

**NANYANG
TECHNOLOGICAL
UNIVERSITY**
SINGAPORE

Analysis of TEM Data using Machine Learning Methods

Muhammed Imran bin Khairul Alam
U1720420H

School of Materials Science & Engineering

**A final year project submitted to the Nanyang Technological University in
fulfilment of the requirement for the degree of Bachelor of Engineering**

Analysis of TEM data using machine learning methods

Muhammed Imran bin Khairul Alam

Supervisor: Asst. Prof Martial Duchamp

Abstract

Liquid phase electron microscopy has various advantages over other in situ microscopy techniques. Due to the nature of the liquid medium during sampling, various types of specimens can be observed, which would not have been suitable in a typical vacuum setup. An additional advantage would be the ability to observe samples without the need for traditional sample preparation. In particular, there is a degree of inaccuracy in the observed morphology of polymer particles when placed under vacuum conditions, without an liquid medium. As such, a liquid medium would directly address this issue and there would be a greater degree of accuracy in its observed morphology. However, liquid phase electron microscopy does come with limitations, which poses a problem for clear and accurate imaging of the samples. These imaging limitations are further exacerbated by the need for low electron doses to preserve the sample. Therefore this paper shall discuss the limitations of liquid phase imaging and the effect of low electron dosage. This paper also presents a comparison of imaging techniques, such as filtering and machine learning methods, that would improve the quality of the raw liquid phase images.

Acknowledgements

I would like to express my gratitude to Professor Martial Duchamp, who has given me the opportunity to pursue my interest in image processing in microscopy applications. He has taught me to be dynamic, flexible and to better adapt my thought process throughout my Final Year Project. He has provided me with constant support and constructive feedback and it has helped me to better shape my project and experiments. I would also like to thank Dr Tay, Jeffrey and Anastasia for their continuous help in supporting my simulations, clarifying many unending doubts and to allow me to better understand the intricacies of conducting electron microscopy experiments.

Contents

1	Introduction	1
1.1	Transmission Electron Microscopy	1
1.1.1	Illumination	2
1.1.2	Objective Lenses	2
1.1.3	Imaging	3
1.1.4	Specimen Preparation	3
1.2	Liquid Phase Electron Microscopy	4
1.3	Limitations of Liquid Phase Electron Microscopy	4
1.4	Noise in Images	5
1.5	Scope	6
2	Literature Review	7
2.1	Filtering	8
2.2	Fourier Transforms	9
2.3	Machine Learning	9
2.3.1	Deep Learning	10
2.3.2	Supervised vs Unsupervised Learning	10
3	Experimental Setup	12
3.1	Dataset	12
3.1.1	Frame Averaging	13
3.1.2	Filtering and Deconvolution Sequence	14
4	Results	18
4.1	Median Filtering, Deblurring and Denoising	18
4.2	Noise2Noise	23
5	Discussion	24
5.1	Filter-based Denoising	24
5.2	Noise2Noise	27
5.3	Conclusion	27
5.4	Future Works	28

Chapter 1

Introduction

The electron microscope was first introduced by Ernst Ruska and Max Knoll in 1931. It was primarily developed due to the limitation of the light microscope as its image resolution is limited to the wavelength of light. This limitation can be demonstrated by equation (1.1), the Rayleigh Criterion, which calculates the smallest resolvable distance of an illuminated specimen, where λ is the wavelength of the radiation source, μ is the refractive index of the viewing medium and β is the semi-angle of collection of the magnifying lens.

$$\delta = \frac{0.61\lambda}{\mu \sin \beta} \quad (1.1)$$

For a typical light microscope, taking the wavelength of green visible light of 550 nm, its maximum resolution is 300 nm. However, this range of resolution does not allow for imaging at the nano scale. In contrast, the wavelength of an electron is significantly shorter than that of light. Hence, the electron microscope was able to reproduce images that were magnified 100s of times greater than a light microscope.

$$\lambda = \frac{1.22}{E^{\frac{1}{2}}} \quad (1.2)$$

Equation (1.2), E represents the energy of the electron in eV. Hence for an electron with an energy of 300 keV, its λ is approximately 0.002 nm. Therefore this would result in a much greater resolution for nano scale specimens.

1.1 Transmission Electron Microscopy

In order to image the morphology of a specimen, Transmission Electron Microscopy (TEM) relies on electrons that pass through the specimen. These electrons are then collected on a fluorescent screen or a charge-coupled device (CCD). Ultimately, the image is digitally displayed for viewing on a monitor. Fig. 1.1 shows the various components of a TEM and can be described in three major parts: the illumination, the condenser

lenses and stage, and the imaging parts.

1.1.1 Illumination

The source of illumination originates from the electron gun. There are three types of electron guns, namely thermionic electron guns, Schottky emission guns and cold field emission guns.

In a thermionic electron gun, the gun filament is usually made of tungsten or a single crystal of lanthanum hexaboride (LaB_6). The cathode of the filament is then heated electrically to emit electrons. These electrons are then accelerated towards the anode due to a high voltage across the cathode and anode. This method of producing electrons is relatively inexpensive since an ultra-high vacuum is not necessary and the replacement cost to filament lifetime ratio is relatively low. However, the brightness produced from the beams are relatively poor. In addition, the lifetime of the filament is approximately 100 hours.

Schottky emission guns emit electrons when there is a lowering of the potential barrier when a strong electric field is applied to a heated metal surface. The gun is made of a single tungsten crystal that is coated with zirconium oxide (ZrO). The brightness produced exceeds that of a thermionic electron gun by a magnitude of three. As such, it is suitable for use in high-resolution imaging. The lifetime of the filament can span across several years. However, the operation of the gun requires an ultra-high vacuum environment.

Cold field emission guns also function by the reduction of the potential barrier due to an applied electric field, similar to the Schottky emission guns. Electrons are able to penetrate the potential barrier via the quantum tunneling effect. However, in contrast to Schottky emission guns, heating of the cathode is not required. [1]

1.1.2 Objective Lenses

Once the electron beam exits the emission gun, it is adjusted by the condenser lenses, C1 and C2. The C1 lens ensures that the electron beam is appropriately focused at the C1 crossover, located between the C1 and C2 lenses. The electron beam then encounters the C2 lens and aperture. A smaller C2 aperture results in a smaller beam size that reaches the specimen. At the same time, the smaller aperture results in a more parallel illumination rays, which are used for TEM imaging.

1.1.3 Imaging

Next, the electron beam reaches the specimen holder, which is held in place on the stage, and passes through the specimen onto the objective lens. The objective lens is then adjusted such that the object plane is the image plane of the lens. This finally produces an image on the viewing screen. [2]

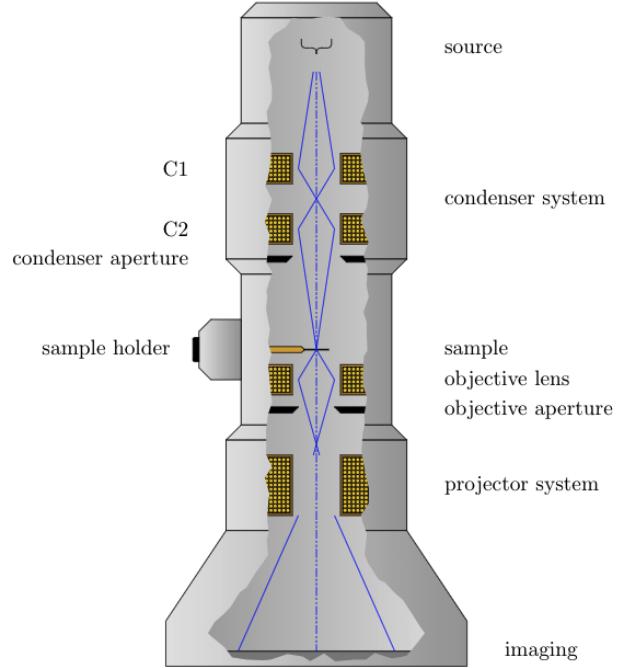


Figure 1.1: Components in a TEM.

1.1.4 Specimen Preparation

Since electrons are used as the source of radiation, the thickness of the prepared specimen should be approximately 100nm. [3] Hence, careful specimen preparation is required to obtain any useful results as too thick of a specimen results in unclear images. Once the specimen's thickness has been appropriately prepared, it is placed on a support grid to be loaded onto the specimen holder, as seen in Fig. 1.2.

Extending the same principles, other methods such as Scanning Transmission Electron Microscopy (STEM), Cryo-EM and in-situ TEM can achieve varying image results, depending on the required analyses. In this paper, the merits and limitations of another TEM method, namely liquid phase electron microscopy (LPEM) are considered.

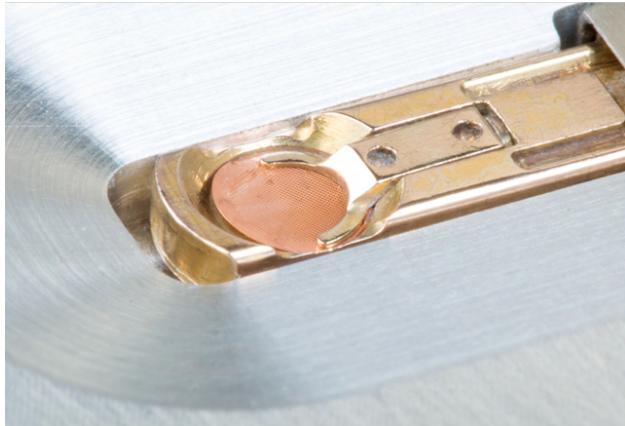


Figure 1.2: TEM specimen holder with a fitted copper support grid. [4]

1.2 Liquid Phase Electron Microscopy

Conventional TEM environments cannot accommodate for liquid or liquid-filled specimen in a vacuum environment. Such an environment would cause changes to its morphology and these changes would not accurately reflect the true state of the specimen. LPEM overcomes this issue by allowing for the specimen to retain its liquid content in-situ, and to provide a liquid medium while being observed.

