

CONVOLUTIONAL NEURAL NETWORK DENOISING OF FOCUSED ION BEAM MICROGRAPHS

Minxu Peng*, Mertcan Cokbas*, Unay Dorken Gallastegi*,
Prakash Ishwar, Janusz Konrad, Brian Kulis, and Vivek K Goyal

Department of Electrical and Computer Engineering, Boston University, USA
{mxpeng, mcokbas, udorken, pi, jkonrad, bkulis, goyal}@bu.edu

ABSTRACT

Most research on deep learning algorithms for image denoising has focused on signal-independent additive noise. Focused ion beam (FIB) microscopy with direct secondary electron detection has an unusual Neyman Type A (compound Poisson) measurement model, and sample damage poses fundamental challenges in obtaining training data. Model-based estimation is difficult and ineffective because of the nonconvexity of the negative log likelihood. In this paper, we develop deep learning-based denoising methods for FIB micrographs using synthetic training data generated from natural images. To the best of our knowledge, this is the first attempt in the literature to solve this problem with deep learning. Our results show that the proposed methods slightly outperform a total variation-regularized model-based method that requires time-resolved measurements that are *not* conventionally available. Improvements over methods using conventional measurements and less accurate noise modeling are dramatic – around 10 dB in peak signal-to-noise ratio.

Index Terms—image denoising, CNN, particle beam microscopy, Neyman Type A distribution

1. INTRODUCTION

Focused ion beam (FIB) microscopy plays a crucial role in imaging fine sample structures at sub-nanometer resolution [1–4]. A focused beam of ions is raster scanned over the sample. The interaction of incident ions with the sample produces secondary electrons (SEs) to be detected. The mean number of SEs per ion reveals information such as the composition and topography of sample components.

Measurements obtained by FIB microscopy are inherently noisy due to randomness in both the numbers of incident ions and the numbers of detected SEs per incident ion. Each of these sources of randomness is conventionally modeled with Poisson distributions, resulting in a Neyman Type A distribution for measurements that is detailed in Section 3. The signal-dependent noise level of this model is illustrated in Fig. 1. Existing image denoising techniques fare poorly on FIB micrographs because they are developed for Poisson measurements or measurements corrupted by additive white Gaussian noise (AWGN).

In many types of imaging, one may improve quality by using averaging to reduce noise variance. This often takes the form of increased acquisition time. In FIB microscopy, this corresponds to increasing the *dose*, i.e., the mean number of incident ions per pixel λ , which may be achieved by increasing the product of beam current and pixel dwell time. However, various studies have shown that sample damage due to sputtering and radiation increases with dose [5–7]. For dose-sensitive samples, this introduces a limit to the image quality that can be achieved without sophisticated data processing. Sample damage also makes it fundamentally infeasible to collect clean-noisy FIB micrograph pairs for training-based methods. Even for samples less prone to damage, acquiring these image pairs would be expensive and time consuming.

The concept of time-resolved (TR) measurement that has recently been introduced to FIB microscopy [8] can be viewed as a way to make measurements more informative without increasing dose. For each pixel, dwell time t is divided into n sub-acquisitions of length t/n , and the number of SEs is counted within each sub-acquisition. Estimation methods developed for TR measurements show significant improvements that are approximately equivalent to making the source produce a deterministic number of ions – *mitigation of source shot noise* [8, 9]. Note that these methods have only been developed for unregularized (pixel-by-pixel) estimation, and they require unconventional data collection. Though measuring n sub-acquisitions at each pixel is feasible, it can increase total imaging time, which can increase cost and susceptibility to sample motion. To achieve similar improvement without TR measurement is thus of interest.

In this work, we use our knowledge of an accurate generative model for FIB microscope data to develop a deep learning-based method that significantly outperforms methods intended for AWGN or Poisson noise removal – without having a database of high-quality FIB micrographs. We adopt a state-of-the-art denoising convolutional neural network (DnCNN) [10], developed mainly for the AWGN denoising task. We train our network on noisy natural images where the noise is synthetically generated according to the FIB microscope model. We also extend the network architecture by adding a VGG-16 network to incorporate perceptual loss. Through extensive experiments, we show that deep learning-based denoising algorithms can be effective on FIB micrographs. Furthermore, we demonstrate that perceptual loss may help to preserve the structure in FIB images during denoising. The deep learning-based methods trained with physically-accurate synthetic noise – operating only on *conventional data* – perform slightly better than total variation-regularized maximum likelihood estimation applied to TR data.

