Title
Applications of Iterative and Deep-Learning Algorithms in 2-D and 3-D Coherent Electron Microscopy

Author
Chang, Dillan Jae

Publication Date
2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Applications of Iterative and Deep-Learning Algorithms in 2-D and 3-D Coherent Electron Microscopy

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Physics

by

Dillan Jae Chang

2023

ABSTRACT OF THE DISSERTATION

Applications of Iterative and Deep-Learning Algorithms
in 2-D and 3-D Coherent Electron Microscopy

by

Dillan Jae Chang
Doctor of Philosophy in Physics
University of California, Los Angeles, 2023
Professor Jianwei Miao, Chair

Significant improvements have been made in pipelines for the acquisition and storage of data from scanning electron imaging experiments. Within these data contain valuable information such as structure, defects and compositions of samples that must first be inverted via an expensive numerical process. A typical scanning CDI imaging setup records measurements in 4-dimensions – 2 from scanning and 2 from scattering dimensions, and with countless microscopy experiments and continuous infrastructure upgrades, the computational demand for facilitating the inversion process and analyzing the results grow at an exponential pace.

A numerical study of the advantages of incorporating ptychographic phase projections in atomic electron tomography is presented. Reconstructed phase images are linear projections of the Coulomb potential and thus are able to image low-Z atoms at a lower electron dose. Advantages of ptychographic AET are presented with numerical simulations of the methodology on a 5-nm zinc-oxide nanoparticle and WS_2/WSe_2 van der Waals heterostructure.

In the field of machine learning, specifically in the field of deep learning, computer scientists were able to effectively leverage Graphics Processing Units to train mathematical

structures that act as universal function approximators. These neural networks shocked the world in their power and versatility: complex image classification, creative text generation, and calculation of heuristics for games previously thought by mankind as impossible.

This thesis facilitates the merging of Fourier microscopy and machine learning via the introduction of a novel deep-learning based algorithm that processes diffraction patterns into meaningful phase image reconstructions. After a systematic explanation of the algorithm, experimental results from scanning electron CDI experiments imaging monolayer graphene, twisted bilayer hexagonal boron nitride, and a gold nanoparticle are presented along with comparisons via traditional methods.

The dissertation of Dillan Jae Chang is approved.

Seth J. Putterman

Pietro Musumeci

Jose Alfonso Rodriguez

Jianwei Miao, Committee Chair

University of California, Los Angeles

2023

To Andrea

Table of Contents

List of Figures

List of Tables

List of Acronyms

CDI	Coherent Diffractive Imaging
HIO	Hybrid Input-Output
ePIE	Extended Ptychographic Iterative Engine
CNN	Convolutional Neural Network
DL-CDI	Deep Learning Coherent Diffractive Imaging
GPU	Graphics Processing Unit
RAM	Random Access Memory
MAP	Maximally Activated Patterns
(HA)ADF	(High Angle) Annular Dark Field
(A)BF	(Annular) Bright Field
STEM	Scanning Transmission Electron Microscopy
FRC	Fourier Ring Correlation
hBN	Hexagonal Boron Nitride
SSB	Single-Side Band
WDD	Wigner Distribution Deconvolution
pAET	Ptychographic Atomic Electron Tomography
FBP	Filtered Back-projection
ART	Algebraic reconstruction technique
GENFIRE	Generalized Fourier Iterative Reconstruction Algorithm
EDX	Energy-dispersive X-ray
SRO	Short-Range Order
MRO	Medium-Range Order

Acknowledgments

All of this work would have been impossible without the help and support from fellow group members, collaborators and Professor Miao. Thank you for providing me an environment where I was truly challenged to my limits and grew in ways that I didn't anticipate when beginning this journey.

I'd like to express special thank you's (in no particular order) to my sunbaes – Mike, Arjun, Dennis and Colum – for your help and company the past few years. Accomplishing something like this thesis becomes possible when one can step on the shoulders of academics like you.

And finally to my wife, Andrea, thank you for your patience throughout the years and helping me understand the greater purpose in what we do. Both of our grit, sweat, and tears were poured onto these pages and I hope that you accept my dedication of this work to you.

Chapter 2.1 is an adapted and partially modified version, including figures, of "Determining the Three-Dimensional Atomic Structure of an Amorphous Solid", Nature **592**, 60 (2021). I acknowledge the coauthors for their contributions: Y. Yang, J. Zhou, F. Zhu, Y. Yuan, D. S. Kim, M. Pham, A. Rana, X. Tian, Y. Yao, S. J. Osher, A. K. Schmid, L. Hu, P. Ercius, and J. Miao. J.M. conceived the idea and directed the project; Y. Yao and L.H. synthesized the samples; J.Z., P.E., A.K.S. and J.M. discussed and/or carried out the experiments; M.P., Y. Yuan, A.R., S.J.O. and J.M. developed the RESIRE algorithm. Y. Yang, F.Z., Y. Yuan, D.J.C., J.Z., D.S.K., X.T. and J.M. performed image reconstruction, atom tracing and classification, analysed the data and/or interpreted the results; J.M., Y. Yang, J.Z. and F.Z. wrote the manuscript.

Chapter 2.2 is an adapted and partially modified version, including figures, of "Three-Dimensional Atomic Packing in Amorphous Solids with Liquid-like Structure", Nat. Mater. **21**, 95 (2022). I acknowledge the coauthors for their contributions: Y. Yuan, D. S. Kim, J. Zhou, F. Zhu, Y. Nagaoka, Y. Yang, M. Pham, S. J. Osher, O. Chen, P. Ercius, A. K. Schmid, and J. Miao. J.M. conceived the idea and directed the project; A.K.S. synthesized the amorphous Ta thin film; Y.N. and O.C. synthesized the amorphous Pd nanoparticles; J.Z., A.K.S., P.E. and J.M. discussed and performed the AET experiments of the amorphous materials; and M.P., Y. Yuan, S.J.O. and J.M. developed the 3D reconstruction algorithm. Y. Yuan, D.S.K., D.J.C., F.Z., Y. Yang and J.M. performed image reconstruction and atom tracing, analysed the data and interpreted the results; D.S.K. and J.M. discussed and performed the MD simulations. Y. Yuan, D.S.K. and J.M. wrote the manuscript.

Chapter 3 is an adapted and partially modified version, including figures, of the article "Ptychographic Atomic Electron Tomography: Towards 3D Imaging of Individual Light Atoms in Materials", Phys. Rev. B, **102**, 174101, (2020). I acknowledge the coauthors for their contributions: D. S. Kim, A. Rana, X. Tian, J. Zhou, P. Ercius, and J. Miao. Reprinted with permission from [1], Copyright (2020) by the American Physical Society.

Chapters 4 and 5.1 are an adapted and partially modified version, including figures, of the article "Deep-Learning Electron Diffractive Imaging", Phys. Rev. Lett. **130**, 016101 (2023). I acknowledge the coauthors for their contributions: C. M. O'Leary, C. Su, D. A. Jacobs, S. Kahn, A. Zettl, J. Ciston, P. Ercius, and J. Miao. Reprinted with permission from [2], Copyright (2023) by the American Physical Society.

This work was primarily supported by the U.S. Department of Energy (DOE), Office of Science, Basic Energy Sciences (BES), Division of Materials Sciences and Engineering under Award No. DE-SC0010378. We also acknowledge the support by STROBE: A National Science Foundation Science & Technology Center under Grant No. DMR 1548924, an Army Research Office MURI grant on Ab-Initio Solid-State Quantum Materials: Design, Production and Characterization at the Atomic Scale (AFOSR Grant No. W911NF1810432) and the NSF DMREF program (Grant No. DMR-1437263). We also thank Gatan, Inc. as well as P. Denes, A. Minor, C. Ophus, J. Joseph, and I. Johnson at LBNL, who contributed to the development of the 4D Camera used for STEM-DPC measurements. The electron ptychography experiments with TEAM I and TEAM 0.5 were performed at the Molecular Foundry, which is supported by the Office of Science, Office of Basic Energy Sciences of the U.S. DOE under Contract No. DE-AC02-05CH11231. We acknowledge the support by the Director, Office of Science, Office of Basic Energy Sciences, Materials Sciences and Engineering Division, of the U.S. DOE within the sp2-Bonded Materials Program (KC-2207) which provided for preliminary TEM and Raman characterization of the h-BN material, and by the van der Waals Heterostructures program (KCWF16) which provided for assembly of the twisted h-BN material.

Vita

2017	B.S. Engineering Physics
	B.S. Computer Science
	Cornell University
2019	M.S. Physics
	University of California, Los Angeles

Publications

- **D. J. Chang**, D. S. Kim, A. Rana, X. Tian, J. Zhou, P. Ercius, and J. Miao, *Ptychographic Atomic Electron Tomography: Towards Three-Dimensional Imaging of Individual Light Atoms in Materials*, Phys. Rev. B **102**, 174101 (2020).

- **D. J. Chang***, C. M. O'Leary*, C. Su, D. A. Jacobs, S. Kahn, A. Zettl, J. Ciston, P. Ercius, J. Miao, *Deep-Learning Electron Diffractive Imaging*, Phys. Rev. Lett. **130**, 016101 (2023).

- **D. J. Chang**, J. Miao, *Extending Deep-Learning Coherent Diffractive Imaging to Complex Objects*, In preparation.

- Y. Yang, J. Zhou, F. Zhu, Y. Yuan, **D. J. Chang**, D. S. Kim, M. Pham, A. Rana, X. Tian, Y. Yao, S. J. Osher, A. K. Schmid, L. Hu, P. Ercius, and J. Miao, *Determining the Three-Dimensional Atomic Structure of an Amorphous Solid*, Nature **592**, 60 (2021).

- Y. Yuan, D. S. Kim, J. Zhou, **D. J. Chang**, F. Zhu, Y. Nagaoka, Y. Yang, M. Pham, S. J. Osher, O. Chen, P. Ercius, A. K. Schmid, and J. Miao, *Three-Dimensional Atomic Packing in Amorphous Solids with Liquid-like Structure*, Nat. Mater. **21**, 95 (2022).

Chapter 1

Introduction

1.1 Atomic Electron Tomography

Atomic Electron Tomography (AET) determines 3D coordinates of individual atoms in materials without assuming crystallinity by combining high resolution tomographic tilt series from scanning transmission electron microscopes (STEM) with powerful iterative tomographic algorithms [3]. Since its inception, AET has successfully imaged grain boundaries, dislocations, point defects, bond distortion, strain tensors, chemical order/disorder and nucleation dynamics with unprecedented 3D detail [4–10]. This section will briefly discuss each step of the methodology and discuss its application in material science at the single-atom level.

1.1.1 Tilt-series Data Acquisition

In order to perform tomography, 2-dimensional projections of a 3-dimensional object must be collected at various projection angles to acquire a tilt-series. For atomic resolution projections, electron microscopes are required due to their small wavelength (finer than the characteristic size of an atom), and other forms of imaging such as scanning probe microscopy can image surface features but fail in structural determination through the sample. As a result, all experimental realizations of AET relied on scanning transmission electron

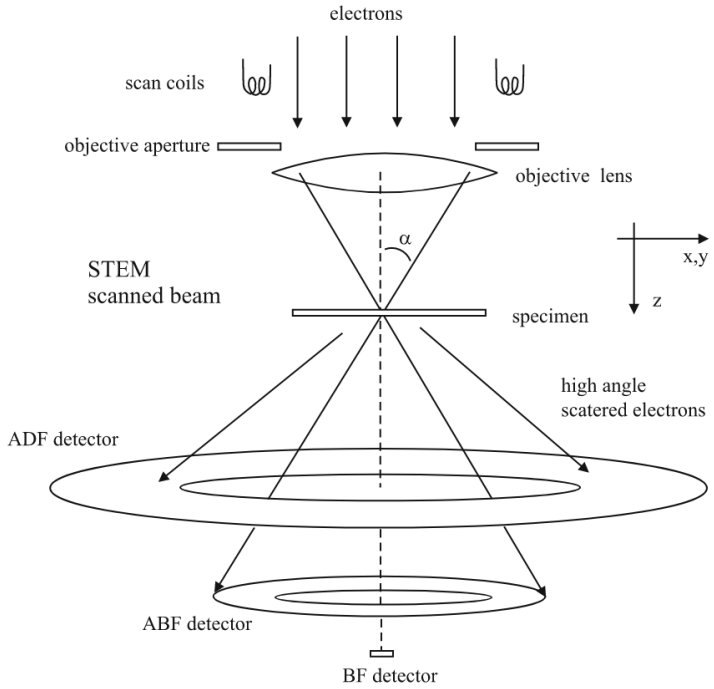

Figure 1.1: Taken from ref [11]. Simplified model (not to scale) of a Scanning Transmission Electron Microscope (STEM). Different detectors are BF = bright field, ADF = annular dark field, ABF = annular bright field. Reproduced with permission from Springer Nature.

microscopy (STEM) whereby an aberration-corrected probe is scanned across a sample and a detector measures the resulting electrons at certain collection angles.

Figure 1.1 shows a simplified model of a STEM, specifically focusing on the detectors. While any of the ADF, ABF or BF detectors can be used to 2D image the specimen, AET employs ADF or sometimes HAADF (High angle annular dark field) detectors. The reason for using dark field detectors is that we would like the atomic features to be recorded as positive semidefinite values on the image, with strongly scattering atoms being recoreded as high signal values. In the presence of scattering features on the specimen, bright field signals tend to attenuate, which render it difficult to use as projections for tomography.

Another reason for employing ADF detectors is due to extensive study on the relationship between atomic number and ADF signals. When imaging single atoms (no atomic columns),

previous studies report a power relation (roughly proportional to $Z^{1.8}$, where Z is atomic number) in ADF contrast [12–14], which is satisfactory for tomographic purposes and important for distinguishing between atom types in AET. Individual atoms with higher atomic numbers could in theory be identified in 3-dimensions by fitting the 3D intensity profile with a Gaussian function and measuring the peak of the curve.

An exact mathematical expression for ADF STEM contrast can be described by the following [11]. In STEM, an aberrated electron probe function $\psi_p(\mathbf{x}, \mathbf{x}_p)$ strikes the specimen at position $\mathbf{x}_p$:

$$\psi_p(\mathbf{x}, \mathbf{x}_p) = A_p \int_0^{k_{\max}} \exp[-i\chi(\mathbf{k}) - 2\pi i \mathbf{k} \cdot (\mathbf{x} - \mathbf{x}_p) d^2\mathbf{k}] \tag{1.1}$$

where $\lambda k_{\max} = \alpha_{\max}$ is the maximum angle in the objective aperture and A_p is a normalization constant chosen to yield

$$\int |\psi_p(\mathbf{x}, \mathbf{x}_p)|^2 d^2\mathbf{x} = 1 \tag{1.2}$$

If the specimen can be approximated as a weak phase, weak amplitube object, the transmitted wave function is given as

$$\psi_t(\mathbf{x}, \mathbf{x}_p) = t(\mathbf{x})\psi_p(\mathbf{x}, \mathbf{x}_p) \tag{1.3}$$

where $t(\mathbf{x})$ is the specimen object function. When this wave function diffracts into the Fraunhofer regime, the function on the detector plane is

$$\Psi_t(\mathbf{k}, \mathbf{x}_p) = \int \exp(2\pi i \mathbf{k} \cdot \mathbf{x}) \psi_t(\mathbf{x}, \mathbf{x}_p) d^2\mathbf{x} \tag{1.4}$$

The annular dark field detector incoherently measures this wave function by integrating between the inner and outer angles, $\lambda k_{D_{\min}}$ and $\lambda k_{D_{\max}}$ respectively, and the result is the

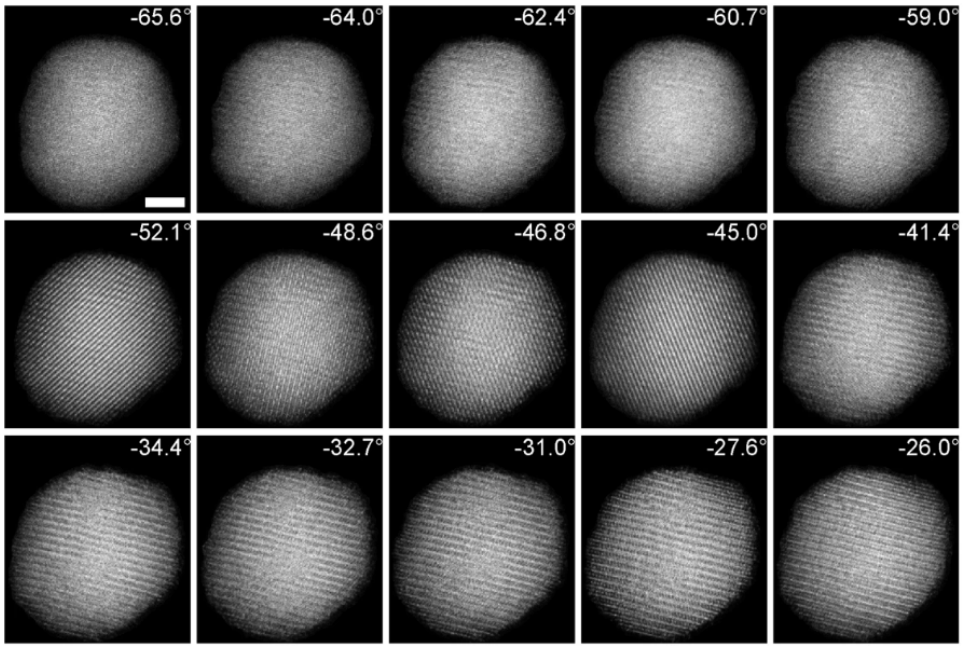

Figure 1.2: Taken from ref [7]. The projection images of an FePt nanoparticle measured using an ADF-STEM. Reproduced with permission from Springer Nature.

final STEM image signal $g(\mathbf{x}_p)$ for one probe position x_p.

$$g(\mathbf{x}_p) = \int |\Psi_t(\mathbf{k}, \mathbf{x}_p)|^2 D(\mathbf{k}) d^2\mathbf{k} \tag{1.5}$$

where $D(\mathbf{k})$ is the detector function.

$$D(\mathbf{k}) = 1, k_{D_{\min}} \leq k \leq k_{D_{\max}} \tag{1.6}$$

$$= 0, \text{otherwise} \tag{1.7}$$

After scanning across 2-dimensions of probe positions and 1 dimension of tilt angles, one can generate a tilt-series of projections for tomography like one shown in Figure 1.2.

1.1.2 Tomography Algorithm

The tilt-series data can be compiled together to reconstruct a singular 3-dimensional object via tomographic algorithms. A common tomographic algorithm commonly used in medical imaging is called filtered back-projection (FPB), a back-projection method combined with a convolution filter to remove blurring [15, 16]. The method works well in cases where there are a high number projections with perfect data, but in the presence of noise and sparsity in projections, reconstruction results from FBP suffer from artifacts. To overcome these limitations, various iterative techniques such as the algebraic reconstruction technique (ART) [17] and simultaneous iterative reconstruction technique (SIRT) [18] have been developed, but they do not fully leverage mutual information shared between projections.

Fourier based algorithms such as equal slope tomography (EST) [19] and generalized Fourier iterative reconstruction algorithm (GENFIRE) [20] reconstruct the volume globally every iteration taking the projection information into consideration simultaneously. As GENFIRE is a generalization of EST for any arbitrary tilt angles, GENFIRE has been used successfully to reconstruct 3-dimensional volumes for AET for the past decade. Figure 1.3 shows the schematic of the GENFIRE algorithm.

The mathematical principle behind GENFIRE is the Fourier slice theorem which states that the 2D Fourier transform of a volumetric projection is equivalent to the slice of the 3D Fourier transform of the volume at that same angle through the origin [16]. This means that if you take the Fourier transform of the tilt-series projections, those values can be enforced in the 3D Fourier transform of the reconstruction volume at voxels along slices through the origin. After the Fourier voxels are replaced with the measured values, one performs the inverse Fourier transform to convert to real space, where one can enforce constraints such as positivity (volume must be filled with positive semidefinite values) and support (volume outside a predetermined region must be 0). After the real space constraint is applied, a 3D

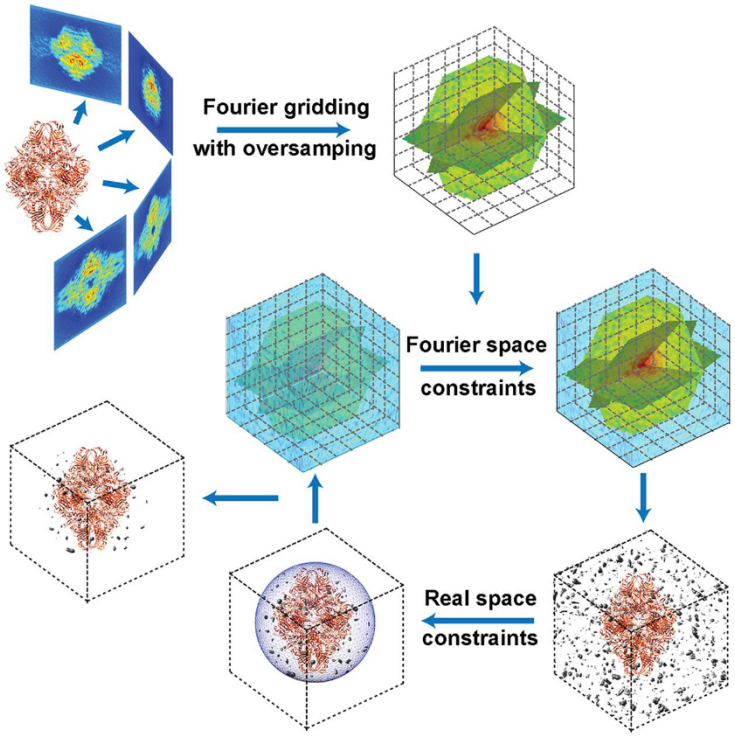

Figure 1.3: Taken from ref [20]. The GENFIRE algorithm. GENFIRE first computes oversampled Fourier slices from a tilt series of 2D projections. The oversampled Fourier slices are used to accurately calculate a small fraction of points on a 3D Cartesian grid based on gridding interpolation. The algorithm then iterates between real and reciprocal space. The support and positivity constraints are enforced in real space, while the small fraction of grid points corresponding to the measured data are enforced in reciprocal space. Error metrics are used to monitor the convergence of the iterative process. After several hundred iterations, the algorithm converges to a 3D structure that is concurrently consistent with the measured data in reciprocal space and the physical constraints in real space. Figure distributed under Creative Commons BY license.

Fourier transform is applied again to conclude one iteration of the algorithm.

The i'th iteration of GENFIRE consists of the following five steps.

1. Apply the inverse FFT to $F_i(\mathbf{k})$ to objtain the i'th real space volume, $\rho_i(\mathbf{r})$.
2. Modify the image by applying the following constraints

$$\rho_i'(\mathbf{r}) = \begin{cases} 0 & (\mathbf{r} \notin S) \cup (\rho_i(\mathbf{r}) < 0) \\ \rho_i(\mathbf{r}) & \text{otherwise} \end{cases} \tag{1.8}$$

   where S represents a support. This step enforces the support and positivity constraints described above.
3. Apply the FFT to ρ_i'(r) to obtain $F_i'(\mathbf{k})$.
4. Compute $F_{i+1}(\mathbf{k})$ by enforcing the Fourier space constraint,

$$F_{i+1}(\mathbf{k}) = \begin{cases} F_{\text{obs}}(\mathbf{k}) & \text{Known voxels} \\ F_i'(\mathbf{k}) & \text{Unknown voxels} \end{cases} \tag{1.9}$$

   $F_{i+1}(\mathbf{k})$ is used for the $(i+1)$'th iteration.
5. Calculate two R-factors, R_k and R_{free}

$$R_k = \frac{\sum_{k_{\text{known}}} |F_{\text{obs}}(k) - F_i(k)|}{\sum_{k_{\text{known}}} |F_{\text{obs}}(k)|} \tag{1.10}$$

$$R_{\text{free}} = \frac{\sum_{k_{\text{withheld}}} |F_{\text{obs}}(k) - F_i(k)|}{\sum_{k_{\text{withheld}}} |F_{\text{obs}}(k)|} \tag{1.11}$$

   where k_{known} represents the known voxels and k_{withheld} is a small number of randomly selected known voxels that are not used in the reconstruction. These error metrics can be used to monitor the convergence and progress of the algorithm.

One important aspect in tomography is keeping the "missing wedge" artifact as low as

possible. Missing wedge artifacts occur due to the fact that tilt stages typically have an angle limit (such as -70 to +70 degrees) resulting in a wedge of missing Fourier data in the reconstruction. This artifact presents as a smear or spread along the direction of missing projections, and is intensified if the projections come with inherent systemic errors or noise.

At the end of a successful execution of AET, from the tilt-series data collection to proper projection alignments and tomography, the practitioner will obtain impressive 3D reconstructions of nanomaterials at atomic resolutions high enough to be able to trace each atom in 3D to picometer precision. For example in 2017, Yang et. al. was able to determine the 3D coordinates of 6,569 iron and 16,627 platinum atoms in an iron-platinum nanoparticle, and correlate chemical order/disorder and crystal defects with material properties at the single-atom level [7]. Since AET does not assume crytallinity at all during reconstruction, the final volume contains enough information to not only pinpoint 3D structure of defects and grains, but also identify the atom species in the FePt nanoparticle allowing determination of various chemical phases within the material [Figure 1.4].

Furthermore, the boundaries of AET reached a new horizon when Zhou et. al. in 2019 was able to use AET to study early-stage nucleation in four dimensions (that is, including time) at atomic resolution [6]. To probe nucleation dynamics, Zhou tracked the same nuclei in each particle at different annealing times (termed common nuclei), and showed five growing, fluctuating and dissolving nuclei shown in Figure 1.5.

1.2 Coherent Diffractive Imaging

Coherent Diffractive Imaging (CDI) is a lensless form of microcopy where the object image is reconstructed from measuring the scattered wavefunction without first being collected by a lens [21]. A typical CDI experiment entails a coherent beam source, henceforth called the probe, striking the object in question, scattering into the far-field regime and then being col-

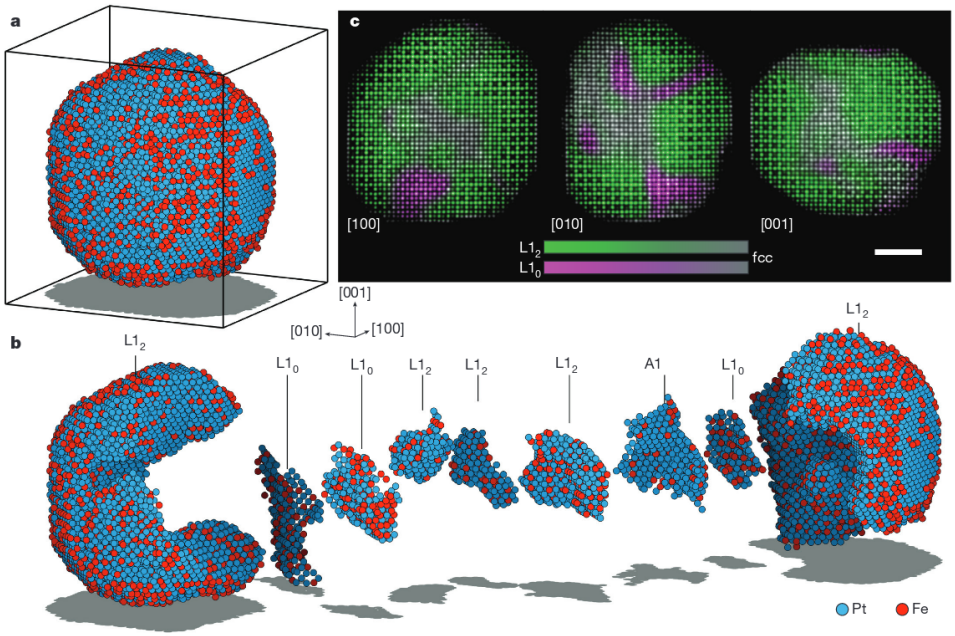

Figure 1.4: Taken from ref [7]. 3D determination of atomic coordinates, chemical species and grain structure of an FePt nanoparticle. a, Overview of the 3D positions of individual atomic species with Fe atoms in red and Pt atoms in blue. b, The nanoparticle consists of two large L12 grains, three small L12 grains, three small L10 grains and a Pt-rich A1 grain. c, Multislice images obtained from the experimental 3D atomic model along the [100], [010] and [001] directions, where several 'L10 grains' (magenta) appearing in the 2D images are deceptive structural information. Colour bars indicate the degree of ordering, from pure L12/L10 to chemically disordered fcc. Scale bar, 2 nm. Reproduced with permission from Springer Nature.

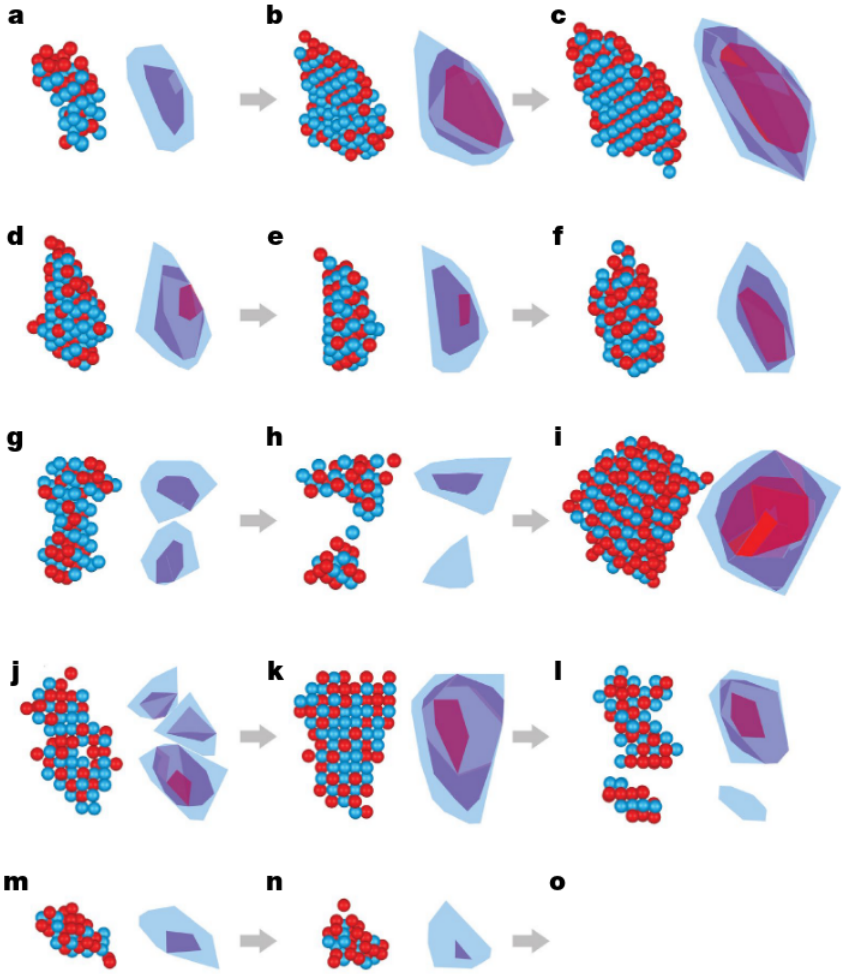

Figure 1.5: Taken from ref [6]. Experimental observation of the same nuclei undergoing growth, fluctuation, dissolution, merging and/or division at 4D atomic resolution. a–c, A representative growing nucleus with an accumulated annealing time of 9 min, 16 min and 26 min, respectively. The atomic models show Fe (red) and Pt (blue) atoms with an order parameter ¡= 0.3, and the 3D contour maps show the distribution of an order parameter of 0.7 (red), 0.5 (purple) and 0.3 (light blue). d–l, Three representative fluctuating nuclei at three annealing times, including merging and dividing nuclei. m–o, A representative dissolving nucleus at three annealing times, which dissolved at 26 min (o). Reproduced with permission from Springer Nature.

lected by a pixel array detector. This form of imaging lends itself to some advantages. First, while lenses are readily available for visible light microscopy, precise optical manipulation for other types of probes such as X-rays is costly and difficult [22]. Second, the resolution obtained by CDI is not limited by careful control of the optics but rather by the largest scattering angle at which information is recorded.

First experimental demonstration of soft X-ray CDI for a non-crystalline sample was performed by Miao et. al. in 1999 [23]. A plane X-ray wave interfered with the sample and the resulting diffraction pattern was measured by a detector. The measurement at the detector can be mathematically represented as the magnitude of the Fourier transform of the object.

$$\Psi(\boldsymbol{k}) = |\mathcal{F}(O(\boldsymbol{r}))| \tag{1.12}$$

where the phase information in momentum space is lost during measurement. A naive approach of simply performing an inverse Fourier transform to recover the object will not work – a problem commonly known as the phase problem.

$$O(\boldsymbol{r}) \neq \mathcal{F}^{-1}\left(|\mathcal{F}(O(\boldsymbol{r}))|\right) \tag{1.13}$$

To successfully recover the image, the diffraction pattern must be oversampled at a frequency finer than the Nyquist interval and then passed through an algorithm that iteratively applies Fourier constraints from experimental measurements [24]. This algorithm, also called the hybrid input-output (HIO) algorithm, is described in detail below [25]:

At the j'th iteration,

1. Fourier transform an trial input object $f_j(\boldsymbol{r})$ to get a Fourier pattern $F_k(\boldsymbol{k})$ in reciprocal space.
2. Generate a new Fourier pattern $F'_j(\boldsymbol{k})$ by using the measured diffraction pattern as the magnitude and the phase of $F'_j(\boldsymbol{k})$ as its phase.

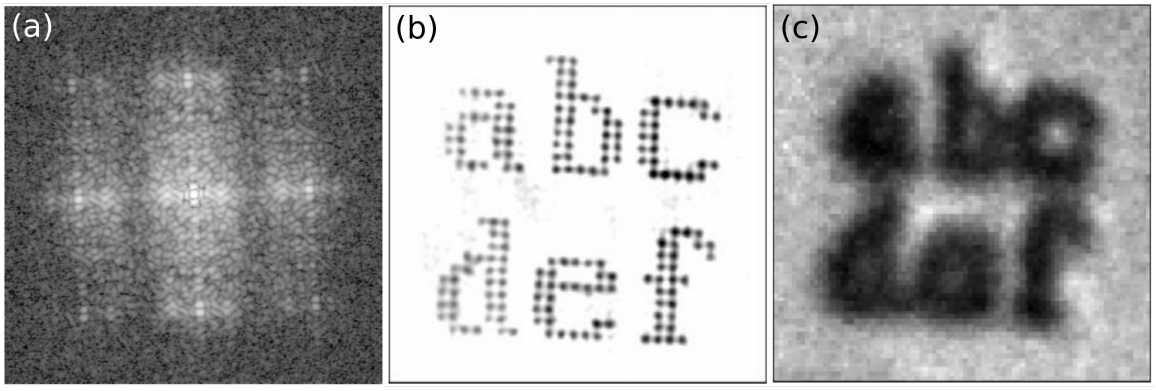

Figure 1.6: Taken from ref [23]. Gold dot X-ray CDI results. (a) Measured diffraction pattern. (b) Phase image of gold dots reconstructed from (a). (c) Optical microscope image of (b). Reproduced with permission from Springer Nature.

3. Inverse Fourier transform $F'_j(\boldsymbol{k})$ to get a new real space object, $f'_j(\boldsymbol{r})$.
4. Generate the $(j+1)$'th object by

$$f_{j+1}(\boldsymbol{r}) = \begin{cases} f'_j(\boldsymbol{r}) & r, f'_j(\boldsymbol{r}) \in S \\ f_j(\boldsymbol{r}) - \beta f'_j(\boldsymbol{r}) & r, f'_j(\boldsymbol{r}) \notin S \end{cases} \tag{1.14}$$

where β represents a constant between 0.5 and 1.0, and S represents pixels inside a finite real space support and the imaginary parts are positive.

By applying this algorithm to the X-ray diffraction pattern shown in Figure 1.6 (a), a high resolution image of the gold dots can be recovered, as shown in Figure 1.6 (b). As a comparison, an optical microscope image of the dots is shown in Figure 1.6 (c).

