Ultrafast Bragg Coherent Diffraction Imaging of Epitaxial Thin Films using Deep Complex-valued Neural Networks

Xi Yu[1*], Longlong Wu[2†], Yuewei Lin[1], Jiecheng Diao[3], Jialun Liu[3], Jörg Hallmann[4], Ulrike Boesenberg[4], Wei Lu[4], Johannes Möller[4], Markus Scholz[4], Alexey Zozulya[4], Anders Madsen[4], Tadesse Assefa[2,5], Emil S. Bozin[2], Yue Cao[6], Hoydoo You[6], Dina Sheyfer[6], Stephan Rosenkranz[6], Samuel D. Marks[7], Paul G. Evans[7], David A. Keen[8], Xi He[2,9], Ivan Božović[2], Mark P.M. Dean[2], Shinjae Yoo[1], Ian K. Robinson[2,3‡]

[1]*Computational Science Initiative, Brookhaven National Laboratory, Upton, NY 11973, USA*

[2]*Condensed Matter Physics and Materials Science Department, Brookhaven National Laboratory, Upton, NY 11973, USA*

[3]*London Centre for Nanotechnology, University College London, London WC1H 0AH, UK*

[4]*European X-Ray Free-Electron Laser Facility, Holzkoppel 4, 22869 Schenefeld, Germany*

[5]*Linac Coherent Light Source, SLAC National Accelerator Laboratory, Menlo Park, CA 94025*

[6]*Materials Science Division, Argonne National Laboratory, Lemont IL 60439, USA*

[7]*Materials Science & Engineering, University of Wisconsin, Madison, WI 53706, USA*

[8]*ISIS Facility, Rutherford Appleton Laboratory, Harwell Campus, Didcot, Oxfordshire OX11 0QX, UK*

[9]*Department of Chemistry, Yale University, New Haven CT 06520, USA.*

Correspondence authors' e-mails: [*]xyu1@bnl.gov, [†]lwu@bnl.gov, [‡]irobinson@bnl.gov

SUPPLEMENTARY NOTE 1: Complex-valued CNN architecture

Layer	Output Shape
CConv2d	[64,32,64,64,2]
CLeakReLU	[64,32,64,64,2]
CConv2d	[64,32,64,64,2]
CLeakReLU	[64,32,64,64,2]
CMaxPool2d	[64,32,32,32,2]
CConv2d	[64,64,32,32,2]
CLeakReLU	[64,64,32,32,2]
CConv2d	[64,64,32,32,2]
CLeakReLU	[64,64,32,32,2]
CMaxPool2d	[64,64,16,16,2]
CConv2d	[64,128,16,16,2]
CLeakReLU	[64,128,16,16,2]
CConv2d	[64,128,16,16,2]
CLeakReLU	[64,128,16,16,2]
CMaxPool2d	[64,128,8,8,2]
CConv2d	[64,256,8,8,2]
CLeakReLU	[64,256,8,8,2]
CConvTrans2d	[64,128,16,16,2]
CConv2d	[64,128,16,16,2]
CLeakReLU	[64,128,16,16,2]
CConv2d	[64,128,16,16,2]
CLeakReLU	[64,128,16,16,2]
CConvTrans2d	[64,64,32,32,2]
CConv2d	[64,64,32,32,2]
CLeakReLU	[64,64,32,32,2]
CConv2d	[64,64,32,32,2]
CLeakReLU	[64,64,32,32,2]
CConvTrans2d	[64,32,64,64,2]
CConv2d	[64,1,64,64,2]

Supplementary Table 1 Complex-valued CNN architecture. The size of input is [64,1,64,64,2], where the first dimension represents the batch size, the second dimension represents the coherent diffraction pattern's intensity channel, the third and fourth dimension represent the heigh and width of image, and the last dimension represents the real and imaginary part. We first down sample the input from 64×64 to 8×8 using encoder and then up-sample to original size using decoder.

