

Removing Imaging Artifacts in Electron Microscopy using an Asymmetrically Cyclic Adversarial Network without Paired Training Data

Tran Minh Quan
UNIST, South Korea
quantm@unist.ac.kr

David Grant Colburn Hildebrand
The Rockefeller University, NY, USA
david@hildebrand.name

Kanggeun Lee
UNIST, South Korea
leekanggeun@unist.ac.kr

Logan A. Thomas
Harvard Medical School, MA, USA
logan.thomas@hms.harvard.edu

Aaron T. Kuan
Harvard Medical School, MA, USA
aaron_kuan@hms.harvard.edu

Wei-Chung Allen Lee
Harvard Medical School, MA, USA
wei-chung_lee@hms.harvard.edu

Won-Ki Jeong
UNIST, South Korea
wkjeong@unist.ac.kr

Abstract

We propose an asymmetrically cyclic adversarial network that performs denoising tasks to improve electron microscopy (EM) image analysis. Deep learning-based denoising methods have typically been trained either with matching pairs of noise-free and noise-corrupted images or by leveraging prior knowledge of noise distributions. Neither of these options is feasible in high-throughput EM imaging pipelines. Our proposed denoising method employs independently acquired noise-free, noise pattern, and noise-corrupted images to automatically learn the underlying noise model and generate denoised outputs. This method is based on three-way cyclic constraints with adversarial training of a deep network to improve the quality of acquired images without paired training data. Its utility is demonstrated for cases where imaging substrates add noise and where acquisition conditions contribute noise. We show that our method, which builds on the concept of CycleGAN, outperforms the current state-of-the-art denoising approaches Noise2Noise and Noise2Void, as well as other learning-based techniques.

1. Introduction

Electron microscopy (EM) is an important imaging technique for the neuroscience field of ‘connectomics’ because it is capable of resolving densely packed neuronal structures that are nanometers in size [15, 26]. Such high-resolution imaging traditionally required human interaction for sample preparation and microscope operation, which served as bot-

tlenecks in the data acquisition process. Recent advances in automatic tissue collection and imaging techniques, such as the automated tape-collecting ultramicrotome (ATUM) [12] and the transmission electron microscope camera array (TEMCA) [1, 24, 34, 37], significantly reduce these acquisition burdens to make peta-scale data collection feasible. With the benefits of these high-throughput techniques, however, comes the introduction of various artifacts that can make data analysis more challenging. For example, the electron absorption properties of some electron-lucent films used with recently developed high-throughput transmission EM (TEM) techniques [9, 17, 10] are not spatially homogeneous, causing coherent background noise (Film noise). Furthermore, charge damage artifacts are sometimes observed with scanning EM (SEM), where electrons can accumulate in the 30–60 nm thickness of sample tissue sections and can cause permanent damage that results in blob-like artifacts alongside Gaussian noise-like corruption (Charge noise). These imaging modalities and corresponding noise examples are illustrated in Fig. 1. Such EM image artifacts are difficult to remove using conventional denoising filters that were developed for specific noise models.

We present a novel semi-supervised learning-based denoising method that learns sources of noise (including artifacts) and removes them from unseen images. The proposed method consists of a three-way asymmetrically cyclic constraints in an adversarial deep network with two generators. One generator approximates the noise model (adding noise to a noise-free image), while the other approximates the inverse of the noise model (removing noise from a noisy image). Unlike other deep learning-based noise removal methods that use supervised training to minimize the dif-

