

CHRISTIAN-ALBRECHTS-UNIVERSITÄT ZU  
KIEL

---

DEPARTMENT OF COMPUTER SCIENCE

Master Thesis

**“Denoising 3D TEM tomography via Advanced  
Neural Radiance Fields”**

Master’s in Computer Science

Mithun Das

SUPERVISORS

**Prof. Dr. Peer Kröger  
Dr.-Ing. Claudius Zelenka**

Kiel, Germany



# Abstract

---

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.



# Acknowledgments

---

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

# Contents

<b>Abstract</b>	<b>iii</b>
<b>Acknowledgments</b>	<b>v</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Background Information</b>	<b>5</b>
2.0.1 Transmission electron microscopy(TEM) . . . . .	5
2.0.2 Neural 3D shape representations . . . . .	7
2.0.3 Novel view Synthesis . . . . .	8
2.0.4 Camera Parameters . . . . .	9
2.0.5 COLMAP . . . . .	11
2.0.6 NeRF (Neural Radiance Field) . . . . .	13
2.0.7 Enhanced Super-Resolution Generative Adversarial Net- works (ESRGAN) . . . . .	16
<b>3 Related Work</b>	<b>17</b>
3.0.1 Inverse problem . . . . .	17
3.0.2 Atomic Resolution . . . . .	19
3.0.3 Noise Modeling . . . . .	20
3.0.4 3D Convolutional Neural Networks . . . . .	21
3.0.5 Denoising . . . . .	22
3.0.6 View Synthesis and Image-based rendering . . . . .	23
<b>4 Datasets</b>	<b>25</b>
4.0.1 TEM Datasets . . . . .	25
4.0.2 STEM Datasets . . . . .	28
4.0.3 Synthetic Datasets . . . . .	29
<b>5 Implementation</b>	<b>31</b>
5.0.1 Setup Overview . . . . .	31
5.0.2 Key Implementation Steps . . . . .	32
5.0.3 Code Availability . . . . .	35
<b>Bibliography</b>	<b>37</b>

# List of Figures

1.1	Comparative Analysis of Original TEM Images and Enhanced Outputs . . . . .	3
2.1	Schematic diagram of Transmission Electron Microscope . . . . .	5
2.2	TEM images of a mesoporous material(MCM-41) at different magnifications [12] . . . . .	6
2.3	DeepSDF's 3D bunny shape representation [13] . . . . .	7
2.4	Multi-view to Novel view synthesis[20] . . . . .	8
2.5	Schematic Representation of Camera Parameters . . . . .	9
2.6	COLMAP camera position extracting from set of 2D images . . . . .	11
2.7	Conceptual Illustration of Neural Radiance Fields (NeRF) [7] . . . . .	13
2.8	Dataset of 2D Images from Various Camera Positions . . . . .	14
2.9	Still Frames from Video Generated Using 2D Images with New Camera Angle . . . . .	14
2.10	ESRGAN Architecture [37] . . . . .	16
3.1	Comparative Reconstruction of a 10 nm Diameter Gold Particle: (A) Using Linear Backprojection in IMOD Software, showing deviation from spherical shape. (B) Using Curvilinear Backprojection in TxBR, accurately representing a spherical particle. [41] . . . . .	17
3.2	Atomic Resolution 3D Imaging of Platinum Nanoparticle: (A) STEM image showing flat twin boundaries. (B) AET reconstruction reveals atomic steps at twin boundaries, with a grain boundary and stacking fault. [50] . . . . .	19
3.3	First row - Original real images, Second row - Noisier versions, Third row - Images after denoising with CNN DAE. [64] . . . . .	21
3.4	TEM Imaging and Denoising of Cadmium Sulfide Nanoparticles: <b>A.</b> Original Image, <b>B.</b> Image with Gaussian Noise, <b>C.</b> Denoised with Average Filter, <b>D.</b> Denoised with Median Filter, <b>E.</b> Denoised with Weiner Filter. [66] . . . . .	22
4.1	TEM Dataset 1 . . . . .	25
4.2	TEM Dataset 2 . . . . .	26
4.3	TEM Dataset 3 . . . . .	27
4.4	TEM Dataset 4 . . . . .	27
4.5	STEM Dataset 1 . . . . .	28
4.6	STEM Dataset 2 . . . . .	28
4.7	Synthetic Dataset . . . . .	29

5.1	Comprehensive Overview of the our Framework: Integrating NeRF and Advanced Denoising Techniques for Enhanced TEM Image Analysis . . . . .	31
-----	---	----



# CHAPTER 1

## Introduction

---

Electron tomography, often known as ET, is now the method of choice for determining the three-dimensional ultra-structure of organelles and cells at nanoscale resolutions. The 3D volume of the specimen can be reproduced by first collecting a tilted set of 2D transmission electron microscopy (TEM) images over a wide-angle range (usually  $\pm 60^\circ$  to  $80^\circ$ ), and then computationally recombining the images. However, the low electron doses that are applicable to biological samples (which are typically less than  $100 \text{ e}^-/\text{\AA}^2$ ) result in extremely poor signal-to-noise ratios (SNR) in the tomograms that are produced as a result [1]. The leading causes of the noise are the stochastic character of the events involving electron scattering and the constraints imposed by electron detection [2]. In addition, flaws in the alignment of the tilt axis, beam-induced specimen deformation, and distortions that are inherent to electron lenses all contribute to the contamination of the TEM data [2]. This leads to very tiny structural features, which are essential for interpreting the sophisticated cellular processes and chemical interactions, to get confused and distorted. Therefore, reducing the amount of noise in tomograms is a critical step in the preprocessing phase that comes before extracting information that has biological significance.

In order to improve 3D ET reconstructions, a number of different denoising algorithms have been implemented. To a certain extent, straightforward linear filters like median filtering, Gaussian smoothing, and anisotropic diffusion filtration can reduce noise, but at the cost of a significant loss of high-resolution information [3]. More enhanced regularization approaches such as total variation (TV) minimization and sparse coding exploit image priors to preserve edges and the rigidity of an image. However, these methods frequently require extensive parameter tuning to strike a balance between the removal of noise and the absorption of detailed information. Deep learning models such as [4] have shown promise for 2D image denoising tasks. However, directly implementing such networks to tomograms slice-by-slice is unable to effectively exploit 3D contextual information and the spatial relationships between coordinates. Although several algorithms are capable of doing block-wise 3D denoising, they are restricted by computational restrictions [4]. Other methods involve training on simulated data, which might not translate well to tomograms taken from actual data. In addition, the majority of currently available deep learning algorithms lack interpretability into the newly acquired features and have difficulty denoising non-uniform noise distributions, which are frequently seen in practice.

While deep learning models like DnCNN [5] have shown promise for 2D image denoising tasks, directly applying such networks to tomograms slice-by-slice fails to fully utilize 3D contextual information and spatial relationships between voxels. Some methods perform block-wise 3D denoising but are limited by computational constraints when scaling to large high-resolution volumes [6]. Other techniques pretrain on simulated data which may not generalize well to real experimental tomograms with complex noise textures [6]. Most existing deep learning approaches also lack interpretability into the learned features and struggle to denoise non-uniform noise distributions as encountered in practice.

More recently, neural radiance fields (NeRF) [7] have demonstrated unprecedented ability to synthesize photorealistic novel views of complex 3D scenes using a continuous volumetric representation. NeRFs learn a 5D radiance field where each 3D coordinate (X,Y,Z) is mapped to an emitted color (R,G,B) and volume density  $\sigma$  using a standard multilayer perceptron (MLP). The key advantages of NeRF over other 3D deep learning representations are:

- Coordinate-based MLPs better capture local spatial relationships compared to convolutional networks
- Continuous scene modeling enables synthesizing views from arbitrary poses

In this work, we propose adapting NeRF for directly denoising 3D volumes reconstructed from tilt series ET data. By training on pairs of noisy input and clean target volumes, the NeRF model may learn to infer superior denoised outputs closely matching the ground truth. The continuous volumetric modeling could outperform other 3D networks that lack such inductive bias. This approach could significantly enhance interpretability of structural details from electron tomograms.

However, directly applying NeRF for reconstructing volumes from noisy TEM tilt series poses significant challenges. The COLMAP structure-from-motion algorithm is utilized by the standard NeRF pipeline in order to perform camera pose estimation for each input image. The high noise levels of TEM projections, on the other hand, can make it difficult for COLMAP to properly establish the viewing angles. Because of this, it is challenging to train NeRF directly on raw TEM pictures that contain noise. Within the scope of this research, we suggest alterations to the NeRF framework that, if implemented, will make it possible to obtain a more accurate camera pose estimation from noisy TEM tilt series, where COLMAP fails to provide any camera poses.

We propose modifications to the NeRF framework to enable more robust camera pose estimation from noisy TEM tilt series. We also investigate training strategies and loss formulations to better condition the model on the noise characteristics of real ET data. This includes using robust loss functions that focus on structure rather than pixel intensities, as well as adaptive sampling and conditioning schemes. By adapting NeRF to handle noisy inputs in this manner, we aim to overcome the limitations of standard NeRF applied directly out-of-the-box to electron tomography volumes. Our noise-aware NeRF model could open new possibilities for high-fidelity 3D denoising and analysis of ET reconstructions.

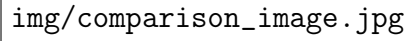
A large rectangular box with a thin black border. Inside the box, the text "img/comparison\_image.jpg" is written in a monospaced font, positioned in the upper-left quadrant of the box.

Figure 1.1: Comparative Analysis of Original TEM Images and Enhanced Outputs

A visual comparison between the original TEM images and the outcomes produced by our suggested method in Figure 1.1. This contrast draws attention to the difficulties caused by the noise in the original data and emphasizes how our method can improve image quality and lead to more precise analysis. Our innovative denoising technique was motivated by the significant reduction in noise and improvement in clarity, as illustrated in the picture. It demonstrates the significant influence that efficient noise reduction can have on TEM picture interpretability, which is necessary for precise 3D reconstruction and later biological investigation.



## CHAPTER 2

# Background Information

---

### 2.0.1 Transmission electron microscopy(TEM)

A beam of electrons is used in transmission electron microscopy (TEM), which generates images of specimens with a resolution that is far higher than that of optical microscopes [8]. In transmission electron microscopy, electrons are emitted by a tungsten filament or field emission source and then accelerated under high voltage (typically 100-300 kV) [9]. Electromagnetic lenses concentrate the electron beam such that it is directed toward the extremely thin sample. Electrons, when they go through the sample, have a variety of interactions with the sample, depending on the density and the thickness of the material. This produces an electron diffraction pattern, which may be interpreted to reveal information about the structure of the material [10].

