Title
Electron Microscopy for the Study of Defect Development in Nanomaterials

Author
Groschner, Catherine Kingston

Publication Date
2021

Peer reviewed|Thesis/dissertation

Electron Microscopy for the Study of Defect Development in Nanomaterials

by

Catherine Kingston Groschner

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering: Materials Science and Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Mary Scott, Chair
Professor Andrew Minor
Professor Joshua S. Bloom

Fall 2021

Electron Microscopy for the Study of Defect Development in Nanomaterials

Abstract

Electron Microscopy for the Study of Defect Development in Nanomaterials

by

Catherine Kingston Groschner

Doctor of Philosophy in Engineering: Materials Science and Engineering

University of California, Berkeley

Professor Mary Scott, Chair

Precisely controlled synthesis of nanostructures is heavily emphasized in the field of nanoscience, in large part due to the desire to control the size, shape, and terminating facets of nanoparticles for applications in catalysis, optics, and medicine. Direct control of the size and shape of solution-grown nanoparticles relies on an understanding of how synthetic parameters alter nanoparticle structures during synthesis. Synthesis conditions rarely yield uniform particles, but the heterogeneity in these populations is hard to quantify especially in respect to atomic structure. Defects are a particularly challenging element to quantify the influence of synthesis on due to the limited number of methods which can give insight into these features. In this thesis we analyze the development of defect structures in multiply twinned metal nanoparticles. We find that non-classical growth mechanisms occur during standard colloidal growth. In order to further study the conditions which cause these non-classical growth modes we analyze methods to automate the analysis of defects in nanoparticles. To this end we propose a machine learning based pipeline for the segmentation and stacking fault classification in nanoparticles. We demonstrate a flexible two step pipeline for analysis of high resolution transmission electron microscopy data, which uses a U-Net for segmentation followed by a random forest for detection of stacking faults. Our trained U-Net is able to segment nanoparticle regions from amorphous background with a Dice coefficient of 0.8 and significantly outperforms traditional image segmentation methods. Using these segmented regions, we are then able to classify whether nanoparticles contain a visible stacking fault with 86% accuracy. We further explore the role of architecture in achieving better segmentation of HRTEM micrographs. These studies in turn illuminate the challenges in generating a large enough dataset for neural network training in electron microscopy. To solve this we evaluate methods to minimize labelling requirements by using automated labeling methods. This strategy recasts the classification problem as a label noise problem, with the fraction of label noise problem. We analyze the impact of label noise rate and training method on the classification error of neural networks for SEM data. We compare the noise resistance of three training methods: standard network training, pretraining with a tiny, label noise-free

dataset, and co-teaching. We find that the pretraining approach yields the most accurate results across label error rates. These developments will help in enabling automated analysis of nanomaterials.

For my family

Contents

List of Figures

List of Tables

Acknowledgments

This dissertation would not have been possible without the support and guidance of my advisor Professor Mary Scott. I cannot imagine any other professor would sit for hours with their graduate student just to teach them the ins and outs of a microscope but Mary did just that. She has helped me become a better researcher and a stronger person, and I am so grateful for her mentorship.

I was privileged to study and work at NCEM and be a part of its incredible community of microscopists. My sincere thanks to Professor Jihan Zhou for teaching me so much about tomography and being my office buddy. I miss our afternoon chats. Thank you to Dr. Karen Bustillo for hours of training, help setting up software, explaining EDS, and so many words of encouragement. Thank you to Chengyu Song for all the tips, tricks, and secrets of the TEAM 0.5 and for countless sample changes. I am so grateful to Dr. Peter Ercius for sharing his seemingly limitless knowledge and encouraging my wacky ML projects. Thank you to Dr. Wolfgang Theis for so many informative discussions on machine learning. A big thanks to Dr. Colin Ophus for much of my knowledge of electron microscopy theory. I think I would have failed my qualifying exam if you hadn't taught me so much about phase contrast.

Thank you to the members of Scott lab and Minor lab. I am so lucky to have had such supportive and knowledgeable colleagues. A big shout out to the early members of Scott lab: Ellis Kennedy, Alex Bruefach, Bengisu Yasar, and Hannah DeVyldere . I am so glad I got to share this journey with you. Thanks to Steven Zeltmann for keeping me caffeinated.

Thanks to all my friends that have kept me going through the ups and downs of grad school. Special acknowledgement to Hanna, Madeline, the Churchill gang, all the housemates at 1937 Francisco St. and the SF/Oakland crew especially our pod: Jessie, Taryn, Danny, and Megan. You guys are the best.

Thank you Erin for helping me to accept good enough. I would not have made it through without you.

Many thanks to my family: Grandma, Grandpa, Uncle Brett, Aunt Mariann, Tori, Alex, Aunt Sue, Dina, Dan, Terri, Mark, and Amy for always cheering me on. You have always believed in me and it has made all the difference.

Thank you to my parents. You have given me every form of support imaginable. You made all this possible. I will never be able to express all my love and gratitude.

And last but not least thank you Spencer. You have brought more joy to my life in the last five years than I ever thought possible. Thank you for being my person. I finished my PhD, will you please marry me now?

Chapter 1

Introduction

Most materials are defined by being made up of a regular, periodic arrangement of constituent atoms. These arrangements of atoms are well known to materials scientists as the crystal structure of the material. Understanding the relationship between this crystal structure and the properties of a material is the core of materials science. However, almost as long as it has been understood that materials have a crystal structure it has been understood that the atomic structure of a material often will have instances where atoms deviate from the overall periodic structure. These instances are known as defects. Defects occur on a variety of length scales from zero-dimensional to three-dimensional. Zero-dimensional defects consist of the addition or subtraction of an atom in the crystal lattice, known as interstitials or vacancies respectively. One-dimensional defects occur when an entire plane of atoms deviate from the crystal structure and are known as dislocations. Two-dimensional defects occur when two orientations of a crystal structure come together such as grain boundaries.

Control of defects has been the crux of many sub-disciplines in materials science because of the influence defects have on material properties. In metallurgy, controlling defects has long been used to increase strength and toughness through cold working and annealing to introduce and control the number of dislocations and grain boundaries [1]. In semiconductors the ability to create defect free ingots enabled the rise of computing [1, 2]. With the rise of nanomaterials the study of defects remains critical. As synthetic control of nanomaterials has inched towards atom by atom control, the interest in controlling defects down to the single vacancy has grown proportionally. The interest in defects in nanomaterials has its root in the critical role defects play in the properties of nanomaterials.

1.1 Role of Defects in Nanomaterials

In order to understand the importance of defects in nanomaterials we must first explore what defines a nanomaterial. A nanomaterial is defined as any material in which one dimension is on the nanometer scale, and is usually less than 100nm [3]. The interest in nanomaterials due to the novel properties which can be achieved by scaling materials down to the nanome-

ter scale. In terms of electronic and optical properties, by limiting the dimensions of the material we create quantum confinement effects which will change the electronic states of the material [4]. Nanomaterials also change the chemical properties of the material by significantly increasing the surface to volume ratio. Surfaces present regions of dangling bonds and these bonds make the material more chemically active [4]. We observe these effects just by reducing the size of the material but inclusion of defects in the crystal lattice will further change these properties.

Just like in bulk materials, defects can influence nanoparticle properties through modifications of the electronic structure. Planar defects and dislocations create broken bonds which create new electronic states which would otherwise be inaccessible. These local electronic states can change the electronic and optical properties of the nanomaterial. An example of this behavior is in the photoemission of CdSe quantum dots [5]. Stacking faults in CdSe quantum dots have been correlated to non-radiative recombination of charge carriers therefore leading to reduced photoluminescence quantum yield due to their creation of trap states [5].

Defects also cause strain in the lattice which subsequently will also influence the band structure and hence both electronic and chemical properties. In terms of electronic property shifts, strain in graphene has been correlated with shifting of dirac point and even destruction of the dirac point [6]. Metal nanoparticles, which are of interest for catalysis, are a prime example of chemical property dependence based on defect based strain. Catalytically active metals tend to be rare and therefore their efficient use has led to interest in providing the most surface area per volume of material which naturally leads to creating nanoparticles of these metals. However the catalytic activity of these nanoparticles is strongly driven by the shape and strain states of the nanoparticle. Studies has shown, for example, that when palladium nanoparticles take on an icosahedral shape, which is highly strained, as opposed to octahedra, which is relatively unstrained, that their ability to catalyze the CO_2 reaction dramatically improves [7]. This increase in catalytic activity is attributed to the defects required to create these shapes because creating an icosahedron requires a central disclination and twenty twin boundaries, all of which leads to a very strained particle. The strained atoms of the particle change the strength the adsorption of key intermediaries [7]. Similarly, the strain in semiconductor nanoparticles can lead to shifts in the photoemissive properties of the nanomaterial. For example, stacking faults have been known to redshift the photoemission of III-V and II-VI semiconductor nanowires [8, 9, 10, 11].

Defects not only directly influence the properties of nanomaterials but also influence properties by directing the growth, and therefore the final shape of solution-synthesized nanomaterials. A famous example of this is the helical or branching nature that nanowires can take on when a screw dislocation lies in the center of the nanowire leads to the creation of an Eshelby twist which is often evidenced by the helical growth of nanowire branches off the screw dislocation containing center nanowire [12]. The Eshelby twist means that as the nanowire continues to grow the material follows the steps in the crystal surface provided by screw dislocations. These steps, like the steps of a spiral staircase, lead to growth of the crystal following the helical pattern of the steps. More recently screw dislocations have

also been suggested to be the cause of morphologically chiral nanoparticles [13]. Since the shape of the particle is critical to how it will interact with its environment whether that is in a device or in a biological application the control over defects during the growth process is critical. However, the presence of defects is not always consistent. Defects can lead to heterogeneity in a nanoparticle sample [14]. This further complicates their study and increase the importance of knowing how to control their presence.

1.2 Synthetic Control of Defects in Nanomaterials

The typical understanding of the development of crystal structure, and consequently defects, in nanomaterials synthesized via wet chemical synthesis is based on classical nucleation and growth theory. From the classical perspective, there are two main variables which control the nucleation and growth process: thermodynamics and kinetics. Depending on the type of defect, it may or may not be thermodynamically favorable. Vacancies and interstitials, for instance, contribute to the entropy of a material system and therefore occur in an equilibrium amount based on the energetics of the system [15, 14]. Dislocations and other 1D and 2D defects increase the energy of the material and are therefore kinetically driven [15].

However, the reality is that many non-classical nucleation and growth processes occur during formation of nanomaterials and their defects, and the growth of a nanoparticle may involve multiple processes. Recent work has begun expounding the non-classical processes occurring between the first stages of nucleation, seed stabilization, and nanoparticle growth. The early heterogeneous nucleation process has now been directly observed by Jeon and coworkers [16]. Their work has supported the idea that nucleation is an extremely dynamic process in which nuclei fluctuate between a disordered and crystalline state until a large enough particle size is achieved which stabilizes the the crystalline structure [16]. After initial nucleation of a stable seed, classical thermodynamics suggests that it will grow by monomer addition into the lowest energy structure [17]. Studies in the past twenty years have shown that is not necessarily the case. Crystal seeds are likely to participate in some form of particle attachment. In fact particle attachment with amorphous particles may also contribute to crystal seed formation creating yet another pathway to nucleation of a nanocrystal [18]. Oriented attachment is one particular case of particle attachment and is important in the development of line defects. During oriented attachment, crystalline seeds may coalesce with the lattice planes of the particles being aligned or misaligned. The misalignment of lattice planes during oriented attachment leads to the formation of dislocations, stacking faults, or twin boundaries [18]. This is one way in which defects may be generated in nanomaterials. Particles growing from seeds containing defects may then develop more defects to relieve the strain caused by templating from the defected structure [19]. In addition, adatom deposition rate, temperature, and other synthetic parameters have also been shown to contribute though the actual mechanisms in these cases are less understood [20, 21, 17]. While great strides have been made, there is still much to be understood about what contributes to defect formation in nanoparticles, particularly in early seed formation.

1.3 Electron Microscopy

All electron microscopes rely on forming an electron beam which is focused with a series of magnetic lens and illuminating the sample in question to create an image. The resolution and available information depends on the specific electron microscopy technique. Each technique changes how the beam is focused and the signal collected from the interaction of the beam with the sample. The three principal forms of electron microscopy for imaging are TEM, STEM, and SEM. Schematics of each technique are shown in Figure 1.1 Scanning electron microscopy (SEM) is the lowest resolution. Its general operating principle is that beam is focused into a probe which is rastered across the sample and then electrons scattered from the surface are collected at each probe position to create an image. The scattered electrons are divided into one of two classes: secondary electrons or backscattered electrons. The difference between the two is how the beam and sample interaction leads to each. Secondary electrons are ejected from their electronic band in the sample due to inelastic scattering interaction with the beam electrons [22]. The contrast from secondary electrons are then due to the orientation of the sample surface to the detector. Backscattered electrons are beam electrons which are essentially reflected out of the sample by being elastically scattered by the sample atoms [22]. Higher atomic number atoms will cause the beam electrons to backscatter more strongly and therefore the contrast in the image is due to variations in the atomic number of the sample elements [22].

Whereas SEM is based on the collection of electrons scattered from the sample surface, both transmission electron microscopy (TEM) and scanning transmission electron microscopy (STEM) are based on having the electron beam transmit through the sample to create a projectional image. SEMs also have much lower spatial resolution, on the order of 0.5nm, while (S)TEM can have a spatial resolution of less than 50pm [23], making TEM or STEM preferable for studies of nanomaterials. The key difference between TEM and STEM imaging is that in TEM the sample is illuminated all at once with a parallel beam of electrons. STEM converges the beam of electrons into a probe which is rastered across the sample. Both methods enable different operational modes which can be used for the study of defects. Bright field (BF) and dark field (DF) TEM imaging are the older TEM based means of imaging defects, and have mostly been used for metal foil samples. BF imaging uses the direct beam to form an image while DF uses a diffracted beam, thereby giving intensity to those regions which gave rise to the diffracted beam. BF or DF TEM of defects relies on the change in the coherent scattering caused by the line or planar defect in order to cause a variation in the diffraction contrast of the image. If there are overlapping defects, interpretation of these images is difficult or even impossible if the combination eliminates the contrast generated. It should also be noted that these techniques do not directly image the defect. The image contrast related to the defect is dependent on the thickness of the sample, the orientation of the sample under the beam, and the position of the defect within the sample. These factors contribute to the imaged "position" of the dislocation being frequently displaced from the true location of the defect in these cases, which may be problematic in the study of nanomaterials such as nanoparticles. Due to these factors high resolution imaging

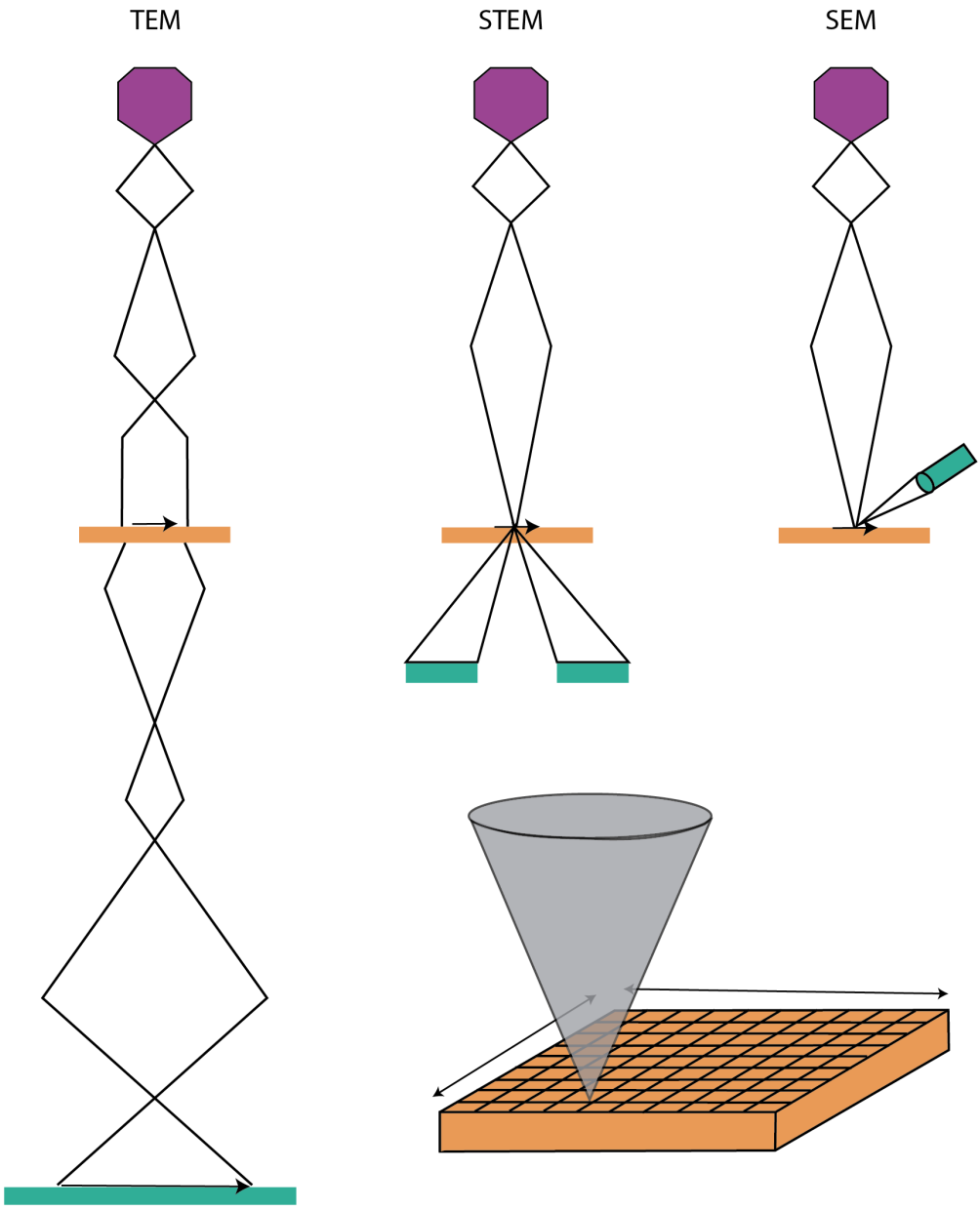


Figure 1.1: Schematics of the imaging systems of a TEM, STEM, and SEM. The purple region represents the electron gun, the oragne region the sample, and the teal region the detector. The subset in the bottom right shows the pixel by pixel rastering of the electron probe which is used both in STEM and SEM.

is usually used for defect analysis in nanomaterials.

There are two common high resolution techniques in electron microscopy, HRTEM and HRSTEM, both of which create an image of the atomic lattice and therefore allow us to

more directly image defects. HRSTEM is an incoherent imaging technique which relies on the elastic scattering of electrons to high angles as the beam is rastered across the sample to create an image. By collecting the electrons scattered to high angles using a high angle annular electron detector (HAADF) we eliminate signals coming from coherently scattered electrons and ensure that all the electrons are incoherently scattered [24]. The incoherent scattering means that the image contrast can be treated as the sum of individual scattering events from each atom in the sample [24]. The scattering is also then dependent on the atomic number of the atoms illuminated and therefore the contrast in the image reflects any compositional variation. As long as the electron probe used to illuminate the sample is smaller than the atomic column spacing then the atomic columns are individually illuminated as the beam is rastered across the sample and an atomic resolution image is created [24]. Since the technique collects those electrons which are scattered to very high angles, the method does require a high intensity electron probe so beam damage to the sample can be a problem. The atomic number dependence also means that the technique does not work well for light elements.

HRTEM is a coherent imaging technique in which the sample is illuminated with a parallel beam of electrons. It allows for imaging of crystal lattice through formation of lattice fringes in the image. These lattice fringes are formed because the scattering of the atomic columns causes a phase shift and subsequent interference in the electron wave. This phase distortion creates a periodicity in the electron wave intensity which are the lattice fringes that are observed [24]. These lattice fringes correspond to the spacing of the atomic planes. If the sample is oriented such that multiple reciprocal vectors lie on the Ewald sphere then multiple lattice fringes in multiple orientations will be observed and this will allow for a representation of the atomic columns [25]. It should be noted that this is not a direct image of the atomic columns but a representation of the atomic columns based on the interference of the electron wave due to the phase shifts induced by the sample and are therefore very sensitive to parameters such as defocus and spherical aberration. If the microscope parameters are controlled correctly and the sample is thin enough then the lattice fringes formed in the image can be used to interpret the atomic structure of the material. Planar defects and grain boundaries will of course change the phase modulation and thus can be imaged with atomic resolution using HRTEM. HRTEM presents the benefit that by illuminating the sample with a parallel beam of electrons the entire sample is imaged at once. This provides a significant speed increase over scanning probe methods and is an important consideration for high throughput imaging.

Development of high throughput imaging is key in order to characterize the products of automated synthesis. With the creation of automated synthesis, larger synthetic parameter spaces can be probed than ever before [26]. But without a way to characterize the output these methods are meaningless. As discussed before, x-ray diffraction methods can provide a way to understand the average structure of these materials, but in order to understand whether defects are being generated during synthesis or to understand the impact of defects in these materials a new high throughput form of electron microscopy is needed to gain statistical information on the prevalence of defects and their relation to synthesis parameters.

On the microscope side, significant efforts have been made to automate the acquisition process. This is a challenging task because software which enables automated collection must be able to track location (in order to ensure the same area is not being repeatedly sampled) and be able to automatically focus to adjust for varying conditions across the SEM or TEM sample. Automated acquisition is now available with commercially available SEMs, enabling huge fields of view by tracking imaging position and stitching individual images together [27]. Development of automated TEM acquisition by the cryo-EM community for biological applications and has lead to open source tools such as SerialEM [28, 29, 30, 31]. As the need for automated TEM has grown though commercial ventures are offering more tools to enable automated acquisition on a range of samples. Most of these methods are not focused on HRTEM and perform better at lower magnifications. Consequently, work is ongoing to create automated collection methods for nanomaterials. Recent demonstrations of automated collection of micrographs of $CaCo_3$ nanoparticles have been particularly promising [32].

Automating the acquisition process is only one of the challenges to automating nanomaterial characterization with electron microscopy. Once data is collected it of course must be analyzed, and this can be one of the most time consuming steps a microscopist faces. Before the rise of machine learning this analysis would be done with conventional image processing steps. Conventional image processing required the creation of a bespoke processing pipeline for each dataset, which is clearly not useful for creating an automated procedure. But with the increased availability of machine learning and deep learning methods, the possibility of creating automated analysis methods which would automate the characterization process grew. Due to the advances made in using deep learning for computer vision tasks it has become the focus of most analysis automation work. Much of this work has been focused on atomic column segmentation from HRSTEM images [33, 34]. However, as previously discussed, HRSTEM is not an ideal means to analyze large volumes of nanomaterials due to the slow acquisition. Therefore means of automating the analysis of HRTEM micrographs needed to be developed. The last few years have seen a growth in this area. The first step to analyzing HRTEM images is to segment out the nanomaterials of interest. We have demonstrated neural networks with good performance on this task using semantic segmentation [35, 36]. Once nanomaterials are identified in the micrograph, then analysis can be done to correlate shape, size, and crystallography as well as defects. This is still an active area of research with many different machine learning methods being explored [36, 37, 38, 39].

1.4 Dissertation Overview

This dissertation will analyze the interplay between synthesis and defects in metal and semiconductor nanoparticles and propose new techniques for the study of said relationship. In Chapter 2, the influence of kinetics is analyzed in relationship to the development of a defected structure in a multiply twinned palladium particle. The defect structure in this case is studied with electron tomography. These results prompt the need for more high through-

put nanoparticle and defect analysis in electron microscopy. Consequently, in Chapter 3 a demonstration of a machine learning pipeline for automating nanoparticle identification and stacking fault quantification is demonstrated. Further improvements in nanoparticle segmentation will be explored in Chapter 4. We demonstrate that improved segmentation can be gained by moving away from the traditional U-Net architecture employed for most segmentation. In Chapter 5 we discuss the limitations of machine learning with respect to the data labeling requirements and propose an automated labeling procedure for geometrically related datasets.

