CNNs vs. Twin Neural Networks in Identifying X-ray Single-Particle Images in Real-Time

Eric Florin, Chuck Yoon

SLAC NATIONAL ACCELERATOR LABORATORY

Introduction

In single particle imaging (SPI), an x-ray free-electron laser (XFEL) pulse hits the target particle in the experiment chamber and diffracts off of it before the particle is destroyed by Coulomb explosion. Diffraction patterns from single-particles are used to reconstruct the three-dimensional (3D) structure of the particle. However, the nanoparticles can often aggregate together resulting in an unusable "multi-hit" image (**Fig. 1**). These multi-hit images must be discarded to correctly determine the structure.

Researchers have resorted to manual particle picking, unsupervised machine learning techniques such as spectral clustering [2] and more recently supervised machine learning using convolutional neural networks (CNNs) [3,4]. Limitation of these CNN papers are two-fold: 1) it has not been shown whether CNNs can handle sparse photon regime, 2) needs retraining for every new experiment.

The focus of this project were to 1) **study CNN performance in sparse photon regime** and 2) **develop a twin neural network architecture to classify single particle images for LCLS deployment**.

Figure 1: Diffraction from single, double, triple, and quadruple PR772 virus hits at LCLS.

Protein Structures & Simulation

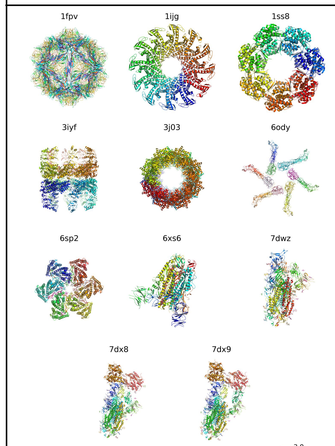

Figure 2: A set of eleven protein structures from Protein Databank (PDB) used in classification. **1)** 1fpv: feline panleukopenia virus (65kDa), **2)** 1ijg: bacteriophage head-tail connector (431kDa), **3)** 1ss8: GroEL (386kDa), **4)** 3iyf: Lidless mm-cpn in open state (890kDa), **5)** 3j03: Lidless mm-cpn in closed state (843kDa), **6)** 6ody: Helicobacter pylori VacA (422kDa), **7)** 6sp2: SERINC from Drosophila melanogaster (354kDa), **8)** 6xs6: SARS-CoV-2 spike (418kDa), **9)** 7dwz: S protein of SARS-CoV-V2 (443kDa), **10)** 7dx8: S protein of SARS-CoV-V2 bound with PD of ACE2 conformation 2 (635kDa), **11)** 7dx9: S protein of SARS-CoV-V2 bound with PD of ACE2 conformation 3 (635kDa)

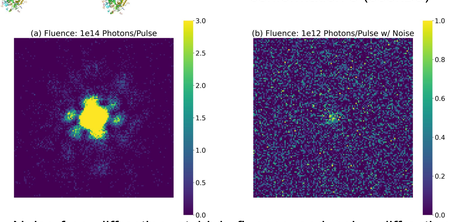

Figure 3: Noise-free diffraction at high fluence and noisy diffraction at low fluence of spike protein SARS-CoV-V2 (**7dx8)**

Simulated diffraction patterns from 11 PDBs (**Fig. 2**) are shown in **Fig. 3**. 220,000 patterns were generated using skopi [1].

Convolutional Neural Networks

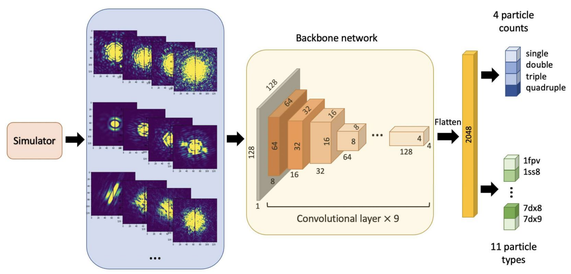

Figure 4: The architecture of the late-branching 10-layer CNN model [7].