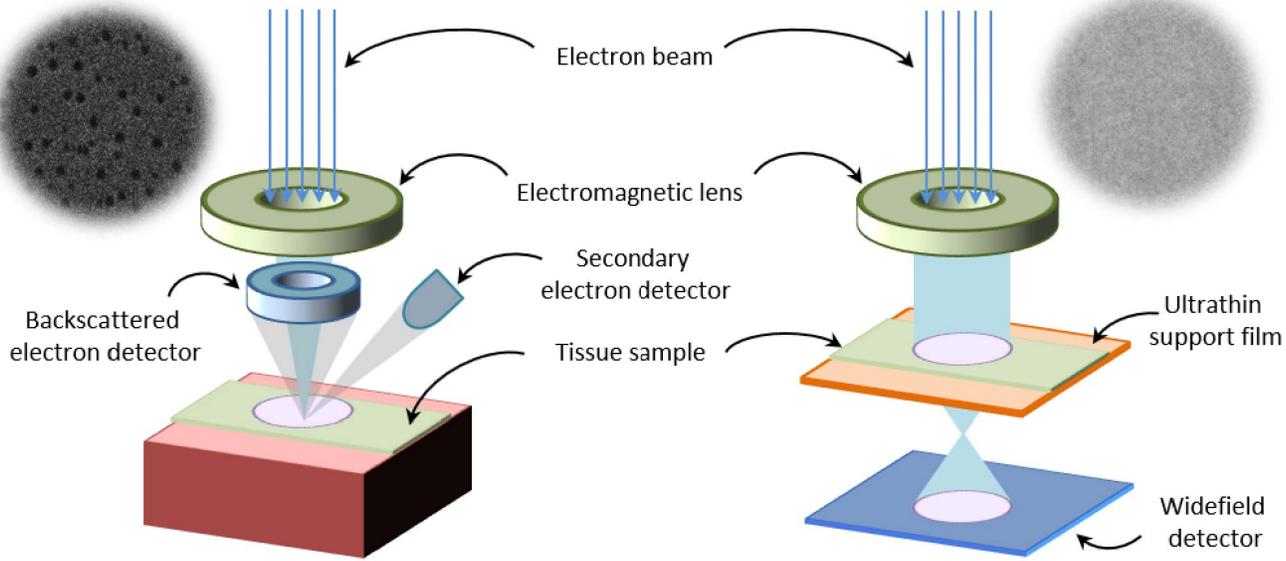


Figure 1: Schematic of SEM (left) and TEM (right) image acquisition workflows accompanied by examples of associated artifacts (noise pattern).

ference between the output of the network and the noise-free image, the proposed method utilizes noise examples from selected blank regions of EM images (that lack biological content) as a prior and constructs a network that can learn to consume or extract that noise pattern. The network is trained without paired ground truth data. This eliminates any need for collecting noise-free and noisy image pairs for ground truth, which is often not practical in automatic imaging workflows, while providing a more general learning-based denoising model that can also handle various unconventional EM imaging artifacts. The results show that the proposed method is more effective at removing EM artifacts than current state-of-the-art denoising methods, including BM3D [5], dictionary-based [3], Noise2Noise [25], and Noise2Void [22]. The proposed method is therefore complementary to existing denoising methods, serving as a flexible solution for end users seeking to reduce effects of noise on their data.

2. Related work

2.1. Spatial filter-based denoising methods

Conventional techniques do not usually harness noise-free samples as ground truth for performing image denoising tasks. Instead, most rely on a common filter design, such as Bilateral filter [35], Non-Local Mean filter [2], BM3D [5], etc. On the other hand, noise reduction problems can also be defined using an optimization scheme, such as that proposed in Anisotropic Diffusion (Perona-

Malik Diffusion) [28], in which Total Variation (TV) [4, 32] either in l_1 or l_2 norms of the results are used as regularizers. Although such methods impose smaller computational burdens by avoiding intensive data and heavy training, they also introduce difficulties such as the need to choose parameters via tuning. Additionally, these methods usually work best for random noise with known distributions.

2.2. Learning-based denoising methods

Modeling images as sparse linear combinations of atoms, as with patch-based dictionary learning K-SVD [6], filter-based dictionary learning, or convolutional sparse coding (CSC) [3], is a denoising approach based on unsupervised learning. Here, a group of atoms, i.e. dictionary, is trained offline (with noise-free data) or blindly online, then used to estimate the noisy image in which the noise model cannot be captured via a sparsity regularizer and results in the noise-free approximation. These methods pose a high computational cost in both training and in estimating the denoised images in which many iterative minimization steps are required for solving pre-defined energy functions.

Deep learning [8, 23] approaches emerged recently with advances in various image processing and computer vision tasks such as classification [14, 18, 21], segmentation [27, 31], localization [7, 13], translation [19, 38], reconstruction [11, 29], denoising [25, 33, 36], and many more. A recent comprehensive overview of EM image restoration techniques can be found in [30]. Among these, randomly augmenting noise-free images with artifacts drawn from

a known distribution (Noise2Noise [25]) makes it possible to cancel out the noise energy without targeting to the noise-free data as long as the input and output are altered by different samples of the same noise type, even though the convergence of training such a deep model cannot be achieved. Alternatively, single blind-spot identity reconstruction (Noise2Void [22]) effectively performs an approximation of the patch image center pixel.