An important detail to note when performing CDI experiments is the requirement for redundancies in measured data in order to perform the reconstruction. In the case of single plane wave CDI experiments, a support window in real space is applied to enforce a small field-of-view, and the diffraction pattern is oversampled. An intuitive explanation for this requirement is that the number of independently measured points must be greater than the number of unknown variables (both magnitude and phase of object in real space) in order

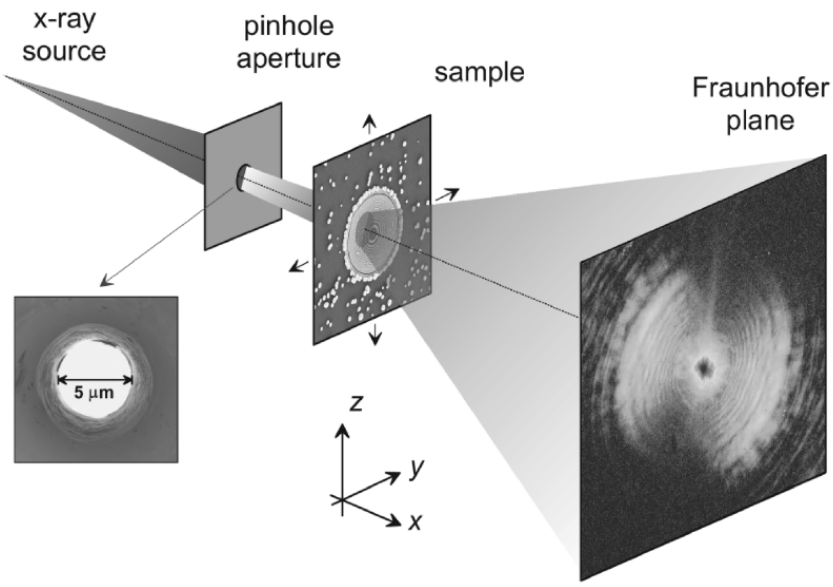

Figure 1.7: Schematic of an X-ray ptychography experiment. Reprinted figure with permission from [29], copyright (2017) by the American Physical Society.

to be recovered by the iterative algorithm. [25, 26]

1.2.1 Ptychography

Ptychography is an implementation of CDI whereby a larger field-of-view of the object is reconstructed by scanning the coherent illumination across the object and iteratively reconstructing the object from numerous diffraction patterns. Although the concept of combining overlapped diffraction patterns can be traced back to Hoppe in 1969 [27], the modern version of ptychography which incorporates iterative algorithms was introduced by Rodenburg in 2004 [28, 29].

Figure 1.7 details the schematic of an X-ray ptychography experimental setup consisting of a probe function striking the sample which can shifted in a gridlike fashion. The measurement in the Fraunhofer plane is mathematically shown as

$$\Psi(\boldsymbol{k}) = |\mathcal{F}(P(\boldsymbol{r})O(\boldsymbol{r} - \boldsymbol{R}))| \tag{1.15}$$

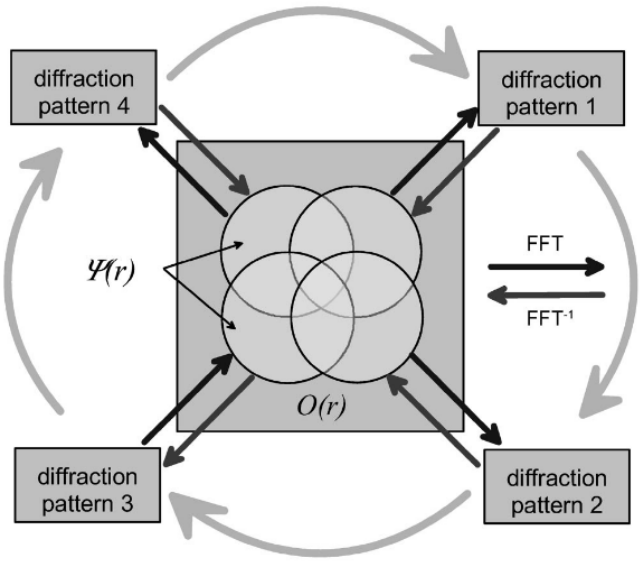

Figure 1.8: Diagram of ptychographic phase retrieval algorithm. Reprinted figure with permission from [29], copyright (2017) by the American Physical Society.

where $\boldsymbol{R}$ is the scan position of the probe on the sample, P is the illuminating probe function, and O is the sample object function. The resulting diffraction data will be 4 dimensions in total – 2 from the diffraction patterns and 2 from the scanning positions.

The diffraction data is then used as an input for an iterative ptychographic phase retrieval algorithm, similar to the HIO algorithm used in single shot experiments. Figure 1.8 demonstrates a high-level diagram of such an algorithm, where each diffraction pattern is enforced for each probe position through repeated Fourier and inverse Fourier transforms. Note the use of overlapped probe illuminations on the sample plane which serves as data redundancies required to overcome the phase problem.

An important extension of ptychographic algorithms emerged from Maiden and Rodenburg in 2009 that allowed simultaneous reconstruction of the object and probe functions from experimental diffraction patterns [30]. Named ePIE (extended ptychographic iterative engine), the algorithm performs complex conjugate gradient descent on both of the probe and object functions for every iteration, and therefore does not require an accurate model

of the probe function in order to obtain successful reconstruction.

The iterative algorithm describing the update from the j'th diffraction pattern, $I_j(\boldsymbol{k})$, is described below:

1. Generate the j'th exit wave guess by multiplying the probe function and the object function at location $\boldsymbol{R_j}$.

$$\psi_j(\boldsymbol{r}) = P(\boldsymbol{r})O(\boldsymbol{r} - \boldsymbol{R}_j) \tag{1.16}$$

2. Apply the measured diffraction by taking the Fourier transform of ψ_j and substituting the magnitude.

$$\Psi_j(\boldsymbol{k}) = \sqrt{I_j(\boldsymbol{k})}\frac{\mathcal{F}[\psi_j(\boldsymbol{r})]}{|\mathcal{F}[\psi_j(\boldsymbol{r})]|} \tag{1.17}$$

3. Calculate the updated exit wave by taking the inverse Fourier transform.

$$\psi'_j(\boldsymbol{r}) = \mathcal{F}^{-1}[\Psi_j(\boldsymbol{k})] \tag{1.18}$$

4. Update the probe and object functions via conjugate gradient descent.

$$O_{new}(\boldsymbol{r} - \boldsymbol{R}_j) = O(\boldsymbol{r} - \boldsymbol{R}_j) + \alpha\frac{P^*(\boldsymbol{r})}{|P(\boldsymbol{r})|^2_{max}}(\psi'_j(\boldsymbol{r}) - \psi_j(\boldsymbol{r})) \tag{1.19}$$

$$P_{new}(\boldsymbol{r}) = P(\boldsymbol{r}) + \beta\frac{O^*(\boldsymbol{r} - \boldsymbol{R}_j)}{|O(\boldsymbol{r} - \boldsymbol{R}_j)|^2_{max}}(\psi'_j(\boldsymbol{r}) - \psi_j(\boldsymbol{r})) \tag{1.20}$$

In the past few decades, ptychography has found wide applications with synchrotron radiation, high harmonic generation, electron and optical microscopy [21, 31–46].

1.2.2 Electron Ptychography

By employing a coherent electron source, ptychography can be performed at sub-Ångstrom resolution, resolving 2D images at the atomic scale [11, 47]. For electron waves, the probe function interacts primarily with the Coulomb potential of the sample, where the specimen transmission function given as

$$t(\boldsymbol{r}) = \exp(i\sigma_e v_z(\boldsymbol{r})) \tag{1.21}$$

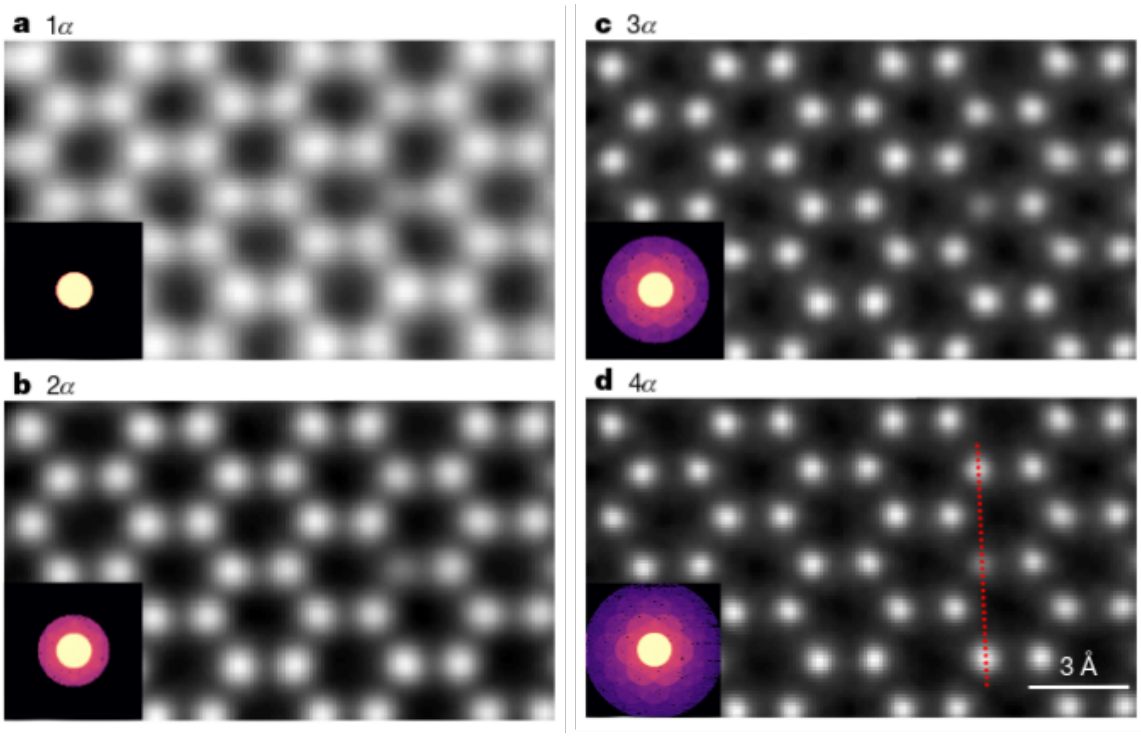

Figure 1.9: Taken from ref [47]. Ptychographic reconstructions with different maximum collection angles. Reproduced with permission from Springer Nature.

where v_z is the 2D projected atomic Coulomb potential and σ_e is the interaction parameter

$$\sigma_e = \frac{2\pi}{\lambda V}\left(\frac{m_0c^2 + eV}{2m_0c^2 + eV}\right) \tag{1.22}$$

where m_0 is the mass of electron and V is the accelerated voltage.

Since the full width at half maximum for isolated atomic potentials range in the order of 0.1 Ångstrom [11], the resolution of electron ptychography experiments will be primarily determined by the maximum collection angle of electrons, $\theta_{\max}$.

$$\text{resolution} = \frac{\lambda}{\sin(\theta_{\max})} \tag{1.23}$$

Figure 1.9 shows the effect of cropping the maximum collection angles on ptychographic

reconstructions [47].

1.3 Convolutional Neural Networks

Neural network based methods attempt to mimic highly-dimensional functions through layers of nonlinear activation functions, inspired by the behavior of neurons. While the first computational model of neural networks emerged in the 1940s [48], the utility of artificial neural networks as tools for non-linear regression and pattern recognition was fully fleshed out in the 1980s with the incorporation of back-proagation algorithms for supervised training [49]. With the spread of graphics processing units (GPUs) designed for parallelization of computational tasks in machine learning tasks, different forms of neural network architectures find widespread use in natural language processing, image recognition, and computer vision applications today [50].

Deep convolutional neural networks (CNNs) which are neural networks with numerous layers of convolution as trainable weights are most apt for image recognition and processing tasks, since information in images tend to relate in limited spatial ranges. One of the most popular types of convolutional neural networks is called the U-net architecture which follows the principle of encoding information into numerous small-size channels, and then decoding the numerous channels into a final low-dimensional representation [51]. The generalizability of CNNs is powerful enough so that CNNs self-trained to learn the evaluation and policy in complex board games such as Go and chess are considerably stronger than the best human players [52]. In this section, I will briefly discuss the fundamental mathematical concepts behind convolutional neural networks.

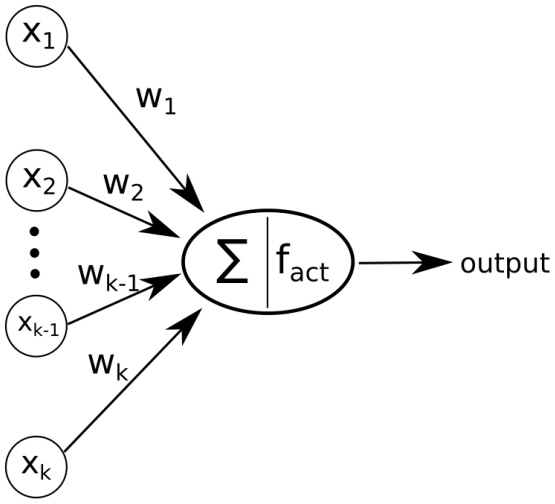

Figure 1.10: A neuron takes a vector of inputs $x_1 ... x_k$, modulates them with trained weights w_1 ... w_k and acts an activation function on the sum of the values.

1.3.1 Neural Networks

In 1943, McCulloch and Pitts introduced the first mathematical model of a biological neuron and hypothesized that a collection of these neurons, or perceptrons, can be networked together to act as logic gates [48]. A diagram of a perceptron, which is the main building block of a neural network, is shown in Figure 1.10.

A single neuron acts as a binary linear separator in the hyperspace defined by the span of x_1 to x_k. Multiple neurons arranged in a layer together allows binary separation in hyperplanes, with the dimension of the hyperplane equal to the number of neurons. The motivation of such a mathematical framework was exciting during a period of technological development in transistors, logic gates, and general computing languages.

One important aspect of a perceptron is in the activation function, f_{act}, which must be a non-linear function.

$$f_{\text{act}}(x+y) \neq f_{\text{act}}(x) + f_{\text{act}}(y) \tag{1.24}$$

The reason for non-linearity is due to the fact that when neurons are arranged vertically as

a series, if the activation function was linear, the entire network would reduce to a series of linear operators, which itself is a linear operator. All of the complexities that neural networks are able to learn derives from the ability of neurons to encapsulate non-linear relations. If 100 layers of neurons arranged in series can be represented as 1 layer, the network simply cannot learn higher-dimensional representations. Typical activation functions include sigmoid and hyperbolic tangent functions that take in real values as input and output a value in the order of unity.

The "learning" of the neuron comes from the determination of the weights that modulate the input vector. If one supervises the learning of the neuron, a series of input vectors X_i are fed into the neuron, and the weights adjust in order to minimize the error between the outputs and the true output values, y_i. The complexity of the learning comes from the number of parameters, or weights, that can be trained as well as the relative arrangement of the neurons in the network.

1.3.2 Backpropagation Algorithm

The computational neural network experienced a boom during the 1980s when numerous (and independent) scientists formulated a systematic way to train the neurons, called the backpropagation algorithm [49, 53–55]. The algorithm defines an objective or error function Err, and performs a gradient descent on it with respect to the neural network weights. Suppose that we have the following neuron $w_{p,q}$ in the following network shown in Figure 1.11. The update value for that neuron is proportional to the derivative of an error function, Err, with respect to that neuron.

$$\Delta w_{p,q} = -\eta \frac{\partial \mathrm{Err}(w_{p,q})}{\partial w_{p,q}} \tag{1.25}$$

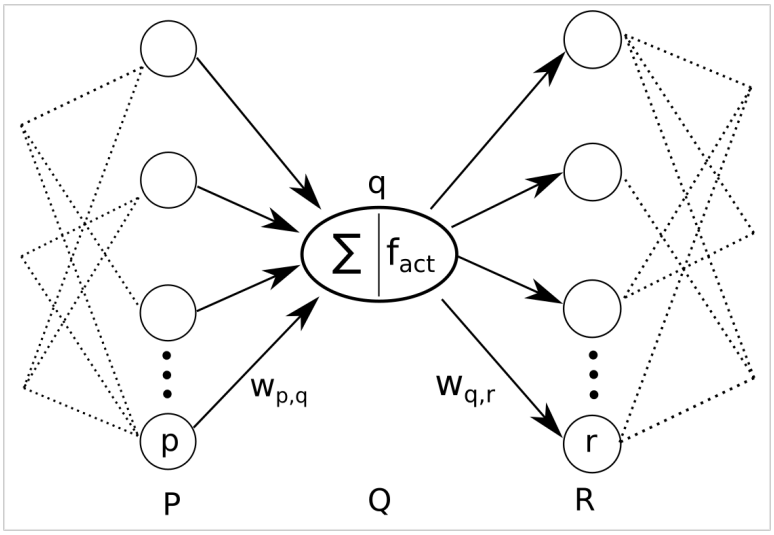

Figure 1.11: A diagram of a neural network. The backpropagation algorithm will attempt to update the weight, $w_{p,q}$.

Taking the chain rule,

$$\frac{\partial \text{Err}(w_{p,q})}{\partial w_{p,q}} = \frac{\partial \text{Err}}{\partial \text{net}_q} \cdot \frac{\partial \text{net}_q}{\partial w_{p,q}} \tag{1.26}$$

where net_q is the sum of all inputs into neuron q from layer P. Simplifying the second term,

$$\frac{\partial \text{net}_q}{\partial w_{p,q}} = \frac{\partial \sum_{p \in P} w_{p,q} o_p}{\partial w_{p,q}} = o_p \tag{1.27}$$

where o_p is the output value from the p'th neuron from layer P. The first term from the chain rule splits as follows (we will call this $-\delta_q$):

$$-\delta_q = \frac{\partial \text{Err}}{\partial \text{net}_q} = \frac{\partial \text{Err}}{\partial o_q} \cdot \frac{\partial o_q}{\partial \text{net}_q} \tag{1.28}$$

The second term is the derivative of the activation function

$$\frac{\partial o_q}{\partial \text{net}_q} = \frac{\partial f_{\text{act}}(\text{net}_q)}{\partial \text{net}_q} = f'_{\text{act}}(\text{net}_q) \tag{1.29}$$

And the first term is

$$\frac{\partial \text{Err}}{\partial o_q} = \sum_{r \in R} \frac{\partial \text{Err}}{\partial \text{net}_r} \frac{\partial \text{net}_r}{\partial o_q} = \sum_{r \in R} \frac{\partial \text{Err}}{\partial \text{net}_r} w_{q,r} \tag{1.30}$$

The derivation comes to a recursive closed form when we see that

$$\frac{\partial \text{Err}}{\partial \text{net}_r} = -\delta_r \tag{1.31}$$

and we come to a final update value for the weight with

$$\Delta w_{p,q} = \eta o_p \delta_q \tag{1.32}$$

where

$$\delta_q = f'_{\text{act}}(\text{net}_h) \cdot \sum_{r \in R} (\delta_r w_{q,r}) \tag{1.33}$$

Thus, to update weight w, we first need to calculate the weight updates in the immediate next layer, and so forth. However, we need a final equation for when the network finally outputs a value to be compared to the true value. If we perform the same chain rule procedure with an L2-norm error function, we derive the final backpropagation formulas.

$$\Delta w_{p,q} = \eta o_p \delta_q \tag{1.34}$$

with

$$\delta_q = \begin{cases} f'_{\text{act}}(\text{net}_q) \cdot (t_q - y_q) (q \text{ outside}) \\ f'_{\text{act}}(\text{net}_q) \cdot \sum_{r \in R} (\delta_r w_{q,r}) (q \text{ inside}) \end{cases} \tag{1.35}$$

where t_q is the true value to fit, and y_q is the output value of the neuron.

1.3.3 Convolutional Neural Networks

There are no limits to the number of neurons in each hidden layer nor the number of layers themselves. As a result, one could design a neural network that could approximate any non-linear and continuous function, as long as enough training data exist. However, it becomes quickly evident that the number of neurons to train blows up significantly if every input vector maps to every next layer of neurons and so forth. While such a neural network (also called a dense neural network) could in theory approximate any function it can represent in finite time [56], the training and predicting with such an architecture becomes practically infeasible.

Neural network approaches to problems typically tend to have input images that are translationally invariant. Meaning, a picture of a dog is a picture of a dog regardless of where the dog is positioned in the image. Therefore, instead of fully connecting every input value to the neural network, it might be prudent to instead train a series of filters that convolute the image at every layer. One immediate advantage of such a network structure is that one could develop a highly-dimensional multi-layer network without blowing up the neuron count that could process an entire image. This architecture which relies on convolution filters instead of direct connections is called a convolutional neural network (CNN).

An example use-case of CNN was demonstrated in 1989 by LeCun et al. [57] where backpropagation algortihm was applied to a convolutional neural network that could classify handwritten digits. The neural network takes an input of 256 units (16-by-16 image) and propagates through three hidden layers, the first two being convolutional layers of kernels, and the last being a fully connected layer. The two convolutional layers allow encoding of the information into lower-dimensional representations which are then parsed by the dense layers.

As the algorithm for updating the weights became highly parallelized, the incorporation

of GPUs during training was developed to great success [58, 59] when a 7 layer CNN trained with 15 million high-resolution images won first place in the ImageNet LSVRC-2010 contest. The resulting neural network consisting of 60 million parameters and 650 thousand neurons could be trained with reasonable time due to efficient GPU implementation.

1.3.4 Residual Layers

One final point that must be mentioned regarding training of CNNs is the vanishing gradient problem that is solved with residual layers [60]. The vanishing gradient problem arises when a neural network with large number of layers eventually fail to update the neural network weights due to the derivative approaching 0 at the earlier layers. This phenomenon can be seen in equation 1.35 where δ_q for inner layers rely on δ_r on the next layers, which is multiplied by a f'_{act} term. For most activation functions, the derivatives tend to be less than 1, so it becomes easily evident that the δ terms quickly vanish to 0 the deeper we backpropagate the errors.

The solution to this problem is by employing residual layers that provide direct connections bypassing the neural network layer, as shown in Figure 1.12. Bypassing a linear term at each layer allows the backpropagation algorithm to ensure that a update value proportional to f'_{act} is applied at each iteration.

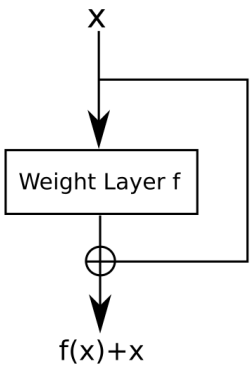

Figure 1.12: Diagram of a residual layer. the input value x bypasses a neural network weight layer f and is directly added to the output.

Chapter 2

Atomic Electron Tomography of Amorphous Solids

2.1 Determining the 3D Atomic Structure of an Amorphous Solid

In this project, we deploy atomic electron tomography to experimentally determine the 3D atomic positions of an amorphous solid. Using a multi-component glass-forming alloy as proof of principle, we quantitatively characterize the short- and medium-range order of the 3D atomic arrangement. We observe that, although the 3D atomic packing of the short-range order is geometrically disordered, some short-range-order structures connect with each other to form crystal-like superclusters and give rise to medium-range order. We identify four types of crystal-like medium-range order—face-centred cubic, hexagonal close-packed, body-centred cubic and simple cubic—coexisting in the amorphous sample, showing translational but not orientational order.

To begin, the samples were synthesized by a carbothermal shock technique with a high cooling rate which was used to create high-entropy-alloy nanoparticles with multi-metal components. Energy-dispersive X-ray (EDX) spectroscopy data show that the nanoparticles are composed of eight elements: Co, Ni, Ru, Rh, Pd, Ag, Ir and Pt. Tomographic tilt series were acquired from seven nanoparticles using an annular dark-field scanning transmission electron microscope operated at 200 kV. To minimize sample drift, four sequential images

per tilt angle were obtained with a dwell time of 3 μs. To monitor any potential damage induced by the electron beam, we took 0° images before, during and after the acquisition of each tilt series and ensured that no noticeable structural change was observed for the seven nanoparticles. The total electron dose of each tilt series was estimated to be between 7×10^5 electrons/$\mathring{A}^2$ and 9×10^5 electrons/$\mathring{A}^2$.

In this study, we focus on the most disordered nanoparticle, from which a tilt series of 55 images was acquired. Although some crystalline features are present the two-dimensional (2D) power spectra calculated from the images show the amorphous halo. After pre-processing and image denoising, the tilt series was reconstructed and the 3D atomic positions were traced and classified. Because the image contrast in the 3D reconstruction depends on the atomic number [12–14], at present AET is only sensitive enough to classify the eight elements into three different types: Co and Ni as type 1; Ru, Rh, Pd and Ag as type 2; and Ir and Pt as type 3. After atom classification, we obtained the 3D atomic model of the nanoparticle, consisting of 8,322, 6,896 and 3,138 atoms of type 1, 2 and 3, respectively. A sample projection, the 2D power spectrum of the 55 projections, two slices of the 3D reconstruction, and a 3D atomic model of the glass-forming nanoparticle is shown in Figure 2.1.

To quantitatively characterize the atomic structure, we employed the local bond orientational order (BOO) parameters to distinguish between the disordered, face-centred cubic (fcc), hexagonal close-packed (hcp) and body-centred cubic (bcc) structures [62]. BOO parameter analysis (shown in Figure 2.2) shows that the local BOO parameters of the majority of atoms deviate markedly from the fcc, hcp and bcc crystal structures, indicating that the atoms are arranged in an amorphous configuration.

Although medium-range order (MRO) in metallic glasses is broadly defined as the nanometre-scale structural organization beyond the short-range order (SRO) [63, 64], in this work we focused on the investigation of the MRO in the framework of the efficient cluster packing

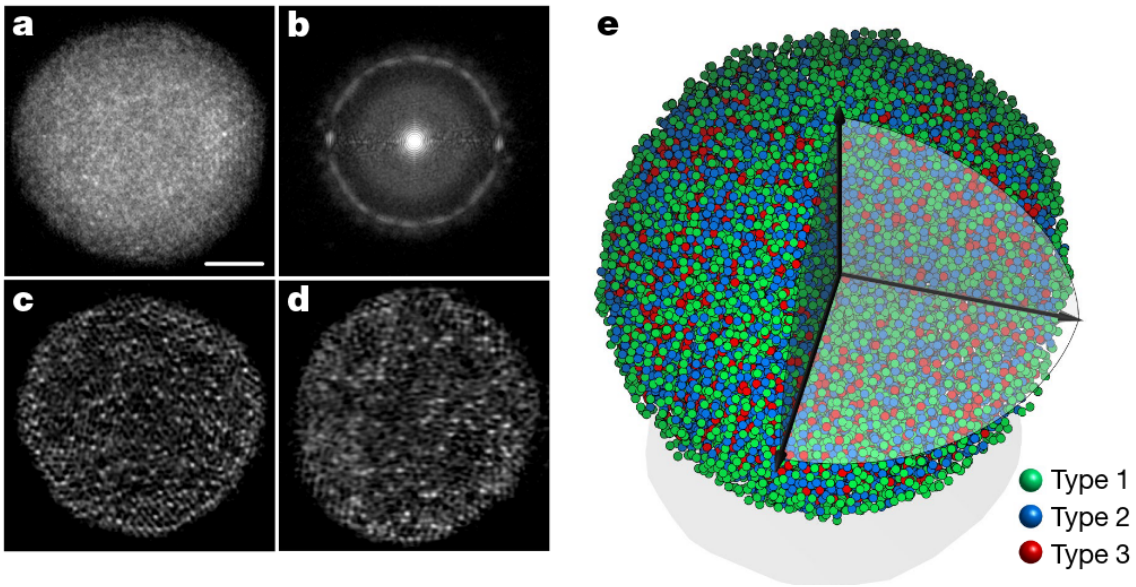

Figure 2.1: Taken from ref [61]. a, Representative experimental image, where some crystalline features are visible. Scale bar, 2 nm. b, Average 2D power spectrum of 55 experimental images (Extended Data Fig. 3), showing the amorphous halo. c, d, Two 2.4-Å-thick slices of the 3D reconstruction in the x–y (c) and y–z (d) plane, where the majority of type-3 atoms (bright dots) are distributed in the second coordination shell. e, Experimental 3D atomic model of the glass-forming nanoparticle.

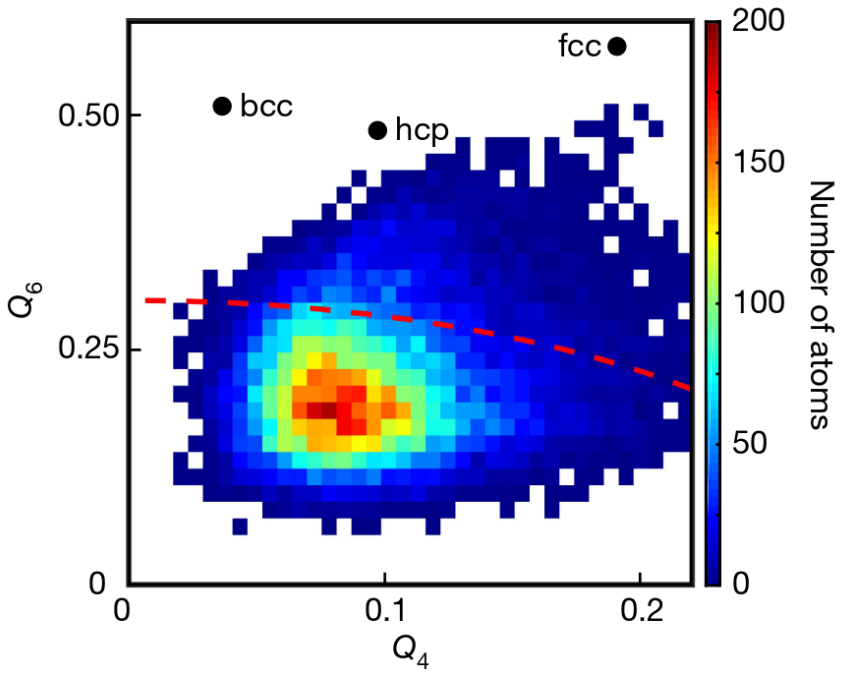

Figure 2.2: Taken from ref [61]. Local BOO parameters of all the atoms in the nanoparticle. According to the criterion of the normalized BOO parameter being ¡0.5 (dashed red curve), 84.54% of the total atoms are disordered.

model [65, 66]. This model hypothesizes that solute atoms are surrounded by randomly positioned solvent atoms to form solute-centre clusters that are densely packed to constitute crystal-like MROs in metallic glasses.

To locate the MRO in the glass-forming nanoparticle, we implemented a breadth-first search algorithm to look for the fcc-, hcp-, bcc-, simple cubic (sc-) and icosahedral-like structures of the solute centres. First, the algorithm identified the solute centres from type-3 atoms based on two criteria: (i) each solute centre must fall within a 0.75 Å radius from an fcc, hcp, bcc or sc lattice point, and (ii) each solute centre must have at least one neighbouring type-3 atom within the second-coordination-shell distance. Second, the identified solute centres were sorted out to generate a queue of the fcc-, hcp-, bcc- or sc-like MRO candidates. Third, starting from the largest MRO candidate (that is, with the most solute centres), each candidate was classified as an MRO if it had at least five or more solute centres and none of the solute centres was already occupied by another MRO. If any solute centres were already occupied, they were removed from the MRO candidate, and the candidate was refitted into the lattice vectors and added back into the queue. If two or more MRO candidates had the same number of solute centres, the one with the smallest error of fitting the solute centres into the lattice vectors was analysed first. This process was repeated until all the MROs were identified, in which each solute centre could only belong to no more than one MRO. Figure 2.3a shows a histogram of the four types of MRO as a function of size, and the inset illustrates the fraction of the solute-centre atoms in the four types of MRO. Figure 2.3b shows the 3D distribution of MROs having eight solute centres or more. To better visualize these MROs, the solute centres are orientated along the fcc, hcp, bcc and sc zone axes (Fig 2.4), showing that the 3D shapes of the MROs are anisotropic.

An attempt was also made to search for icosahedral-like MROs. The breadth-first search algorithm was used to find the MROs that fall within a 0.75 Å radius from the 12 vertices of an icosahedron. Because the icosahedron cannot be periodically packed in three dimensions,

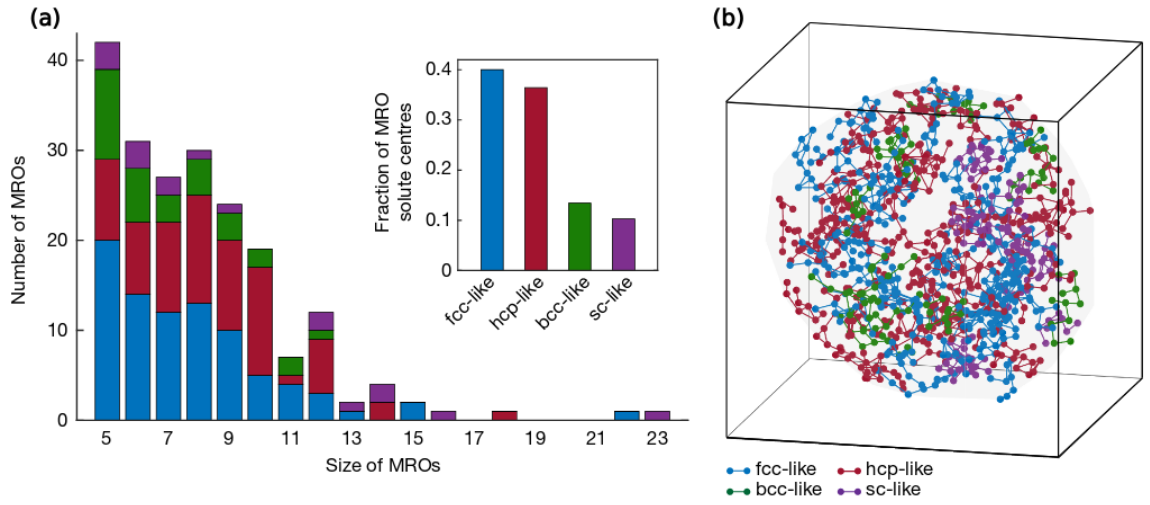

Figure 2.3: Taken from ref [61]. a) Histogram of the four types of MRO—fcc- (blue), hcp(red), bcc- (green) and sc-like (purple)as a function of size (that is, the number of solute centres). The total number of fcc-, hcp-, bcc- and sc-like MROs is 85, 71, 31 and 17, respectively. The inset shows the fraction of solute-centre atoms in the four types of MRO. b) Distribution of the four types of MRO with eight solute-centre atoms or more, where the central region lacks large MROs.

only the nearest-neighbour vertices were searched, making the largest possible MRO have 13 solute centres (one central solute centre plus 12 nearest neighbours). After performing the search, the resulting possible MROs had a mean value of 3.9, meaning that on average each solute centre was connected to only three others when constrained to an icosahedron within the second coordination shell. Furthermore, although the largest possible MRO had seven solute centres, none of these solute centres formed five-fold symmetry. We also repeated this analysis with a 1Åradius cut-off; the mean value of solute centres became 4.5, the largest MRO had eight solute centres, and there were 19 five-fold symmetries.