Within LPEM, there are several types of imaging techniques which can be broadly classified into open or closed cell. Open cell techniques require a high pressure within the specimen chamber relative to the rest of the microscope, which is achieved through differential pumping. However, due to the nature of differential pumping, atmospheric pressure cannot be obtained within the specimen chamber. As a result, solvents such as isopropanol and water cannot be used as the liquid medium since it would evaporate.

In contrast, the closed cell approach encapsulates the specimen between two windows, separated by a spacer and O-rings. This creates a cell that is a few hundred nanometres thick and is able to accommodate higher pressures than open cell approaches and independent of the pressure in the chamber. [5]

1.3 Limitations of Liquid Phase Electron Microscopy

Despite its merits, LPEM has several drawbacks that limit the resultant image resolution. Firstly, the thickness of the liquid in the cell is inversely correlated to the resolution of the resultant image. As such, a thicker liquid would decrease the image resolution. Next, for closed cells in particular, variations in liquid thickness across the cell is caused by the difference in pressure between the interior and exterior of the cell. Uneven

thickness would result in inaccurate depths, which may lead to misinterpretation of results. Similarly, thick windows of a liquid cell can reduce the image resolution due to electron scattering.

In addition, trapped bubbles also pose a problem since liquid flow is hindered. These bubbles are occasionally stuck within the liquid cell, or along the liquid medium's path of flow. As such, this may pose a problem for in-situ experiments that require the flow of the medium. Lastly, when imaging biological samples, a high electron dose would result in cellular damage due to radiolysis. Hence a low electron dose is required to ensure the specimen is not damaged. However this would result in significantly poorer image quality due to a weaker electron beam. In [6], polymer vesicles are imaged in a LPEM environment. However, due to the low dose required for preserving the specimen, the images contain high levels of noise. [7] Therefore a solution is needed to improve the resultant image resolution in order for it to be interpretable.

1.4 Noise in Images

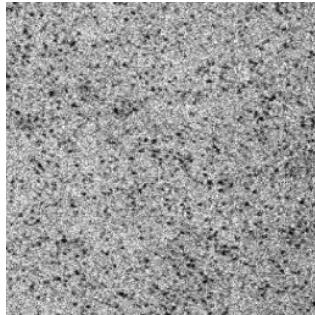


Figure 1.3: TEM image of Cadmium Sulfide nanoparticles with Gaussian noise.

Noise present in images are represented by pixel values which do not accurately reflect the information captured by the image sensor. These can be caused by insufficient specimen illumination, background artefacts or even poor image sensor resolution. [8] As a result, this is seen as a reduction in the quality of interpretable information. A good denoising technique should preserve edges of objects present in the image, while effectively removing the noise. [9] In the case of TEM, the presence of noise may result in misidentifying particles, misinterpreting the relative size of particles or even missing particles in the worst case scenario. Fig. 1.3, the TEM image of Cadmium Sulfide nanoparticles have been largely masked by the noise present. Hence it is difficult to distinguish an actual nanoparticle from noisy pixels which is critical for further analysis. [10]

Assuming that the value of each pixel in an image is independent, each pixel value is directly proportional

to the number of illuminating particles that was detected by pixel's corresponding sensor area over a given time interval. However the intensity of the illumination source is not constant and the exact number of particles that are detected at a time interval follows a probability distribution. The probability distribution can be approximated to a Poisson distribution given that the illumination source has constant brightness over a time interval. Hence the probability of detecting f number of particles detected at a single pixel over a given time interval is shown in Equation 1.3.

$$p(f) = \frac{\langle f \rangle^f e^{-\langle f \rangle}}{f!} \quad \text{for } f \in \mathbb{Z}^+ \quad (1.3)$$

A common noise pattern observed in images is additive Gaussian noise, which follows a Gaussian distribution and is an additive noise that is independent of the underlying pixel value. This can be introduced during image acquisition, where the source of illumination is poor. Noise n is shown in Equation 1.4, where x is the observed image and t is the underlying clean image.

$$x = t + n \quad (1.4)$$

1.5 Scope

In this report, the problem of denoising low-dose, highly noisy LPSEM images is considered. Firstly, the nature of such image datasets are explored. Various denoising techniques such as filtering and unsupervised machine learning are then considered to determine its fit to the initial denoising problem. Lastly, these techniques are tested and the results are analysed.

Chapter 2

Literature Review

Images can be represented as a matrix of individual pixel intensities. Though pixel values can take any value, they are typically normalised to a range of 0 to 255 in order to be visually displayed. Singular pixel values or a neighbourhood of pixels can represent information that is visually interpretable. For example, a linear edge of an object can be represented by a sharp difference in neighbouring pixel values, as shown in Fig. 2.1. As such, pixel values can be manipulated by applying filters or other non-linear mathematical functions to produce a different visual effect.

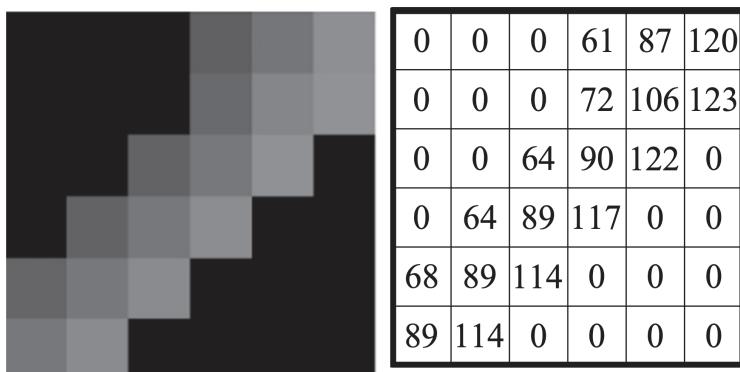


Figure 2.1: Pixel value representation of an edge. [11]

Image denoising has been a field that has been extensively explored, with various algorithms and methods employing varying degrees of image filters [12] and even convolutional neural networks (CNNs) [13] to identify key features in the images. The majority of the methods can be distinguished based on the use of a priori, noise-free reference images, a combination of the two, or no prior knowledge or reference image at all.

2.1 Filtering

A simple average filter uses a moving neighbourhood window that averages the pixels within the window. The averaged value then replaces the pixel value at the centre of the window. This can be seen in Fig. 2.2, where the sum of pixel values in the red window is 270 and the average value of the 3x3 window is 90. Therefore the pixel is assigned a new value of 90. As a result, the edges of objects in the image are blurred, giving the image a smoothed effect.

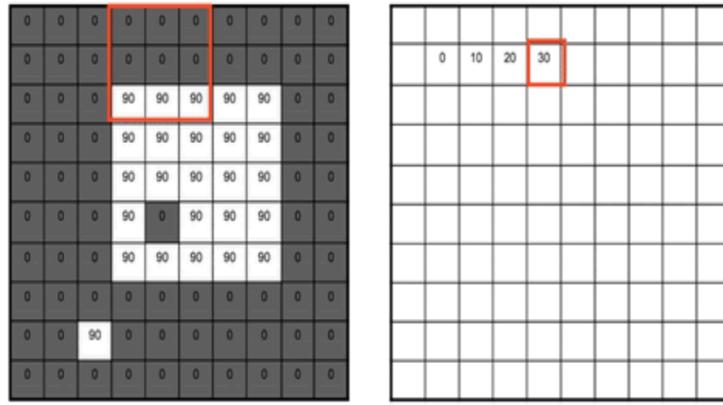


Figure 2.2: A 3x3 moving average window and its corresponding outcome. [14]

Similarly, Gaussian smoothing applies a window to the the image. The window applied is a discrete approximation of a 2D Gaussian distribution. Equation 2.1 expresses the 2D Gaussian distribution, where x and y are the coordinates of the pixel value and σ is the standard deviation of the distribution.

$$F(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2.1)$$

In contrast, a median filter is a non-linear filter that is effective in preserving edges of objects. It is also effective in handling outliers, hence removing extremely high or low pixel values. Fig. 2.3 depicts a 3x3 median filter on the top left hand corner of the input matrix. Once the nine values in the window are sorted, the median value is chosen to replace the existing pixel value, as shown in the output matrix. [15] proposes a combination of a median and a progressive deconvolution filter on LPEM images that first removes the noise present in the image and subsequently restoring the sharpness of edges present.

Filtering techniques tend to use a priori, which determines the level of noise present. A priori includes estimating the standard deviation of Gaussian noise present in the image. In particular, median filtering with a small window size may not effectively remove noise with high variance. On the other hand, a large

Input						Output					
1	4	0	1	3	1	1	4	0	1	3	1
2	2	4	2	2	3	2	1	1	1	1	3
1	0	1	0	1	0	1	1	1	1	2	0
1	2	1	0	2	2	1	1	1	1	1	2
2	5	3	1	2	5	2	2	2	2	2	5
1	1	4	2	3	0	1	1	4	2	3	0

Sorted:0,0,1,1,1,2,2,4,4

Figure 2.3: A 3x3 median filter and its corresponding outcome. [16]

window size may blur the image. [17] Therefore testing the filter window size is an important step in testing the suitability of a filter.

2.2 Fourier Transforms

In a Fourier transformation of an image, the pixel values of the spatial domain is converted to amplitudes in the frequency domain. By doing so, the difference in values of the low frequency peaks and the high frequency noise can be distinguished, allowing for a threshold to be determined to remove the noise.

A benefit to using Fourier Transform is that the original data is not removed. [18] However, the nature of Fourier Transforms, that is to decompose a signal into a weighted sum of sinusoids, lacks spatial localisation. As a result the origin of a frequency cannot be mapped back to the original signal present in the image. [19]

2.3 Machine Learning

Machine learning solutions can be loosely defined as models that learn to produce an optimal solution from a dataset and its labels or generalising towards a pre-defined objective function. In the field of computer vision, repeated patterns in images can be exploited by machine learning algorithms that contribute towards a general solution. This can be particularly useful when such patterns are either not easily observable by humans or the said patterns are complex in nature, rendering it too tedious to produce a complete ruleset.