2.3. Difference from the proposed and related work

The key difference that sets the proposed method apart from related deep learning methods that do not require ground truth data is that others assume knowledge of a mathematical model that describes the noise patterns (c.f. prior distribution in Monte-Carlo based Steins Unbiased Risk Estimator (MC-SURE) [33] and Noise2Noise [25] and complementary distribution in learning the approximated identity of a blind-spot pixel in Noise2Void [22]). Instead, we exploit the attributes of blank areas in real EM images, which are similarly corrupted by noise and where signal predominantly arises from noise sources, and leverage them as conditions for performing denoising. To the best of our knowledge, this work presents the first denoiser that can reconstruct noise-free images while simultaneously recovering the input noise.

3. Method

3.1. Data preparation

Data for two example cases were prepared for our experiments: one is applied to Film noise and is intra-EM type (i.e. TEM for both noise-free and noise-corrupted images), while the other is applied to Charge noise and is inter-EM type (i.e. TEM for noise-free image and SEM for noise-corrupted image).

Selection of intra-type Film noise samples: Thousands of thin sections of chemically fixed and stained mouse cortex tissue were collected onto two different support film substrates compatible with TEM: pioloform or LUXfilm®. The TEM_{DR5} dataset was acquired from mouse visual cortex tissue sections collected onto pioloform support film. The TEM_{PPC} dataset was acquired from mouse posterior parietal cortex tissue sections collected onto LUXfilm® support film (Luxel Corporation). Pioloform support films are thinner and more fragile, but contribute little noise to the images. LUXfilm® support films are more robust and better suited for automatic TEM workflows [9, 17, 10], but can add substantial noise to the images. TEM images were acquired at $4.3 \times 4.3 \times \sim 40 \text{ nm}^3 \text{vx}^{-1}$ resolution using a modified JEOL 1200CX system. To model the Film noise, we also captured images of LUXfilm® support films lacking

tissue sections. By superimposing Film noise onto noise-free images (from TEM_{DR5}) via pixel-wise multiplication, we created synthetic noise-corrupted images for validating the trained network. These synthetic noise-corrupted images were not used for training the proposed model.

Selection of inter-type Charge noise samples: Images were acquired from 5–7 days post-fertilization larval zebrafish brain tissue with both TEM and SEM methods. The TEM_{ZB} images were captured at a resolution of $4.0 \times 4.0 \times \sim 40 \text{ nm}^3 \text{vx}^{-1}$ using a modified JEOL 1200CX system. The SEM_{ZB} images were captured at a resolution of $4.0 \times 4.0 \times \sim 60 \text{ nm}^3 \text{vx}^{-1}$ using a FEI Magellan XHR 400L system [16].

3.2. Overview of the proposed method

Fig. 2 presents an overview of the architecture for the proposed denoising method. The input is a collection of three types of images: noise-free EM images (such as TEM_{ZB} or TEM_{DR5}), pure noise patterns from blank regions cut out of noise-corrupted EM images (Film or Charge noise patterns), and noise-corrupted EM images (such as SEM_{ZB} or TEM_{PPC}). Note that these three input images are acquired independently and we do not use paired training data.

The generator \mathbf{G} in the proposed model consists of two end-to-end convolutional-autoencoder networks: \mathbf{G}_1 synthesizes a *fake noise-corrupted* image I_f from a 2-channel image, i.e. a merge between the noise-free image C_t and the noise pattern N_t patches, which were sampled randomly from larger tiles in the corresponding noise-free image and noise pattern training datasets; \mathbf{G}_2 decomposes the input image I_t , sampled randomly from real noise-corrupted EM training images, into a 2-channel image that is a concatenation of the *fake noise-free* image C_f and the *fake noise pattern* N_f . Mathematically, the above descriptions can be informally defined as follows, in which \mathbf{G}_1 and \mathbf{G}_2 are considered being universal approximate functions:

$$\begin{aligned} I_f &= \mathbf{G}_1(C_t, N_t) \\ C_f, N_f &= \mathbf{G}_2(I_t) \end{aligned} \quad (1)$$

$$\begin{aligned} C_r, N_r &= \mathbf{G}_2(I_f) \\ I_r &= \mathbf{G}_1(C_f, N_f) \end{aligned} \quad (2)$$