* Denotes equal contribution. This work was supported in part by the U.S. National Science Foundation under Grant No. 1815896. The authors gratefully acknowledge that the computational work reported in this paper was performed on the Shared Computing Cluster administered by Boston University’s Research Computing Services (URL: www.bu.edu/tech/support/research/). The authors thank Professor Ashok Cutkosky for insightful discussions. Corresponding author: Minxu Peng.

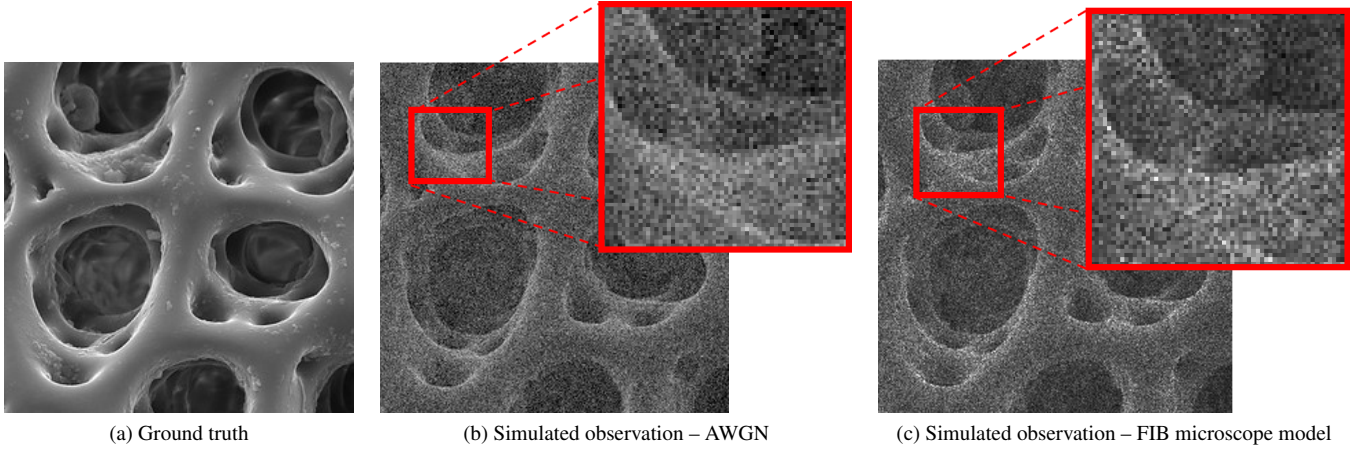


Fig. 1: Examples of different noise models on a microscopic image. (a) A clean image. (b) A simulation with additive white Gaussian noise. (c) A simulation with a physically accurate model for FIB microscopy. Note that the noise variance is signal dependent to a greater extent than with Poisson data.

2. RELATED WORK

Traditional methods. Various techniques have been proposed to recover a clean image from a noisy observation based on a measurement model and a regularizer or prior. For instance, total variation (TV) regularization [11] is very broadly successful in reducing noise without excessive smoothing of sharp edges. We include a comparison with a TV-regularized method here.

Deep neural network for image denoising. Deep neural networks [10, 12–14] have received increasing attention for the past decade. DnCNN [10] combines a very deep neural network [15], batch normalization [16] and residual learning [17] for a series of computer vision tasks, including image denoising for Gaussian noise, super-resolution, and deblurring. DnCNN is based on the assumption that the residual mapping is easier to learn than the original signal of interest, where residual is defined as the difference between the noisy observation and ground truth image. The network employs a single residual unit to output the residual image and exhibits a boost of performance from the contribution of batch normalization and residual learning.

Perceptual loss. Pixelwise difference losses have achieved successes in recovering images from distortions. However, they fail to capture the perceptual quality of images and often result in over-smoothed reconstructions. Perceptual loss [18] quantifies perceptual differences between images in a feature space and can be used to increase the spatial-structure fidelity between the ground truth and restored images. It is commonly used in style transfer [19] and super resolution [18] and can be adapted to image denoising [20]. A pre-trained convolutional neural network (e.g., VGG-16 [15]) is employed to extract high-level features of the output from the denoising neural network to enhance the perceptual appearance with respect to the ground truth image.