Chapter 2

Development of Defects in Multiply Twinned Palladium[1]

2.1 Introduction

A nanomaterial system in which defects play a critical role are multiply twinned particles (MTPs). MTPs are ubiquitous in solution-grown nanoparticle populations of face-centered cubic (fcc) metals. The ability to improve catalytic activity by controlling the exposed surface facet and strain states of the MTPs is a major driving force to understand their evolution during synthesis [40, 41, 41, 21, 42, 43, 44, 45]. A large number of tunable synthetic parameters are known to influence nanoparticle shape, such as reaction temperature, reducing agent concentration, and capping agents. These can be employed to direct the structure of nanomaterials [20, 46]. The primary multiply-twinned structures that appear in fcc metal nanoparticle populations are decahedra and icosahedra, which have an idealized structure described by assemblies of 5 or 20 tetrahedral subunits respectively, with subunits joined by twin boundaries of close-packed $\langle 111 \rangle$-terminated surface facets. However, other multiply twinned structures are commonly observed, either as transient structures during synthesis [47, 48] or final reaction products [49, 50].

Studies of MTP synthesis have identified many possible growth pathways to generate MTPs [21, 48, 47, 51]. Many efforts to create size and shape controlled MTPs center around controlling the population of nanometer-sized crystal seeds, which can uniformly grow into larger particles with the same shape [20]. However, other structural evolution pathways are known to occur during colloidal growth of MTPs, such as successive twinning and oriented attachment [48, 52]. The successive twinning growth process is unique to MTPs and refers to the additive growth of new tetrahedra to multiply twinned structures. This process can evolve single tetrahedra into a more complex MTP by island-to-tetrahedron growth on one of the facets of the single tetrahedra.

[1]This chapter is a version of P. Pelz and C. Groschner, A. Bruefah, A. Satariano, C. Ophus, and M.C. Scott. "Simultaneous Successive Twinning Captured by Atomic Electron Tomography". Submitted.

The details of the successive twinning process are not clear. Several examples of successive twinning have been observed using electron microscopy in both static and liquid cell experiments [47, 51, 48]. The decahedron to icosahedron successive twinning growth pathway is particularly difficult to quantitatively characterize using two dimensional imaging methods given the complex overlapping crystal grain structure of these MTPs. Some studies have indicated that icosahedra can evolve from fully formed decahedra [47, 51], while others claim icosahedra generated through successive twinning require a partially formed decahedra [53]. Furthermore, the role of local defects and surface structures play during successive twinning have not been identified.

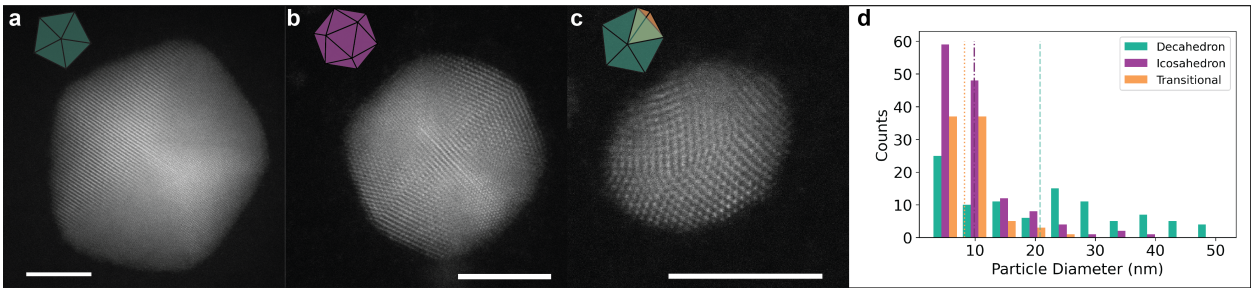


Figure 2.1: Representative images of the structures observed and histogram from the aliquot study. a) Sample decahedron. b) Sample icosahedron. c) Sample particle transitioning between decahedron and icosahedron. d) Histogram of the populations observed for the different structures as a function of size. Scale bar represents 5 nm

.

Electron microscopy characterization has been critical in our understanding of MTP particle stability and growth. However, most previous experimental electron microscopy studies are limited to two dimensional imaging, such that many of the multi-tetrahedron stuctures present during successive twinning appear ambiguous. Atomic electron tomography (AET) is a method that utilizes high-resolution scanning transmission electron microscopy (HR-STEM) datasets to reconstruct the atomic-scale 3D structure of materials. Previous AET studies have resolved the structure, local defects and strain in icosahedral and decahedral metal nanoparticles [54, 55], making it an ideal tool to resolve the complicated crystal structure of nanoparticles undergoing successive twinning.

In this work, we applied conventional HR-STEM to a population of Pd nanoparticles to determine the frequency of appearance of decahedra, icosahedra, and multiply twinned particles undergoing successive twinning, which we will refer to as multi-tetrahedron particles. AET was then used to measure the atomic-scale structure of a representative particle undergoing successive twinning. AET revealed significant structural disorder within the particle. The more than 20,000 atomic coordinates in 3D provided by AET analysis were further classified according to their local fcc or hexagonally close-packed (hcp) environment. This classification enabled atomic-scale 3D visualization of a simultaneous successive twin-

ning process, where the particle was midway through a transformation from decahedron to icosahedron. We also observe stacking faults and other defects within the crystal grains of the particle, and a region of crystalline disorder on the surface of the particle.

2.2 Methods

Sample preparation

All reagents used in this synthesis were purchased from Sigma Aldrich. The Pd nanoparticle reaction was prepared based on the method reported by Lim *et al.* with slight modifications [20]. We used the conditions reported to synthesize a population of primarily decahedral particles. Briefly, a 15 mL three-necked flask was rinsed 3 times with MilliQ water and acetone, then dried. The flask was equipped with a reflux condenser and teflon-coated magnetic stir bar. A 4 mL aqueous solution containing poly(vinyl pyrrolidone) (PVP, 55,000 MW) and citric acid was transferred to the flask and heated to 90°C while stirring using a heating mantle. Upon reaching 90°C, A 1.5 mL solution of Sodium tetrachloropalladate was rapidly added to the flask. For the aliqout study, samples were taken at 1, 3, 6, 8, 10, and 24 hours. The final product was isolated by adding a 300 µL aliqout of acetone to 100 µL nanoparticle solution in a clean microcentrifuge tube and centrifuged for 30 minutes at 13,000 rpm. The supernatant was decanted and the pellet was re-suspended in 300 µL ethanol. The wash and rinse was repeated 2 times, and the purified particles were stored in MilliQ water. Aqueous solution of Pd particles was deposited via nebulizer on a SiN window. The tomography study utilized a sample taken after incubating for 8 hours.

Data Acquisition

Several tomographic tilt series were acquired from Pd nanoparticles using the TEAM 0.5 microscope and TEAM stage [56] at the National Center for Electron Microscopy in the Molecular Foundry. Images were acquired at 200 kV in ADF-STEM mode with a 25 mrad convergence semi-angle (resulting in a probe size of 0.8 Å), 41 mrad and 210 mrad detector inner and outer semi-angles, and a beam current of 4 pA. The tilt series was collected at 49 angles with a tilt range of 64 to −61 degrees. Two images per tilt angle were measured with 3 µs dwell time with a scan rotation of 0 and 90 degrees to minimize image blurring. Owing to imperfections in the calibration of the x- and y- scanning coils in the microscope's STEM scanning system, an additional correction was applied to the images to ensure square pixels. This scan distortion was measured using a standard sample under the same imaging conditions, and corrected using the method described by Ophus *et al.* [57].

Image denoising

The drift-corrected images were denoised with the BM3D algorithm [58], with Anscombe variance-stabilizing transform and its inverse applied before and after denoising [59].

Tomographic reconstruction

After denoising, the 49 images were projected onto the tilt axis, and aligned with sub-pixel cross-correlation among the 1D-curves. A constant intensity scaling factor was fitted to the 1D curves to account for nonlinearities in the intensity at high tilt angles. Then the images were aligned to neighboring tilts with 2D subpixel cross-correlation. From the initially-aligned tilt series a 3D reconstruction was performed using the fast adaptive shrinkage-thresholding algorithm (FASTA), an accelerated gradient algorithm with adaptive stepsize for faster convergence [60]. To compute the forward and backward projections, we used the generalized ray transform interface of the Operator Discretization Library [61] in a 3D parallel-beam Euler geometry with an GPU-accelerated backend of the ASTRA tomography toolbox [62]. To increase the accuracy of the projections, we used the trilinear interpolation feature of the ASTRA library to compute the forward and inverse Ray-transforms. The code is freely available at this url. To minimize the translational and angular misalignments, we use a projection matching approach [63] with simulated annealing, where all three Euler angles are varied by a randomly picked value in the range of -0.5 to 0.5 degrees and the calculated projection error compared with the current projection error after a full reconstruction. The lowest-error angles are then used as new initial angles for the next tomographic reconstruction. This process is repeated for 10 outer iterations and the random Euler angle perturbation reduced linearly every iteration. Using this approach, the reconstruction converged to an R-factor of 8.1 %.

Atom tracing

The 3D atomic positions of the Pd atoms were determined using the following procedure based on the code provided by Ren *et al.* [64]. (I) all local intensity maxima were identified from the 3D reconstruction and added to a candidate list. From the initial candidate list, peaks which were within a minimum distance of 2.2 Å of a higher-intensity peak were deleted. (II) The initial list of peak positions was refined by fitting a 3D Gaussian function to each peak after subtracting neighboring peaks within a maximum radius of 4 Å. Using this initial atom candidate list, we added, refined, and merged new unidentified peaks for 4 iterations in the following order: (III) Subtract the fitted Gaussians of all current peak candidates from the reconstruction volume. (IV) Add new candidate peaks over an intensity threshold of 50 to the candidate list. (V) Refine the positions of all atom candidates as in (II) for 4 iterations. (VI) Merge peaks that are closer than a minimum distance of 2.2 Å after position refinement. (VII) Refine the positions of all atom candidates as in (II) for 4 iterations. (VIII) go to (III) if iterations not done. (IX) A final set of 4 positions refinement iterations as in

step (II) yielded the final set of coordinates of 22412 candidate atoms. Previous work then used an atom flipping procedure to eliminate low-intensity atom candidates whose addition does not decrease the experimental error.

We found that some low-intensity atoms were in the center of the particle, therefore we included additional features to eliminate non-atoms from the atom candidate list. For each atom candidate we computed the radial distribution function (RDF) and split it up into 7 radial regions corresponding to sections between the peaks of the total RDF. The sectioned RDF together with the fitted peak intensity, sigma, and voxel intensities of the peak in a 2 Å radius then formed a X-dimensional feature vector for each atom candidate. We then used the Uniform Manifold Approximation and Projection (UMAP) [65] method to project the feature vector onto a 2-dimensional manifold, shown in Appendix 1 (Fig. A.4). On one end of this manifold we identified disordered surface atoms with low intensity. We then used a Bayesian Gaussian Mixture Model [66] to classify the candidate list in the 2-dimensional manifold into two classes of potential atoms and non-atoms. Then we ranked the candidates according to their probability to belong to the non-atom class. To select a threshold for including atoms in the non-atom class and we performed ADF-STEM simulations (see the following section) of our tilt series with different probability thresholds for including atoms in the non-atom class. The ADF-STEM simulations were then compared with the Fourier Ring Information (FRI) criterion [67] to find the threshold that extracts the largest amount of information from the experimental data. The results are shown in Appendix 1 Figure 3. We found that a threshold of 99.994 % extracted the maximum amount of FRI from the data.

We then finalized the set of atoms by performing the following atom-flipping procedure to determine if an atom within the bracket from 99.99 % to 99.997 % should be added. For each atom that falls in this probability range we compute a tilt series with and without that atom and compare the resulting FRI, using a simple linear image formation model, with the image being a linear sum of 3D Gaussian distributions for each atom, and the standard deviation determined from the atom fitting procedure. If addition of an atom increases the FRI we include it in the list of atoms, otherwise we exclude it. This procedure yields a final set of 20632 Pd atoms in the particle.

STEM simulations and tracing precision

For atom classification and to evaluate the precision of the atom tracing procedure, we re-created the tilt series of 49 projections with the refined experimental Euler angles from the traced coordinates with quantum mechanical STEM simulations using the PRISM algorithm[68] implemented in the Prismatic simulation software [69]. A total of 49 cubic super cells of size 11 nm was created. The final atomic model was placed within the super cells. Individual super cells were divided into 2 Å slices along the beam direction and sampled with a pixel size of 6.2 pm in the transverse direction. The experimental parameters of 200 kV high tension, 25 mrad convergence semi-angle, 41 mrad and 210 mrad detector inner and outer semi-angles, 0 mm C_3 aberration were used for the simulation. We employed a Fourier

interpolation factor of 10 in the PRISM algorithm and simulated 8 frozen phonon configurations. We matched the probe step to the reconstruction voxel size of 25 pm. Each simulated ADF-STEM image was convolved with a Gaussian function to simulate incoherent source spread and other incoherent effects and minimize the difference to experimental images. A 3D volume was then reconstructed from the simulated tilt series with the FASTA algorithm described above and a new model was obtained by using the same atom tracing procedure. The new atomic coordinates were rotated and translated to minimize the global position deviations between the models and then atoms were matched between the models and a root mean square displacement calculated between the models, with a radial search cutoff of 1.5 Å around each atom. A histogram of the atomic deviation between the common pairs is shown in Appendix 1 Figure 4, indicating a mean deviation of 32.2 pm. This is slightly higher than the deviation of previous AET studies [59], which we attribute to the relatively low number of projection measurements available relative to the size of the reconstructed particle.

Atom Classification

To classify traced atoms according to their crystallographic coordination, we first generated polyhedra with 12 vertices $\mathbf{p}_j$ arranged in both fcc and hcp stacking geometries with nearest neighbor spacing equal to the mean measured value of 2.93 Å. These polyhedra are rotated to 1026 orientations roughly evenly spaced on 1/12th of the unit sphere. For each atomic site, these polyhedra were rotated using a matrix $\mathbf{m}$ to minimize the total distance between their (ideal) coordinates and the nearest relative atomic site using the iterative closest point (ICP) algorithm [70]. Finally we compute an order parameter for each polyhedra at each site s_k equal to

$$s_k = \sum_{j=1}^{12} \max\left(1 - \frac{|\mathbf{r}_j - \mathbf{r}_k - \mathbf{m}\,\mathbf{p}_j|}{d_{\max}}, 0\right), \tag{2.1}$$

where $\mathbf{r}_j$ is the position of the j'th neighboring coordinate to site k at position $\mathbf{r}_k$, and $d_{\max}$ is maximum allowed distance of a site from an ideal position, which we set equal to half the average nearest neighbor distance of 1.47 Å. This cost function can generate values of 0 to 12, where a value of 12 indicates perfect alignment between the polyhedral template. We keep only the maximum order parameter (best agreement) for both fcc and hcp polyhedra, and use these values to determine the local ordering.

2.3 Results and Discussion

Population Statistics of Pd Nanoparticles

To understand the relationship between MTP size and structure, we first analyzed Pd nanoparticle populations using HR-STEM. Following previous work [20], we used an aqueous synthesis known to produce decahedral particles that employs polyvinylpyrrolidone (PVP)

as a stabilizing agent and citrate as a reducing agent and capping agent. To broadly capture size and shape statistics, incubation time was varied from 1 hour to 24 hours (See methods section). Population statistics of the synthesized nanoparticles were manually determined from 691 particles identified in the HR-STEM data. Figure 2.1 summarizes the results of the HR-STEM study. The results confirm the presence of decahedra, iscoahedra, and other particles that appear to be undergoing a successive twinning process. Figures 2.1a-c show sample micrographs of these structures. The size and shape distributions presented in Figure 2.1d show a clear size trend between the three structures. The icosahedra and successive twinning structures have size distributions skewed towards smaller sizes, with average sizes of 9.87 nm $\pm$ 0.57 nm and 8.28 nm $\pm$ 0.41 nm, respectively. The decahedra show a more uniform distribution with an average size of 20.8 nm $\pm$ 1.39 nm . The synthesis conditions used for this study were a relatively high concentration of reducing agent as well as a stabilizing agent (PVP), and should result primarily in decahedra [46]. PVP, however, primarily interacts with nanoparticles larger than 10 nm [71]. If the PVP interaction is critical to stabilizing growing decahedra this may explain why multi-tetrahedron and icosahedron particles are primarily observed at sizes below approximately 10 nm.

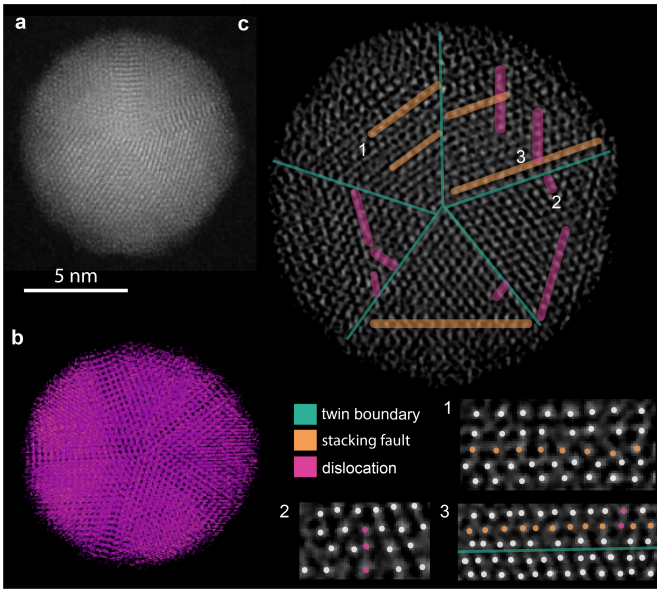


Figure 2.2: a) Sample STEM micrograph from the tilt series. b) Volumetric rendering of the tomographic reconstruction. c) An example slice through the atomic volume normal to the five-fold axis with various defects highlighted.

Atomic Electron Tomography of a Pd MTP

To obtain a detailed understanding of the three-dimensional atomic structure of an MTP undergoing successive twinning, we performed AET on a representative Pd particle. We used aberration-corrected STEM to collect a tilt series of 49 images of an approximately 10 nm diameter particle, with tilt angles ranging from -61 to 64° (Appendix 1 Fig. A.1). Images taken before and after the tilt series acquisition indicate that the structure of the particle did not change during imaging (Appendix 1 Fig. A.2). While the shape and crystal structure of the particle were largely consistent with a decahedron (Fig. 2.2a), several images in the tilt series indicate the presence of additional crystalline grains in the nanoparticle, suggesting that the particle is undergoing successive twinning (Appendix 1 Fig. A.1). After denoising and aligning the tilt series (Methods), we reconstructed the volume, shown in Figure 2.2b, using a FASTA based reconstruction algorithm (Methods). To assess the consistency between the final reconstructed volume and the input projections, we define the R-factor as the pixel-wise difference of the absolute values of the measured and calculated projections from the reconstructed volume, normalized by the intensities of the measured projections. The R-factor of the final reconstruction was 8.1%, which is consistent with other reported AET reconstructions [72].

After the final reconstruction volume was obtained, atomic positions were determined using an iterative 3D Gaussian fitting procedure. Atomic locations were included or excluded based on unsupervised clustering of the atoms based on atom intensities and radial distribution function of the atoms. Subsequently we determined the atom set that maximizes the Fourier Ring Information [67] between a linear image generated from the atom positions and the measured projections (see Methods for a detailed description). Using this method, we determined the atom positions of 20632 atoms, shown in Figure 2.3a.

A qualitative analysis of the structures present in the reconstructed volume, shown in Figure 2.2b, reveals a large number of defects in the structure. To illustrate the large number of both stacking faults and dislocations in the particle a 0.25 Å-thick slice through the volume with defects highlighted is presented in Figure 2.2. Many of the stacking faults and associated dislocations are observed adjacent to the twin boundaries, which was also observed in with molecular dynamics simulations of five-fold twinned metal nanowires [73]. However, defects and disorder are not restricted to regions around the twin boundaries and can be found within the tetrahedral subunit bulk, as shown in defect 3 in Figure 2.2 c). The number of stacking faults and edge dislocations observed in the reconstructed particle is much higher than in previous studies [19, 55], and we also see many more defects than predicted for nanomaterials of this size [73]. The stacking fault and twinning energy of fcc metals is in general low [74], implying facile formation of stacking faults and twins during nanoparticle growth. Therefore, a possible explanation for the high number of defects in the particle is that the stacking fault mobility is slow compared to the rate of adatom addition to the particle.

Ideal fcc tetrahedra cannot be tiled into a decahedron or icosahedron in a way that is space filling [75]. MTP decahedra and icosahedra, consequently, must contain a high degree

of internal strain because the crystal lattice must accommodate the missing volume [76, 77]. Defects, such as stacking faults, are one mechanism for stress relief in MTPs [77]. Therefore, it is likely that the presence of the observed defects is in part due to stress relief, and is consistent with inhomogeneous strain within the particle.

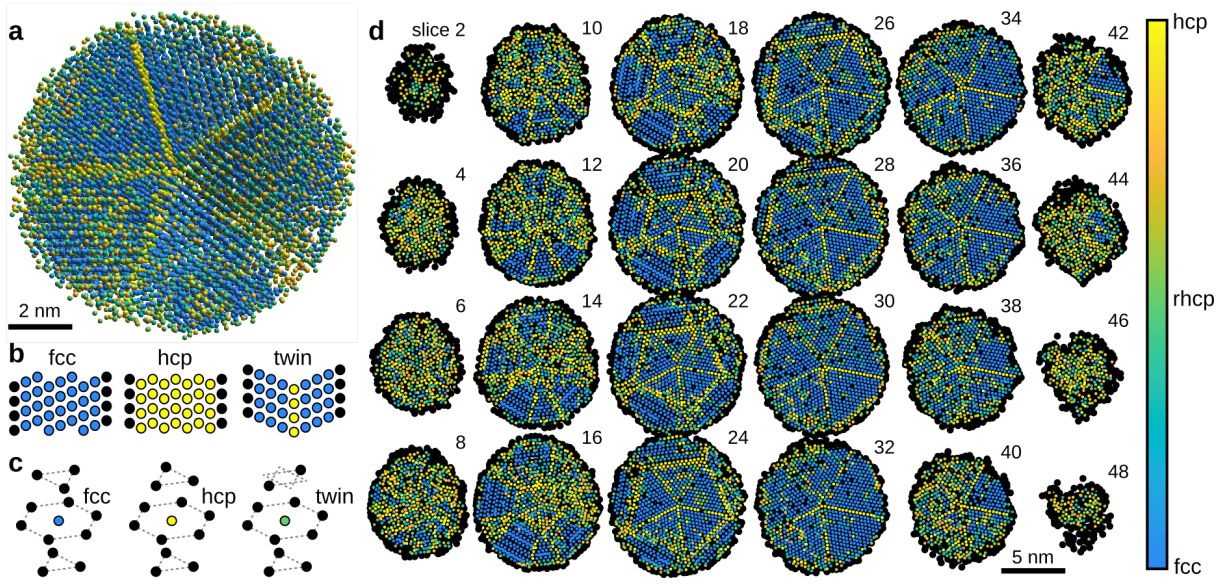


Figure 2.3: Crystallographic ordering of the atoms in the Pd MTP. (a) Traced atom positions with coordination numbers of 12, colored by fcc to hcp ordering. (b) Atomic stacking arrangement of fcc, hcp, and a twinned structure view from the side. (c) 3D atomic neighborhood of sides with fcc, hcp, and rhcp ordering. (d) Slices throughout the MTP showing the crystallographic ordering for each atom with at least 10 neighboring sites.

Crystal Structure of Successively Twinning Particle

To better understand the crystal structure in the reconstructed particle we classified the traced coordinates according to their crystallographic coordination. First, all sites were sorted according to their coordination number using a cutoff radius of 3.75 Å. Atoms with coordination numbers below 10 were classified as surface sites, or highly disordered regions if they were fully contained in the nanoparticle. Most remaining sites in the particle bulk were primarily arranged into 4-atom tetrahedra. The most common structural packing of these tetrahedra were either fcc or hcp ordering. However, a large degree of disorder was present in many regions of the MTP on length scales larger than 4-atom clusters.

