

Generative models

Anders Krogh

Generative neural networks

Aim:

Model the distribution of your data, $P(x)$, by mapping from a low-dimensional representation space

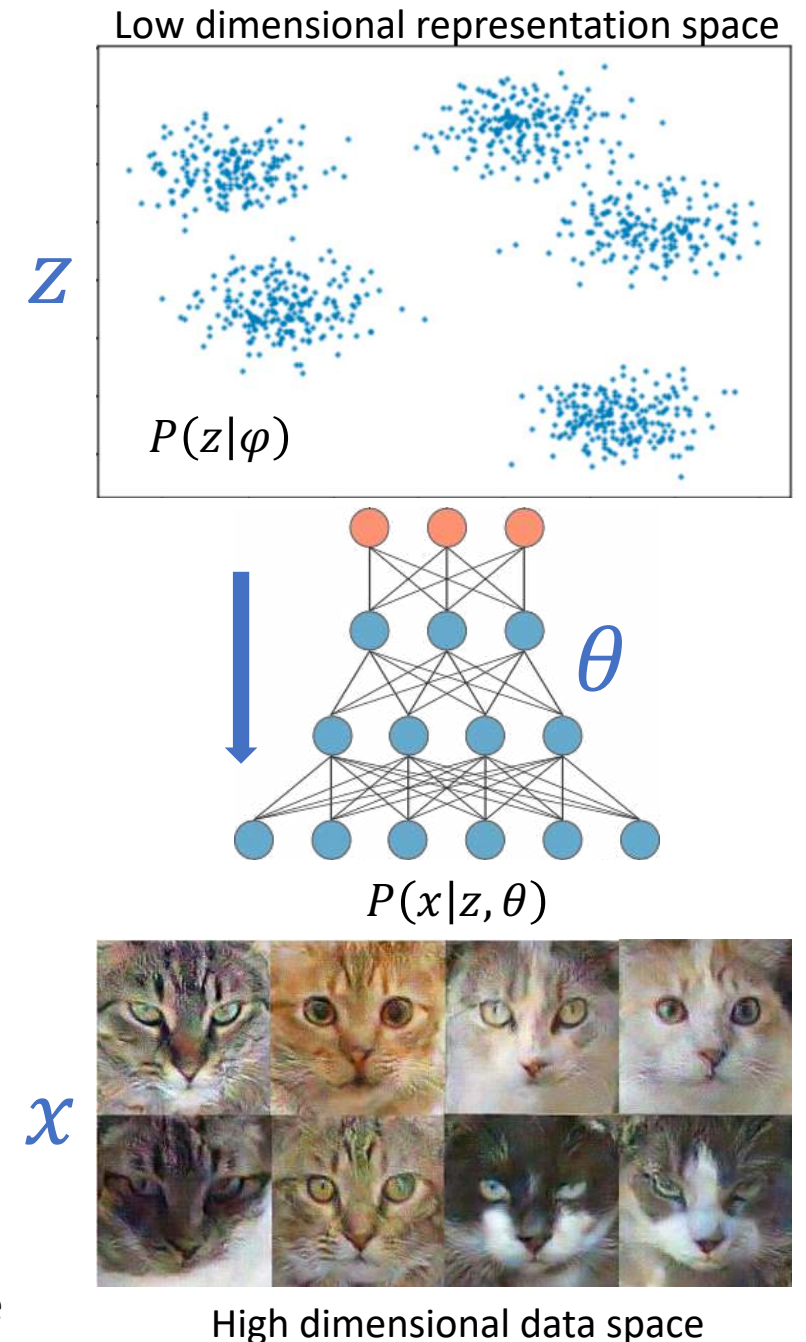
Ingredients:

- Distribution over representations: $P(z|\varphi)$
- Decoder (neural network): $P(x|z, \theta)$ *

Generative:

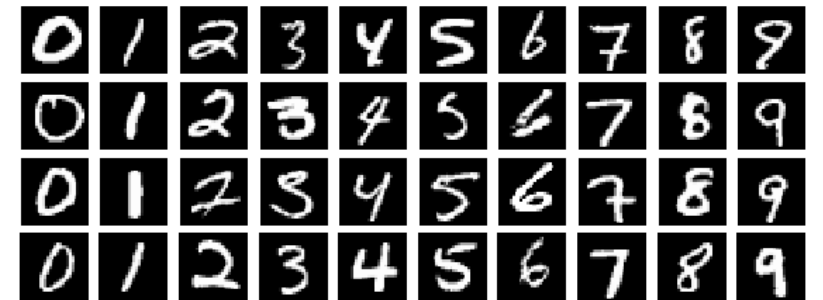
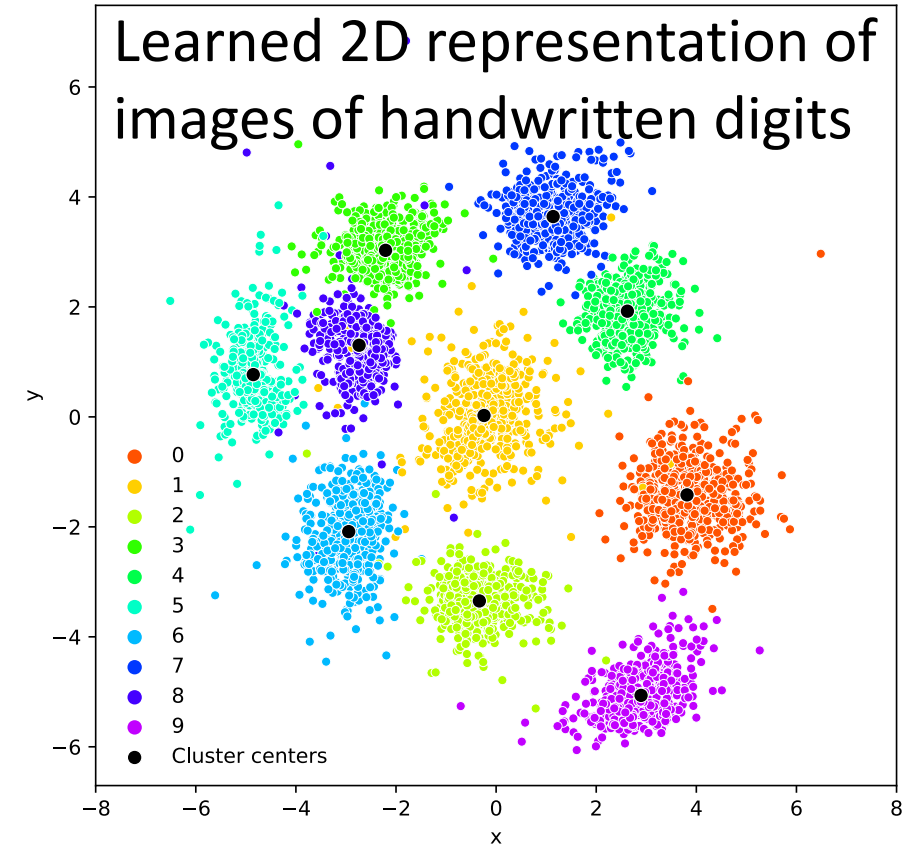
1. Sample random z from $P(z|\varphi)$
2. Sample random x from $P(x|z, \theta)$

*) Neural network may for instance output a mean of a normal with fixed variance



Why generate images of fake cats?

- Model learns a low dimensional representation
- Related to manifold learning
 - UMAP
 - tSNE
- This may be useful for other puposes
 - Representations useful for other tasks
 - Grouping of tumor types, stocks, cats,...
 - Imputation
 - Interpolation



Images from MNIST

Latent space interpolation

Copied from
“Generative Modeling with Variational Auto
Encoder (VAE). Understanding the intuition
behind Variational Autoencoder”
By Fathy Rashad