SUPPLEMENTARY NOTE 2: Complex-valued Convolution and Activation Function

We first introduce the complex convolutional operation. Let $I = I_r + iI_i$, as a complex-valued input, where I_r and I_i are the real and imaginary part of I. Given a complex-valued convolutional filter $k = k_r + ik_i$, the complex-valued convolutional operation on a complex-valued input and a complex-valued filter can be described as follows:

$$I * k = (I_r + iI_i) * (k_r + ik_i) = (I_r * k_r - I_i * k_i) + i(I_r * k_i + I_i * k_r), \tag{1}$$

here, we can split a complex-valued convolutional operation into 2 separate real-valued convolutions (*i.e.*, k_r, k_i), and calculate the output based on the Eq. (1). Similarly, for complex-valued transpose convolution operation, we can just replace k_r and k_i as real-valued transpose convolution operation. The main text Fig. 2(b) illustrates the complex-valued convolutional operation.

In real-valued neural network, rectified linear unit (ReLU) is the most widely used activation function. However, for a complex signal, the real or imaginary component could be negative in some case. Accordingly, we employ Leak ReLU activation function, which has a small slope for negative values, on both real and imaginary part. The definition of Leak ReLU and complex Leak ReLU are shown as follows:

$$LeakyReLU(x) = \begin{cases} x, \ x \geq 0 \\ negative\ slope \ \times x, \ otherwise, \end{cases} \tag{2}$$

$$CLeakyReLU(z) = LeakyReLU(z_r) + i * LeakyReLU(z_i), \tag{3}$$

where $z = z_r + z_i$, and we set $negative\ slope = 0.2$.

The definition of complex-valued Max-pooling becomes:

$$CMaxpool(z) = Maxpool(z_r) + i * Maxpool(z_i). \tag{4}$$

SUPPLEMENTARY NOTE 3: Definition of χ^2 error and SSIM

We calculated the χ^2 error for the modulus of the diffraction pattern in reciprocal space defined as:

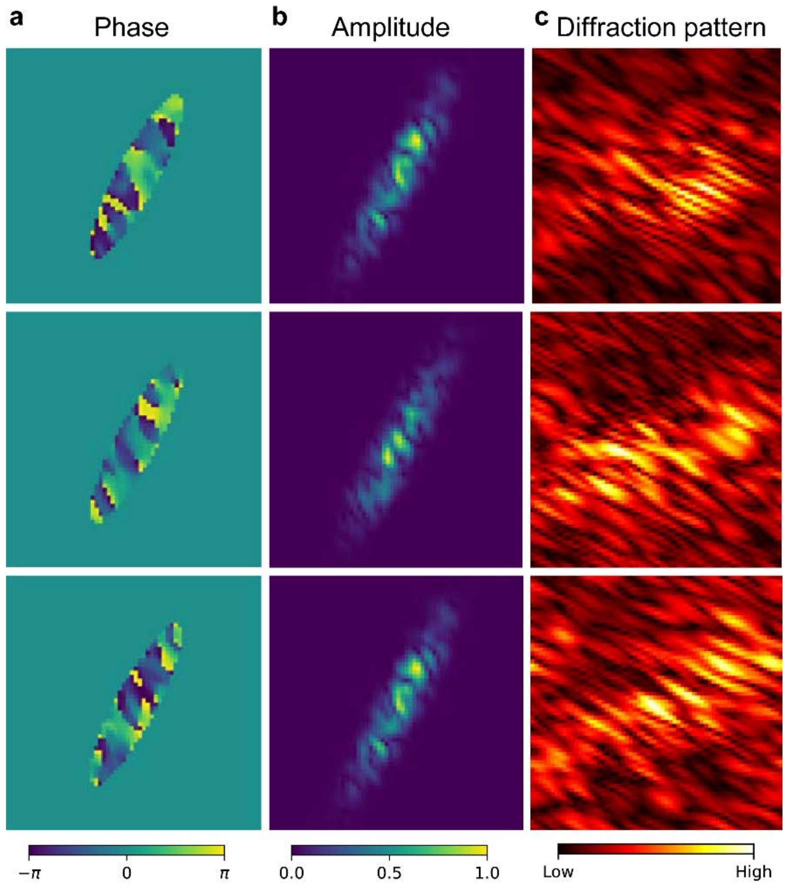
$$\chi^2 = \frac{\sum\left(\sqrt{I_\mathrm{e}} - \sqrt{I_\mathrm{m}}\right)^2}{\sum I_\mathrm{m}}, \tag{5}$$

where I_e is the reconstructed X-ray diffraction intensity and I_m is the true or experimental diffraction intensity.