3. DATA MODEL AND MODEL-BASED ESTIMATION

In this section, we introduce the measurement model for FIB microscopy with noiseless direct SE detection. Compared with commercially prevalent use of scintillators and photomultiplier tubes, direct SE detection offers higher signal-to-noise ratio [21] and is

easier to model. This measurement model underlies both model-based reconstruction and our generation of synthetic noisy images for training and testing of learning-based reconstruction. We assume a square micrograph with J pixels and use $\mathcal{J} = \{1, \dots, \sqrt{J}\}^2$ to denote the pixel index set. Here the focus is on the SE yield image $\boldsymbol{\eta} \in [0, \infty)^{\mathcal{J}}$ and measurements that allow estimation of this image. These are related to network inputs and outputs in Section 5.2.

Measurement model. Our model for measurements is separable across the pixels, so we present it here without pixel indexing. During a fixed dwell time t , ion incidences follow a Poisson process. The number of incident ions M is a hidden Poisson random variable with mean $\lambda = \Lambda t$, where Λ represents the rate of incident ions per unit time. Dose λ is assumed to be equal for all the pixels. Incident ion i causes Z_i number of SEs to be detected. Each Z_i is Poisson distributed with mean η , which is called the SE yield. Both uses of Poisson distributions are well-established in the particle beam microscopy literature [22]. The total detected SEs $Q = \sum_{i=1}^M Z_i$ is a Neyman Type A random variable with probability mass function

$$P_Q(q; \lambda | \eta) = \frac{e^{-\lambda} \eta^q}{q!} \sum_{m=0}^{\infty} \frac{(\lambda e^{-\eta})^m m^q}{m!}, \quad (1)$$

mean

$$\mathbb{E}[Q] = \lambda \eta, \quad (2)$$

and variance

$$\text{Var}(Q) = \lambda \eta (\eta + 1). \quad (3)$$

Note that both mean and variance depend on η , making the noise neither additive nor white. Furthermore, the relationship between mean and variance is substantially different than for Poisson-distributed data, assuming η is not too small. Combining measurements across pixels gives $Q \in \mathbb{N}^{\mathcal{J}}$. A comparison of AWGN and this microscopic noise model can be seen in Figure 1.

If each pixel dwell time t is split over n sub-acquisitions, at each pixel we obtain an SE count vector of length n with independent entries $V^{(k)}$, $k \in \{1, \dots, n\}$, distributed according to (1) with λ replaced by λ/n . When λ/n is small, the measurement model for the TR vector can be approximated well by using only the $m = 0$ and $m = 1$ terms in (1). The counts from the sub-acquisitions

together comprise the full SE count above, so

$$V_j^{(1)} + V_j^{(2)} + \dots + V_j^{(n)} = Q_j,$$

for any pixel index $j \in \mathcal{J}$. Thus, a TR dataset is $\mathbf{V} \in \mathbb{N}^{J \times \{1, \dots, n\}}$, where summing over the last dimension of the tensor gives non-TR dataset $\mathbf{Q}$.

Model-based estimation. At any pixel, scaling the detected SE counts Q by total dose λ yields the unbiased estimator

$$\hat{\eta}_{\text{conventional}} = Q/\lambda, \quad (4)$$

which has mean-squared error (MSE) $\eta(\eta + 1)/\lambda$. The $(\eta + 1)$ factor in the MSE is rooted in the $(\eta + 1)$ factor in (3) by which the variance of a Neyman Type A random variable exceeds the variance of a Poisson random variable with the same mean. This cost of randomness of the number of incident ions can be mitigated by TR measurement [8, 9]. The pixelwise separable maximum likelihood (ML) estimator using the TR data,

$$\hat{\eta}_{\text{TR}} = \arg \max_{\eta \in [0, \infty)} \prod_{k=1}^n P_Q(v^{(k)}; \lambda/n | \eta), \quad (5)$$

has been shown empirically to realize most of the performance improvement from TR measurement that is predicted by Fisher information [9].