In order to determine the large scale structure of the MTP, we used a polyhedral matching algorithm inspired by Larsen *et al.* [78] to determine the degree of fcc and hcp ordering of each atomic shell. In this dataset, we found that polyhedral matching produced much more

robust identification of the differently ordered MTP regions than other methods such as local bond order parameters [79] or common neighbor analysis [80] (Appendix 1). A detailed description of the classification protocol is found in the Methods section. Briefly, for each valid atomic site we define polyhedra with vertices that correspond to fcc or hcp stacking geometries and fit them to the measured positions of the surrounding 12 atoms, depicted in Figure 2.3c. After rotating the polyhedra to minimize the distance between the ideal and measured atomic positions of the 12 nearest neighbors, we compute an order parameter for both fcc and hcp ordering (Methods). We then keep only the maximum order parameter for both fcc and hcp polyhedra. These values are used to determine the local ordering.

The results of this classification algorithm are shown in Figure 2.3. The 3D atomic coordinates with 12 nearest neighbors are shown in Figure 2.3a, where each site is colored according to the difference between the hcp and fcc order parameters defined above, from -1 to 1. The colors are strongly bimodal, indicating that the majority of sites possess either highly ordered hcp or fcc arrangements. The overall 5-fold symmetry of the nanoparticle is immediately obvious, though a significant amount of disorder is present both on the particle surface and in the bulk structure.

To visualize the structural ordering of the entire structure, we have plotted every other atomic slice of the particle in Figure 2.3d, where the atoms are again colored by the difference between the hcp and fcc order parameters. Surface atoms are shown in black, as many of these sites do not have enough nearest neighbors to distinguish between fcc and hcp ordering. Figure 2.3b shows the two endpoint order parameters, the ideal fcc and hcp structures, viewed from the side. Figure 2.3b also shows a third structure motif which appears in many locations in Figure 2.3d, a twinned fcc structure, where a single line of atoms possessing high hcp ordering separates two fcc grains with mirrored structures. Figure 2.3c shows the 3D atomic arrangement of the two ideal polyhedra with fcc and hcp ordering, as well as a third structure which has the same order parameter for both structures (10 out of 12 possible sites agree with each class of polyhedron). This structure is labeled as random hexagonal closed packed (rhcp), following other studies which have used this label for packing falling between fcc and hcp [81].

The twinning structure obtained reveals that the reconstructed particle contains a core decahedron, with additional, partially formed tetrahedra on top. The resultant structure is consistent with a partially formed icosahedron (Figure 2.4a.) Considering the arrangement of partially formed tetrahedral subunits atop the decahedral core of the reconstructed nanoparticle, shown in Figure 2.4a, the particle seems to be captured in the midst of successively twinning, and has partially transformed into an icosahedron. There are ten additional partially formed tetradedra, forming two rows above the core decahedron. The tetrahedra closer to the decahedron are more fully formed. These nucleating tetrahedral units are highlighted in Figure 2.4a.

Prior observations of successive twinning have been in particles on the 100 nm scale and have shown tetrahedron by tetrahedron growth [49, 47, 51] illustrated schematically in Figure 2.4b, pathway A. However, the successive twinning we observe, illustrated in Figure 2.4b pathway B, is better described as a simultaneous process, where the tetrahedra comprising

the center portion of the icosahedra grow at the same time. This process more closely resembles the successive twinning process predicted for metal nanoclusters, where coordinated formation of hcp layer across a decahedral surface leads to simultaneous twinning [82, 83]. It is known that decahedra have more hcp adatom sites than fcc [82], making hcp island growth, and therefore the growth of stacking faults and twin boundaries, more probable, especially if there is limited surface diffusion. Our observations also directly confirm addition of multiple tetrahedra to a decahedral particle as a route towards an icosahedral structure.

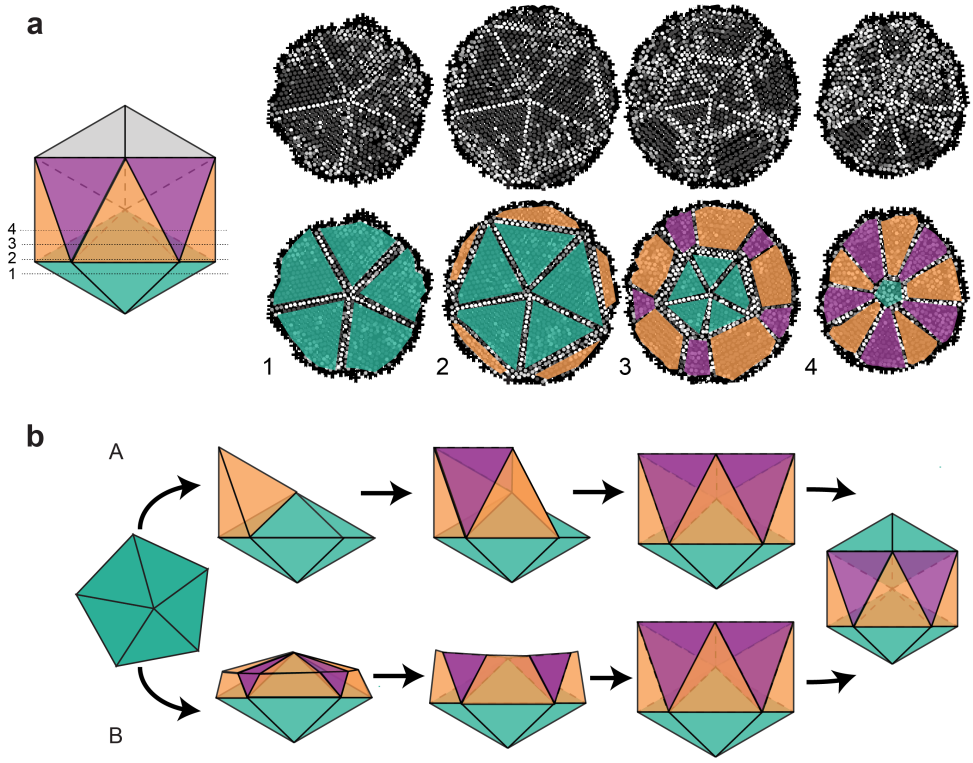


Figure 2.4: a) Sample slices from the labeled atomic coordinates shown in Figure 2.3 labeled with how they would correspond to slices through an icosahedron. b) Schematic of the two pathways observed for the successive twinning process.

Disorder close to the surface

In addition to illustrating the twinning and defect structures present, our traced atomic coordinate classification protocol also reveals a large degree of disorder close to the bottom and top surface of the Pd nanoparticle. An ideal decahedral particle would exhibit ⟨111⟩ terminated surface facets. The top and bottom surface slices shown in the first and last column of Figure 2.3c show that the particle does not have ⟨111⟩ terminating facets. Instead,

this surface of the particle contains a mixture of coordination types, including a significant fraction of random hexagonally close-packed (rhcp) coordinated atomic sites whose relation to fcc and hcp stacking is shown in Figure 2.3a. rhcp stacking has previously been observed during palladium nucleation and is further evidence of the non-equilibrium structure of these growing MTP particles [21]. The surface disorder is also an indication that surface diffusion on the particle is relatively slow [84].

2.4 Conclusion

In this study we have determined the 3D atomic positions of over 20,000 atoms in a multiply twinned palladium nanoparticle. We have found that the structure derived from the aqueous synthesis contains substantially more defects than would be expected from previous computational and experimental studies. We directly observe a simultaneous successive twinning process wherein a decahedral particle is transitioning directly to an icosahedron. Based on our HR-STEM studies, we suspect this process would only occur in small decahedral nanoparticles under our reaction conditions, as after a certain size the structural directing PVP will have a stronger influence on growth.

The complex structure observed has implications for nanoparticle functionality. The lack of $\langle 111 \rangle$ terminating facets will significantly alter catalytic reactivity in a Pd particle. Similarly, the high number of defects in the interior structure of the particle will change the surface strain states of the particle, which have will also affect catalytic activity [40, 7]. Therefore, the combination of HR-STEM and AET used in this study provide unique insight into the structure and functionality of MTPs. Future work will need to determine how the synthetic parameters influence the development of this defected transient structure.

Chapter 3

Automated Nanoparticle and Defect Analysis[1]

3.1 Introduction

In the last chapter we saw the influence of synthetic pathway on the development of defects in nanoparticle systems. It is clear from that study that there is a need for rapid, but accurate, image analysis in electron microscopy studies of nanomaterials. With the advent of fast, high efficiency electron detectors and automated imaging protocols [30, 31], incorporating electron microscopy into high throughput materials design efforts [85] is increasingly feasible. These new and upcoming capabilities strongly motivate automated methods to extract relevant structural features, such as nanoparticle size, shape, and defect content, from high resolution transmission electron microscopy HRTEM) data to link these features to bulk properties and study the influence of heterogeneity on bulk behavior [86, 5, 87]. In general, protocols which outperform classical image analysis and do not require time-consuming manual analysis are needed. Given recent advances in image interpretation using deep learning [88, 89], segmentation *via* convolutional neural networks (CNN), along with other machine learning techniques, are promising routes toward automatic interpretation of HRTEM micrographs. Here, we demonstrate a two-step pipeline to detect and classify regions of interest in HRTEM micrographs. This tool uses a convolutional neural net to identify crystalline regions (nanoparticles) from an amorphous background in the images, and then feeds individual regions of interest into a random forest classifier to detect whether or not they contain a stacking fault [90].

Previous work has applied deep learning to atomic resolution images of nanomaterials from high resolution scanning transmission electron microscopy and HRTEM, but this has focused on atomic column segmentation or atomic column tracking and specific structure

[1]This chapter is a version of C. Groschner, C. Choi, and M.C. Scott. "Machine Learning Pipeline for Segmentation and Defect Identification from High-Resolution Transmission Electron Microscopy Data." *Microscopy and Microanalysis*, 27, 3, p.549–556 (2021).

analysis [33, 37, 34, 91, 92, 93, 94, 95, 96, 97, 98]. Other work has focused on lower resolution TEM micrographs, where amplitude contrast is the dominant contrast mechanism [99, 100, 101, 102, 103, 104, 105]. Less is known about how to identify isolated regions of interest in HRTEM data and classify the structures therein. Here, we aimed to first identify nanoparticle regions, isolate them, and then classify the nanoparticles according to their structure.

We have focused on conventional HRTEM as it is a dose-efficient and high frame rate imaging mode and therefore useful for fast data acquisition on a wide range of materials. However, segmentation in HRTEM is particularly challenging as the contrast between the substrate that the nanomaterial sits on and the nanomaterial itself can be very low. Because traditional image processing techniques are highly error prone for these types of images, we have implemented CNN for semantic segmentation, which is segmentation of the image on a pixel by pixel basis. Our CNN has a modified U-Net architecture [89] and can accurately segment a diverse population of nanoparticles despite a small number of training images. Segmentation is demonstrated on both gold and cadmium selenide nanoparticles. After segmentation, individual nanoparticle regions can be isolated and fed directly into existing python tools to extract size and shape statistics. To detect the presence of defects in nanoparticle regions, we implement a random forest classifier. We demonstrated the ability of the random forest classifier to detect stacking faults in the CdSe subset of identified nanoparticles. Both the CNN and classifier demonstrate state of the art performance at their respective tasks. While this work focuses on HRTEM images of nanoparticles supported on a carbon substrate, in principle the tool can be used to detect any regions of crystallinity in HRTEM data, and the simple random forest classifier is designed to be easily retrained to detect a variety of types of defects, making this a flexible pipeline suited for a variety of image analysis tasks.

3.2 Methods

Strategy

The overall goal of this work was to create a tool for automated structural analysis of nanoparticles from HRTEM data. In order to achieve this, we created a two-step pipeline for accurate segmentation and defect classification. The first stage of the analysis pipeline was to train a U-Net convolutional neural network to segment particle regions. The details of the architecture are shown in Figure 3.1. After segmentation, we isolated individual nanoparticle regions using a series of morphological closings and openings. These isolated regions were used for feature generation and training of a random forest classifier to detect stacking faults.

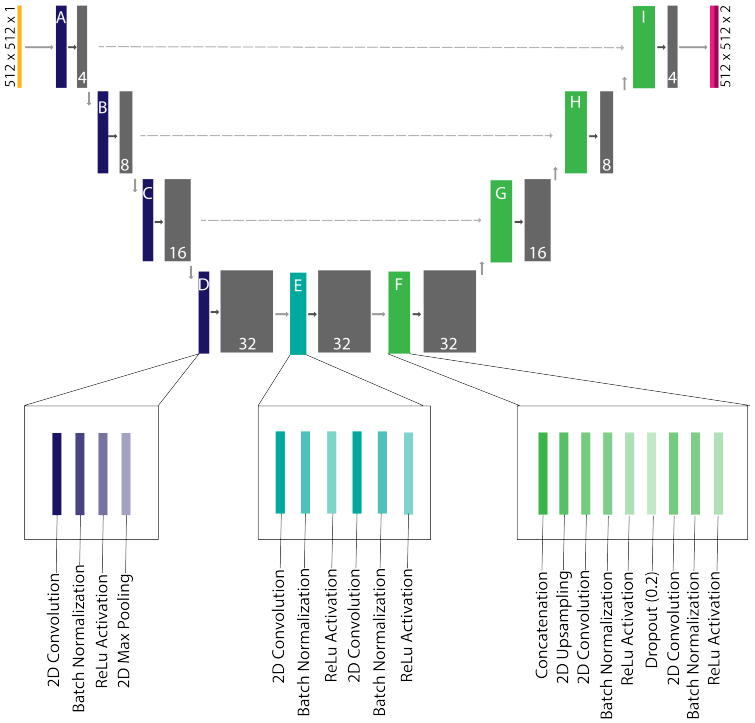


Figure 3.1: The U-Net style architecture of CNN implemented for segmentation of nanoparticle regions. Blocks colored gray represent the output features and are labeled with the number of features created. Residual blocks A-D contain a 2D convolutional layer, batch normalization layer, rectified linear unit (ReLu) activation layer, and a 2D max pooling layer. Residual block E contains a 2D convolutional layer, batch normalization layer, and ReLu activation layer, which is repeated once. Residual blocks F-I contain a 2D upsampling layer, concatenation layer, a 2D convolution, batch normalization, and ReLu activation followed by a dropout layer and another set of 2D convolution, batch normalization, and ReLu activation layers. The final layer is a softmax activation layer which outputs two segmentation maps one for each class, background and particle.

Data Collection and Preprocessing

We collected 46 1024×1024 micrographs of CdSe nanoparticles and 13 4096×4096 micrographs of Au nanoparticles using an aberration corrected TEM at 300kV. To reduce the computational demands, we sliced each micrograph into four or sixteen 512×512 images. To establish a ground truth for segmentation the data was manually labeled by an experienced researcher into either a particle class or background class. The ground truth segmentation maps were created using the MATLAB labeler application [106]. All the micrographs were then normalized by first applying a median filter with a 3×3 kernel which served to remove spurious X-rays. After filtering, the images were normalized.

Significant portions of the raw input images were carbon background, which meant that when sliced into 512×512 images only 28% of images contained nanoparticles. Therefore, any image that did not contain particle class pixels was discarded from the training and test set. The remaining data was split so that 129 images were used in the training set and 43 images were in the test set. Images were rotated and flipped to augment the dataset to more than 1000 images which were then split 4:1 into the training and validation sets.

Network Architecture

The architecture implemented was based on the U-Net architecture developed by Ronneberger *et al.* following the same convolution-deconvolution structure [89]. We chose the U-Net architecture because it had previously been successful on segmentation of gray scale micrographs [89]. The architecture also is designed to combine high and low spatial frequency features using concatenation of features between layers [89]. Sampling across feature space is critical in HR-TEM images where key features are high frequency lattice fringes and low frequency particle edges. We modified the standard U-Net structure for accurate segmentation and to prevent overfitting.

Network Training

Training was performed using the Savio GPU cluster at U.C. Berkeley using a compute node with two Intel Xeon E5-2623 v3 CPU cores and an NVIDIA K80 GPU. The node had 64 GB of available RAM. The model was built using Keras with a Tensorflow backend [107]. Training used 129 sample images, which then fed into a Keras image augmentation generator [107], which randomly rotated and flipped the images. Each epoch, or complete pass through the training set, contained 1,000 samples created from rotating the original processed 129 micrographs and split between the training and validation sets in a 4:1 ratio. Training was stopped once the loss on the validation set did not decrease by 0.001 after two epochs. The model was limited to train for a maximum of ten epochs. The model used categorical cross-entropy for the loss function and Adam as the optimizer with a learning rate of 1×10^{-4} [108, 109]. Batch size was set to 20, where batch size defines how many samples from the training set are propagated through the network before updating the model weights.

Testing

A holdout test dataset of 43 images was reserved for final testing after model training had completed. The predicted particle segmentation map was thresholded to 0.5 prior to metric computation because this represented the best tradeoff between precision and recall. Precision and recall versus threshold curves for the network both on Au and CdSe data are provided in Appendix 2. Reported metrics are the results from these test sets.

Alternate Methods for Segmentation

Two standard segmentation methods were applied to the dataset to compare to the U-Net segmentation. Standard thresholding was applied using Otsu's method [110] to threshold out the nanoparticles based on intensity. Fourier filtering was also used to segment out crystalline regions. For this filtering, the Fourier transform of the image was multiplied by an annulus to eliminate frequencies related to the lattice structure of the material. The inverse Fourier transform of the filtered image was then smoothed and thresholded to create a segmentation map.

Random Forest Classifier

After applying the neural network, the resulting segmentation masks were used to isolate particle regions. To limit the number of agglomerated regions input into the random forest, we implemented morphological filtering on predicted segmentation maps using the scikit-image python package [111]. Regions which were too large to be independent particles were discarded before stacking fault detection. The random forest was developed using the scikit-learn python package [66]. A total of 329 particle regions were in the training set and 163 regions were in the test and validation set. The training set had 71 examples of stacking faults, 156 examples of no stacking fault, 80 examples particle not on the zone axis, 60 examples of agglomerations, and 43 empty regions. The training and validation set were manually labeled. The random forest had 500 decision trees, the criterion for tree splitting was gini impurity, which calculates the probability of classifying the data incorrectly. A limited hyperparameter serach was performed during random forest development. This hyperparameter search lead to the selection of gini impurity and 500 decision trees because these parameters were found to have the best performance. The feature set consisted of the mean, standard deviation, and center of mass of the radially integrated Fourier transform of the particle region and the mean and standard deviation of the nanoparticle image in real space. This feature set was selected such that validation accuracy and true positive rate for each class were highest. These results lead to a small and informationally dense feature set based on summary statistics.

3.3 Results and Discussion

Evaluation Metrics for Segmentation

To evaluate the neural network's ability to segment nanoparticles from the micrograph we used three evaluation metrics: (1) Dice coefficient (DICE), also known as F1-score, (2) precision, and (3) recall. We look at three metrics because each offers a difference weighting of error types. Dice coefficient is a standard metric for segmentation tasks and measures the union of positive pixels between the true and predicted segmentation masks normalized by the total number of positive pixels in the two groups. The Dice coefficient can be a very

strict metric for small particle regions because of the weighting of true positive pixels, shown in Equation 3.1, where TP stands for true positives, FP for false positives and FN for false negatives [112]. Fewer positive pixels overall means small errors can have a strong effect on the metric. Therefore it is also important to consider other metrics, such as precision and recall. Precision essentially measures how successful the neural network is at only selecting true particles. Recall provides information on how well the network can find particles.

$$DICE = \frac{2TP}{2TP + FP + FN} \tag{3.1}$$

Performance of U-Net Segmentation and Comparison with Standard Image Processing Methods

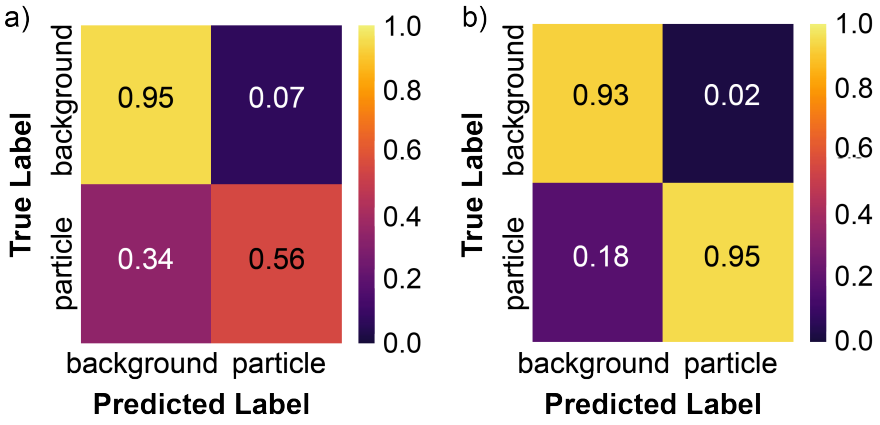


Figure 3.2: The confusion matrices for the trained U-Net on (a) CdSe test set and (b) for the Au test set.

We find segmentation via U-Net style CNN to be the most effective way to segment micrographs as compared to standard methods. Overall, the Dice coefficient across the test set was 0.8. Separating the test set according to sample material, the network achieved a Dice coefficient for Au nanoparticle micrographs of 0.89 while the CdSe nanoparticles were segmented with a Dice coefficient of 0.59, which is on par with other state of the art segmentation procedures [113]. Further metrics are provided in Table 3.1 and the confusion matrices presented in Figure 3.2. Sample segmentations for both are shown in Figure 3.3. CdSe particles clearly are the much more difficult segmentation case, both for human and computer labelers. CNN based segmentation is likely limited due to the limited signal relative to background and the size of the particle regions (limiting the context for the neural network to learn). We explored whether training solely on CdSe would improve the segmentation but we found that training on a diverse set of materials actually improved results. These results are presented in the Appendix 2. Despite a lower Dice coefficient on the CdSe data, the

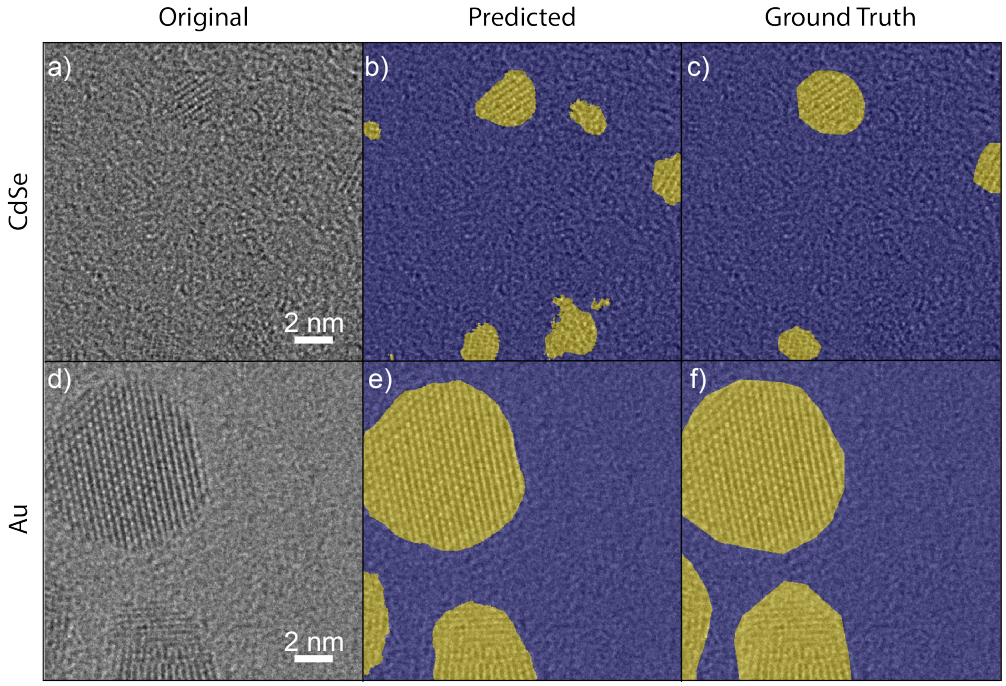


Figure 3.3: The top row, a–c, of the figure shows a sample micrograph, segmentation from the network and ground truth segmentation from the CdSe dataset. The second row, d–f, of the figure shows a sample micrograph, segmentation from the network and ground truth segmentation from the Au dataset. For all segmentation maps, yellow represents regions predicted to be in the particle class, blue regions predicted as background.

network created by the demonstrated training procedure results in a network which tends to under predict particle regions, as can be seen in Figure 3.2(a), from the low false positive rate for the particle class. This bias is preferred to over prediction of particles which would make later classification challenging.