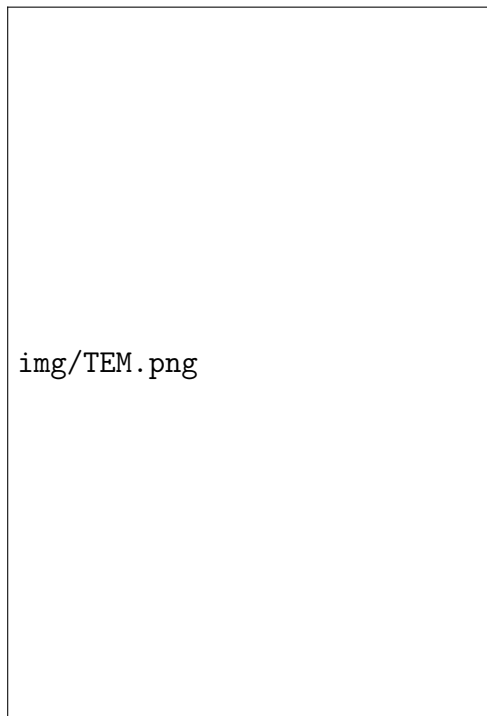


Figure 2.1: Schematic diagram of Transmission Electron Microscope

Source: <https://anapath.ch/electron-microscopy-2/>

Additional lenses concentrate the transmitted electrons so that they may be captured as an image on a detector or camera [9]. The transmission electron microscope (TEM) may provide magnifications of up to 2 million times [9], which enables the viewing of structures and details on a scale as tiny as a nanometer or an angstrom. Because of this, it is an extremely useful instrument for study in the fields of materials science, cell biology, molecular structure analysis, and semiconductors [9]. Imaging mode and diffraction mode are the major modes of operation for the transmission electron microscope (TEM) [11]. The image that is created by the transmitted electrons is used by the imaging mode. It is possible to examine either the diffraction pattern or the image depending on how the magnetic lenses are adjusted. The electron diffraction patterns are the primary focus of the diffraction mode, which focuses on the crystal structure [11]. The preparation of samples is an essential part of TEM. To facilitate electron transmission, specimens must have a thickness of between 50 and 100 nanometers (nm)[11]. Staining with substantial amounts of heavy metal salts is required for biological and polymer materials to produce contrast. Imaging of hydrated materials is possible because of specialized methods such as cryo-TEM, which vitrifies the samples [11].

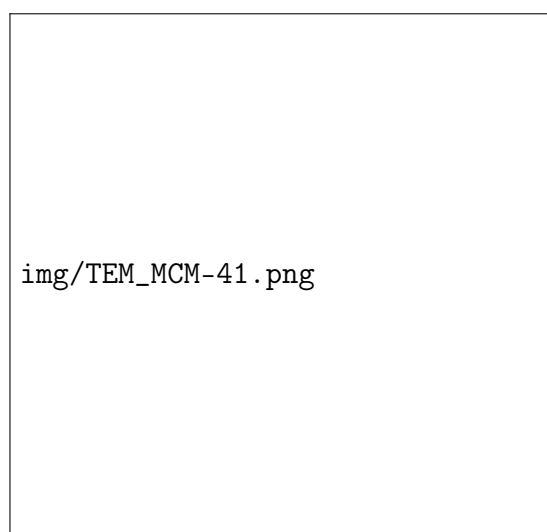


Figure 2.2: TEM images of a mesoporous material(MCM-41) at different magnifications [12]

There is a possibility that radiation will destroy sensitive specimens, which is one of the TEM's limitations [8]. Imaging of living biological samples is likewise not possible due to the vacuum environment. Nevertheless, transmission electron microscopy continues to be an essential instrument for high-resolution structural characterization in both the physical and biological sciences [8]. In this study, transmission electron microscopy (TEM) was used to examine Janus-like particles that were created from block copolymers. Transmission electron microscopy (TEM) gives the resolution and contrast necessary to clearly examine the nanostructure morphology and surface topology of the Janus particles [10].

## 2.0.2 Neural 3D shape representations

Major innovations in deep learning have enabled neural networks to automatically represent and render complex 3D shapes, which was not previously feasible [13]. Neural implicit models can effectively represent shapes by mapping 3D coordinates to occupancy probabilities, signed distance values, or view-dependent radiance [14]. This contrasts with traditional explicit surface and volumetric representations like meshes and voxels.

Early works focused on learning continuous signed distance functions for representing 3D surfaces on synthetic shape datasets [15]. Subsequent techniques aimed to relax the dependence on 3D supervision by formulating differentiable rendering losses that could be optimized using only 2D images [16]. However, these approaches were limited to simplistic and smooth shapes.

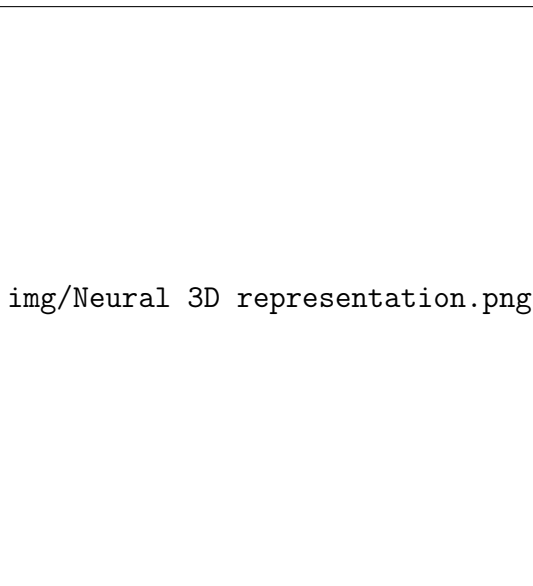


Figure 2.3: DeepSDF’s 3D bunny shape representation [13]

More recently, coordinate-based neural radiance fields have achieved significant improvements in modeling complex real-world 3D scenes [7]. By representing scene properties like volume density and view-dependent emitted radiance as continuous 5D functions, photorealistic novel views of intricate scenes can be rendered.

In this work, we investigate leveraging the capabilities of modern neural 3D representations to reconstruct and denoise volumes from transmission electron microscopy (TEM) tilt series. By training these networks to map from noisy TEM observations to cleaner target volumes in a self-supervised manner, they may learn specialized priors relevant for electron microscopy. The coordinate-based modeling may also better capture critical local context compared to other 3D approaches. This could significantly enhance interpretability of fine structural details from TEM tomograms.

### 2.0.3 Novel view Synthesis

The act of creating fresh photographic perspectives on a subject from one or more input photos is called "view synthesis". This may be done with either a single image or many images. This allows to create unique synthetic viewpoints using only a little amount of photographic data. View synthesis is useful in a variety of contexts, including virtual reality, augmented reality, and the reconstruction of three-dimensional models[17]. Many different techniques have been used for view synthesis. The multi-view stereo approach builds a three-dimensional reconstruction of a scene by piecing together a few photographs obtained with a variety of cameras [18][17]. Then, this model may be displayed from any perspectives. Image-based rendering distorts and interpolates pixels depending on the original inputs to infer new viewpoints [19]. These methods concentrate on identifying correspondences between different pictures.

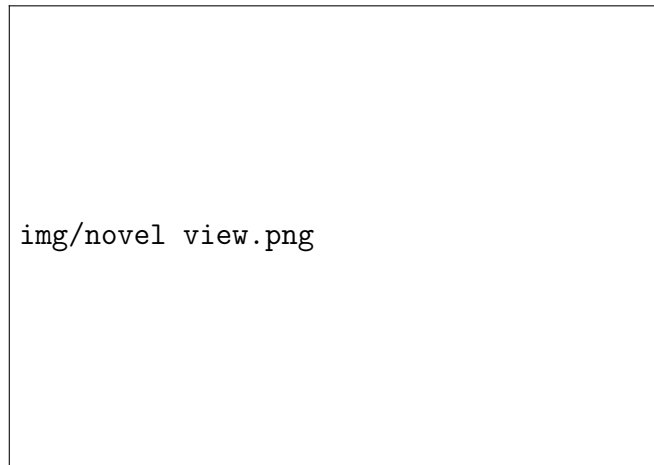


Figure 2.4: Multi-view to Novel view synthesis[20]

The most recent deep learning algorithms develop an implicit representation of the image generation process using neural networks. The neural rendering algorithms directly produce unique views by making predictions about the values of pixels based on the attributes of the scene that they have learnt. Neural radiance fields (NeRF) [7] are a method for efficiently encoding a scene as a continuous five-dimensional function that maps three-dimensional coordinates to volume density and view-dependent brightness[7]. The continuous volumetric scene representation that NeRF provides has made it possible to do photorealistic view synthesis with only a few photos. The capacity to implicitly infer a three-

dimensional structure and appearance from just two-dimensional supervision is the primary benefit offered by neural view synthesis systems. Because of this, formal three-dimensional modeling or estimate is not required. These learning-based systems continue to increase the realism and flexibility of new view creation across a wide variety of applications, including augmented reality, virtual tourism, and 3D photography [21].



## 2.0.4 Camera Parameters

The geometric and optical properties of a camera are referred to as its camera parameters. These parameters define how a camera constructs a picture from the 3D world [20]. Understanding the process of picture generation as well as the tasks involved in 3D computer vision relies heavily on an accurate representation of these factors.

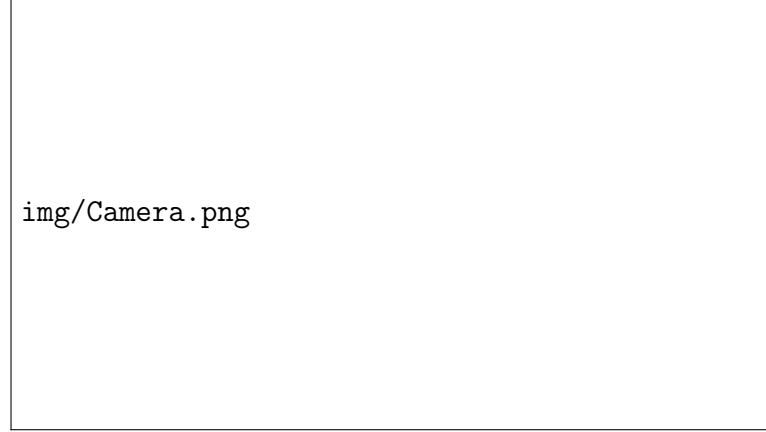


Figure 2.5: Schematic Representation of Camera Parameters. Source: OpenMVG Documentation.

**Intrinsic Parameters** are those that are unique to a camera and are not affected by the scene:

- **Focal Length ( $f$ ):** The distance from the optical center to the image plane when the image is sharp. A primary component that determines both the field of view and the magnification [22]. When dealing with non-square pixels, the x and y axes may have unique values.
- **Principal Point ( $c_x, c_y$ ):** The coordinates ( $c_x, c_y$ ) of the image center on the sensor plane. It accounts for lenses that are not perfectly aligned [22].
- **Skew Coefficient ( $\alpha$ ):** A rotation of the axis between the pixel grid and the sensor that considers non-rectangular pixel shapes [22]. Produces a shearing transformation when applied.
- **Distortion Coefficients:** This model simulates optical distortions such as radial, tangential, and narrow prism effects. Radial is the most noticeable and gives an impression like a barrel or pincushion [22].

The intrinsic camera matrix  $K$  can be represented as:

$$K = \begin{bmatrix} f_x & \alpha & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (2.1)$$

where  $f_x$  and  $f_y$  are the focal lengths expressed in pixel units.

**Extrinsic Parameters** are determined by the position of the camera in relation to the world:

- **Rotation Matrix ( $R$ )**: A 3x3 matrix describing the camera's orientation in a world coordinate system [23]. Represented by a sequence of rotations (Euler angles).
- **Translation Vector ( $\mathbf{T}$ )**: A vector defining the position of the camera's origin in world coordinates [23].