Regularized estimation without TR measurement is made difficult by the series form of (1) and the non-convexity of its negative logarithm. With TR measurement, provided λ/n is sufficiently small, the derivative of $\log P_Q(v^{(k)}; \lambda/n | \eta)$ with respect to η can be approximated efficiently. This allows us to compute the TV-regularized ML estimator for a full SE yield image,

$$\arg \min_{\eta \in [0, \infty)^J} \left[- \sum_{j \in \mathcal{J}} \sum_{k=1}^n \log P_Q(v_j^{(k)}; \lambda/n | \eta_j) + w_{\text{TV}} \|\eta\|_{\text{TV}} \right], \quad (6)$$

where w_{TV} is the regularization weight on the 2D TV norm $\|\cdot\|_{\text{TV}}$.

4. DEEP LEARNING ARCHITECTURES AND LOSS FUNCTIONS

To reconstruct the clean image $\mathbf{y} \in \mathbb{R}^{\mathcal{J}}$ from noisy observation $\mathbf{x} \in \mathbb{R}^{\mathcal{J}}$, we use a feed-forward convolutional neural network called DnCNN [10] as the backbone architecture, which can be seen in Figure 2. We modified the original network architecture for our method. We pass the predicted image $\hat{\mathbf{y}}$ through a ‘sigmoid’ function to ensure that $\hat{\mathbf{y}} \in [0, 1]^{\mathcal{J}}$. In the remainder of this section, we discuss three different types of loss functions that we combined separately with DnCNN [10].

4.1. Direct Loss

One natural way to apply a deep learning method to an image denoising problem is to estimate the clean image directly from the noisy image. In this approach, loss is measured on the difference between the clean image and estimated image. We use the MSE metric:

$$L(\mathbf{y}, \hat{\mathbf{y}}) = \|\mathbf{y} - \hat{\mathbf{y}}\|_F^2. \quad (7)$$

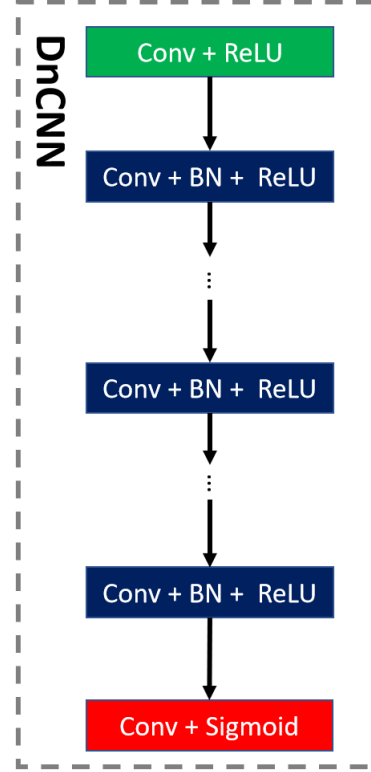


Fig. 2: Denoising network architecture proposed by Zhang et al. [10]. The network consists of 17 layers: the first layer is composed of a convolutional layer followed by a ReLU; the intermediate 15 layers are composed of a convolutional layer followed by batch normalization and a ReLU; and the last layer consists solely a convolutional layer. The output of the network is modified to be either an estimated clean image or an estimated residual image.

4.2. Residual Loss

In DnCNN [10], the loss function is designed to estimate the residual between the noisy and clean image instead of estimating the image directly. Let the ground-truth residual $\mathbf{r}$ be defined as

$$\mathbf{r} = \mathbf{x} - \mathbf{y}, \quad (8)$$

and let $\hat{\mathbf{r}} = \mathbf{x} - \hat{\mathbf{y}}$ be the estimated residual. The residual loss function is calculated as the MSE between the ground-truth and estimated residuals:

$$L(\mathbf{r}, \hat{\mathbf{r}}) = \|\mathbf{r} - \hat{\mathbf{r}}\|_F^2. \quad (9)$$

4.3. Perceptual Loss

In [18], Johnson et al. proposed a perceptual loss network. This loss network is used to extract descriptive features of the images. From different intermediate layers of this loss network, features with different characteristics can be extracted. For each of the chosen layers, we compute the MSE between the feature vectors of the recovered image from DnCNN [10] and the clean image. Perceptual loss is then defined as

$$L(\mathbf{y}, \hat{\mathbf{y}}) = \sum_{i=1}^{\ell} w_i \|F_i(\mathbf{y}) - F_i(\hat{\mathbf{y}})\|^2, \quad (10)$$

where ℓ denotes the number of different layers chosen to extract features from the loss network, $F_i(\cdot)$ denotes the feature vector coming from the i th layer of the loss network, and w_i represents its weight. We select the weights empirically. For our application, we choose VGG-16 [15] as our loss network. The layers for extracting the feature vectors are kept the same as in [18].