There are two different paths for training. The 'upward' direction encourages the reconstruction of input noise-free image and noise pattern after going through \mathbf{G}_1 and \mathbf{G}_2 . The 'downward' path acts similarly but in reverse order, as it attempts to reassemble the input noise-corrupted image by subsequently proceeding over \mathbf{G}_2 and \mathbf{G}_1 . These serve as strong constraints for three independently generated input images.

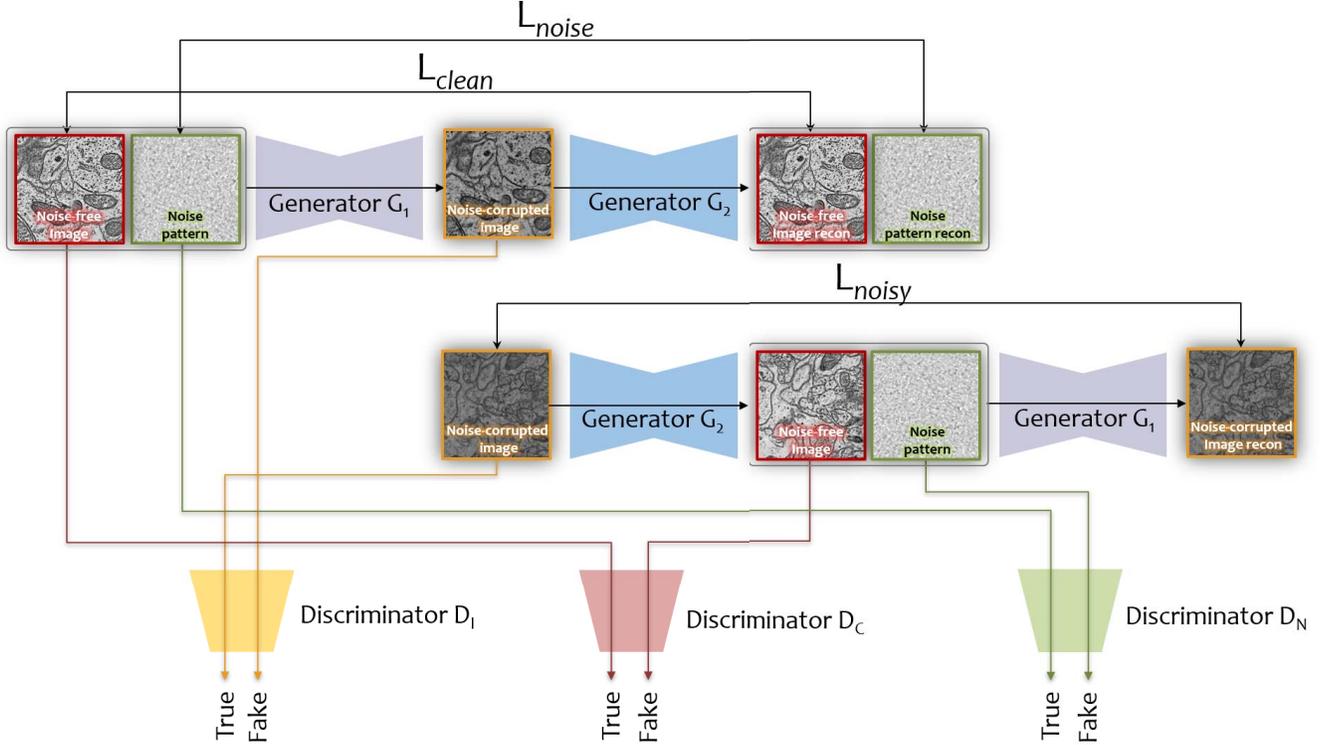


Figure 2: The architecture of the proposed model.