	DICE	Precision	Recall
Combined Dataset	0.8	0.82	0.78
CdSe Data	0.59	0.56	0.62
Au Data	0.89	0.95	0.84

Table 3.1: Performance metrics for test sets

In order to determine if the network was more successful than standard image processing techniques, we segmented the micrographs using Otsu's method [110] and using Fourier filtering [114, 115]. More information for each segmentation method is supplied in the

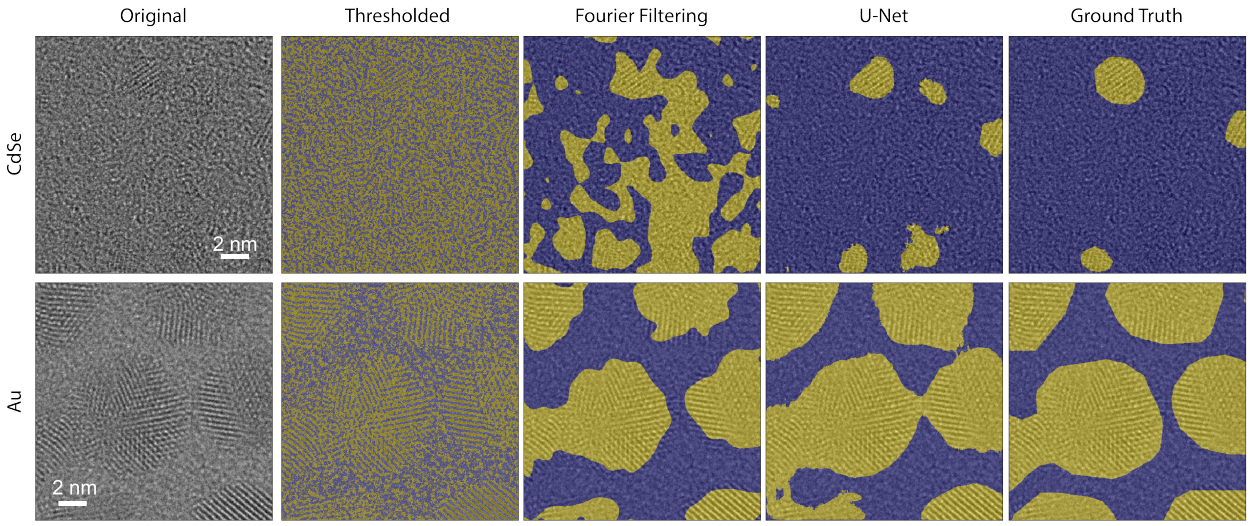


Figure 3.4: Sample micrographs of CdSe and Au particles and the resulting segmentation maps when segmented by Otsu's method, Fourier filtering, and U-Net. For all segmentation maps, yellow represents regions predicted to be in the particle class, blue regions predicted as background.

Methods section. We found that basic thresholding on the complete CdSe test set achieved a Dice coefficient of 0.21 and Fourier filtering achieved a Dice coefficient of 0.33. We note that this is significantly worse than using Otsu's method to threshold the Au test set, which leads to a Dice coefficient of 0.45 or using Fourier filtering which leads to a Dice coefficient 0.78. The Dice coefficients for the combined test set was 0.33 and 0.52 for thresholding via Otsu's method and Fourier filtering respectively. This takes into account optimizing the radius of the annulus mask in Fourier filtering which impacts the segmentation quality. Sample segmentations for each method are shown in Figure 3.4. Clearly, on all accounts, the neural network performs significantly better than standard segmentation methods, but it is a particular improvement for the CdSe dataset, where there is little amplitude contrast. In addition, beyond the improved segmentation, using a neural network also has the benefit of not hand tuning parameters for each image, which are critical in standard processing methods. Therefore, for this type of TEM data, U-Net provides a better, and more high throughput, means of particle segmentation.

CdSe Particle Statistics from Random Forest Classification

The primary reason to develop a highly accurate segmentation technique for HRTEM is to enable high throughput analysis of crystal structures. Segmented regions of interest can be used in further classification models and automated statistics. Statistics on size and shape can be provided directly from segmentation results. Segmented regions can also be passed

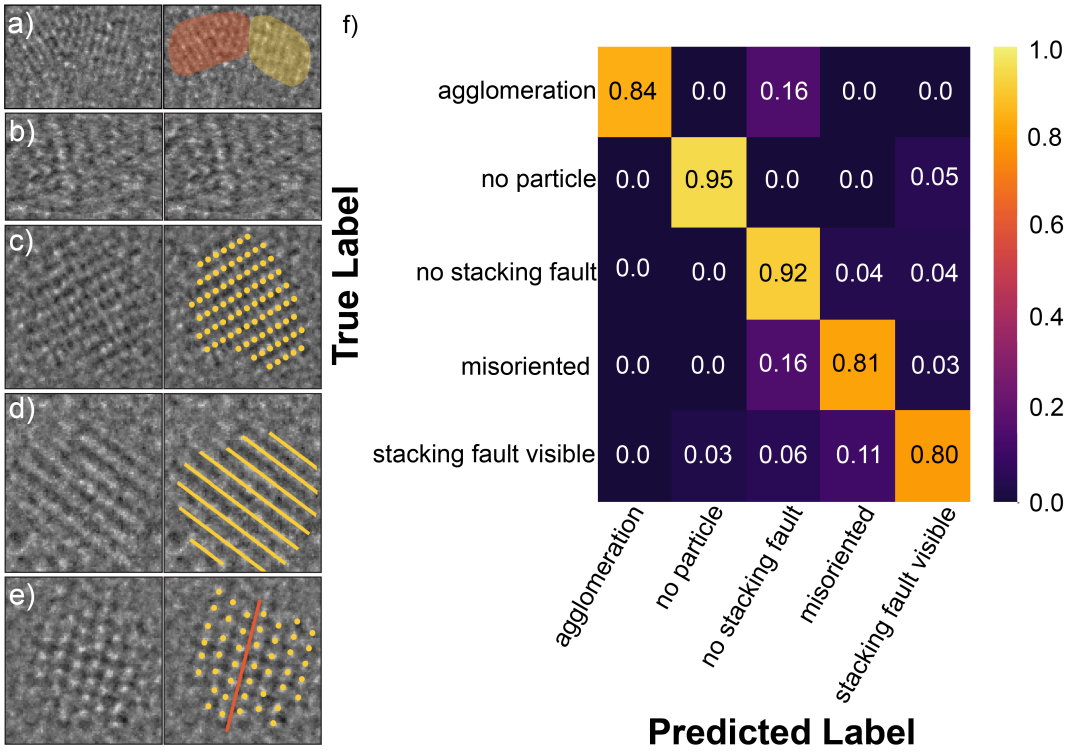


Figure 3.5: Sample micrographs, original on the left and annotated on the right, of the five classes predicted: a) agglomeration (the two particle regions are highlighted in orange and yellow respectively), b) no particle, c) no stacking fault (atomic columns, marked in yellow, are aligned), d) misoriented (atomic planes, marked by yellow lines, are visible but atomic columns are not), and e) stacking fault visible (atomic columns, marked in yellow, can be seen to be offset across the stacking fault, marked in orange). f) Confusion matrix for the random forest classifier.

to classifiers for structure or defect detection as we will demonstrate.

Stacking faults in CdSe are of interest due to the influence they can have on particle growth and therefore shape of CdSe nanoparticles [116]. Here we show that a simple random forest can be used to detect stacking faults. A random forest was chosen for several reasons, including the excellent performance of ensemble methods as well as the speed and simplicity to train [117]. Combined with high model interpretability, makes this classifier a good choice for fast iteration and rapid development.

The random forest demonstrated is used to categorize the CdSe nanoparticles into five different classes: (1) particle has a stacking fault visible, (2) particle does not have stacking fault visible, (3) particle does not have atomic column contrast (misoriented), (4) region is an agglomeration of particles, and (5) region is background. Sample micrographs for each class are shown in Figure 3.5(a–e). We included the last category to handle the few cases in

which the neural network segments a region incorrectly, since perfect segmentation cannot be guaranteed from our network. The feature set we developed for the random forest is extremely simple and consists only of the mean, standard deviation, and center of mass of the radially integrated Fourier transform of the particle region and the mean and standard deviation of the nanoparticle image in real space. Figure 3.5(f) shows the confusion matrix for the random forest classifier on the test data set. The confusion matrix provides the fraction of correct and incorrect labels for each class. The random forest achieves 86% class balanced accuracy compared to the expert labels given. The random forest predicted that 52% of particles were in the correct orientation to observe a stacking fault, and of those 38% of particles contain a stacking fault. This can be compared to the ground truth of 49% of particles which were in the correct orientation to observe a stacking fault and the 35% of those particles that contained a stacking fault. As a consequence, we can see that using an extremely limited feature space still yields good classification of particles.

In summary, this two step pipeline offers a flexible method to identify crystalline regions of interest and classify individual regions according to known features. The segmentation portion of the pipeline outperforms standard segmentation methods, even when those methods are optimized for individual datasets. The easily retrainable random forest classifier enables detection of known features within the identified regions. We note that the outputted individual regions could also be input into unsurpervised classifiers for detection of previously unidentified features. Further work could also aim to implement more complex structure and defect classification methods[118, 91]. One particularly useful future area of work would be to develop zone axis classification alongside defect identification, which would aid in the interpretation of observed defects.

Coupled with existing size and shape analysis tools, the output of the segmentation network and classifier demonstrated here offer a way to determine statistical distributions of features of interest, such as size, shape and defect presence, and importantly, allow one to detect correlations between these features. This capability will be critical to offer insight into nanomaterial population evolution during during high throughtput synthesis studies, or to identify key structural features that influence bulk properties in combinatorial synthesis efforts.

3.4 Conclusion

Here we have demonstrated a method of flexible automated analysis for nanoparticles in HRTEM by combining a neural network for segmentation and a random forest for classification tasks. We have shown that a U-Net architecture can far outperform traditional image processing techniques for segmentation of HRTEM images, while a simple classification tool can accurately classify structural features using a very limited number of parameters. Breaking apart the segmentation and classification tasks leads to an accurate tool with limited data labelling and feature engineering requirements. Moreover, this pipeline provides a flexible open source tool as the base for further analysis and classification tools for local atomic

structure. Such quantitative, automated analysis will have significant implications for a broad range of nanoparticle synthesis, structure-property relationship, and combinatorial synthesis studies.

Chapter 4

Architecture Optimization for HRTEM Segmentation

4.1 Introduction

We have already demonstrated one potential pipeline that includes segmentation of HRTEM micrographs. However, improvements in segmentation can certainly be made and as deep learning is becoming a key element of Transmission Electron Microscopy (TEM) analysis, it is important to identify best methods. Most TEM segmentation procedures rely on using the U-Net architecture proposed by Ronneberger and coworkers [89]. However, a plethora of architectures have been proposed for segmentation of real images. There has been an effort in the biomedical community to understand the impact of network architecture on segmentation [119]. However, this area is still just beginning to be developed for segmentation of TEM micrographs for materials science [120].

The segmentation problem is fundamental to enabling automated microscopy studies, therefore, we chose to analyze network architectures for the segmentation task. We analyze five different architectures. The first three are variations on the U-Net architecture which was used for segmentation in the previous chapter. U-Net consists of a contracting and expanding path and any number of convolutional neural network architectures can be put in as the contracting pathway, also referred to as the backbone of the network. We test three U-Net variations. The first is the standard implementation of U-Net with a VGG16 as the backbone [121]. The second is U-Net with a ResNet-18 backbone [122] and the third uses MobileNet as the backbone [123]. The layers which make up these three backbones are shown in Figure 4.1. We then endeavored beyond U-Net. Several segmentation architectures have been proposed since U-Net and we explore the use of two - what we refer to as an atrous network which relies on atrous convolutional layers as described by Chen *et al.* [124], and the feature pyramid network (FPN) developed by Lin *et al.* [125]. Atrous convolution promises improvements based on more efficiently sampling the image space by increasing the strides between convolutions as shown in Figure 4.2. FPN may improve segmentation based on the

combination of features from different levels of down sampling as shown in Figure 4.3.

We find that the ResNet-18 based U-Net and the FPN perform the best. All networks tested far outperform the standard, VGG16-based U-Net. We find that the FPN actually appears to outperform the labels provided and suggests that the current results are limited not by architecture but by the labeling of the data.

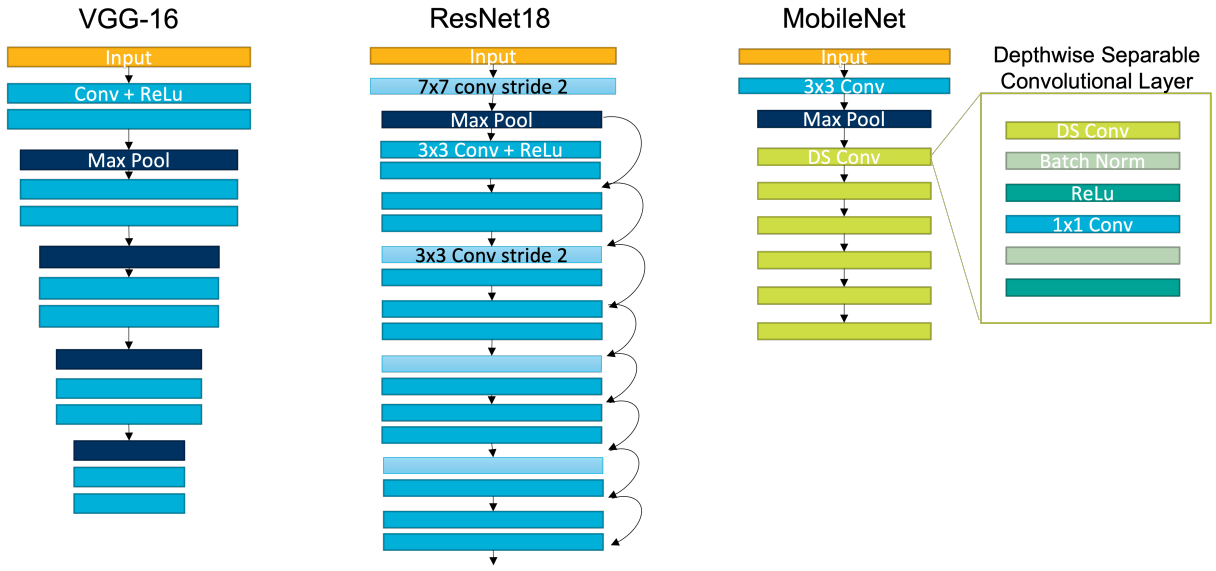


Figure 4.1: The convolutional neural network architectures used as different backbones for testing variations on the U-Net architecture.

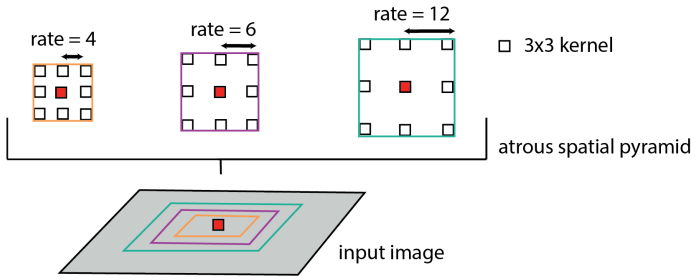


Figure 4.2: Schematic of the atrous convolutions.

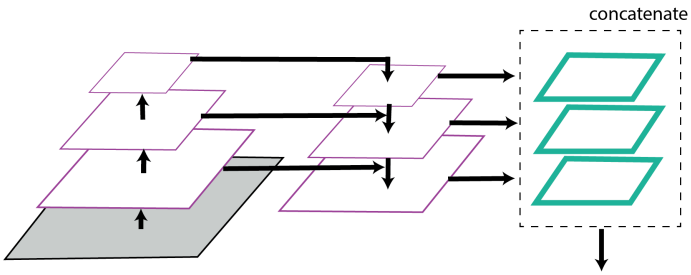


Figure 4.3: Schematic of the feature pyramid network.

4.2 Methods

Data Collection and Preprocessing

We collected CdSe micrographs using the TEAM 0.5 aberration corrected microscope at an operating voltage of 300 kV. CdSe particles were synthesized via the WANDA synthesis robot at the Molecular Foundry [26]. This solution was then diluted with hexane and dropcast on 200 mesh ultra-thin carbon grids. Forty six 1024×1024 images were taken by hand. Labels for this dataset were manually created using the MATLAB labeler app [106]. The CdSe particles being smaller and a lower atomic number material provided a lower contrast and, therefore, lower signal to background dataset.

Au nanoparticle data was collected using an FEI Themis with image aberration correction operated at 300 kV. Au nanoparticles in phosphate buffered saline were purchased from Sigma Aldrich. The Au nanoparticle solution was diluted with water and dropcast on to 200 mesh carbon grids. Prior to drop casting, carbon grids were cleaned using a plasma cleaner. Thirteen 4026×4026 TEM micrographs were collected using SerialEM. SerialEM is a software which enables automated electron microscopy imaging [30], and was used for automated focusing in order to speed up image collection. The autofocusing function also helped ensure that the range of defocus was within approximately 100 nm of Gaussian focus.

Labels for network training were created by hand using the MATLAB labeler application [106]. All the micrographs were then normalized by first applying a median filter with a 3×3 kernel which served to remove spurious X-rays. After filtering, the images were normalized by the group maximum.

The training set consisted of 129 images. Images in the training set were rotated and flipped to augment the dataset which were then split 4:1 into the training and validation sets.

Network Architectures Implementation and Training

We implemented the variations of the U-Net architecture using the segmentation models package [126]. The three variations changed the layers in the contracting path, also known as the backbone of the the U-Net. The three backbones implemented were the VGG16 [121], ResNet-18 [122], and MobileNet [123] architectures. Each architecture was separately trained three times in order to understand the average performance. Each training consisted of 20 epochs with 100 steps per epoch. The model used a combined categorical cross-entropy and dice coefficient loss function and Adam as the optimizer with a learning rate of 1×10^{-3} ([108, 109]).

We also implemented the feature pyramid network [125] using the segmentation models package [126]. It was implemented using a ResNet18 backbone. The feature pyramid network was trained with the same data and loss function as the modified U-Nets but with Adadelta as the optimizer with a learning rate of 1×10^{-3}. The atrous network was implemented using Keras [107] and was based on the architecture presented by Chen *et al.* [124]. The atrous network was trained with the same data but using categorical crossentropy loss and Adam as the optimizer with a learning rate of 5×10^{-5}. All hyperparameter changes were made to optimize performance of the different architectures. All training was done using a GPU through Google Colaboratory.

4.3 Results

	Dice Coefficient	Intersection over Union
U-Net VGG16	0.74	0.61
U-Net ResNet-18	0.82	0.71
U-Net MobileNet	0.80	0.69
Atrous Net	0.80	0.67
FPN	0.82	0.72

Table 4.1: Dice coefficient and intersection over union for each of the architectures on the hold out test set.

The results for the various architectures is shown in Table 4.1. Many of the architectures perform fairly well, but the VGG16-based U-Net clearly performs the worst. Of the U-Net based architectures the ResNet-18 performs the best. MobileNet achieves performance close to that of the ResNet-18. Sample segmentation maps from the three U-Net based architectures are shown compared to the ground truth map overlaid on the original image in Figure

4.4. While the ResNet-18 achieves the highest metrics, it is clear from a visual inspection that the MobileNet does better at recovering the small particle regions. In the example shown the MobileNet based U-Net actually does better at identifying all potential particle regions than ResNet-18 based U-Net.

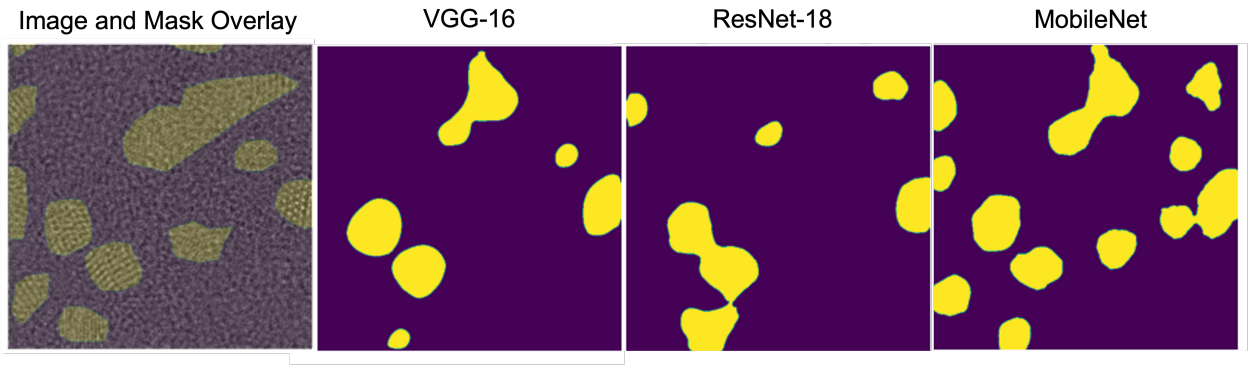


Figure 4.4: Example segmentation maps predicted by each of the U-Net based networks.

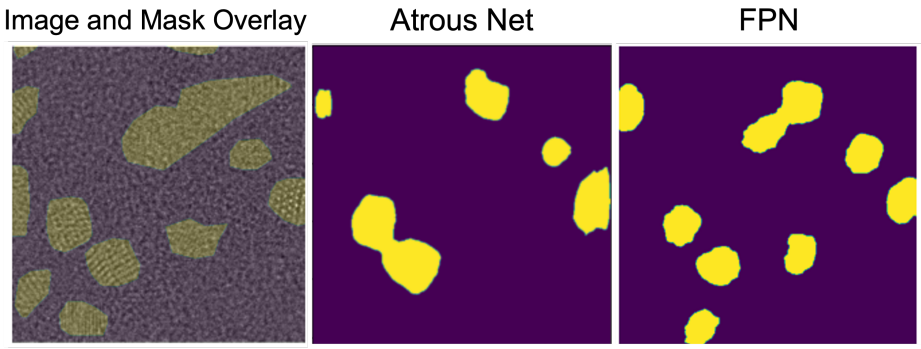


Figure 4.5: Example segmentation maps predicted by the FPN and atrous networks.

Of the non-U-Net based architectures the FPN performs the best. It also achieves the best performance of all the architectures, slightly outperforming the ResNet-18 based U-Net. The atrous net also does reasonably well and performs similarly to the MobileNet based U-Net. Analyzing the sample segmentation maps from the atrous and FPN also lead to interesting observations. Atrous net seems to miss several regions. FPN, by contrast, can be seen to find regions which were missed by the human labeler.

Figure 4.6: Example segmentation map predicted where FPN outperforms human made ground truth segmentation map. a) Human generated segmentation map overlaid micrograph. b) FPN generated segmentation map. c) Yellow box indicates the particle region which FPN successfully identified that the human labeler missed.

4.4 Discussion

What becomes clear from these results is that improvements can be made in HRTEM segmentation by moving away from the original, VGG16, based architecture. ResNet-18 provides the best performance of all the U-Net architectures. It's performance is closely followed by that of the MobileNet based U-Net and may outperform MobileNet simply because it has almost twice the number of trainable parameters. However, the ResNet-18 based U-Net also contains residual units (essentially skips in the contracting network) that provides high resolution information deeper into the network that the MobileNet does not. This would suggest that ResNet would perform better than MobileNet on the small particle regions which from the sample segmentation maps in Figure 4.4, is not necessarily the case. This may suggest that the increase in parameters, and not the architecture changes, are the reason for the improved performance over MobileNet. MobileNet does achieve very good results relative to the number of parameters. It is impossible to say, based on these results, why MobileNet does so well but it should be noted that both ResNet-18 and MobileNet eliminate all but the first max-pooling step and perform well, whereas the VGG16 contains a max-pooling layer after every two convolutional layers. This may suggest that higher performance on HRTEM data can be achieved with less max-pooling. The ability of MobileNet based U-Net to find smaller particle regions also raises the question as to whether using 1×1 convolutions like the MobileNet backbone uses as opposed to the 3×3 size convolutions makes the network more sensitive. It is likely that this is dependent on the pixelsize and future work should look into this dependency on the physics of the system.