5. DATA AND EXPERIMENTS

5.1. Datasets

Due to lack of a FIB micrograph denoising dataset, we use regular camera images to train our network. We used the Berkeley Segmentation Data Set and Benchmarks 500 (BSDS500) [23] for training. We use 400 images for training and the remaining 100 images for validation. We use the techniques introduced in the next section to distort the images for training, validation and testing. Distorted images will serve as the input, and the clean images will serve as the ground truth.

Our testing is two-fold. The first testing phase is quantitative using two benchmark datasets: the Berkeley segmentation dataset (BSD68) [24] composed of 68 images and the 12-image dataset used in [10]. The second testing phase is qualitative using the microscopic image shown in Figure 3. For the chosen microscopic image we also provide quantitative results.

5.2. Data Generation and Experimental Setup

In this work, three sets of experiments vary the training loss, vary the synthetic generation of training data, and provide model-based reconstructions for comparison. In all trained networks, we use the same network parameters (number of layers, number of channels, batch normalization, etc.) as the original DnCNN [10]. For optimization, we use Adam optimizer with a learning rate of 5×10^{-5} . We schedule a learning rate decay at every 20 epochs by a factor of 0.2. Overall, we use 100 epochs to train each model. Experiments are performed on a PC with Intel(R) Core(TM) i7-6700K CPU at 4.00 Ghz and Nvidia GeForce GTX 1070. Training a single network takes approximately 16 hours on a GPU.

The first set of experiments compares the three networks discussed in Section 4, all trained with noisy images generated following the data model described in Section 3. We refer to these methods as **CNN-Direct**, **CNN-Residual**, and **CNN-Perceptual**. Each clean image $y \in [0, 1]$ is scaled through

$$\eta = 6y + 2 \quad (11)$$

to give ground-truth SE yield values in the interval $[2, 8]$ to match the range suggested in [1]. Given the SE yield image η , an image of total detected SEs Q is generated for ion dose $\lambda = 20$ following the model described in (1). The final noisy image x is $Q/320$, with the scaling chosen so that most pixels in most realizations are in the range $[0, 1]$.

The second set of experiments is designed to demonstrate the impact of lacking an accurate noise model for FIB microscopy. We present results only with the direct loss of Section 4.1; the residual and perceptual loss give worse quantitative performance and no additional insights. We maintain the image scaling of (11) and the dose of $\lambda = 20$, but we model the data generation in three ways. If the source beam provided λ incident particles deterministically, the number of SEs would follow a Poisson distribution with mean $\lambda\eta$. Without the physical justification of a deterministic source beam, one might also use a Poisson model out of naivety or because it

is a suitable model for scanning electron microscopy. We refer to the denoising method trained with Poisson data as **CNN-Direct-Poisson**. Two additional networks are trained using images distorted by AWGN. Let $\bar{\eta}$ denote the mean over the pixels of the SE yield image η . Then training with AWGN variance $\lambda\bar{\eta}$ is called **CNN-Direct-AWGN-Poisson** because it matches the average variance to Poisson-distributed data, and training with AWGN variance $\lambda\bar{\eta}(1 + \bar{\eta})$ is called **CNN-Direct-AWGN-Neyman** because it matches the variance to (3).

The final set of experiments simulates performance of the model-based methods. The values of η and λ are unchanged from above. We refer to the estimator in (4) as **Conventional**. The total dose λ is split over $n = 100$ sub-acquisitions for TR methods. We refer to the estimator in (5) as **TR-Unregularized** and to the estimator in (6) as **TR-TV**. The value of $w_{TV} = 1.4$ is optimized to minimize MSE.