The discriminators \mathbf{D} include \mathbf{D}_I , \mathbf{D}_C , \mathbf{D}_N for recognizing noise-free images C , noise patterns N , and noise-corrupted images I , respectively, and attempt to differentiate between the real instances from the acquired data and the fake results generated by \mathbf{G} . It is worth noting that although \mathbf{G}_1 and \mathbf{G}_2 share the architecture of a convolutional-autoencoder and the same binary classifier structure is used to construct \mathbf{D}_C , \mathbf{D}_N , \mathbf{D}_I , each has its own variable scope and will therefore be updated differently during the training. The inner configurations, such as number of filters, residual bottlenecks, etc., were set identical to those of CycleGAN [38]. The entire system involves training \mathbf{G} and \mathbf{D} adversarially until a balance is reached at the convergence stage.

3.3. Loss definition

We sought to train \mathbf{D} so as to maximize the probability of assigning the correct *true* or *false* label to noise-free, noise pattern, and noise-corrupted images. The objective function for \mathbf{D} can be interpreted as maximizing the log-likelihood for estimating the conditional probability, where the image comes from: $\mathbf{D}_I(I_t) = \mathbf{D}_I(\mathbf{G}_1(C_t, N_t)) = 0$ (*fake*) and $\mathbf{D}_I(I) = 1$ (*real*). Simultaneously, generator \mathbf{G}_1 is trained to minimize $[\log(1 - \mathbf{D}_I(I_t))]$, or

$[\log(1 - \mathbf{D}_I(\mathbf{G}_1(C_t, N_t)))]$. This can be addressed by formally defining an adversarial loss L_{adv} and solving its min-max problem:

$$L_{adv} = [\log(1 - \mathbf{D}_C(C_f))] + [\log \mathbf{D}_C(C_t)] \\ + [\log(1 - \mathbf{D}_N(N_f))] + [\log \mathbf{D}_N(N_t)] \quad (3) \\ + [\log(1 - \mathbf{D}_I(I_f))] + [\log \mathbf{D}_I(I_t)]$$

In extreme cases with large enough resources and data, C_f features could be omitted from the network weights instead of generating from the input. To prevent this, we introduced an additional data consistency loss constraint, L_{cyc} , which is a combination of each input image L_{clean} , L_{noise} , L_{noisy} in a cyclic fashion, such that the noise-free image can be added to or multiplied by the noise pattern to yield the noise-corrupted image and, inversely, the noise-corrupted image can be separated into a noise-free image and noise pattern. In practice, distance metrics such as mean square error (MSE), mean absolute error (MAE), and others can be used to implement L_{cyc} . In our experiments, we employed MAE (ℓ_1 distance). Note that the cyclic loss influences only the generator \mathbf{G} , not the discriminator \mathbf{D} .

$$\begin{aligned}
L_{noise-free} &= \mathbf{d}(C_t, C_r) \\
L_{noise} &= \mathbf{d}(N_t, N_r) \\
L_{image} &= \mathbf{d}(I_t, I_r)
\end{aligned}
\tag{4}$$

$$L_{cyc} = L_{noise-free} + L_{noise} + L_{image} \tag{5}$$

In summary, the proposed system involves two sub-networks which are trained adversarially to minimize the following loss:

$$L_{total} = L_{adv} + \lambda L_{cyc} \tag{6}$$

where λ was used as a weighting factor between loss constraints and set to $\lambda = 10$ for all experiments described here. The entire framework was implemented using the TensorPack¹ system-oriented programming wrapper for the TensorFlow² library.

3.4. Training and Testing specifications

Training phase: The proposed model accepts a specific Field of View (FoV) size of 512×512 , but the acquired EM images are larger. Therefore, we randomly sampled an image patch with size 512×512 for each kind of dataset (noise-free, noise pattern, and noise-corrupted) to train the proposed GAN method. This strategy helps avoid overfitting issues because the training instances are renewed for each iteration. For training, we used the Adam optimizer [20] with an initial learning rate of $1e^{-4}$ that decreased monotonically over 500 epochs.

Testing phase: To obtain denoised result images, we deployed the trained model on subdivided noise-corrupted test images in the form of overlapping 512×512 patches with a step stride of 256. The prediction of each patch was then multiplied with a Gaussian weight and the final result obtained by dividing the weighted estimate by the total per-pixel weight. This resulted in Gaussian blending between patches, effectively avoiding windowing artifacts with a naive subdivision approach.

