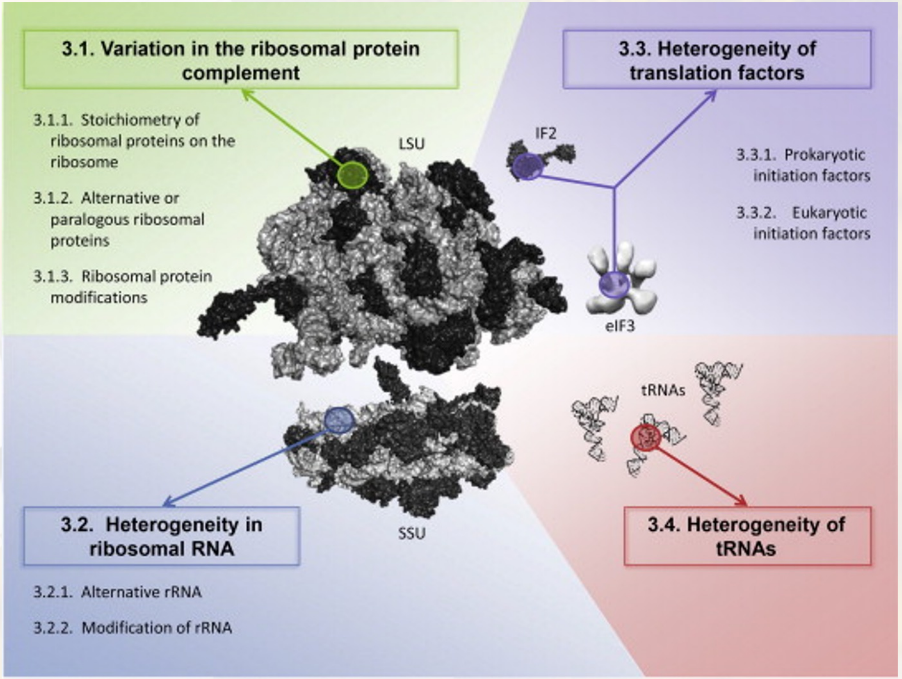
CryoDRGN: Reconstruction of Heterogeneous Cryo-EM Structures Using Neural Networks

Ellen D. Zhong, Tristan Bepler, Bonnie Berger, and Joseph H. Davis

Present by Kevin Cao

Heterogeneity

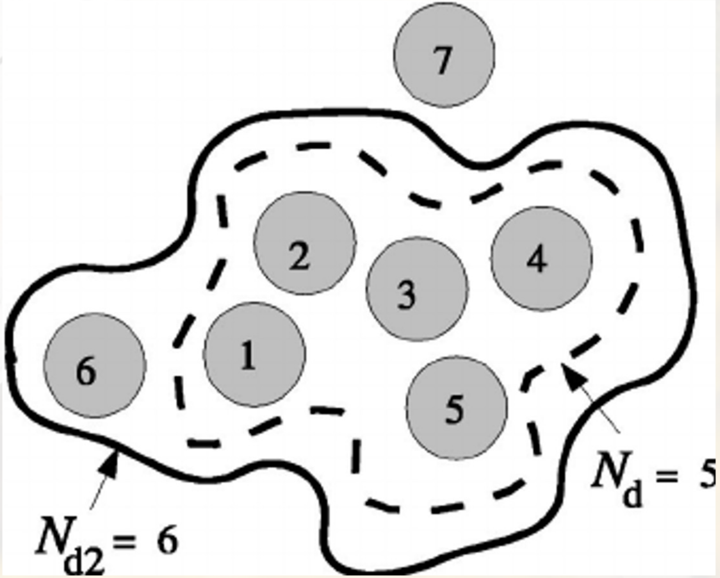
In protein reconstruction field, structure heterogeneity refers to **the protein can have possible states or structures**, and it is natural that these structures changes are **continuous**.



https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.sciencedirect.com%2Fscience%2Farticle%2Fpii%2FS0300908414003952&psig=AOvVaw33u8THC6jbNvn_FUPxMDTH&ust=1695516666020000&source=images&cd=vfe&opi=89978449&ved=0CBEQjhxqFwoTCNi71_fBv4EDFQAAAAAdAAAAABA0