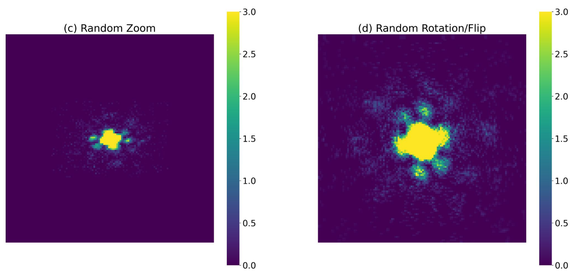

Figure 5: Data augmentations technique to introduce diversity in training data.

Table 1: Benchmark CNN performance on test dataset with no noise applied.

Benchmark			
Model	**Test Accuracy**	**Test Count Accuracy**	**Test Particle Accuracy**
3-layer Multi-output CNN (Late)	85.4	93.4	90.3
5-layer Multi-output CNN (Late)	85.1	93.4	90.4
10-layer Multi-output CNN (Late)	90.4	97.3	91.6
18-layer Multi-output CNN (Late)	84.3	95.4	86.3
Multi-output CNN (Early)	87.8	94.9	91.4
ResNet 18	27.6	65.3	39.3
VGG16	89.3	98.3	89.9

Table 2: CNN performance on test dataset with gaussian noise.

With Data Augmentation (with gaussian noise set to 0.3)			
Model	**Test Accuracy**	**Test Count Accuracy**	**Test Particle Accuracy**
3-layer Multi-output CNN (Late)	48.9	75.4	61.6
5-layer Multi-output CNN (Late)	51.5	77.9	64.1
10-layer Multi-output CNN (Late)	50.5	79.9	60.4
18-layer Multi-output CNN (Late)	5.3	56.6	9.3
Multi-output CNN (Early)	7.0	77.6	8.9
ResNet 18	58.0	78.4	70.4
VGG16	60.7	82.5	70.6

The CNNs [5] use a multi-output architecture which can **predict both protein type and number of particles** in the diffraction pattern (**Fig. 4**). A benchmark was taken to see if model accuracies were high for noise-free images (**Table 1**). Then, the CNNs were tested at lower fluences with fluence jitter. Poisson and Gaussian noise were applied to the images (**Table 2**). Data augmentation was used to significantly increase the diversity of the training dataset and also simulate changes in the detector distance/wavelength (**Fig. 5**). Results are **better than an expert human.**

Twin Neural Network

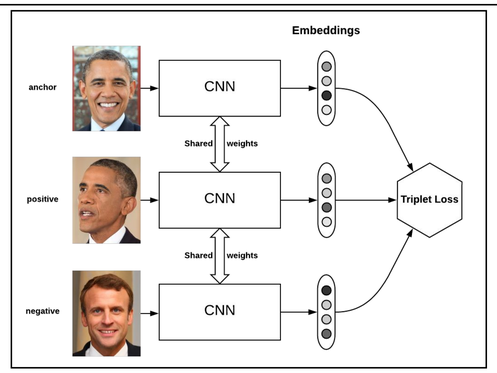

Figure 6: The general architecture of a twin neural network.
https://omoindrot.github.io/triplet-loss

The **twin neural network was developed for one-shot learning** and is used by Google for facial recognition (Fig. 4) [6]. Twin network tries to minimize the triplet loss function (Fig. 5 & 6). Training dataset contained 5,000 diffraction images, for each of the four particle counts (single, double, triple, and quadruple) of each PDB. The network was tested by changing the number of embeddings (by starting at 2 and increasing by powers of 2) and by changing the alpha values (starting at 0.0 and increasing by 0.1). The result was that the twin neural network was able to **classify single-hit images very effectively** when the embedding dimension was 32, no matter the alpha value used. For lower embedding dimensions, however, an optimizing alpha value helped achieve better accuracy (Fig. 7).

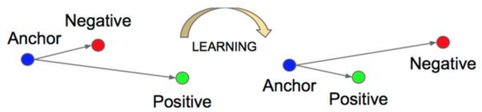

Figure 7: Example of embedding calculations being adjusted using triplet loss in model training.

$$d_p = \sqrt{\frac{\sum_{i=0}^{N-1}(f(a_i) - f(p_i))^2}{N}} \qquad d_n = \sqrt{\frac{\sum_{i=0}^{N-1}(f(a_i) - f(n_i))^2}{N}}$$

$$Loss = max(d_p - d_n + \alpha, 0)$$

Figure 8: Triplet loss function used to train model.