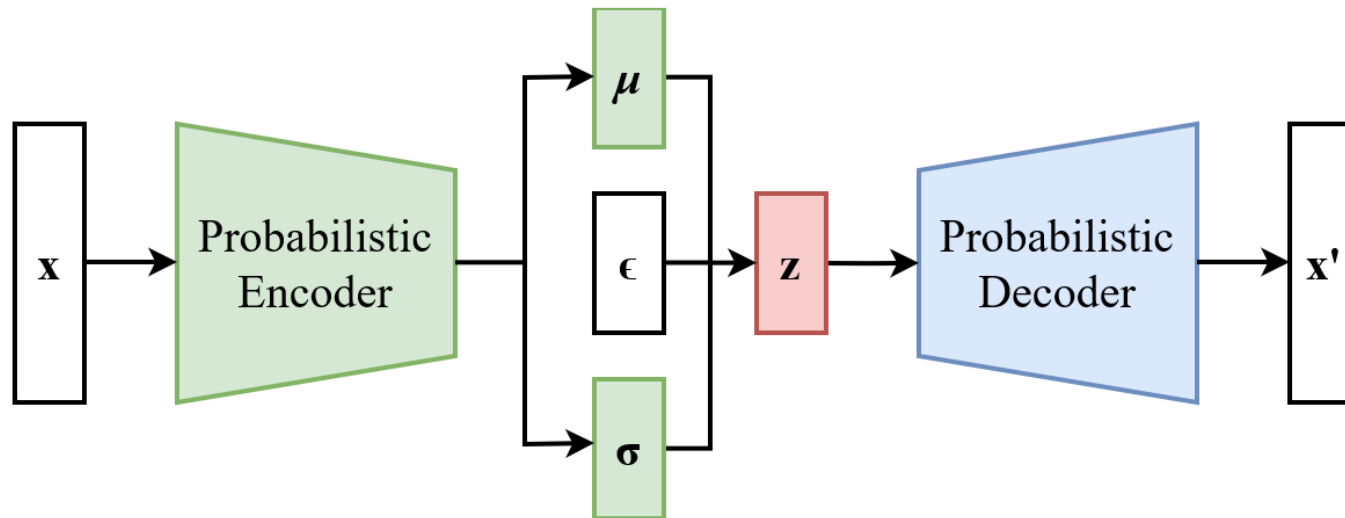
[https://medium.com/vitrox-
publication/generative-modeling-with-variational-
auto-encoder-vae-fc449be9890e](https://medium.com/vitrox-publication/generative-modeling-with-variational-auto-encoder-vae-fc449be9890e)

Similar: App that can predict how you will look
when you get old



VAE: Variational Autoencoder

- Training of generative models require additional “tricks”
- One method is to train an **encoder** together with the **decoder** to give a distribution $P(z|x)$ to sample from
- This is called a variational autoencoder



- **Generative adversarial network (GAN)** is another type

Generative models for gene expression

Bioinformatics, 36(16), 2020, 4415–4422

doi: 10.1093/bioinformatics/btaa293

Advance Access Publication Date: 16 May 2020

Original Paper

OXFORD

Gene expression

scVAE: variational auto-encoders for single-cell gene expression data

Christopher Heje Grønbech^{1,2,3,*}, Maximillian Fornitz Vording³, Pascal N. Timshel⁴, Casper Kaae Sønderby¹, Tune H. Pers⁴ and Ole Winther^{1,2,3}

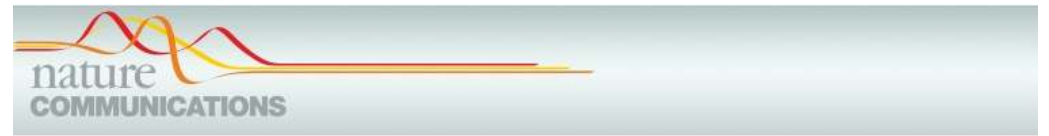
nature|methods

ARTICLES

<https://doi.org/10.1038/s41592-019-0494-8>

scGen predicts single-cell perturbation responses

Mohammad Lotfollahi^{1,2}, F. Alexander Wolf^{1*} and Fabian J. Theis^{1,2,3*}



ARTICLE

<https://doi.org/10.1038/s41467-019-14018-z>

OPEN

Realistic in silico generation and augmentation of single-cell RNA-seq data using generative adversarial networks

Mohamed Marouf^{1,5}, Pierre Machart^{1,5}, Vikas Bansal¹, Christoph Kilian^{1,2}, Daniel S. Magruder^{1,3}, Christian F. Krebs² & Stefan Bonn^{1,4*}

Bioinformatics, 36(11), 2020, 3418–3421

doi: 10.1093/bioinformatics/btaa169

Advance Access Publication Date: 16 March 2020

Original Paper

OXFORD

Gene expression

Interpretable factor models of single-cell RNA-seq via variational autoencoders

Valentine Svensson^{1,*}, Adam Gayoso², Nir Yosef^{2,3,4} and Lior Pachter^{1,5}



ARTICLE

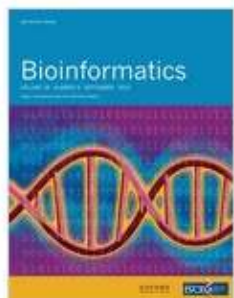
<https://doi.org/10.1038/s41467-021-26017-0>

OPEN

Check for updates

VEGA is an interpretable generative model for inferring biological network activity in single-cell transcriptomics