We calculated the partial PDFs of all the fcc-, hcp-, bcc- and sc-like solute centres in the glass-forming nanoparticle, and their corresponding maximum peak positions are at 4.62, 4.77, 4.82 and 3.88 Å, respectively (Fig 2.5). These peak positions represent the average nearest-neighbour distances of the solute centres in the four crystal-like MROs, and the

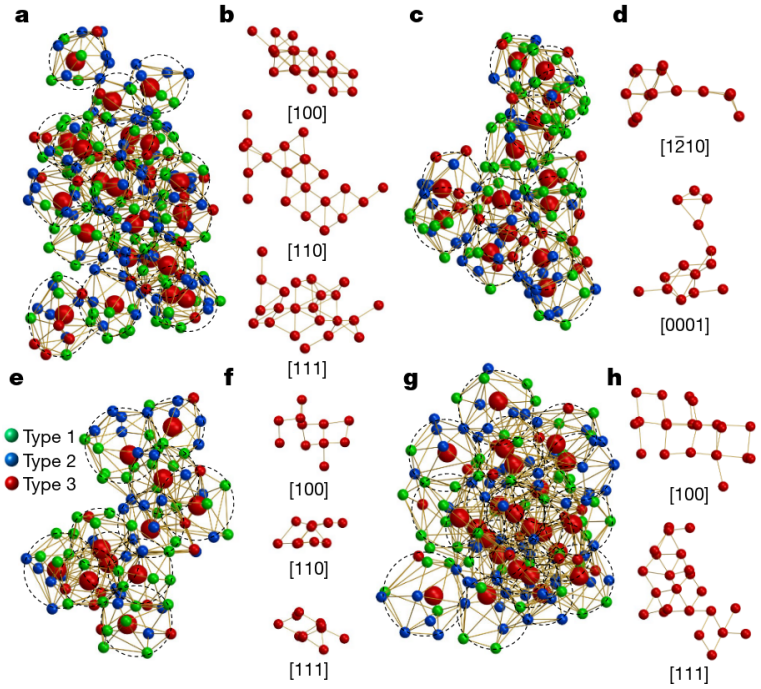

Figure 2.4: Taken from ref [61]. Representative fcc- (a), hcp- (c), bcc- (e) and sc-like (g) MROs, consisting of 22, 14, 11 and 23 solute centres (large red spheres), respectively, where the individual solute-centre clusters (dashed circles) are randomly oriented. To better visualize the crystal-like MROs, the solvent atoms have been removed and the solute centres are orientated along the fcc (b), hcp (d), bcc (f) and sc (h) zone axes, showing that the MROs have anisotropic 3D shapes and strongly deviate from the crystal lattices.

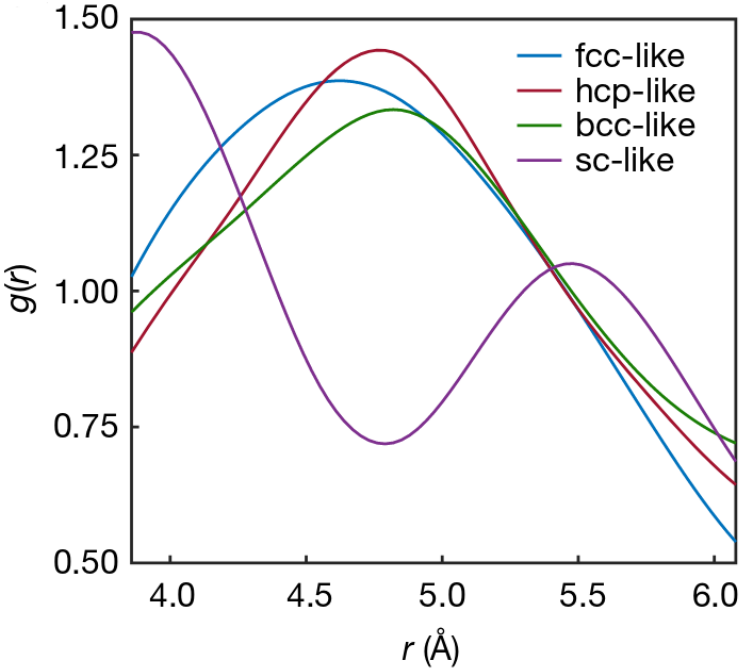

Figure 2.5: Taken from ref [61]. Partial PDFs of the fcc-, hcp-, bcc- and sc-like solute centres in the glassforming nanoparticle, where the maximum peak positions are located at 4.62, 4.77, 4.82 and 3.88 Å, respectively.

broadened peaks signify the severe deviation from the crystal lattices. Compared with the other three partial PDFs, the partial PDF of the sc-like MROs has two peaks, and the ratio of the second to the first peak position is about 2 (Fig 2.5, purple curve), which corresponds to the ratio of the diagonal to the side length of a square. The shorter nearest-neighbour distance of the sc-like MROs compared to the other three crystal-like MROs indicates that the sc-like solute-centre clusters are more closely connected with their neighbours.

Our quantitative analysis of the SRO and MRO in a multi-component glass-forming nanoparticle provides direct experimental evidence to support the general framework of the efficient cluster-packing model [65, 66]; that is, solute-centre clusters are densely packed in some parts of the sample to form crystal-like MROs. Through AET, we observed coexistence of fcc-, hcp-, bcc- and sc-like MROs in the glass-forming nanoparticle. By quantifying their length, volume and 3D structure, we found that the MRO not only has a large variation in length and volume, but also severely deviates from the crystal lattices.

2.2 3D Atomic Packing in Amorphous Solids with Liquid-Like Structure

In 1952, Frank hypothesized that icosahedral order is the prevalent atomic motif in monatomic liquids [67]. A polytetrahedral packing model has been proposed to explain the three-dimensional (3D) atomic structure of monatomic liquids and amorphous materials [68], however, no experimental method could directly determine the 3D atomic packing of liquids and amorphous materials due to the lack of long-range order.

In this project we advance AET to reveal the 3D atomic structure of an amorphous Ta thin film and two amorphous Pd nanoparticles that are not metallic glasses but have a liquid-like structure. We observed that pentagonal bipyramids are the main atomic motifs in the monatomic amorphous materials. Instead of assembling icosahedra, most pentagonal

bipyramids closely connect with each other to form pentagonal bipyramid networks (PBNs) that extend to the medium-range order.

The AET experiments were conducted with a scanning transmission electron microscope (STEM) in annular dark-field (ADF) mode. Tomographic tilt series were acquired from an amorphous Ta thin film and two Pd nanoparticles, which were synthesized by physical vapour deposition and colloidal chemistry with ligand engineering, respectively.

Tomographic tilt series of the Ta thin film and Pd nanoparticles were acquired using the TEAM I microscope at the National Center for Electron Microscopy. The microscope was operated in ADF mode with an electron energy of 300 keV. A low-exposure acquisition scheme was adopted for our data acquisition34. When measuring an image at a tilt angle, a nearby nanoparticle or a feature in the Ta film was used to align and focus the image, thus reducing the unnecessary electron dose to the sample under study. At each tilt angle, three sequential images were taken with a dwell time of 3 μs to minimize the dose rate and drift distortion in each image. To further mitigate the beam damage to the samples, the total electron dose of each tilt series was optimized to be about 3.4–6.0 $\times 10^5$ electrons/Å2.

After image preprocessing, each tilt series was reconstructed by tomography, and the 3D coordinates of individual atoms were traced and refined to produce an experimental atomic model. Figure 2.6a,b show the experimental 3D atomic model of the Ta thin film and two Pd nanoparticles (named Pd1 and Pd2), respectively. To quantify the disorder, we calculated the bond orientational order parameters for all the atoms [62]. We observed that 20.1%, 2.2% and 1.8% of the atoms form crystal nuclei on the surface of the Ta thin film and the Pd1 and Pd2 nanoparticles, respectively (grey atoms in Fig 2.6a,b,)

After excluding these nuclei, we plot the PDFs of the disordered atoms (Fig 2.7), which exhibit similar shapes despite the different chemical composition and synthesis methods of the samples. As a comparison, the PDF of a Ta liquid obtained by the molecular dynamics (MD) simulation is shown as a dotted curve in Fig 2.7, in which the peak and valley positions

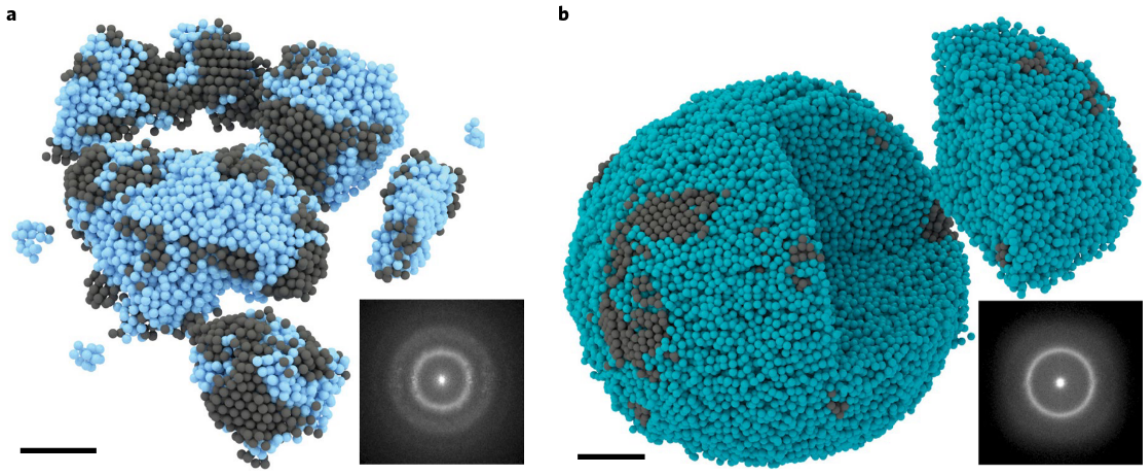

Figure 2.6: Taken from ref [69]. a,b, Experimental 3D atomic model of an amorphous Ta film (a) and a Pd nanoparticle (Pd1; b) with surface crystal nuclei in grey. The two insets show the average 2D power spectra of the experimental images for the Ta film and Pd1 nanoparticle, where the amorphous halo is visible. Scale bars, 2 nm.

agree with those of the Ta thin film and two Pd nanoparticles.

To quantitatively characterize the tetrahedra in the amorphous materials (Methods), we used the distortion parameter, defined as $\delta = e_{\max}/e_{\text{avg}} - 1$, where $e_{\max}$ and e_{avg} are the maximum and average edge lengths of each tetrahedron, respectively. With $\delta > 0.2$, more than 96.8% of the atoms in the samples form tetrahedra. By sharing faces, these tetrahedra constitute four main motifs: triplets, quadrilateral, pentagonal and hexagonal bipyramids (Figure 2.8). The four motifs are represented by three-, four-, five- and six-fold skeletons, which are formed by connecting the centroids of the tetrahedra (coloured lines in Figure 2.8). Analysis shows that of the population of the four motifs in the three samples, pentagonal bipyramids are the most abundant atomic motif. This observation can be explained by the fact that the atomic packing in pentagonal bipyramids requires less distortion than in the other motifs [68].

Since a tetrahedron and a pentagonal bipyramid represent the densest packing of four and seven atoms, respectively, we correlated polytetrahedral packing with the local mass

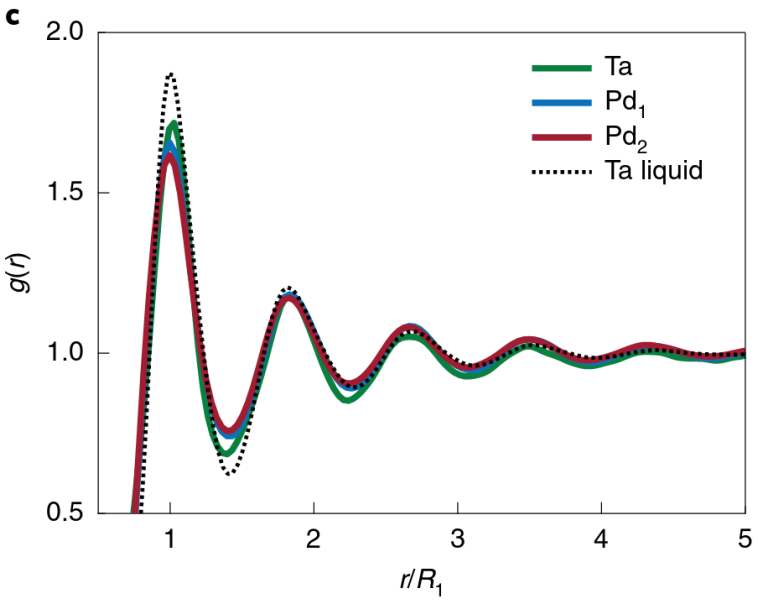

Figure 2.7: Taken from ref [69]. PDFs of the Ta film (green), two Pd nanoparticles (Pd1 in blue and Pd2 in red) and an MD-simulated Ta liquid at 5,200 K (dotted curve), where g(r) is the PDF and r/R1 is the distance (r) normalized by the first peak position (R1).

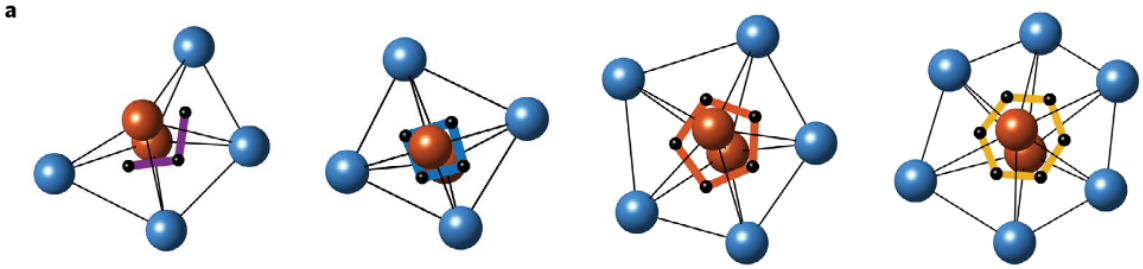

Figure 2.8: Taken from ref [69]. The four most populated atomic motifs (triplets, quadrilateral, pentagonal and hexagonal bipyramids) in the three samples, where the capping atoms are in brown and the ring atoms are in blue, connected by the bonds (black lines). The four motifs are represented by a three- (purple), four- (blue), five- (orange) and six-fold (yellow) skeleton, which connects the centroid (black dot) of each tetrahedron.

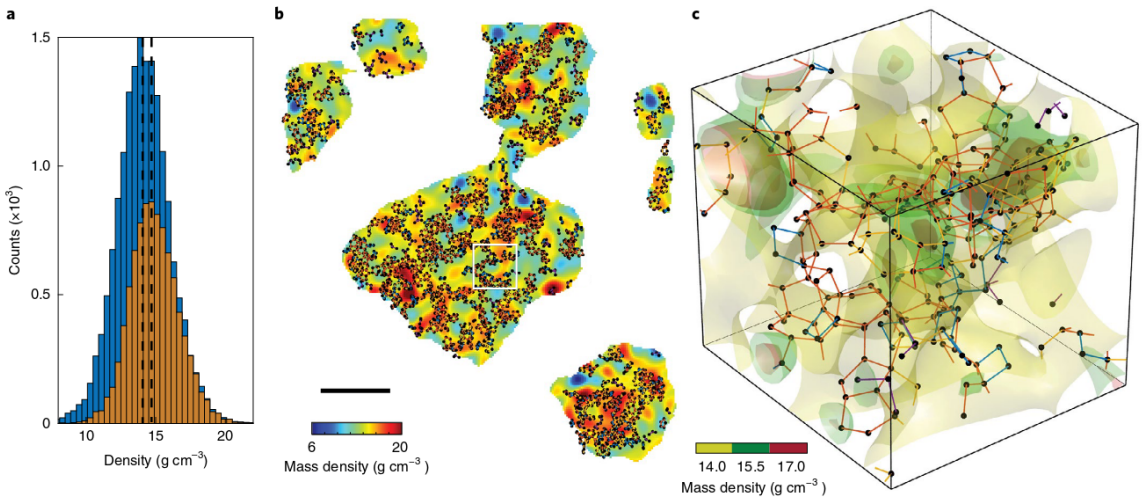

Figure 2.9: Taken from ref [69]. a, Mass density distribution in the regions of the amorphous Ta film with (orange) and without (blue) polytetrahedral packing, where polytetrahedral packing increases the average mass density by 3% (dashed lines). b, A slice through the Ta film shows the local mass density heterogeneity (colour) overlaid with polytetrahedral packing (black). Scale bar, 2 nm. c, A 3D surface rendering of local mass density heterogeneity magnified from the square region in b, which is overlaid with three- (purple), four- (blue), five- (orange) and six-fold (yellow) skeletons.

density of the amorphous materials. Figure 2.9a shows the mass density distribution in the regions of three samples with and without polytetrahedral packing, where the average mass density increases with polytetrahedral packing. We also observed 3D local mass density heterogeneity in the amorphous materials. A slice through each sample shows the local mass density heterogeneity overlaid with polytetrahedral packing (Figure 2.9b). A magnified region in each sample reveals that 3D local mass density heterogeneity is strongly correlated to the atomic packing of the four motifs in the three samples (Figure 2.9c).

We observed that a number of pentagonal bipyramids link to each other by sharing four or five atoms with their neighbours (Figure 2.10a,b), which we define as vertex or edge sharing of the five-fold skeletons, respectively. We found that while 63.5% of pentagonal bipyramids in the three samples do not share any vertex with their neighbours, the majority of them (72.5%) have at least one edge-sharing neighbour.

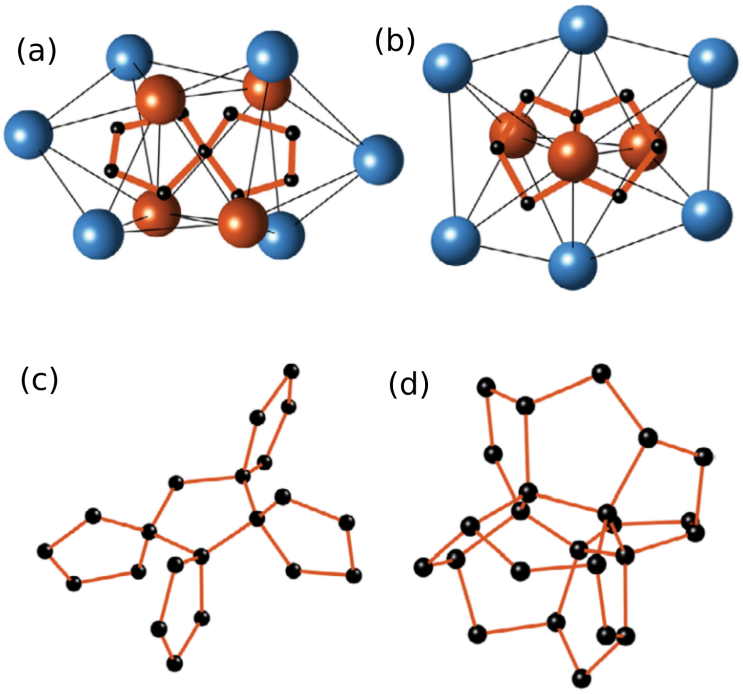

Figure 2.10: Taken from ref [69]. a,b, Vertex and edge sharing of the five-fold skeletons, which share four and five atoms with their neighbours, respectively. The capping atoms are in brown. c,d, Pentagonal bipyramid clusters with the most vertex- (c) and edge-sharing neighbours (d) in the three amorphous materials.

Figure 2.10c,d shows two pentagonal bipyramid clusters with the most vertexand edge-sharing neighbours, respectively, where the larger cluster is formed by edge sharing. These results indicate that edge sharing of the five-fold skeletons is a more dominant feature in the packing of pentagonal bipyramids. We then investigated if these pentagonal bipyramids form icosahedra. An icosahedron requires the tight packing of 12 pentagonal bipyramids by edge sharing. Due to geometric frustration [68], the five-fold skeletons in an icosahedron form a regular dodecahedron with the dihedral angle of 116.57°. But we observed that the dihedral angles between two edge-sharing pentagonal bipyramids peak at 120.7° in the three amorphous samples, a value that is close to 120° in the absence of geometric frustration. This observation further confirms that the pentagonal bipyramids assemble only partial icosahedra.

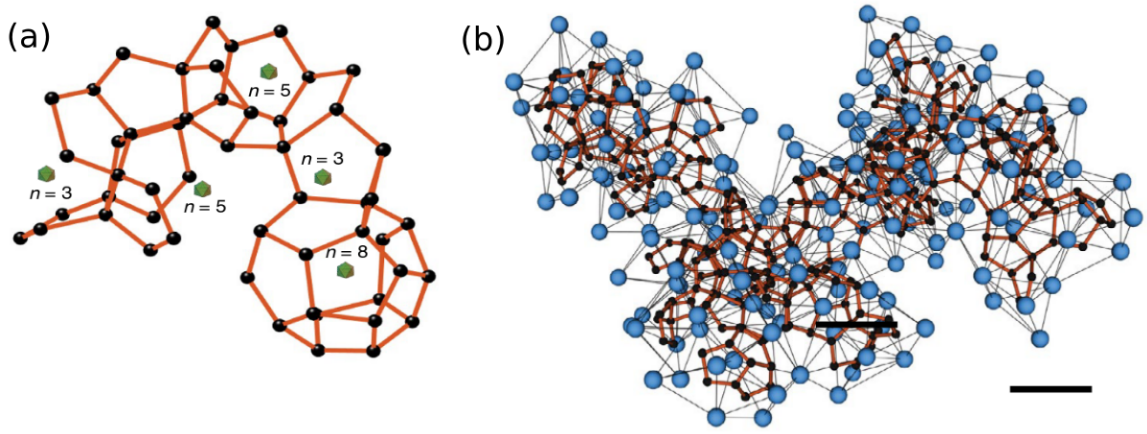

Figure 2.11: Taken from ref [69]. a, Five-fold skeleton of a representative PBN, containing five partial icosahedra with n = 3, 5, 5, 3 and 8. The centre of each partial icosahedron is labelled by an icosahedron (green). b, The largest PBN in the amorphous materials, containing 135 pentagonal bipyramids formed by 165 Ta atoms (blue balls). The black lines represent the bonds between the Ta atoms and the black dots and orange lines show the PBN with five-fold skeletons. Scale bar, 4 Å.

Instead of assembling icosahedra, most pentagonal bipyramids with edge-sharing skeletons form PBNs in the amorphous samples. Figure 2.11a shows a representative PBN comprising five partial icosahedra, and the largest PBN was found in the Ta thin film, consisting of 135 pentagonal bipyramids formed by 165 atoms with an end-to-end length of 2.83 nm (Figure 2.11b).

Our experimental results provide fundamental insight into the 3D atomic packing of monatomic amorphous materials and liquids. Although there are several crystal nuclei on the surface of the Ta thin film and two Pd nanoparticles, we verified that the crystal nuclei have a minimal impact on the structural disorder of the samples based on the following observations. After removing the crystal nuclei, the peak and valley positions of the PDFs of the amorphous materials are in good agreement with those of the MD-simulated Ta liquid (Figure 2.7).

Our experimental and MD simulation results further reveal that pentagonal bipyramids are the prevalent atomic motif and form nanometre-scale PBNs in monatomic amorphous

materials and liquids. During the quench from a liquid to a metallic glass state, the PBNs quickly grow in size and assemble more icosahedra, indicating the PBNs can trap the atoms in a local minimum (a non-crystalline or glass state) instead of a global minimum (a crystalline state). Looking forward, the ability to determine the 3D atomic structure of amorphous thin films is expected to greatly expand the applicability of AET to a broad range of technologically relevant materials such as quasicrystals.

Chapter 3

Incorporating Phase Projections in Atomic Electron Tomography

In this chapter, we will demonstrate the combination of ptychography and atomic electron tomography as an effective method for low dose imaging of individual low-Z atoms in three-dimensions, and argue for the utility of incorporating coherent electron images in tomography for imaging samples in the physical sciences.

3.1 Ptychographic Atomic Electron Tomography

All projection measurements in AET so far have been limited to incoherent electron scattering methods such as (High Angle) Annular Dark Field (ADF/HAADF), whereas imaging methods from coherent beam sources have made significant strides in recent decades – including ptychography. Conventionally, ptychography is an imaging method in 2 dimensions, but it has also been extended to recover 3D vertical information from a single projection by modeling the object as multiple slices of phase objects [70]. However, the vertical resolution of this method is limited by $dz = \lambda/(2\sin^2(\theta_{\max}/2))$ compared to the lateral resolution limitation of $dx = \lambda/\sin(\theta_{\max}$, where λ is the wavelength of the probe and $\theta_{\max}$ is the maximum detector collection angle [71]. As an example, a STEM experiment with an 80 keV electron

probe and 80 mrad maximum collection angle would yield a theoretical maximum vertical resolution of 1.3 nm compared to its lateral counterpart of 0.39 Å. Since atomic resolution typically requires near-ångstrom resolution, single-projection multi-slice ptychography alone would not be able to recover signals from individual atoms in 3D. Due to this limitation, ptychographic data has to be measured at multiple tilt angles and combined tomographically to achieve high vertical resolution in 3D phase images [36].

As powerful single-electron pixel-array detectors that can achieve deep sub-Ångstrom resolution become commercially available, ptychographic AET (pAET) has evolved from a hypothetical idea to an experimental possibility. We use numerical experiments to demonstrate pAET as a feasible method for low-dose 3D imaging of individual light atoms by performing multislice simulations on a zinc oxide nanoparticle and a vertical WS2/WSe2 van der Waals (vdW) heterostructure.

3.2 Simulation of pAET on ZnO Nanoparticle

An atomic model of a spherical Wurtzite ZnO nanoparticle with a diameter of 5 nm was generated with 1% randomly dispersed oxygen vacancies. This model was used to generate 0.2-Å-thick projected potential slices for multislice simulation (probe sampling: 0.1 Å; sample sampling: 0.1 Å; maximum detector angle: 104 mrad) [72, 73] using tabulated Hartree-Fock approximations of atomic potentials [11, 74]. An aberration-corrected STEM probe was simulated using the parameters of the TEAM 0.5 (electron energy: 80 keV; probe semi-convergence angle: 30 mrad; C1: 0 nm; C3: 900 nm; C5: -622 µm; probe size: 0.88 Å; probe step size: 0.4 Å). By tilting the ZnO nanoparticle from -70° to +70°, we calculated 29 data sets of diffraction patterns at 29 tilt angles. The reason that we chose a small tilt range was to demonstrate that pAET can tolerate a large missing wedge. For each tilt angle, eight frozen phonon configurations at room temperature were obtained and averaged for the

data set. Unlike conventional AET tilt series where each tilt image is a 2D ADF or HAADF image, a ptychographic tilt series consists of 4D data sets: 2 dimensions from scanning the sample in real space, and 2 from the diffraction in momentum space. To represent real experimental conditions, the 4D data sets were corrupted with simulated Poisson noise by sampling independent electron events from the ideal diffraction patterns. The diffraction patterns were sampled to represent an effective dose per projection of $3.0 \times 10^4 e^-/\mathring{A}^2$, thus simulating a total electron radiation dose of $8.7 \times 10^5 e^-/\mathring{A}^2$. Figure 3.5(a) shows a 2D logarithmic heat map of the average diffraction pattern from one tilt series projection.

The overlap of adjacent probes (44%) gives enough redundancy to iteratively solve for both object and probe functions using the extended ptychographic iterative engine (ePIE) algorithm [30]. Due to high Poisson noise in individual diffraction patterns, we used a small update parameter (0.01) for the object function to prevent overfitting the noise in reconstructions. Figure 3.5(b) shows an example of a ptychographic phase reconstruction on the left hand side, and the ADF image (detector inner angle: 30 mrad; outer angle: 90 mrad) generated from the same diffraction data set on the right hand side. 3D reconstructions of both ptychographic and ADF projections were performed using REal Space Iterative Reconstruction (RESIRE) – a powerful tomographic algorithm that iteratively minimizes the error between the measured and calculated projections using the gradient descent [61].

Isosurface rendering of the 3D reconstruction performed with ptychographic projections are shown in Figure 3.5(c), with a magnified isosurface rendering of the volume's core in Figure 3.5(d). 3D rendering of individual oxygen atoms can be observed in the magnified figure. Figure 3.5(e) shows 2.0-Å-thick central slice of the 3D volume reconstructed with ptychographic phase projections on the left hand side, and with ADF-STEM projections on the right hand side, with a magnified image shown in Fig. 3.5(f). The smaller blobs found in the ePIE/RESIRE reconstruction correspond to individual oxygen atoms, which were not resolved using ADF/RESIRE. Because RESIRE does not assume any periodicity while

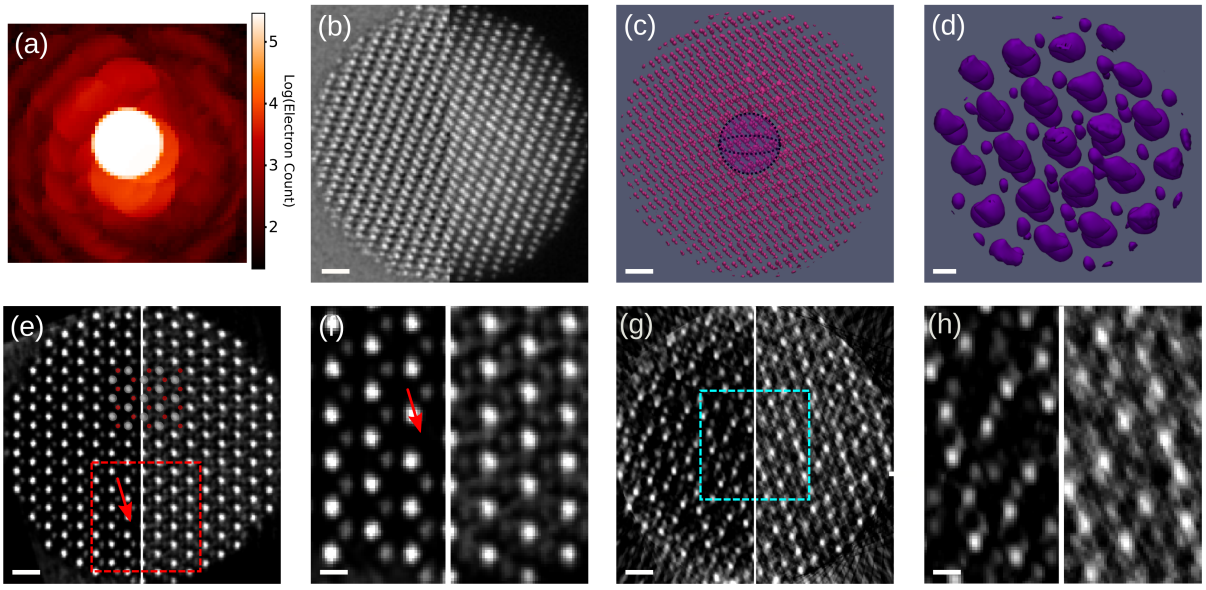

Figure 3.1: Taken from [1]. Numerical experiment on ptychography- and ADF-STEM-based atomic electron tomography of a 5 nm ZnO nanoparticle. (a) 2D logarithmic heat map of the average of 15625 diffraction patterns from a tilt series. The diffraction patterns simulate a dose per projection of $3.0 \times 10^4 e/\mathring{A}^2$. (b) A representative ptychographic phase projection (left) and an ADF-STEM projection (right). Both projections were reconstructed with the same diffraction patterns seen in (a). (c) 3D isosurface rendering of the volume after tomographic reconstruction of ptychographic phase projections. (d) Magnified isosurface rendering of the core of the volume in (c). Individual oxygen atoms rendered as smaller spheres are observed. Note that the zinc atoms look disproportionate large due to the isosurface rendering effect. (e) A 2.0-Å-thick central slice of the volume in the [0001] direction reconstructed with ptychographic phase projections (left) and with ADF-STEM projections (right). (f) A magnified image of (e), indicating an oxygen defect with a red arrow. (g) A 2.0-Å-thick slice through the missing wedge direction of the reconstruction performed with ptychographic projections (left) and with ADF projections (right). (h) A magnified image of (g), demonstrating greater missing wedge artifacts in the reconstruction when performing tomography with ADF-STEM projections. Scale bars in (b), (c), (e) and (g) indicate 5 Å, scale bar in (e) indicates 1 Å, and scale bars in (f) and (h) indicate 2 Å.

performing 3D reconstruction, individual oxygen atom defects placed in the original model were also able to be resolved in the pAET reconstruction as indicated by the red arrow in Fig. 3.5(f). Furthermore, the improved quality in pAET reconstructions is especially evident when looking at a slice through the missing wedge direction, shown in Fig. 3.5(g), and magnified in Fig. 3.5(h).

3.3 Z-Contrast in pAET

To better understand the nature of individual low-Z atom contrast with pAET, we performed multislice simulations of individual atoms with varying atomic number to measure their relative contrast. Similar calculations have been performed analytically for phase contrast [75], and numerically for ADF-STEM and bright field STEM contrast [11]. Multislice simulations of individual atoms ranging from C (Z = 6) to Xe (Z = 54) were performed with an aberration-corrected probe (semi-convergence angle: 24 mrad) of varying energies (60 keV, 120 keV, and 200 keV) and probe step size of 0.4 Å. Ptychographic phase projections were reconstructed with ePIE and ADF-STEM projections were reconstructed by integrating the diffraction patterns from 24 mrad to 120 mrad. Contrast per atom was defined as the height of the fitted 2D Gaussian function and were normalized by setting the atomic contrast of Xe to unity. Finally, projected atomic potentials were calculated and fitted as 2D Gaussians to measure ideal relative atomic contrasts. The results of these calculations are shown in Fig. 3.2.

Note that the contrasts from the potentials are not strictly monotonic as a function of atomic number, as variations in the filling of electron shells create fluctuations reflected by the Hartree-Fock approximations. We found that the contrast from ADF-STEM was monotonic as a function of the atomic number, supporting previous studies that report a power relation (roughly proportional to Z1.8) [12–14]. Such a relation can be tolerated while conducting

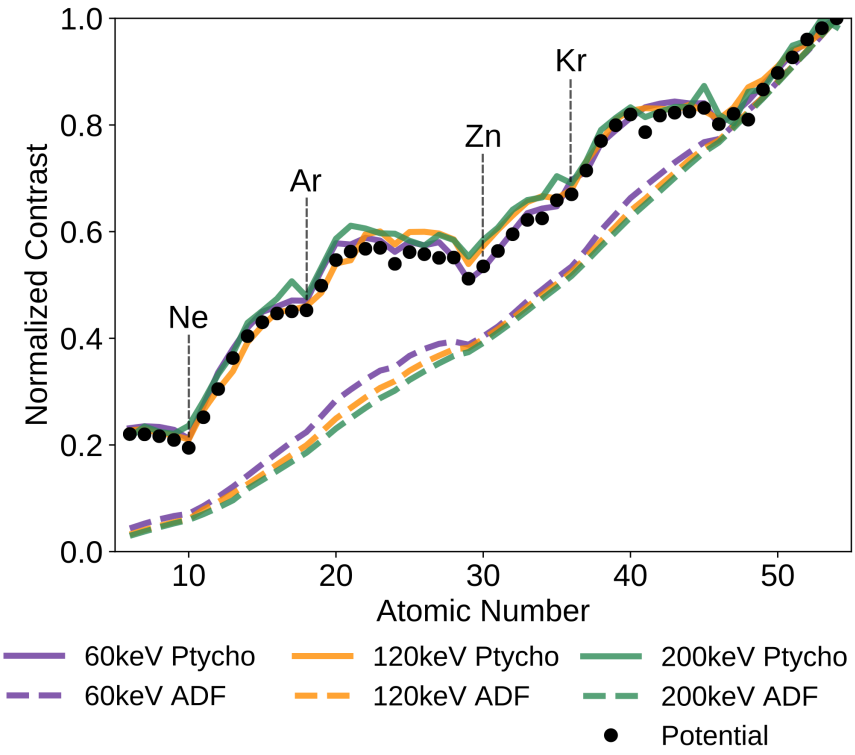

Figure 3.2: Taken from [1]. Quantitative comparison of normalized atomic contrast from C to Xe between ptychography and ADF for three different electron beam energies (60 keV, 120 keV, and 240 keV). Normalized peak heights from Hartree-Fock potential estimations are also plotted. Atoms where valence shells become fully filled (Ne, Ar, Zn, and Kr) are indicated to explain fluctuations in the potential peaks.

tomography in metallic samples, but this makes it more difficult to simultaneously image individual low and high Z atoms, such as in metallic oxides, due to a large ratio of the relative signal ($S_{Zn}/S_O = 7.75$). In contrast, ptychography was not only able to recover a higher relative signal for low-Z atoms compared to ADF-STEM ($S_{Zn}/S_O = 2.48$), but also was sensitive enough to recover the aforementioned fluctuations in atomic potentials for all three beam energies. This advantage of using phase signals for tomography has also been numerically demonstrated with multi-slice simulation of high-resolution transmission-electron microscopy images as input projections [76].

3.4 Linearity in Projections

Successful tomography requires that the input projections need to be a sum of some monotonic response to a physical property along the direction of projection – a requirement aptly named as the projection requirement [13, 77]. Although perfect linearity is ideal, power law atomic contrast in ADF/HAADF-STEM projections is sufficient to locate and identify individual atoms in materials, but it requires a relatively high electron dose. In contrast, ptychography reconstructs the phase induced in the transmitted beam by the Coulomb potential, which acts as a linear response to atoms. Therefore, we should expect a higher degree of linearity in AET projections reconstructed with ptychography.