2.3.1 Deep Learning

Due to the nature of image pixels that can be represented as matrices, the patterns and correlations present within the image are complex. A common architecture employed in deep learning models is convolutional neural networks. These can be seen as layers that iteratively recognise and extract features in the dataset. Pooling layers subsequently reduce the dimensionality of the features without removing the features themselves. Finally, the output layer produces the intended output. In the context of image denoising, the model should ideally produce a clean and noise-free image. Fig. 2.4 shows a typical layout of a convolutional neural network, where an input layer is followed by several layers of convolutional and pooling layers before eventually producing discrete outputs.

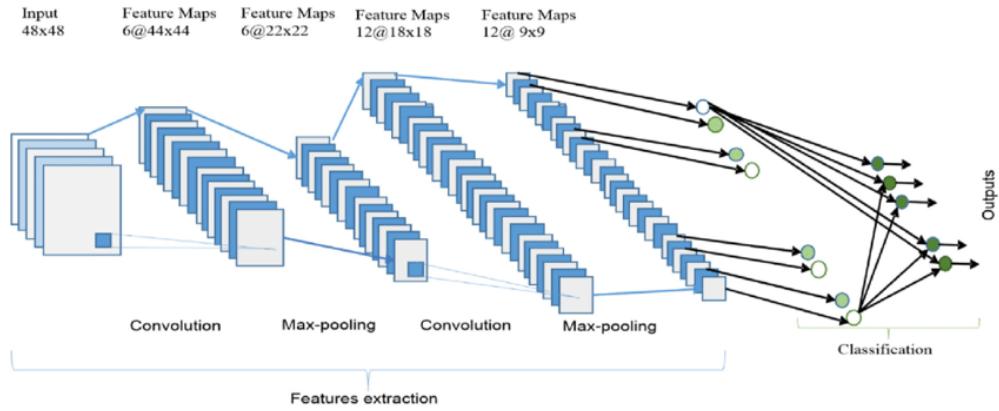


Figure 2.4: A typical convolutional neural network architecture. [20]

Another common architecture used in deep learning models is the autoencoder. Due to the nature of its encoding-decoding structure, the model would ideally reduce the input data's dimensions and simultaneously remove noise. The autoencoder is made up of four main components: the encoder, bottleneck, decoder and the loss calculated between the input and output. The encoder initially reduces the input dimensions. The bottleneck is then achieved as the lowest dimensions of the input. Subsequently the decoder reconstructs the encoded representation and finally the calculated loss is backpropagated to train the network. Fig. 2.5 shows the sequential layers present in an autoencoder, with the compressed representation as the bottleneck.

2.3.2 Supervised vs Unsupervised Learning

Supervised learning utilises labeled input and output data in order to predict accurate outcomes on unseen data. The model iteratively improves its solution by comparing its predictions to that of the given input data, thereby reducing the error between its predictions and the given outputs. The reliance on labeled data is inherently a problem for many domains where such data cannot be obtained. In the context of TEM

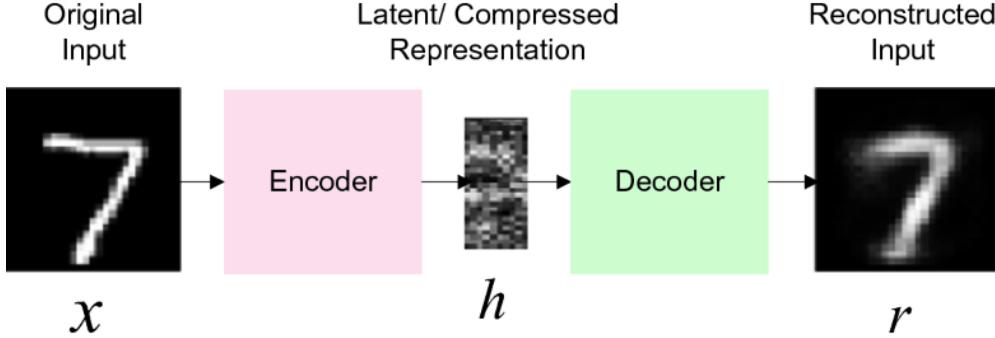


Figure 2.5: A typical autoencoder structure. [21]

images, it is often expensive to obtain additional data. In addition, there is also an added difficulty in obtaining clean reference images of specimens. [22] presents a denoising network which utilises pairs of clean and noisy images in order to denoise electron microscopy images. However such pairs are unavailable when taking images in an LPEM setting.

In contrast, unsupervised learning does not require labeled inputs and their corresponding outputs. Instead, it relies on recurring patterns or structures present in the given dataset. However, some validation of the dataset and the model’s outcome must be present in order to gain any meaningful insights. For example, a TEM image dataset containing noisy images must be validated to have consistent information. Similarly validation of the results must be carried out in order to recognise that unwanted visual artefacts are not added to the image. Such validation can only be carried out with the relevant domain knowledge.

Lehtinen had initially proposed to train a deep learning model on pairs of noisy images which contained the same underlying information, instead of the conventional noisy-clean image pairs. [23] This relied on having a large dataset of training pairs and the noise approximating to a Gaussian distribution, where the model would eventually learn to estimate the correct pixel values in a noisy image. Subsequently, [24] extended the idea to a self-supervised solution, where training pairs are instead taken from patches of a single image. As such, it removes the previous underlying assumption of the noise distribution in the images. Taking the idea of self-supervised methods a step further, [25] includes the prior knowledge of the point spread function of a light microscope in their deep learning model. However, the results from these newer models have proven to produce poorer results than their supervised counterparts.

Chapter 3

Experimental Setup

The datasets used in the simulations contain LPEM digital micrographs of poly(ethylene glycol)-b-poly(lactide) (PEG-PLA) vesicles that were subjected to changes in temperature over a time period. The experiments conducted aimed to determine the structural response of the PEG-PLA vesicles at different temperatures. However, the exact temperature ramp prior to the dataset acquisition is unclear. Hence the methods described in this paper aim to denoise the acquired data, such that changes in the vesicle structure, if any, could be clearly determined.

3.1 Dataset

Each file in the initial dataset was stored in Gatan's proprietary format, dm4. Subsequently, the image signal was extracted from the files, which was represented as a 2D-array. Images were analysed using Gatan's Digital Micrograph software, Python Hyperspy package and MATLAB visualisation tools. Without editing the images, it was not possible to carry out any image analysis as the noise present in the images is significantly greater than the information present. Fig. 3.1 shows a raw image from the dataset that has no observable artefacts.

The dataset consists of a video sequence captured over 2 minutes 48 seconds at a rate of 25 frames per second and the temperature was kept constant at 50°C. Each frame had a magnification of 10000 and the was exposed using an electron source of 300kV for 0.04s. All images in the dataset had an original resolution of 4096x4096 pixels and were acquired via a JEM-ARM300F GRAND ARM Atomic Resolution Electron Microscope. In all images, the signal-to-noise ratio was extremely poor. Hence no direct analysis could be carried out. As seen in Fig. 3.2a, there are extreme outlier pixel values in a single frame which make it

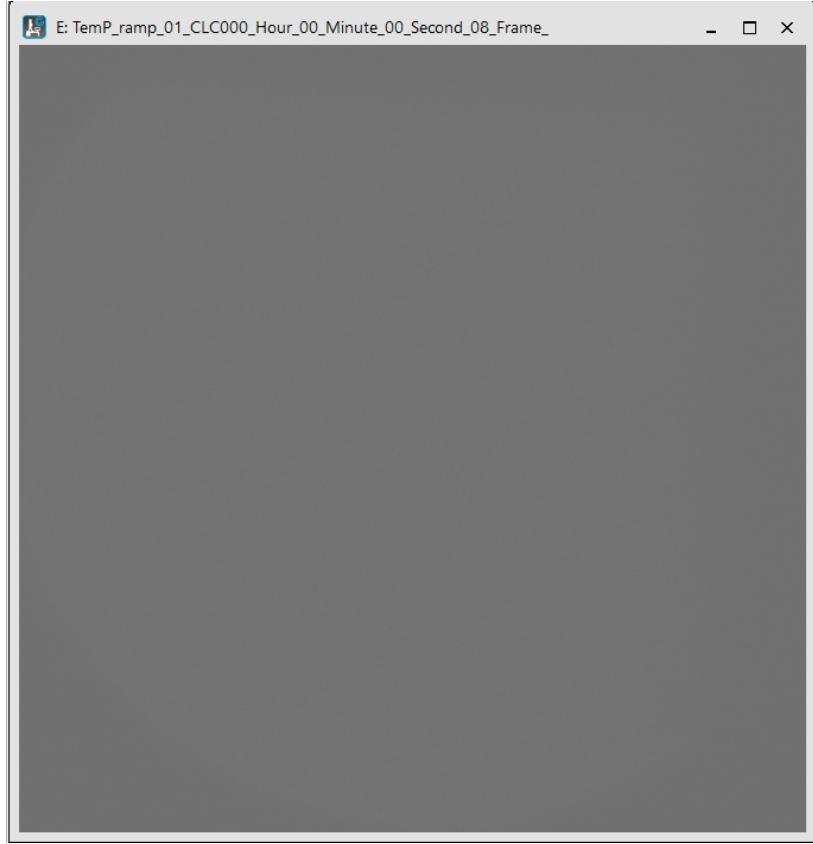


Figure 3.1: No visible information observable on a raw dataset image. Taken from Gatan's Digital Micrograph software interface.

difficult to analyse frames individually. Fig. 3.2b shows very limited information within the crop as there is no clear distinction between signal and noise pixels.