The extrinsic parameters are combined into a 3x4 matrix  $[R \mid \mathbf{T}]$ , defining the transformation from world coordinates to camera coordinates.

Together, intrinsic and extrinsic parameters form the camera projection matrix  $P$ , which is responsible for mapping 3D world points into 2D picture coordinates:

$$P = K \times [R \mid \mathbf{T}] \quad (2.2)$$

For computer vision applications like posture estimation, 3D reconstruction, and unique view synthesis, accurate assessment of these parameters is essential.

Applications in augmented reality [24], autonomous navigation [24], and computational photography [24] heavily rely on precise camera calibration. Adapting camera models to new modalities such as light field imaging continues to be an active area of research.

### 2.0.5 COLMAP

COLMAP is an open-source pipeline that uses structure-from-motion (SfM) and multi-view stereo (MVS) to generate 3D models from 2D images [25]. Through solid correspondence construction, global optimization, and volumetric fusion, it features state-of-the-art reconstructions.

- **Feature Extraction and Matching**

First, appearance-based image features that can be paired between views are found and described. Based on local gradients, SIFT is frequently used to locate scale- and rotation-invariant key points [26]. Each key point has a high-dimensional descriptor vector that is insensitive to noise, perspective, and illumination [26]. Based on similarity measures like Euclidean or

cosine distance, an effective closest neighbor search matched characteristics between image pairings. Uncertain matches can be eliminated with the ratio test [26]. Outlier matches that are inconsistent with a single 3D point are eliminated by geometric verification using RANSAC [27].

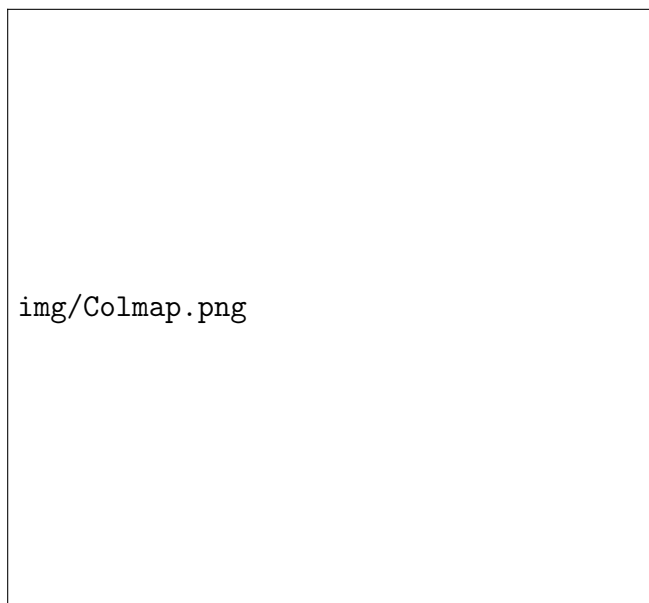


Figure 2.6: COLMAP camera position extracting from set of 2D images

- **Incremental Structure from Motion (SfM)**

In an incremental SfM method, the registered 2D-2D matches create initial sparse 3D point clouds [28]. An initial point cloud is plotted using an initial image pair. Which views to update next are efficiently chosen by robust visibility constraints [28]. With points recursively mapped from fresh views [29], camera poses are predicted using a Straight Linear Transform within a RANSAC cycle.

- **Global SfM Optimization**

Utilizing bundle adjustment, the progressive reconstruction is globally improved to simultaneously improve camera poses and 3D point coordinates.

Scale drift is reduced with regularization. Bundle adjustment reduces the top view error between the positions of anticipated and actual 2D features in all perspectives [30]. This enhances accuracy and comprehensiveness.

- **Multi-View Stereo (MVS) Depth Map Estimation**

The optimal cameras and points start the estimate of the multi-view stereo depth map. Using photo consistency metrics such as normalized cross correlation between distorted picture patches, dense correspondence is created in each view [31]. Accuracy is improved by regularization using filtering such as Gaussian smoothing [32]. The per-view depth maps that are constructed include geometric detail.

- **Surface Reconstruction**

When creating a final 3D surface mesh, volumetric fusion methods such as screening Poisson reconstruction are utilized to merge the depth data to produce the mesh [33]. It accomplishes this by interpolating an indicator function to provide a continuous and smooth surface. Additional post-processing steps, such as graph cuts-based optimization [34], may be utilized to improve details even further. Realism and color are added when texturing with the use of input images.

### 2.0.6 NeRF (Neural Radiance Field)

Neural radiance fields (NeRF) are a recent breakthrough technique for novel view synthesis and 3D scene modeling using implicit neural representations [7]. NeRF represents a scene as a continuous 5D radiance field (3D position + 2D viewing direction) using a multilayer perceptron (MLP).



img/NERF.png

Figure 2.7: Conceptual Illustration of Neural Radiance Fields (NeRF) [7]

---

#### Algorithm 1 Training Neural Radiance Field (NeRF)

---

```

1: Initialize MLP with random weights.
2: for each training iteration do
3:   Select a subset of input images.
4:   for each selected image do
5:     Sample a set of rays passing through image pixels.
6:     for each ray do
7:       Sample points along the ray.
8:       Query MLP for color and density at each point.
9:       Calculate rendered color of the ray using volume rendering.
10:    end for
11:    Calculate loss between rendered and actual pixel colors.
12:    Update MLP weights to minimize the loss.
13:  end for
14: end for

```

---

The MLP maps each (x,y,z) location in space to an RGB color value and volume density scalar. The color indicates the emitted radiance, while density encodes occlusion. Querying this MLP at sampled points along camera rays enables volumetric ray marching to render novel views. The integral of density \* color approximates the total radiance along each ray. Compared to discrete voxel grids or meshes, the continuous coordinate-based modeling better captures smooth variations in structure, appearance, and lighting. The MLP can represent complex scenes in a memory-efficient compact latent code rather than an explicit 3D model. Adjusting MLP weights based on rendered and real image differences allows optimizing the scene representation.

Key advantages of the NeRF approach include:

- Coordinate MLPs effectively model local relationships.

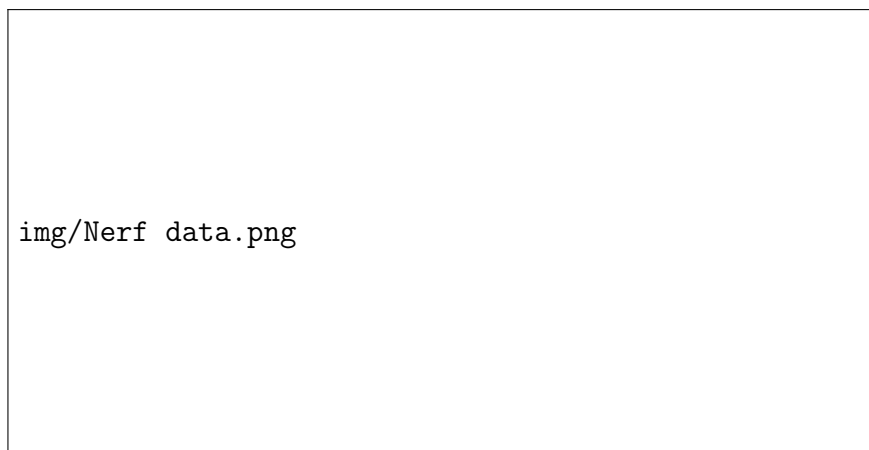


Figure 2.8: Dataset of 2D Images from Various Camera Positions

- Continuous representation enables high quality view interpolation.
- Volume density handles complex occlusion effects.
- View-direction encoding models view-dependent phenomena like highlights.

NeRF obtains impressive results in reconstructing 3D scenes from only a sparse set of input views (e.g. 20-50 images). However, it relies on accurate camera pose estimation and clean photographic inputs. The utility for noisy domains like biomedical imaging is still being explored. For training purposes, you will need the camera parameters for the images that have been provided. These camera parameters are typically computed by SfM tools like COLMAP.



Figure 2.9: Still Frames from Video Generated Using 2D Images with New Camera Angle

## Self-Supervised Learning in NeRF for TEM Data

Self-supervised learning has emerged as a powerful paradigm for training neural networks when labeled data is scarce. Without requiring manual annotations, self-supervision facilitates flexible utilization of intrinsic structure within the data itself to guide representation learning [35]. We investigate a self-supervised framework for training neural radiance fields on transmission electron microscopy (TEM) data that is tailored to handle key imaging characteristics.

Conventionally, NeRF models are trained using datasets of posed 2D images with associated camera parameters. This enables supervision via view reconstruction losses between rendered and real images [7]. However obtaining accurate poses can be extremely challenging for modalities like TEM where standard structure-from-motion approaches fail due to low signal-to-noise and missing angular coverage. Self-supervision circumvents this issue by instead leveraging principles of geometry [36] and appearance consistency [36] across the input views for supervision.

Specifically, we propose training NeRF to synthesize target images from the tilt series using only a subset of the remaining images. This requires inferring a consistent underlying 3D representation aligning with all images. We randomly sample training directions, and predict images along the held-out directions by volumetric raymarching through the MLP scene representation. The reconstruction loss comparing predicted and true images acts as self-supervision for 3D reasoning. By learning to predict complete volumes from extremely sparse noisy data in this manner, NeRF implicitly learns priors tuned for electron tomography.

The continuous coordinate-based modeling of NeRF may capture finer-grained 3D relationships compared to voxel or 2D approaches [15]. Global context aggregation can also inform local details. We analyze design choices around ray sampling patterns, reconstruction objectives, conditioning schemes and model capacity tailored for this self-supervised formulation. Experiments validate our approach reconstructing higher fidelity volumes from few input tilts compared to state-of-the-art alternatives. This demonstrates the potential of self-supervised coordinate-based networks to better address key challenges for computational TEM analysis.

### 2.0.7 Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN)

An overview of Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN) [37] shows that this technology is a major step forward in the super-resolution of images. This deep learning model is an essential tool in industries that need high-resolution imaging because it is made to boost image resolution beyond what the sensor can handle. ESRGAN is a development of the previous Super-Resolution GAN (SRGAN) [38], but it has undergone significant improvements that enable it to produce images with finer features and more realistic textures [39].



The image placeholder shows the ESRGAN Architecture diagram, which is a complex neural network structure. It typically includes an encoder-decoder architecture with skip connections, residual blocks (RRDBs), and a discriminator. The diagram illustrates how the network takes a low-resolution input and produces a high-resolution output by learning to generate fine details.

Figure 2.10: ESRGAN Architecture [37]

The implementation of Residual-in-Residual Dense Blocks (RRDB) to enhance the model's depth and parameters and make optimization simpler is one of ESRGAN's primary advantages over SRGAN. Additionally, it introduces the Relativistic GAN (RaGAN) loss, which helps to improve texture details and sharpen edges. Furthermore, ESRGAN uses a perceptual loss that makes use of pretrained features from VGG networks, which improves the super-resolved images' perceived quality [37].

Especially useful in Transmission Electron Microscopy (TEM) context, ESRGAN may infer high-resolution images four times larger than the input [37]. The more realistic textures of the model are essential for improving medical images and computational microscopy. The ability of ESRGAN to synthesize fine details is utilized as a post-processing step for processed TEM data in our thesis study. In order to improve the quality of TEM images, where the clarity of microscopic structures and details is crucial, this application is essential.