Figure 3.3(a) shows the graphical representation of the multislice simulation used to test linearity in ADF/HAADF-STEM images. N Si atoms separated by a distance d = 3 Å were placed co-linearly in the path of the electron beam (energy: 80 keV; semi-convergence angle α: 24 mrad), and was sampled with probe step size of 0.4 Å. Three different projections of atomic columns were calculated using five different inner and outer ADF detector angle combinations. After fitting each projection of atom columns to a 2D Gaussian, the contrast per atom was measured as the height of the 2D Gaussian divided by the number of atoms.

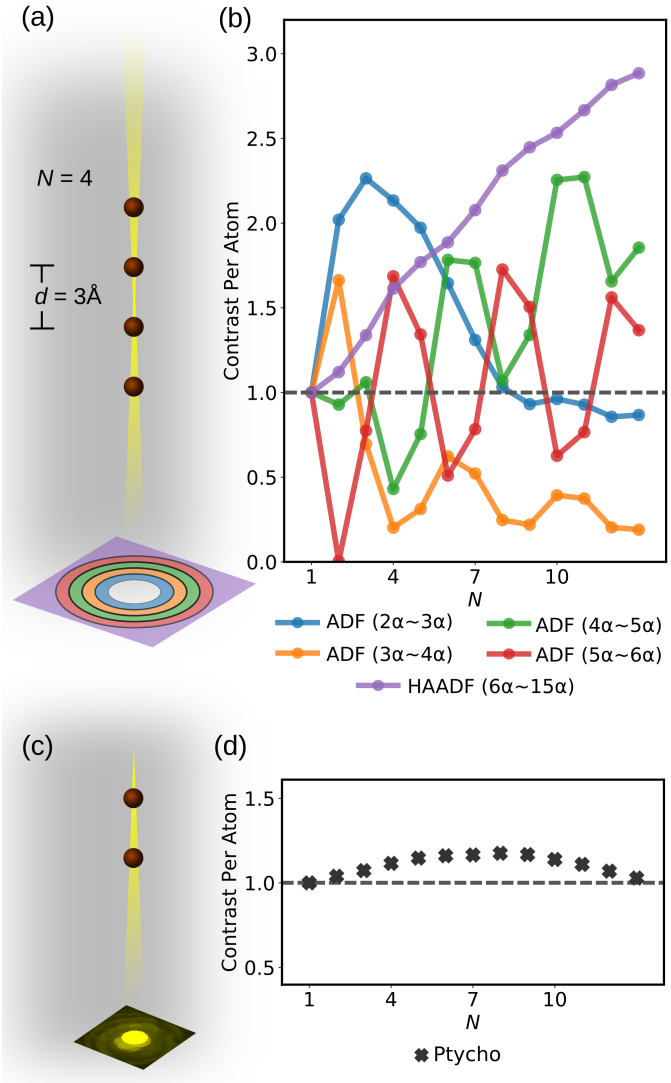

Figure 3.3: Taken from [1]. Contrast per atom when imaging along the zone-axis with ADF/HAADF and ptychography. (a) An experimental schematic of multislice simulation of an 80 keV electron beam probe and 24 mrad semi-convergence angle imaging a column of N Si atoms separated by distance of 3 Å. Five different combinations of inner and outer angles were simulated. (b) Contrast per atom in atomic columns when measured with various ADF angles. The horizontal gray line plotted at 1 indicates the ideal linear projection. (c) Similar experimental schematic as (a) except the substitution of ADF detectors with a pixel array detector, allowing for ptychographic reconstruction. (d) Contrast per atom in columns when reconstructed with ptychography.

The result of this calculation is shown in Fig. 3.3(b), along with a horizontal line at 1 indicating contrast per atom from ideal linear projections. We observed significant deviation from the linear contrast when the collection angles were smaller than 6α as channeling effects due to low electron beam energy distorted the images of columns.

Furthermore, the measured atomic contrast was highly sensitive to the collection angles in this regime due to multiple scattering. Only when the measured inner angle was higher than 6α did we observe monotonicity in the contrast per atom. This might suggest HAADF-STEM as a more suitable method than ADF-STEM when performing AET using low energy beams, but the electron dose in real HAADF-STEM experiments required to sufficiently overcome Poisson noise makes it a less reliable choice when imaging dose sensitive materials. For a comparison, we performed a similar numerical experiment by replacing the ADF detector with a pixel array detector for ptychography, shown in Fig. 3.3(c), and plotted results from contrast calculations in Fig. 3.3(d). Ptychography was able to maintain a more linear contrast per atom than ADF/HAADF-STEM.

3.5 pAET in 2D Materials

Lastly, the method of pAET can also be used to image individual low-Z atoms in geometries other than nanoparticles such as thin films and vdW heterostructures. The vdW interaction between the top and bottom layers mediates various types of coupling across the interface in vdW heterostructures. It has been reported that the quantum properties of vdW heterostructures are highly tunable with vertical stacking through moiré potentials, and that different stacking alignments can cause drastic changes in exciton excitation as well as other quantum properties such as superconductivity and correlated insulator states [78–82]. The capability to precisely determine the 3D coordinates and chemical species of individual atoms combined with ab initio calculations is anticipated to reveal unprecedented details about the

correlation between the atomic structure of vdW heterostructures and their exotic quantum properties.

Using identical electron probe parameters as those used in Fig. , multislice simulation was performed on an atomic model of vertical WS2/WSe2 van der Waals heterostructure, with its two tungsten layers separated by 6.54 Å and tiled by 12.5°. Figure 3.4(a) shows 2.0-Å-thick slices of every atomic layer of the heterostructure when reconstructed with ePIE and RESIRE. Despite severely corrupting the diffraction patterns by simulating a total electron dose of $5.1 \times 10^4 e/\mathring{A}^2$, every atomic layer including the two sulfur layers are resolved. However, when tomography was performed on ADF-STEM projections with an equivalent electron dose, as shown in Fig. 3.4(b), the signals from the sulfur layers attenuate below a level at which individual atoms are traceable.

3.6 Mn3O4 Nanowire pAET Experiment

An experiment demonstrating pAET was attempted on a ptychographic tilt-series dataset of a Mn3O4 nanowire. The sample was loaded on a tomographic single-tilt holder (Fischione Model 2000) and a tilt series from -70 to 70 degrees at 2 degree increments was collected. The imaging was performed with electrons at 300 kV, with convergence angle at 11.5 mrad, scan step size of 0.822 Angstroms (10Mx, 256x256 pixels, FOV of 210-by-210 Å2), and an estimated defocus of -60 nanometers. The defocus was a nominal value from the microscope, with the focal point of the probe lying between the sample and the detector. During the tilt-series acquisition, the nanowire experienced mild damage which could result in errors during the alignment process. The before-after image of the nanowire during the experiment can be found in Figure 3.5.

Due to the large defocus of the probe, the illumination function could not be approximated as a single coherent source, but rather as a decoherent source modeled as 8 decompo-

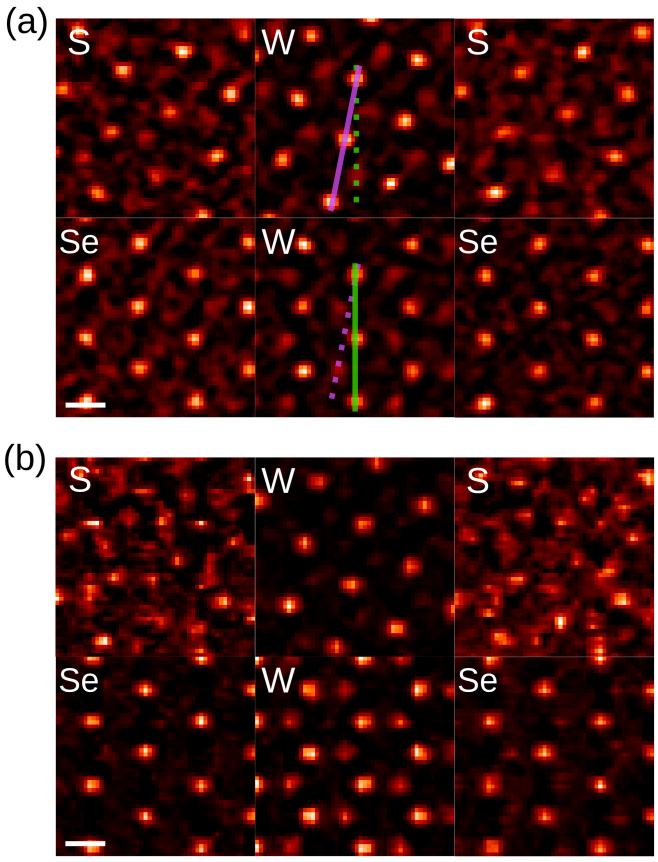

Figure 3.4: Taken from [1]. Numerical experiment on ptychography- and ADF-STEM-based AET of a vertical WS2/WSe2 van der Waals heterostructure. (a) 2.0-Å-thick slices of each atomic layer when simulated with a total electron dose of $5.1 \times 10^4 e/\mathring{A}^2$ and reconstructed with ePIE and RESIRE. Tilt angle of 12.5° between the two tungsten layers was recovered, indicated by colored lines. (b) 2.0-Å-thick slices of the same experiment as (a) when reconstructed with ADF and RESIRE. Scale bars, 2 Å.

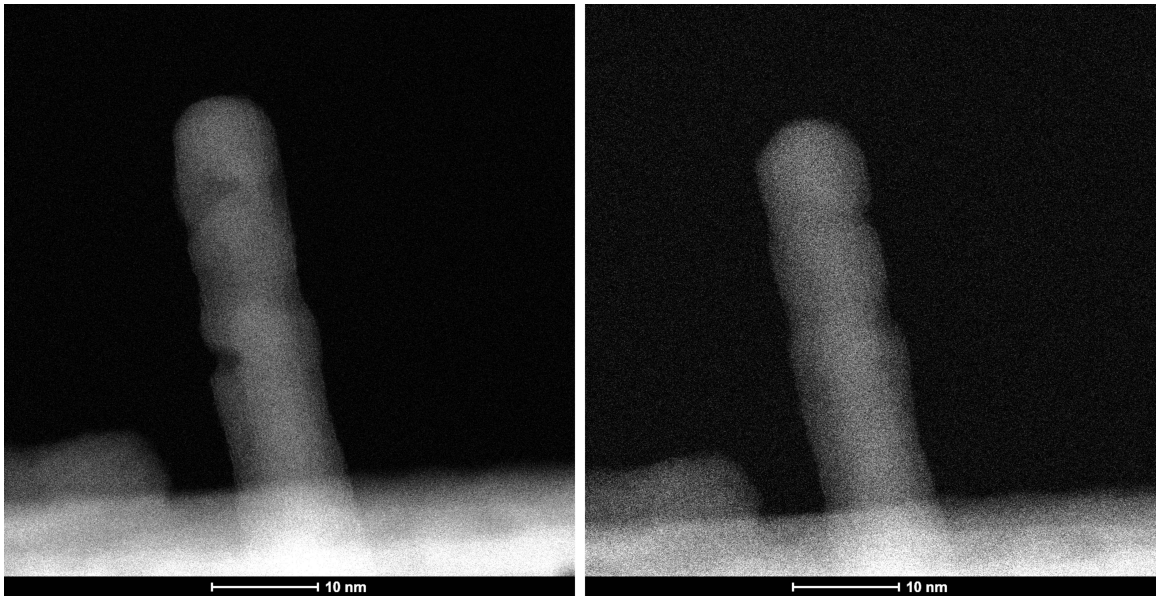

Figure 3.5: Before and after images of the Mn3O4 nanowire during the pAET experiment. Mild damage along the sides of the nanowire can be observed after the tilt series.

sitions. Multiple-probe ptychography was therefore employed with alpha of 0.1 and beta of 0.1. Since we observed sample damage along the sides of the sample, the tip of the nanowire was used to reconstruct ptychographic projections, as shown in Figure 3.6. Despite the probe function having diameters in the order of nanometers, multiple-probe ptychography was able to resolve fine atomic features within the projections.

To align the projections, the background behind the nanowire was first subtracted by approximating the phase values at the boundary as a Laplace surface. Then, since the tilt direction is near parallel to the direction of the nanowire, the projection of the nanowire along the perpendicular direction (also called the common line) should be consistent throughout the tilt-series. This was used to first align the nanowire along the parallel direction by minimizing the error between the common lines. After this alignment, the location of the axis of rotation was approximated by a process called center-of-mass alignment, where the axis of rotation can be set at the center of mass of the object to maintain a self-consistent axis between projections.

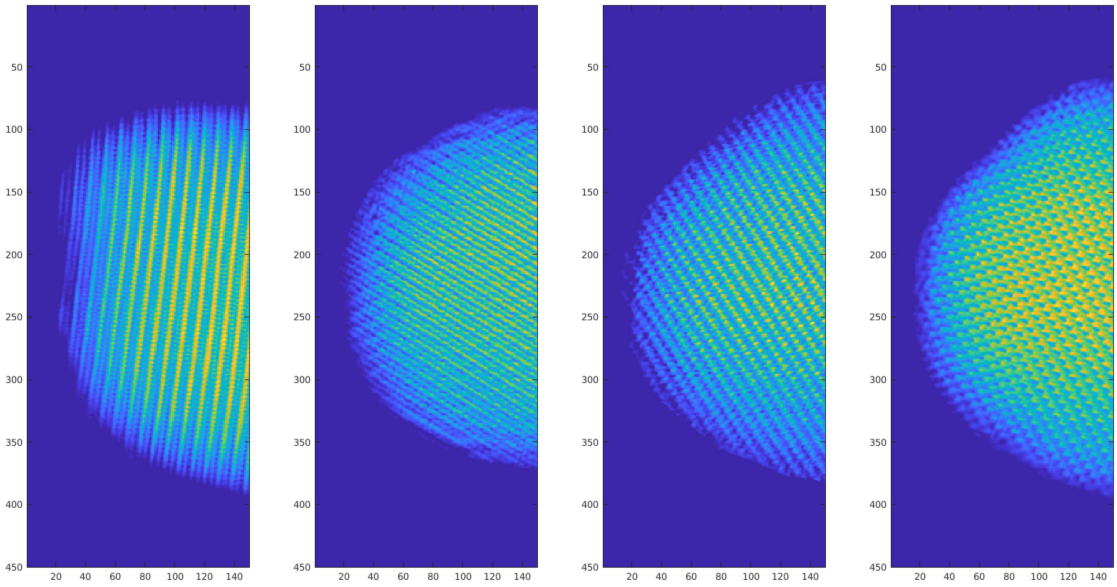

Figure 3.6: Tilt-series ptychographic projections of the tip of the nanowire. Representative projections at 4 different tilt angles are shown.

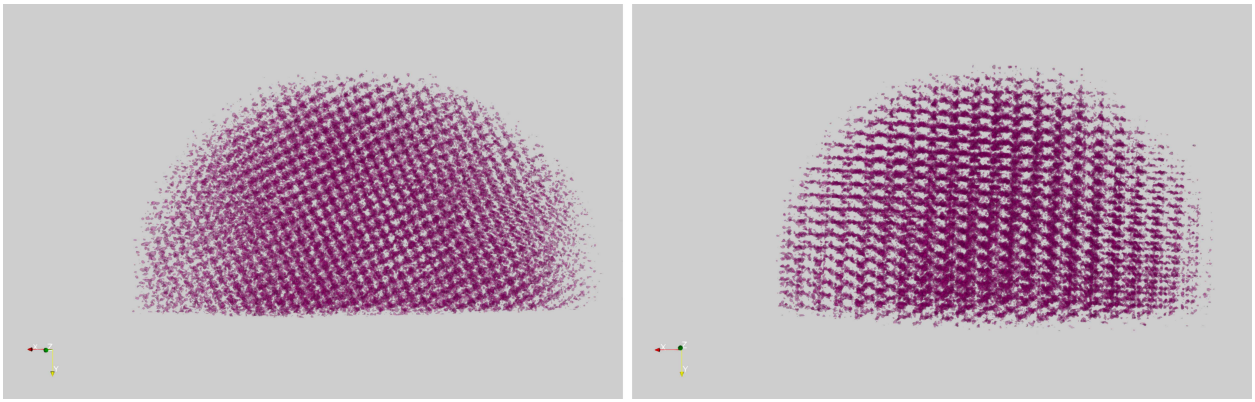

Figure 3.7: Projections of the reconstructed nanowire along crystallographic zone-axes. Atomic columns of manganese are resolved, consistent with a predicted tetragonal crystal structure.

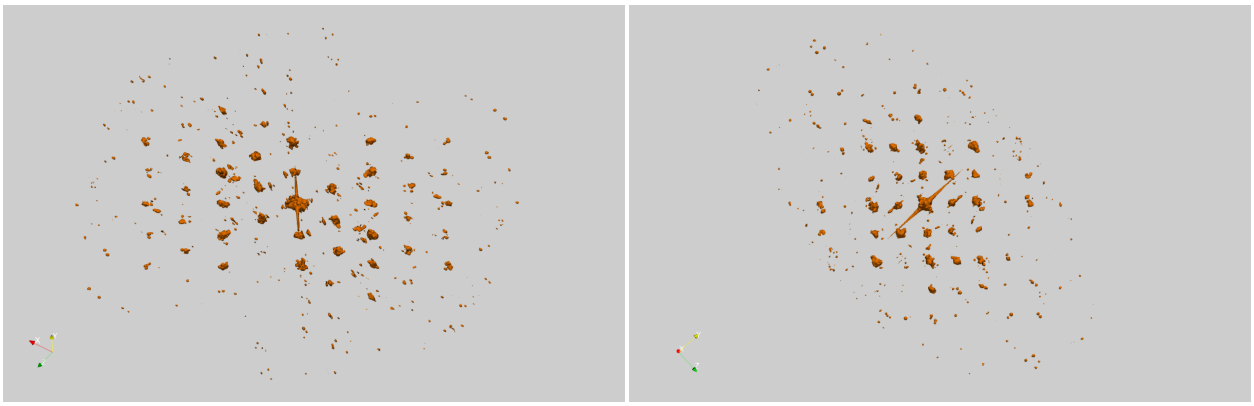

Figure 3.8: Projections of the 3D Fourier transform of the volumetric pAET reconstruction.

Once the projections were fully aligned, GENFIRE was used to reconstruct the 3D volume of the nanowire, as shown in Figure 3.7. The reconstruction resolved individual atomic columns of manganese consistent with the predicted crystal structure for Mn3O4, with clear Bragg peaks in the 3D Fourier transform as shown in Figure 3.8. Even though the signals from the manganese atoms were high, peaks where oxygen atoms should be were not resolved from the reconstruction. A possible reason for this could be the large defocus of the probe incident on the sample. Increasing the variables to solve by necessitating multiple probes and increasing the field-of-view of the phase object per diffraction pattern could result in the algorithm converging at local minima during ptychography.

Chapter 4

Deep Learning CDI Algorithm

4.1 Summary and Motivation for Deep Learning CDI

The remarkable generalizability of CNNs naturally begs the question of their potential in lensless phase imaging. The first foray into deep learning enabled lensless imaging was in 2017 by Sinha et al. that attempted to reconstruct a phase image from measurements made with near-field diffraction [83].

Figure 4.1 shows the results (column V) of a fully trained convolutional neural network that predicts the original phase object image (column II) from diffraction pattern (column III). A spatial light modulator was used to generate 10000 diffraction patterns with which the neural network was trained. The neural network represents an end-to-end solution to inverse problems, even if the forward function is destructive (since only the magnitude of diffraction patterns are measured). Furthermore, the inversion is performed without any iterative methods since the computational load required to solve the inverse is performed before instead of after the diffraction measurement.

However, a high degree of spatial correlation exists between a phase image and its diffraction measurement in the near-field. Yet, most diffractive experiments for X-rays and electrons measure the scattering wave function in the far-field, where the spatial correlation between

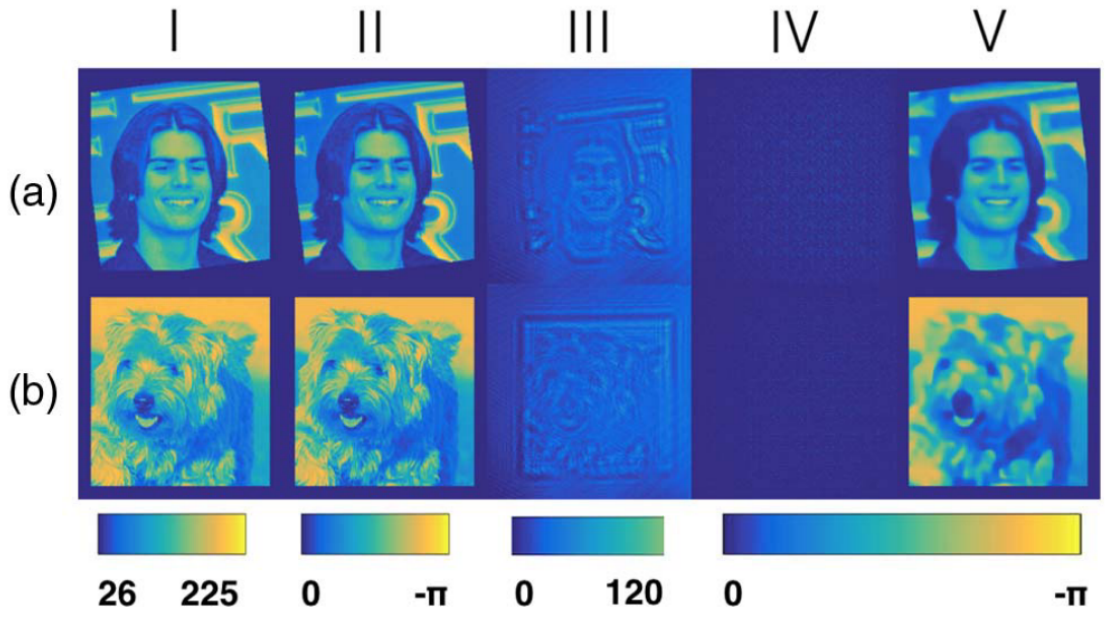

Figure 4.1: Deep neural network trained to reconstruct from near-field diffraction. Rows (a) and (b) show results from networks trained with Faces-LFW and ImageNet datasets, respectively. (i) Randomly selected image. (ii) Image converted to phase object. (iii) Near-field diffraction measured on CMOS. (iv) Neural network output before training. (v) Neural network output after training. Reprinted with permission from [83] ©The Optical Society.

diffraction pattern and object is near non-existant. Is it possible for CNNs to generalize this inverse process?

Cherukara et al. in 2020 demonstrated that an CNN based inversion to far-field diffraction is possible via a X-ray diffraction experiment on etched tungsten test pattern [84], as shown in Figure 4.2. Like the optical near-field experiment, a significant portion of experimental diffraction data as well as the ground truth object image to reconstruct had to be generated *a priori* to training. It seemed that even though the generalizability of CNNs was powerful enough to pattern match diffraction patterns in the far field, any serious application of deep learning based lensless imaging required huge training datasets. A neural network trained to image tungsten by training with tungsten diffraction data could only predict images of tungsten, severely limiting the general applicability of the method.

CDI and ptychography are powerful imaging methods that don't require precise optics and experimental parameters to obtain high resolution reconstructions algorithmically. The paradigm of a coherent illumination source being measured in far-field regime after interacting with a sample is general and thus finds itself in active use and development in various fields. However, the algorithms required to iteratively solve for the object are computationally expensive and depend on skilled microscopists to adjust for parameters and interpret the data. Furthermore, ptychography datasets are measured and stored in 4 dimensions, and the sheer volume of data manipulated before final reconstruction acts as a bottleneck in the image recovery process.

Over the past decades, active development in deep learning methods have demonstrated capability in non-iteratively solving complex and non-linear inverse problems, including lensless imaging. By front-loading the computation required prior to image reconstruction instead of during, CNNs have demonstrated real time reconstruction of complex objects with no further fine tuning of parameters. However, deep learning methods require a large training dataset, which currently can only be obtained via extensive experiments, rendering the

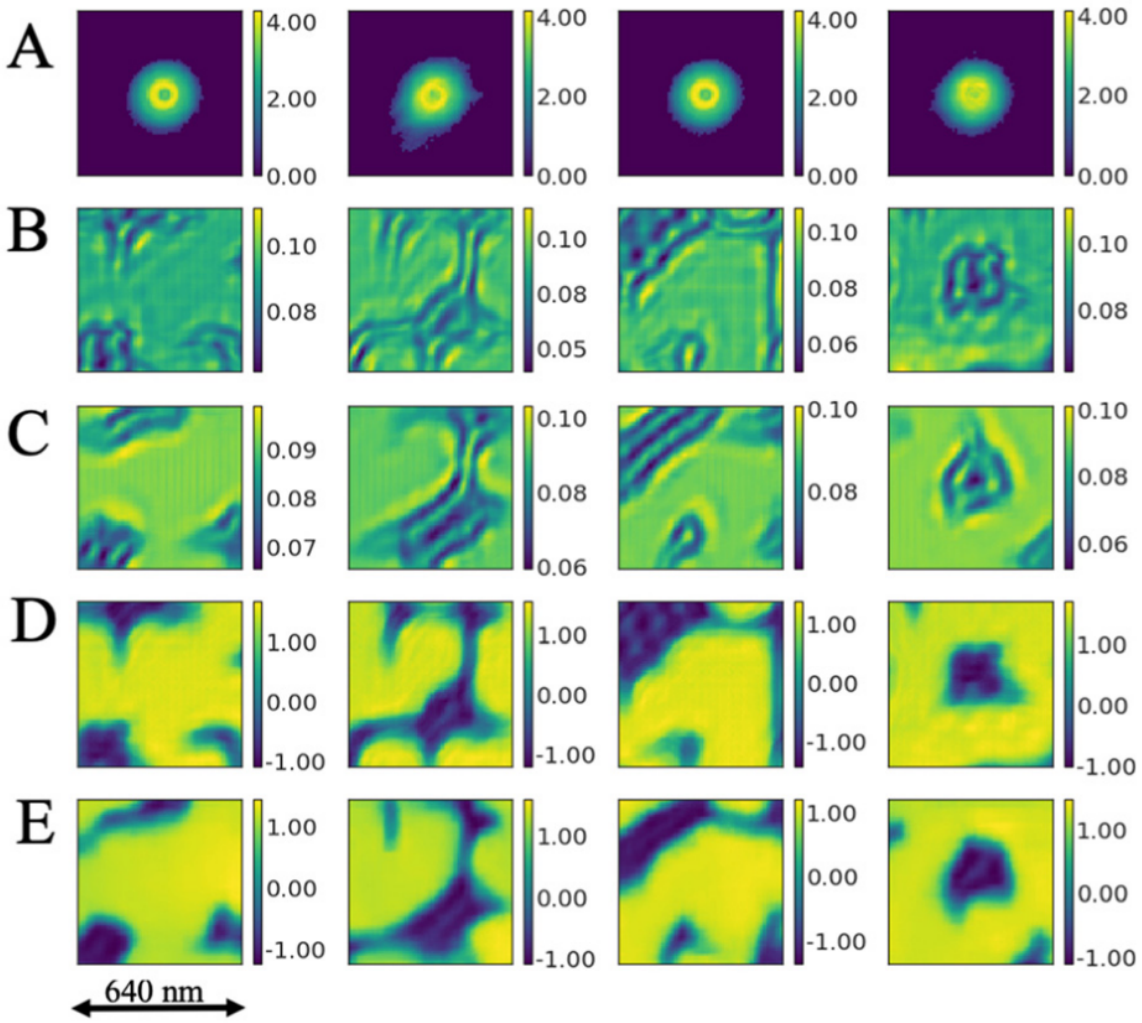

Figure 4.2: Single shot neural network results trained with tungsten ptychography data. (a) Input diffraction patterns for the CNN. (b,c) reconstructed amplitude by ePIE and CNN, respectively. (d,e) reconstructed phase by ePIE and CNN, respectively. Reprinted from [84], with the permission of AIP Publishing.

intent of building a real-time reconstruction system pointless.

In this chapter, we report our process of deep learning CDI whereby neural networks trained with *only* simulated diffraction data are powerful enough to reconstruct complex objects from 4D data directly. Once fully trained, our neural networks leverage their throughput capabilities to enable real-time processing of diffraction data without significant loss in reconstruction quality when compared to conventional ptychographic methods. We believe this advancement in leveraging deep-learning capabilities in lensless imaging embodies a pivotal step in microscopy methodologies in the information age.

4.2 Algorithm Overview

The algorithm is primarily divided into three parts:

1. Generate augmented training dataset via simulating diffraction patterns using random stock images.

2. Train a CNN to perform phase retrieval. The final CNN will predict phase patches from individual diffraction patterns.

3. Stitch retrieved phase patches together to generate a final large field-of-view phase image.

4.3 Augmented Training Data Generation

Deep learning CDI with augmented data begins with the forward propagation of a coherent source. Figure 4.3(a) shows a diagram of a typical electron ptychography setup where a focused coherent source illuminates an object. The resulting wave function is propagated to the far field and only the square of the magnitude of the wave function is measured by

a pixel array detector. Mathematically, this forward process relates the object function and the measurement by

$$M(\boldsymbol{k}) = |\mathcal{F}(P(\boldsymbol{r})O(\boldsymbol{r}))|$$

where $M(\boldsymbol{k})$ is the magnitude of the Fourier transform ($\mathcal{F}$), $P(\boldsymbol{r})$ and $O(\boldsymbol{r})$ are the complex probe and object functions, respectively. Since the phase of the Fourier transform is lost during measurement, the inverse of this forward process is nonlinear. Figure 4.3(b) shows the process of converting a random stock image found on the internet into a pure phase object, which is illuminated by a probe function to produce an exit wave. The magnitude and phase of the probe function are estimated based on the defocus and aberration of the electron optics.

Even though some probe parameters can be measured directly by the detector (such as semi-convergence), most parameters that dictate the illumination function at the sample plane is lost when only measuring the magnitude of the Fourier transform. For example, the defocus of the probe in real space is a quadratic term in the phase of the Fourier transform, which is irrecoverable unless there are redundancies in the data to recover iteratively. While the phase modulation in Fourier space can be any arbitrary function, it's advantageous to approximate this function using multipole expansion in scattering angle. In this thesis, we approximate all probe functions upto second order abberation:

$$\begin{aligned}\chi(\alpha_x, \alpha_y) = \frac{2\pi}{\lambda}[&\frac{1}{2}C_{10}\alpha^2 + \frac{1}{2}C_{12a}(\alpha_x^2 - \alpha_y^2) + C_{12b}\alpha_x\alpha_y + \frac{1}{3}C_{21a}\alpha_x\alpha^2 + \\ &\frac{1}{3}C_{21a}\alpha_y\alpha^2 + \frac{1}{3}C_{23a}\alpha_x(\alpha_x^2 - 3\alpha_y^2) + \frac{1}{3}C_{23b}\alpha_y(\alpha_y^2 - 3\alpha_x^2)]\end{aligned} \tag{4.1}$$

where α_x and α_y are scattering angles in x and y respectively, and λ is the wavelength of the coherent source. The seven constant terms in the equation refer to actual physical variables that can be calibrated directly from the electron microscope – C_{10} is the defocus, C_{12} is

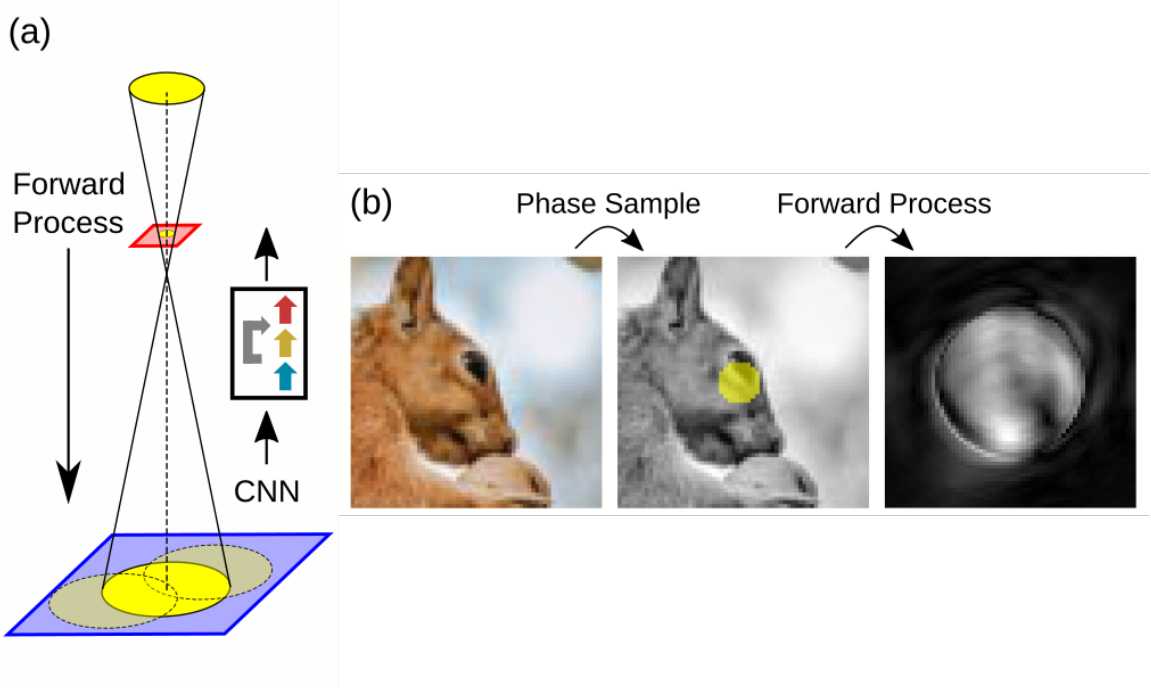

Figure 4.3: Taken from ref [2]. (a) The forward process in electron ptychography that transforms a phase object to overlapped diffraction patterns. A CNN is trained to directly invert this process. (b) Random stock images from the internet are used as phase objects to calculate the diffraction patterns with the forward process illustrated in (a), where the yellow circle indicates the illuminated function of one scanning position.

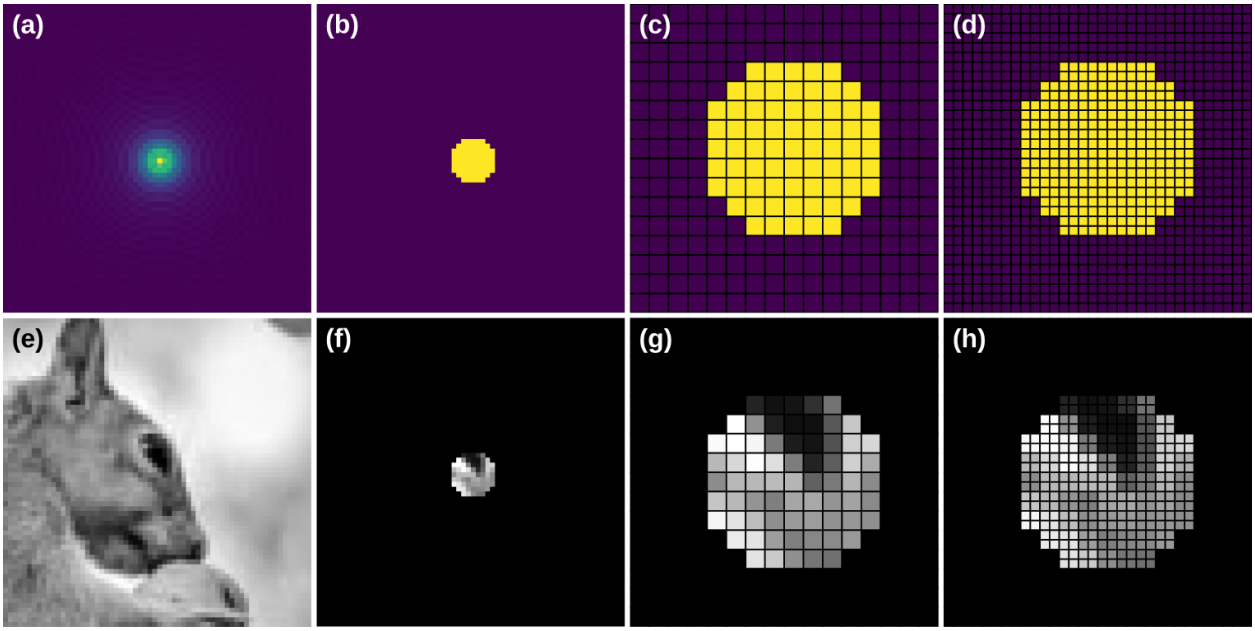

Figure 4.4: Taken from ref [2]. Procedure for generating real-space support and training data. (a) The magnitude of a probe function. (b) Real-space support based on the probe function. The phase object will only be predicted within this support, and any prediction outside of the support will not be used for stitching. (c) Cropped image of (b), removing non-trainable regions. Grids show individual pixels. (d) Oversampling (c) by a factor of 2. Cropping and oversampling in real space are optional steps to match the input/output dimensions of the CNN. (e) An example phase object. (f) A phase patch in the support predicted by the CNN. (g) Result of (f) after cropping non-trainable regions. (h) Result of (g) after oversampling the sample by a factor of 2. Any of (f), (g), or (h) can be used as training datasets for the CNN output given proper post-processing (padding and down-sampling) in the experiment.

twofold astigmatism, C_{21} is coma, and C_{23} is threefold astigmatism [11].