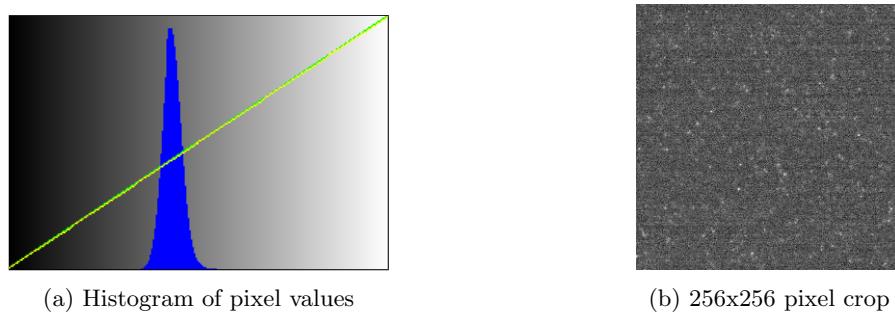


Figure 3.2: Unedited raw Frame 1 at Minute 0, Second 8 from video sequence.

3.1.1 Frame Averaging

Given that each frame in the video sequence was exposed to a very small dose for a short time period, each individual frame would not effectively capture sufficient information on its own. Therefore in order to

represent any significant signal within an image, several images can be combined via averaging. The flow rate of the liquid is sufficiently slow to allow averaging consecutive frames without losing significant information in the resultant images given the fast rate of image capture. In addition, averaging over multiple frames would reduce the overall Gaussian noise [26]. Hence a moving average window was applied to consecutive frames in order to obtain a more coherent dataset, as shown in Fig. 3.3.

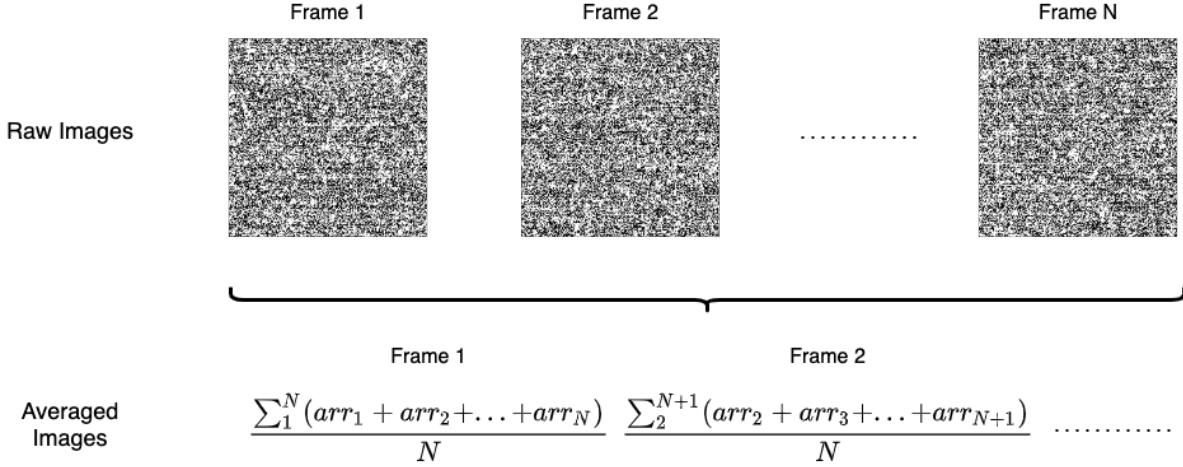


Figure 3.3: Moving average of video sequence frames to generate averaged frames dataset

However, the number of averaged frames can be modified in order to reduce the occurrence of noise in the image. As the size of the average window increases, the resultant image would contain information that is increasingly blurred, considering the objects in the frame are in motion. On the other hand, selecting a window size that is too small would not significantly reduce the noise, and the perceived motion between consecutive frames would be disjointed. In addition, 0.1% of the highest and lowest pixel values of the averaged image were removed to further reduce noise. Therefore a range of window sizes were used and their corresponding histograms are shown in Fig. 3.4. With an increase in the window size, the pixel values are more concentrated nearer to the mean pixel value. As such, when the pixels in the resultant image are normalised to a range between 0 and 255, the pixel intensity values are increasingly spread out as depicted by increasingly sparse images. All images were saved in .tiff format to prevent image quality loss through file compression.

3.1.2 Filtering and Deconvolution Sequence

In [15], a median filter, denoising step and a deblurring function is applied to an LPEM input image in order to obtain a clear image. However, the exposure time of the frame used in the paper is more than 2 times

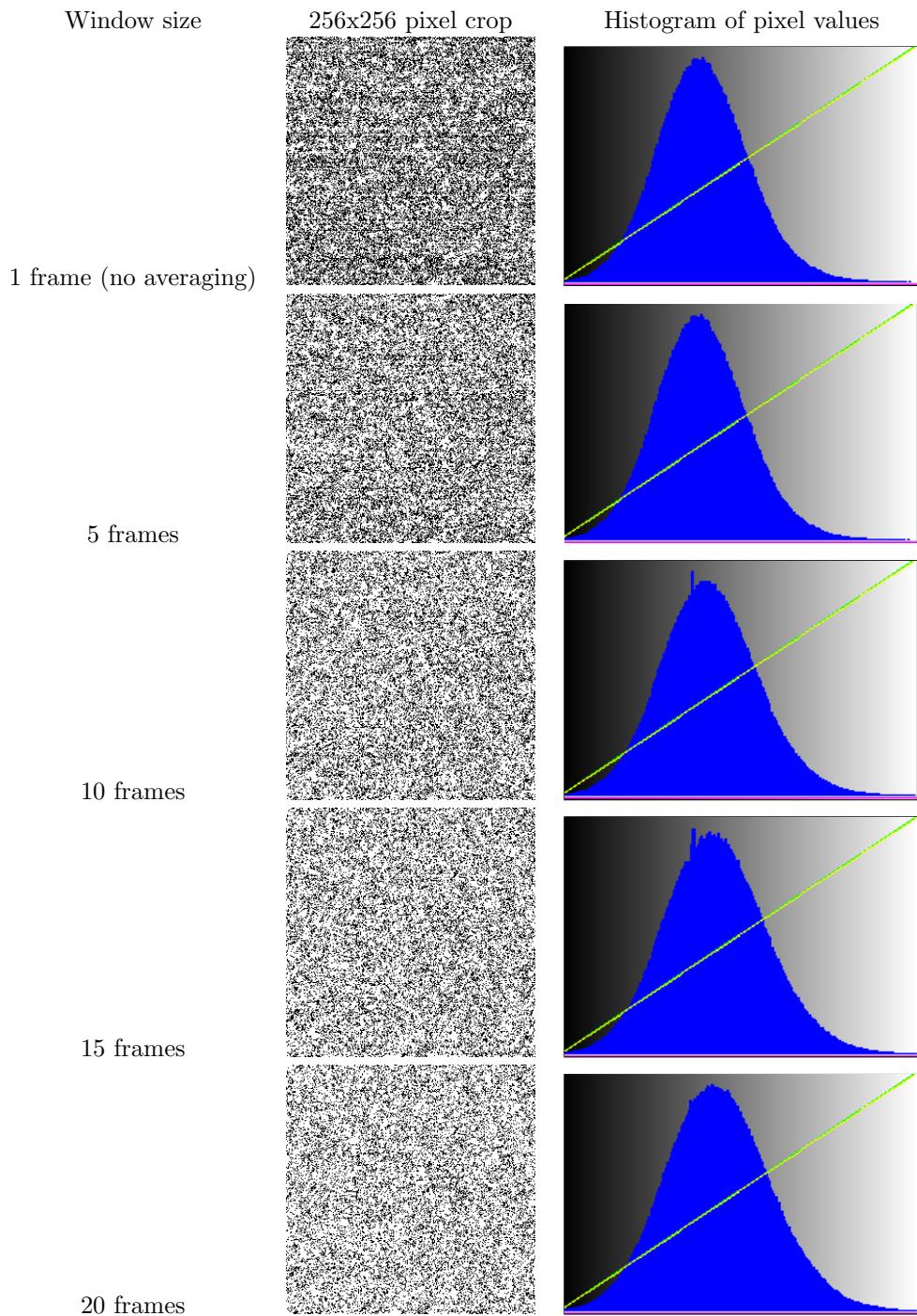


Figure 3.4: Resultant image and pixel value histogram when an average window is applied.

longer than an individual frame in the video sequence dataset. Therefore varying parameters were tested in order to achieve denoised images.

The initial median filtering step is followed by a patch normalisation. In this case, median filtering of various filter sizes are tested, and whether the inclusion of the patch normalisation improves the intermediate image. Fig. 3.5 shows the outcome of median testing on the various averaged frames. Subsequently, Progressive Image Denoising (PID) [27] is applied, where noise is separated from the signal, the noise is estimated and dual-domain image denoising [28] is applied to both the spatial and Fourier domains of the image. Lastly, a local deconvolution is applied to a combination of the median filtered image and the PID processed image.