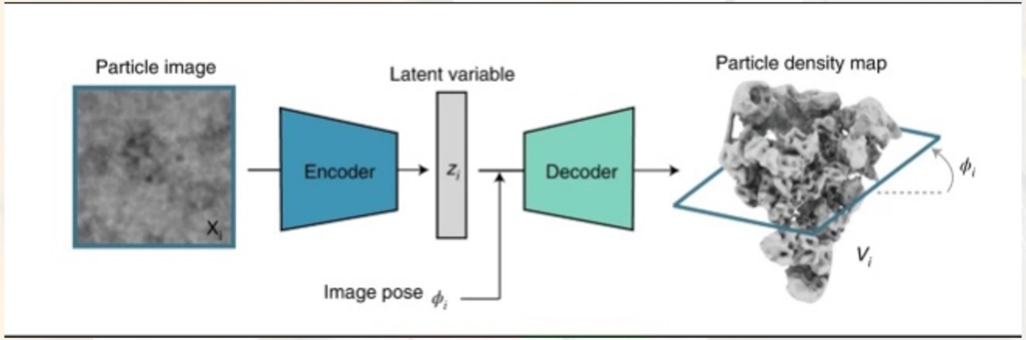
Heterogeneity

Before CryoDRGN, due to the intrinsic flexibility and large conformational changes when proteins performs their function, classical reconstruction methods like **discrete clustering** always fail to capture the full range.



https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.researchgate.net%2Ffigure%2FDiscrete-clusters-defined-according-to-Stillingers-criterion-Ref-8-with-cluster_fig2_7831700&psig=AOvVaw3wGSs-1UizhqxRB19QzBr7&ust=1695518309616000&source=images&cd=vfe&opi=89978449&ved=0CBEQjhxqFwoTCIDE3lbIv4EDFQAAAAAdAAAAABBo

High-level Idea of CryoDRGN



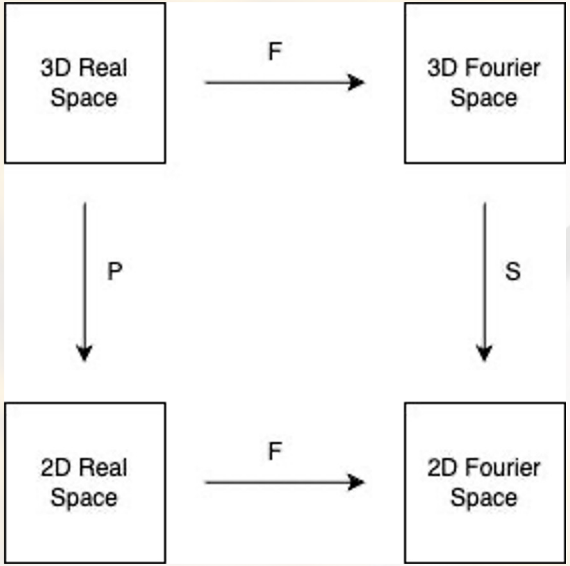
- CryoDRGN is an end-to-end learning framework for a generative model over 3D volume using an image-encoder, volume-decoder neural network structure.
- The inputs are 2D cryo-EM images and the output is volume density map for the protein need reconstruction.

https://www.nature.com/articles/s41592-020-01049-4

Background Information

- Fourier Slice Theorem
- Image Formation Model
- VAE (Variational Autoencoder)

Fourier Slice Theorem



P: Projection
S: Slice
F: Fourier Transformation

FPV = SFV

Image Formation Model - Real Space

$$X(r_x, r_y) = g * \int_{\mathbb{R}} V(R^T \mathbf{r} + t)\, dr_z + noise \qquad \mathbf{r} = (r_x, r_y, r_z)^T$$

g: PSF (Point Spread Function)

V: 3D Volume

R: Rotation Group in R^3

t: In-plane Translation

Image Formation Model - Fourier Space

$$\hat{X}(k_x, k_y) = \hat{g} S(t) A(R) \hat{V}(k_x, k_y) + \epsilon$$

g_hat: CTF (Contrast Transfer Function)

A(R)V_hat(k_x, k_y): a slice in Fourier Space

S(t): phase shift function in Fourier Space

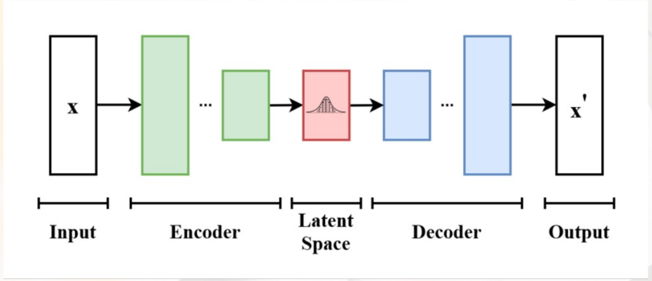
ε: a Gaussian noise

Image Formation Model - Fourier Space

Under this model, the probability of observing an image with specific pose is:

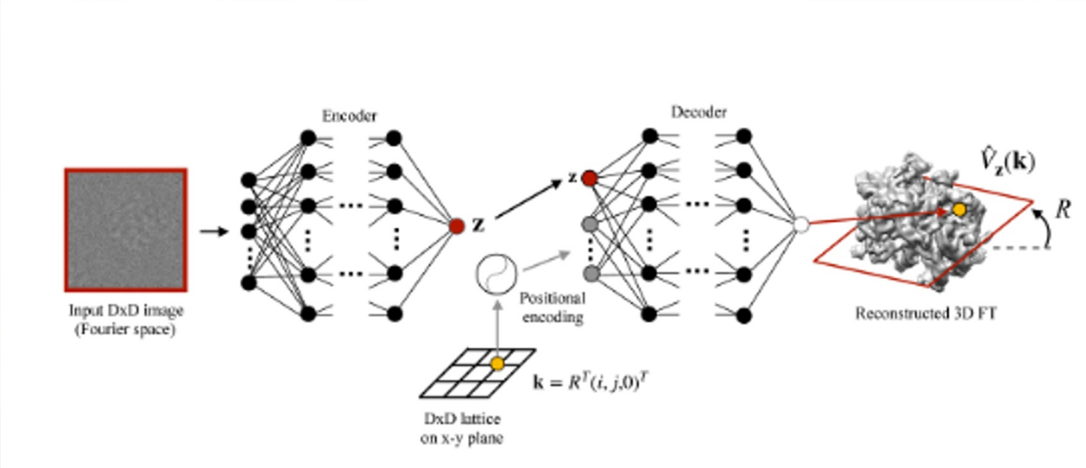
$$p(\hat{X}|\phi, \hat{V}) = p(\hat{X}|R, t, \hat{V}) = \frac{1}{Z} \exp\left(\sum_l \frac{-1}{2\sigma_l^2} \left| \hat{g}_l A_l(R)\hat{V} - S_l(t)\hat{X}_l \right|^2 \right)$$

Variational Autoencoder



- Encoder: map input to low-dimension latent variable *z*, and output mean and variance of *z*.
- Latent Space: a middle layer between Encoder and decoder for sampling from the distribution of *z*.
- Decoder: map each data point generated in latent space to the original input space.
- ELBO (Evidence Lower Bound): the object function to maximize.

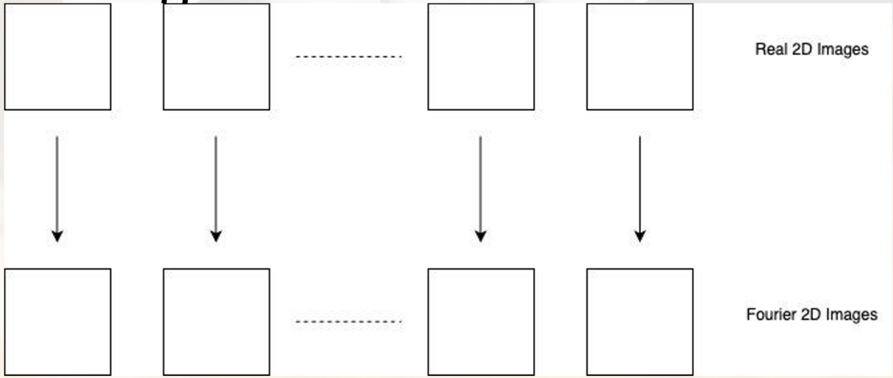
Technical Details within CryoDRGN pipeline:



- Input 2D Image Processing
- Inference model
- Generative model in Decoder
- PE (Positional Encoding)

Input 2D Image Processing:

- The raw noisy cryo-EM images are 2D projections of the real 3D protein.
- By Fourier Slice Theorem, 2D Fourier transformation mapped them to slices into 3D Fourier Space which go through the origin.



Inference model:

- Inference for latent variable z: Follow the standard VAE network to get a probabilistic encoder $q_\xi(z|\hat{X})$

- Inference for pose [R, t]. Perform a global search in

$$SO(3) \times \mathbb{R}^2$$

for the maximum likelihood pose.

Generative model in Decoder:

$$\log p(\hat{X}|R, t, z) = \log p(\hat{X}'|R, z) = \sum \log p_\theta(\hat{V}|R^T c_0^{(i)}, z)$$

$$\hat{X}' = S(-t)\hat{X}$$

X'_hat: the centered image, which eliminates the effect of in-plane translation.

C_0: a vector of 3D coordinates of a fixed lattice spanning on the x-y plane on the x-y plane to represent the unoriented coordinates of an image's pixels

Positional Encoding:

$$pe^{(2i)}(k_j) = sin(k_j D\pi(2/D)^{2i/D}),\ i = 1, ..., D/2; k_j \in k$$

$$pe^{(2i+1)}(k_j) = cos(k_j D\pi(2/D)^{2i/D}),\ i = 1, ..., D/2; k_j \in k$$

- K denotes frequency in Fourier Space.
- For each frequency, position encoding is needed for better fitting results of the generation model.
- The aim of PE is similar to that in NeRF paper.