Figure 4.4 shows the procedure for generating phase snapshots to train the CNN along with visual explanations for the relationship between a probe and a real-space support. Cropping non-trainable regions or oversampling the trainable regions in real space allows the phase snapshots to match the input/output array sizes of the CNN, as shown in Fig. 4.4(f-h).

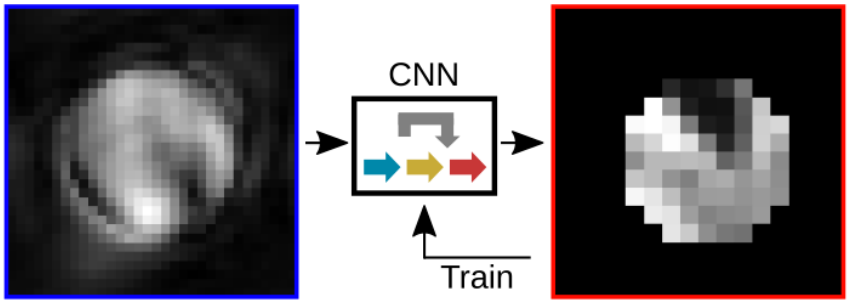

Figure 4.5: Taken from ref [2]. The CNN is trained to directly map from the square root of the diffraction patterns to the phase patches within the illuminated areas.

4.4 Training CNN for Phase Retrieval

The square root of the noisy diffraction intensity is used to train CNNs with an L1-norm loss function to recover the phases in the illuminated area (named a phase patch) [Fig. 4.5]. As the diffraction intensity is corrupted with Poisson noise, a low learning rate of 1.010^{-4} and a high dropout rate of 0.2 are applied at every layer to prevent overfitting of noise while training with stochastic gradient descent.

In our experience, using randomly generated stock photos from the internet provides a rich source of entropy within the images to sufficiently train the CNNs without imposing any regularizations [85]. Detailed examples of phase image generation, forward process and CNN performance as validation data can be found in Figure 4.6.

4.4.1 CNN Architecture

The architecture of the CNN is an encoder-decoder architecture Figure 4.7, more commonly known as U-net [51], with skip connections between corresponding tensor sizes to prevent vanishing gradient issues. A schematic of the residual layers is shown in Figure 4.8, with skip connections acting as concatenations that provide direct throughput within the architecture.

Regarding space complexity, one unique aspect of CNNs is their constant memory use

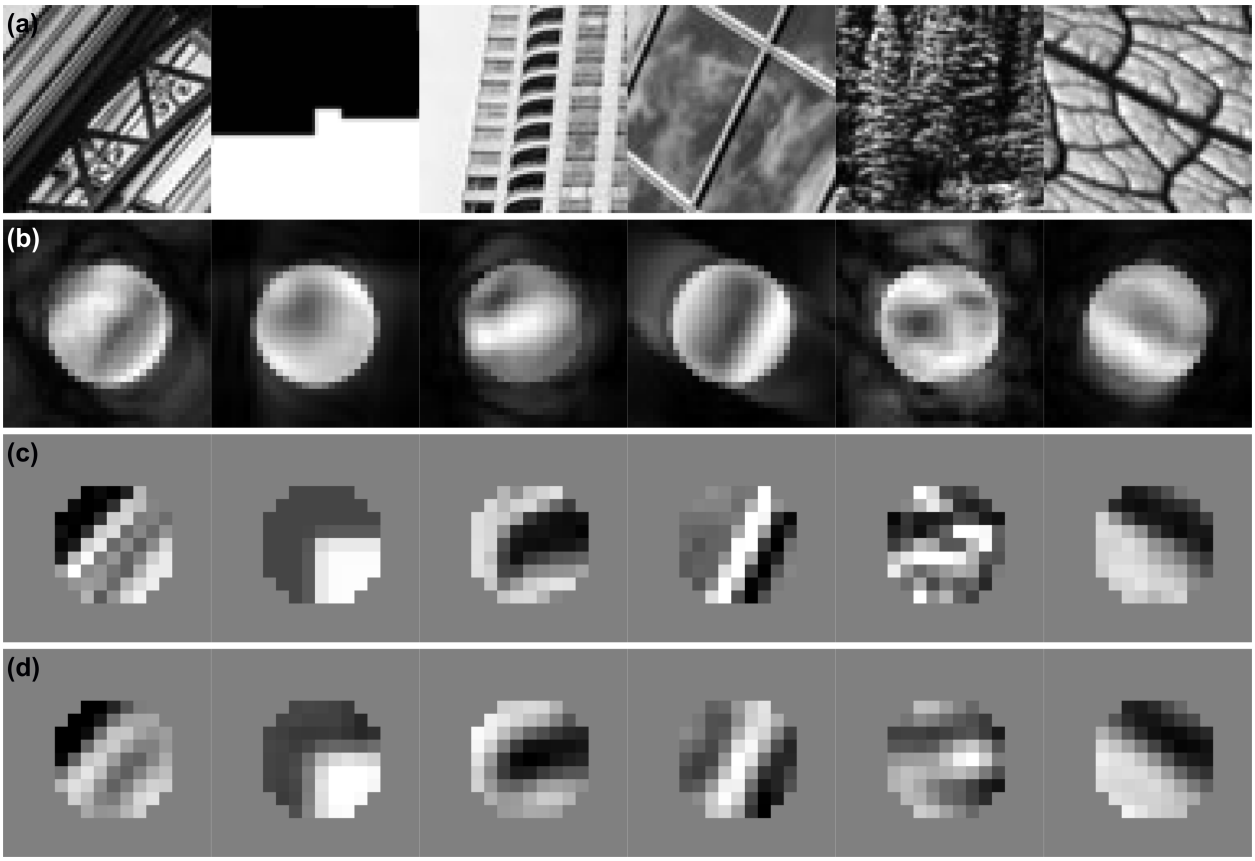

Figure 4.6: Taken from ref [2]. Examples of augmented data generation and CNN performance. (a) Random images from the internet used as pure phase objects to train a CNN. (b) Diffraction patterns generated from the phase objects in (a). (c) Perfect phase patches within the illuminated areas. (d) Corresponding phase patches independently retrieved from the square root of the diffraction patterns by a trained CNN without any iteration.

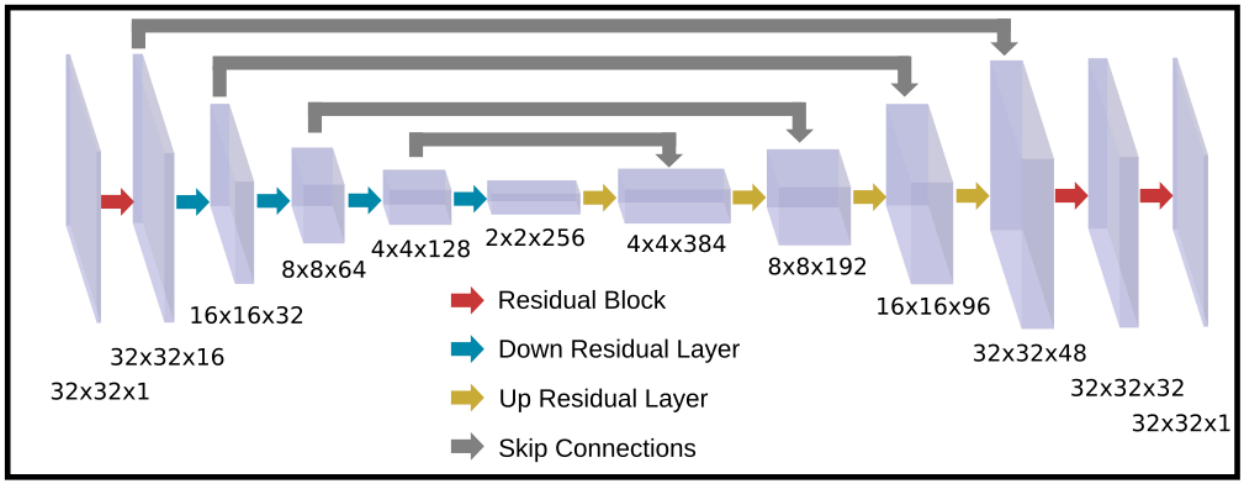

Figure 4.7: Taken from ref [2]. Detailed schematic of the CNN architecture.

as a function of input sizes. When training the CNNs for larger detector arrays, the total number of trainable parameters remain constant in the GPU RAM, demonstrating efficient scaling in memory. As a result, since the training data as well as diffraction patterns scale linearly with array sizes, runtime and space complexities scale linearly for both training and predictions.

4.4.2 Impact on Training Data Size

We performed a quantitative study of the effect of the training data size on CNN phase retrieval, shown in Figure 4.9. We observed that the final L1 norm values as well as the phase reconstructions didn't significantly improve after around 250,000 training sample sizes. While the required training data size depends on the complexity of the probe function (such as defocus and aberration), we found that 250,000 was sufficient for accurate phase reconstructions for all three electron ptychographic experiments.

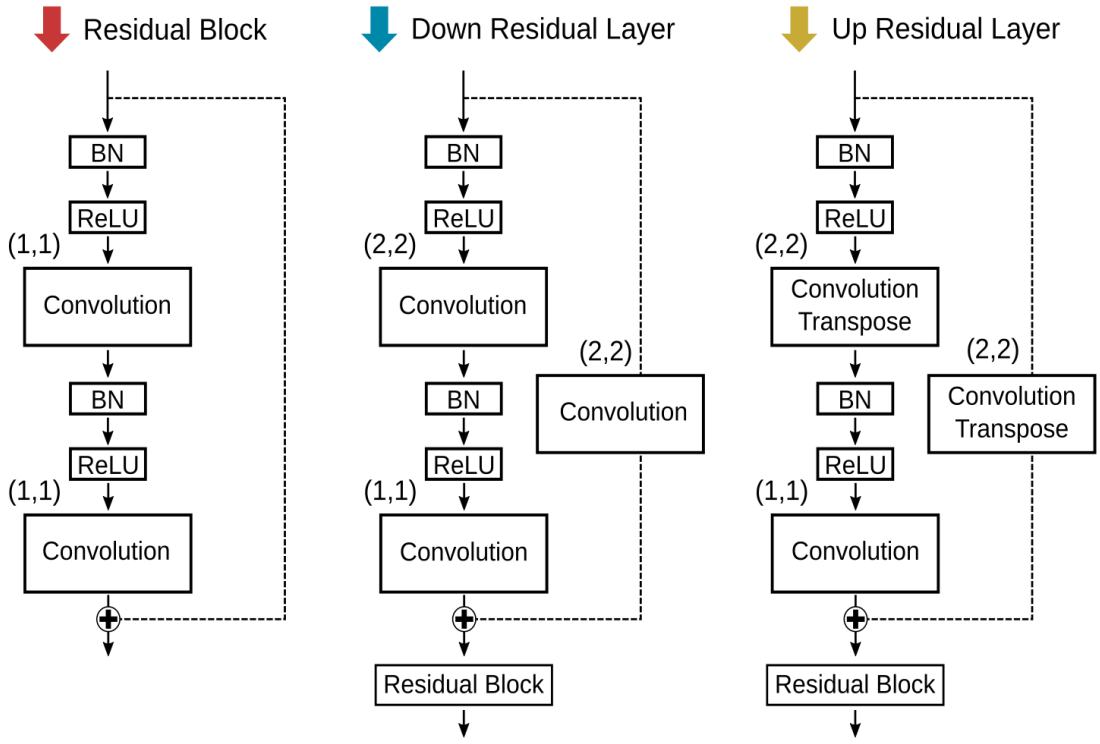

Figure 4.8: Taken from ref [2]. Schematic of residual layers used in the CNN architecture. Different strides used in the convolution filters are shown in parentheses. All filters are size 3×3.

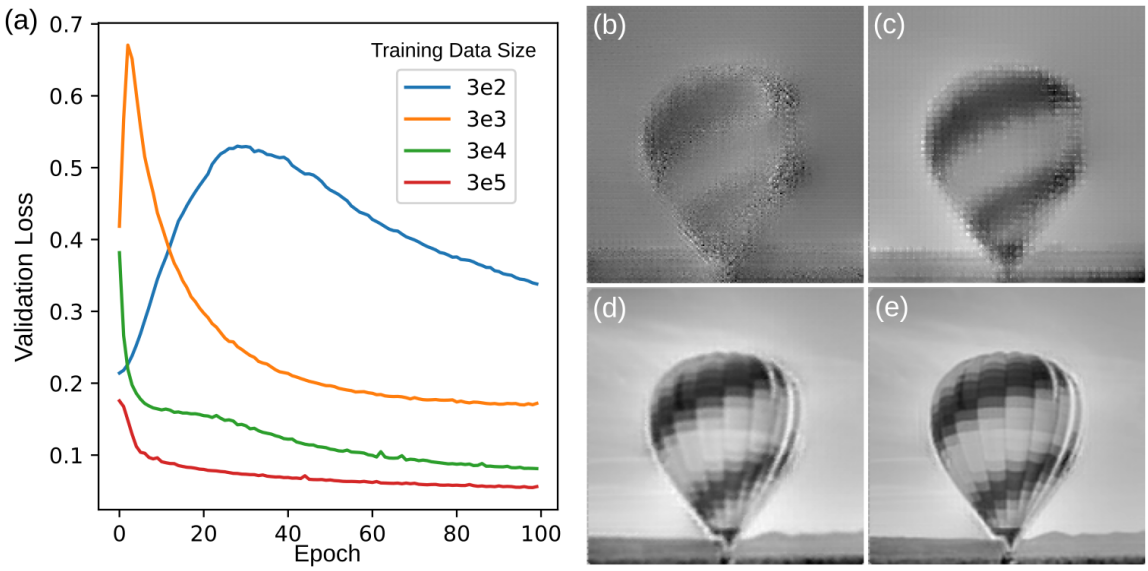

Figure 4.9: Taken from ref [2]. Effect of training data size on CNN phase retrieval. (a) L1 validation loss when trained with various training data sizes. (b-e) Stitched phase reconstructions after 100 epochs of training with 3e2, 3e3, 3e4, and 3e5 training data sizes, respectively.

4.4.3 Bypassing the Phase Problem

Note that only from the diffraction intensity, one cannot distinguish a function $f(\boldsymbol{r})$ from its twin, $f^*(-\boldsymbol{r})$, which is known as the twin image problem [25]. CNNs solve this problem by using the constraint of the probe function. When training CNNs, the function taken to the far-field is modeled as $f(\boldsymbol{r}) = P(\boldsymbol{r})O(\boldsymbol{r})$, whereas its twin function is represented by

$$f^*(-\boldsymbol{r}) = P(\boldsymbol{r})\frac{P^*(-\boldsymbol{r})}{P(\boldsymbol{r})}O^*(-\boldsymbol{r})$$

where $P(\boldsymbol{r})$ is pre-determined in the current work. Since the inequality

$$P^*(-\boldsymbol{r}) \neq P(\boldsymbol{r})$$

usually holds due to the aberration and/or defocus of the probe, CNNs are trained to recover $O(\boldsymbol{r})$, but eliminate the non-physical solution of $\frac{P^*(-\boldsymbol{r})}{P(\boldsymbol{r})}O^*(-\boldsymbol{r})$ as demonstrated in our numerical simulations and experimental results.

4.5 Stitching Algorithm

The fully trained CNN is then used to directly map from the magnitudes of the Fourier transform to phase patches without any iteration, which are combined to form a phase image by a stitching algorithm [Fig. 4.10]. As the zero frequency of the phase is irrecoverable from CNNs, two adjacent phase patches differ in the overlapping region by an overall phase shift. The stitching algorithm is therefore executed by minimizing an error function E,

$$E = \sum_{i,j \neq i} \left[\mu(\phi_i S_{i,j}) - \mu(\phi_j S_{j,i})\right]^2$$

where ϕ_i and ϕ_j are the i'th and j'th phase patch, respectively, and $S_{i,j}$ is the overlapping region between the two phase patches, and μ returns the mean value. The error function is minimized by gradient descent and a well-stitched phase image usually requires only a few iterations.

This chapter shows experimenal DL-CDI results performed using scanning transmission electron microscopes. Diffraction patterns of three vastly different samples collected on different detectors and electron microscopes were successfully processed into phase image reconstructions.

4.6 Twisted Hexagonal Boron Nitride Experiment

The first sample consists of two 5-nm-thick hexagonal boron nitride (hBN) flakes with a twisted interface. The experiment was conducted on the TEAM I double-corrected S/TEM

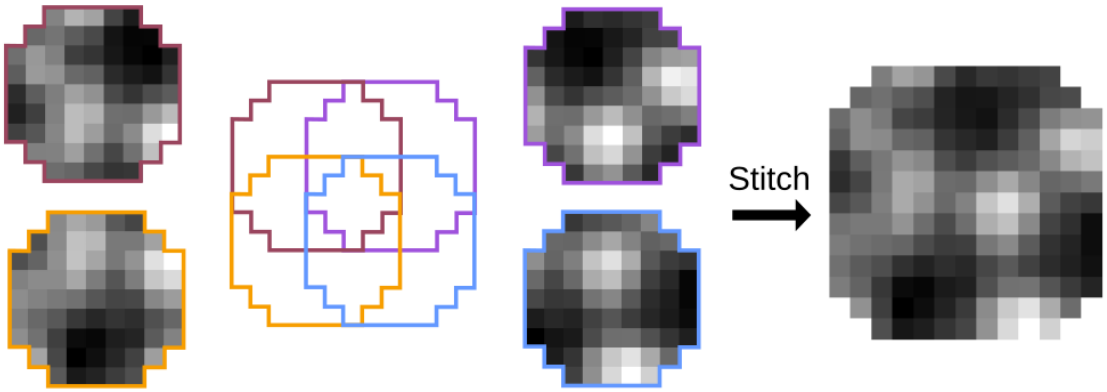

Figure 4.10: Taken from ref [2]. Four representative phase patches directly retrieved by the CNN. Based on the overlapping regions, all the phase patches are stitched together to form a phase image.

Parameter	Value
Microscope	TEAM I
Voltage (kV)	300
Convergence semi-angle (mrad)	17.1
Scanning step size (Å)	0.25
Number of probe positions	512×512
Dwell time per probe position (ms)	0.870
Pixelated detector	Gatan K3
Number of detector pixels (before binning)	512×512
Dose ($e^{-}\mathring{A}^{-2}$)	4.5×10^{6}

Table 4.1: Experimental parameters of CDI datasets for hBN experiment.

instrument at the National Center for Electron Microscopy (NCEM). Additional experimental parameters can be found in Table 4.1.

The microscope was equipped with a Gatan K3 pixelated detector, which operated in electron-counting mode and was binned (×2) and windowed (×2) to 512×512 pixels. The diffraction patterns were further binned to 32×32 pixels in post-processing. Figure 4.11(a) shows a representative diffraction pattern with the presence of heavy noise. Conventional STEM imaging modes such as annular dark-field (ADF) imaging produce images with a poor signal-to-noise ratio [Fig. 4.11(b)].

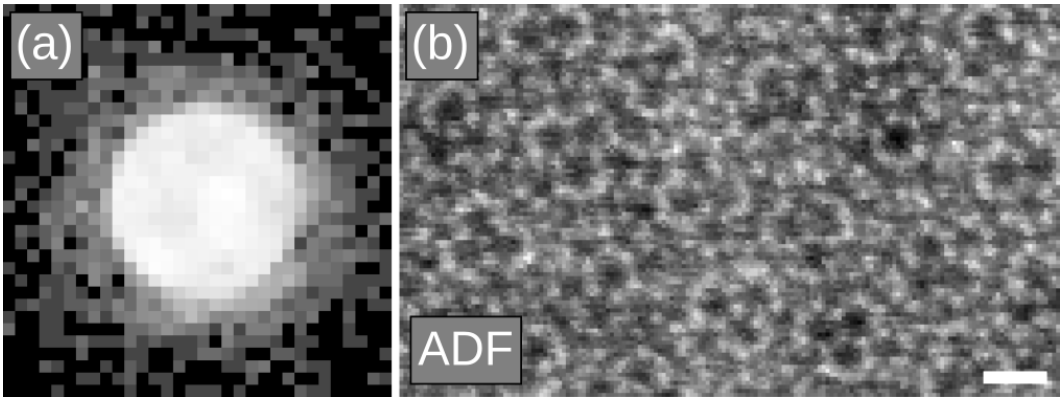

Figure 4.11: Taken from ref [2]. (a) A representative diffraction pattern measured from two 5-nm-thick hBN flakes with a twisted interface (displayed in a logarithmic scale). (b) An ADF-STEM image of the hBN sample generated by integrating the diffraction intensity outside the bright-field disk at each scanning position.

Parameter	Value
# detector pixels (after cropping + binning)	32
Maximum collection angle after cropping (mrad)	34.2
Angular pixel size (mrad/pixel)	1.07
Training sample size	250,000
Learning rate	1.e-4
# of trainable parameters	9,445,408
Epochs trained	100
Prediction time (s per 1000 phase patches)	0.519

Table 4.2: Computational parameters used in deep learning CDI for the hBN experiment.

To generate the training dataset, a probe function was analytically generated by parameterizing the aberration function using the aforementioned seven parameters. After training with 250,000 simulated diffraction patterns, the CNN was used to retrieve the individual phase patches from the experimental data in real time on standard commercial GPUs (0.5 s per 1000 phase patches on an NVIDIA GTX 1060 6gb). Computational parameters used for deep learning CDI can be found in 4.2.The phase patches of the twisted hBN sample were then stitched together to form a phase image [Fig. 2(c) and (d)]. As a comparison, the phase reconstruction of the same experimental data by ePIE is shown in Fig. 2(e) and (f).

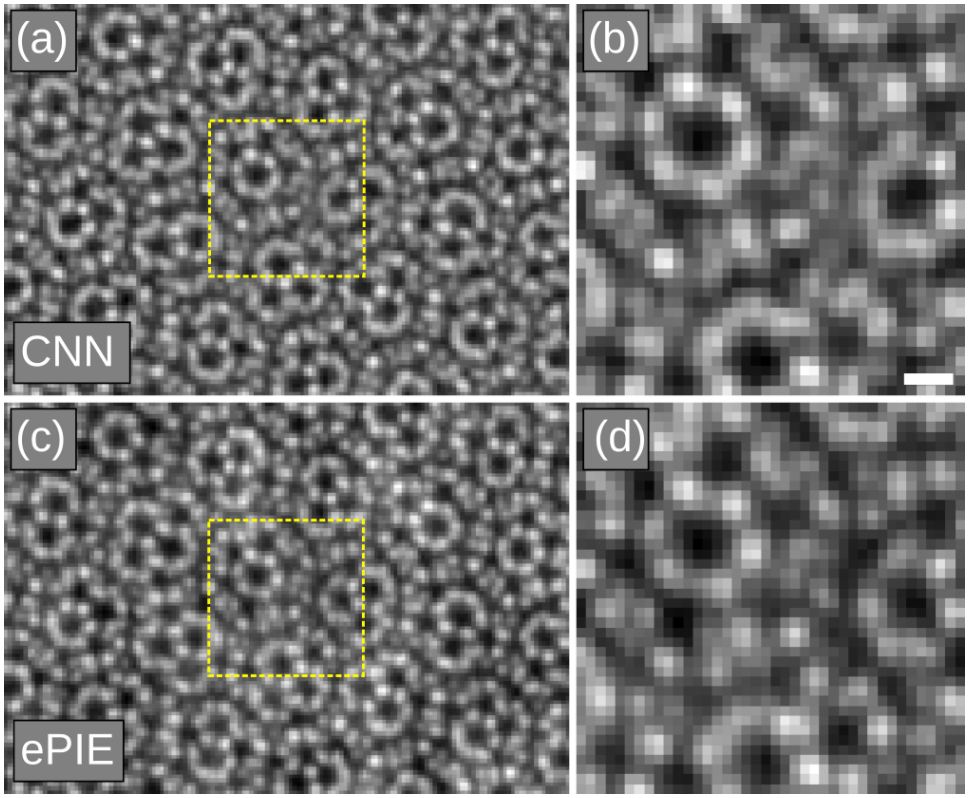

Figure 4.12: Taken from ref [2]. (a) Phase image of the twisted hBN interface retrieved by a trained CNN. (b) Magnified view of the dotted square in (a). (c) Phase reconstruction of the twisted hBN sample reconstructed by ePIE. (d) Magnified view of the dotted square in (c).

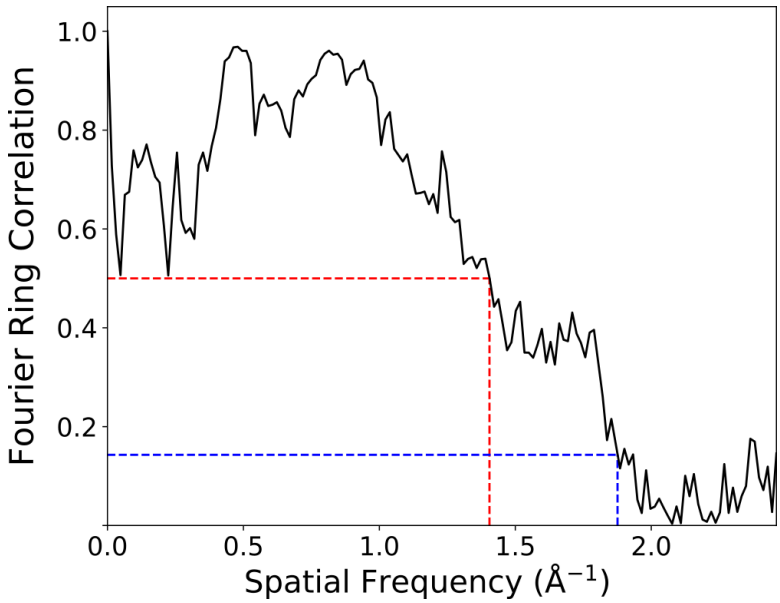

Figure 4.13: Taken from ref [2]. Fourier ring correlation between Figure 4.12(a) and 4.12(c), where the red and blue dashed lines indicate a resolution of 0.71 and 0.53 Å, based on the cutoff criteria of FRC = 0.5 and 0.143, respectively.

Fourier ring correlation (FRC) between the CNN and ePIE reconstructions shows a good agreement between the two methods [Fig. 4.13]. The slight difference at low spatial frequencies of the FRC curve is likely because the stitching algorithm can combine the phase patches together more uniformly than ePIE. Based on the cutoff criteria of FRC = 0.5 and 0.143 [86], the resolution of the CNN phase image was quantified to be 0.71 and 0.53 Å, respectively [red and blue dashed lines in Fig. 4.13], demonstrating that both methods consistently reconstructed the diffraction signal beyond the bright-field disk.

4.7 Monolayer Graphene Experiment and Data Sparsity in Scanning Positions

Next, we used an experimental data set of monolayer graphene to investigate the tolerance of the stitching algorithm to varying overlap of the diffraction patterns. The ptychographic data

Parameter	Value
Microscope	JEOL ARM200F
Voltage (kV)	80
Convergence semi-angle (mrad)	31.5
Scanning step size (Å)	0.13
Number of probe positions	512×512
Dwell time per probe position (ms)	0.250
Pixelated detector	JEOL 4DCanvas
Number of detector pixels (before binning)	66 × 66
Dose ($e^{-}\mathring{A}^{-2}$)	1.4×10^6

Table 4.3: Experimental parameters of CDI datasets for the monolayer graphene experiment.

Parameter	Value
# detector pixels (after cropping + binning)	32
Maximum collection angle after cropping (mrad)	38.9
Angular pixel size (mrad/pixel)	1.21
Training sample size	250,000
Learning rate	1.e-4
# of trainable parameters	9,445,408
Epochs trained	100
Prediction time (s per 1000 phase patches)	0.523

Table 4.4: Computational parameters used in deep learning CDI for the monolayer graphene experiment.

set was acquired using a JEOL 4DCanvas pixelated detector installed on a JEM-ARM200F probe-corrected microscope, as reported elsewhere [87]. The probe function was analytically generated with a second order aberration function as implemented in the hBN experiment, and a CNN was trained by 250,000 simulated diffraction patterns. Additional experimental and computational parameters can be found in Tables 4.3 and 4.4.

Phase patches independently retrieved by the CNN from the experimental diffraction patterns were stitched together to form a phase image [Fig. 4.14(a)], which is in a good agreement with the ePIE reconstruction of the same data set [Fig. 4.14(e)]. To study the performance of the stitching algorithm with varying overlap of the diffraction patterns, every

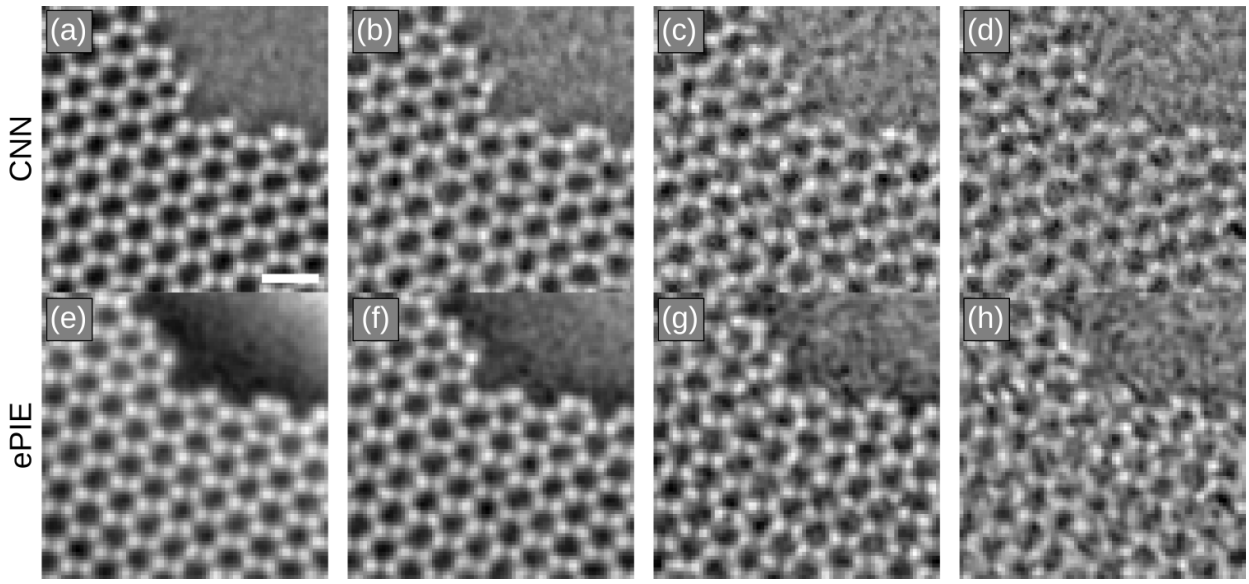

Figure 4.14: Taken from ref [2]. Deep learning phase retrieval of the experimental monolayer graphene data with varying overlap of the diffraction patterns. (a-d) Stitched phase images of the monolayer graphene retrieved by a trained CNN with a scanning step size of 0.132 Å, 0.264 Å, 0.396 Å, and 0.528 Å, respectively. (e-h) Phase images of the monolayer graphene reconstructed by ePIE with the same scanning step sizes in (a-d). Scale bar, 4 Å.

other diffraction pattern in both x and y scanning directions was used to reconstruct the phase image in Fig. 4.14(b), doubling the scanning step size from 0.132 Å to 0.264 Å and quartering the effective electron dose. Additionally, the stitched phase images of the CNN by taking every three (0.396 Å step size) and four (0.528 Å step size) diffraction patterns are shown in Fig. 4.14(c) and (d), respectively. Note that the phase images in Fig. 4.14(a-d) were recovered using the same trained CNN, as the forward process did not change when taking data with varying scanning step sizes. As a comparison, the corresponding phase images reconstructed by ePIE are shown in Fig. 4.14(e-h), demonstrating that deep learning and ePIE provide similar reconstruction qualities with varying overlap of the diffraction patterns.

Parameter	Value
Microscope	TEAM 0.5
Voltage (kV)	300
Convergence semi-angle (mrad)	17.1
Scanning step size (Å)	0.26
Number of probe positions	512×512
Dwell time per probe position (ms)	0.044
Pixelated detector	4D Camera
Number of detector pixels (before binning)	576 × 576
Dose ($e^{-}\mathring{A}^{-2}$)	4.6 × 104

Table 4.5: Experimental parameters of CDI datasets for the 5nm Au nanoparticle experiment.

Parameter	Value
# detector pixels (after cropping + binning)	32
Maximum collection angle after cropping (mrad)	53.4
Angular pixel size (mrad/pixel)	1.67
Training sample size	250,000
Learning rate	1.e-4
# of trainable parameters	9,445,408
Epochs trained	100
Prediction time (s per 1000 phase patches)	0.458

Table 4.6: Computational parameters used in deep learning CDI for the Au nanoparticle experiment.

4.8 5nm Au Nanoparticle Experiment

Finally, a ptychographic experiment was conducted on a 5 nm gold nanoparticle to examine the performance of CNNs with strongly scattering atoms. The data set was acquired using the 4D Camera (576×576 pixels) installed on the TEAM 0.5 double-corrected S/TEM instrument at NCEM. Additional experimental and computational parameters can be found in Tables 4.5 and 4.6.

Figure 4.15(a) shows a representative diffraction pattern after binning, where heavy noise is evident. An ADF-STEM image was generated by integrating the intensity outside the

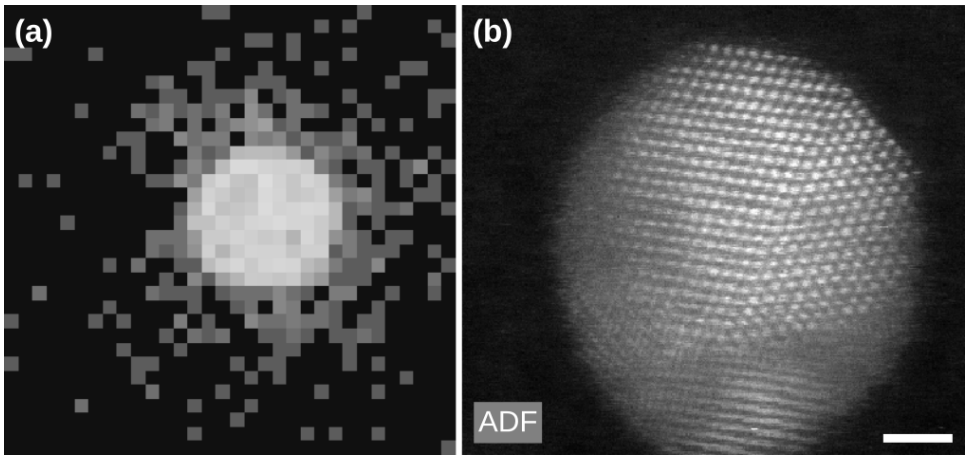

Figure 4.15: Taken from ref [2]. (a) Representative diffraction pattern acquired from a 5-nm gold nanoparticle (displayed in a logarithmic scale). (b) ADF-STEM image generated by integrating the diffraction intensity outside the bright-field disk at each scanning position. Scale bar, 1nm.

bright-field disk [Fig. 4.15(b)], corresponding to an incoherent image of the sample.