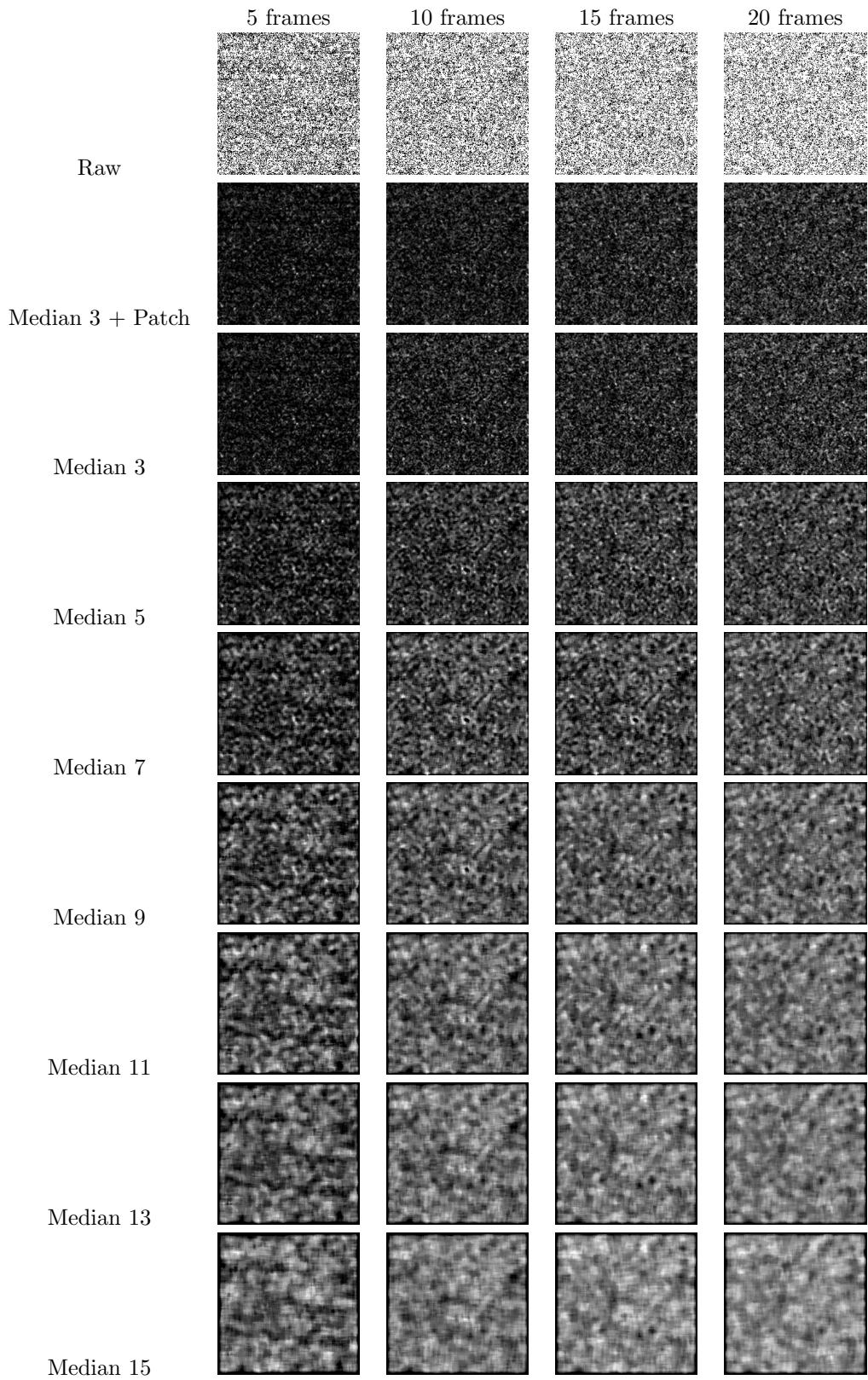


Figure 3.5: Visual comparison of different median filter sizes on the various averaged frame sequences.

Chapter 4

Results

4.1 Median Filtering, Deblurring and Denoising

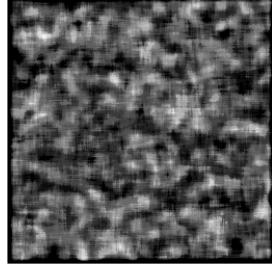
The method used in this section follows the pipeline suggested in [15]. From Fig. 3.5, a 5 frame averaged window was chosen as the dataset to be used subsequently with Progressive Image Denoising. In addition, the setting of a median filter of size 11x11 was chosen. The combination of settings had been chosen due to the considerations:

- As the average frame window size increased, the information represented in the final image also increases. However, an increased window size also increased the blur of the motion of artefacts present in the frame.
- A median filter of size 9x9 or smaller had no significant improvement in sharpening the edges detected on the image. Conversely, a median filter of size 13x13 or larger introduced a blurring effect.

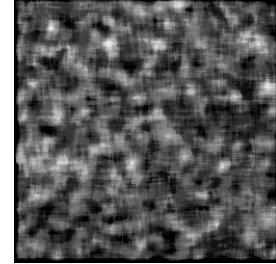
From this point on, the median filter refers to a median filter of size 11x11 and the averaged frames are a result of averaging 5 frames unless otherwise stated.

However, since the initial averaged frames were still noisy, it was necessary to confirm if the edges shown in one cropped frame would be different from that of another cropped frame. Fig. 4.1 shows the result of a median filter applied to averaged frames. The edges observable in both images are different, hence it can be visually concluded that the median filter is not introducing artefacts into the images.

The images were then fed into the PID function which was executed on MATLAB. It was observed that running the function fewer than 30 iterations had poor denoising, while more than 30 iterations had

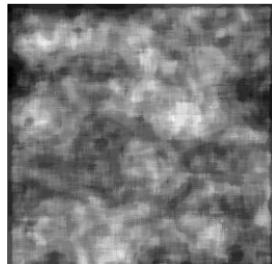


(a) 1st crop area

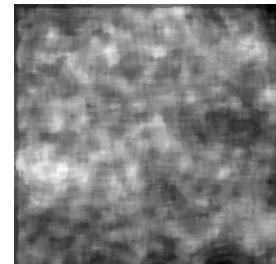


(b) 2nd crop area

Figure 4.1: Median filter applied to 256x256 crops from averaged frames of Minute 0, Second 8 to Minute 0, Second 10.



(a) 1st crop area



(b) 2nd crop area

Figure 4.2: PID applied to median filtered crops, as seen in 4.1.

insignificant denoising effect.

Lastly, both the median filtered and PID processed images were fed into the next step, where size and stride of the deconvolution window for the subsequent deblurring and local deconvolution functions are important parameters to be determined through experiments. From Fig. 4.3, a larger kernel size applied during the deblurring step leads to a greater blurring effect of edges present in the image. However, too small a kernel size would result in 'ring' effects around the edges, which could create unwanted artefacts in the final deconvolution step. Hence a kernel size of 4 was chosen.

Next, the settings for the local deconvolution function were tested. First, the stride of the window was compared, keeping a constant width of 16 pixels. Fig. 4.4 shows the different images based on the settings. Stride size 4 and above produce images with a 'hatching' effect, which hinders any visual interpretation of the results. Conversely, stride size 1 produces an overly sharp image where the edges of the perceived artefacts may not be accurate. Therefore a stride of size 2 was chosen. The width of the window was then compared in Fig. 4.5. Widths 28 pixels and greater displayed blurring of the artefact's edges and a decreasing overall contrast. Conversely widths 20 pixels and smaller displayed visible noise which masked the edges of the

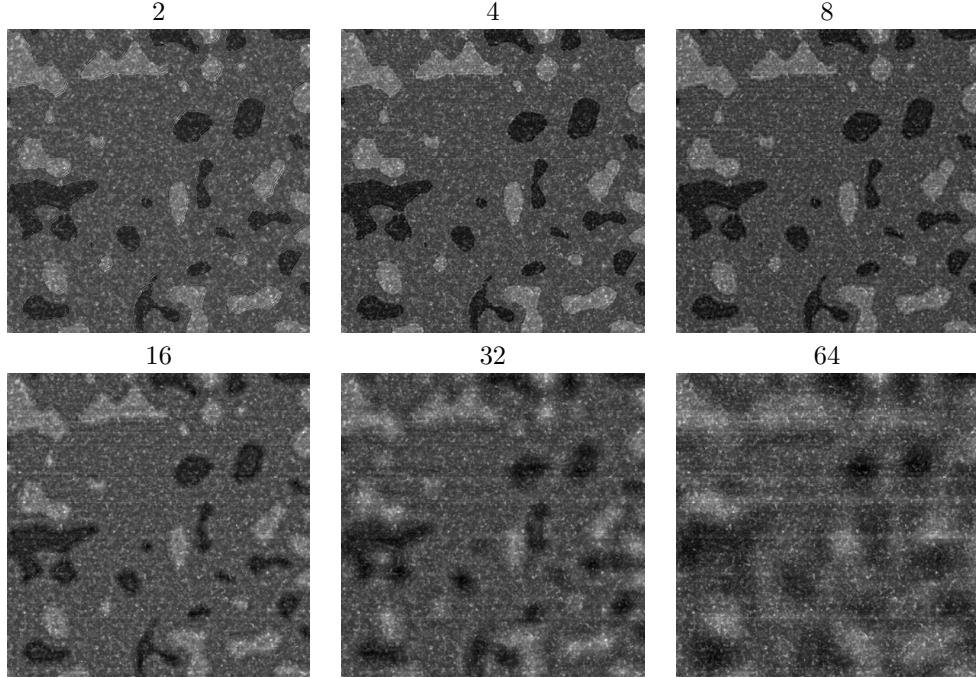


Figure 4.3: Different kernel sizes used for the deblurring function.

artefacts. Therefore a deconvolution window of width 26 pixels was chosen.

The final parameters chosen for the sequence are as follows:

- Average frame window: 5 frames
- Median window size: 11x11 pixels
- Number of PID iterations: 30
- Kernel size for deblurring: 4x4 pixels
- Stride for deconvolution window: 2 pixels
- Deconvolution window size: 26x26 pixels

Fig. 4.6 shows the raw image, median filtered image, PID and deblurred processed image and the local deconvolution processed image. The right column of images represents the final denoised images. The denoising process was completed in approximately 1 hour 30 minutes.