In terms of non-U-Net based architectures the atrous net does not appear to provide any significant advantages. It is a bit surprising that the atrous net should perform similarly to U-Nets since the sampling of the image is so different from these architectures. FPN performs the best of all architectures analyzed. The FPN combines representations from

different spatial resolutions, since the HRTEM data has information at many length scales maybe the improved performance is related to this sampling. Especially in an image where spatial frequencies are so important it would make sense that an architecture would need to predict based on sampling many different parts of Fourier space.

It should also be noted that the FPN has been found to outperform the labels, such as in the sample image shown in Figure 4.6, where it can be seen that it finds a particle region that the human labeler missed. With this example in mind and since all the networks seem to be approaching similar metrics the question becomes whether the networks are held back by error in labeling. In this case it is unclear since the best we can do is the human labeled data. However, we will analyze the ability of networks to handle label error in the next chapter and future work using computationally generated HRTEM images may be able to help elucidate on this point for these specific architectures.

4.5 Conclusions

We have demonstrated that significant improvements can be made over the traditional U-Net architecture for the segmentation of HRTEM images. We have found that feature pyramid networks provide the most promising results for HRTEM segmentation. We have found evidence that this architecture actually outperformed the human segmentation maps. Future work will need to illuminate the exact mechanism which explains this improved performance.

Chapter 5

Removing Labeling Requirements for Machine Learning Based Nanoparticle Analysis[1]

5.1 Introduction

In Chapter 3 we have shown that diversifying datasets can lead to improved performance across micrographs. This suggests, as is well known, that larger electron microscopy datasets would provide better neural network analysis. Often the limiting step in creating larger datasets is in the manual labeling. Automated labeling provides an obvious solution but is fraught with its own challenges. We have also found in Chapter 4 that analysis of network performance can be limited by label error in the dataset. Chiral nanomaterials provide an interesting test bed for analyzing automated labeling systems and how to handle label error.

There is growing interest in inorganic chiral nanomaterials for application in optoelectronics and biomimetics [127, 128, 129]. Specific parameters during wet chemical synthesis of chiral nanomaterials [130, 13, 131, 132, 133] can induce a large degree of structural variety. Of particular importance, synthesis parameters can favor one handedness over another. For example, despite their underlying chiral crystal structure, Tellurium (Te) nanoparticles can have different ratios of certain chiralities depending on synthesis conditions [130, 13]. These variations are induced by many factors including thermodynamic versus kinetic growth pathways, and differences in interactions of chiral organic molecules with small clusters of atoms during synthesis. Therefore, precise tuning of chirality alongside size via wet chemical synthesis is yet to be obtained in many systems. To do this, one must first determine the influence of the many synthetic parameters, such as temperature, precursor concentration, or concentration and type of structure-directing chiral ligands, on the outcome population.

[1]This chapter is a version of C. Groschner, A.J. Pattison, A. Ben-Moshe, A.P. Alivisatos, W. Theis, and M.C. Scott. "Classifying Handedness in Chiral Nanomaterials Using Label Label Error-Robust Deep Learning". Submitted.

This motivates the development of methods to classify handedness in chiral nanoparticle populations with the goal of determining the influence of these synthetic parameters.

While circular dichroism (CD) measurements are sensitive to chirality in Te nanoparticles, it is very challenging to extract quantitative information about the abundance of each handedness, as the molar CD of these materials is unknown and can only be estimated [130, 13]. Scanning electron microscopy (SEM), in contrast, can be used to unambiguously determine the handedness of morphologically chiral nanomaterials [134, 135, 136]. SEM is sensitive to surface topology and can directly determine morphological chirality and handedness, unlike (scanning) transmission electron microscopy ((S)TEM) methods which sums information along the beam direction such that faceting information can be lost and therefore requires that multiple images be used to determine handedness [137]. High-throughput SEM imaging is therefore a particularly promising way to measure the size and handedness of large populations of chiral nanoparticles to better understand the role of synthetic variables on outcome populations. However, determining particle statistics by hand from high-throughput data is extremely laborious and time consuming. Due to the increasing ease of implementing neural networks for image analysis [138, 107, 139], deep learning is a promising replacement for manual analysis. Yet, deep learning is known for requiring large training datasets, which still need manual labeling, meaning that the application of deep learning to chirality studies could also be prohibitively time consuming due to expert labeling requirements.

Given that synthesis routes yielding chiral materials often favor one handedness over the other, we have found that it is possible to label the handedness of all the particles in the dataset by first labeling them with the dominant handedness, and then mirroring these images to create a dataset labeled with the opposite handedness. This process is demonstrated in Figure 5.1). Labeling all images with the majority handedness of course leads to a dataset with a specific fraction of erroneous labels that is equal to the fraction of particles which did not have the dominant handedness. For synthesis conditions which yield almost exclusively one handedness, this automated mirror labeling strategy is very successful, but for other conditions this can yield a significant number of mislabeled images. The question then becomes how to extend this automated labeling method across synthesis conditions, such that the accuracy of networks is not hampered by mislabeled data.

Generating accurate models from erroneously labeled datasets, also known as noisy datasets in the machine learning community , has been an expanding area of research. Deep learning is known to be able to memorize random inputs during training [140], and thus has the potential to memorize noise. What makes true learning, and thus generalizability, possible is that before memorizing real data, repeated motifs within the data are learned by the network [141]. This means that networks can recognize predictive features without directly memorizing inputs. Deep neural networks are thus able to learn the true signal from very noisy datasets [142]. To further extend the label error tolerance of deep neural networks several strategies have been implemented. These methods can be broadly separated into four types of approaches: 1) noise robust techniques [143, 142, 144, 143], 2) label cleaning methods [145, 146], 3) noise estimation [147, 148], and 4) preprocessing pipelines [149]. This paper will focus on noise robust and label cleaning strategies. Previous research into these

methods focus on how these methods perform on very large datasets of real images with a large number of classes. Microscopy data, in contrast, contains fewer classes but also much smaller datasets than standard computer vision datasets.

In this work, we focus on noise-robust and label cleaning methods that do not require the creation of a large labeled dataset or other supervision. First, we will examine the performance of a standard convolutional neural network when trained for a binary classification task on a small, idealized SEM dataset. The architecture for the network is shown in Figure 5.2a. It consists of two convolutional and max pooling blocks with a final dropout and dense unit . Previous work has shown that the noise type often influences network resilience to noise, but these studies have focused on multi-class data with isolated random or label flip noise [142, 146, 149, 150]. This SEM dataset has both label flip and feature-dependent noise, the label noise being correlated with image features. For example the minority enantiomer may be generally smaller particles, providing an image feature which is correlated with label noise, but cannot be directly controlled, which adds complexity to our task. For our experiments, we create an idealized dataset to limit these features that are noise correlated. To create the idealized dataset we take a manually, accurately labeled dataset and mirror a certain fraction of segmented particles to create the specified error fraction as is described in Figure 5.1 so that we can model different possible synthesis outcomes. We first seek to understand how a standard convolutional neural network handles this label noise, then compare its performance to more complex noise handling strategies. As part of this comparison we explore using the co-teaching training procedure developed by Han *et al.* [146]. The co-teaching method uses the difference in loss between two networks trained in tandem on the noisy dataset to clean data and then further train its neighbor model. For consistency, the two tandem networks used in our study of co-teaching consist of the same architecture as the standard CNN we tested, as shown in Figure 5.2a. A schematic of the co-teaching training procedure is provided in Figure 5.2c. Finally, we propose a new method inspired by noise estimation and curriculum learning [151, 152, 150, 153]. In this method, we use a very small clean dataset to start network learning before training on the noisy dataset. The neural network architecture is the same as the standard CNN shown in Figure 5.2a and the training procedure is illustrated in Figure 5.2b. . This new method leverages the memorization effect, in order to push networks toward learning relevant features even at high error rates [154] while avoiding data loss by cleaning. From these experiments, we show that, for small datasets, the label cleaning and curriculum based strategies achieve similar accuracy rates, but that curriculum style training leads to the most consistent results across noise rates. We then use as-synthesized micrographs of left and right handed particles to explore the way in which real data containing feature correlated noise then impacts the performance of the three aforementioned methods. We find that once feature correlated nouse is added only the co-teaching method maintains high accuracy. We hope the exploration of these methods will enable wider application of weak labeling and therefore deep learning for the chiral materials community and the materials community more broadly.

5.2 Methods

Dataset Generation

Datasets of known label error (known percentages of mislabeled images) were generated by starting with a clean dataset of right handed images. Particles were segmented by hand using MATLAB data labeler software prior to dataset generation. A set percentage (varying from 0 to 50%) of the images were randomly selected and mirrored, as shown in Figure 5.1. All images in the dataset were then labeled as having right handedness. Then, all the images were mirrored, creating a dataset in which all images are labeled as left handed. This procedure ensures that the specific fraction would always be facing the opposite direction from its label. All image augmentation for training was performed after this operation. We applied this method to two datasets. The first was the number 7 samples from the Keras implementation of the MNIST data set. We chose to use the number 7 since it is a chiral shape. The second dataset consisted of right handed Te chiral nanoparticles which were manually segmented and labeled by an expert. The complete dataset consists of 1,914 right handed particles which are mirrored to create a set of 3,828 total left and right handed nanoparticle images. 80% of the data is used in the training set, 10% in the validation set, and 10% in the test set. During training the images are augmented using a combination of 180° rotations, 5° rotations, up to 30% zoom of the image, and 10% shearing. The augmentations are used to create a dataset of 91,840 images per epoch (training round) for training Keras networks, and 97,984 images for PyTorch networks. The difference in number of images is due to differences in augmentation implementation between Keras and PyTorch.

Standard Convolutional Neural Network Classifier

The neural networks were developed with Keras. A simple convolutional neural network was implemented for both the MNIST and Te nanoparticle datasets; which contained two convolutional residual units, a schematic is shown in Figure 5.2a. Each residual unit contained the convolutional layer, a ReLu layer, and a max-pooling layer. After the residual units, the features are flattened, and passed to a dense layer, dropout, and final dense layer. The logits from the dense layer are then passed to a softmax activation function. Training was done with a categorical crossentropy loss function, and adadelta optimizer. The learning rate was 0.001, the batch size was 32, and ran for 100 epochs. Model check pointing was used so that model weights were only saved if the validation loss had decreased.

Two Step Training Classifier

The network implemented for two step training was also developed with Keras and employed the same neural network architecture as the standard CNN. A schematic is shown in Figure 5.2b. The training of the network consists of two distinct steps. First, the network is trained on an extremely small set of images with error free labels. For this first step we used 10

left handed particle images and 10 right handed particle images, which are augmented to represent 288 samples per epoch. The network is trained on the tiny, clean training set for ten epochs. The network is then trained in the same way as the standard convolutional neural network for 100 epochs.

Co-teaching Classifier

Networks were developed using PyTorch. The code and procedure was adapted from work done by Han *et al.* [146]. The co-teaching method creates two identical CNNs. Two separate batches of data are given on each network for training. The data whose loss is below a certain threshold then gets passed to the other network for training and data above is removed [146]. A schematic of the procedure is shown in Figure 5.2c. In this way data is cleaned, under the assumption that noisy labels will lead to higher loss. The architecture of the two CNNs was changed to match the CNN architecture used in the Keras-based handedness classifier. The only difference was the addition of an average pooling layer to compensate for differences in layer implementations. The models were trained on the same dataset. The learning rate was 0.001, the batch size was 128, and the maximum number of epochs was 40.

Image Acquisition

Tellurium nanoparticles suspended in an aqueous solution were dropcast onto a silicon wafer. Micrographs were acquired on FEI Helios G4 UX at 2kV using a through-lens detectors with a working distance of 2 mm. Images were collected using Maps 2.5 Software, which collected a grid of approximately 1200 images of collections of nanoparticles (approximately 1400-by-900 pixels each). This large-scale acquisition was stitched together using the same Maps 2.5 software. Nanoparticles were manually segmented from the larger images to make the small input images for the neural networks. Sample images are provided in the supplementary materials.

Synthesis of Te Nanoparticles

The synthesis follows a methodology developed by Ben-Moshe *et al.* [130], with some modifications. 5.5 ml water, 15 mg TeO2 and 20 µl NaOH (1M in H2O) were stirred vigorously in a 20 ml glass vial (at room temperature), before 2.5 ml of hydrazine hydrate (80% solution) were added in one go. 25 seconds after the addition of hydrazine, 1 ml of a 100 mM solution of D-penicillamine (adjusted to pH 11 using NaOH solution) was added in one go. The reaction was stopped after three hours, by diluting twice with a 100 mM SDS solution, followed by repeated cycles of cleaning using centrifugation (6000 RPM, 10 minutes) and dispersion in water.

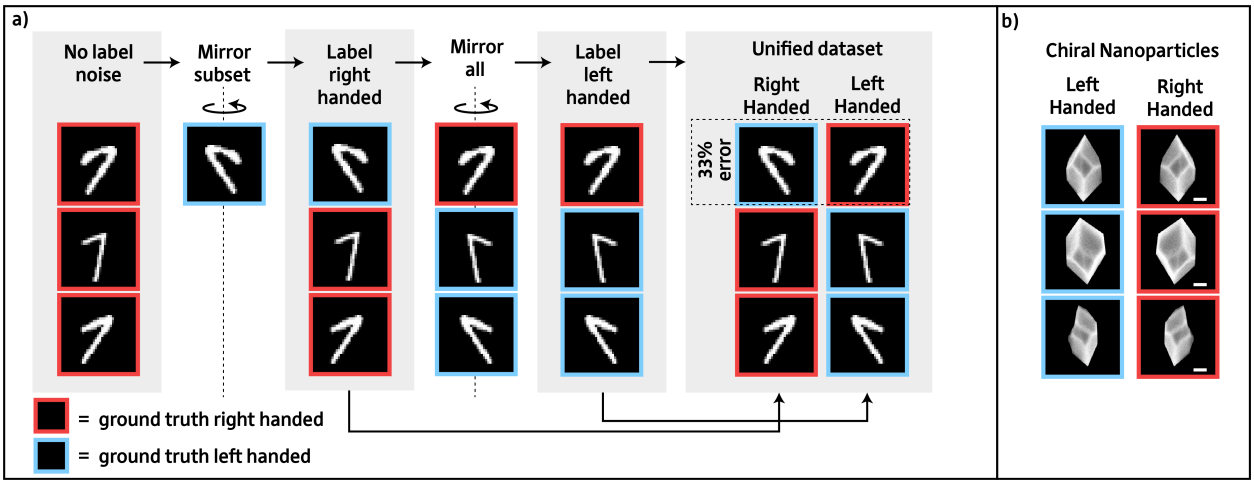


Figure 5.1: The mirror labeling scheme used for data labeling with sample images. a) Schematic demonstrating the mirror labeling procedure by which training datasets are automatically labeled and the error rate in the dataset is controlled on the MNIST seven dataset. b) Examples of images for both classes from the chiral nanoparticle datasets. Scale bars, 100nm.

5.3 Results

Analysis Methods

For each technique we assessed the networks' performance on a clean ground truth test set with 50% left and right handed images after being trained on noisy datasets with different label error fractions. We also analyzed the distribution of error between the two classes. We did this by plotting the histogram of incorrectly labeled left handed particles and number of incorrectly labeled right handed particles in the test dataset. Dramatically misclassifying one but not the other label could lead to a significant skew in the calculated ratio between classes, a metric which is very important in studying chirality. We finally evaluate the ability of the networks to recover the fraction of particles belonging to the minority population, to simulate the use of these methods for data from real synthesis procedures. For this analysis we use a test set that has the same mislabel fraction as the training set. Using this second test set, we plot $f(x) = n_{\text{left}}(x)/n_{\text{total}}$, where x is the mislabeled fraction of the training set, $n_{\text{left}}(x)$ is the total number of particles labeled left handed by the network for that test set, and n_{total} is the total number of particles in the test set. All networks were trained on noisy datasets. The noise generation process is outlined in the methods section and illustrated in Figure 5.1a. For each method we trained three separate networks using that method and plot the average performance metric and the standard error.

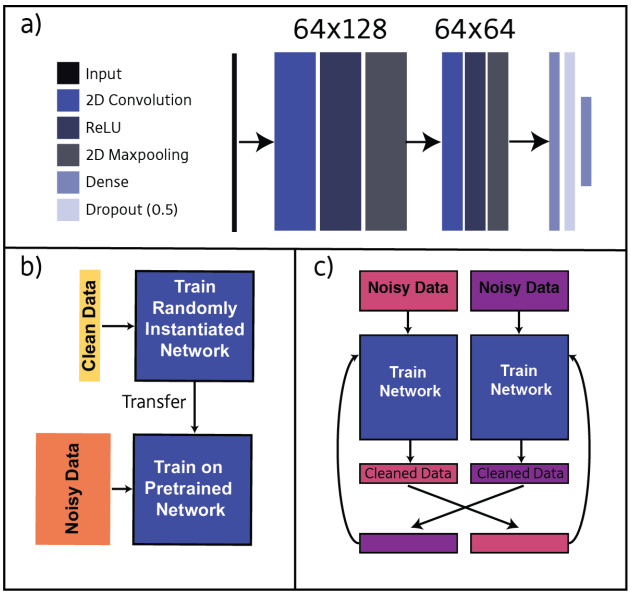


Figure 5.2: Schematics for the three methods studied. All networks use the architecture shown in part a. a) Schematic of convolutional neural network architecture implemented with features by height and width labeled. b) Schematic of two-step training procedure. Randomly instantiated network is trained for one training round, i.e. epoch, using clean data before network weights are transferred and trained using noisy dataset. c) Schematic of co-teaching training procedure. Two networks are trained in parallel and after each training round the data is cleaned by only passing data with low enough loss to the twin network.

MNIST Sevens

Standard Convolutional Neural Network

Before testing our three noise-robust neural network approaches, we first tested a standard convolutional neural net's ability to learn to classify chiral images in the presence of varying amounts of label noise using images of the number seven from the MNIST dataset. The training set for the model contained 10024 images and the test set 1258 images. The architecture used is shown in Figure 5.2a. Figure 5.3a shows that a standard convolutional neural network can learn the classes accurately up to 42% label error. Figure 5.3b shows the percent of misclassified particles which were left handed vs right handed demonstrating that there is minimal bias, in terms of handedness, in the classification error as well. Figure 5.3c shows that, given an enantiomerically skewed population, the neural network for that amount of label noise is able to correctly predict the fraction of sevens which have that minority handedness. All these metrics suggest that the mirror labeling procedure works and even a standard CNN can handle the label noise presented.

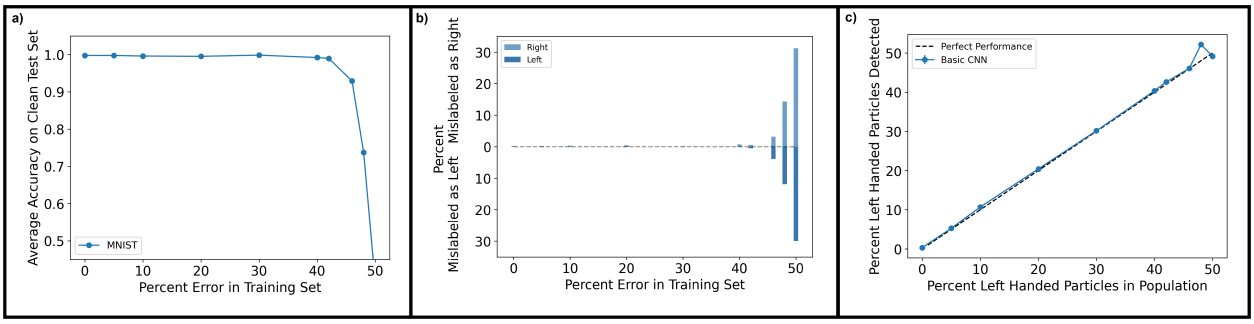


Figure 5.3: Figures quantifying the performance of the standard CNN on the MNIST training sets. a) Accuracy of standard CNN on the MNIST seven clean ground truth test set given increasing error in the training set. b) The percent of sevens erroneously labeled as left or right handed in the test set. c) Predicted fraction of left handed sevens for a population where label noise is specified by the percent of left handed sevens using the network which corresponds to training with that percent of label noise .

Te Chiral Nanoparticles

Standard Convolutional Neural Network

We then applied the same mirror labeling and training protocol from the MNIST dataset to the SEM images of chiral Te nanoparticles. Sample images of the chiral particles are shown in Figure 5.1b. The training set for the chiral nanoparticle models contained 3062 images, before augmentation, and the test set 382 images. Further information on augmentation is presented in the Methods section. Augmentation was used to ensure that handedness was not correlated with arbitrary images features, e.g. if right handed particles tended to have highest intensity on the right hand side of the image, random 90 degree rotations were implemented during augmentation to destroy this arbitrary correlation and ensure networks learn on truly predictive features. For the standard CNN on SEM images, the clean test set accuracy begins to degrade after 10% label error, as can be seen in Figure 5.4a. Reasonable accuracy is still recovered up to 30% label error. In order to examine the maximum amount of noise that can be handled we increased our sampling close to 50% flip error - close to random labeling. Figure 5.4b gives the percent of misclassified particles which were left handed vs right handed. This shows that the network on average is rather inconsistent in whether it misclassifies one handedness more than the other at high label error. Figure 5.4c shows the ability of the network, for an enantiomerically skewed population, to recover the correct fraction of the mislabeled portion of the population. The standard CNN is able to accurately predict the fraction of the minority population (and therefore mislabeled in the training set) up to 40% label error.

Co-teaching

Accuracy results for different label error rates using the co-teaching training method are shown in Figure 5.4a. We see that at very low label error the standard CNN achieves higher accuracy than co-teaching. However, the co-teaching method is able to achieve much better accuracy than the standard CNN when the label error increases past 10%. One benefit of the co-teaching network is that from 0% to approximately 40% label error the recovered accuracy is very consistent. One promise of the co-teaching network is that datasets with large amounts of label error can still be used for accurate training. However, in Figure 5.4a we see that there are still limits on the amount of label error which can be tolerated. We still observe a dramatic reduction in accuracy beyond 40% label error. As shown in Figure 5.4b, the co-teaching network does not misclassify one handedness more than the other at low label error rates, but at label error rates above 42% the network appears to be significantly over predicting left handed particles. This skew in misclassification is also reflected in the fraction of left handed particles detected. Figure 5.4c shows that the co-teaching networks on average, badly underestimate the number of minority handed particles past 42% label error.

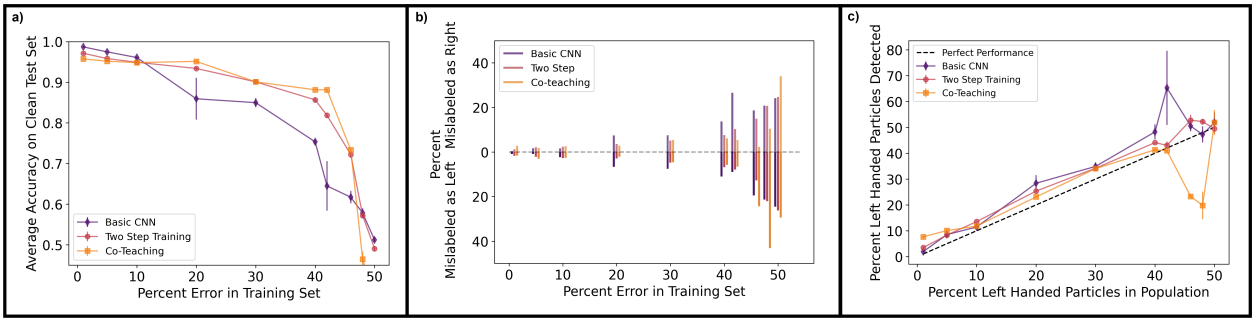


Figure 5.4: Figures quantifying the performance of the three methods on the Te training sets. a) Average accuracy of networks from each technique on the clean ground truth test set of Te SEM data given increasing label error in the Te training set. b) The percent of right and left handed particles which are misclassified by each network c) Predicted fraction of left handed Te particles for a population with a specified percent of left handed particles using the network which corresponds to training with that percent of left handed particles.