Our thesis tackles the problem of enhancing the resolution and perceived quality of TEM pictures through the implementation of ESRGAN. Our goal in using ESRGAN is to improve these images' diagnostic utility and aesthetic appeal so that they may be used for more in-depth examination and interpretation. One example of how cutting-edge deep learning techniques are being used in practice to improve scientific imaging is the incorporation of ESRGAN into the pipeline for processing TEM pictures. This method advances the science of microscopy while also establishing a standard for the use of super-resolution techniques in other fields where improved picture quality and resolution are necessary.



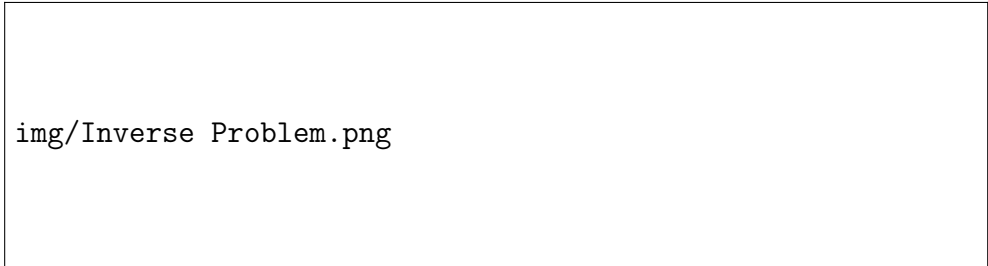
## CHAPTER 3

# Related Work

---

### 3.0.1 Inverse problem

The process of reconstructing 3D structures from TEM tilt series is a challenging and ill-posed inverse problem. In this process, a volume is reconstructed from finite and noisy 2D projections by inverting the advanced visualization model, which projects the 3D structure to the observed dimensions [40]. This allows the volume to be recovered from the 2D projections. Traditional algorithms, such as back projection, are incapable of fully addressing the ill-posed Ness of the problem because of its structural dimension.



img/Inverse Problem.png

Figure 3.1: Comparative Reconstruction of a 10 nm Diameter Gold Particle: (A) Using Linear Backprojection in IMOD Software, showing deviation from spherical shape. (B) Using Curvilinear Backprojection in TxBR, accurately representing a spherical particle. [41]

Regularized iterative reconstruction approaches have been created as a means of overcoming restrictions that are unique to analytical methods [42]. These strategies integrate past knowledge in order to limit the space available for solutions. Total variation (TV), a technique for regularization that creates reconstructions with smoothed intensity variations while keeping edges [42], is a technique that is extensively utilized. Compressive sensing is an additional method that makes use of sparsity in transform domains such as gradients or wavelets [43]. In addition, patch-based sparse coding methods have been utilized in order to discover an exhaustive collection of local basis functions for the purpose of denoising [44]. However, when the amount of data grows larger, these methods experience a rise in the computational burdens they must bear, and choosing the appropriate regularization parameters becomes a process that is not easy to do.

The inclusion of a missing wedge, which results from the limited tilt range, low signal-to-noise ratios, and high volumes of the 3D data [45], further complicates the TEM reconstruction process. In order to overcome these obstacles, more robust regularization strategies that are specifically adapted to the imaging physics of TEM need to be developed. Estimating the point spread function of the microscope using model-based methods has been suggested as a way to invert the blurring effects of the instrument [45].

To enhance the quality of the reconstruction, one method that does so by capitalizing on the self-similarity that exists between blocks is known as non-local means filtering [41]. In recent years, learning-based algorithms, such as dictionary learning and deep convolutional networks [44], have demonstrated promising results as post-processing filters or end-to-end reconstruction methods.

In the realm of TEM reconstruction, the development of more recent deep 3D representations, such as generative adversarial networks and neural radiance/volumetric fields [46], has made room for the introduction of new prospects. These methods involve training networks to map real noisy TEM projections to cleaner target volumes, which enables the networks to implicitly encode appropriate priors for accurate reconstruction. However, in order for these methods to work, a substantial amount of training data that covers the entire spectrum of viewing angles is required. Despite these developments, modeling the process of TEM image creation and adapting regularization limitations are still open difficulties that need to be addressed. There is a possibility that hybrid approaches, which combine model-based reconstruction with learned regularization techniques, could provide a more resilient solution.

### 3.0.2 Atomic Resolution

Atomic-resolution transmission electron microscopy (TEM) has emerged as a revolutionary technique for materials characterization by directly imaging individual atoms [47]. Modern TEMs use aberration correctors and monochromators that eliminate chromatic blurring and lens flaws to achieve sub-angstrom resolution [48]. Advanced detectors and highly stable devices have created new opportunities for quantitative analysis [49].

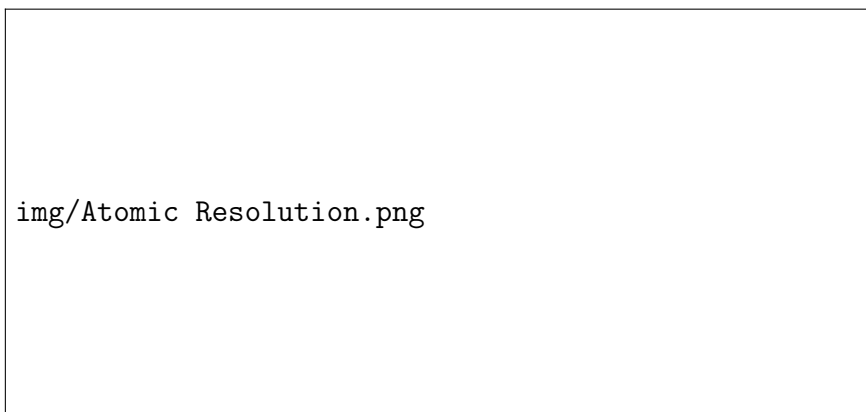


Figure 3.2: Atomic Resolution 3D Imaging of Platinum Nanoparticle: (A) STEM image showing flat twin boundaries. (B) AET reconstruction reveals atomic steps at twin boundaries, with a grain boundary and stacking fault. [50]

On atomic structures, many imaging techniques offer complimentary information. Atomic number contrast images are created through high angle annular dark field (HAADF) imaging, where the intensity scales with  $Z^2$  [50]. Elemental distributions are mapped by atomically resolved energy dispersive X-ray (EDX) spectroscopy. High resolution TEM can image light elements and even depict atom columns. For high precision data, scanning TEM (STEM) raster scans a focused probe [51]. High beam currents within a tiny probe are necessary for atomically detailed imaging, though. As a result, the electron dosages exceed what many materials can tolerate in terms of radiation [52]. The fundamentally probabilistic electron-sample interactions thus mask important atomic organization details with quantum noise [53]. Robust denoising techniques designed for TEM are necessary to obtain quantitative data.

Through statistical post-processing, techniques like multi-frame averaging [51] and principal component analysis enhance signal-to-noise. Sparse regularization techniques take advantage of structural redundancy to reduce noise. Compact representations for denoising are discovered through dictionary learning [50]. Convolutional neural networks have most recently demonstrated the ability to learn potent priors from atomistic image simulations [50]. In general, the developments that have been made in aberration corrected TEM have effectively actualized single-atom sensitivity and precision [53]. Researchers now have capabilities never before seen to unearth new insights through quantitative atomic-scale characterization. These skills are made possible with the assistance of specific denoising techniques, which help researchers overcome resolution restrictions imposed by noise.

### 3.0.3 Noise Modeling

Accurately modeling and characterizing noise is critical for developing effective reconstruction and denoising methods for transmission electron microscopy (TEM) images. Multiple studies have investigated the noise properties and sources in TEM.

- **Shot Noise**

One of the primary sources of noise is shot noise stemming from the quantum nature of electrons and the stochastic process of electron-sample interaction. Reconstruction and denoising techniques for transmission electron microscopy (TEM) images require precise noise characterization and modeling. The sources and characteristics of noise in TEM have been the subject of numerous investigations [54][55]. The number of electrons scattered from each part of the specimen fluctuates, leading to signal-dependent shot noise. Robust statistical distributions capture this behavior.

- **Detector Noise**

TEM detectors also introduce additional noise from readout electronics and amplification [56]. On CCD cameras, dark current shot noise and readout noise are present. Scintillator-photomultiplier detectors show signal-dependent Poisson noise characteristics [57]. Accurate detector models enable simulation of cumulative noise.

- **Beam Current Noise**

Fluctuations in beam current and brightness over time also contribute noise in TEM imaging [58]. Monitoring beam current during acquisition allows normalization to reduce this noise [58] [59]. But residual fluctuations persist and should be incorporated into models.

- **Noise Texture**

The microscope point spread function and optical transfer function modulate the texture of noise in the images. Accurately modeling these effects based on system parameters enables generating realistic synthetic noise for training machine learning models [60].

- **Multiresolution Modeling**

Noise also exhibits signal-dependency and non-stationarity over spatial frequencies [61][62]. Variance stabilization using multiresolution transforms has been proposed to normalize noise over different scales [63].

Overall, rigorous characterization and modeling of the multiple noise sources and their interactions is key to developing optimized TEM reconstruction and restoration techniques. Both model-based and learning-based methods benefit from accurate noise models matched to real TEM imaging.

### 3.0.4 3D Convolutional Neural Networks

3D Convolutional Neural Networks (CNNs) are a powerful tool that leverage the unique characteristics of 3D context and convolution operations to perform a wide array of tasks such as segmentation, classification, and reconstruction of volumetric data. Among the various architectures available, the 3D U-Net has demonstrated exceptional performance, particularly in the field of medical image analysis, by utilizing encoder-decoder convolutions [64].

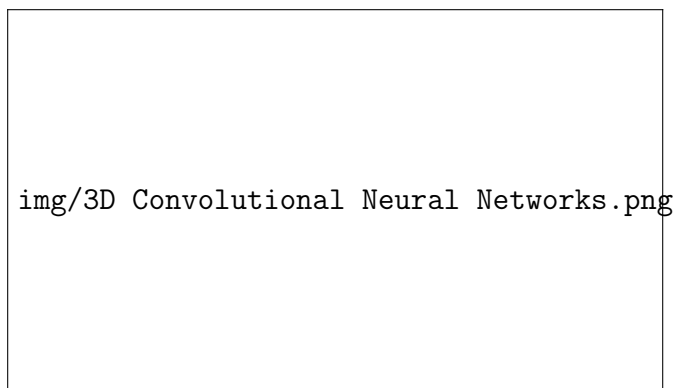


Figure 3.3: First row - Original real images, Second row - Noisier versions, Third row - Images after denoising with CNN DAE. [64]

One of the key benefits of 3D CNNs is their ability to act as data-driven filters that can denoise Transmission Electron Microscopy (TEM) volumes, thus enhancing the interpretability of these volumes. For instance, 3D convolutional autoencoders that have been specifically trained to reconstruct TEM data can effectively serve as noise suppression filters [65]. The process of applying 3D CNN denoising before proceeding with coordinate-based Multilayer Perceptron (MLP) modeling may significantly help condition the data, making it more suitable for further analysis.