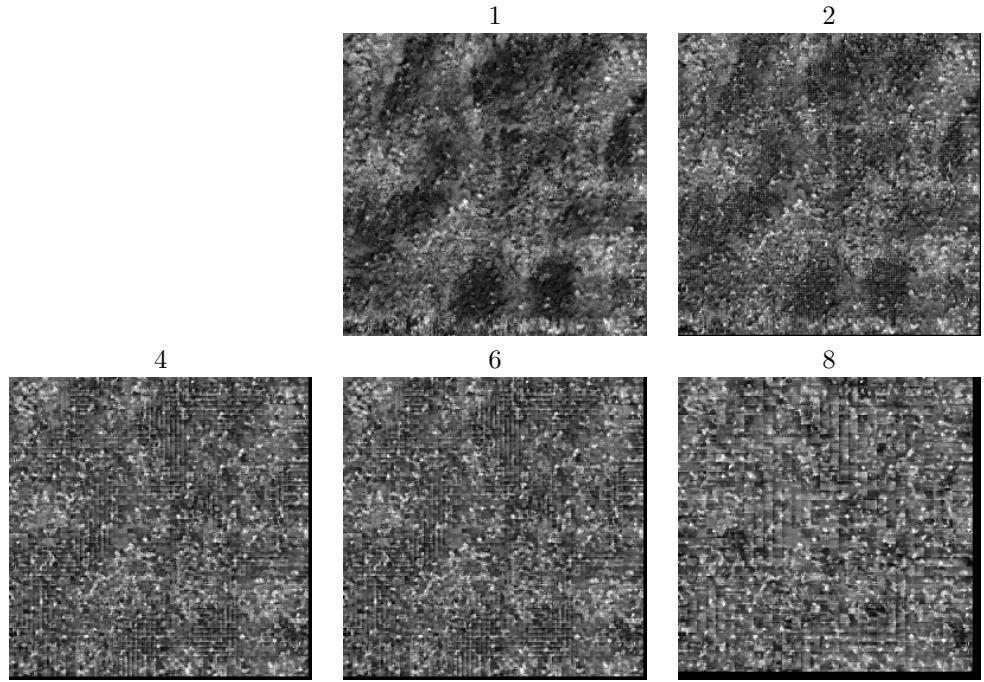


Figure 4.4: Comparison of stride settings used for the local deconvolution function. Width of deconvolution window was kept constant at 16 pixels.

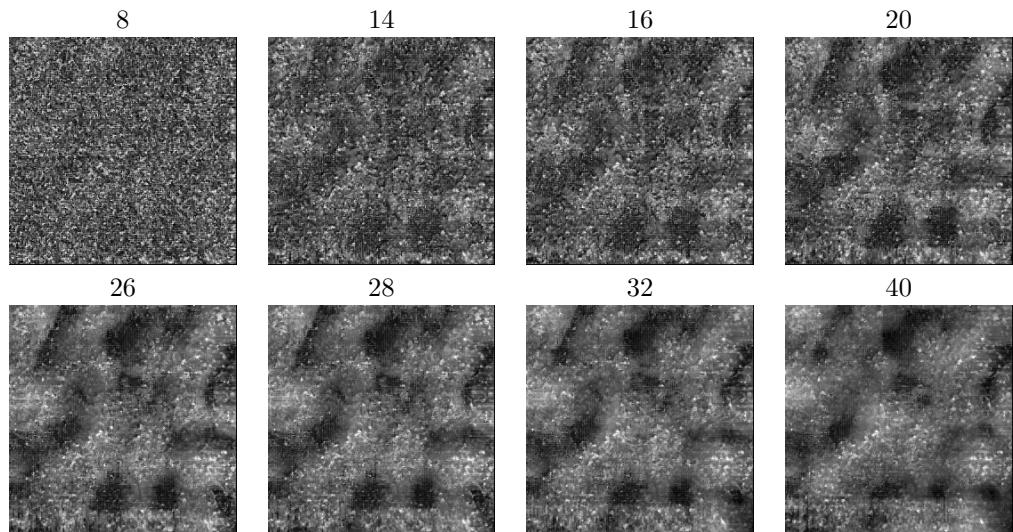


Figure 4.5: Comparison of window size width settings used for the local deconvolution function. Stride size of deconvolution window was kept constant at 2 pixels.

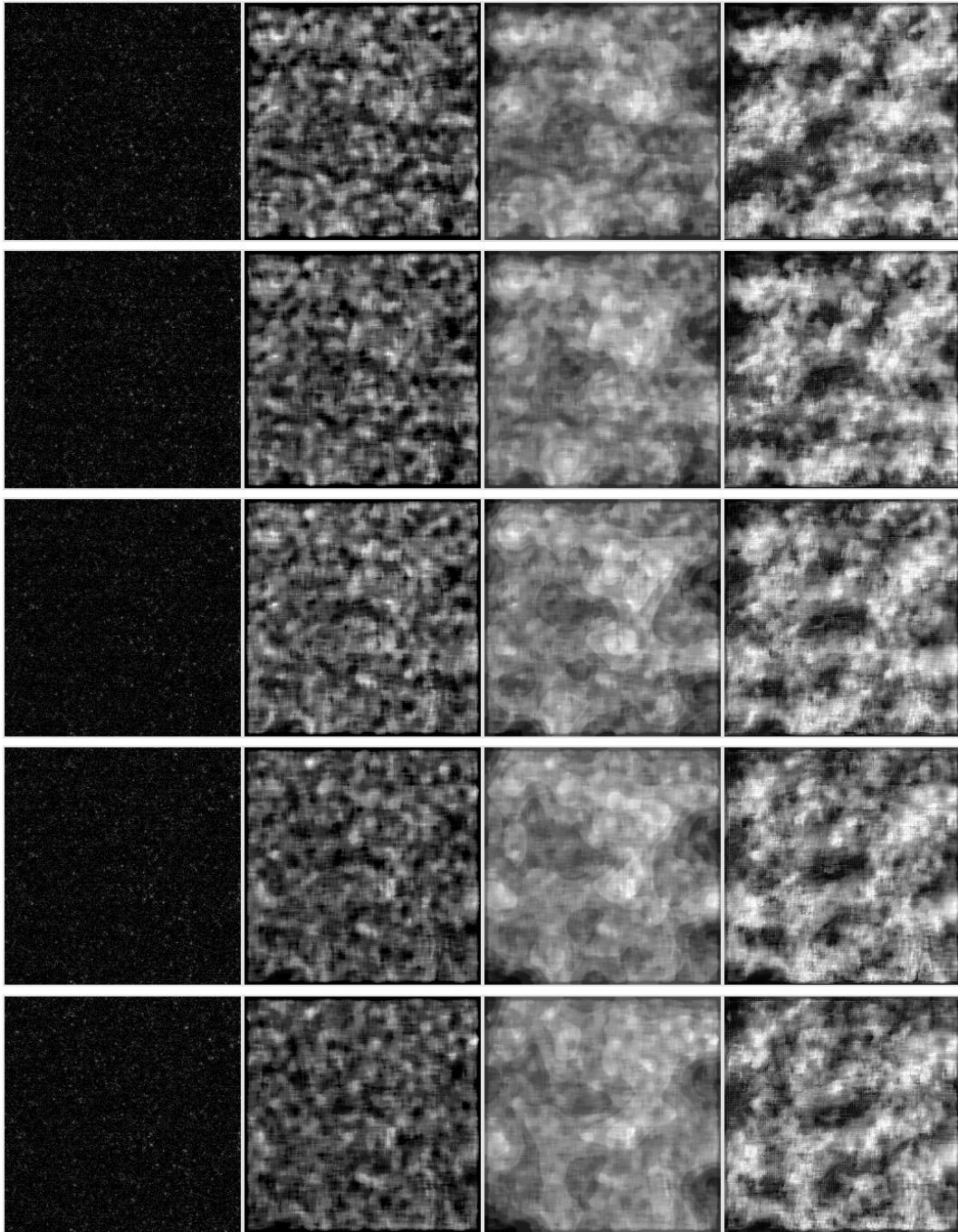


Figure 4.6: Collated images from raw image, median filtered, PID and deblurring, and local deconvolution from filter-based denoising. Each row represents consecutive average windows (ie. Frames 1-5, 2-6, 3-7, 4-8, 5-9) at Minute 0, Second 8 - 10.

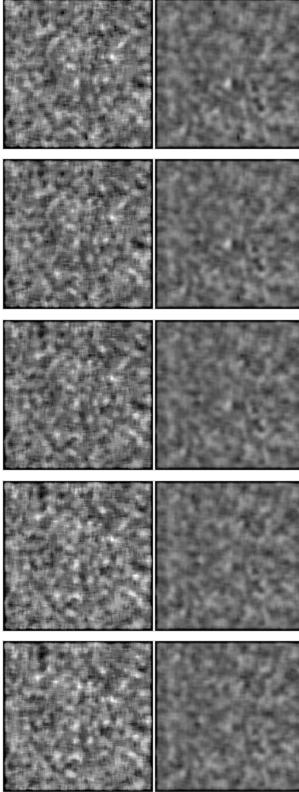


Figure 4.7: Collated images of the input median filtered and output denoised images from trained Noise2Noise model. Each row represents consecutive average windows (ie. Frames 1-20, 2-21, 3-22, 4-23, 5-24) at Minute 0, Second 8 - 10.

4.2 Noise2Noise

In [23], noisy paired images were used to train a denoising model. However the assumption made was to have pairs of images that had the same underlying information but a different set of noise pixel representation. However, since there is liquid flow in the images of the dataset, it is impossible to achieve exact pairs of images with the same underlying information. As such, the median filter of size 11 was applied to the dataset of 20 average frames.

A training set of 20 pairs of averaged frames were used for training the model. The model was written in PyTorch and was trained for 200 epochs at a fixed learning rate of 0.001. The training process took 10 minutes to complete. Subsequently, 5 median images in Fig. 4.6 were tested and the results shown in Fig. 4.7 shows the subsequent results.

Chapter 5

Discussion

5.1 Filter-based Denoising

The denoised images produced from the filter-based method, as seen in the right most column in Fig. 4.6, have revealed several artefacts. The thicker lines and the particle-like shapes can be interpreted as structurally evolved PEG-PLA vesicles, which do break down into such structures under applied heat. The advanced breakdown of the structure observed in the images could be explained by prior structural degradation of the vesicles. It is known that the temperature of the specimen was ramped up from room temperature to 50°C. However the duration of the temperature ramp is unknown, and it is possible that the vesicles had significantly degraded into micellar structures before the video sequence was recorded.

In addition, Fig. 5.1 and Fig. 5.2 are results from applying the same process to the video dataset at the 1 minute mark and at the end of the video at 2 minutes 45 seconds. The three results of each unique time period show that the denoised images have similar patterns but are not identical in arrangement across the frame. This could be due to the movement of the micellar structures from the flow of the liquid medium or from Brownian motion. Should the denoising process introduce additional artefacts onto the images, these additional artefacts would be consistently present across all processed images. Hence this is further evidence that the denoising process is not merely introducing artefacts into the image, but it is identifying existing artefacts based on the noisy input images.