Two Step Training

Similar to co-teaching two step training creates a network which performs well up to about 40% label error. Co-teaching and the standard CNN sometimes slightly outperform the two step method in terms of accuracy but the two step training creates perhaps the most consistent results. This is reflected in the fact that this training procedure leads to the most consistent detection of minority particles. Unlike the other methods, it does not drastically

over or under estimate the population of left handed particles at any point, as seen in Figure 5.4c. This is a consequence of the fact that the training procedure does not lead to a high bias against one class or the other at any training label error, as can be seen in 5.4b.

Application to As-Synthesized Particles

We then went beyond the ideal and applied the methods discussed to a collection of as-synthesized Te particles. The difference between the as-synthesized and ideal case is that in the ideal case we start with all right handed particles so particles all "left handed" particles are mirror images of right handed particles. The as-synthesized dataset is made up of particle images from a sample that contained 22% left handed and 78% right handed particles. The preparation of these particles can be found elsewhere [13]. We created a training set of 3542 particle images. The labeling procedure was then used on these particles (without the error creation step since left handed particles are now part of the sample). The results are shown in Table 5.1. The performance of these networks does not match the ideal dataset. We see that the standard CNN and two step training procedure fall short of acceptable accuracy. Only the co-teaching method maintains high accuracy.

	Average Accuracy As-Synthesized
Standard CNN	0.65 ± 0.049
Two Step Training	0.66 ± 0.002
Co-teaching	0.91 ± 0.002

Table 5.1: The average accuracy and standard error on the clean, as-synthesized balanced test set for each of the three techniques.

5.4 Discussion

Performance comparison between MNIST and nanoparticle trained neural nets

A comparison of the accuracy on the clean test set indicates that the application of a CNN to the MNIST data results in higher accuracy across most noise levels than the results from training with the chiral nanoparticle dataset. The accuracy of the network compared to the chiral nanoparticle dataset suggests that the network is fitting a simpler feature space than the chiral nanoparticle case. It should be noted that the MNIST input image is much smaller than the Te, meaning that by keeping the network size the same we have more convolutional kernels relative to the number of pixels in the MNIST network. In addition, to the human

observer, it is easy to see that there are only a few prototypical examples for a number seven in the MNIST dataset and few deviations from these prototypes. By inspecting data in Figure 5.1a and 5.1b, it is clear that while most parts of the seven images contribute to the chiral shape, the chiral particles have many features which are unrelated to the chirality of the particle. In fact, only the center facets of the chiral particles determine a particle's handedness. Therefore, since we have held all hyperparameters constant while training on each dataset, it is reasonable to interpret at least part of the difference in accuracy is due to the difference in richness of the feature space. Furthermore, learning a more complex feature space with fewer features actually contributing to the classification task is likely the reason for the lower accuracy of CNN trained on the chiral nanoparticle dataset.

The limited feature space of the MNIST dataset vs the chiral nanoparticles was by design, for the purpose of comparing label noise robustness of the network when applied to datasets with and without "hard examples", those samples which are harder for the network to learn. This was important to consider since previous work has found evidence that CNNs treat hard examples and noise in a similar fashion [155, 140]. By proving the noise robust nature of our proposed standard CNN on the MNIST dataset , we could isolate any challenges when applying the same mirror labeling scheme to actual chiral materials which have many more possible representations. We hypothesize that this increased feature space is responsible for the lower performance of the networks on chiral nanoparticle data, particularly the standard CNN, in the presence of high label error compared to the network trained on MNIST. This highlights that noise robust properties of standard neural networks are heavily dependent on the richness of the feature space.

Comparison of noise-robust architectures and training procedures on ideal SEM data

Our results on the ideal SEM dataset suggests that the best method to employ depends on label error and the type of error tolerated. In general two step training and co-teaching show enhanced robustness to label error, as they achieve a high classification accuracy, until approximately 40% label noise and above.

The standard CNN does not demonstrate the same label error robustness as the other two methods. This deviates from the results of the standard CNN on the MNIST sevens and suggests that for the more complex feature space presented by the Te dataset, the noise robustness of a standard CNN is not as consistent.

Co-teaching provides a way to explore label cleaning methods. For very low label error the co-teaching achieves the lowest accuracy, which is likely connected to throwing away hard examples during training [155, 140]. However, as training error increases co-teaching becomes the most accurate training method. However, once training label error is too high the accuracy of the co-teaching method rapidly deteriorates and shows strong skew in which handedness is misclassified, unlike the two-step method which declines in accuracy but does not show an increase in misclassification skew. It is likely that since the co-teaching method

includes no prior knowledge about left and right handed representations, past a certain label error rate it is no longer able to distinguish the label error from variations in handedness, therefore causing its performance to plummet.

The two step method achieves almost the same level of accuracy as the standard CNN at low training label error but is also able to sustain high levels of accuracy with more than 40% label error like the co-teaching method. This is interesting because we do not use this data to infer any extra information or weighting schemes when training on data with label error unlike some other pretraining methods [153, 151]. To our knowledge, other work in this field has used small supervised training sets to train label error estimation layers and other corrective measures but not as a transfer learning procedure to initialize the network. The results we present, therefore, suggest that benefits can also be reaped by utilizing curated, tiny, manually labeled datasets to help the network learn the correct relationships between key features. We hypothesize that this method is exploiting the fact that networks are prone, in early training rounds, to learn simple features as opposed to memorizing data [154, 141]. While we rely on the early training rounds learning easy representations from the tiny dataset to direct learning class features correctly, it is likely that by initializing the network with a very small dataset, we are limiting representations the network is sensitive to. This points to an inherent tradeoff in this method between labeling requirements and generalization, which should be explored in future work. This may also explain why we achieve slightly lower accuracy when training label error is low.

The difference in the misclassification bias between the co-teaching and two step method is what truly distinguishes these two techniques in this application. Previous work has shown that pre-training a network with an unsupervised dataset leads to better generalization due to the pre-training acting as regularizer [156]. Though we use a supervised training set for pre-training, the difference in misclassifcation bias between the two networks supports that particularly at high levels of training noise, the first training round is constraining the network to more balanced learning. This is a key finding for the application proposed since being able to consistently recover handedness ratios is vital for understanding the influence of synthesis parameters on the development material handedness.

Comparison of Ideal to As-Synthesized Dataset

The difference in performance of the three methods on the ideal dataset versus the as-synthesized dataset highlights the importance of the alternate features present between the as-synthesized left and right handed particles. Analyzing the as-synthesized images (shown in Appendix 3, Figure 3), it is evident that the growth which led to minority left handed particles in the sample also caused those left handed particles to have a smaller chiral facet. The particles are also are smaller overall. Therefore in the as-synthesized case we not only have a complex feature space to learn (as was seen in the ideal vs MNIST comparison) but we also have image features that are directly correlated with the label error. This feature correlated label error clearly has important implications for the methods explored. It seems that the chiral facet is crucial because resizing all right and left handed particles to be the

same size did not improve performance. This result is surprising since the training was implemented with augmentations to enlarge and shrink the particles so that the chiral facet would take on a number of sizes for each handedness. In essence this should have eliminated any size dependence. The fact that augmentation did not enable learning in the standard CNN and two-step training method suggests that the relationship between the rest of the particle and the chiral facet also has a relation to handedness. It should also be noted that we did not augment to control for contrast variations when testing with this dataset so future work should also consider the impact of this factor. Overall, these results suggest that when the noise has alternative relationships beyond the mirror operation that co-teaching is the best method. Future work will need to be cognizant of these differences between ideal datasets and their actual application.

Implications for experimental applications

The as-synthesized analysis shows that for most experimental applications the co-teaching method is the best approach. The as-synthesized data set had 22% minority handedness present corresponding to 22% label error. This is a relatively high label error and suggests that the co-teaching method would be able to handle experimental conditions [13]. The co-teaching method should, therefore, be used anytime that it cannot be confirmed whether the synthesis conditions lead to variations in size (or other particle features).

It should be noted that it is conceivable that there could be experimental conditions which may resemble the ideal dataset. If the samples are true enantiomers and therefore only varying by a mirror operation, as in the ideal dataset, then there are two clear use cases that require choosing between these networks for practical applications of the red mirror black labeling approach in the chiral nanomaterial community. For synthesis routes where one handedness is much more favored than the other, training a standard CNN is most likely to yield the most accurate population statistics because under this condition our mirror labeling technique will yield low label error . For chirality studies where synthesis conditions lead to a more balanced set of left and right handed particles, and therefore higher label error, the two step neural network is the best option since this network provides accurate results and little bias in class prediction across populations.

There are several future use cases for the three methods developed here. The automated mirror labeling system described can be used not only in the realm of SEM chiral image analysis, but for any imaging method where the chiral structure leads to geometrically related images. Outside of chirality classification, the label-error robust methods developed have a wide range of possible use cases across the microscopy community. While the characteristics of the label error will vary on a case by case basis, this work provides a starting point for the use of label error robust methods for a variety of microscopy tasks.

In summary, these results give important insight for leveraging automatic labeling systems such as ours on other microscopy datasets. We see automated labeling as a potential avenue for analysis in the chirality community but also among other materials where datasets are too large to realistically be manually labeled. We have shown the tradeoffs between the

richness of the feature space and label error robustness of CNNs, which should be taken into account when using these techniques for other studies.

5.5 Conclusion

We have shown that both label error robust and label cleaning methods perform well for small datasets. However, we find that both two step training and co-teaching far outperform the inherent label error robustness of a standard CNN. These are important findings for deep learning studies synthesis of chiral materials, and also extend to a wide array of binary classification problems in the materials community. By demonstrating conditions under which accurate classifiers can be created without manual labeling we hope to expand the usefulness of deep learning for materials characterization.

Code Availability

Code for this study is available at `https://doi.org/10.5281/zenodo.5548317`.

Data Availability

All image and label data that support the findings of this study are available on Zenodo: `https://doi.org/10.5281/zenodo.5009042`.

Chapter 6

Conclusions

6.1 Summary

In this dissertation we have shown the large impact synthetic parameters can play on the development of defects and the type of tools needed to understand this defect development. To this end, in Chapter 2 we have analyzed the formation of defects during non-classical growth modes in palladium nanoparticles. We have shown a simultaneous twinning behavior that can convert a decahedral particle into an incosahedral particle. We are limited in our study of multiply twinned palladium nanoparticles by the inability to characterize a large number of particles and therefore understand exactly which conditions contributed to our observation of simultaneous twinning. To address these issues we have proposed to develop a high throughput analysis method for HRTEM micrographs of nanoparticles. We have presented the results of these efforts in Chapter 3. In Chapter 3 we demonstrate the ability to use a combination of deep learning and random forest classifiers to segment out nanoparticles from HRTEM micrographs and then use a random forest classifier to classify whether or not the nanoparticle contains a stacking fault. In developing this pipeline we discovered the need to have nanoparticles made of different materials in the dataset due to the different image features different materials create in the TEM. We then also explored ways to improve segmentation based on neural network architecture, as discussed in Chapter 4. This led to the finding that the feature pyramid network showed the highest performance on our mixed material dataset. It also highlighted trouble with manual labeling of data. This highlights a need to find alternative methods to hand data labeling, since this would limit the use of deep learning for other electron microscopy analysis tasks. To address this issue, in Chapter 5, we looked into automated labeling techniques which would recast incorrectly labeled images as a label noise problem. We analyzed three methods for combating label noise: standard convolutional neural networks, co-teaching neural networks, and a two-step training approach based on curriculum learning. We find that different relationships between label noise and image features dictate which method performs the best. If features help predict label noise then co-teaching performs best, while if image feature and label noise

is unrelated the two-step procedure performs best.

6.2 Future Work

Using machine learning and deep learning to probe defects in nanomaterials is still a nascent area of research. Future work will need to further develop benchmarks to understand when networks generalize well to HRTEM data so that architecture and other changes can be easily compared. This is critical since, as we have seen, the data the network is being applied to is critical to evaluation. We demonstrate noise robust methods to allow automated labeling for one analysis task - chirality classification. It would be interesting to expand this work to other tasks such as segmentation. Beyond these directly related tasks to the work presented there are still many components to fully automating the analysis procedure. One notable task which would be very useful in the future is automated zone-axis classification. This would assist researchers to know if the defects being searched for would even be visible in the imaged orientation of the particle. All of these tools would then contribute to studying the development of defects during aliquot studies of nanomaterials. These developments will increase synthetic control of nanomaterials by better understanding heterogeneity across samples, bringing us one step closer to atom by atom material design.

Bibliography

[1] W D Callister and D G Rethwisch. *Materials Science and Engineering: An Introduction*. John Wiley & Sons, 2010. ISBN: 9780470419977.

[2] J. Czochralski. "Ein neues Verfahren zur Messung der Kristallisationsgeschwindigkeit der Metalle". In: *Zeitschrift für Physikalische Chemie* 92U.1 (1918), pp. 219–221. DOI: doi:10.1515/zpch-1918-9212.

[3] G Carneiro et al. *Deep Learning and Data Labeling for Medical Applications: First International Workshop, LABELS 2016, and Second International Workshop, DLMIA 2016, Held in Conjunction with MICCAI 2016, Athens, Greece, October 21, 2016, Proceedings*. Lecture Notes in Computer Science. Springer International Publishing, 2016. ISBN: 9783319469768.

[4] Amretashis Sengupta and Chandan Kumar Sarkar. "Introduction". In: *Introduction to Nano: Basics to Nanoscience and Nanotechnology*. Ed. by Amretashis Sengupta and Chandan Kumar Sarkar. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, pp. 1–6. ISBN: 978-3-662-47314-6. DOI: 10.1007/978-3-662-47314-6_1. URL: https://doi.org/10.1007/978-3-662-47314-6_1.

[5] Noah J. Orfield et al. "Correlation of Atomic Structure and Photoluminescence of the Same Quantum Dot: Pinpointing Surface and Internal Defects That Inhibit Photoluminescence". In: *ACS Nano* 9.1 (Jan. 27, 2015), pp. 831–839. DOI: 10.1021/nn506420w.

[6] Jen-Hsien Wong, Bi-Ru Wu, and Ming-Fa Lin. "Strain Effect on the Electronic Properties of Single Layer and Bilayer Graphene". In: *The Journal of Physical Chemistry C* 116.14 (2012), pp. 8271–8277. ISSN: 1932-7447. DOI: 10.1021/jp300840k. URL: https://doi.org/10.1021/jp300840k.

[7] Hongwen Huang et al. "Understanding of Strain Effects in the Electrochemical Reduction of CO2: Using Pd Nanostructures as an Ideal Platform". In: *Angewandte Chemie International Edition* 56.13 (2017), pp. 3594–3598. DOI: https://doi.org/10.1002/anie.201612617.

[8] D. Spirkoska et al. "Structural and optical properties of high quality zinc-blende/wurtzite GaAs nanowire heterostructures". In: *Phys. Rev. B* 80 (24 2009), p. 245325. DOI: 10.1103/PhysRevB.80.245325.

[9] R. Liu et al. "Luminescence from stacking faults in gallium nitride". In: *Applied Physics Letters* 86.2 (2005), p. 021908. DOI: 10.1063/1.1852085.

[10] B. D. Liu et al. "Local defect-induced red-shift of cathodoluminescence in individual ZnS nanobelts". In: *Nanoscale* 6 (21 2014), pp. 12414–12420. DOI: 10.1039/C4NR04464A.

[11] Auston G. Butterfield, Lucas T. Alameda, and Raymond E. Schaak. "Emergence and Control of Stacking Fault Formation during Nanoparticle Cation Exchange Reactions". In: *Journal of the American Chemical Society* 143.4 (2021). PMID: 33492956, pp. 1779–1783. DOI: 10.1021/jacs.0c13072.

[12] Matthew J. Bierman et al. "Dislocation-Driven Nanowire Growth and Eshelby Twist". In: *Science* 320.5879 (2008), pp. 1060–1063. DOI: 10.1126/science.1157131.

[13] Assaf Ben-Moshe et al. "The chain of chirality transfer in tellurium nanocrystals". In: *Science* 372.6543 (2021), pp. 729–733. ISSN: 0036-8075. DOI: 10.1126/science.abf9645.

[14] Zeger Hens and Jonathan De Roo. "Atomically Precise Nanocrystals". In: *Journal of the American Chemical Society* 142.37 (2020), pp. 15627–15637. DOI: 10.1021/jacs.0c05082.

[15] D. J. Bacon and D. Hull. *Introduction to Dislocations (5th ed.)* Butterworth-Heinemann, 2011.

[16] Sungho Jeon et al. "Reversible disorder-order transitions in atomic crystal nucleation". In: *Science* 371.6528 (2021). DOI: 10.1126/SCIENCE.AAZ7555.

[17] Younan Xia et al. "Seed-Mediated Growth of Colloidal Metal Nanocrystals". In: *Angewandte Chemie International Edition* 56.1 (2017), pp. 60–95. DOI: https://doi.org/10.1002/anie.201604731.

[18] James J. De Yoreo et al. "Crystallization by particle attachment in synthetic, biogenic, and geologic environments". In: *Science* 349.6247 (2015). DOI: 10.1126/SCIENCE.AAA6760.

[19] H. Hofmeister. "Lattice defects in decahedral multiply twinned particles of palladium". In: *Zeitschrift fur Physik D Atoms, Molecules and Clusters* 19.4 (1991), pp. 307–310. ISSN: 1431-5866. DOI: 10.1007/BF01448317.

[20] Byungkwon Lim, Yujie Xiong, and Younan Xia. "A Water-Based Synthesis of Octahedral, Decahedral, and Icosahedral Pd Nanocrystals". In: *Angewandte Chemie International Edition* 46.48 (2007), pp. 9279–9282. DOI: https://doi.org/10.1002/anie.200703755.

[21] Younan Xia et al. "Shape-controlled synthesis of metal nanocrystals: simple chemistry meets complex physics?" In: *Angewandte Chemie (International Ed. in English)* 48.1 (2009), pp. 60–103. ISSN: 1521-3773. DOI: 10.1002/anie.200802248.

[22] Joseph Goldstein et al. *Scanning Electron Microscopy and X-Ray Microanalysis : A Text for Biologists, Materials Scientists, and Geologists*. New York, NY, UNITED STATES: Springer, 1992. ISBN: 9781461304913.

[23] URL: https://www.thermofisher.com/us/en/home/materials-science/learning-center/applications/sem-tem-difference.html.

[24] Brent Fultz and James M. Howe. *Transmission electron microscopy and diffractometry of materials*. Springer., 2013.

[25] David Bernard Williams and C. Barry Carter. *Transmission electron microscopy: a textbook for materials science*. Springer, 2009.

[26] Emory M. Chan et al. "Reproducible, High-Throughput Synthesis of Colloidal Nanocrystals for Optimization in Multidimensional Parameter Space". In: *Nano Letters* 10.5 (May 12, 2010), pp. 1874–1885. DOI: 10.1021/nl100669s.

[27] URL: https://www.thermofisher.com/us/en/home/electron-microscopy/products/software-em-3d-vis/maps-software.html.

[28] David N. Mastronarde. "Automated electron microscope tomography using robust prediction of specimen movements". In: *Journal of Structural Biology* 152.1 (2005), pp. 36–51. ISSN: 1047-8477. DOI: https://doi.org/10.1016/j.jsb.2005.07.007.

[29] Chris Meyer et al. "Nion Swift: Open Source Image Processing Software for Instrument Control, Data Acquisition, Organization, Visualization, and Analysis Using Python." In: *Microscopy and Microanalysis* 25.S2 (2019), pp. 122–123. DOI: 10.1017/S143192761900134X.

[30] Martin Schorb et al. "Software tools for automated transmission electron microscopy". In: *Nature Methods* (May 13, 2019). DOI: 10.1038/s41592-019-0396-9.

[31] James LeBeau, Abinash Kumar, and Matthew Hauwiller. "A Universal Scripting Engine for Transmission Electron Microscopy". In: *Microscopy and Microanalysis* 26.S2 (2020), pp. 2958–2959. DOI: 10.1017/S1431927620023338.

[32] Alex Lin et al. "High-throughput Characterization of CaCO3 Mineralization in Genetically Engineered Organisms". In: *Microscopy and Microanalysis* 27.S1 (2021), pp. 66–67. DOI: 10.1017/S1431927621000842.

[33] Maxim Ziatdinov et al. "Deep Learning of Atomically Resolved Scanning Transmission Electron Microscopy Images: Chemical Identification and Tracking Local Transformations". In: *ACS Nano* 11.12 (Dec. 26, 2017), pp. 12742–12752. DOI: 10.1021/acsnano.7b07504.

[34] Maxim A. Ziatdinov et al. "Atomic mechanisms for the Si atom dynamics in graphene: chemical transformations at the edge and in the bulk". In: 2019.

[35] Catherine Groschner and Mary Scott. "Architecture Optimization and Interpretability in Neural Networks for HRTEM Segmentation". In: *Microscopy and Microanalysis* 26.S2 (2020), pp. 1476–1477. DOI: 10.1017/S1431927620018255.

[36] Catherine K. Groschner, Christina Choi, and Mary C. Scott. "Machine Learning Pipeline for Segmentation and Defect Identification from High-Resolution Transmission Electron Microscopy Data". In: *Microscopy and Microanalysis* 27.3 (2021), pp. 549–556. DOI: 10.1017/S1431927621000386.

[37] Jacob Madsen et al. "A Deep Learning Approach to Identify Local Structures in Atomic-Resolution Transmission Electron Microscopy Images". In: *Advanced Theory and Simulations* 1.8 (Aug. 2018). DOI: 10.1002/adts.201800037.

[38] Chia-Hao Lee et al. "Deep Learning Enabled Strain Mapping of Single-Atom Defects in Two-Dimensional Transition Metal Dichalcogenides with Sub-Picometer Precision". In: *Nano Letters* 20.5 (2020), pp. 3369–3377. DOI: 10.1021/acs.nanolett.0c00269.

[39] Chia-Hao Lee et al. "Deep Learning Enabled Atom-by-Atom Analysis of 2D materials on the Million-Atom Scale". In: *Microscopy and Microanalysis* 27.S1 (2021), pp. 904–906. DOI: 10.1017/S1431927621003482.

[40] Sang-Il Choi et al. "A Comprehensive Study of Formic Acid Oxidation on Palladium Nanocrystals with Different Types of Facets and Twin Defects". In: *ChemCatChem* 7.14 (2015), pp. 2077–2084. ISSN: 1867-3899. DOI: https://doi.org/10.1002/cctc.201500094.

[41] Yifan Li et al. "Structure-Sensitive CO2 Electroreduction to Hydrocarbons on Ultrathin 5-fold Twinned Copper Nanowires". In: *Nano Letters* 17.2 (2017), pp. 1312–1317. ISSN: 1530-6984. DOI: 10.1021/acs.nanolett.6b05287.

[42] Sankar Narayanan et al. "Strain Hardening and Size Effect in Five-fold Twinned Ag Nanowires". In: *Nano Letters* 15.6 (2015), pp. 4037–4044. ISSN: 1530-6984. DOI: 10.1021/acs.nanolett.5b01015.

[43] Hui Li et al. "Fluorescence Enhancement of Silver Nanoparticle Hybrid Probes and Ultrasensitive Detection of IgE". In: *Analytical Chemistry* 83.23 (2011), pp. 8945–8952. ISSN: 0003-2700. DOI: 10.1021/ac201574s.

[44] Brendan Pietrobon and Vladimir Kitaev. "Photochemical Synthesis of Monodisperse Size-Controlled Silver Decahedral Nanoparticles and Their Remarkable Optical Properties". In: *Chemistry of Materials* 20.16 (2008), pp. 5186–5190. ISSN: 0897-4756. DOI: 10.1021/cm800926u.