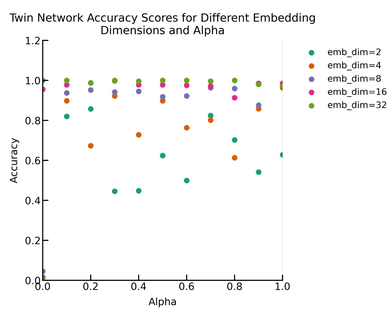

Figure 9: Twin network accuracy scores with gaussian noise set to 0.15

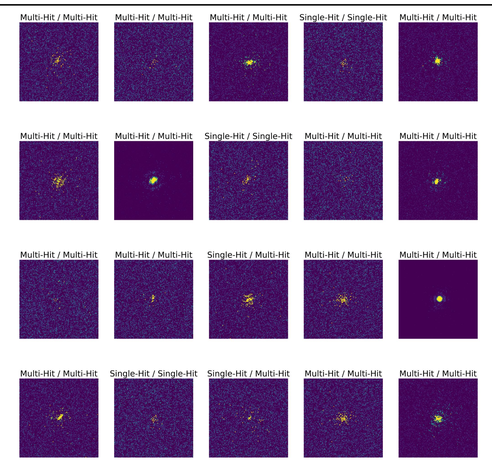

Figure 10: Classification of single-particle images and multi-hit images from twin network, labeled **Prediction / Ground Truth.**

References:

[1] Skopi: Single particle imaging simulation package.
https://github.com/chuckie82/skopi

[2] C.H. Yoon, et al, Unsupervised classification of single-particle X-ray diffraction snapshots by spectral clustering, Optics express, 2011.

[3] Y. Shi, K. Yin, X. Tai, H. DeMirci, A. Hosseinizadeh, B. G. Hogue, H. Li, A. Ourmazd, P. Schwander, I.A. Vartanyants, C. H. Yoon, A. Aquila, and H. Liu. Evaluation of the performance of classification algorithms for XFEL single-particle imaging data. IUCrJ, 6(2):331–340, 2019.

[4] A. Ignatenko, D. Assalauova, S. A. Bobkov, L. Gelisio, A. B. Teslyuk, V. A. Ilyin, and I. A. Vartanyants. Classification of diffraction patterns in single particle imaging experiments performed at x-ray free-electron lasers using a convolutional neural network. Machine Learning: Science and Technology, 2(2):025014, Feb 2021.

[5] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 770–778, 2016.

[6] F. Schroff, D. Kalenichenko, and J. Philbin. Facenet: A Unified Embedding for Face Recognition and Clustering. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, 7-12 June 2015, 815-823.

[7] X. Cai and E. Liu. Classification of X-ray Free Electron Laser Single Particle Images Using Multiple Output Convolutional Neural Networks, 2021.

Conclusions

Key takeaways from this project are as follows. First, the 10-layer multi-output CNN from [7] is somewhat effective at classifying images as single particle or multi-hit when noise was introduced to the diffraction images, but it is not as accurate as when no noise was added to the images. This could become an issue if it is used to classify single particle images from noisy experimental data. Second, the twin neural network can be seen as a viable option for classifying single particle images from multi-hit images, given the right feature vector dimension and alpha value used in model training.

In regards to future work, we hope to train and test the 10-layer multi-output CNN and twin neural network on experimental data taken from a variety of detectors and beam parameters. We hope to train and test both models on more PDBs used in the project so far, especially since proteins have a variety of shapes and sizes. We also hope to explore different types of CNNs and embedding models for the twin neural network to further expand the classification capabilities.

Acknowledgments

We would like to thank Ariana Peck for allowing us to use her cmtip software package to generate the diffraction images used to test the twin neural network, Frédéric Poitevin and Monarin Uervirojnangkoorn for their help in understanding the math behind diffraction resolution, and Shawn Cai and Enci Liu for allowing us to use their datasets and thumbnail image generation notebook to make the thumbnail images necessary for model training/testing.