The probe function was analytically generated with a second order aberration function as for the hBN and graphene cases, and a CNN was trained by 250,000 simulated diffraction patterns. Figure 4.16(c) and (d) show the phase images retrieved by the CNN and ePIE, respectively, exhibiting consistent atomic features along the zone axis. The FRC curve of the two images indicates a resolution of 0.70 and 0.55 Å with the cutoff criteria of FRC = 0.5 and 0.143, respectively [Fig. 4.17]. The dip in the low spatial frequencies in the FRC curve is consistent with the observation in the hBN result [Fig. 4.13].

4.9 Comparison vs Other Methods

4.9.1 Tolerance of Shot Noise vs ePIE

To investigate CNNs and the stitching algorithm under various shot noise conditions, we simulated diffraction patterns with three different orders of the noise and observed that the final stitched phase reconstructions using the CNNs are consistent with those obtained by

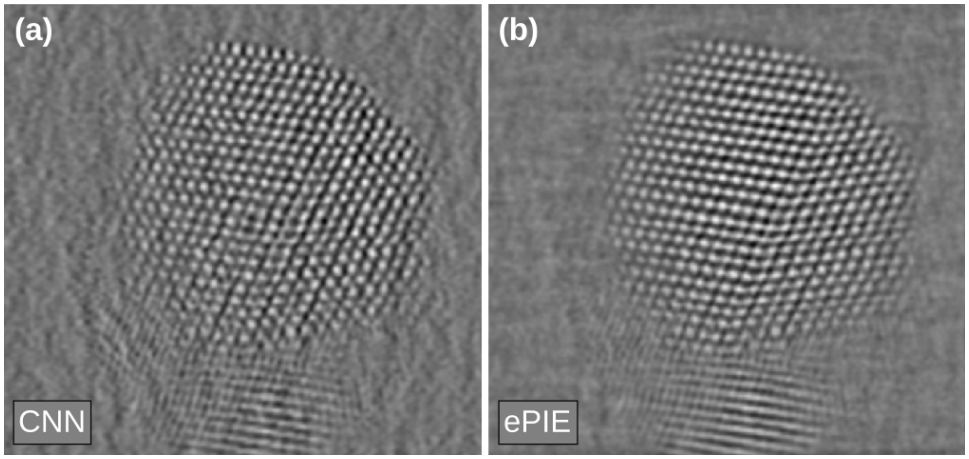

Figure 4.16: Taken from ref [2]. (a) Phase image of the gold nanoparticle retrieved by a trained CNN. (b) Phase image of the nanoparticle reconstructed by ePIE.

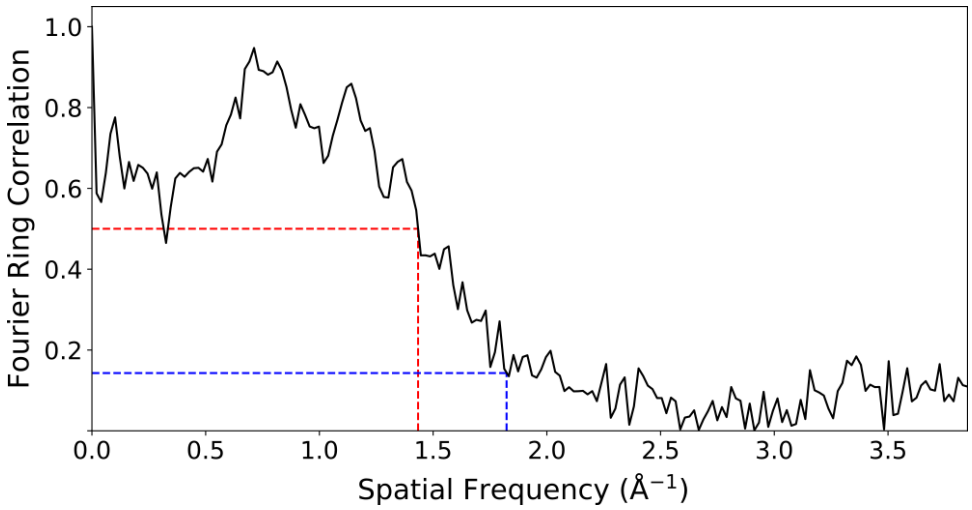

Figure 4.17: Taken from ref [2]. FRC curve between Figure 4.16(a) and (b), where the red and blue dashed lines indicate a resolution of 0.70 and 0.55 Å, based on the cutoff criteria of FRC = 0.5 and 0.143, respectively.

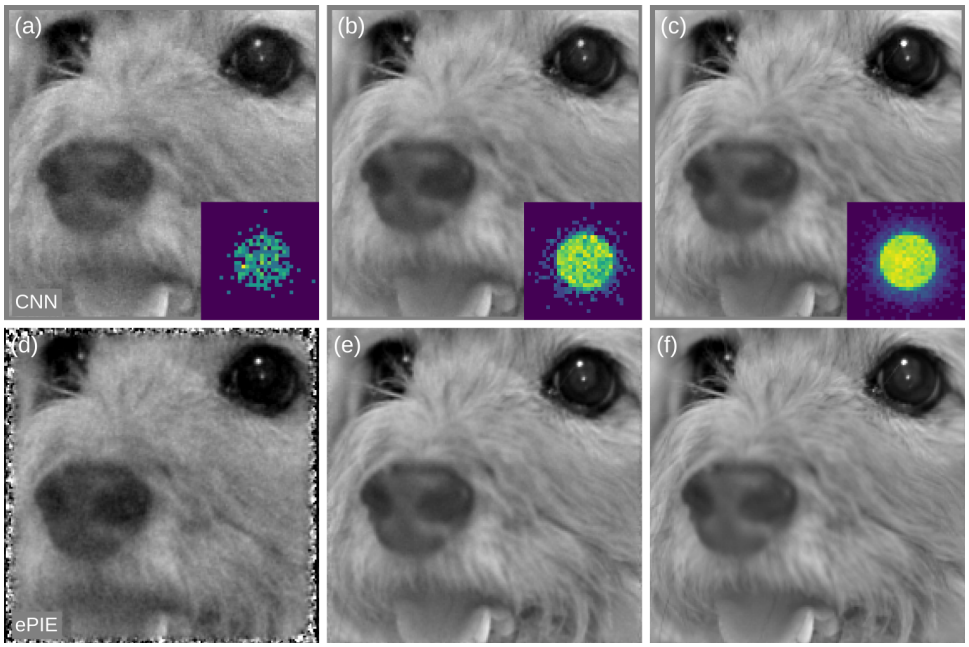

Figure 4.18: Taken from ref [2]. Comparison between CNNs and ePIE under various shot noise conditions. (a-c) Stitched CNN phase reconstructions from diffraction patterns with 3e2, 3e3, and 3e4 events per pattern, respectively. Insets show representative diffraction patterns. (d-f) Phase reconstructions using ePIE from diffraction patterns with 3e2, 3e3, and 3e4 events per pattern, respectively.

the extended ptychographic iterative engine (ePIE) [Fig. 4.18].

4.9.2 Comparison vs Non-Iterative Methods

CNNs and iterative ptychographic methods utilize the sampling in the detector plane to reconstruct the phase, whereas SSB and WDD methods perform an integration in the detector plane and utilize the sampling from the probe scanning plane. As a result, small scan step sizes (typically < 0.5 Åin STEM) are required to sample at the Nyquist limit for SSB and WDD methods. In contrast, CNNs and ePIE methods can be implemented with larger scan step sizes to achieve larger fields of view, as long as there is sufficient overlap between adjacent probe positions, which can be tuned by defocusing the probe.

To illustrate this point, we performed numerical simulations to compare CNNs with non-

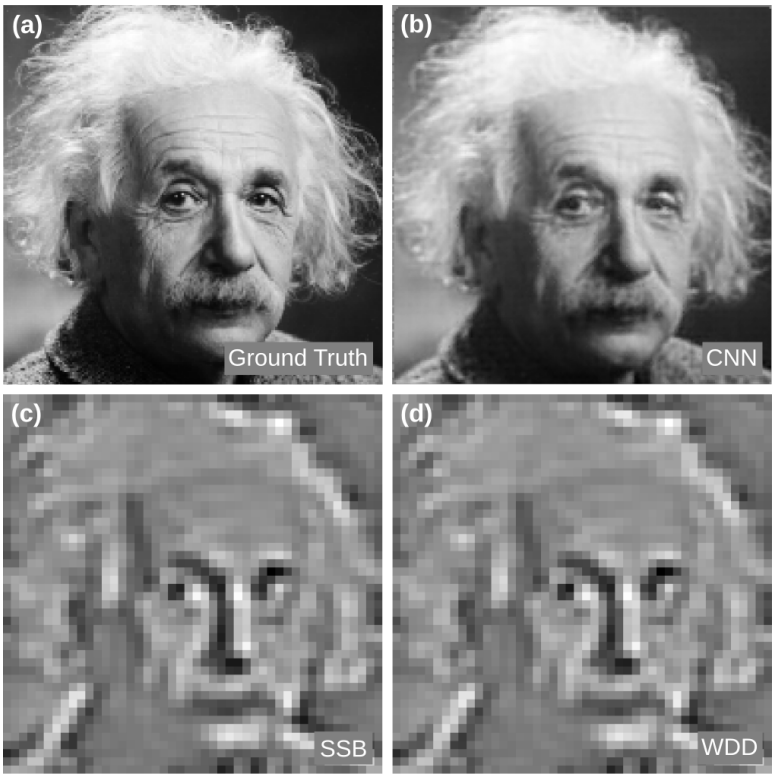

Figure 4.19: Taken from ref [2]. Comparison of phase reconstructions using CNNs and non-iterative methods. Diffraction patterns were generated by a defocused probe of 6 pixels in diameter and a scanning step size of 4 pixels to show the difference in reconstruction quality. (a) Ground truth image simulated as a phase object. (b-d) Phase image reconstructions by CNNs, Single Side Band ptychography, and Wigner-Distribution Deconvolution, respectively.

iterative methods. We generated a set of diffraction patterns using a defocused probe of 6 pixels in diameter and a scanning step size of 4 pixels. Figure 4.19 shows the phase reconstructions by CNNs, single side band ptychography, and Wigner-distribution deconvolution. While CNNs faithfully reconstruct the phase image, the phase images by the non-iterative methods are degraded due to the use of a defocused probe. Furthermore, CNNs and ePIE can achieve higher spatial resolutions than the non-iterative methods as the resolution of the former is only limited by the spatial frequency of the diffraction signal.

4.10 Limitations of DL-CDI

While numerous experiments above demonstrate similar reconstruction qualities between DL-CDI and ePIE, some characteristics about convolutional neural networks yield sub-par reconstruction qualities for certain types of diffractive experiments. An example of one is the electron ptychography experiment performed on MoS2 and collected on the electron microscope pixel-array detector (EMPAD), as shown in Figure 4.20.

The EMPAD is capable of collecting electron signals with a dynamic range of 1000000:1, and thus are able to collect upto 5 times the semi-convergance angle. The sample diffraction patterns shown in Figures 4.20(b) and (c) are in logarithmic scale, and the final ptychography resolution is ultimately dose-dependent as there aren't enough electron events to overcome significant Poisson noise in high angles. Fourier transform of signals tend to decay exponentially as a function of frequency, explaining this phenomenon.

A convolutional neural network based algorithm employs a series of linear convolutions and deconvolutions to encode and decode in the latent space, but how do those capabilities scale when the input data varies exponentially? Figure 4.21(a) show the DL-CDI phase reconstruction of MoS2 from the ptychographic experiment shown in Figure 4.20. Compared to the ePIE reconstruction from the same data shown in Figure 4.21(b), DL-CDI fails to interpret the high resolution information measured at high scattering angles. The resolution that DL-CDI is able to resolve at involves interpreting measurements around the bright field disk and proximal dark field scattering and the relationship between. Thus, for dose-limited and high scattering angle CDI experiments, traditional iterative methods should be employed for high resolution reconstructions.

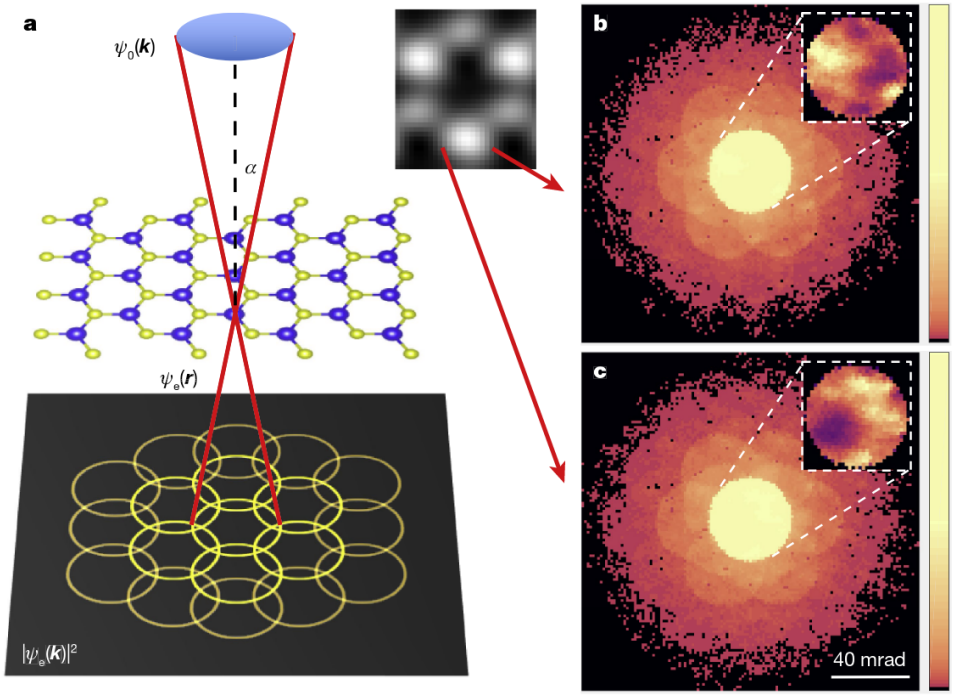

Figure 4.20: Taken from ref [47]. (a) At each scan position, the incident probe ($\psi_0(k)$) is focused on the sample and the entire diffraction pattern of the exit wave ($|\psi_e(k)|^2$) is recorded by the EMPAD. The blue and yellow atoms represent molybdenum and sulfur atoms in the object plane. ψ_0 and ψ_e refer to the incident and exit wavefunctions respectively; r is the $(x,\ y)$ positional coordinate in the real-space plane; and k is the $(k_x,\ k_y)$ wavenumber coordinate in the conjugate momentum-space plane. (b, c) Averaged diffraction pattern intensity (on a logarithmic scale) from the electron beam at the marked scan positions near a molybdenum column. Insets show the intensity (on a linear scale) of the bright-field disks. The substantial intensity differences at large scattering angles provide contrast information for ADF imaging and are essential for resolution enhancement in ptychography. Reproduced with permission from Springer Nature.

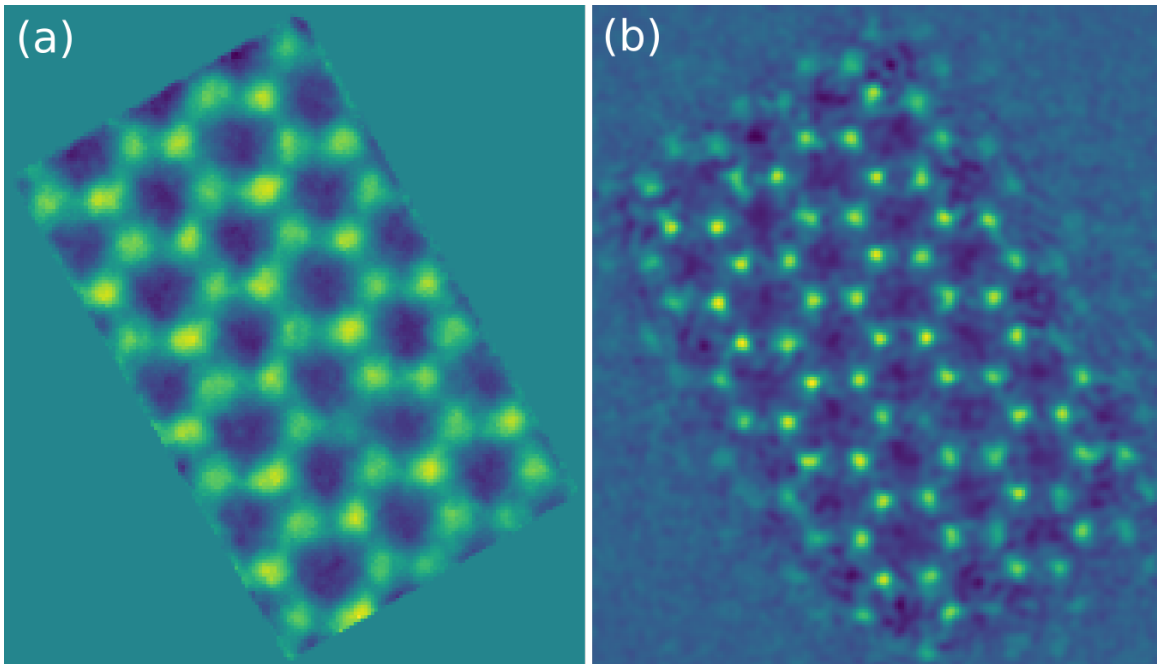

Figure 4.21: Ptychographic imaging of MoS2. (a) DL-CDI phase reconstruction of MoS2. (b) ePIE phase reconstruction from the same data.

4.11 Implementation of DL-CDI on Tensorflow

This section will briefly explain the programming methods used in implementing DL-CDI using Tensorflow in Python. The complete source code along with a sample dataset can be found in `https://github.com/dillanchang/deep-phase-imaging`.

The main method for the code used in training the neural network begins as such:

```
1  import numpy as np
2  import tensorflow as tf
3  from tensorflow.keras import layers, models, callbacks,
4    optimizers
5  from tensorflow.keras import backend as K
6
7  SAVE_PATH          = './data/1_train/weights'
8  NUM_VALIDATION     = 16384
9  SHUFFLE_BUFFER_LEN = 65536
10 NUM_EPOCHS         = 100
11 LEARNING_RATE      = 0.0001
12 DR_RATE            = 0.05
13 BATCH_SIZE         = 32
```

```
14
15 def main():
16   x = np.load('./data/0_gen-train-data/x.npy')
17   y = np.load('./data/0_gen-train-data/y.npy')
18   y = y[..., np.newaxis]
19   train_set = tf.data.Dataset.from_tensor_slices((
20     x[NUM_VALIDATION:x.shape[0],:,:,:],
21     y[NUM_VALIDATION:y.shape[0],:,:,:] ))
22   train_set = train_set.shuffle(SHUFFLE_BUFFER_LEN)
23     .batch(BATCH_SIZE)
24   test_set  = tf.data.Dataset.from_tensor_slices((
25     x[0:NUM_VALIDATION,:,:,:],
26     y[0:NUM_VALIDATION,:,:,:] ))
27   test_set  = test_set.batch(BATCH_SIZE)
28
29   model = create_model()
30
31   opt = optimizers.SGD(learning_rate=LEARNING_RATE)
32   model.compile(optimizer=opt,loss=nts_in_mask)
33   model.summary()
34
35   checkpoints = callbacks.ModelCheckpoint('%s/{epoch:03d}.hdf5'
36     %SAVE_PATH, save_weights_only=False, verbose=1,
37     save_freq="epoch")
38
39   history = model.fit(train_set,epochs=NUM_EPOCHS,verbose=1,
40     callbacks=[checkpoints], validation_data=test_set)
```

The program loads the simulated dataset as `x` and `y`, separates them into training and validation datasets, and begins training at line 36. The call towards generating the architecture of the neural network is done in line 29. The `create_model()` function is shown below.

```
1  def create_model():
2    input_layer = layers.Input((32,32,4))
3    data_layer  = input_layer[:,:,:,0:3]
4    mask_layer  = input_layer[:,:,:,3]
5    mask_layer  = mask_layer[..., np.newaxis]
6
7    cnn   = data_layer
8
9    cnn_a = add_residual_block(16,cnn)
10   conn1 = layers.Dropout(rate=DR_RATE)(cnn_a)
11   cnn   = layers.Dropout(rate=DR_RATE)(cnn_a)
12
13   # ... more down-residual layers
14
```

```
15    cnn   = add_down_residual_layer(256,cnn)
16    cnn   = layers.Dropout(rate=DR_RATE)(cnn)
17
18    # ... more up-residual layers
19
20    cnn   = add_up_residual_layer(32,cnn)
21    cnn   = layers.Dropout(rate=DR_RATE)(cnn)
22    cnn   = layers.Concatenate(axis=3)([cnn,conn1])
23
24    cnn   = add_residual_block(32,cnn)
25    cnn   = add_residual_block(1,cnn)
26
27    cnn   = layers.Concatenate(axis=3)([cnn,mask_layer])
28
29    model = models.Model(input_layer,cnn)
30    return model
```

Every three lines of the code define a layer within the convolutional network and inserts a dropout layer for the downpath and the skip connections. The dimensions of the tensors can easily be changed here to modify the number of trainable parameters. The implementation of the down-residual layer, for example, is straight forward.

```
1  def add_down_residual_layer(output_dim,input_layer):
2    layer = add_residual_downsampling_block(output_dim,input_layer)
3    layer = add_residual_block(output_dim,layer)
4    return layer
5
6  def add_residual_block(output_dim,input_layer):
7    layer = input_layer
8    layer = layers.BatchNormalization()(layer)
9    layer = layers.ReLU()(layer)
10   layer = layers.Conv2D(output_dim, (3,3), padding='same')(layer)
11   layer = layers.BatchNormalization()(layer)
12   layer = layers.ReLU()(layer)
13   layer = layers.Conv2D(output_dim, (3,3), padding='same')(layer)
14   residual_layer = layers.Conv2D(output_dim, (3,3),
15   padding='same')(input_layer)
16   layer = layers.Add()([layer,residual_layer])
17   return layer
18
19 def add_residual_downsampling_block(output_dim,input_layer):
20   layer = input_layer
21   layer = layers.BatchNormalization()(layer)
22   layer = layers.ReLU()(layer)
23   layer = layers.Conv2D(output_dim, (3,3), strides=(2,2),
24   padding='same')(layer)
25   layer = layers.BatchNormalization()(layer)
```

```
26    layer = layers.ReLU()(layer)
27    layer = layers.Conv2D(output_dim, (3,3), padding='same')(layer)
28    residual_layer = layers.Conv2D(output_dim, (3,3), strides=(2,2),
29    padding='same')(input_layer)
30    layer = layers.Add()([layer,residual_layer])
31    return layer
```

Once the training is complete, the final neural network is saved on the local device, which can be used to predict phase snapshots from experimental diffraction patterns by loading the model and running the `model.predict()` function.

Chapter 5

Extending DL-CDI

5.1 Probe Function Recovery

CNN based methods for phase recovery relied on an *a priori* knowledge of the probe function before training, yet iterative methods such as ePIE fully incorporates probe function retrieval during the reconstruction process. In this section, I will discuss the potential of leveraging the generality of neural networks in order to recover the defocus and aberration parameters for electron probe functions. We do this by training a single CNN with a sufficiently large database of randomly generated stock images and probe functions with the goal of recovering the probe function directly from experimental diffraction patterns.

To begin, we first need to redo the training portion by simulating not only the diffraction pattern, but also varying the probe function along some parameters to encompass a wide range of possible variables that experimentalists might encounter during experimentation. The range of defocus and aberration parameters with which the probe functions were generated can be found in Table 5.1.

To compensate for both weakly and strongly scattering samples, a modified loss function

Parameter	Value
Binned detector dimensions (pixels)	32×32
Probe semi-convergence angle (pixels)	[3.2,12.8)
$C_{10}\alpha_{\max}^2/\lambda$ (probe defocus)	[-10.0,10)
$C_{12a}\alpha_{\max}^2/\lambda$ (probe twofold astigmatism)	[-10.0,10)
$C_{12b}\alpha_{\max}^2/\lambda$ (probe twofold astigmatism)	[-10.0,10)
$C_{21a}\alpha_{\max}^3/\lambda$ (probe coma)	[-10.0,10)
$C_{21b}\alpha_{\max}^3/\lambda$ (probe coma)	[-10.0,10)
$C_{23a}\alpha_{\max}^3/\lambda$ (probe threefold astigmatism)	[-10.0,10)
$C_{23b}\alpha_{\max}^3/\lambda$ (probe threefold astigmatism)	[-10.0,10)
RMSE of phase in mask	[0.,0.75)
Events per diffraction pattern	[3000,10000)
CNN input dimensions	$32 \times 32 \times 3$
Training sample size	1,500,000
Learning rate	1.e-4
Dropout rate	0.05
Epochs trained	100
Batch size	32
# of trainable parameters	9,445,988

Table 5.1: Experimental parameters of data sets used for deep learning CDI.

was used to normalize the L1 loss with the overall variance of the phase within the mask.

$$\text{loss} = \frac{\sum |y_{\text{pred}} - y_{\text{true}}|}{\sum |y_{\text{true}}| + \delta} \tag{5.1}$$

where y_{pred} is the CNN phase patch prediction within the region, and y_{pred} is the ground truth of the phase within the region, and δ was set at 0.01.

Of course, the definition of the mask must be standardized for different probe functions. The loss function above is summed over the region through which 90% of the power from the probe interacts with the sample. Test data generation and prediction by this universal CNN can be found in Figure 5.1. The utility of this neural network becomes apparent immediately, since one fully trained neural network can be used for a wide range of electron experiments, without ever training again. For all experimental results in this section, only one neural network was used to reconstruct all phase images.

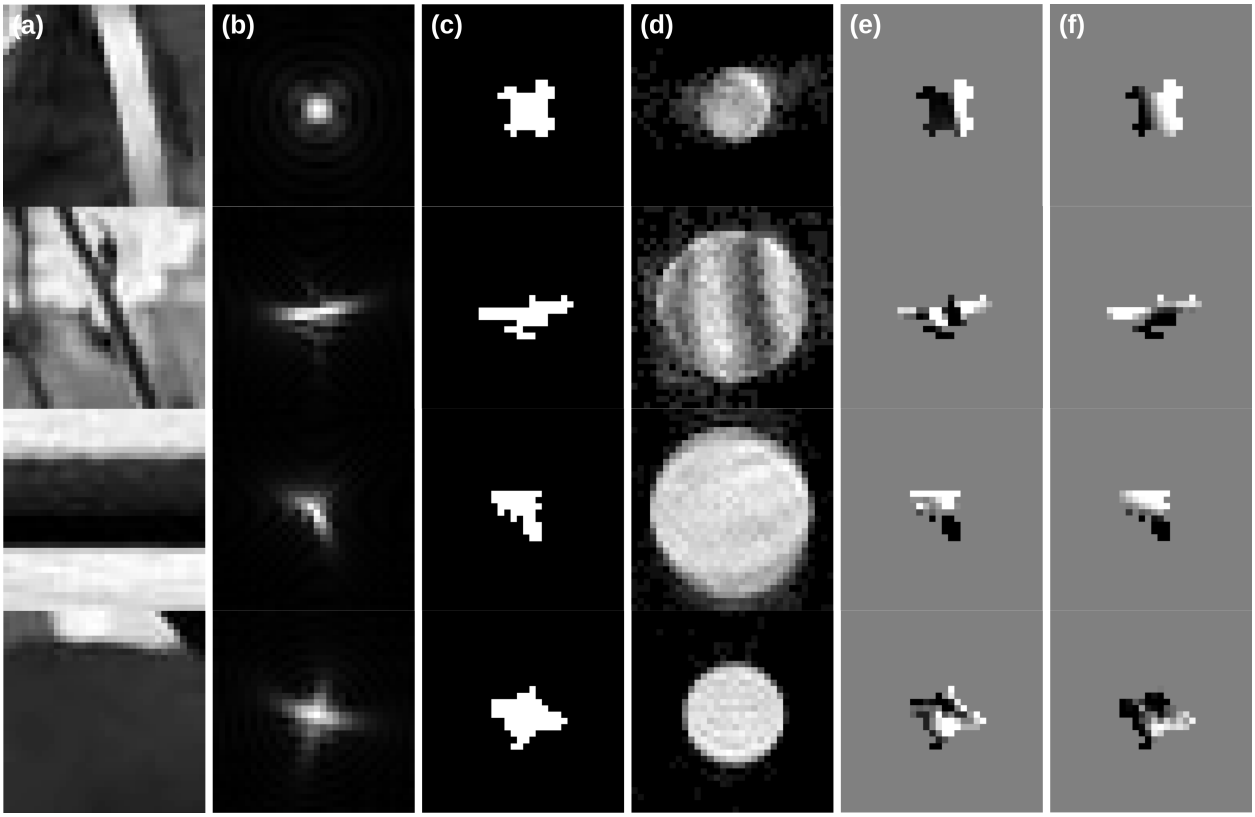

Figure 5.1: Examples of augmented data generation and CNN performance. (a) Random images from the internet used as pure phase objects to train a CNN. b) Real-space amplitudes of randomly generated probe functions parametrized by seven parameters. c) Real-space support based on the probe function. The phase object will only be predicted within this support, and any prediction outside of the support will not be used for stitching. d) Diffraction patterns generated from the phase objects in (a) and probe functions in (b). (e) Perfect phase patches within the illuminated areas. (f) Corresponding phase patches independently retrieved from the square root of the diffraction patterns by a trained CNN without any iteration.

One major advantage of the CNN taking the probe function as an input variable is that the feed-forward capabilities of neural networks can be directly leveraged to perform recovery of the probe function from experimental diffraction patterns. Since the CNN was trained with a range of probe functions parametrized by 7 defocus and aberration parameters, a trial probe function can be input along with pairs of adjacent diffraction patterns to evaluate the probe function by measuring agreements in the overlapped regions, as shown in Figure 5.2(a). This error signal is a useful metric to systematically update the probe parameters through either gradient descent or stochastic optimization methods. Since the CNN can achieve high throughput in phase snapshot evaluations, and measuring the error within overlapped regions can be parallelized, the probe function can be quickly recovered by explicitly using the overlap conditions. For all electron experiments, 100 pairs of diffraction patterns were used with a random optimization algorithm of 30 trial probe functions per iteration, and a well-defined probe function was recovered within 10 iterations. Figures 5.2(b) and 5.2(c) show the probe function before and after probe recovery for the experimental hBN sample, and Figure 5.2(d) shows the probe function recovered independently from ePIE.

To see the effects of the probe retrieval, Figures 5.3(a) and 5.3(c) show samples of the object function recovered from the CNN using probe functions from 5.2(b) and 5.2(c), respectively, with closeups in Figures 5.3(b) and 5.3(d), respectively.

With the universal CNN, we can reperform phase reconstructions of experimental diffraction patterns from previous chapters. After performing probe retrieval using the CNN, we can input the solved probe function as an input, retrieve the phase snapshots and perform the same stitching algorithm to obtain results in Figure 5.4. The same process was also performed on the Au nanoparticle sample with similar results (Figure 5.5), showing consistent ability to perform probe retrieval in various sample conditions.

However, the reconstruction process using the universal CNN ran into issues due to the out-of-time events associated with data acquisition from the detector initially used,

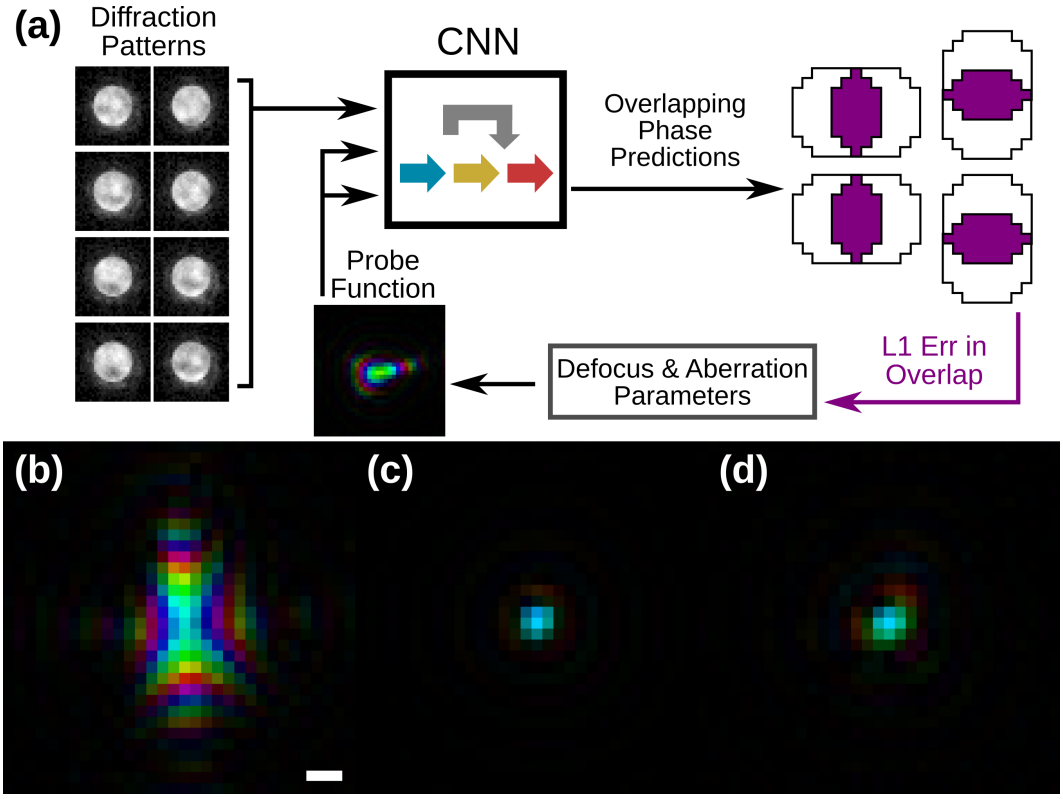

Figure 5.2: Taken from ref [2]. Diagram and results of probe recovery procedure with results for the experimental hBN data. a) The CNN makes predictions from pairs of adjacent diffraction patterns with a trial probe function. The cumulative L1 error within the overlapped regions is used as the fitness function to optimize the defocus and aberration parameters for the next iteration of the probe function. b) Initial guess of the probe function for the hBN sample. c) Resulting probe function after CNN probe retrieval process. d) Probe reconstruction from ePIE.

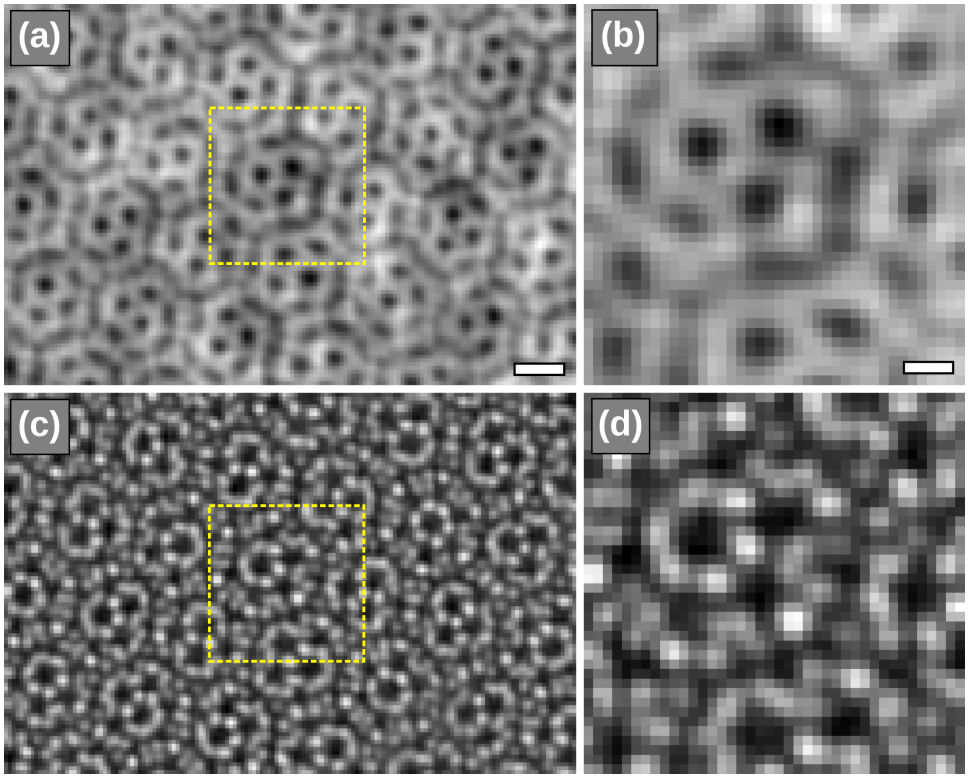

Figure 5.3: Taken from ref [2]. (a) Phase image of the twisted hBN interface retrieved by the CNN with the initial probe function. (b) Magnified view of the dotted square in (a). (c) Phase image of the twisted hBN interface retrieved by the CNN with the probe function in Figure 5.2(c). (d) Magnified view of the dotted square in (c).

