Generating single-cell microscopy images from the Human Protein Atlas

Using hierarchical VQ-VAE and autoregressive priors to generate images from discrete latent representations.

Jeppe Thagaard

INTRO

- Images visualizing proteins in cells are commonly used for biomedical research, and these cells could hold the key for the next breakthrough in medicine.
- Human cells express different spatio-temporal distributions of proteins.
- Can we learn to generate biological plausible images?

METHODS

- Discrete latent representation via hierarchical vector quantized variational autoencoder (VQ-VAE-2) [1,2].

$$\mathcal{L}(\mathbf{x}, D(\mathbf{e})) = ||\mathbf{x} - D(\mathbf{e})||_2^2 + ||sg[E(\mathbf{x})] - \mathbf{e}||_2^2 + \beta||sg[\mathbf{e}] - E(\mathbf{x})||_2^2$$

- Double latent priors via autoregressive PixelSNAIL [3] with self-attention trained individually.

RESULTS

- VQ-VAE-2 can compress microscopy images to discrete latent representations and reconstruct images well but still blurry.

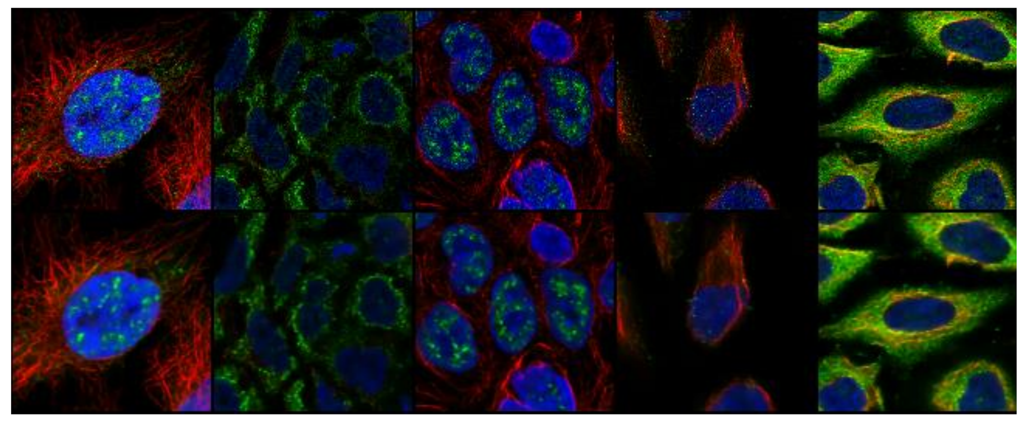

Figure 1: Reconstructions from VQ-VAE-2. Top: Original. Bottom: Reconstruction.

- The autoregressive generative network learns a strong prior over the latent codes.
- Generated samples are shown in figure 2.

FUTURE DIRECTIONS

- Use 256x256 or higher resolution as input to avoid too low latent dimensions (32x32,16x16).
- Extend class conditional prior as in original paper to control image generation based on label.

DTU

VISIOPHARM®

VQ-VAE with autoregressive priors can generate **single-cell microscopy images** with protein localization differences.