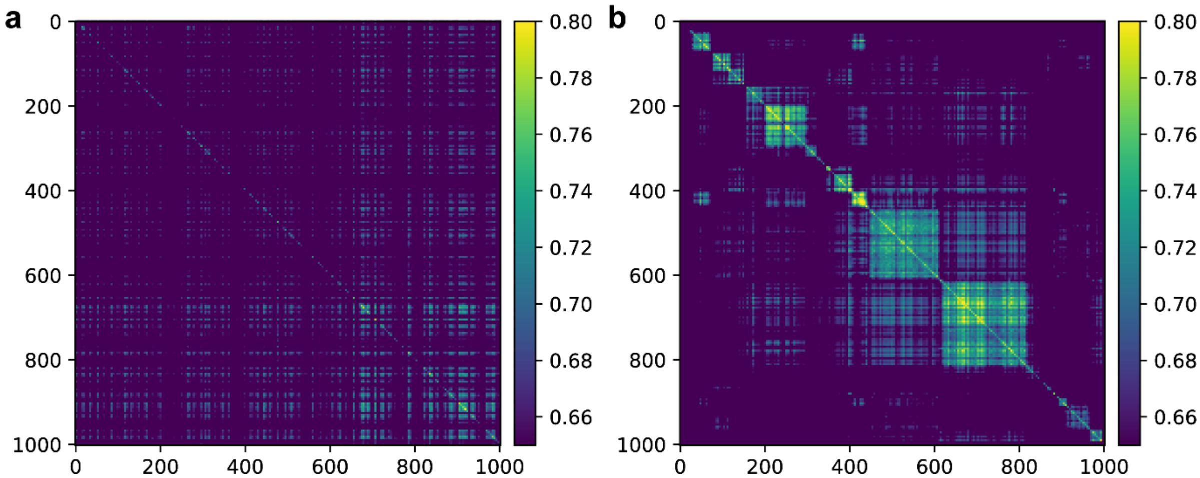
The SSIM index between image arrays $\boldsymbol{x}$ and $\boldsymbol{y}$ is defined:

$$SSIM = \frac{\left(2\mu_x\mu_y + c_1\right)\left(2\sigma_{xy} + c_2\right)}{\left(\mu_x^2 + \mu_y^2 + c_1\right)\left(\sigma_x^2 + \sigma_y^2 + c_2\right)}, \tag{6}$$