Combining Volumetric CNNs and Coordinate MLPs for Improved Performance In recent years, hybrid methods that merge the functionalities of volumetric CNNs and coordinate MLPs have been developed, demonstrating potential for improved reconstruction outcomes by leveraging their complementary strengths [46] [44]. In such a hybrid approach, volumetric CNN encoders first aggregate global context from the 3D input data. Subsequently, coordinate-based MLP decoders model local relationships at each individual location.

This unique combination enables the joint learning of multi-scale representations, in which the CNN provides top-down semantic guidance, while the MLP preserves the bottom-up spatial details. In the context of TEM data, this dual approach could effectively capture both anatomical priors and the fine structural variations that are typically present. The global-local modeling provided by this combination may enable accurate reconstruction from sparse, noisy tilt series projections, thereby potentially revolutionizing the way we handle and interpret such data.

### 3.0.5 Denoising

Reducing noise in transmission electron microscopy (TEM) images is critical for enabling accurate reconstruction and analysis. However, the low electron doses used in TEM result in extremely low signal-to-noise ratios. Conventional linear filters like Gaussian smoothing remove noise at the expense of blurred structural details. More advanced model-based methods are not robust to non-Gaussian noise encountered in TEM.

A variety of denoising methods have been developed for TEM images, including median filtering, Wiener filtering, wavelet transform-based denoising, and deep learning-based denoising [66]. Median filtering is a simple and efficient method, but it can blur image edges. Wiener filtering can preserve image edges, but it can be computationally expensive. Wavelet transform-based denoising is a good compromise between efficiency and image quality [66].

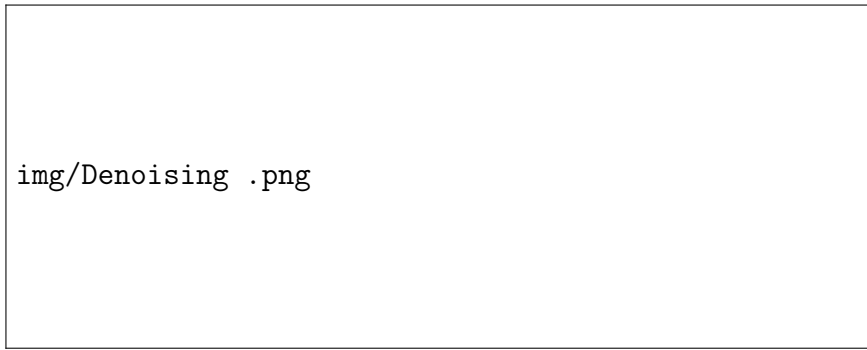


Figure 3.4: TEM Imaging and Denoising of Cadmium Sulfide Nanoparticles: **A.** Original Image, **B.** Image with Gaussian Noise, **C.** Denoised with Average Filter, **D.** Denoised with Median Filter, **E.** Denoised with Weiner Filter. [66]

For denoising, sparse coding techniques take advantage of priors in the image such as non-local self-similarity and local sparsity. Sparsifying transforms are used by methods such as K-SVD [67] and BM3D to group comparable patches and filter noise. Computational expenses do not scale well with the magnitude of medical images, despite being effective. There are difficulties in choosing regularization parameters optimally.

Deep learning approaches have recently shown great promise for image denoising by learning data-driven filters [44]. Convolutional networks trained as discriminators between clean and noisy image patches can implicitly model complex image priors. Recurrent inference further boosts quality[44]. Autoencoder architectures directly optimize reconstruction loss [44]. Multi-image network training leverages complementary information across tilt series.

Applying and tailoring deep denoisers to 3D TEM data could significantly enhance reconstruction quality from noisy tilt projections. The high capacity of deep networks may better capture noise characteristics compared to hand-crafted models. Overall, learned denoising provides new opportunities to overcome resolution limits imposed by noise in TEM imaging.

### 3.0.6 View Synthesis and Image-based rendering

View synthesis refers to rendering novel views of a scene from limited input images. Traditional image-based rendering (IBR) approaches extrapolate new perspectives by warping and interpolating input images based on estimated dense correspondence [19]. These methods rely on accurate multi-view stereo matching.

More recent learning-based strategies use neural networks to implicitly infer scene structure and appearance for high quality view synthesis [68]. Neural radiance fields (NeRF) have shown promising results by encoding scenes as continuous 5D radiance fields with MLPs [7]. The coordinate-based volumetric scene representation better captures smooth variations compared to voxels.

Applying view synthesis techniques to TEM could significantly improve 3D volume reconstruction from sparse 2D tilt series projections. The continuous modeling of NeRF may effectively represent structural priors and noise characteristics to enable generating new perspectives. However, scaling to large TEM datasets remains challenging.

Among view synthesis techniques, NeRF provides an appealing approach for TEM reconstruction by leveraging MLPs to model local relationships in a coordinate-based volumetric domain. In this work, we propose adapting NeRF to learn mappings between real noisy TEM tilt inputs and target volumes in a self-supervised manner. The noise-aware coordinate modeling could address key challenges in TEM imaging.





## CHAPTER 4

# Datasets

---

This chapter provides an overview of the datasets utilized for experimental validation in this work. Three main types of data are used: real transmission electron microscopy (TEM) images, real scanning TEM (STEM) images, and synthetically generated volumes. Details on the source, composition, imaging parameters, and pre-processing of each dataset are provided. Both real and synthetic data are essential for robust training, evaluation and analysis of the proposed methods.

### 4.0.1 TEM Datasets

**Source for all TEM datasets:** CAU Technical Faculty (Synthesis and Real Structure Group)

- **Dataset 1**

<b>Total Image</b>	10
<b>Dimensions</b>	639 X 639 to 1496 X 1496
<b>Format</b>	BMP
<b>Image Size</b>	20KB to 280KB

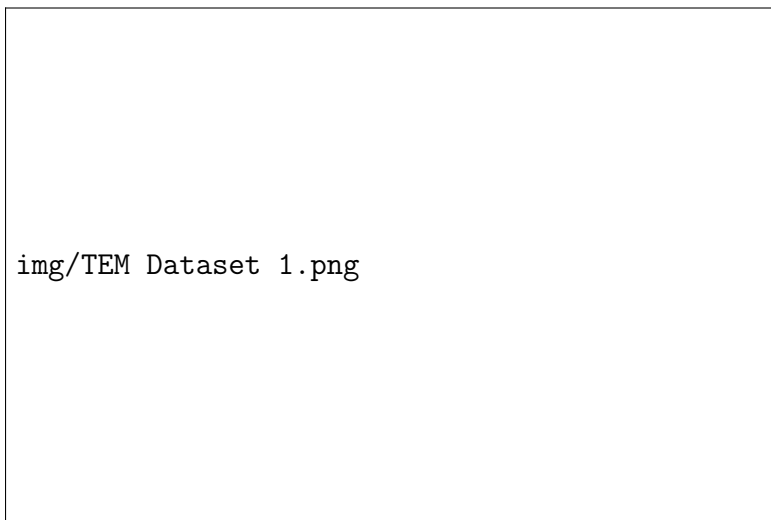


Figure 4.1: TEM Dataset 1

**Listing 1** Resize all images to match the dimensions of the first image

---

```

1      import cv2
2      import numpy as np
3
4      for idx, file_name in enumerate( grayscale_files ):
5          grayscale_img =
            ↳ cv2.imread( os.path.join( folder_path,
            ↳ file_name ), cv2.IMREAD_GRAYSCALE )
6
7          # If it's the first image, store its dimensions
8          if idx == 0:
9              first_image_height, first_image_width, _ =
            ↳ enhanced_img.shape
10
11         enhanced_img = cv2.resize( enhanced_img,
            ↳ (first_image_width, first_image_height) )

```

---

To prepare the dataset for NeRF, which requires JPG format, a script converts all BMP images in the folder to JPG. Given the dataset's ten images of varying dimensions, preprocessing is vital to standardizing their shape and aligning all images to the sizes of the first image. This uniformity is crucial for reducing noise and achieving consistent, reliable experiment results.

- **Dataset 2**

<b>Total Image</b>	23
<b>Dimensions</b>	2048 X 2048
<b>Format</b>	BMP
<b>Image Size</b>	123KB

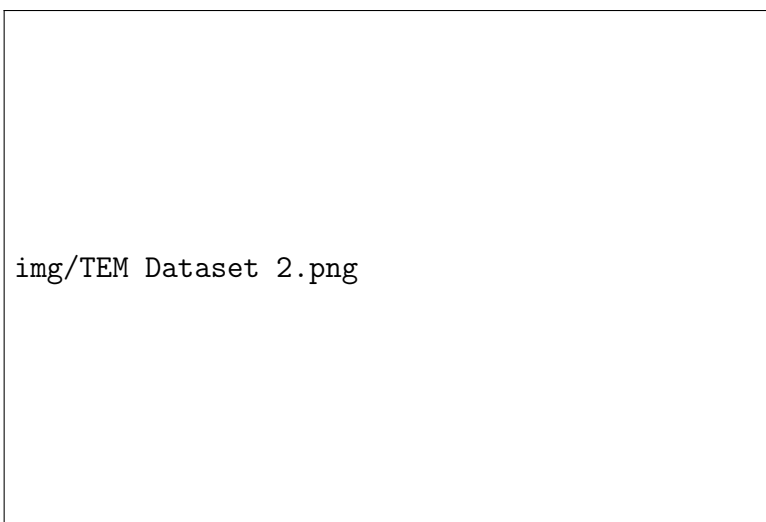


Figure 4.2: TEM Dataset 2

- Dataset 3

<b>Total Image</b>	48
<b>Dimensions</b>	256 X 256
<b>Format</b>	BMP
<b>Image Size</b>	350KB

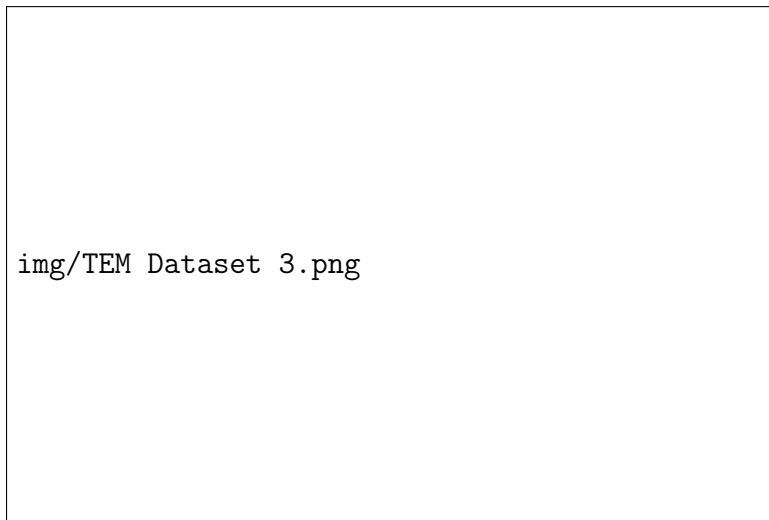


Figure 4.3: TEM Dataset 3

- Dataset 4

<b>Total Image</b>	20
<b>Dimensions</b>	1421 X 1421
<b>Format</b>	JPG
<b>Image Size</b>	350KB