6. RESULTS AND DISCUSSION

In this section, we present qualitative and quantitative results for all the experiments discussed in Section 5.2. Table 1 shows quantitative results on methods tested on noisy images corrupted following the model in (1) using peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM) metrics. The best-performing algorithm with respect to both evaluation metrics is CNN-Direct. We also observe that the performance of CNN-Residual is very close to CNN-Direct. Recalling that the Neyman Type A distribution gives signal-dependent variance, this finding suggests that learning based on residual loss is less effective when the residual is not signal-independent. Moreover, both CNN-Direct and CNN-Residual outperform CNN-Perceptual in the stated metrics. However, these metrics do not always align with human perception. For instance, based on the qualitative results in Figure 3, it is arguable that CNN-Direct oversmooths somewhat. Depending on the application, loss of details such as textures and edges may be not acceptable. For assessing how well the texture is preserved, we can focus on the red-framed section of the images and any detailed part of this mesh structure. Perceptual loss may be the best method in terms of preserving the texture and edges.

Among the three methods representing less accurate modeling of the FIB microscopy data, CNN-Direct-AWGN-Poisson achieves the worst performance. This modeling accounts for neither the particle nature of the data nor the correct signal-dependent variance model (3). Slightly better performance – though still more than 10 dB worse than CNN-Direct – is obtained with CNN-Direct-Poisson. Both CNN-Direct-Poisson and CNN-Direct-AWGN-Poisson underestimate the variance of the measurements by a large margin. By matching the average variance to (3), CNN-Direct-AWGN-Neyman achieves performance within about 2 dB to networks that are trained under the Neyman Type A model, which demonstrates the significance of the correct noise knowledge.

The improvement of TR-Unregularized over Conventional reconstruction is by about 6 dB, which is roughly consistent with previous theoretical and empirical results [8, 9]. Including TV regularization significantly improves the quantitative and visual performances; the PSNR increases by more than 7 dB and the SSIM is doubled. The TR-TV method achieves performance comparable with the best deep learning method, with PSNR difference less than 0.7 dB. While the use of TV regularization is not generally competitive with deep learning-based methods, here it is not a comparison based on the same input data; time-resolved measurements are fundamentally more informative.

Table 1: Performance comparison of the proposed algorithms on BSDS68 and 12-image dataset using PSNR and SSIM. The highest values for PSNR and SSIM for each dataset are shown in boldface. The last two columns are the performances for a single representative microscopic image in Figure 3. Rows of the table are different denoising algorithms specified in Section 5.2. The first three rows are for the networks that are discussed. The following three rows are for networks trained with less accurate synthetic noise. The last rows are for model-based algorithms.

	BSDS-68		12-image		Microscopic Image	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
CNN-Direct	27.57	0.844	28.19	0.885	28.58	0.889
CNN-Residual	27.53	0.842	28.11	0.882	28.52	0.886
CNN-Perceptual	24.99	0.793	25.58	0.829	25.83	0.823
CNN-Direct-Poisson	16.64	0.367	15.77	0.327	17.04	0.353
CNN-Direct-AWGN-Poisson	16.45	0.398	15.73	0.349	17.23	0.382
CNN-Direct-AWGN-Neyman	25.70	0.815	26.10	0.854	26.61	0.860
Conventional	14.28	0.201	13.67	0.180	15.01	0.173
TR-Unregularized	20.05	0.384	19.73	0.352	20.36	0.355
TR-TV	27.31	0.753	27.91	0.792	27.94	0.788