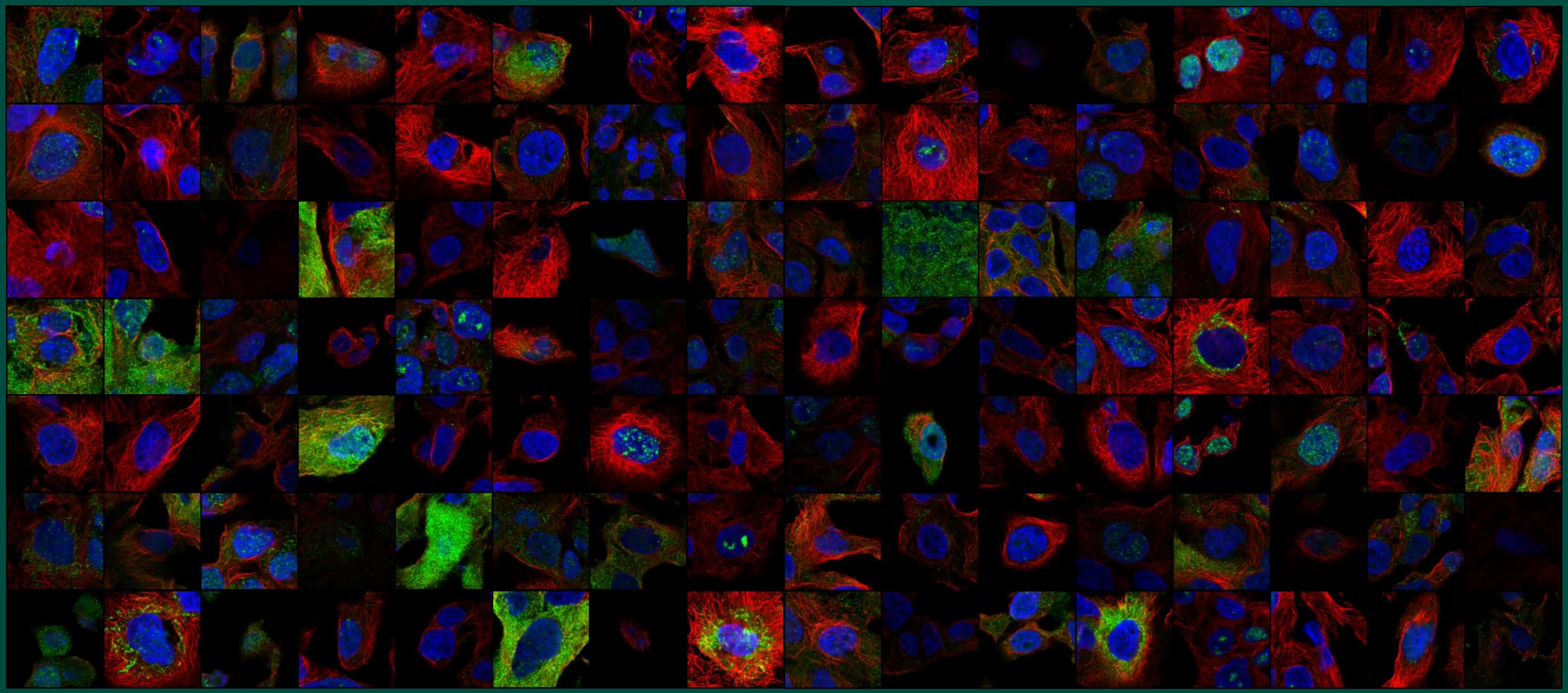

Figure 2: Representative generated samples.

DATA:

- 157K single-cell crops from Human Protein Atlas images using semi-automatic screening using cell-inpainting [5] and UMAP
- 24 unique protein localization labels