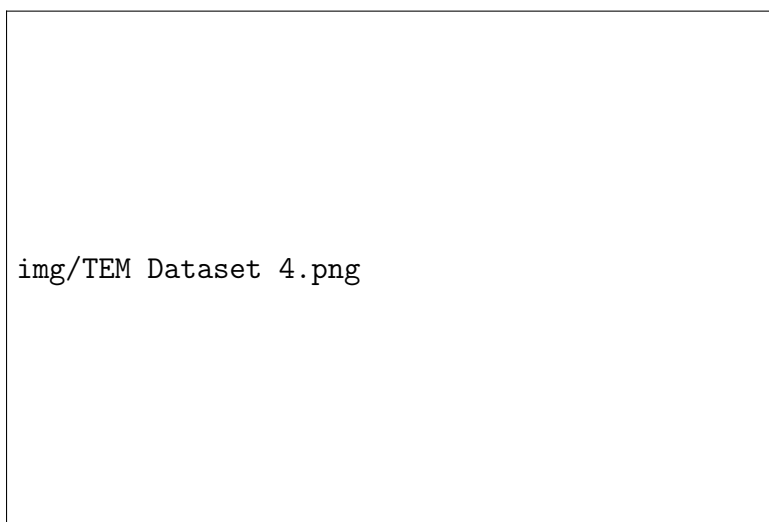


Figure 4.4: TEM Dataset 4

### 4.0.2 STEM Datasets

**Source for all STEM datasets:** CAU Technical Faculty (Synthesis and Real Structure Group)

- **Dataset 1**

<b>Total Image</b>	17
<b>Dimensions</b>	1024 X 1024
<b>Format</b>	JPG
<b>Image Size</b>	1MB

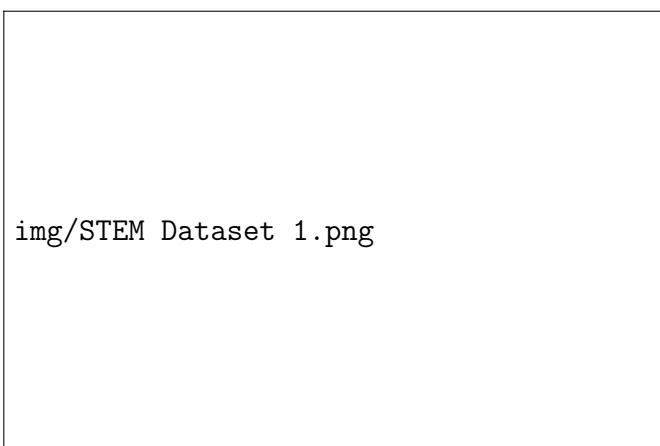


Figure 4.5: STEM Dataset 1

- **Dataset 2**

<b>Total Image</b>	16
<b>Dimensions</b>	1024 X 1024
<b>Format</b>	JPG
<b>Image Size</b>	80KB

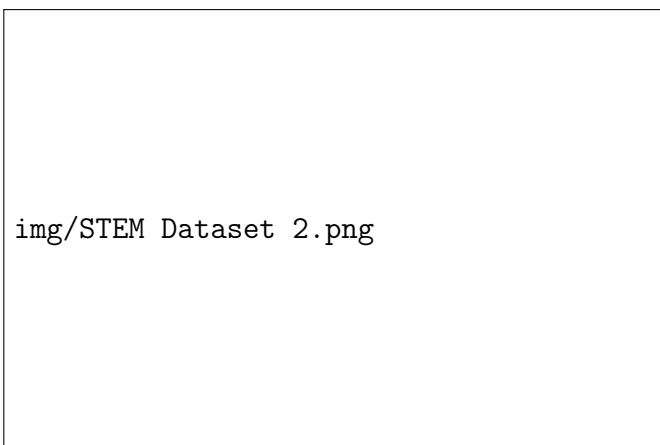


Figure 4.6: STEM Dataset 2

### 4.0.3 Synthetic Datasets

- Dataset 1

<b>Source</b>	Blender Generated Data
<b>Total Image</b>	120
<b>Dimensions</b>	800 X 800
<b>Format</b>	JPG
<b>Image Size</b>	80KB

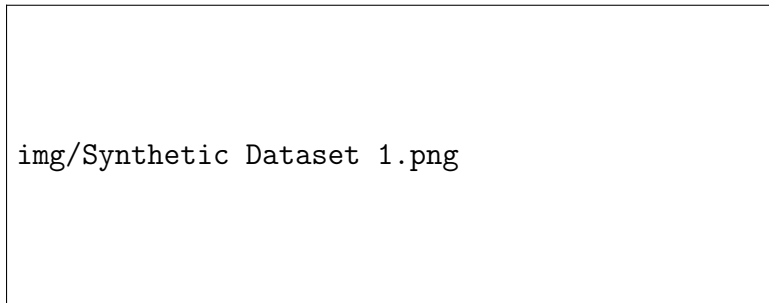


Figure 4.7: Synthetic Dataset

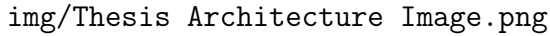


## CHAPTER 5

# Implementation

---

This section provides a comprehensive overview of the key steps taken to implement our integrated framework for denoising and analysis of TEM images, as depicted in Figure 5.1. We build upon recent advances in Neural Radiance Fields, specifically the NeRF-MM architecture along with techniques from NanoNeRF and NeRF-Dark, to address the noise and imaging challenges associated with TEM data. It outlines the development of these methods, specifically aimed at addressing denoise TEM Images. The implementation process is structured to illustrate each step, from initial environment setup to final execution. Emphasis is placed on key parts of the code and their contributions to the overall methodology, ensuring a comprehensive understanding of the applied techniques and how they integrate to achieve the objectives of this thesis.



img/Thesis Architecture Image.png

Figure 5.1: Comprehensive Overview of the our Framework: Integrating NeRF and Advanced Denoising Techniques for Enhanced TEM Image Analysis

### 5.0.1 Setup Overview

The NeRF-MM implementation was conducted within a Python environment, leveraging crucial libraries such as Torch and NumPy. Initial steps included setting the `CUDA_LAUNCH_BLOCKING=1` environment variable for synchronized CUDA operations and cloning the NeRF-MM repository from GitHub. To ensure the reproducibility of results, seed values for random, NumPy, and Torch were carefully set, and the deterministic behavior of PyTorch’s backend (CUDNN) was enabled. The implementation extensively used core modules and utility functions from the NeRF-MM library, notably `encode_position` and `volume_rendering`, which play a vital role in positional encoding and volume operations within the neural radiance field framework.

### 5.0.2 Key Implementation Steps

The primary steps involved in the NeRF-MM implementation were as follows:

1. **Data Preprocessing:** The data preparation phase involved importing and processing the dataset to meet the specific requirements of our implementation. A significant challenge encountered was the format of the TEM images; they were initially in BMP format and grayscale. To align with our implementation’s needs, these images were converted to either JPG or PNG format and transformed into RGB color space. Additionally, the dataset contained images of varying dimensions, which could potentially lead to confusion during model training. To address this, a Python script was employed to preprocess the images, ensuring uniformity in shape and size across each dataset. This preprocessing step was crucial to maintain consistency in the data and facilitate effective training of the model.
2. **Model Architecture:** We enhanced the TinyNerf neural network class from NerfMM[69] to assimilate advantageous features from NanoNerf [70] and Nerf-Dark [71] models. This forms the backbone for learning an optimized radiance field from noisy TEM data.

Specifically, we updated TinyNerf by modifying the number and width of layers, activation functions, and other hyperparameters based on the configurations found to work best in NanoNerf [70] and Nerf-Dark [71]. This helps account for the sparse view sampling and noise characteristics in TEM images.

The adapted TinyNerf architecture manages positional and directional encoding inputs to efficiently compute scene density and RGB color outputs. The adjustments enhance model capacity to capture detail and lighting variations in the radiance field from limited input views.

By synergizing recent neural design improvements, our TinyNerf implementation facilitates more effective representation learning for high-fidelity novel view synthesis from challenging TEM data.

```

1      class TinyNerf(nn.Module):
2          def __init__(self, pos_in_dims, dir_in_dims, D):
3              super(TinyNerf, self).__init__()
4              # Sequential layers definition
5              self.layers0 = nn.Sequential(
6                  nn.Linear(pos_in_dims, D), nn.ReLU(),
7                  # ... [other layers] ...
8                  nn.Linear(D, D), nn.ReLU()
9              )
10             # Definition of density and RGB layers
11             self.fc_density = nn.Linear(D, 1)
12             self.fc_feature = nn.Linear(D, D)
13             self.rgb_layers = nn.Sequential(nn.Linear(D
14             ↪ + dir_in_dims, D//2), nn.ReLU())
15             self.fc_rgb = nn.Linear(D//2, 3)

```



```

15
16     def forward(self, pos_enc, dir_enc):
17         # Forward pass implementation
18         rgb = self.fc_rgb(self.rgb_layers(x))
19         return torch.cat([rgb, density], dim=3)

```

3. **Training and Optimization:** The training and optimization of the NeRF-MM model were critical stages of the implementation. This process involved setting up and training a TinyNerf neural network model, fine-tuning camera pose parameters, and optimizing the focal length. A critical optimization technique implemented was an early stopping mechanism. The training process monitored the updates in camera parameters and halted the training prematurely if these parameters ceased to significantly change after a certain number of epochs. This approach ensured efficiency by avoiding unnecessary computations when the model had converged.

```

1     # Training Loop
2     for epoch_i in range(N_EPOCH):
3         L2_loss = train_one_epoch(...)
4         train_psnr = mse2psnr(L2_loss)
5         writer.add_scalar('train/psnr', train_psnr,
6             ↪ epoch_i)
7         # Early stopping based on camera parameter
8         ↪ updates
9         if check_for_convergence(learned_c2ws):
10             break
11         # Update steps and learning rate scheduling
12         scheduler_nerf.step()
13         scheduler_focal.step()
14         scheduler_pose.step()
15         # Check for convergence is a hypothetical function

```

Additionally, the training loop involved calculating the mean squared error (MSE) loss and peak signal-to-noise ratio (PSNR) for performance evaluation. The learning rate was adjusted using a staircase exponential decay scheduler, which contributed to the effective convergence of the model.

4. **Rendering and Visualization:** This phase focused on rendering 3D views using a spiral camera trajectory, inspired by `nerf_pl`, and visualizing camera poses. The NeRF model rendered novel viewpoints, dynamically representing the scene. To optimize computational efficiency, images were resized without compromising visual quality. Rendered images and depth maps were saved as GIFs for detailed scene analysis. Additionally, camera pose visualization through animated plots provided insights into the model's spatial understanding, dynamically illustrating camera movement over training epochs.