[45] Xue Wang et al. "Pd@Pt Core–Shell Concave Decahedra: A Class of Catalysts for the Oxygen Reduction Reaction with Enhanced Activity and Durability". In: *Journal of the American Chemical Society* 137.47 (2015), pp. 15036–15042. ISSN: 0002-7863. DOI: 10.1021/jacs.5b10059.

[46] Yujie Xiong and Younan Xia. "Shape-controlled synthesis of metal nanostructures: The case of palladium". In: *Advanced Materials* 19.20 (2007), pp. 3385–3391. ISSN: 09359648. DOI: 10.1002/adma.200701301.

[47] Mark R Langille, Jian Zhang, and Chad A Mirkin. "Plasmon-Mediated Synthesis of Heterometallic Nanorods and Icosahedra". In: *Angewandte Chemie International Edition* 50.15 (2011), pp. 3543–3547.

[48] Xiaoming Ma et al. "Unveiling Growth Pathways of Multiply Twinned Gold Nanoparticles by In Situ Liquid Cell Transmission Electron Microscopy". In: *ACS Nano* 14.8 (2020), pp. 9594–9604. ISSN: 1936-0851. DOI: 10/ghhxft.

[49] H Hofmeister. "Habit and internal structure of multiply twinned gold particles on silver bromide films". In: *Thin Solid Films* 116.1-3 (1984), pp. 151–162.

[50] Herbert Hofmeister. "Shape variations and anisotropic growth of multiply twinned nanoparticles". In: *Zeitschrift für Kristallographie* 224.11 (2009), pp. 528–538. DOI: doi:10.1524/zkri.2009.1034. URL: https://doi.org/10.1524/zkri.2009.1034.

[51] Mark R Langille et al. "Stepwise evolution of spherical seeds into 20-fold twinned icosahedra". In: *Science* 337.6097 (2012), pp. 954–957.

[52] Miao Song et al. "Oriented attachment induces fivefold twins by forming and decomposing high-energy grain boundaries". In: *Science* 367.6473 (2020), pp. 40–45. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.aax6511.

[53] Masaharu Tsuji et al. "Stepwise growth of decahedral and icosahedral silver nanocrystals in DMF". In: *Crystal growth & design* 10.1 (2010), pp. 296–301.

[54] M. C. Scott et al. "Electron tomography at 2.4-ångström resolution". In: *Nature* 483.7390 (2012), pp. 444–447. ISSN: 1476-4687. DOI: 10/gddk92.

[55] Bart Goris et al. "Measuring Lattice Strain in Three Dimensions through Electron Microscopy". In: *Nano Letters* 15.10 (2015), pp. 6996–7001. ISSN: 1530-6984. DOI: 10.1021/acs.nanolett.5b03008.

[56] Peter Ercius et al. "Operation of TEAM I in a User Environment at NCEM". In: *Microscopy and Microanalysis* 18.4 (2012), pp. 676–683. ISSN: 1435-8115, 1431-9276. DOI: 10.1017/S1431927612001225.

[57] Colin Ophus, Jim Ciston, and Chris T. Nelson. "Correcting nonlinear drift distortion of scanning probe and scanning transmission electron microscopies from image pairs with orthogonal scan directions". In: *Ultramicroscopy* 162 (2016), pp. 1–9. ISSN: 0304-3991.

[58] A. Danielyan, V. Katkovnik, and K. Egiazarian. "BM3D Frames and Variational Image Deblurring". In: *IEEE Transactions on Image Processing* 21.4 (2012), pp. 1715–1728. ISSN: 1057-7149. DOI: 10.1109/TIP.2011.2176954.

[59] Yongsoo Yang et al. "Deciphering chemical order/disorder and material properties at the single-atom level". In: *Nature* 542.7639 (2017), pp. 75–79. ISSN: 1476-4687. DOI: 10/f9n48t.

[60] Tom Goldstein, Christoph Studer, and Richard Baraniuk. "A Field Guide to Forward-Backward Splitting with a FASTA Implementation". In: *arXiv:1411.3406 [cs]* (2014). URL: http://arxiv.org/abs/1411.3406.

[61] Jonas Adler et al. *odlgroup/odl: ODL 0.7.0*. Zenodo, 2018. DOI: 10.5281/zenodo.1442734. URL: https://zenodo.org/record/1442734.

[62] Wim van Aarle et al. "The ASTRA Toolbox: A platform for advanced algorithm development in electron tomography". In: *Ultramicroscopy* 157 (2015), pp. 35–47. ISSN: 0304-3991. DOI: 10/f7q2jk.

[63] Joachim Dengler. "A multi-resolution approach to the 3D reconstruction from an electron microscope tilt series solving the alignment problem without gold particles". In: *Ultramicroscopy* 30.3 (1989), pp. 337–348. ISSN: 0304-3991. DOI: 10/dx9ssk.

[64] David Ren et al. "A multiple scattering algorithm for three dimensional phase contrast atomic electron tomography". In: *Ultramicroscopy* 208 (2020), p. 112860.

[65] Leland McInnes, John Healy, and James Melville. "UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction". In: *arXiv:1802.03426 [cs, stat]* (2020). URL: http://arxiv.org/abs/1802.03426.

[66] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[67] Marin van Heel and Michael Schatz. "Information: to Harvest, to Have and to Hold". In: *arXiv:2009.03223 [physics]* (2020). URL: http://arxiv.org/abs/2009.03223.

[68] Colin Ophus. "A fast image simulation algorithm for scanning transmission electron microscopy". In: *Advanced Structural and Chemical Imaging* 3.1 (2017), p. 13. ISSN: 2198-0926. DOI: 10/gft3gf.

[69] Luis Rangel DaCosta et al. "Prismatic 2.0 - Simulation Software for Scanning and High Resolution Transmission Electron Microscopy (STEM and HRTEM)". In: *Micron* 3.1 (2021), p. 15. ISSN: 2198-0926. DOI: 10/gh5pp2.

[70] Hans Martin Kjer and Jakob Wilm. "Evaluation of surface registration algorithms for PET motion correction". B.S. thesis. Citeseer, 2010.

[71] Kallum M. Koczkur et al. "Polyvinylpyrrolidone (PVP) in nanoparticle synthesis". In: *Dalton Transactions* 44.41 (2015), pp. 17883–17905. ISSN: 14779234. DOI: 10.1039/c5dt02964c.

[72] Yao Yang et al. "Determining the three-dimensional atomic structure of an amorphous solid". In: *Nature* 592.7852 (2021), pp. 60–64. ISSN: 1476-4687. DOI: 10.1038/s41586-021-03354-0.

[73] Ya Zhou and Kristen A. Fichthorn. "Internal Stress-Induced Orthorhombic Phase in 5-Fold-Twinned Noble Metal Nanowires". In: *The Journal of Physical Chemistry C* 118.32 (2014), pp. 18746–18755. ISSN: 1932-7447. DOI: 10.1021/jp505334x.

[74] NM Rosengaard and Hans Lomholt Skriver. "Calculated stacking-fault energies of elemental metals". In: *Physical Review B* 47.19 (1993), p. 12865.

[75] B. G. Bagley. "A Dense Packing of Hard Spheres with Five-fold Symmetry". In: *Nature* 208.50115011 (1965), pp. 674–675. ISSN: 1476-4687. DOI: 10.1038/208674a0.

[76] Shozo Ino. "Stability of multiply-twinned particles". In: *Journal of the Physical Society of Japan* 27.4 (1969), pp. 941–953.

[77] A. Howie and L. D. Marks. "Elastic strains and the energy balance for multiply twinned particles". In: *Philosophical Magazine A* 49.1 (1984), pp. 95–109. ISSN: 0141-8610. DOI: 10.1080/01418618408233432.

[78] Peter Mahler Larsen, Søren Schmidt, and Jakob Schiøtz. "Robust structural identification via polyhedral template matching". In: *Modelling and Simulation in Materials Science and Engineering* 24.5 (2016), p. 055007.

[79] Wolfgang Lechner and Christoph Dellago. "Accurate determination of crystal structures based on averaged local bond order parameters". In: *The Journal of chemical physics* 129.11 (2008), p. 114707.

[80] Daniel Faken and Hannes Jónsson. "Systematic analysis of local atomic structure combined with 3D computer graphics". In: *Computational Materials Science* 2.2 (1994), pp. 279–286.

[81] Benjamin Harke et al. "Three-dimensional nanoscopy of colloidal crystals". In: *Nano letters* 8.5 (2008), pp. 1309–1313.

[82] Francesca Baletto and Riccardo Ferrando. "Island adsorption and adatom diffusion on 3D non-crystalline silver nanoclusters". In: *Surface Science* 490.3 (2001), pp. 361–375. ISSN: 0039-6028. DOI: https://doi.org/10.1016/S0039-6028(01)01355-3.

[83] Francesca Baletto and Riccardo Ferrando. "Structural properties of nanoclusters: Energetic, thermodynamic, and kinetic effects". In: *Rev. Mod. Phys.* 77 (1 2005), pp. 371–423. DOI: 10.1103/RevModPhys.77.371. URL: https://link.aps.org/doi/10.1103/RevModPhys.77.371.

[84] Xiaohu Xia et al. "On the role of surface diffusion in determining the shape or morphology of noble-metal nanocrystals". In: *Proceedings of the National Academy of Sciences* 110.17 (2013), pp. 6669–6673. DOI: 10.1073/pnas.1222109110.

[85] Anubhav Jain et al. "Commentary: The Materials Project: A materials genome approach to accelerating materials innovation". In: *APL Materials* 1.1 (July 2013), p. 011002. DOI: 10.1063/1.4812323.

[86] Shuo Zhao et al. "Influence of Atomic-Level Morphology on Catalysis: The Case of Sphere and Rod-Like Gold Nanoclusters for CO $_2$ Electroreduction". In: *ACS Catalysis* 8.6 (June 2018), pp. 4996–5001. DOI: 10.1021/acscatal.8b00365.

[87] Haider I. Rasool et al. "Measurement of the intrinsic strength of crystalline and polycrystalline graphene". In: *Nature Communications* 4.1 (Dec. 2013). DOI: 10.1038/ncomms3811.

[88] Christian Szegedy, Alexander Toshev, and Dumitru Erhan. "Deep Neural Networks for Object Detection". In: *Advances in Neural Information Processing Systems 26*. Ed. by C. J. C. Burges et al. Curran Associates, Inc., 2013, pp. 2553–2561.

[89] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Ed. by Nassir Navab et al. Vol. 9351. Cham: Springer International Publishing, 2015, pp. 234–241. DOI: 10.1007/978-3-319-24574-4_28.

[90] Catherine K. Groschner, Christina Choi, and M. C. Scott. "Methodologies for Successful Segmentation of HRTEM Images via Neural Network". In: (2020). arXiv: 2001.05022. URL: http://arxiv.org/abs/2001.05022.

[91] Artem Maksov et al. "Deep learning analysis of defect and phase evolution during electron beam-induced transformations in WS2". In: *npj Computational Materials* 5.1 (2019), p. 12. DOI: 10.1038/s41524-019-0152-9.

[92] Jakob Schiøtz et al. "Identifying Atoms in High Resolution Transmission Electron Micrographs Using a Deep Convolutional Neural Net". In: *Microscopy and Microanalysis* 24.S1 (2018), pp. 512–513. DOI: 10.1017/S1431927618003057.

[93] Jakob Schiøtz et al. "Using Neural Networks to Identify Atoms in HRTEM Images". In: *Microscopy and Microanalysis* 25.S2 (2019), pp. 216–217. DOI: 10.1017/S1431927619001818.

[94] Sergei V. Kalinin et al. "Lab on a beam—Big data and artificial intelligence in scanning transmission electron microscopy". In: *MRS Bulletin* 44.7 (2019), pp. 565–575. DOI: 10.1557/mrs.2019.159.

[95] Jiadong Dan, Xiaoxu Zhao, and Stephen J. Pennycook. "A machine perspective of atomic defects in scanning transmission electron microscopy". In: *InfoMat* 1.3 (2019), pp. 359–375. ISSN: 2567-3165. DOI: 10.1002/inf2.12026.

[96] Benedykt R. Jany, Arkadiusz Janas, and Franciszek Krok. "Automatic microscopic image analysis by moving window local Fourier Transform and Machine Learning". In: *Micron* 130.July 2019 (2020), p. 102800. ISSN: 09684328. DOI: 10.1016/j.micron.2019.102800. URL: https://doi.org/10.1016/j.micron.2019.102800.

[97] Georg Daniel Förster et al. "A deep learning approach for determining the chiral indices of carbon nanotubes from high-resolution transmission electron microscopy images". In: *Carbon* 169 (2020), pp. 465–474. ISSN: 00086223. DOI: 10.1016/j.carbon.2020.06.086. arXiv: 2003.11505. URL: https://linkinghub.elsevier.com/retrieve/pii/S0008622320306552.

[98] Nik Dennler et al. "Learning-based Defect Recognition for Quasi-Periodic Microscope Images". In: *arXiv* (2020), pp. 1–19. arXiv: 2007.01309.

[99] D. J. Groom et al. "Automatic segmentation of inorganic nanoparticles in BF TEM micrographs". In: *Ultramicroscopy* 194.June (2018), pp. 25–34. DOI: 10.1016/j.ultramic.2018.06.002.

[100] Max Frei and Frank Einar Kruis. "Fully automated primary particle size analysis of agglomerates on transmission electron microscopy images via artificial neural networks". In: *Powder Technology* 332 (2018), pp. 120–130. ISSN: 1873328X. DOI: 10.1016/j.powtec.2018.03.032. arXiv: 1806.04010.

[101] Christine R. Laramy et al. "High-Throughput, Algorithmic Determination of Nanoparticle Structure from Electron Microscopy Images". In: *ACS Nano* 9.12 (2015), pp. 12488–12495. ISSN: 1936086X. DOI: 10.1021/acsnano.5b05968.

[102] Wei Li, Kevin G. Field, and Dane Morgan. "Automated defect analysis in electron microscopic images". In: *npj Computational Materials* 4.1 (2018). ISSN: 20573960. DOI: 10.1038/s41524-018-0093-8.

[103] Gokhan Güven and Ayse Betul Oktay. "Nanoparticle detection from TEM images with deep learning". In: *26th IEEE Signal Processing and Communications Applications Conference, SIU 2018* (2018), pp. 1–4. DOI: 10.1109/SIU.2018.8404468.

[104] Graham Roberts et al. "Deep Learning for Semantic Segmentation of Defects in Advanced STEM Images of Steels". In: *Scientific Reports* 9.1 (2019), pp. 1–12. ISSN: 20452322. DOI: 10.1038/s41598-019-49105-0.

[105] Jiwon Yeom et al. "Segmentation of Experimental Datasets Via Convolutional Neural Networks Trained on Phase Field Simulations". In: *SSRN Electronic Journal* (2020). DOI: 10.2139/ssrn.3680400.

[106] *Get Started with the Image Labeler - MATLAB & Simulink*. URL: https://www.mathworks.com/help/vision/ug/get-started-with-the-image-labeler.html (visited on 05/30/2019).

[107] François Chollet et al. *Keras*. https://keras.io. 2015.

[108] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. http://www.deeplearningbook.org. MIT Press, 2016.

[109] Diederik P Kingma and Jimmy Lei Ba. "Adam: A Method for Stochastic Optimization". In: *arXiv* (2015), pp. 1–15. arXiv: arXiv:1412.6980v9.

[110] Nobuyuki Otsu. "A Threshold Selection Method from Gray-Level Histograms". In: *IEEE Transactions on Systems, Man, and Cybernetics* SMC 9.1 (1979), pp. 62–66. ISSN: 0018-9472. DOI: 10.1109/TSMC.1979.4310076.

[111] Stéfan van der Walt et al. "scikit-image: image processing in Python". In: *PeerJ* 2 (June 2014), e453. ISSN: 2167-8359. DOI: 10.7717/peerj.453. URL: https://doi.org/10.7717/peerj.453.

[112] Abdel Aziz Taha and Allan Hanbury. "Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool". In: *BMC Medical Imaging* 15.1 (2015). ISSN: 14712342. DOI: 10.1186/s12880-015-0068-x.

[113] Brian L. DeCost et al. "High Throughput Quantitative Metallography for Complex Microstructures Using Deep Learning: A Case Study in Ultrahigh Carbon Steel". In: *Microscopy and Microanalysis* 25.1 (2019), pp. 21–29. DOI: 10.1017/S1431927618015635.

[114] P.R. Buseck, Y. Epelboin, and A. Rimsky. "Signal Processing of High-Resolution Transmission Electron Microscope Images Using Fourier Transforms". In: *Acta Cryst.* A44 (1988), pp. 975–986.

[115] A. F. De Jong, W. Coene, and D. Van Dyck. "Image processing of HRTEM images with non-periodic features". In: *Ultramicroscopy* 27.1 (1989), pp. 53–65. ISSN: 03043991. DOI: 10.1016/0304-3991(89)90200-3.

[116] Xiaogang Peng et al. "Shape control of CdSe nanocrystals". In: *Nature* 404.6773 (2000), pp. 59–61. ISSN: 00280836. DOI: 10.1038/35003535.

[117] Geron, Aurelien. *Hands-On Machine Learning with Scikit-Learn & TensorFlow*. 1st ed. Sebatopol: O'Reilly Media, 2017.

[118] Wei Li, Kevin G Field, and Dane Morgan. "Automated defect analysis in electron microscopic images". In: *npj Computational Materials* 4.1 (2018), pp. 1–9.

[119] G Carneiro et al. *Deep Learning and Data Labeling for Medical Applications: First International Workshop, LABELS 2016, and Second International Workshop, DLMIA 2016, Held in Conjunction with MICCAI 2016, Athens, Greece, October 21, 2016, Proceedings*. Lecture Notes in Computer Science. Springer International Publishing, 2016. ISBN: 9783319469768. URL: https://books.google.com/books?id=p6U3DQAAQBAJ.

[120] James P Horwath et al. "Understanding important features of deep learning models for segmentation of high-resolution transmission electron microscopy images". In: *npj Computational Materials* 6.1 (2020), p. 108. ISSN: 2057-3960. DOI: 10.1038/s41524-020-00363-x. URL: https://doi.org/10.1038/s41524-020-00363-x.

[121] Karen Simonyan and Andrew Zisserman. *Very Deep Convolutional Networks for Large-Scale Image Recognition*. 2015. arXiv: 1409.1556 [cs.CV].

[122] Kaiming He et al. "Deep Residual Learning for Image Recognition". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016.

[123] Andrew G. Howard et al. *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications*. 2017.

[124] Liang-Chieh Chen et al. "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 40.4 (2018), pp. 834–848. DOI: 10.1109/TPAMI.2017.2699184.

[125] Tsung-Yi Lin et al. "Feature Pyramid Networks for Object Detection". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017.

[126] Pavel Yakubovskiy. *Segmentation Models*. `https://github.com/qubvel/segmentation_models`. 2019.

[127] Wei Ma et al. "Attomolar DNA detection with chiral nanorod assemblies". In: *Nature Communications* 4.May (2013). ISSN: 20411723. DOI: 10.1038/ncomms3689.

[128] Peng-peng Wang, Shang Jie Yu, and Min Ouyang. "Assembled Suprastructures of Inorganic Chiral Nanocrystals and Hierarchical Chirality". In: *Journal of the American Chemical Society* 139.17 (2017), pp. 6070–6073. ISSN: 15205126. DOI: 10.1021/jacs.7b02523.

[129] Hongyu Zhang et al. "Engineering of chiral nanomaterials for biomimetic catalysis". In: *Chemical Science* 11.48 (2020), pp. 12937–12954. ISSN: 20416539. DOI: 10.1039/d0sc03245j.

[130] Assaf Ben-Moshe et al. "Enantioselective control of lattice and shape chirality in inorganic nanostructures using chiral biomolecules". In: *Nature Communications* 5 (2014). ISSN: 20411723. DOI: 10.1038/ncomms5302.

[131] Peng-peng Wang et al. "Cooperative expression of atomic chirality in inorganic nanostructures". In: *Nature Communications* 8.1 (2017), p. 14312. ISSN: 2041-1723. DOI: 10.1038/ncomms14312. URL: https://doi.org/10.1038/ncomms14312.

[132] Wei Ma et al. "Chiral Inorganic Nanostructures". In: *Chemical Reviews* 117.12 (2017), pp. 8041–8093. DOI: 10.1021/acs.chemrev.6b00755.

[133] Uri Hananel et al. "Enantiomeric Control of Intrinsically Chiral Nanocrystals". In: *Advanced Materials* 32.41 (2020). ISSN: 15214095. DOI: 10.1002/adma.201905594.

[134] Hyeohn Kim et al. "γ-Glutamylcysteine- and Cysteinylglycine-Directed Growth of Chiral Gold Nanoparticles and their Crystallographic Analysis". In: *Angewandte Chemie - International Edition* 59.31 (2020), pp. 12976–12983. ISSN: 15213773. DOI: 10.1002/anie.202003760.

[135] Nam Heon Cho et al. "Cysteine Induced Chiral Morphology in Palladium Nanoparticle". In: *Particle and Particle Systems Characterization* 36.5 (2019), pp. 1–5. ISSN: 15214117. DOI: 10.1002/ppsc.201900062.

[136] Ali Rafiei Miandashti et al. "Experimental and Theoretical Observation of Photothermal Chirality in Gold Nanoparticle Helicoids". In: *ACS nano* 14.4 (2020), pp. 4188–4195. ISSN: 1936086X. DOI: 10.1021/acsnano.9b09062.

[137] Zhuoya Dong and Yanhang Ma. "Atomic-level handedness determination of chiral crystals using aberration-corrected scanning transmission electron microscopy". In: *Nature Communications* 11.1 (2020), pp. 1–6. ISSN: 20411723. DOI: 10.1038/s41467-020-15388-5. URL: http://dx.doi.org/10.1038/s41467-020-15388-5.

[138] Jeffrey Mark Ede. "Review: Deep Learning in Electron Microscopy". In: *Machine Learning: Science and Technology* (2020). URL: http://iopscience.iop.org/article/10.1088/2632-2153/abd614.

[139] Adam Paszke et al. "PyTorch: An Imperative Style, High-Performance Deep Learning Library". In: *Advances in Neural Information Processing Systems 32*. Ed. by H Wallach et al. Curran Associates, Inc., 2019, pp. 8024–8035. URL: http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf.

[140] Chiyuan Zhang et al. "Understanding deep learning requires rethinking generalization". In: *5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings* (2017). arXiv: 1611.03530.

[141] Devansh Arpit et al. "A closer look at memorization in deep networks". In: *arXiv* (2017). ISSN: 23318422.

[142] David Rolnick et al. "Deep learning is robust to massive label noise". In: *arXiv* (2017). ISSN: 23318422. arXiv: 1705.10694.

[143] Yisen Wang et al. "Symmetric cross entropy for robust learning with noisy labels". In: *arXiv* (2019), pp. 322–330. ISSN: 23318422.

[144] Giorgio Patrini et al. "Making deep neural networks robust to label noise: A loss correction approach". In: *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017* 2017-Janua (2017), pp. 2233–2241. DOI: 10.1109/CVPR.2017.240. arXiv: 1609.03683.

[145] Sunil Thulasidasan et al. "Combating label noise in deep learning using abstention". In: *arXiv* (2019). ISSN: 23318422.

[146] Bo Han et al. "Co-teaching: Robust training of deep neural networks with extremely noisy labels". In: *Advances in Neural Information Processing Systems* 2018-Decem.Nips (2018), pp. 8527–8537. ISSN: 10495258. arXiv: 1804.06872.

[147] Eric Arazo et al. "Unsupervised label noise modeling and loss correction". In: *36th International Conference on Machine Learning, ICML 2019* 2019-June (2019), pp. 465–474. arXiv: 1904.11238.

[148] Yair Dgani, Hayit Greenspan, and Jacob Goldberger. "Training a neural network based on unreliable human annotation of medical images". In: *Proceedings - International Symposium on Biomedical Imaging* 2018-April.Isbi (2018), pp. 39–42. ISSN: 19458452. DOI: 10.1109/ISBI.2018.8363518.

[149] Davood Karimi et al. "Deep learning with noisy labels: Exploring techniques and remedies in medical image analysis". In: *arXiv* (2019), pp. 1–22. ISSN: 23318422. arXiv: 1912.02911.