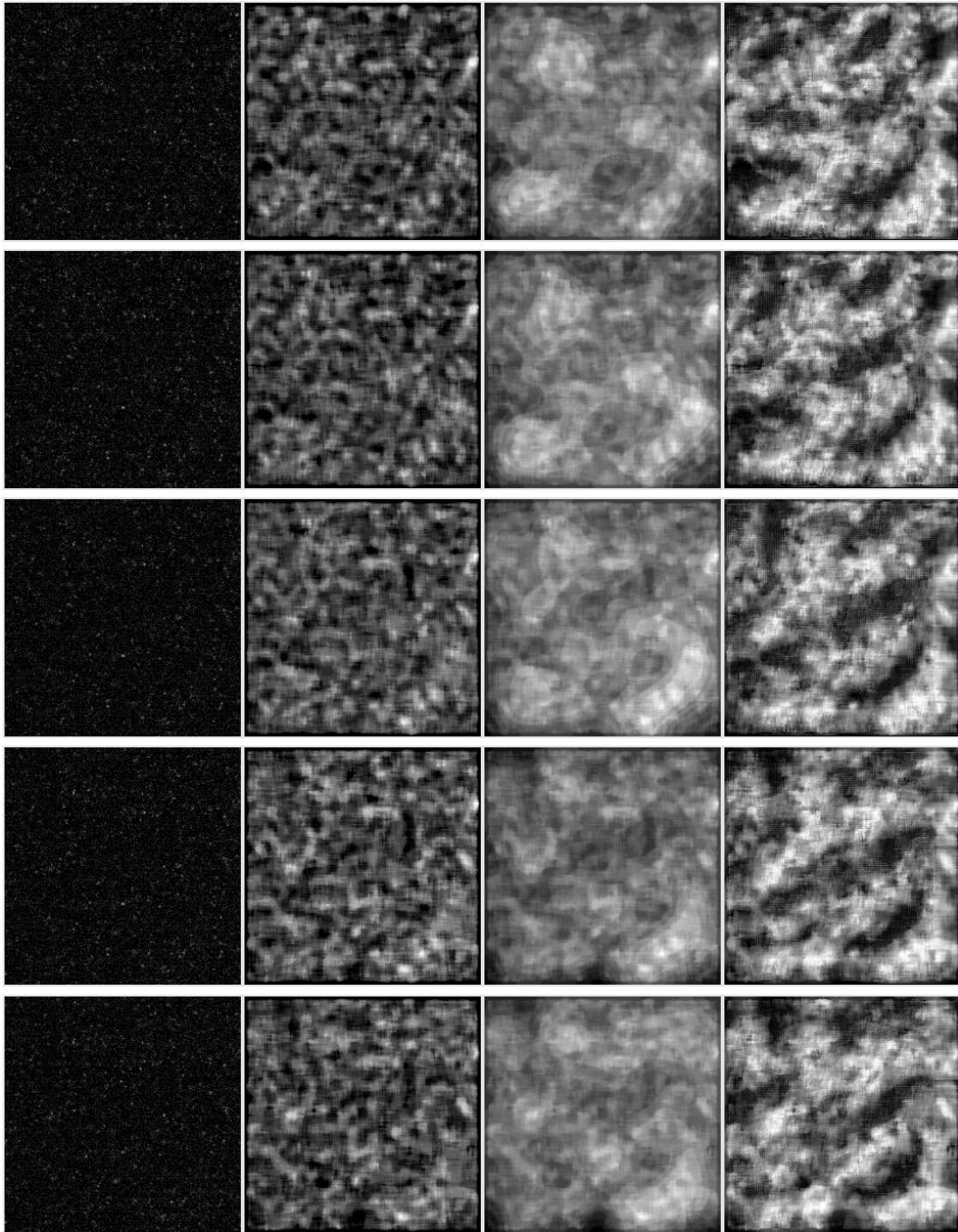


Figure 5.1: Collated images from raw image, median filtered, PID and deblurring, and local deconvolution from filter-based denoising. Each row represents consecutive average windows (ie. Frames 1-5, 2-6, 3-7, 4-8, 5-9) at Minute 1, Second 0 - 2.

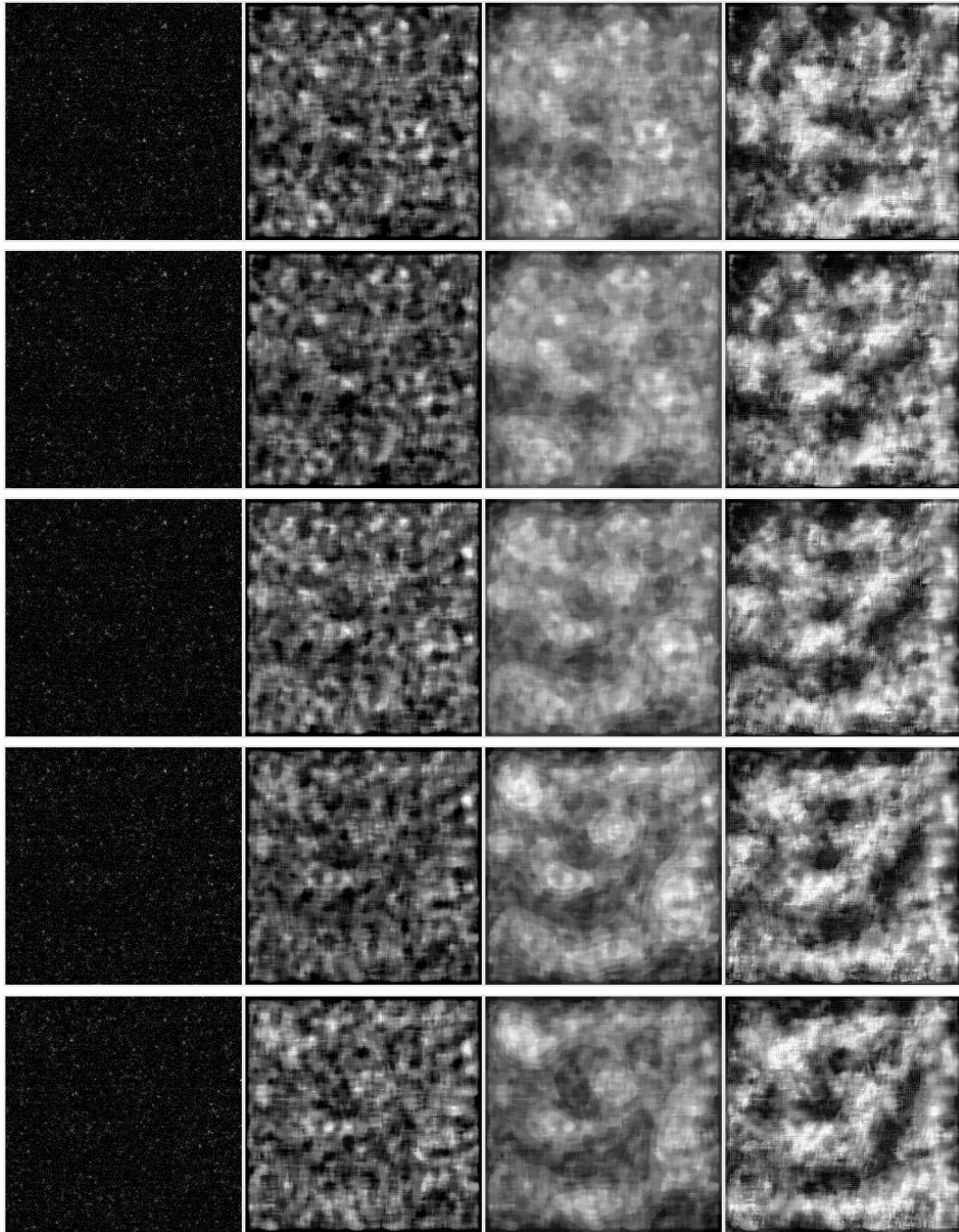


Figure 5.2: Collated images from raw image, median filtered, PID and deblurring, and local deconvolution from filter-based denoising. Each row represents consecutive average windows (ie. Frames 1-5, 2-6, 3-7, 4-8, 5-9) at Minute 2, Second 45 - 47.

5.2 Noise2Noise

In contrast, the images produced from the Noise2Noise model did not present a significant improvement in the image quality. The failure of the model could be attributed to both the extreme noise level present in the images, and the lack of proper paired images that contained the same underlying information. In the training dataset provided, 20 frames were averaged to produce consecutive noisy image pairs. Despite the relatively large average window size, it is possible that the averaged images did not contain sufficiently similar underlying information to be properly denoised. In addition, the large average window size could possibly have blurred any significant edges that had been present in the individual frames, hence the underlying information could have already been blurred prior to denoising. If a smaller average window size had been used, the underlying information of consecutive average frames would not be similar, rendering it impossible for the model to recognise the pixel value patterns across the image pairs. In contrast, this method would be better suited for other TEM imaging datasets since the specimen is typically static and this would be further helped by a high rate of image acquisition. In particular, Cryo-EM utilises static specimen, hence a video sequence of such a data would possibly be suitable for this denoising model.

5.3 Conclusion

In conclusion, traditional supervised machine learning methods are not applicable for denoising extremely noisy LPEM images since clean and noise-free images are unavailable. Therefore filter-based and unsupervised machine learning techniques were explored. A combination of filtering techniques and deconvolutions prove to be better suited than unsupervised learning in denoising extremely noisy LPEM images. Median filtering and PID are effective in retrieving structural information from such frames. However, such methods would not be successful without first averaging consecutive individual frames in order to increase the probability of information-containing pixels and decreasing the probability of noisy pixels appearing. The local deconvolution step has also helped in sharpening the edges of artefacts and identifying their structure. In comparison, unsupervised machine learning has not yielded such success and there is no significant improvement in the image quality. This could be due to the fact that the unsupervised model relies on pairs of noisy images with the same underlying information. Due to the fast movement of the specimen in the liquid medium during image acquisition, such pairs could not be obtained through frame averaging without losing significant information. Therefore hand-tuned filter-based denoising techniques are a better option for denoising extremely noisy images.