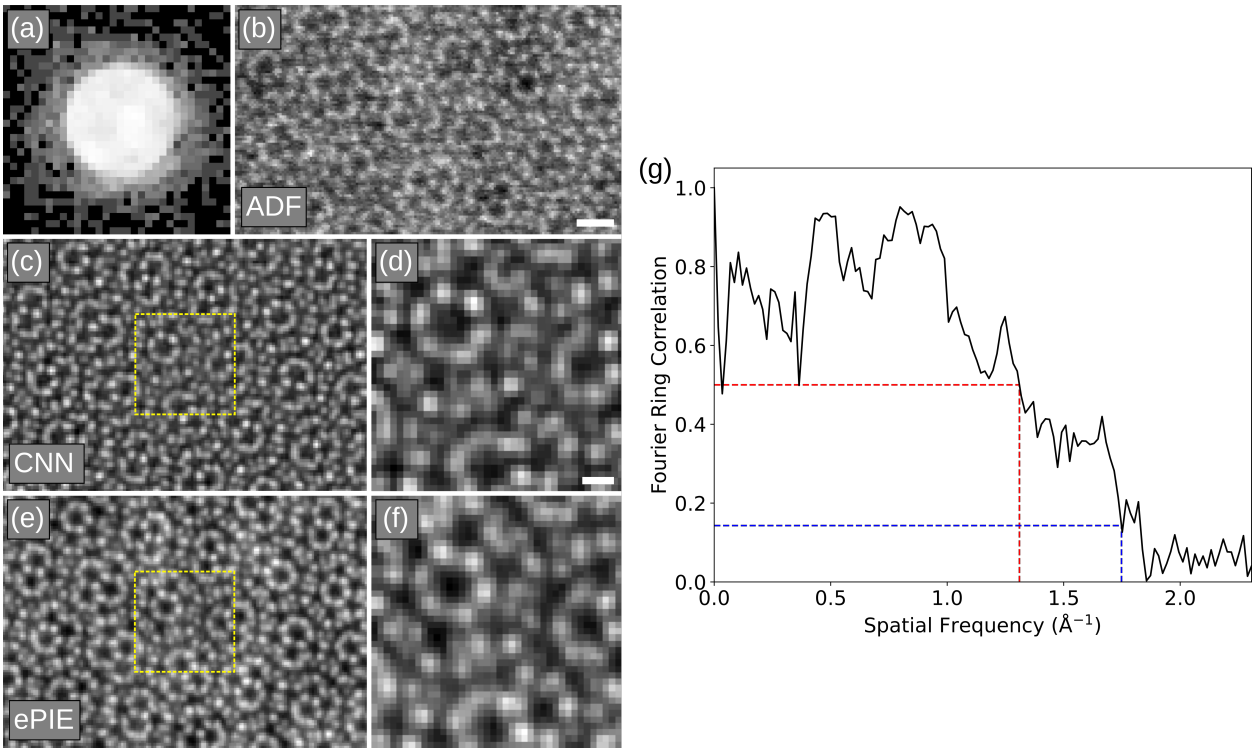

Figure 5.4: Diagram and results of probe recovery procedure with results for the experimental hBN data. a) The CNN makes predictions from pairs of adjacent diffraction patterns with a trial probe function. The cumulative L1 error within the overlapped regions is used as the fitness function to optimize the defocus and aberration parameters for the next iteration of the probe function. b) Initial guess of the probe function for the hBN sample. c) Resulting probe function after CNN probe retrieval process. d) Probe reconstruction from ePIE e) Stitched CNN reconstruction of hBN sample using the probe function in (b). f) Stitched CNN reconstruction of hBN sample using the probe in (c). g) ePIE reconstruction of hBN sample. Scale bars, 1 Å (b); and 2 Å (e).

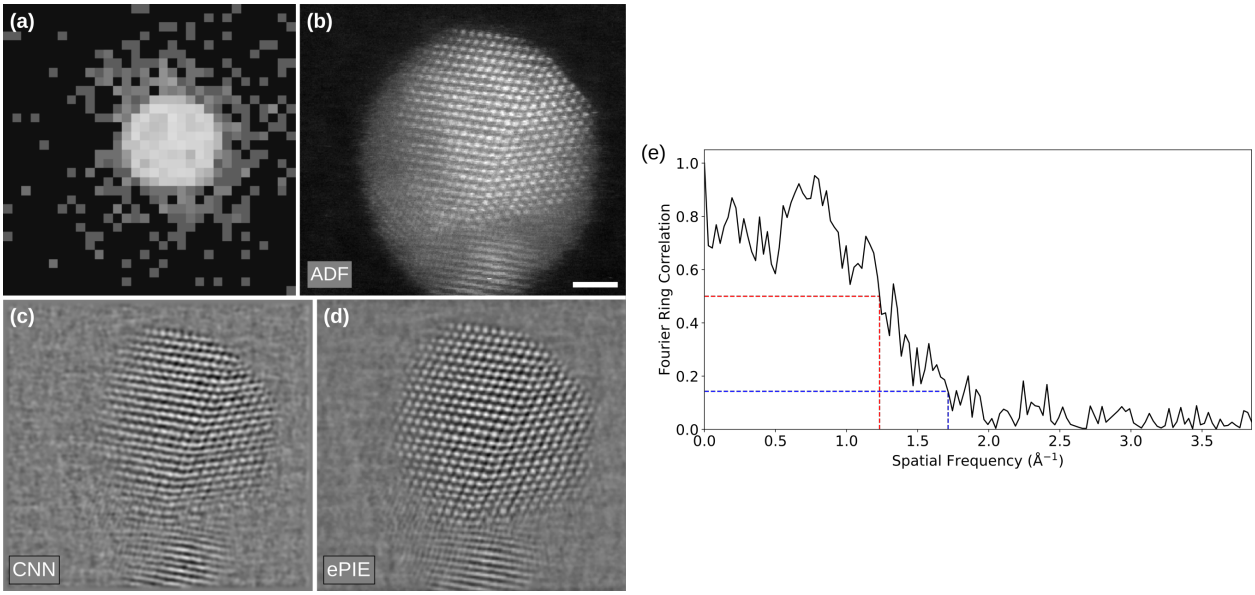

Figure 5.5: Deep learning phase retrieval of the experimental gold nanoparticle data. (a) Representative diffraction pattern acquired from a 5-nm gold nanoparticle (displayed in a logarithmic scale). (b) ADF-STEM image generated by integrating the diffraction intensity outside the bright-field disk at each scanning position. (c) Phase image of the gold nanoparticle retrieved by a trained CNN. (d) Phase image of the nanoparticle reconstructed by ePIE. (e) FRC curve between (c) and (d), where the red and blue dashed lines indicate a resolution of 0.81 and 0.58 Å, based on the cutoff criteria of FRC = 0.5 and 0.143, respectively. Scale bar, 1 nm.

Parameter	Value
Microscope	Nion UltraSTEMTM 100
Voltage (kV)	70
Convergence semi-angle (mrad)	32
Scanning step size (Å)	0.3125
Maximum collection angle after cropping (mrad)	46.5
Angular pixel size (mrad/pixel)	1.45
# probe positions	128×128
Dwell time per probe position (ms)	1.
Pixelated detector	Hamamatsu ORCA SCMOS
# detector pixels (before binning)	256×256
Dose (e- Å-2)	3.2×10^6

Table 5.2: Experimental parameters of data sets used for bilayer graphene experiment.

resulting in a strip of background counts. Although this did not affect reconstruction of the object when training a unique CNN for the experiment, the new method that utilizes one universal neural network for all experiments was sensitive to these background counts. Upon changing to a detector whose data does exhibit out-of-time events, the probe recovery step was successful. On the new detector, we performed reconstruction using a bilayer graphene sample, which still satisfies the thin sample (i.e. weak-phase-object) requirement. Reconstruction with various step sizes in bilayer graphene ptychographic data can be found in Figure 5.6, along with corresponding experimental variables in Table 5.2.

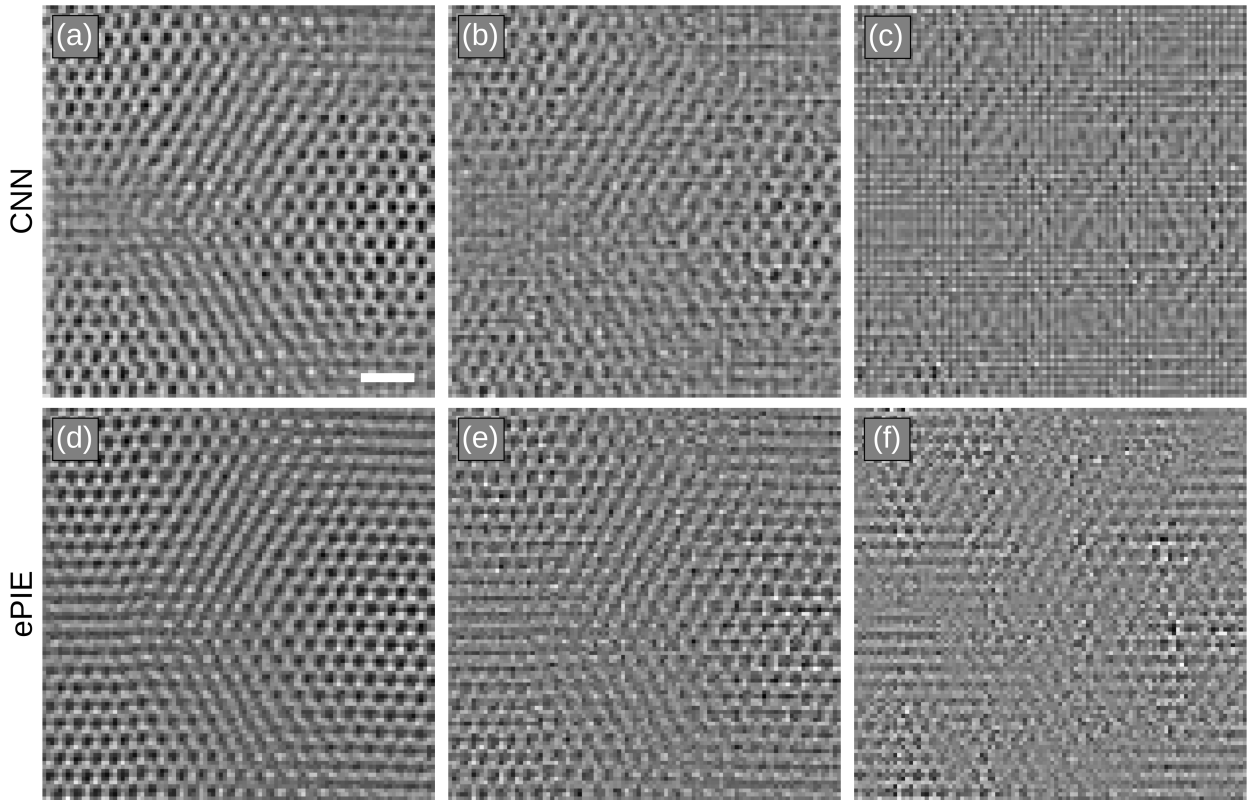

Figure 5.6: Deep learning phase retrieval of the experimental twisted bilayer graphene data with varying overlap of the diffraction patterns. (a-c) Stitched phase images of the twisted bilayer graphene retrieved by a trained CNN with a scanning step size of 0.313 Å, 0.625 Å, and 0.938 Å, respectively. (d-f) Phase images of the twisted bilayer graphene reconstructed by ePIE with the same scanning step sizes in (a-c). Scale bar, 5 Å.

5.2 Complex Object Recovery

So far, DL-CDI assumed a pure phase object with negligible absorption from the sample. As a result, the training dataset consisted of diffraction patterns scattered by pure phase objects only and significantly reduced the domain space which the neural network predicts from. However under real experimental conditions, wave functions interact with the sample by both refraction and absorption, so a generalized algorithm that can reconstruct both the phase and magnitude is important, especially for X-ray and optical ptychography experiments. While previous studies have demonstrated complex object CDI reconstructions via CNNs, these methods are only able to predict within a limited range of object functions by either limiting the geometries of the sample [88, 89], or by capitalizing on an inherent correlation between the magnitude and phase of the object function [84].

In this section, we present an extension of DL-CDI that enables reconstruction of complex object functions from ptychographic data without assuming any relation between the phase and magnitude of the sample. Complex object reconstructions from atomic-resolution experimental STEM ptychography data will be presented to demonstrate the efficacy of the method. Furthermore, results from numerical experiments and analyses of the neural networks will be shown to better understand the method's performance in non-ideal conditions.

The method of complex DL-CDI first requires the training of a CNN that maps a diffraction pattern to its corresponding object functions. The training dataset for the CNN can be generated by forward simulating with a probe and a numerical object function. An extensive collection of experimental data from a specific CDI setup is not required to properly train the inverse process. The first experimental realization of DL-CDI reconstructed pure phase objects where the magnitude signal was approximated as negligible, and the training procedure for such a CNN was straightforward due to the diffraction pattern being a direct result from phase fluctuations only. However, extending the methodology to encompass both the

phase and the magnitude becomes a harder problem to solve since the neural network needs to predict the correct phase function given any possible magnitude function, and vice versa.

Figure 5.7(a) shows the architecture of the neural network used for complex object reconstruction, along with detailed schematics of the layers in Figure 5.7(b). Previous studies of complex object reconstruction via CNN incorporated architectures that share the encoding layers before branching out to two decoding arms that reconstruct the phase and the magnitude separately [84, 88]. However, we observed superior reconstruction quality from CNN architectures that completely separate the encoding and decoding for the phase and the magnitude reconstructions.

An important advantage of this architecture can be seen from the loss function used during training.

$$\text{Loss} = |\phi_{\text{true}} - \phi_{\text{pred}}| + |\log(\mu_{\text{true}}) - \log(\mu_{\text{pred}})| \tag{5.2}$$

where ϕ_{true} and μtrue are the ground truth phase and magnitudes, respectively, and ϕ_{pred} and μ_{pred} are the CNN predictions of the phase and magnitudes, respectively. The CNN is trained to retrieve the logarithm of the magnitude signal due to the fact that magnitude functions must be positive definite. Usually, additive terms in the loss function need to be modulated by an incorporation of some constant factor controlling the impact of one term over the other. However, since the two branches of our architecture are completely separable, the back-propagation of error from one side does not update any parameters in the other. In short, the independence of each branch ensures that the user is not forced to subjectively trade off accuracy between the two image outputs.

Once the CNN is fully trained, it is able to return snapshots of the phase and magnitude images within a support defined by the probe illumination function. A separate stitching method must be performed on these overlapping snapshots to reconstruct a large field-of-view image from a matrix of diffraction patterns. The stitching algorithm for the phase image is

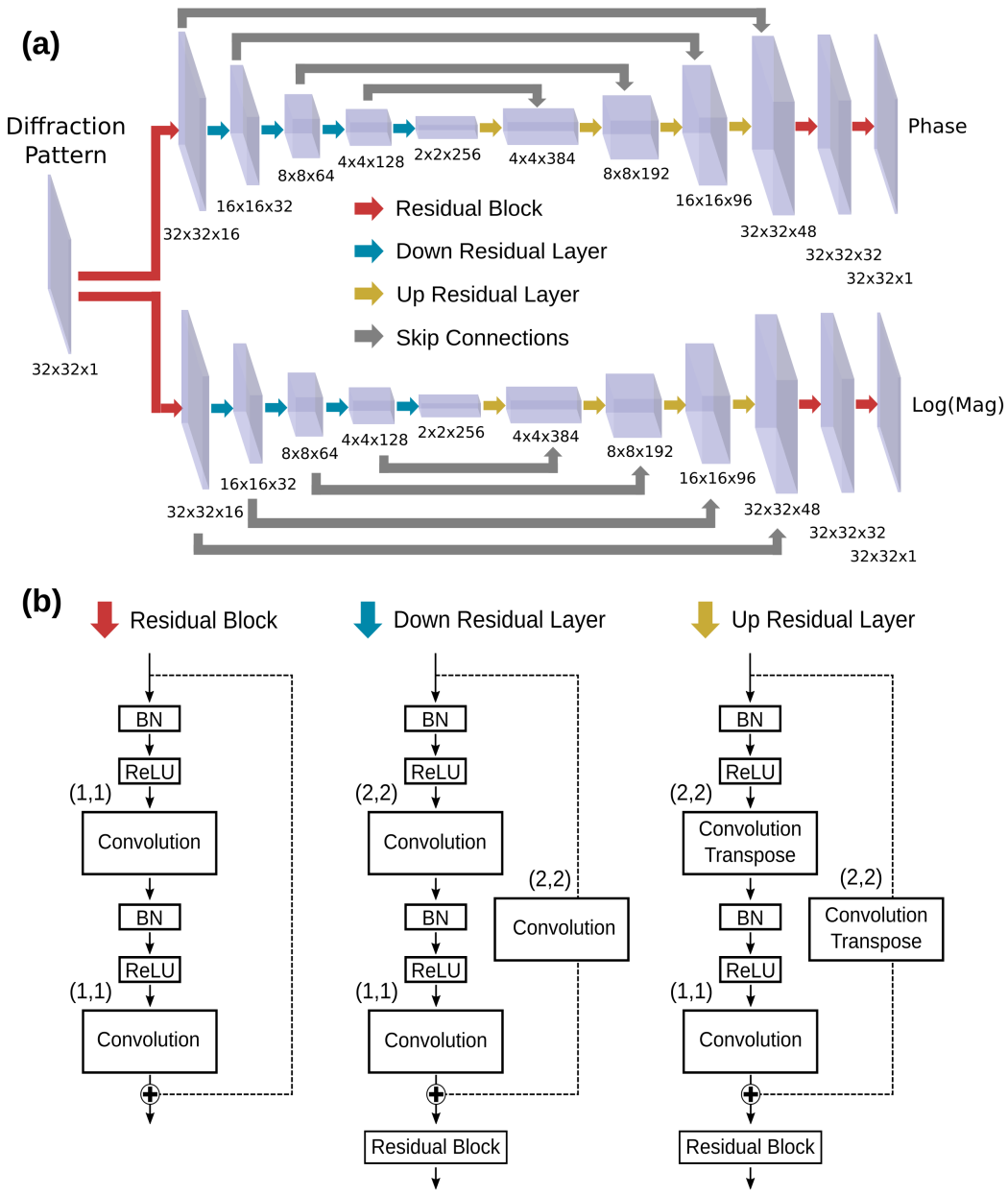

Figure 5.7: (a) Convolutional neural network architecture for complex object reconstruction in DL-CDI. Two entirely separable networks are trained from diffraction patterns simulated with a complex object and no encoded layers are shared between the two branches. The CNN is trained to return the logarithm of the magnitude function to ensure that the magnitude is positive definite. (b) Schematic of residual layers used in the CNN architecture. Different strides used in the convolution filters are shown in parentheses. All filters are size 3×3.

described in a previous section [2], where an iterative algorithm is necessary to merge adjacent phase images with differing overall phase constants. However, if the flux of the illumination function is known a priori, the attenuation of the probe from the sample is measurable at the detector and therefore an iterative procedure is not required to stitch the magnitude snapshots. Instead, calculating the mean of the logarithms (therefore the geometric mean) of the magnitude predictions from all pertaining snapshots at a given position is sufficient in stitching the magnitude function.

To examine the performance of the CNN in complex object reconstructions from experimental data, we performed a ptychographic experiment on the 5 nm gold nanoparticle ptychographic data presented in previous sections. Like before, a probe function was analytically generated with a second order aberration function, and a CNN was trained from 1,500,000 simulated diffraction patterns from 1,500,000 pairs of uncorrelated stock images for the phase and magnitude. Figures 5.8(a) and 5.8(b) show the phase and the magnitude reconstructions from the CNN, respectively, and comparison reconstructions of the phase and the magnitude from ePIE are shown in Figures 5.8(c) and 5.8(d), respectively. The CNN was able to independently reconstruct the fluctuations in magnitudes along the zone axis, and the minimum magnitude values at the atomic columns were consistent between CNN and ePIE at around 0.7 to 0.8. A slight difference in the reconstruction of the background was observed in the phase images, with ePIE demonstrating a lower sensitivity to the amorphous carbon background.

Next, we performed experimental complex object reconstructions on the hBN experimental ptychographic data presented in previous sections. Like the gold nanoparticle sample, a probe function was analytically generated up to second order aberration to train a CNN with 1,500,000 pairs of uncorrelated images their corresponding diffraction patterns. Figures 5.9(a) and 5.9(b) show the reconstructed phase images from CNN and ePIE, respectively, and figures 5.9(c) and 5.9(d) show the reconstructed magnitude images from CNN and ePIE,

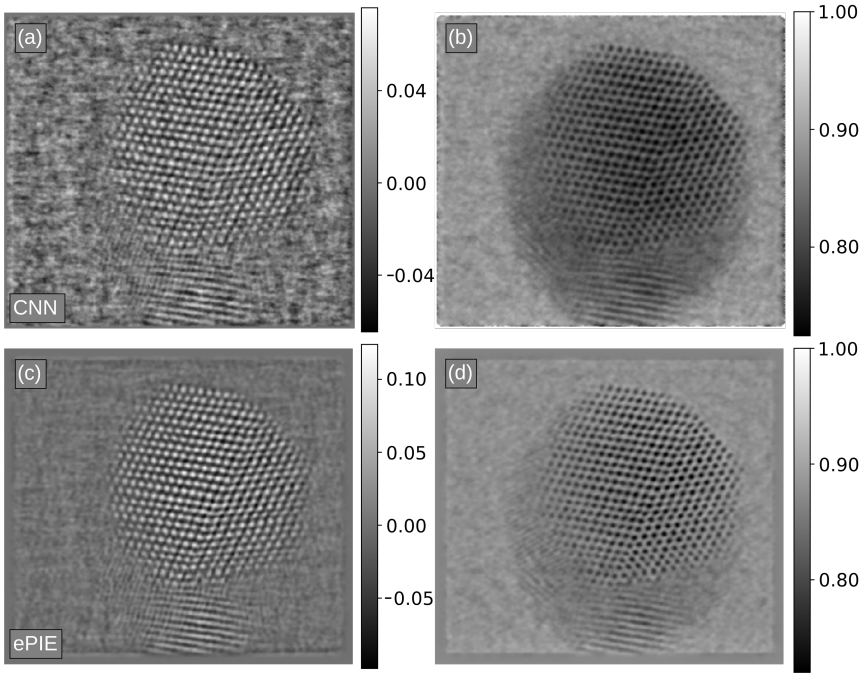

Figure 5.8: Deep learning complex object phase retrieval of the experimental data of a gold nanoparticle. (a) Stitched phase and (b) stitched magnitude reconstructions by a trained CNN. (c) Phase and (d) magnitude reconstructions by ePIE.

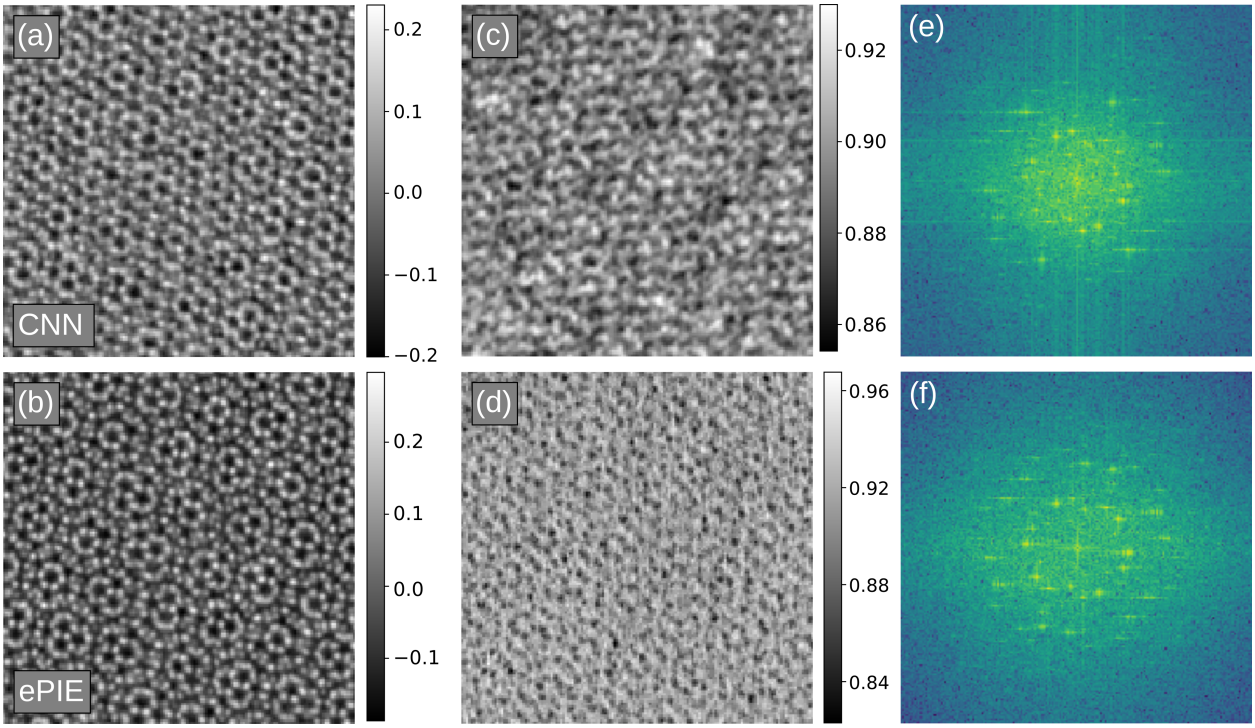

Figure 5.9: Deep learning complex object phase retrieval of the experimental data of a twisted hBN interface. (a), (b) Reconstructed phase images from CNN and ePIE, respectively. (c), (d) Reconstructed magnitude images from CNN and ePIE, respectively. (e), (f) 2D Fourier transform of the magnitude reconstructions in (c) and (d), respectively.

respectively. The magnitude reconstructions differed at high spatial frequencies, but from the plots of the 2D Fourier transform of the magnitude reconstructions as shown in figures 5.9(e) and 5.9(f), we were able to observe a high degree of similarity in the Fourier peaks.

A numerical study on the effect of shot noise on DL-CDI complex reconstructions was performed by simulating a ptychographic experiment on a complex object comprised of two images. Three sets of diffraction patterns were generated with varying degrees of shot noise: one with infinite dose, one with 1e4 events per pattern, and one with 1e3 events per pattern. A CNN was trained with 1,500,000 infinite dose diffraction patterns simulated from a probe function known a priori, with dropout layers of frequency at 0.2 and a learning rate of 1e-4 for 100 iterations. After fully training, the CNN was used to generate phase and magnitude snapshots which were stitched together to reconstruct images shown in Figures 5.10(a),

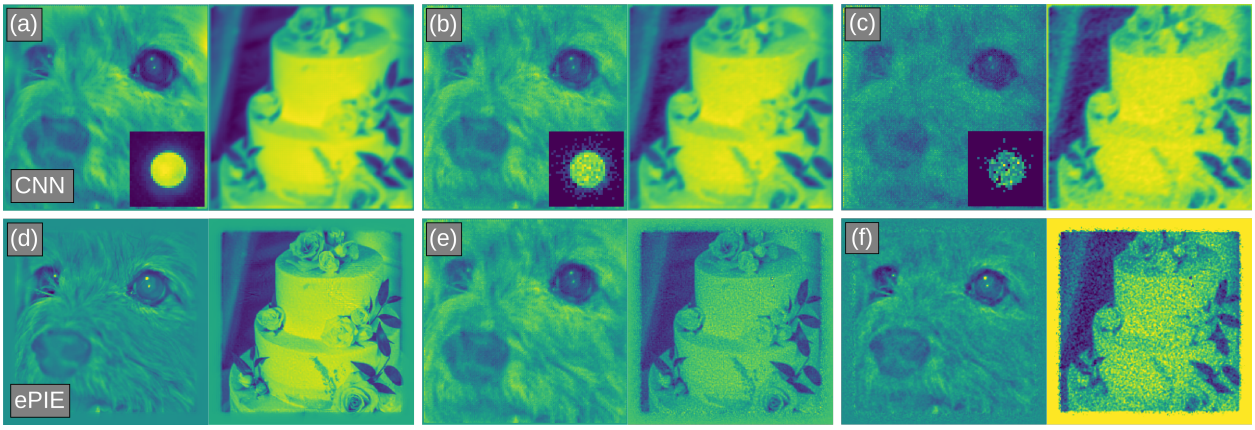

Figure 5.10: (a), (b), (c) Phase (left) and magnitude (right) DL-CDI reconstructions from a simulated ptychographic dataset with doses of infinite, 1e4 and 1e3 events per pattern, respectively. Insets show a representative diffraction pattern. (d), (e), (f) ePIE reconstructions from the same simulated ptychographic dataset with doses of infinite, 1e4 and 1e3 events per pattern, respectively.

5.10(b), and 5.10(c) at different shot noise levels. Representative diffraction patterns for each ptychographic dataset are shown in the insets. We observed robustness in the CNN in reconstructing both the phase and magnitude images from highly corrupted diffraction patterns, even at a low dose of 1000 patterns per diffraction pattern. As a comparison, ePIE reconstructions from the same datasets are shown in Figures 5.10(d), 5.10(e) and 5.10(f).

One important note from this numerical study was the use of just a singular CNN trained with infinite dose diffraction patterns to predict phase and magnitude snapshots for a range of shot noise levels in diffraction patterns. An attempt was made in simulating and training with diffraction patterns corrupted by shot noise (assuming that the electron dose is known a priori), but observed poorer quality of reconstructions than from a CNN trained with infinite dose simulated patterns with high-rate (0.2) dropout layers. This may suggest that best method for compensating for stochastic processes encountered during experimentation is applying conventional techniques used to suppress over-fitting such as batch normalization, dropout layers, and regularization [90, 91].

To better understand the internal mechanisms of CNN-based ptychographic methods,

we analyzed neural networks by visualizing maximally activated patterns (MAPs). MAPs are input patterns that maximize activation in a specific layer within the network and are computed by performing gradient descent on the layer outputs with respect to the input pattern [92–94]. To formalize mathematically, let θ represent our neural network weights, and let $h_{ij}(\theta, \mathbf{x})$ be the activation of a given unit i from a given layer j with $\mathbf{x}$ as an input vector. Calculating the MAP for a given input would be finding $\mathbf{x^*}$ such that

$$\mathbf{x^*} = \arg \max_{\mathbf{x} s.t. ||\mathbf{x}||=\rho} h_{ij}(\theta, \mathbf{x}) \tag{5.3}$$

where ρ enforces a bounded norm.

The patterns that emerge from this maximization process enable qualitative visualization of how a neural network selects important features. The first half of the CNN architecture shown in Figure 5.7 is a series of convolutional layers followed by downsampling layers to encode an input image into 256, 2-by-2 feature representations. For our analysis, MAPs were calculated at this output layer from two CNNs trained with different probe functions to explore the effect of physical experimental conditions in network performance.

Figure 5.11(a) shows a sample focused probe function used to train a CNN for DL-CDI. MAPs at the 2-by-2 feature representation layer for this CNN is shown in Figure 5.11(b). We observed a lack of fine features in the patterns, indicating that the phase retrieval process does not heavily rely on local operations. Instead, smooth patterns with a low number of peaks were present which were consistent with typical focused probe STEM diffraction patterns consisting of diffraction discs and lack of fine features.

Next, another CNN was trained with a de-focused probe function, shown in Figure 5.11(c), which would increase the field-of-view per diffraction pattern. Figure 5.11(d) shows 25 MAPs at the 2-by-2 feature representation layer for this CNN. Usually, de-focusing a probe function results in stripes and finer features in the diffraction patterns, therefore if the

CNN was able to encapsulate these features, the MAPs would comprise of patterns with finer features than those from a focused probe. Instead we observed that the CNN was unable to encode these features and some filters at the 2-by-2 feature layer were turned off entirely, evident from completely dark MAPs. This suggests that increasing the number of trainable parameters to compensate for the increased complexity is not a productive response and that the archetypal CNN structure might be a sub-optimal choice for encoding and decoding measurements in the Fourier domain.

In conclusion, we extended the generality of deep-learning based CDI by modifying the architecture to compensate for both the phase and magnitude of the object function. Training a CNN with two completely independent branches allows for a loss function that does not trade-off quality between the phase and the magnitude while not assuming any correlation. The method was applied on experimental ptychographic data on twisted hexagonal boron nitride (hBN) flakes and a gold nanoparticle and demonstrated consistent complex object reconstructions when compared to conventional iterative ptychographic methods. A numerical experiment on the method's tolerance of shot noise demonstrated that computational methods to reduce over-fitting could be sufficient in compensating for stochastic processes present in experimental conditions.

Finally, an analysis was performed on the CNN by calculating maximally activated patterns for filters within the network. Visualizing MAPs of comparable neural networks showed under-fitting of features for a CNN tasked with learning a more complex inverse function. While deep convolutional neural networks represent the state-of-the-art for computer vision tasks, encoding and decoding measurements of CDI experiments in Fourier space might be non-ideal for a network architecture primarily designed for understanding real-space relations within images.

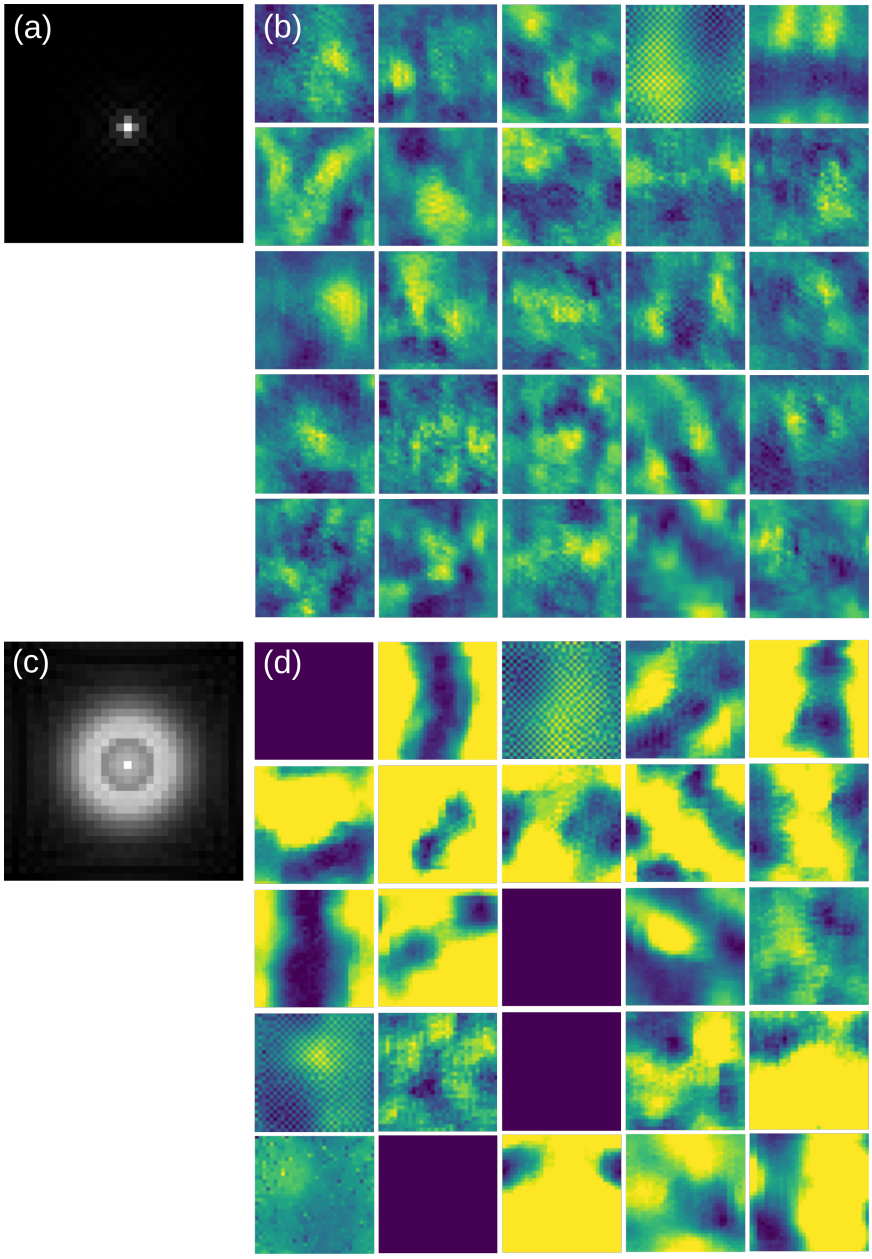

Figure 5.11: Maximally activated patterns (MAPs) from CNNs trained with varying defocus. (a) A focused probe function used to train a CNN for DL-CDI. (b) 25 representative MAPs from CNN trained with probe function in (a). (c) A de-focused probe function. (d) 25 representative MAPs from CNN trained with probe function in (c).