[150] Benoît Frénay and Michel Verleysen. "Classification in the presence of label noise: A survey". In: *IEEE Transactions on Neural Networks and Learning Systems* 25.5 (2014), pp. 845–869. ISSN: 21622388. DOI: 10.1109/TNNLS.2013.2292894.

[151] Dan Hendrycks et al. "Using trusted data to train deep networks on labels corrupted by severe noise". In: *Advances in Neural Information Processing Systems* 2018-Decem.Nips (2018), pp. 10456–10465. ISSN: 10495258. arXiv: 1802.05300.

[152] Jonathan Krause et al. "The unreasonable effectiveness of noisy data for fine-grained recognition". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9907 LNCS (2016), pp. 301–320. ISSN: 16113349. DOI: 10.1007/978-3-319-46487-9_19. arXiv: 1511.06789.

[153] Andreas Veit et al. "Learning From Noisy Large-Scale Datasets With Minimal Supervision". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017.

[154] Hansi Yang et al. "Searching to Exploit Memorization Effect in Learning from Corrupted Labels". In: *CoRR* abs/1911.02377 (2019). arXiv: 1911.02377. URL: http://arxiv.org/abs/1911.02377.

[155] Vitaly Feldman. "Does Learning Require Memorization? A Short Tale about a Long Tail". In: *arXiv* (2019), pp. 954–959. ISSN: 23318422.

[156] Dumitru Erhan et al. "The difficulty of training deep architectures and the effect of unsupervised pre-training". In: *Journal of Machine Learning Research* 5 (2009), pp. 153–160. ISSN: 15324435.

Appendix A

Appendix 1

A.1 Tilt Series

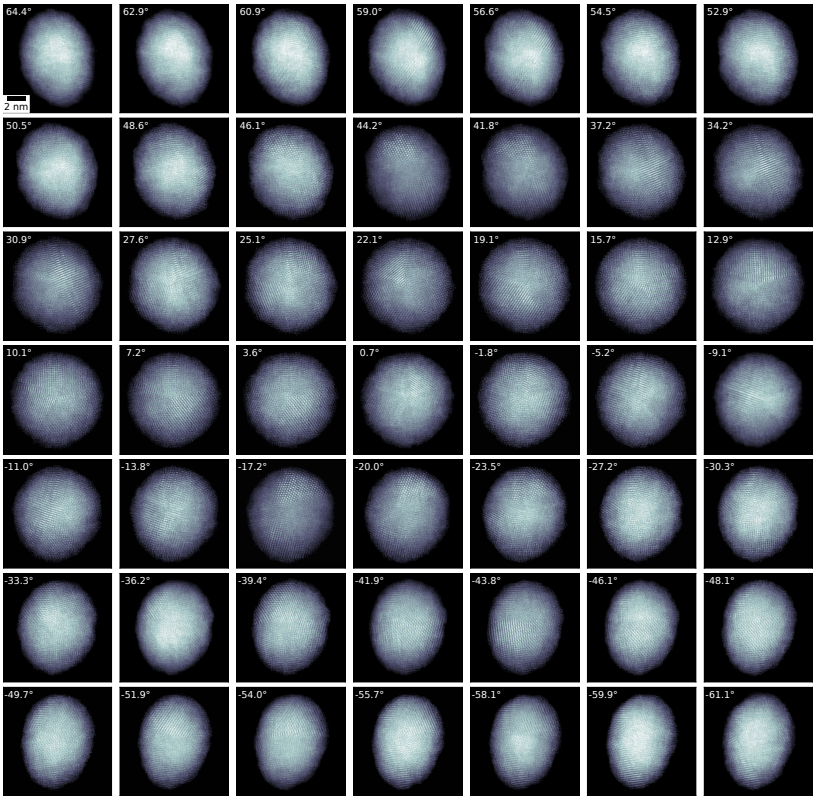


Figure A.1: Tomographic tilt series of the multiply-twinned Pd nanoparticle. The 49 projection images with a tilt range from 64 to −61 degrees (show at top left of each panel) were measured using ADF-STEM.

A.2 Before and After Tomography

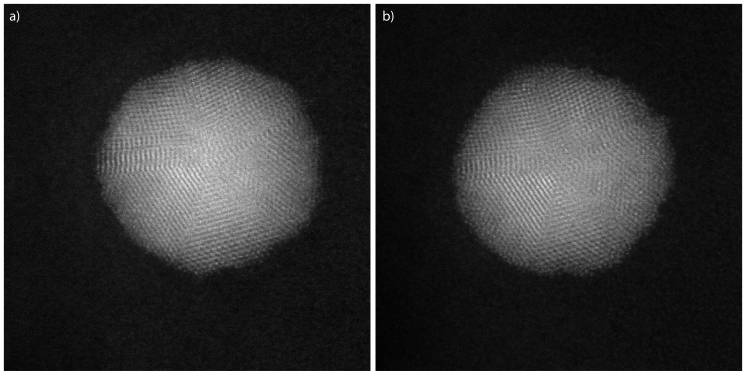


Figure A.2: a) Particle before and (b) after tomography experiment.

A.3 Atom Tracing

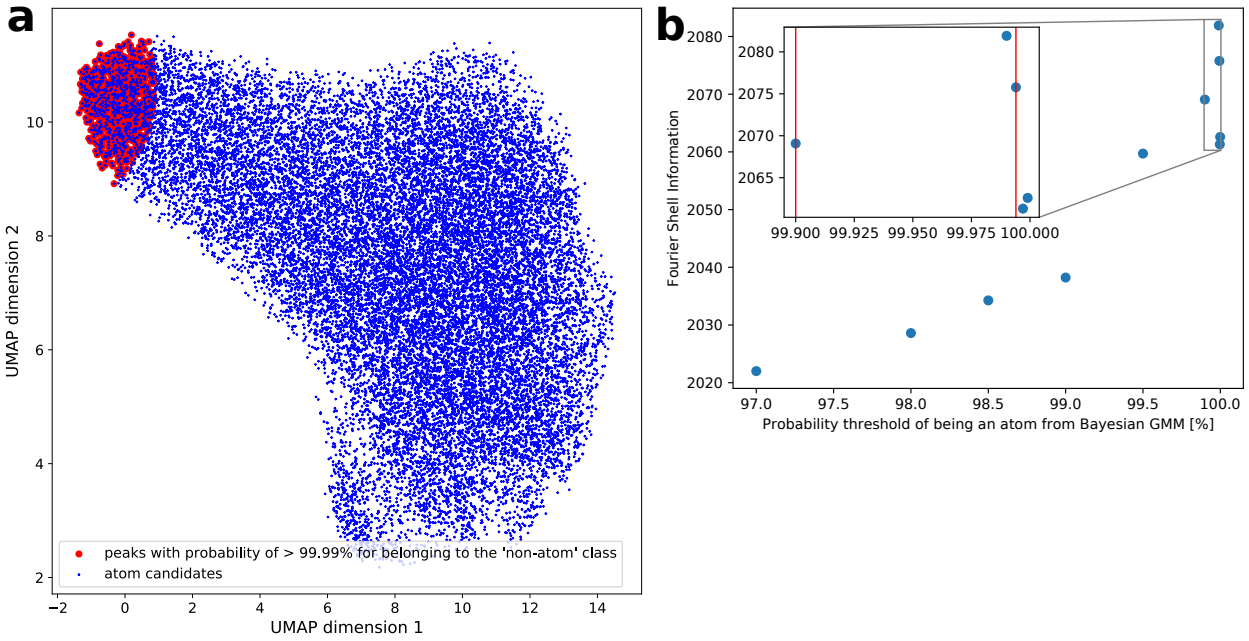


Figure A.3: a) Two-dimensional manifold of the peak candidates obtained with UMAP. b) Fourier shell information of PRISM-simulated tilt series performed with different probability thresholds for the atom candidate list. The probability of and atom candidate not being an atom was obtained from a two-class Bayesian Gaussian Mixture model fitted to the manifold in a). The threshold range between the red lines was identified as the range that extracts the most information from the experimental data.

A.4 STEM Simulations and Tracing Precision

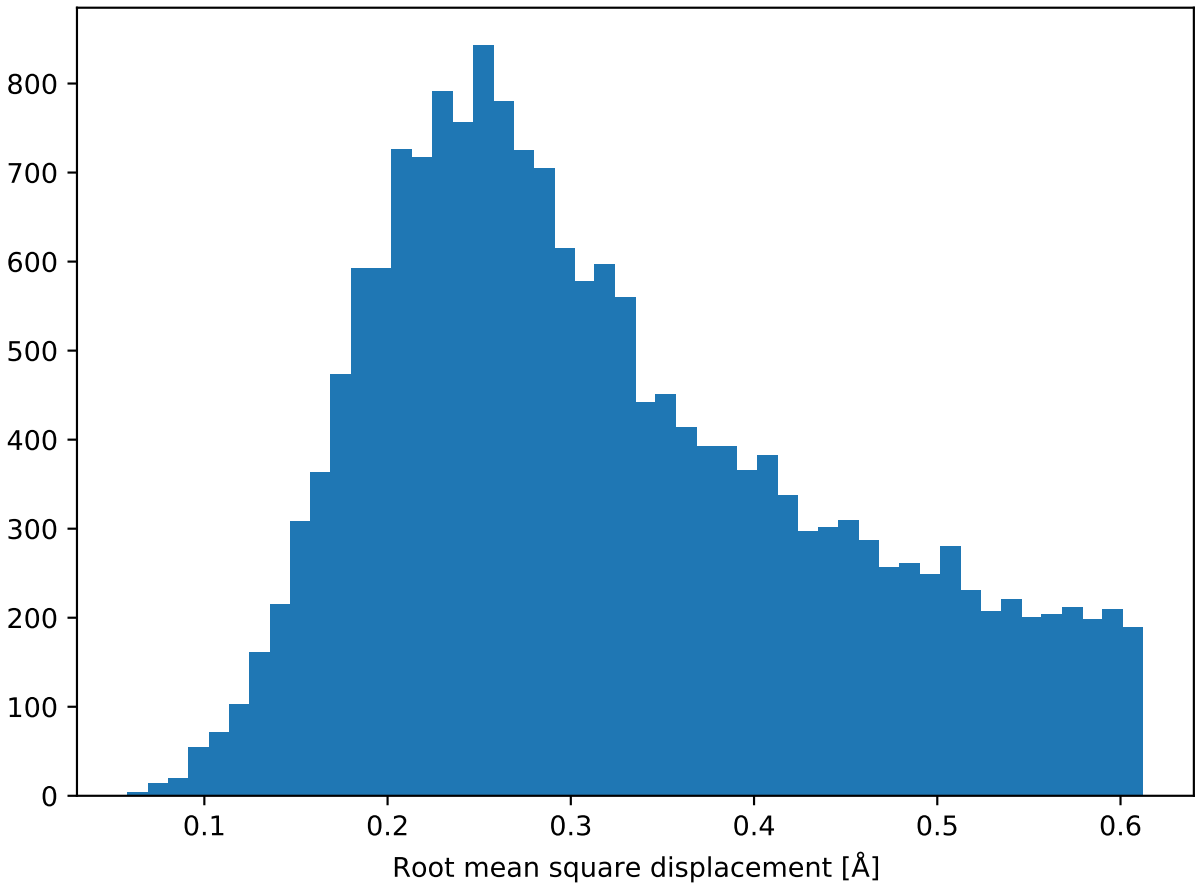


Figure A.4: Histogram of the difference in atomic positions between the experimentally determined atomic model and that obtained from multi-slice simulations.

A.5 Ordering Parameter Comparison

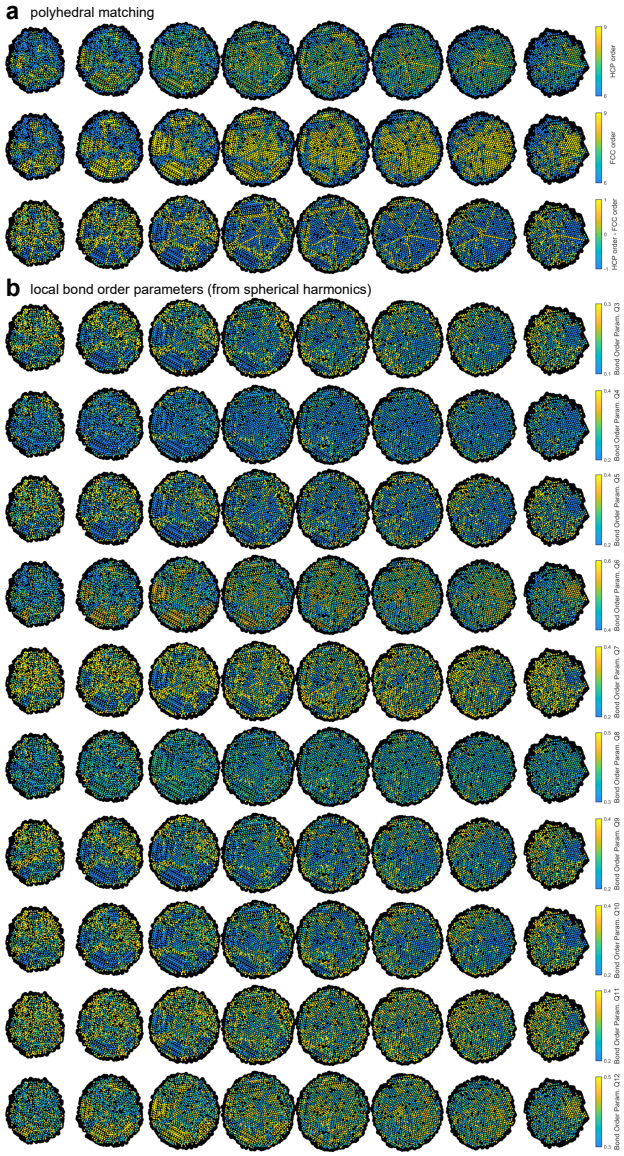


Figure A.5: Comparison of methods to determine local ordering. a) Classification of hcp vs. fcc ordering in volume slices using a polyhedral matching algorithm [78]. b) Classification of fcc vs hcp ordering in volume slices using spherical harmonics.

Appendix B

Appendix 2

B.1 Results From Different Training Sets

Segmentation was found to improve when a more varied dataset was used. Results for training solely on CdSe compared to CdSe and Au are presented in Table B.1 and in Figure B.1.

	DICE	Precision	Recall
U-Net with only CdSe Data	0.43	0.44	0.42
U-Net with Combined Data	0.59	0.56	0.62

Table B.1: Performance metrics for U-Net on CdSe data when the U-Net was trained with different datasets. The same number of images were used for each training instance.

B.2 Threshold Determination

Precision versus threshold and recall versus threshold was plotted for both the Au nanoparticle and CdSe nanoparticle segmentation. Based on the plots the optimal threshold was chosen. The plots are shown in Figure B.2

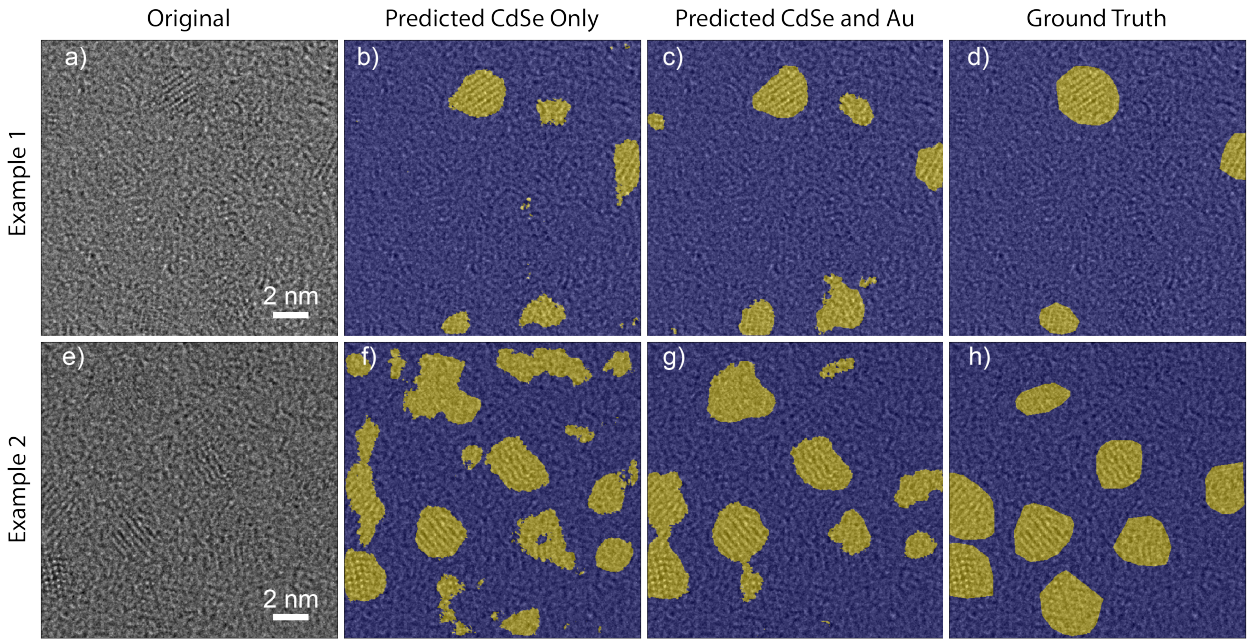
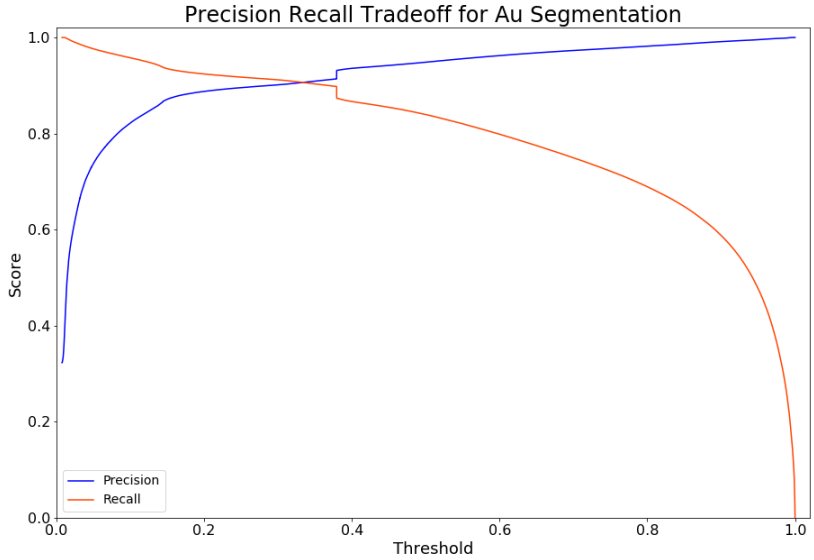
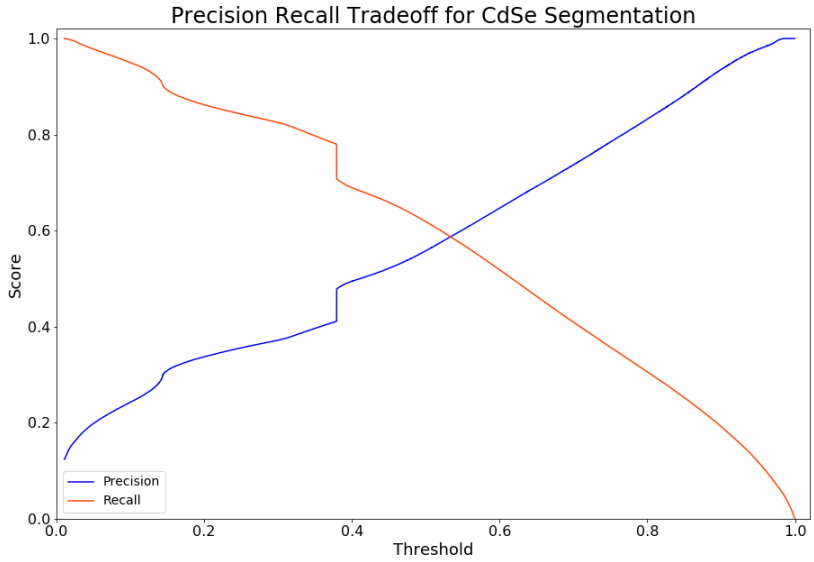


Figure B.1: Two sample micrographs of CdSe particles and the resulting segmentation maps when trained with only CdSe data and when trained with Au and CdSe data.



(a)



(b)

Figure B.2: Precision and recall vs threshold for (a) segmentation of Au nanoparticles and (b) segmentation of CdSe nanoparticles.

Appendix C

Appendix 3

C.1 Section of the Maps 2.5 Acquired SEM Micrograph

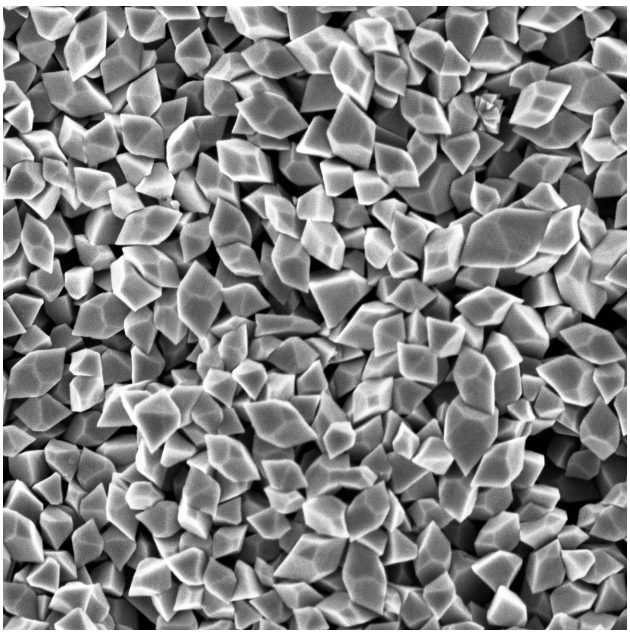


Figure C.1: Sample SEM micrograph of Te nanoparticles acquired with Maps 2.5 software.

C.2 Sample SEM Images Ideal Dataset with Mirroring

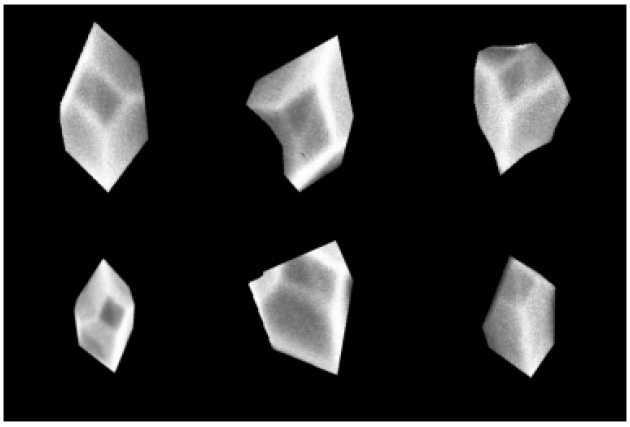


Figure C.2: Sample segmented Te nanoparticles from SEM micrographs from the ideal dataset.

C.3 Sample SEM Images As-Synthesized Dataset

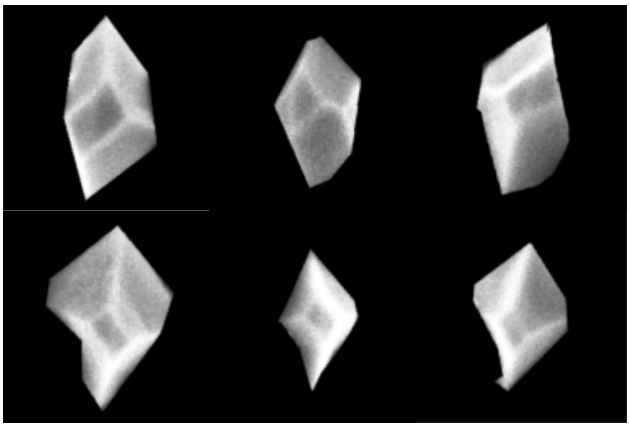


Figure C.3: Sample segmented Te nanoparticles from SEM micrographs from the as-synthesized dataset.