5. **Post-processing:** After generating the GIFs, key post-processing steps were applied to all frames for quality enhancement. This process includes:

- (a) **Enhancing with ESRGAN:** Image enhancement was achieved using the ESRGAN model. This method focuses on upscaling and improving image resolution. The pre-trained ESRGAN model from TensorFlow Hub was utilized for this enhancement.

```

1      # Load a pre-trained model from TensorFlow Hub
2
3      def enhance_image(image_path, save_path):
4          # Convert to TensorFlow tensor and
5          ↪ preprocess
6          rgb_image = tf.cast(rgb_image, tf.float32)
7          ↪ / 255.0
8          rgb_image = tf.expand_dims(rgb_image, 0)
9          # Apply the model
10         enhanced_image = model(rgb_image)
11         enhanced_image = tf.squeeze(enhanced_image,
12         ↪ 0)
13
14         os.makedirs(os.path.dirname(save_path),
15         ↪ exist_ok=True)
16
17         return grayscale_output

```

- (b) **Denoising:** This step involves applying various denoising techniques to reduce noise artifacts. Methods such as wavelet, Gaussian blur, median filter, bilateral filter, and non-local means were used. Each technique was applied to the frames and the denoised images were saved in respective subfolders.

```

1
2      denoising_methods = {
3          'wavelet': denoise_wavelet,
4          'gaussian_blur': lambda img:
5          ↪ cv2.GaussianBlur(img, (5, 5), 0),
6          'median_filter': lambda img: median(img),
7          'bilateral_filter': lambda img:
8          ↪ cv2.bilateralFilter(img, 9, 75, 75),
9          'non_local_means': lambda img:
10         ↪ denoise_nl_means(img, h=1.15 *
11         ↪ np.std(img))
12     }

```

These techniques significantly improve the visual quality of the images, rendering them more suitable for in-depth analysis.

### 5.0.3 Code Availability

The complete source code for this project, including all scripts and detailed documentation, is available on GitHub. It can be accessed at:

<https://github.com/Mithunjack/Thesis-NeRFs>.



# Bibliography

---

- [1] Achilleas S. Frangakis. “It’s noisy out there! A review of denoising techniques in cryo-electron tomography”. In: *Journal of Structural Biology* 213 (4 Dec. 2021), p. 107804. ISSN: 1047-8477. DOI: 10.1016/J.JSB.2021.107804.
- [2] David C. Joy. “Noise and Its Effects on the Low-Voltage SEM”. In: *Biological Low-Voltage Scanning Electron Microscopy* (Nov. 2008), pp. 129–144. DOI: 10.1007/978-0-387-72972-5\_4. URL: [https://link.springer.com/chapter/10.1007/978-0-387-72972-5\\_4](https://link.springer.com/chapter/10.1007/978-0-387-72972-5_4).
- [3] Achilleas S. Frangakis and Reiner Hegerl. “Noise reduction in electron tomographic reconstructions using nonlinear anisotropic diffusion”. In: *Journal of structural biology* 135 (3 2001), pp. 239–250. ISSN: 1047-8477. DOI: 10.1006/JSBI.2001.4406. URL: <https://pubmed.ncbi.nlm.nih.gov/11722164/>.
- [4] Vicente González-Ruiz and Jose Jesus Fernández. “FlowDenoising: Structure-preserving denoising in 3D electron microscopy (3DEM)”. In: *SoftwareX* 23 (July 2023), p. 101413. ISSN: 2352-7110. DOI: 10.1016/J.SOFTX.2023.101413.
- [5] Yunjin Chen and Thomas Pock. “Trainable Nonlinear Reaction Diffusion: A Flexible Framework for Fast and Effective Image Restoration”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39 (6 June 2017), pp. 1256–1272. ISSN: 01628828. DOI: 10.1109/TPAMI.2016.2596743.
- [6] Xin Deng and Pier Luigi Dragotti. “Deep Convolutional Neural Network for Multi-modal Image Restoration and Fusion”. In: ().
- [7] Ben Mildenhall et al. “NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis”. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 12346 LNCS (Mar. 2020), pp. 405–421. ISSN: 16113349. DOI: 10.1007/978-3-030-58452-8\_24. URL: <https://arxiv.org/abs/2003.08934v2>.
- [8] R. F. Egerton, P. Li, and M. Malac. “Radiation damage in the TEM and SEM”. In: *Micron* 35 (6 Aug. 2004), pp. 399–409. ISSN: 0968-4328. DOI: 10.1016/J.MICRON.2004.02.003.

- [9] Baptiste Gault et al. “Estimation of the Reconstruction Parameters for Atom Probe Tomography”. In: *Microscopy and Microanalysis* 14 (4 Aug. 2008), pp. 296–305. ISSN: 1431-9276. DOI: 10.1017/S1431927608080690. URL: <https://dx.doi.org/10.1017/S1431927608080690>.
- [10] C. Y. Tang and Z. Yang. “Transmission Electron Microscopy (TEM)”. In: *Membrane Characterization* (Jan. 2017), pp. 145–159. DOI: 10.1016/B978-0-444-63776-5.00008-5.
- [11] Marc Adrian et al. “Cryo-electron microscopy of viruses”. In: *Nature* 1984/308:5954 308 (5954 1984), pp. 32–36. ISSN: 1476-4687. DOI: 10.1038/308032a0. URL: <https://www.nature.com/articles/308032a0>.
- [12] Alexis Loiseau et al. “Silver-Based Plasmonic Nanoparticles for and Their Use in Biosensing”. In: (2019). DOI: 10.3390/bios9020078. URL: [www.mdpi.com/journal/biosensors](http://www.mdpi.com/journal/biosensors).
- [13] Jeong Joon Park et al. “DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation”. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 2019-June (Jan. 2019), pp. 165–174. ISSN: 10636919. DOI: 10.1109/CVPR.2019.00025. URL: <https://arxiv.org/abs/1901.05103v1>.
- [14] Lars Mescheder et al. “Occupancy Networks: Learning 3D Reconstruction in Function Space”. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 2019-June (Dec. 2018), pp. 4455–4465. ISSN: 10636919. DOI: 10.1109/CVPR.2019.00459. URL: <https://arxiv.org/abs/1812.03828v2>.
- [15] Vincent Sitzmann et al. “Implicit Neural Representations with Periodic Activation Functions”. In: *Advances in Neural Information Processing Systems* 33 (2020), pp. 7462–7473.
- [16] Zhirong Wu et al. “3D ShapeNets: A deep representation for volumetric shapes”. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 07-12-June-2015 (Oct. 2015), pp. 1912–1920. ISSN: 10636919. DOI: 10.1109/CVPR.2015.7298801.
- [17] Weihao Xia and Jing-Hao Xue. “A Survey on 3D-aware Image Synthesis”. In: (). URL: <https://weihaox.github..>
- [18] Steven M. Seitz et al. “A comparison and evaluation of multi-view stereo reconstruction algorithms”. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 1 (2006), pp. 519–526. ISSN: 10636919. DOI: 10.1109/CVPR.2006.19.
- [19] Shenchang Eric Chen and Lance Williams. “View Interpolation for Image Synthesis”. In: *Seminal Graphics Papers: Pushing the Boundaries, Volume 2* (Aug. 2023), pp. 423–432. DOI: 10.1145/3596711.3596757. URL: <https://dl.acm.org/doi/10.1145/3596711.3596757>.
- [20] Richard Hartley and Andrew Zisserman. “Multiple View Geometry in Computer Vision, Second Edition”. In: (2000). URL: [www.cambridge.org/9780521540513](http://www.cambridge.org/9780521540513).

- [21] Jiemin Fang et al. “Fast Dynamic Radiance Fields with Time-Aware Neural Voxels TiNeuVox (ours) 1 min 4 mins 8 mins Sparse Time-View Input Images Time Synthesis View Synthesis Fast Training for Time-View Synthesis D-NeRF Training Time”. In: (). DOI: 10.1145/3550469.3555383. URL: <https://doi.org/10.1145/3550469.3555383>.
- [22] Janne Heikkila and Olli Silven. “Four-step camera calibration procedure with implicit image correction”. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (1997), pp. 1106–1112. ISSN: 10636919. DOI: 10.1109/CVPR.1997.609468.
- [23] Zhengyou Zhang. “A flexible new technique for camera calibration”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22 (11 Nov. 2000), pp. 1330–1334. ISSN: 01628828. DOI: 10.1109/34.888718.
- [24] Zhengyou Zhang et al. “A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry”. In: *Artificial Intelligence* 78 (1-2 Oct. 1995), pp. 87–119. ISSN: 0004-3702. DOI: 10.1016/0004-3702(95)00022-4.
- [25] Johannes L Schönberger and Jan-Michael Frahm. “Structure-from-Motion Revisited”. In: (). URL: <https://github.com/colmap/colmap>.
- [26] David G. Lowe. “Object recognition from local scale-invariant features”. In: *Proceedings of the IEEE International Conference on Computer Vision* 2 (1999), pp. 1150–1157. DOI: 10.1109/ICCV.1999.790410.
- [27] David G. Lowe. “Distinctive image features from scale-invariant keypoints”. In: *International Journal of Computer Vision* 60 (2 Nov. 2004), pp. 91–110. ISSN: 09205691. DOI: 10.1023/B:VISI.0000029664.99615.94/METRICS. URL: <https://link.springer.com/article/10.1023/B:VISI.0000029664.99615.94>.
- [28] Martin A. Fischler and Robert C. Bolles. “Random sample consensus”. In: *Communications of the ACM* 24 (6 June 1981), pp. 381–395. ISSN: 15577317. DOI: 10.1145/358669.358692. URL: <https://dl.acm.org/doi/10.1145/358669.358692>.
- [29] Noah Snavely et al. “Photo Tourism: Exploring Photo Collections in 3D”. In: (). URL: [www.cc.gatech.edu/4d-cities](http://www.cc.gatech.edu/4d-cities).
- [30] Jared Heinly et al. “Reconstructing the World\* in Six Days \*(As Captured by the Yahoo 100 Million Image Dataset)”. In: ().
- [31] Bill Triggs et al. “Bundle Adjustment-A Modern Synthesis”. In: ().
- [32] Silvano Galliani et al. “Massively Parallel Multiview Stereopsis by Surface Normal Diffusion”. In: ().
- [33] Gabriele Facciolo, Carlo de Franchis, and Enric Meinhardt. “MGM: A Significantly More Global Matching for Stereovision”. In: (Dec. 2015), pp. 90.1–90.12. DOI: 10.5244/C.29.90.
- [34] Michael Kazhdan, Johns Hopkins University, and Hugues Hoppe. “Screened Poisson Surface Reconstruction”. In: *ACM Trans. Graph. NN, N, Article NN (Month YYYY)* (). DOI: 10.1145/XXXXXXX.YYYYYYY. URL: <http://doi.acm.org/10.1145/XXXXXXX.YYYYYYY>.