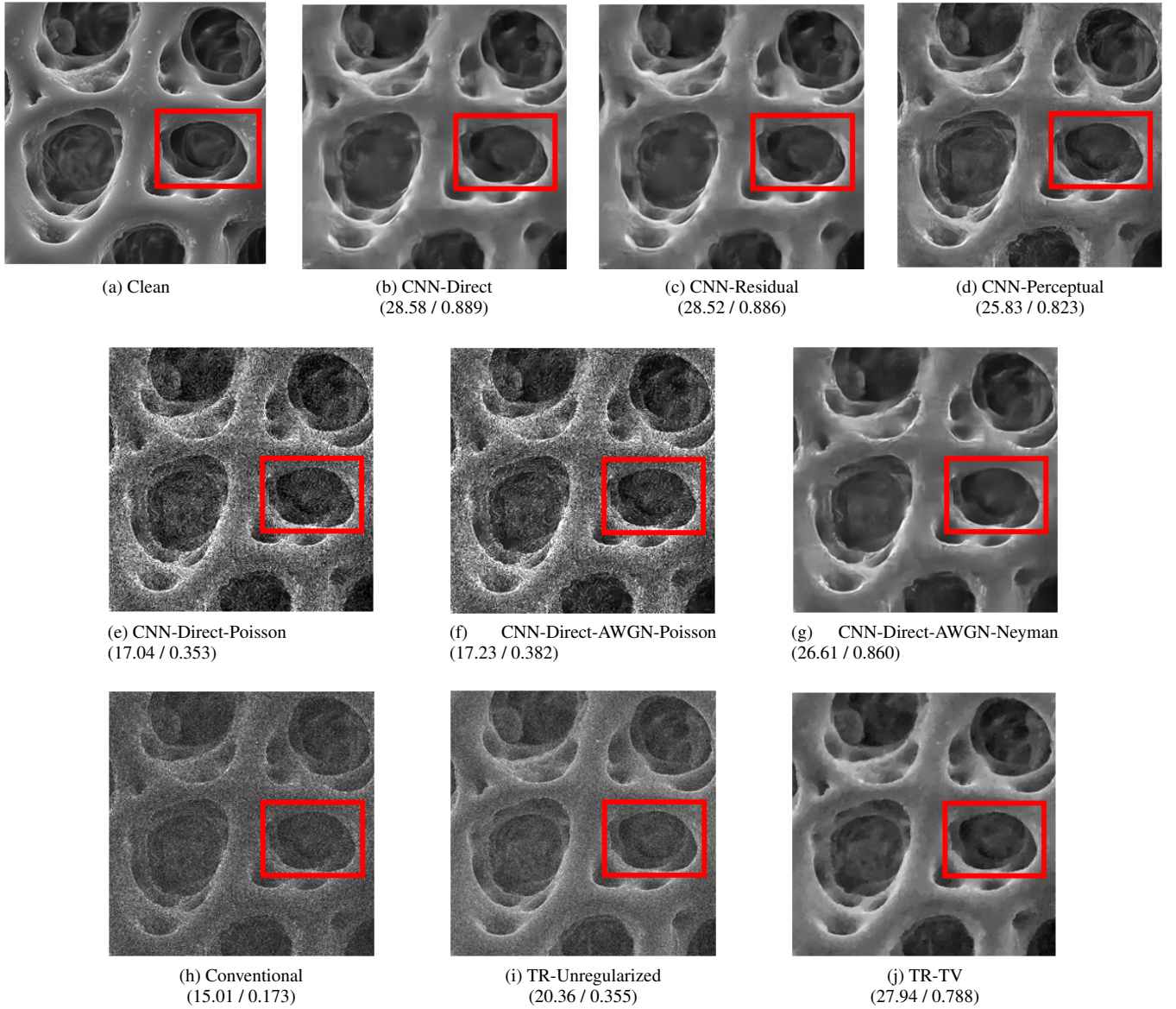


Fig. 3: Qualitative comparison of the algorithms on a microscopic image. Numbers in parentheses are PSNR in dB and SSIM values.

7. CONCLUSION

In this paper, we demonstrated how to apply deep learning-based image denoising algorithms on FIB microscopic image denoising despite a lack of training data from that imaging modality. We show that deep learning algorithms can achieve similar performance to a TV-regularized model-based method that requires unconventional time-resolved data. The original versions of the deep learning methods that we adapted were mainly tested for AWGN. In our work, we used the compound Poisson noise model that is physically accurate for FIB microscopy. We studied how the performance degrades under less accurate modeling of the noisy observations. Results suggest that CNN-Direct and CNN-Residual perform the best in terms of PSNR and SSIM. Furthermore, we show that perceptual loss can be used to preserve the structure.

We have provided results only for the ion dose level of $\lambda = 20$, which combines with the SE yield to determine the noise level. Among our areas of future work is to demonstrate the robustness of the proposed methods to a range of noise levels. In addition, data following a Neyman Type A distribution displays spatially variant noise related to the ground truth value, as shown in (3). The employed DnCNN neural network lacks the flexibility to deal with spatially variant noise, while we want to design a network that can handle such noise. Finally, our results reinforce the merit of having time-resolved data. We want to extend our work to be able to incorporate time-resolved measurements into learned-denoiser technology.