Chapter 6

Summary

We demonstrated deep learning CDI at sub-Angström resolution using CNNs trained only with stock images from the internet. The CNNs were then coupled with a stitching algorithm to directly retrieve the phase images of twisted hBN, monolayer graphene, and a gold nanoparticle from experimental electron diffraction patterns. Quantitative analysis using FRC curves indicates that the phase images recovered by deep learning CDI have comparable quality to those reconstructed by ePIE. The spatial resolution of these phase images was quantified by the FRC to be in the range of 0.71-0.53 Å.

Compared with iterative algorithms such as ePIE that use overlapping regions to reconstruct the phase information, deep learning CDI independently recovers phase patches at different scanning positions by CNNs and then stitches them together to form a phase image, which can be implemented in real-time with a computational time orders of magnitude faster than iterative algorithms. In contrast to non-iterative ptychographic methods that require small scan step sizes (typically ¡0.5 Å) to sample at the Nyquist limit, deep learning CDI can be implemented with larger scan step sizes to achieve larger fields of view, which can be tuned by defocusing the probe. Additionally, it can reach higher spatial resolutions than the non-iterative methods as the resolution of deep learning CDI is only limited by the spatial frequency of the diffraction signal.

We further expanded the method of DL-CDI to allow probe function recovery directly from diffraction patterns and to compensate for varying absorption in the sample. To recover the probe function, we were able to train a CNN that takes the probe function as an input variable. This CNN can then be used to iteratively solve for the experimental probe function by searching across 7 defocus and aberration parameters. Redoing the electron DL-CDI experiments with this approach showed promising results of a universal neural network that only needed to be trained once to make predictions for a variety of experiments.

Furthermore, by employing multiple channels in the CNN, this thesis laid out the foundation for complex objection function recovery. Complex objects that contain information about both the phase induced but also the intensity attenuation require two independent channels that invert information encoded in the diffraction pattern. Simulation and experimental results indicate the feasibility of this method and is expected to play a pivotal role in coherent diffractive imaging for a wide range of modalities.

With further development, we expect that deep learning CDI could become an important tool for real-time, atomic-scale imaging of a wide range of samples across different disciplines. One potential application of incorporating phase imaging in scanning transmission electron microscopy is leveraging improved signal from low-Z atoms and linearity in atomic projections. Ptychographic reconstructions from 4D-STEM data, which can be accelerated via deep learning, can offer significant advantages when performing electron tomography at the atomic scale. As high quality pixel array electron detectors become more commercially viable and data storage and management become more streamlined, we envision these computational methods as powerful tools for studying the 3D atomic structures of low-Z and radiation sensitive materials such as transition metal oxides, functional 2D heterostructures and glasses.

Bibliography

[1] Dillan J. Chang et al. "Ptychographic atomic electron tomography: Towards three-dimensional imaging of individual light atoms in materials". en. In: *Physical Review B* 102.17 (Nov. 2020), p. 174101. ISSN: 2469-9950, 2469-9969. DOI: 10.1103/PhysRevB.102.174101. URL: https://link.aps.org/doi/10.1103/PhysRevB.102.174101 (visited on 02/21/2021).

[2] Dillan J. Chang et al. "Deep-Learning Electron Diffractive Imaging". en. In: *Physical Review Letters* 130.1 (Jan. 2023), p. 016101. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.130.016101. URL: https://link.aps.org/doi/10.1103/PhysRevLett.130.016101 (visited on 01/08/2023).

[3] J. Miao, P. Ercius, and S. J. L. Billinge. "Atomic electron tomography: 3D structures without crystals". en. In: *Science* 353.6306 (Sept. 2016), aaf2157–aaf2157. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.aaf2157. URL: https://www.sciencemag.org/lookup/doi/10.1126/science.aaf2157 (visited on 02/20/2020).

[4] M. C. Scott et al. "Electron tomography at 2.4-ångström resolution". en. In: *Nature* 483.7390 (Mar. 2012), pp. 444–447. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature10934. URL: http://www.nature.com/articles/nature10934 (visited on 02/20/2020).

[5] Chien-Chun Chen et al. "Three-dimensional imaging of dislocations in a nanoparticle at atomic resolution". en. In: *Nature* 496.7443 (Apr. 2013), pp. 74–77. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature12009. URL: http://www.nature.com/articles/nature12009 (visited on 02/20/2020).

[6] Jihan Zhou et al. "Observing crystal nucleation in four dimensions using atomic electron tomography". en. In: *Nature* 570.7762 (June 2019), pp. 500–503. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-019-1317-x. URL: http://www.nature.com/articles/s41586-019-1317-x (visited on 02/20/2020).

[7] Yongsoo Yang et al. "Deciphering chemical order/disorder and material properties at the single-atom level". en. In: *Nature* 542.7639 (Feb. 2017), pp. 75–79. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature21042. URL: http://www.nature.com/articles/nature21042 (visited on 02/20/2020).

[8] Rui Xu et al. "Three-dimensional coordinates of individual atoms in materials revealed by electron tomography". en. In: *Nature Materials* 14.11 (Nov. 2015), pp. 1099–1103. ISSN: 1476-1122, 1476-4660. DOI: 10.1038/nmat4426. URL: http://www.nature.com/articles/nmat4426 (visited on 02/20/2020).

[9] Xuezeng Tian et al. "Correlating the three-dimensional atomic defects and electronic properties of two-dimensional transition metal dichalcogenides". en. In: *Nature Materials* (Mar. 2020). ISSN: 1476-1122, 1476-4660. DOI: 10.1038/s41563-020-0636-5. URL: http://www.nature.com/articles/s41563-020-0636-5 (visited on 07/01/2020).

[10] Bart Goris et al. "Three-Dimensional Elemental Mapping at the Atomic Scale in Bimetallic Nanocrystals". en. In: *Nano Letters* 13.9 (Sept. 2013), pp. 4236–4241. ISSN: 1530-6984, 1530-6992. DOI: 10.1021/nl401945b. URL: https://pubs.acs.org/doi/10.1021/nl401945b (visited on 10/28/2020).

[11] Earl J. Kirkland. *Advanced Computing in Electron Microscopy*. en. Cham: Springer International Publishing, 2020. ISBN: 978-3-030-33259-4 978-3-030-33260-0. DOI: 10.1007/978-3-030-33260-0. URL: http://link.springer.com/10.1007/978-3-030-33260-0 (visited on 07/25/2022).

[12] Stephen J. Pennycook and Peter D. Nellist, eds. *Scanning Transmission Electron Microscopy*. en. New York, NY: Springer New York, 2011. ISBN: 978-1-4419-7199-9 978-1-4419-7200-2. DOI: 10.1007/978-1-4419-7200-2. URL: http://link.springer.com/10.1007/978-1-4419-7200-2 (visited on 02/25/2020).

[13] P.A. Midgley and M. Weyland. "3D electron microscopy in the physical sciences: the development of Z-contrast and EFTEM tomography". en. In: *Ultramicroscopy* 96.3-4 (Sept. 2003), pp. 413–431. ISSN: 03043991. DOI: 10.1016/S0304-3991(03)00105-0. URL: https://linkinghub.elsevier.com/retrieve/pii/S0304399103001050 (visited on 02/25/2020).

[14] James M. LeBeau et al. "Quantitative Atomic Resolution Scanning Transmission Electron Microscopy". en. In: *Physical Review Letters* 100.20 (May 2008), p. 206101. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.100.206101. URL: https://link.aps.org/doi/10.1103/PhysRevLett.100.206101 (visited on 02/25/2020).

[15] Joachim Frank. *Three-Dimensional Electron Microscopy of Macromolecular Assemblies*. Oxford University Press, Mar. 2006. ISBN: 978-0-19-518218-7. DOI: 10.1093/acprof:oso/9780195182187.001.0001. URL: https://academic.oup.com/book/6747 (visited on 02/24/2023).

[16] Avinash C. Kak and Malcolm Slaney. *Principles of Computerized Tomographic Imaging*. en. Society for Industrial and Applied Mathematics, Jan. 2001. ISBN: 978-0-89871-494-4 978-0-89871-927-7. DOI: 10.1137/1.9780898719277. URL: http:

//epubs.siam.org/doi/book/10.1137/1.9780898719277 (visited on 02/24/2023).

[17] Richard Gordon, Robert Bender, and Gabor T. Herman. "Algebraic Reconstruction Techniques (ART) for three-dimensional electron microscopy and X-ray photography". en. In: *Journal of Theoretical Biology* 29.3 (Dec. 1970), pp. 471–481. ISSN: 00225193. DOI: 10.1016/0022-5193(70)90109-8. URL: https://linkinghub.elsevier.com/retrieve/pii/0022519370901098 (visited on 02/24/2023).

[18] Peter Gilbert. "Iterative methods for the three-dimensional reconstruction of an object from projections". en. In: *Journal of Theoretical Biology* 36.1 (July 1972), pp. 105–117. ISSN: 00225193. DOI: 10.1016/0022-5193(72)90180-4. URL: https://linkinghub.elsevier.com/retrieve/pii/0022519372901804 (visited on 02/24/2023).

[19] Jianwei Miao, Friedrich Förster, and Ofer Levi. "Equally sloped tomography with oversampling reconstruction". en. In: *Physical Review B* 72.5 (Aug. 2005), p. 052103. ISSN: 1098-0121, 1550-235X. DOI: 10.1103/PhysRevB.72.052103. URL: https://link.aps.org/doi/10.1103/PhysRevB.72.052103 (visited on 02/24/2023).

[20] Alan Pryor et al. "GENFIRE: A generalized Fourier iterative reconstruction algorithm for high-resolution 3D imaging". en. In: *Scientific Reports* 7.1 (Dec. 2017), p. 10409. ISSN: 2045-2322. DOI: 10.1038/s41598-017-09847-1. URL: http://www.nature.com/articles/s41598-017-09847-1 (visited on 02/20/2020).

[21] J. Miao et al. "Beyond crystallography: Diffractive imaging using coherent x-ray light sources". en. In: *Science* 348.6234 (May 2015), pp. 530–535. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.aaa1394. URL: http://www.sciencemag.org/cgi/doi/10.1126/science.aaa1394 (visited on 05/29/2019).

[22] Craig Hoffman and Ronald Driggers, eds. *Encyclopedia of Optical and Photonic Engineering, Second Edition*. en. 0th ed. CRC Press, Sept. 2015. ISBN: 978-1-4398-5099-2. DOI: 10.1081/E-EOE2. URL: https://www.taylorfrancis.com/books/9781351247177 (visited on 07/19/2022).

[23] Jianwei Miao et al. "Extending the methodology of X-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens". en. In: *Nature* 400.6742 (July 1999), pp. 342–344. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/22498. URL: http://www.nature.com/articles/22498 (visited on 02/20/2020).

[24] J. R. Fienup. "Reconstruction of an object from the modulus of its Fourier transform". en. In: *Optics Letters* 3.1 (July 1978), p. 27. ISSN: 0146-9592, 1539-4794. DOI: 10.1364/OL.3.000027. URL: https://opg.optica.org/abstract.cfm?URI=ol-3-1-27 (visited on 07/20/2022).

[25] J. Miao, D. Sayre, and H. N. Chapman. "Phase retrieval from the magnitude of the Fourier transforms of nonperiodic objects". en. In: *Journal of the Optical Society of America A* 15.6 (June 1998), p. 1662. ISSN: 1084-7529, 1520-8532. DOI: 10.1364/JOSAA.15.001662. URL: https://www.osapublishing.org/abstract.cfm?URI=josaa-15-6-1662 (visited on 04/04/2021).

[26] R. H. T. Bates. "Fourier phase problems are uniquely solvable in more than one dimension. I: Underlying theory". In: *Optik (Stuttg.)* 61.3 (1982), pp. 247–262.

[27] W. Hoppe. "Beugung im inhomogenen Primärstrahlwellenfeld. I. Prinzip einer Phasenmessung von Elektronenbeugungsinterferenzen". In: *Acta Crystallographica Section A* 25.4 (July 1969), pp. 495–501. ISSN: 0567-7394. DOI: 10.1107/S0567739469001045. URL: http://scripts.iucr.org/cgi-bin/paper?S0567739469001045 (visited on 02/20/2020).

[28] J. M. Rodenburg and H. M. L. Faulkner. "A phase retrieval algorithm for shifting illumination". en. In: *Applied Physics Letters* 85.20 (Nov. 2004), pp. 4795–4797. ISSN: 0003-6951, 1077-3118. DOI: 10.1063/1.1823034. URL: http://aip.scitation.org/doi/10.1063/1.1823034 (visited on 02/20/2020).

[29] J. M. Rodenburg et al. "Hard-X-Ray Lensless Imaging of Extended Objects". en. In: *Physical Review Letters* 98.3 (Jan. 2007), p. 034801. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.98.034801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.98.034801 (visited on 05/29/2019).

[30] Andrew M. Maiden and John M. Rodenburg. "An improved ptychographical phase retrieval algorithm for diffractive imaging". en. In: *Ultramicroscopy* 109.10 (Sept. 2009), pp. 1256–1262. ISSN: 03043991. DOI: 10.1016/j.ultramic.2009.05.012. URL: https://linkinghub.elsevier.com/retrieve/pii/S0304399109001284 (visited on 05/29/2019).

[31] Martin Dierolf et al. "Ptychographic X-ray computed tomography at the nanoscale". en. In: *Nature* 467.7314 (Sept. 2010), pp. 436–439. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature09419. URL: http://www.nature.com/articles/nature09419 (visited on 05/29/2019).

[32] Si Gao et al. "Electron ptychographic microscopy for three-dimensional imaging". en. In: *Nature Communications* 8.1 (Dec. 2017), p. 163. ISSN: 2041-1723. DOI: 10.1038/s41467-017-00150-1. URL: http://www.nature.com/articles/s41467-017-00150-1 (visited on 05/29/2019).

[33] H. Yang et al. "Simultaneous atomic-resolution electron ptychography and Z-contrast imaging of light and heavy elements in complex nanostructures". en. In: *Nature Communications* 7.1 (Nov. 2016), p. 12532. ISSN: 2041-1723. DOI: 10.1038/ncomms12532. URL: http://www.nature.com/articles/ncomms12532 (visited on 02/21/2020).

[34] M.J. Humphry et al. "Ptychographic electron microscopy using high-angle dark-field scattering for sub-nanometre resolution imaging". en. In: *Nature Communications* 3.1 (Jan. 2012), p. 730. ISSN: 2041-1723. DOI: 10.1038/ncomms1733. URL: http://www.nature.com/articles/ncomms1733 (visited on 02/21/2020).

[35] Peng Wang et al. "Electron Ptychographic Diffractive Imaging of Boron Atoms in LaB6 Crystals". en. In: *Scientific Reports* 7.1 (Dec. 2017), p. 2857. ISSN: 2045-2322. DOI: 10.1038/s41598-017-02778-x. URL: http://www.nature.com/articles/s41598-017-02778-x (visited on 02/21/2020).

[36] Peng Li and Andrew Maiden. "Multi-slice ptychographic tomography". en. In: *Scientific Reports* 8.1 (Dec. 2018), p. 2049. ISSN: 2045-2322. DOI: 10.1038/s41598-018-20530-x. URL: http://www.nature.com/articles/s41598-018-20530-x (visited on 06/29/2020).

[37] Arjun Rana et al. "Potential of Attosecond Coherent Diffractive Imaging". en. In: *Physical Review Letters* 125.8 (Aug. 2020), p. 086101. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.125.086101. URL: https://link.aps.org/doi/10.1103/PhysRevLett.125.086101 (visited on 10/28/2020).

[38] Liqi Zhou et al. "Low-dose phase retrieval of biological specimens using cryo-electron ptychography". en. In: *Nature Communications* 11.1 (Dec. 2020), p. 2773. ISSN: 2041-1723. DOI: 10.1038/s41467-020-16391-6. URL: http://www.nature.com/articles/s41467-020-16391-6 (visited on 01/26/2021).

[39] Claire Donnelly et al. "Three-dimensional magnetization structures revealed with X-ray vector nanotomography". en. In: *Nature* 547.7663 (July 2017), pp. 328–331. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature23006. URL: http://www.nature.com/articles/nature23006 (visited on 12/26/2021).

[40] Mirko Holler et al. “High-resolution non-destructive three-dimensional imaging of integrated circuits”. en. In: *Nature* 543.7645 (Mar. 2017), pp. 402–406. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature21698. URL: http://www.nature.com/articles/nature21698 (visited on 12/26/2021).

[41] David A. Shapiro et al. “Chemical composition mapping with nanometre resolution by soft X-ray microscopy”. en. In: *Nature Photonics* 8.10 (Oct. 2014), pp. 765–769. ISSN: 1749-4885, 1749-4893. DOI: 10.1038/nphoton.2014.207. URL: http://www.nature.com/articles/nphoton.2014.207 (visited on 12/26/2021).

[42] Guoan Zheng, Roarke Horstmeyer, and Changhuei Yang. “Wide-field, high-resolution Fourier ptychographic microscopy”. en. In: *Nature Photonics* 7.9 (Sept. 2013), pp. 739–745. ISSN: 1749-4885, 1749-4893. DOI: 10.1038/nphoton.2013.187. URL: http://www.nature.com/articles/nphoton.2013.187 (visited on 12/26/2021).

[43] Dennis F. Gardner et al. “Subwavelength coherent imaging of periodic samples using a 13.5 nm tabletop high-harmonic light source”. en. In: *Nature Photonics* 11.4 (Apr. 2017), pp. 259–263. ISSN: 1749-4885, 1749-4893. DOI: 10.1038/nphoton.2017.33. URL: http://www.nature.com/articles/nphoton.2017.33 (visited on 07/25/2022).

[44] Junjing Deng et al. “Correlative 3D x-ray fluorescence and ptychographic tomography of frozen-hydrated green algae”. en. In: *Science Advances* 4.11 (Nov. 2018), eaau4548. ISSN: 2375-2548. DOI: 10.1126/sciadv.aau4548. URL: https://www.science.org/doi/10.1126/sciadv.aau4548 (visited on 07/25/2022).

[45] Yuan Hung Lo et al. “X-ray linear dichroic ptychography”. en. In: *Proceedings of the National Academy of Sciences* 118.3 (Jan. 2021), e2019068118. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.2019068118. URL: https://pnas.org/doi/full/10.1073/pnas.2019068118 (visited on 07/25/2022).

[46] Aaron Michelson et al. “Three-dimensional visualization of nanoparticle lattices and multimaterial frameworks”. en. In: *Science* 376.6589 (Apr. 2022), pp. 203–207. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.abk0463. URL: https://www.science.org/doi/10.1126/science.abk0463 (visited on 07/25/2022).

[47] Yi Jiang et al. “Electron ptychography of 2D materials to deep sub-ångström resolution”. en. In: *Nature* 559.7714 (July 2018), pp. 343–349. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-018-0298-5. URL: http://www.nature.com/articles/s41586-018-0298-5 (visited on 05/29/2019).

[48] Warren S. McCulloch and Walter Pitts. “A logical calculus of the ideas immanent in nervous activity”. en. In: *The Bulletin of Mathematical Biophysics* 5.4 (Dec. 1943), pp. 115–133. ISSN: 0007-4985, 1522-9602. DOI: 10.1007/BF02478259. URL: http://link.springer.com/10.1007/BF02478259 (visited on 07/25/2022).

[49] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. “Learning representations by back-propagating errors”. en. In: *Nature* 323.6088 (Oct. 1986), pp. 533–536. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/323533a0. URL: http://www.nature.com/articles/323533a0 (visited on 07/25/2022).

[50] Jürgen Schmidhuber. “Deep learning in neural networks: An overview”. en. In: *Neural Networks* 61 (Jan. 2015), pp. 85–117. ISSN: 08936080. DOI: 10.1016/j.neunet.2014.09.003. URL: https://linkinghub.elsevier.com/retrieve/pii/S0893608014002135 (visited on 12/27/2021).

[51] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. “U-Net: Convolutional Networks for Biomedical Image Segmentation”. en. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Ed. by Nassir Navab et al. Vol. 9351. Series Title: Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015, pp. 234–241. ISBN: 978-3-319-24573-7 978-3-319-24574-4. DOI: 10.1007/978-

3-319-24574-4_28. URL: http://link.springer.com/10.1007/978-3-319-24574-4_28 (visited on 12/28/2021).

[52] David Silver et al. "Mastering the game of Go without human knowledge". en. In: *Nature* 550.7676 (Oct. 2017), pp. 354–359. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature24270. URL: http://www.nature.com/articles/nature24270 (visited on 07/25/2022).

[53] David E. Rumelhart, ed. *Parallel distributed processing. 1: Foundations / David E. Rumelhart*. eng. 12. print. Cambridge, Mass: MIT Pr, 1999. ISBN: 978-0-262-68053-0 978-0-262-18120-4.

[54] P. J. Werbos. "Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences". PhD Thesis. Harvard University, 1974.

[55] Andreas Zell. *Simulation neuronaler Netze*. ger. 4., unveränd. Nachdr. München Wien: Oldenbourg, 2003. ISBN: 978-3-486-24350-5.

[56] Frank Rosenblatt. "PRINCIPLES OF NEURODYNAMICS. PERCEPTRONS AND THE THEORY OF BRAIN MECHANISMS". In: *American Journal of Psychology* 76 (1963), p. 705.

[57] Y. LeCun et al. "Backpropagation Applied to Handwritten Zip Code Recognition". en. In: *Neural Computation* 1.4 (Dec. 1989), pp. 541–551. ISSN: 0899-7667, 1530-888X. DOI: 10.1162/neco.1989.1.4.541. URL: https://direct.mit.edu/neco/article/1/4/541-551/5515 (visited on 02/27/2023).

[58] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. "ImageNet classification with deep convolutional neural networks". en. In: *Communications of the ACM* 60.6 (May 2017), pp. 84–90. ISSN: 0001-0782, 1557-7317. DOI: 10.1145/3065386. URL: https://dl.acm.org/doi/10.1145/3065386 (visited on 02/26/2023).

[59] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. "Deep learning". en. In: *Nature* 521.7553 (May 2015), pp. 436–444. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature14539. URL: http://www.nature.com/articles/nature14539 (visited on 02/26/2023).

[60] Kaiming He et al. "Deep Residual Learning for Image Recognition". In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV, USA: IEEE, June 2016, pp. 770–778. ISBN: 978-1-4673-8851-1. DOI: 10.1109/CVPR.2016.90. URL: http://ieeexplore.ieee.org/document/7780459/ (visited on 02/26/2023).

[61] Yao Yang et al. "Determining the three-dimensional atomic structure of an amorphous solid". en. In: *Nature* 592.7852 (Apr. 2021), pp. 60–64. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-021-03354-0. URL: http://www.nature.com/articles/s41586-021-03354-0 (visited on 05/05/2021).

[62] Paul J. Steinhardt, David R. Nelson, and Marco Ronchetti. "Bond-orientational order in liquids and glasses". en. In: *Physical Review B* 28.2 (July 1983), pp. 784–805. ISSN: 0163-1829. DOI: 10.1103/PhysRevB.28.784. URL: https://link.aps.org/doi/10.1103/PhysRevB.28.784 (visited on 02/26/2023).

[63] Jinwoo Hwang et al. "Nanoscale Structure and Structural Relaxation in Zr 50 Cu 45 Al 5 Bulk Metallic Glass". en. In: *Physical Review Letters* 108.19 (May 2012), p. 195505. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.108.195505. URL: https://link.aps.org/doi/10.1103/PhysRevLett.108.195505 (visited on 02/26/2023).

[64] Akihiko Hirata et al. "Direct observation of local atomic order in a metallic glass". en. In: *Nature Materials* 10.1 (Jan. 2011), pp. 28–33. ISSN: 1476-1122, 1476-4660. DOI:

10.1038/nmat2897. URL: https://www.nature.com/articles/nmat2897 (visited on 02/26/2023).

[65] Daniel B. Miracle. "A structural model for metallic glasses". en. In: *Nature Materials* 3.10 (Oct. 2004), pp. 697–702. ISSN: 1476-1122, 1476-4660. DOI: 10.1038/nmat1219. URL: https://www.nature.com/articles/nmat1219 (visited on 02/26/2023).

[66] Daniel B. Miracle. "A Physical Model for Metallic Glass Structures: An Introduction and Update". en. In: *JOM* 64.7 (July 2012), pp. 846–855. ISSN: 1047-4838, 1543-1851. DOI: 10.1007/s11837-012-0359-4. URL: http://link.springer.com/10.1007/s11837-012-0359-4 (visited on 02/26/2023).

[67] "Supercooling of liquids". en. In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 215.1120 (Nov. 1952), pp. 43–46. ISSN: 0080-4630, 2053-9169. DOI: 10.1098/rspa.1952.0194. URL: https://royalsocietypublishing.org/doi/10.1098/rspa.1952.0194 (visited on 02/26/2023).

[68] David R. Nelson and Frans Spaepen. "Polytetrahedral Order in Condensed Matter". en. In: *Solid State Physics*. Vol. 42. Elsevier, 1989, pp. 1–90. ISBN: 978-0-12-607742-1. DOI: 10.1016/S0081-1947(08)60079-X. URL: https://linkinghub.elsevier.com/retrieve/pii/S008119470860079X (visited on 02/26/2023).

[69] Yakun Yuan et al. "Three-dimensional atomic packing in amorphous solids with liquid-like structure". en. In: *Nature Materials* 21.1 (Jan. 2022), pp. 95–102. ISSN: 1476-1122, 1476-4660. DOI: 10.1038/s41563-021-01114-z. URL: https://www.nature.com/articles/s41563-021-01114-z (visited on 01/08/2022).

[70] A. M. Maiden, M. J. Humphry, and J. M. Rodenburg. "Ptychographic transmission microscopy in three dimensions using a multi-slice approach". en. In: *Journal of the Optical Society of America A* 29.8 (Aug. 2012), p. 1606. ISSN: 1084-7529, 1520-8532.

DOI: 10.1364/JOSAA.29.001606. URL: https://www.osapublishing.org/abstract.cfm?URI=josaa-29-8-1606 (visited on 05/29/2019).

[71] Akihiro Suzuki et al. "High-Resolution Multislice X-Ray Ptychography of Extended Thick Objects". en. In: *Physical Review Letters* 112.5 (Feb. 2014), p. 053903. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.112.053903. URL: https://link.aps.org/doi/10.1103/PhysRevLett.112.053903 (visited on 03/10/2020).

[72] Earl J. Kirkland, Russell F. Loane, and John Silcox. "Simulation of annular dark field stem images using a modified multislice method". en. In: *Ultramicroscopy* 23.1 (Jan. 1987), pp. 77–96. ISSN: 03043991. DOI: 10.1016/0304-3991(87)90229-4. URL: https://linkinghub.elsevier.com/retrieve/pii/0304399187902294 (visited on 02/21/2020).

[73] J. M. Cowley and A. F. Moodie. "The scattering of electrons by atoms and crystals. I. A new theoretical approach". In: *Acta Crystallographica* 10.10 (Oct. 1957), pp. 609–619. ISSN: 0365-110X. DOI: 10.1107/S0365110X57002194. URL: http://scripts.iucr.org/cgi-bin/paper?S0365110X57002194 (visited on 02/21/2020).

[74] Alan Pryor, Colin Ophus, and Jianwei Miao. "A streaming multi-GPU implementation of image simulation algorithms for scanning transmission electron microscopy". en. In: *Advanced Structural and Chemical Imaging* 3.1 (Dec. 2017), p. 15. ISSN: 2198-0926. DOI: 10.1186/s40679-017-0048-z. URL: https://ascimaging.springeropen.com/articles/10.1186/s40679-017-0048-z (visited on 02/21/2020).

[75] Claire B. Eisenhandler and Benjamin M. Siegel. "Imaging of Single Atoms with the Electron Microscope by Phase Contrast". en. In: *Journal of Applied Physics* 37.4 (Mar. 1966), pp. 1613–1620. ISSN: 0021-8979, 1089-7550. DOI: 10.1063/1.1708575.

URL: http://aip.scitation.org/doi/10.1063/1.1708575 (visited on 02/28/2020).

[76] David Ren et al. "A multiple scattering algorithm for three dimensional phase contrast atomic electron tomography". en. In: *Ultramicroscopy* 208 (Jan. 2020), p. 112860. ISSN: 03043991. DOI: 10.1016/j.ultramic.2019.112860. URL: https://linkinghub.elsevier.com/retrieve/pii/S030439911930052X (visited on 03/27/2020).

[77] Peter W. Hawkes. "The Electron Microscope as a Structure Projector". en. In: *Electron Tomography*. Ed. by Joachim Frank. Boston, MA: Springer US, 1992, pp. 17–38. ISBN: 978-1-4757-2165-2 978-1-4757-2163-8. DOI: 10.1007/978-1-4757-2163-8_2. URL: http://link.springer.com/10.1007/978-1-4757-2163-8_2 (visited on 02/26/2020).

[78] Chenhao Jin et al. "Observation of moiré excitons in WSe2/WS2 heterostructure superlattices". en. In: *Nature* 567.7746 (Mar. 2019), pp. 76–80. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-019-0976-y. URL: http://www.nature.com/articles/s41586-019-0976-y (visited on 04/10/2020).

[79] Yuan Cao et al. "Unconventional superconductivity in magic-angle graphene superlattices". en. In: *Nature* 556.7699 (Apr. 2018), pp. 43–50. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature26160. URL: http://www.nature.com/articles/nature26160 (visited on 04/13/2020).

[80] Guorui Chen et al. "Evidence of a gate-tunable Mott insulator in a trilayer graphene moiré superlattice". en. In: *Nature Physics* 15.3 (Mar. 2019), pp. 237–241. ISSN: 1745-2473, 1745-2481. DOI: 10.1038/s41567-018-0387-2. URL: http://www.nature.com/articles/s41567-018-0387-2 (visited on 04/13/2020).

[81] Chendong Zhang et al. “Interlayer couplings, Moiré patterns, and 2D electronic superlattices in MoS $_2$ /WSe $_2$ hetero-bilayers”. en. In: *Science Advances* 3.1 (Jan. 2017), e1601459. ISSN: 2375-2548. DOI: 10.1126/sciadv.1601459. URL: https://advances.sciencemag.org/lookup/doi/10.1126/sciadv.1601459 (visited on 04/23/2020).

[82] Kha Tran et al. “Evidence for moiré excitons in van der Waals heterostructures”. en. In: *Nature* 567.7746 (Mar. 2019), pp. 71–75. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-019-0975-z. URL: http://www.nature.com/articles/s41586-019-0975-z (visited on 04/23/2020).

[83] Ayan Sinha et al. “Lensless computational imaging through deep learning”. en. In: *Optica* 4.9 (Sept. 2017), p. 1117. ISSN: 2334-2536. DOI: 10.1364/OPTICA.4.001117. URL: https://www.osapublishing.org/abstract.cfm?URI=optica-4-9-1117 (visited on 10/28/2020).

[84] Mathew J. Cherukara et al. “AI-enabled high-resolution scanning coherent diffraction imaging”. en. In: *Applied Physics Letters* 117.4 (July 2020), p. 044103. ISSN: 0003-6951, 1077-3118. DOI: 10.1063/5.0013065. URL: http://aip.scitation.org/doi/10.1063/5.0013065 (visited on 02/25/2021).

[85] Mo Deng et al. “On the interplay between physical and content priors in deep learning for computational imaging”. en. In: *Optics Express* 28.16 (Aug. 2020), p. 24152. ISSN: 1094-4087. DOI: 10.1364/OE.395204. URL: https://www.osapublishing.org/abstract.cfm?URI=oe-28-16-24152 (visited on 10/28/2020).

[86] Peter B. Rosenthal and Richard Henderson. “Optimal Determination of Particle Orientation, Absolute Hand, and Contrast Loss in Single-particle Electron Cryomicroscopy”. en. In: *Journal of Molecular Biology* 333.4 (Oct. 2003), pp. 721–745. ISSN: 00222836.

DOI: 10.1016/j.jmb.2003.07.013. URL: https://linkinghub.elsevier.com/retrieve/pii/S0022283603010222 (visited on 10/28/2022).

[87] Colum M. O'Leary et al. "Contrast transfer and noise considerations in focused-probe electron ptychography". en. In: *Ultramicroscopy* 221 (Feb. 2021), p. 113189. ISSN: 03043991. DOI: 10.1016/j.ultramic.2020.113189. URL: https://linkinghub.elsevier.com/retrieve/pii/S0304399120303314 (visited on 10/28/2022).

[88] Mathew J. Cherukara, Youssef S. G. Nashed, and Ross J. Harder. "Real-time coherent diffraction inversion using deep generative networks". en. In: *Scientific Reports* 8.1 (Dec. 2018), p. 16520. ISSN: 2045-2322. DOI: 10.1038/s41598-018-34525-1. URL: http://www.nature.com/articles/s41598-018-34525-1 (visited on 10/28/2020).

[89] Zhong Zhuang et al. "Practical Phase Retrieval Using Double Deep Image Priors". In: (2022). Publisher: arXiv Version Number: 1. DOI: 10.48550/ARXIV.2211.00799. URL: https://arxiv.org/abs/2211.00799 (visited on 04/26/2023).

[90] Sergey Ioffe and Christian Szegedy. "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift". In: (2015). Publisher: arXiv Version Number: 3. DOI: 10.48550/ARXIV.1502.03167. URL: https://arxiv.org/abs/1502.03167 (visited on 04/25/2023).

[91] Nitish Srivastava et al. "Dropout: A Simple Way to Prevent Neural Networks from Overfitting". In: *Journal of Machine Learning Research* 15.56 (2014), pp. 1929–1958. URL: http://jmlr.org/papers/v15/srivastava14a.html.

[92] Aravindh Mahendran and Andrea Vedaldi. "Visualizing Deep Convolutional Neural Networks Using Natural Pre-images". en. In: *International Journal of Computer Vision* 120.3 (Dec. 2016), pp. 233–255. ISSN: 0920-5691, 1573-1405. DOI: 10.1007/s11263-

016-0911-8. URL: http://link.springer.com/10.1007/s11263-016-0911-8 (visited on 04/25/2023).

[93] Matthew D. Zeiler and Rob Fergus. "Visualizing and Understanding Convolutional Networks". en. In: *Computer Vision – ECCV 2014*. Ed. by David Fleet et al. Vol. 8689. Series Title: Lecture Notes in Computer Science. Cham: Springer International Publishing, 2014, pp. 818–833. ISBN: 978-3-319-10589-5 978-3-319-10590-1. DOI: 10.1007/978-3-319-10590-1_53. URL: http://link.springer.com/10.1007/978-3-319-10590-1_53 (visited on 04/25/2023).

[94] Dumitru Erhan et al. "Visualizing Higher-Layer Features of a Deep Network". In: *Technical Report, Univeristé de Montréal* (Jan. 2009).