5.4 Future Works

As an improvement to the results presented, future works can include the combination of several denoising techniques in order to achieve a better overall result. For example, a prior data preprocessing step using PID could potentially reduce the noise in the images before being fed into an unsupervised denoising model. Alternatively, if time and cost permits, noisy image sample pairs of metallic nanoparticles such as copper or gold nanoparticles can be placed in the same LPEM environment but with a higher electron dose. A similar video recording can be taken. By doing so, the validity of the tested methods can be further confirmed, since the exact shape of the nanoparticles are known before conducting the experiment. Lastly, a slow and controlled rate of flow in an LPEM would allow the observed specimen to move slowly across the image. Coupling this slow rate of movement with a high frame rate of image acquisition would possibly improve the results from fully unsupervised denoising techniques such as Noise2Noise, since pairing consecutive frames would better represent the same underlying information.

References

- [1] “Elements of a Transmission Electron Microscope,” in *Transmission Electron Microscopy: Physics of Image Formation*, New York, NY: Springer New York, 2008, pp. 75–138, ISBN: 978-0-387-34758-5. DOI: 10.1007/978-0-387-40093-8_4. [Online]. Available: https://doi.org/10.1007/978-0-387-40093-8_4.
- [2] D. B. Williams and C. B. Carter, *Transmission Electron Microscopy, A Textbook for Materials Science*. Springer, 2009.
- [3] S. Sadamatsu, M. Tanaka, K. Higashida, and S. Matsumura, “Transmission electron microscopy of bulk specimens over 10 μm in thickness,” vol. 162, pp. 10–16, Mar. 2016. DOI: 10.1016/j.ultramic.2015.09.001. [Online]. Available: <https://doi.org/10.1016/j.ultramic.2015.09.001>.
- [4] *Cryo transfer tomography holder – model 2550*, <https://www.eden-instruments.com/en/in-operando-equipments/fischione/cryo-transfer-tomography-holder-model-2550/>, Accessed: 09-07-2021.
- [5] S. Pu, C. Gong, and A. W. Robertson, “Liquid cell transmission electron microscopy and its applications,” *Royal Society Open Science*, vol. 7, no. 1, p. 191204, Jan. 2020. DOI: 10.1098/rsos.191204. [Online]. Available: <https://doi.org/10.1098/rsos.191204>.
- [6] A. Ianiro, H. Wu, M. M. J. van Rijt, M. P. Vena, A. D. A. Keizer, A. C. C. Esteves, R. Tuinier, H. Friedrich, N. A. J. M. Sommerdijk, and J. P. Patterson, “Liquid–liquid phase separation during amphiphilic self-assembly,” vol. 11, no. 4, pp. 320–328, Feb. 2019. DOI: 10.1038/s41557-019-0210-4. [Online]. Available: <https://doi.org/10.1038/s41557-019-0210-4>.
- [7] L. E. Franken, K. Grünwald, E. J. Boekema, and M. C. A. Stuart, “A technical introduction to transmission electron microscopy for soft-matter: Imaging, possibilities, choices, and technical developments,” *Small*, vol. 16, no. 14, p. 1906198, Apr. 2020. DOI: 10.1002/smll.201906198. [Online]. Available: <https://doi.org/10.1002/smll.201906198>.

- [8] K. Yadav, D. Mohan, and A. S. Parihar, “Image detection in noisy images,” in *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*, 2021, pp. 917–923. doi: 10.1109/ICICCS51141.2021.9432243.
- [9] L. Fan, F. Zhang, H. Fan, and C. Zhang, “Brief review of image denoising techniques,” vol. 2, no. 1, Jul. 2019. doi: 10.1186/s42492-019-0016-7. [Online]. Available: <https://doi.org/10.1186/s42492-019-0016-7>.
- [10] H. S. Kushwaha, S. Tanwar, K. Rathore, and S. Srivastava, “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.
- [11] J. Wang, J. Bi, L. Wang, and X. Wang, “A non-reference evaluation method for edge detection of wear particles in ferrograph images,” vol. 100, pp. 863–876, Feb. 2018. doi: 10.1016/j.ymssp.2017.08.014. [Online]. Available: <https://doi.org/10.1016/j.ymssp.2017.08.014>.
- [12] N. Ofir, M. Galun, S. Alpert, A. Brandt, B. Nadler, and R. Basri, “On detection of faint edges in noisy images,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 4, pp. 894–908, Apr. 2020. doi: 10.1109/tpami.2019.2892134. [Online]. Available: <https://doi.org/10.1109/tpami.2019.2892134>.
- [13] J. M. Ede, “Deep learning in electron microscopy,” *Machine Learning: Science and Technology*, vol. 2, no. 1, p. 011004, Mar. 2021. doi: 10.1088/2632-2153/abd614. [Online]. Available: <https://doi.org/10.1088/2632-2153/abd614>.
- [14] *Tutorial 1: Image filtering*, <https://ai.stanford.edu/~syyeung/cvweb/tutorial1.html>, Accessed: 04-03-2021.
- [15] G. Marchello, C. D. Pace, A. Duro-Castano, G. Battaglia, and L. Ruiz-Pérez, “End-to-end image analysis pipeline for liquid-phase electron microscopy,” *Journal of Microscopy*, vol. 279, no. 3, pp. 242–248, Mar. 2020. doi: 10.1111/jmi.12889. [Online]. Available: <https://doi.org/10.1111/jmi.12889>.
- [16] R. Raju, T. Maul, and A. Bargiela, “New image processing pipelines for membrane detection,” vol. 3, no. 1, pp. 15–23, Jan. 2015. doi: 10.12792/jiiae.3.15. [Online]. Available: <https://doi.org/10.12792/jiiae.3.15>.
- [17] M. A. Gungor and I. Karagoz, “The effects of the median filter with different window sizes for ultrasound image,” IEEE, Oct. 2016. doi: 10.1109/compcomm.2016.7924761. [Online]. Available: <https://doi.org/10.1109/compcomm.2016.7924761>.

- [18] *Fast fourier transform (fft) background*, https://www.l3harrisgeospatial.com/docs/_backgroundfast-fouriertransform.html, Accessed: 01-06-2021.
- [19] N. Bonnet and P. Vautrot, “Image analysis: Is the fourier transform becoming obsolete?,” vol. 8, no. 1, pp. 59–75, 1997. DOI: [10.1051/mmm:1997106](https://doi.org/10.1051/mmm:1997106). [Online]. Available: <https://doi.org/10.1051/mmm:1997106>.
- [20] M. Z. Alom, T. M. Taha, C. Yakopcic, S. Westberg, P. Sidike, M. S. Nasrin, M. Hasan, B. C. V. Essen, A. A. S. Awwal, and V. K. Asari, “A state-of-the-art survey on deep learning theory and architectures,” vol. 8, no. 3, p. 292, Mar. 2019. DOI: [10.3390/electronics8030292](https://doi.org/10.3390/electronics8030292). [Online]. Available: <https://doi.org/10.3390/electronics8030292>.
- [21] K. Pawar and V. Z. Attar, “Assessment of autoencoder architectures for data representation,” in, Springer International Publishing, Oct. 2019, pp. 101–132. DOI: [10.1007/978-3-030-31756-0_4](https://doi.org/10.1007/978-3-030-31756-0_4). [Online]. Available: https://doi.org/10.1007/978-3-030-31756-0_4.
- [22] J. M. Ede and R. Beanland, “Improving electron micrograph signal-to-noise with an atrous convolutional encoder-decoder,” vol. 202, pp. 18–25, Jul. 2019. DOI: [10.1016/j.ultramic.2019.03.017](https://doi.org/10.1016/j.ultramic.2019.03.017). [Online]. Available: <https://doi.org/10.1016/j.ultramic.2019.03.017>.
- [23] J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, and T. Aila, “Noise2Noise: Learning image restoration without clean data,” in *Proceedings of the 35th International Conference on Machine Learning*, J. Dy and A. Krause, Eds., ser. Proceedings of Machine Learning Research, vol. 80, PMLR, Jul. 2018, pp. 2965–2974. [Online]. Available: <https://proceedings.mlr.press/v80/lehtinen18a.html>.
- [24] A. Krull, T.-O. Buchholz, and F. Jug, “Noise2void - learning denoising from single noisy images,” IEEE, Jun. 2019. DOI: [10.1109/cvpr.2019.00223](https://doi.org/10.1109/cvpr.2019.00223). [Online]. Available: <https://doi.org/10.1109/cvpr.2019.00223>.
- [25] A. S. Goncharova, A. Honigmann, F. Jug, and A. Krull, “Improving blind spot denoising for microscopy,” in *Computer Vision – ECCV 2020 Workshops*, Springer International Publishing, 2020, pp. 380–393. DOI: [10.1007/978-3-030-66415-2_25](https://doi.org/10.1007/978-3-030-66415-2_25). [Online]. Available: https://doi.org/10.1007/978-3-030-66415-2_25.
- [26] A. Papini, “A new algorithm to reduce noise in microscopy images implemented with a simple program in python,” vol. 75, no. 3, pp. 334–342, Sep. 2011. DOI: [10.1002/jemt.21062](https://doi.org/10.1002/jemt.21062). [Online]. Available: <https://doi.org/10.1002/jemt.21062>.

- [27] C. Knaus and M. Zwicker, “Progressive image denoising,” vol. 23, no. 7, pp. 3114–3125, Jul. 2014. DOI: 10.1109/tip.2014.2326771. [Online]. Available: <https://doi.org/10.1109/tip.2014.2326771>.
- [28] ——, “Dual-domain image denoising,” IEEE, Sep. 2013. DOI: 10.1109/icip.2013.6738091. [Online]. Available: <https://doi.org/10.1109/icip.2013.6738091>.