Applying convolutional neural networks to model electron density maps.

Po-Nan Li, Saulo de Oliveira, Henry van den Bedem
{liponan,sauloho,vdbedem}@slac.stanford.edu

Introduction

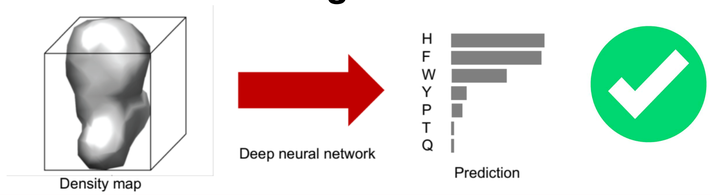
- To achieve high data throughput, we need fast and automated solutions to process electron densities generated at SLAC.

Cryo-EM SSRL LCLS

- **Our goal** is to develop a fully-automated Deep Learning (DL) pipeline to build molecular structure models from electron densities.

Our approach

1. Backbone tracing / box placement.
2. **Amino acid recognition**



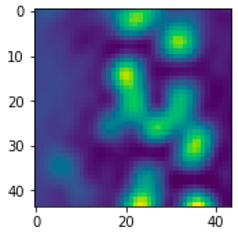
3. Pose prediction (atomic coordinates)

Bibliography:

[1] Xu, Kui, et al. "A²-Net: Molecular Structure Estimation from Cryo-EM Density Volumes." arXiv preprint arXiv:1901.00785 (2019).
[2] He, Kaiming, et al. "Identity mappings in deep residual networks." European conference on computer vision. Springer, Cham, 2016.
[3] Chen, Yu, et al. "Adversarial posenet: A structure-aware convolutional network for human pose estimation." Proceedings of the IEEE International Conference on Computer Vision. 2017.
[4] Image from: https://cvlab.epfl.ch/research/research-surv/research-surv-human-pose-estimation/

Data processing

- We simulated **crystallographic** electron densities for 10,000+ protein structures selected from the PDB.
- Data was simulated in the 1.4-1.6 Å resolution range.
- We then selected 18,831 residues from this data set.



- We created an 11Å x 11Å x 11Å grid around each residue using a step size of 0.25Å (44 x 44 x 44).
- We computed (scaled) electron density value for each grid point.

Training	Val.	Test
16783	1024	1024

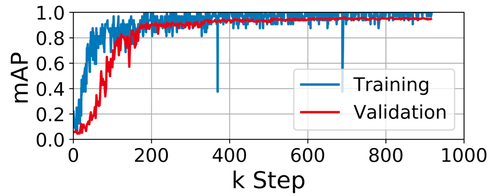
Network architecture

Type	Filters	Size	Output	
input			41x41x41	
conv	128	3x3x3/1	41x41x41	
max		2x2x2/2	20x20x20	
conv	256	3x3x3/1	20x20x20	
max		2x2x2/2	10x10x10	
conv	512	3x3x3/1	10x10x10	
conv	256	1x1x1/1	10x10x10	
conv	512	3x3x3/1	10x10x10	
max		2x2x2/2	5x5x5	
conv	1024	3x3x3/1	5x5x5	x11
conv	512	1x1x1/1	5x5x5	
conv	256	1x1x1/1	5x5x5	
max		5x5x5/5	1x1x1	
conv	20	1x1x1/1	1x1x1	
softmax				

Architecture of our conv. neural network. The highlighted block is residual and repeats 11 times.

- We trained a 3D convolutional neural network (ResNet-11) with 29 learnable layers[1,2].
- Each conv. layer is followed by a ReLU non-linear activation layer.
- Channel-wise batch-normalization and L2 weight decay are used to prevent over-fitting.

Results



Training and validation mean Average Precision (mAP). Validation was monitored in real-time at every 1024 steps.

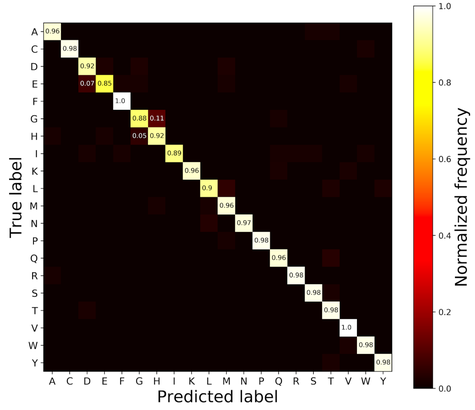
Architecture	mAP
ResNet-11	0.840
ResNet-11 + DA	0.948
Resnet-11 + DA + WD	**0.951**

We used our validation set to optimize hyper-parameters, where:

DA: data augmentation.

{0°, 90°, 180°, 270°} rotations about the *x, y, z* axes during training.

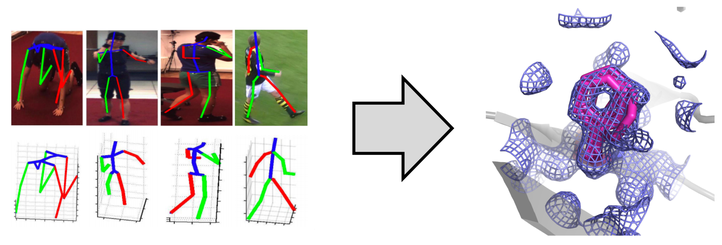
WD: weight decay.



Confusion matrix on validation data set. Frequencies are normalized based on the number of *True* labels per amino acid type. Values lower than 0.05 were omitted.

What's next?

- **Currently:** we are working on pose prediction[3].



- **Next:** trace the backbone into density.
- **Last:** other resolution ranges and Cryo-EM.