Results:

- Unsupervised Homogeneous Reconstruction
- Heterogeneous Reconstruction with Pose Supervision
- Unsupervised Heterogeneous Reconstruction

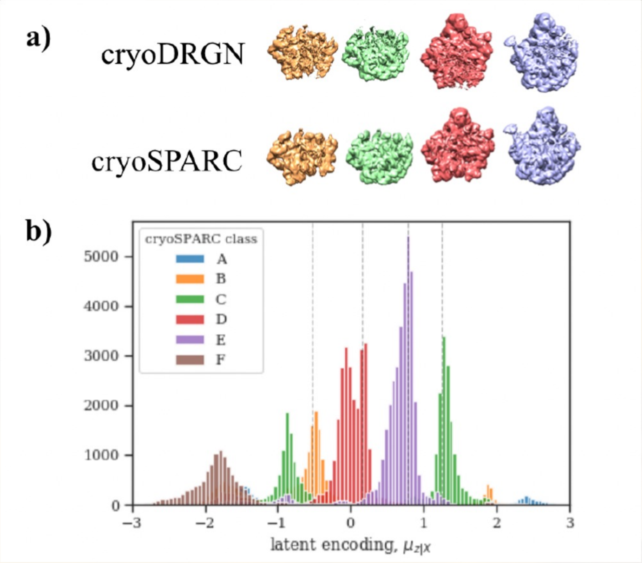
Unsupervised Homogeneous Reconstruction

Method	Dataset	
	No Noise	SNR=0.1
cryoSPARC	0.0009 / 0.47	0.002 / 0.64
cryoDRGN	0.0004 / 0.27	0.003 / 0.38

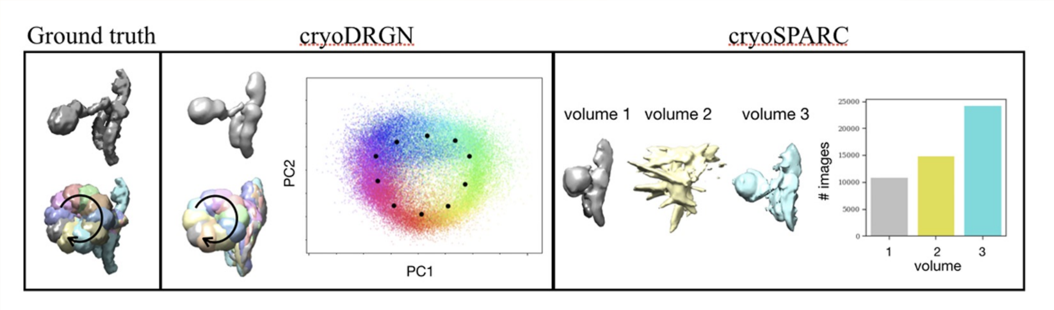
CryoDRGN has a better or nearly equal rotation/translation error compared with cryoSPARC.

Heterogeneous Reconstruction with Pose Supervision

The volumes generated by cryoDRGN aligns well with those generated by cryoSPARC at four peaks of the latent distribution.



Unsupervised Heterogeneous Reconstruction



Compared with 3-class cryoSPARC, the continuous reconstruction results fit the dynamic of ground truth better.

Unsupervised Heterogeneous Reconstruction

Dataset	cryoDRGN	cryoDRGN+tilt	cryoSPARC
Linear 1D motion	2.50(0.62)	2.35(0.36)	3.60(2.27)
Linear 2D motion	4.44(2.50)	2.93(1.02)	6.90(3.77)
Circular 1D motion	4.05(2.40)	2.63(0.74)	4.87(2.17)
Discrete 10 class	4.95(3.16)	2.58(1.00)	5.69(5.15)

CryoDRGN has lower average and standard deviation compared with cryoSPARC.

Conclusion

- CryoDRGN is a novel deep network-based model reconstructing a continuous 3D volume via Fourier space.
- It provides a reasonable method to find out the heterogeneity hidden behind the noisy and randomly posed input 2D image.
- However, reconstructing a single heterogeneous particle still takes a lot of time, future improvements can be implemented.

Questions?

Quizzes

1. What is the advantage of cryoDRGN compared with discrete clustering methods?
   It can fit continuous conformation changes better.

2. Which important theorem we encountered in a previous lecture lays a significant theoretical foundation for this paper?
   Fourier Slice Theorem

Thank you for your Listening