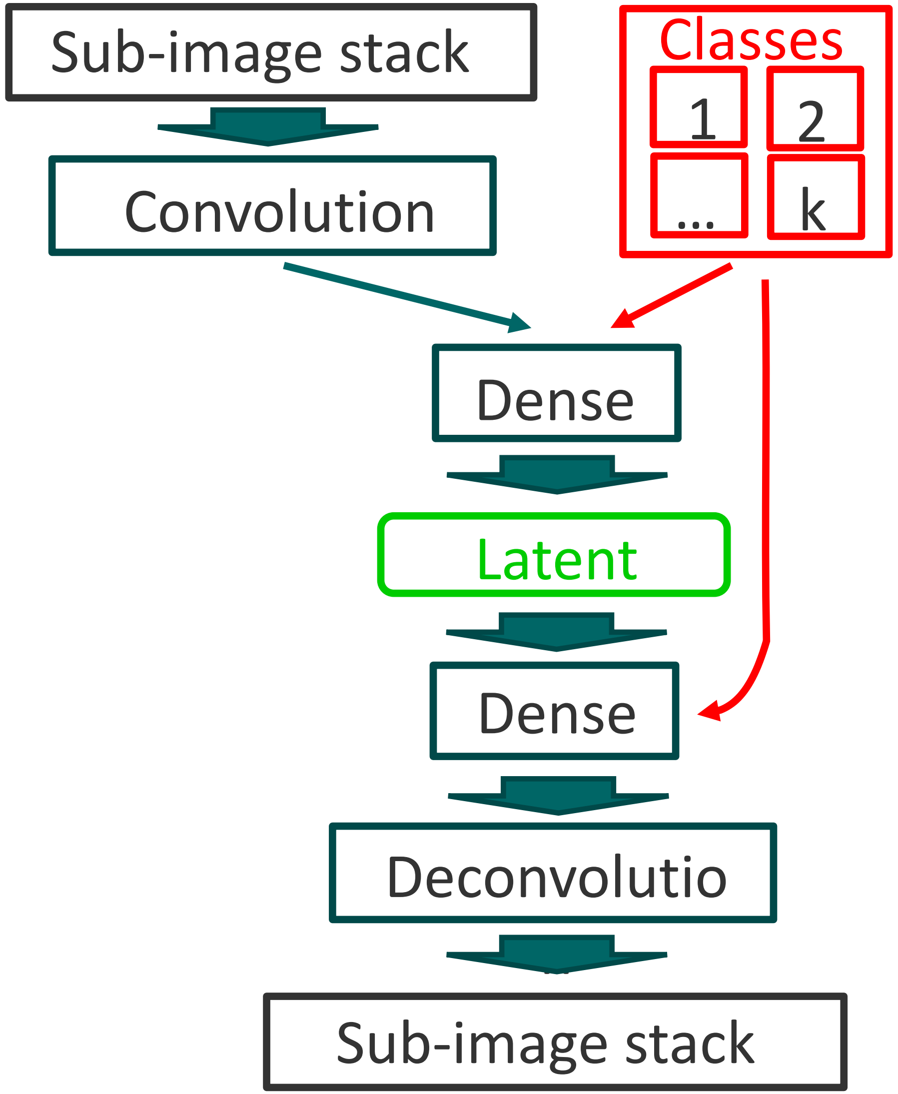
https://rtoledo.me/post/2021-05-31-edit-face-attributes-using-vae/edit-face-attributes-using-vae/

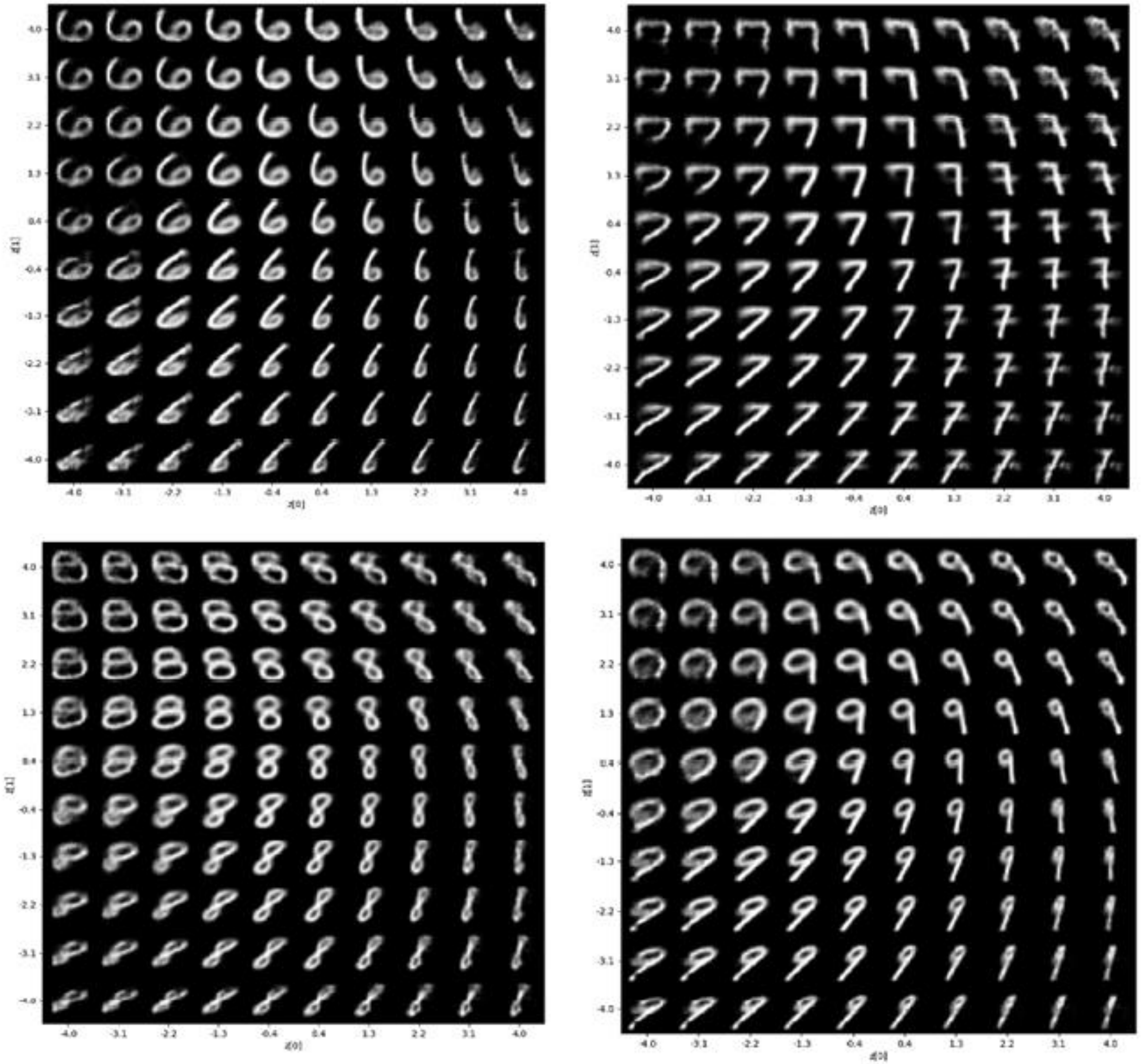
Changing Attributes



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