8. REFERENCES

- [1] J. Notte, R. Hill, S. McVey, L. Farkas, R. Percival, and B. Ward, "An introduction to helium ion microscopy," *Microscopy and Microanalysis*, vol. 12, no. S02, pp. 126–127, 2006.
- [2] D. Emmrich, A. Beyer, A. Nadzeyka, S. Bauerdick, J. Meyer, J. Kotakoski, and A. Götzhäuser, "Nanopore fabrication and characterization by helium ion microscopy," *Appl. Phys. Lett.*, vol. 108, no. 16, p. 163103, 2016.
- [3] M. S. Joens, C. Huynh, J. M. Kasuboski, D. Ferranti, Y. J. Sigal, F. Zeitvogel, M. Obst, C. J. Burkhardt, K. P. Curran, S. H. Chalasani, L. A. Stern, B. Goetze, and J. A. J. Fitzpatrick, "Helium ion microscopy (HIM) for the imaging of biological samples at sub-nanometer resolution," *Sci. Rep.*, vol. 3, no. 3514, 2013.
- [4] T. Wirtz, O. De Castro, J.-N. Audinot, and P. Philipp, "Imaging and analytics on the helium ion microscope," *Ann. Rev. Analytical Chem.*, vol. 12, pp. 523–543, 2019.
- [5] J. Orloff, L. W. Swanson, and M. Utlaut, "Fundamental limits to imaging resolution for focused ion beams," *J. Vacuum Sci. Technol. B*, vol. 14, no. 6, pp. 3759–3763, 1996.
- [6] R. Livengood, S. Tan, Y. Greenzweig, J. Notte, and S. McVey, "Subsurface damage from helium ions as a function of dose, beam energy, and dose rate," *J. Vacuum Sci. Technol. B*, vol. 27, no. 6, pp. 3244–3249, 2009.
- [7] V. Castaldo, C. W. Hagen, P. Kruit, E. Van Veldhoven, and D. Maas, "On the influence of the sputtering in determining the resolution of a scanning ion microscope," *J. Vacuum Sci. Technol. B*, vol. 27, no. 6, pp. 3196–3202, 2009.
- [8] M. Peng, J. Murray-Bruce, K. K. Berggren, and V. K. Goyal, "Source shot noise mitigation in focused ion beam microscopy by time-resolved measurement," *Ultramicroscopy*, vol. 211, p. 112948, 2020.
- [9] M. Peng, J. Murray-Bruce, and V. K. Goyal, "Time-resolved focused ion beam microscopy: Modeling, estimation methods, and analyses," *IEEE Trans. Comput. Imaging*, vol. 7, pp. 547–561, 2021.
- [10] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, 2017.
- [11] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlin. Phenom.*, vol. 60, no. 1-4, pp. 259–268, 1992.
- [12] K. Zhang, W. Zuo, and L. Zhang, "FFDNet: Toward a fast and flexible solution for CNN-based image denoising," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4608–4622, 2018.
- [13] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2018, pp. 9446–9454.
- [14] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep CNN denoiser prior for image restoration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2017, pp. 3929–3938.
- [15] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [16] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Int. Conf. Mach. Learn.* PMLR, 2015, pp. 448–456.
- [17] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2016, pp. 770–778.
- [18] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Proc. Europ. Conf. Comput. Vis.* Springer, 2016, pp. 694–711.
- [19] L. A. Gatys, A. S. Ecker, and M. Bethge, "A neural algorithm of artistic style," *arXiv preprint arXiv:1508.06576*, 2015.
- [20] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M. K. Kalra, Y. Zhang, L. Sun, and G. Wang, "Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss," *IEEE Trans. Medical Imaging*, vol. 37, no. 6, pp. 1348–1357, 2018.
- [21] S. Yamada, T. Ito, K. Gouhara, and Y. Uchikawa, "Electron-count imaging in SEM," *Scanning*, vol. 13, no. 2, pp. 165–171, 1991.
- [22] F. Timischl, M. Date, and S. Nemoto, "A statistical model of signal-noise in scanning electron microscopy," *Scanning*, vol. 34, no. 3, pp. 137–144, 2012.
- [23] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 33, no. 5, pp. 898–916, May 2011.
- [24] S. Roth and M. J. Black, "Fields of experts," *Int. J. Comput. Vis.*, vol. 82, no. 2, pp. 205–229, 2009.