- [35] Alexander Kolesnikov et al. *Revisiting Self-Supervised Visual Representation Learning*. 2019. URL: <https://github.com/google/revisiting-self-supervised>.
- [36] Chiyu Max Jiang et al. “Local Implicit Grid Representations for 3D Scenes”. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (Mar. 2020), pp. 6000–6009. ISSN: 10636919. DOI: 10.1109/CVPR42600.2020.00604. URL: <https://arxiv.org/abs/2003.08981v1>.
- [37] Xintao Wang et al. “ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks”. In: (). URL: <https://github.com/xinntao/ESRGAN..>
- [38] Christian Ledig et al. “Photo-realistic single image super-resolution using a generative adversarial network”. In: *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017* 2017-January (Nov. 2017), pp. 105–114. DOI: 10.1109/CVPR.2017.19.
- [39] Yochai Blau et al. “The 2018 PIRM Challenge on Perceptual Image Super-resolution”. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 11133 LNCS (Sept. 2018), pp. 334–355. ISSN: 16113349. DOI: 10.1007/978-3-030-11021-5\_21. URL: <https://arxiv.org/abs/1809.07517v3>.
- [40] Ruoqian Lin et al. “TEMImageNet Training Library and AtomSegNet Deep-Learning Models for High-Precision Atom Segmentation, Localization, Denoising, and Super-Resolution Processing of Atomic-Resolution Images”. In: (Dec. 2020). URL: <https://arxiv.org/abs/2012.09093v2>.
- [41] Albert Lawrence et al. “Transform-based backprojection for volume reconstruction of large format electron microscope tilt series”. In: *Journal of structural biology* 154 (2 May 2006), pp. 144–167. ISSN: 1047-8477. DOI: 10.1016/J.JSB.2005.12.012. URL: <https://pubmed.ncbi.nlm.nih.gov/16542854/>.
- [42] Christian Widmer et al. “GRED: Graph-Regularized 3D Shape Reconstruction from Highly Anisotropic and Noisy Images”. In: *Signal, Image and Video Processing* 8 (1 Sept. 2013), pp. 41–48. ISSN: 18631711. DOI: 10.1007/s11760-014-0694-8. URL: <https://arxiv.org/abs/1309.4426v1>.
- [43] C. O.S. Sorzano et al. “Transfer function restoration in 3D electron microscopy via iterative data refinement”. In: *Physics in medicine and biology* 49 (4 Feb. 2004), pp. 509–522. ISSN: 0031-9155. DOI: 10.1088/0031-9155/49/4/003. URL: <https://pubmed.ncbi.nlm.nih.gov/15005161/>.
- [44] Kai Zhang et al. “Learning Deep CNN Denoiser Prior for Image Restoration”. In: *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017* 2017-January (Apr. 2017), pp. 2808–2817. DOI: 10.1109/CVPR.2017.300. URL: <https://arxiv.org/abs/1704.03264v1>.
- [45] Jose Jesus Fernandez. “Computational methods for electron tomography”. In: *Micron* 43 (10 Oct. 2012), pp. 1010–1030. ISSN: 0968-4328. DOI: 10.1016/J.MICRON.2012.05.003.



- [46] Gaby N. Moawad et al. “Augmented Realities, Artificial Intelligence, and Machine Learning: Clinical Implications and How Technology Is Shaping the Future of Medicine”. In: *Journal of Clinical Medicine* 9 (12 Dec. 2020), pp. 1–7. ISSN: 20770383. DOI: 10.3390/JCM9123811. URL: /pmc/articles/PMC7761251/%20/pmc/articles/PMC7761251/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7761251/.
- [47] Kazuaki Kawahara et al. “Atomic-resolution STEM image denoising by total variation regularization”. In: *Microscopy (Oxford, England)* 71 (5 Oct. 2022), pp. 302–310. ISSN: 2050-5701. DOI: 10.1093/JMICRO/DFAC032. URL: https://pubmed.ncbi.nlm.nih.gov/35713554/.
- [48] Ondrej L. Krivanek et al. “Atom-by-atom structural and chemical analysis by annular dark-field electron microscopy”. In: *Nature* 464 (7288 Mar. 2010), pp. 571–574. ISSN: 1476-4687. DOI: 10.1038/NATURE08879. URL: https://pubmed.ncbi.nlm.nih.gov/20336141/.
- [49] Daniel S. Wastl, Alfred J. Weymouth, and Franz J. Giessibl. “Optimizing atomic resolution of force microscopy in ambient conditions”. In: *Physical Review B - Condensed Matter and Materials Physics* 87 (24 Mar. 2013). DOI: 10.1103/PhysRevB.87.245415. URL: http://arxiv.org/abs/1303.5204%20http://dx.doi.org/10.1103/PhysRevB.87.245415.
- [50] Jianwei Miao, Peter Ercius, and Simon J.L. Billinge. “Atomic electron tomography: 3D structures without crystals”. In: *Science* 353 (6306 Sept. 2016), aaf2157–aaf2157. ISSN: 0036-8075. DOI: 10.1126/SCIENCE.AAF2157. URL: http://science.sciencemag.org/.
- [51] Ian MacLaren and Quentin M. Ramasse. “Aberration-corrected scanning transmission electron microscopy for atomic-resolution studies of functional oxides”. In: *International Materials Reviews* 59 (3 Apr. 2014), pp. 115–131. ISSN: 09506608. DOI: 10.1179/1743280413Y.0000000026. URL: https://www.tandfonline.com/doi/abs/10.1179/1743280413Y.0000000026.
- [52] David A. Muller. “Structure and bonding at the atomic scale by scanning transmission electron microscopy”. In: *Nature Materials* 2009 8:4 8 (4 2009), pp. 263–270. ISSN: 1476-4660. DOI: 10.1038/nmat2380. URL: https://www.nature.com/articles/nmat2380.
- [53] Martin Linck et al. “Chromatic Aberration Correction for Atomic Resolution TEM Imaging from 20 to 80 kV”. In: *Physical Review Letters* 117 (7 Aug. 2016), p. 076101. ISSN: 10797114. DOI: 10.1103/PHYSREVLETT.117.076101/FIGURES/5/MEDIUM. URL: https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.117.076101.
- [54] K. Ishizuka and N. Uyeda. “A new theoretical and practical approach to the multislice method”. In: *urn:issn:0567-7394* 33 (5 Sept. 1977), pp. 740–749. ISSN: 0567-7394. DOI: 10.1107/S0567739477001879. URL: //scripts.iucr.org/cgi-bin/paper?a14220.
- [55] P. D. Nellist and S. J. Pennycook. “Incoherent imaging using dynamically scattered coherent electrons”. In: *Ultramicroscopy* 78 (1-4 June 1999), pp. 111–124. ISSN: 03043991. DOI: 10.1016/S0304-3991(99)00017-0.

- [56] J. Barthel and A. Thust. “Aberration measurement in HRTEM: implementation and diagnostic use of numerical procedures for the highly precise recognition of diffractogram patterns”. In: *Ultramicroscopy* 111 (1 Dec. 2010), pp. 27–46. ISSN: 1879-2723. DOI: 10.1016/J.ULTRAMIC.2010.09.007. URL: <https://pubmed.ncbi.nlm.nih.gov/21111264/>.
- [57] Adrian I. Ruskin, Zhiheng Yu, and Nikolaus Grigorieff. “Quantitative characterization of electron detectors for transmission electron microscopy”. In: *Journal of structural biology* 184 (3 Dec. 2013), pp. 385–393. ISSN: 1095-8657. DOI: 10.1016/J.JSB.2013.10.016. URL: <https://pubmed.ncbi.nlm.nih.gov/24189638/>.
- [58] A. R. Faruqi and G. McMullan. “Electronic detectors for electron microscopy”. In: *Quarterly reviews of biophysics* 44 (3 Aug. 2011), pp. 357–390. ISSN: 1469-8994. DOI: 10.1017/S0033583511000035. URL: <https://pubmed.ncbi.nlm.nih.gov/21524337/>.
- [59] Xiahan Sang and James M. LeBeau. “Revolving scanning transmission electron microscopy: Correcting sample drift distortion without prior knowledge”. In: *Ultramicroscopy* 138 (Mar. 2014), pp. 28–35. ISSN: 0304-3991. DOI: 10.1016/J.ULTRAMIC.2013.12.004.
- [60] Kazuo Ishizuka. “Contrast transfer of crystal images in TEM”. In: *Ultramicroscopy* 5 (1-3 1980), pp. 55–65. ISSN: 0304-3991. URL: [https://www.academia.edu/18916287/Contrast\\_transfer\\_of\\_crystal\\_images\\_in\\_TEM](https://www.academia.edu/18916287/Contrast_transfer_of_crystal_images_in_TEM).
- [61] Naomi Falsini et al. “Halide Perovskites Films for Ionizing Radiation Detection: An Overview of Novel Solid-State Devices”. In: *Sensors 2023, Vol. 23, Page 4930* 23 (10 May 2023), p. 4930. ISSN: 1424-8220. DOI: 10.3390/S23104930. URL: <https://www.mdpi.com/1424-8220/23/10/4930/html> 20<https://www.mdpi.com/1424-8220/23/10/4930>.
- [62] Lewys Jones et al. “Smart Align—a new tool for robust non-rigid registration of scanning microscope data”. In: *Advanced Structural and Chemical Imaging* 1 (1 Dec. 2015), pp. 1–16. ISSN: 21980926. DOI: 10.1186/S40679-015-0008-4/FIGURES/11. URL: <https://ascimaging.springeropen.com/articles/10.1186/s40679-015-0008-4>.
- [63] Jérôme Boulanger, Charles Kervrann, and Patrick Bouthemy. “Space-time adaptation for patch-based image sequence restoration”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29 (6 June 2007), pp. 1096–1102. ISSN: 01628828. DOI: 10.1109/TPAMI.2007.1064. URL: [https://www.researchgate.net/publication/6397823\\_Space-Time\\_Adaptation\\_for\\_Patch-Based\\_Image\\_Sequence\\_Restoration](https://www.researchgate.net/publication/6397823_Space-Time_Adaptation_for_Patch-Based_Image_Sequence_Restoration).
- [64] Özgün Çiçek et al. “3D U-net: Learning dense volumetric segmentation from sparse annotation”. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9901 LNCS (2016), pp. 424–432. ISSN: 16113349. DOI: 10.1007/978-3-319-46723-8\_49/TABLES/3. URL: [https://link.springer.com/chapter/10.1007/978-3-319-46723-8\\_49](https://link.springer.com/chapter/10.1007/978-3-319-46723-8_49).
- [65] Lovedeep Gondara. “Medical image denoising using convolutional denoising autoencoders”. In: ().

- [66] Himmat S. Kushwaha et al. “De-noising Filters for TEM (Transmission Electron Microscopy) Image of Nanomaterials”. In: *2012 Second International Conference on Advanced Computing Communication Technologies*. 2012, pp. 276–281. DOI: 10.1109/ACCT.2012.41.
- [67] Michael Elad and Michal Aharon. “Image denoising via sparse and redundant representations over learned dictionaries”. In: *IEEE Transactions on Image Processing* 15 (12 Dec. 2006), pp. 3736–3745. ISSN: 10577149. DOI: 10.1109/TIP.2006.881969.
- [68] Leonard McMillan and Gary Bishop. “Plenoptic modeling”. In: (1995), pp. 39–46. DOI: 10.1145/218380.218398. URL: <https://dl.acm.org/doi/10.1145/218380.218398>.
- [69] Zirui Wang et al. “NeRF–: Neural Radiance Fields Without Known Camera Parameters”. In: (Feb. 2021). URL: <https://arxiv.org/abs/2102.07064v4>.
- [70] Naama Pearl, Tali Treibitz, and Simon Korman. “NAN: Noise-Aware NeRFs for Burst-Denoising”. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2022-June* (Apr. 2022), pp. 12662–12671. ISSN: 10636919. DOI: 10.1109/CVPR52688.2022.01234. URL: <https://arxiv.org/abs/2204.04668v2>.
- [71] Ben Mildenhall et al. “NeRF in the Dark: High Dynamic Range View Synthesis from Noisy Raw Images”. In: (Nov. 2021). URL: <http://arxiv.org/abs/2111.13679>.