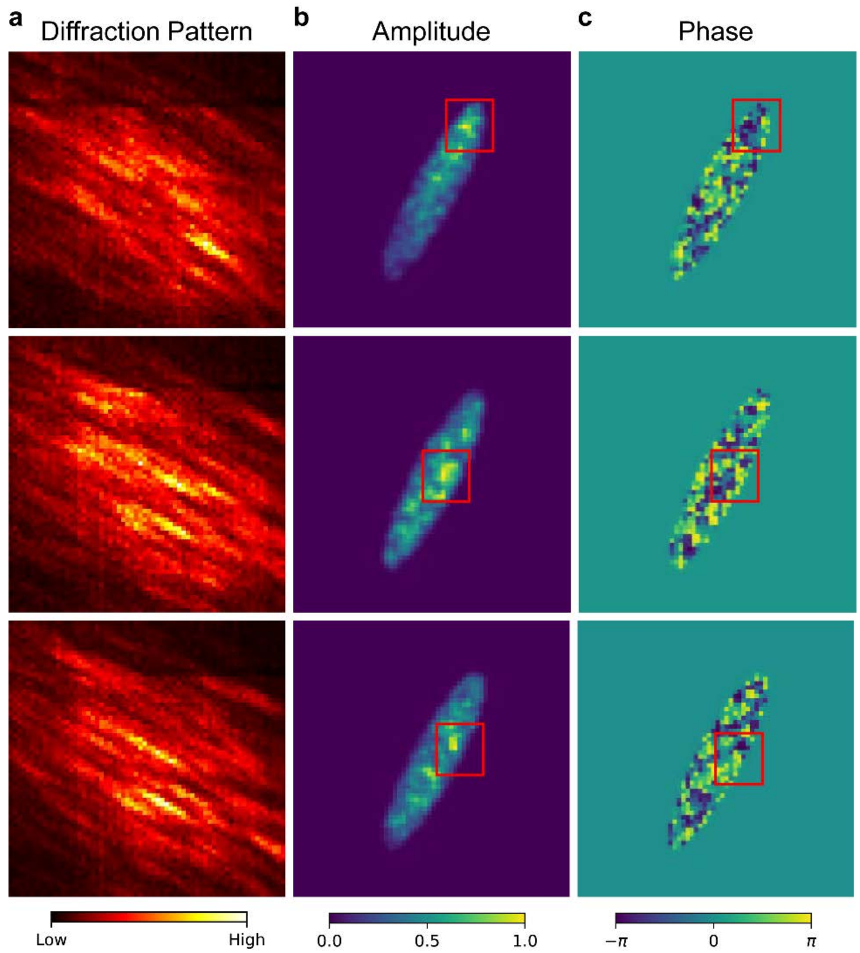
where $\mu_{\boldsymbol{x}}$ and $\mu_{\boldsymbol{y}}$ denote the pixel sample mean of $\boldsymbol{x}$ and $\boldsymbol{y}$. $\sigma_{\boldsymbol{x}}^2$, $\sigma_{\boldsymbol{y}}^2$, and $\sigma_{\boldsymbol{xy}}$ represent the variance of $\boldsymbol{x}$ and variance of $\boldsymbol{y}$**,** and covariance of $\boldsymbol{x}$ and $\boldsymbol{y}$. c_1 and c_2 are $(k_1L)^2$ and $(k_2L)^2$, respectively, where L is the dynamic range of pixel-values. We used the default values $k_1 = 0.01$ and $k_2 = 0.03$.



Supplementary Figure 1 Synthetic data samples for pre-trained model on the real XFEL experimental data. a synthetic phase. **b** synthetic amplitude. **c** synthetic diffraction pattern. We choose the Gaussian support as the ellipsoid shape with semi-major 18 and semi-minor 4, and rotation with 28 degrees. As can be seen, the synthetic diffractions are similar to the real XFEL experimental data, and we use such synthetic data to train a model as the initiation and continue training on real XFEL experimental data.



Supplementary Figure 2 2D correlation coefficient heatmaps of real experimental XFEL data. We calculate correlation coefficient for these 1002 average images and group them into 20 different clusters. The pre-processing real XEFL data is obtained by averaging each cluster. We first calculate the correlations between 1002 average images shown in **a**, and then utilizing the hierarchical clustering method to find hierarchy within the data and ordering the data in clusters. In specific, we employ the dissimilarity matrix computed by $d(x, y) = 1 - |\rho_{x,y}|$. After computing the dissimilarity matrix, we can group data hierarchically according to their dissimilarity, we visualize it in **b** and obtain 20 different clusters with threshold equal to 0.7.



Supplementary Figure 3: predicted amplitude and phase in real space for three different real experimental XFEL test data. a the experimental XFEL diffraction pattern samples**.** The red rectangle highlights the large value regions of the **b** predicted amplitude and the corresponding regions of **c** predicted phase. We observe the movement of large value regions in the predicted amplitude, but they are not related in the phase, which indicates that movement of the beam rather than internal sample fluctuations.