Lucas Senige¹, Ioannis Anastopoulos¹, Hongxu Ding^{1&2} & Joshua Stuart^{1&3}



Volume 39, Issue 9
September 2023

JOURNAL ARTICLE

The Deep Generative Decoder: MAP estimation of representations improves modelling of single-cell RNA data

Viktoria Schuster, Anders Krogh 

Bioinformatics, Volume 39, Issue 9, September 2023, btad497, <https://doi.org/10.1093/bioinformatics/btad497>

Genome Biology

Method | [Open access](#) | Published: 16 November 2023

N-of-one differential gene expression without control samples using a deep generative model

Iñigo Prada-Luengo, Viktoria Schuster, Yuhu Liang, Thilde Terkelsen, Valentina Sora & Anders Krogh 

Genome Biology **24**, Article number: 263 (2023) | [Cite this article](#)



bioRxiv

THE PREPRINT SERVER FOR BIOLOGY

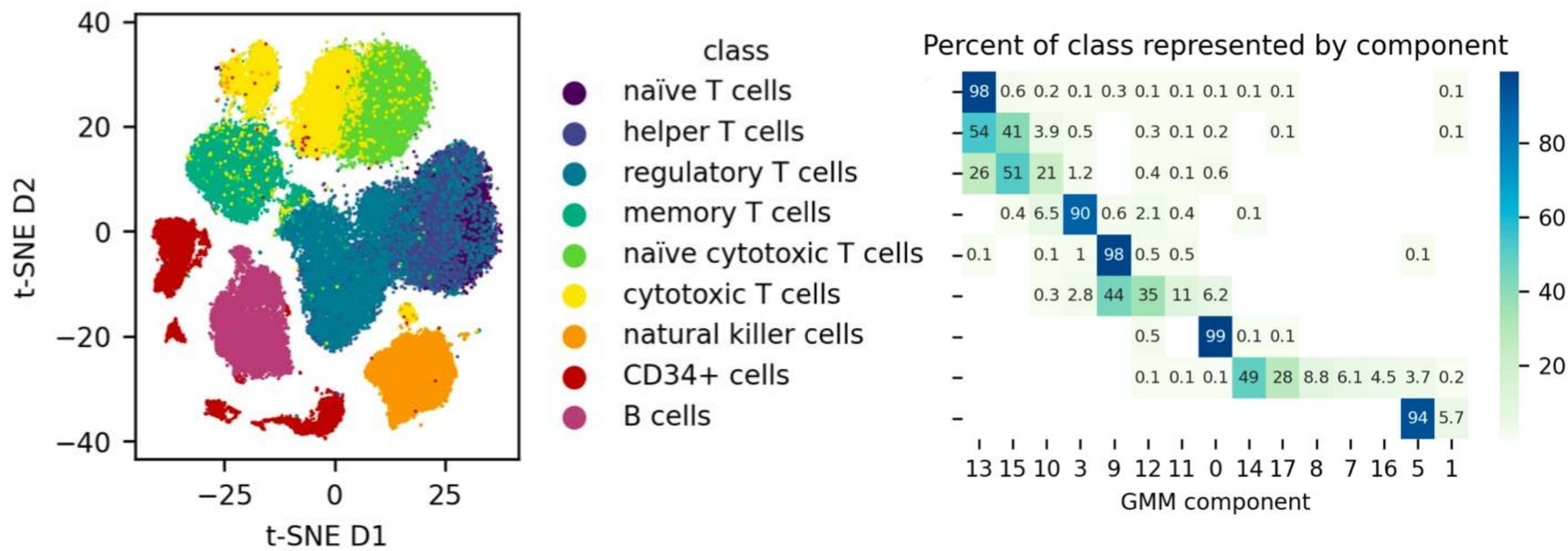
multiDGD: A versatile deep generative model for multi-omics data

 Viktoria Schuster,  Emma Dann,  Anders Krogh,  Sarah A. Teichmann

doi: <https://doi.org/10.1101/2023.08.23.554420>

This article is a preprint and has not been certified by peer review [what does this mean?].

A Deep Generative model applied to a single-cell dataset from peripheral blood mononuclear cells*



Schuster & Krogh: <https://arxiv.org/abs/2110.06672>

*) Data from Zheng, et al. Nat Commun. 2017 Apr;8(1):14049. <https://doi.org/10.1038/ncomms14049>.

Sub-clustering for marker genes