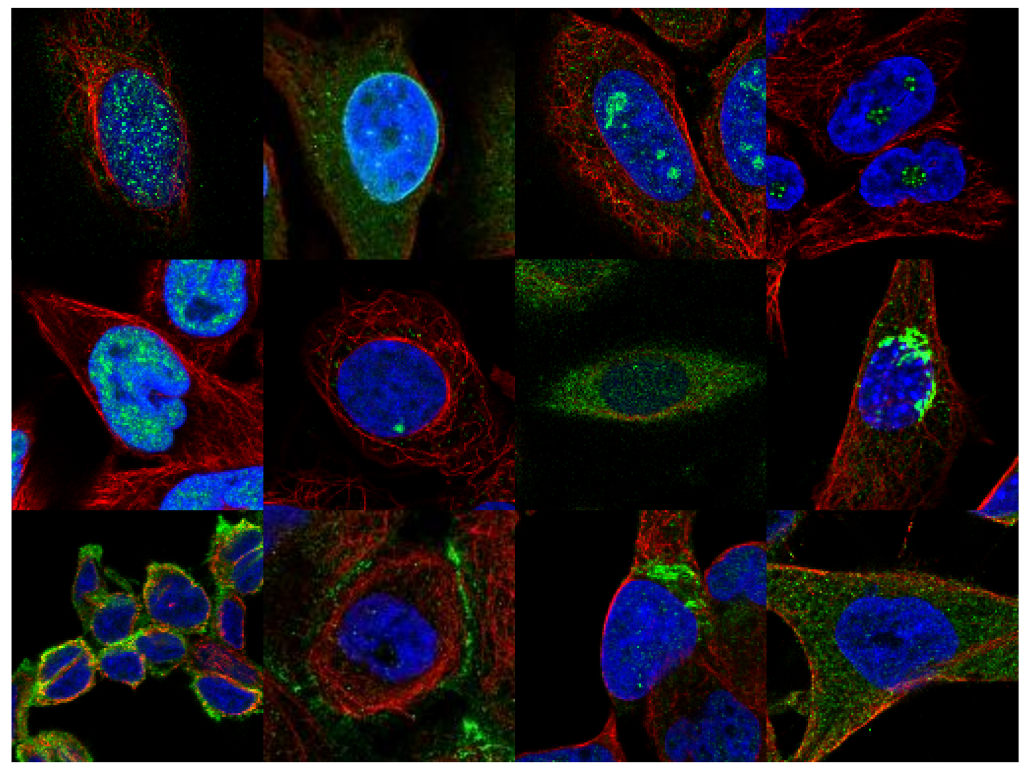

Figure 3: Examples of original images. From top left: Nucleoplasm, Nuclear membrane, Nucleoli, Nucleoli fibrillar center, Nuclear speckles, Endoplasmic reticulum, Golgi apparatus, Plasma membrane, Cell junctions, Mitochondria, Cytosol.

MODEL DETAILS:

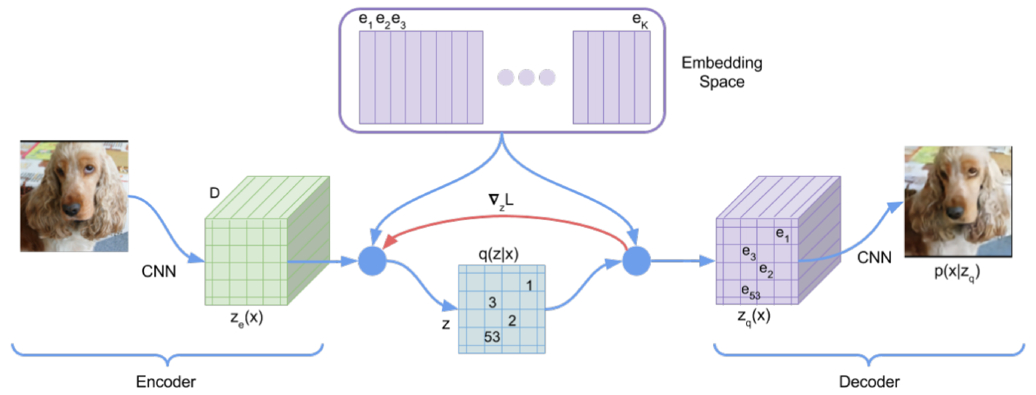

Figure 4: VQ-VAE encodes image to discrete distances to codebook (K=512). Figure from [2]

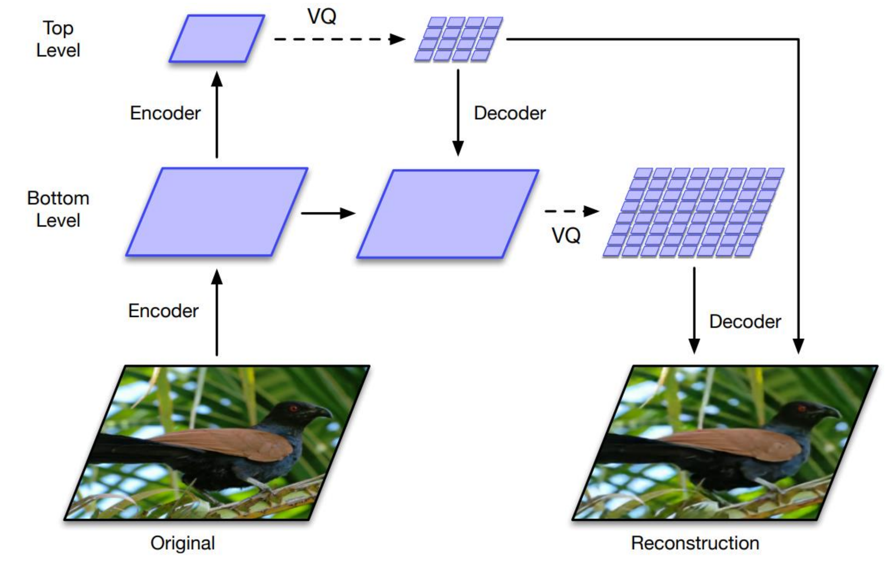

Figure 5: Encoder and decoder training. VQ-VAE-2 use two hierarchical encoding levels e_{top} and e_{bottom}, where the latter is conditioned on e_{top}. Figure from [1]

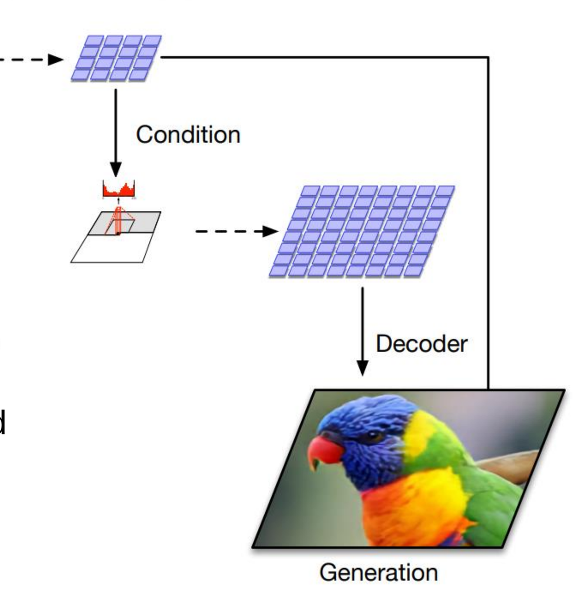

Figure 6: Image generation. Two autoregressive models generates the top and bottom priors. The latter is conditioned on the top encoding. Both discrete encodings are forwarded in the original VQ-VAE decoder to generate the image. Figure from [1].

REFERENCES:

[1] Razavi, A., Oord, A. V. D., & Vinyals, O. (2019). Generating Diverse High-Fidelity Images with VQ-VAE-2. *arXiv preprint arXiv:1906.00446*.
[2] van den Oord, A., & Vinyals, O. (2017). Neural discrete representation learning. In *Advances in Neural Information Processing Systems* (pp. 6306-6315).
[3] Chen, X., Mishra, N., Rohaninejad, M., & Abbeel, P. (2017). Pixelsnail: An improved autoregressive generative model. *arXiv preprint arXiv:1712.09763*.
[4] Regev, A., Teichmann, S. A., Lander, E. S., Amit, I., Benoist, C., Birney, E., ... & Clevers, H. (2017). Science forum: the human cell atlas. *Elife*, *6*, e27041.
[5] Lu, A., Kraus, O. Z., Cooper, S., & Moses, A. M. (2018). Learning unsupervised feature representations for single cell microscopy images with paired cell inpainting. bioRxiv, 395954.