4. Results

4.1. Experiment setup

Table 1 summarizes the experiments we conducted to assess the performance of the proposed method. There are three types of noise that we considered: Gaussian noise, Charge noise, and Film noise. Even though Gaussian noise is not the main target of the proposed denoising method, it serves as a standard with which to compare other existing denoising methods specifically designed for that noise model.

¹<https://github.com/tensorpack/>

²<http://www.tensorflow.org/>

Table 1: Specifications for our experiment cases.

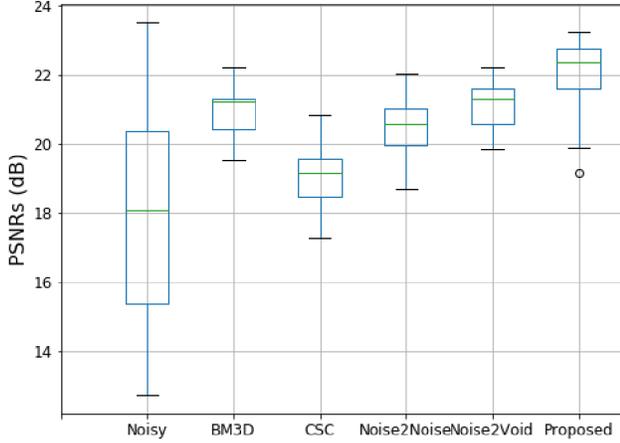
	Noise-Free Images	Noise Type	Noisy Images (Types)
1	TEM _{ZB}	Gaussian	TEM _{ZB} + Gaussian (Synthetic)
2	TEM _{ZB}	Charge	TEM _{ZB} + Charge (Synthetic)
3	N/A	Charge	SEM _{ZB} (Real)
4	TEM _{DR5}	Film	TEM _{DR5} × Film (Synthetic)
5	N/A	Film	TEM _{PPC} (Real)

We conducted both quantitative and qualitative evaluations. For quantitative evaluations (cases 1, 2, and 4), we generated synthetic noise-corrupted images by adding or multiplying either synthetic Gaussian noise ($\mu = 0.0$, $\sigma = 0.05$) or real Charge and Film noise patterns to clean TEM images (TEM_{ZB} and TEM_{DR5}), applied denoising methods (proposed, BM3D [5], CSC [3], Noise2Noise [25], and CycleGAN [38]), and compared the results with the ground truth noise-free images using the Peak-Signal-to-Noise-Ratio (PSNR) (Fig. 3). For cases 3 and 5, ground truth noise-free images were not available, so denoised results were visually compared. Each experiment included 128 test images.

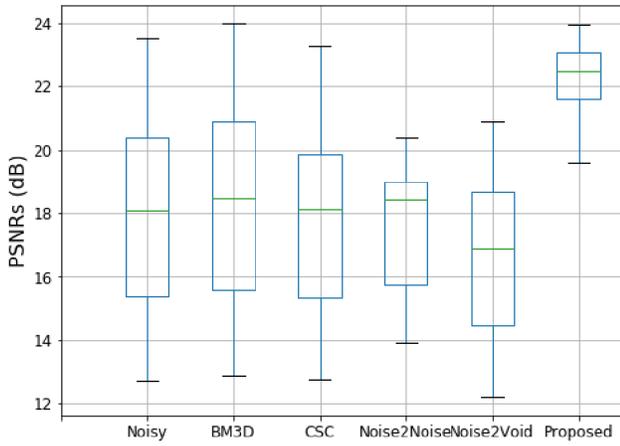
4.2. Quantitative and qualitative evaluations

Case 1 (TEM_{ZB} and Gaussian noise): In this case involving Gaussian additive noise (Fig. 3a), the proposed approach outperformed Noise2Noise [25] and CycleGAN [38] by small margins because it inherited both benefits from a prior distribution (noise model assumed Gaussian) and a hard complementary distribution (input noise reconstruction). CSC [3] was less able to recover the noise-free inputs. Therefore, the resulting reconstructed and the ground truth images were more similar with the proposed method than with other approaches (Fig. 4). Hence, we conclude that training a denoiser with a hard constraint to recover the input noise delivers competitive output image quality in situations with either incoherent noise or additive Gaussian noise.

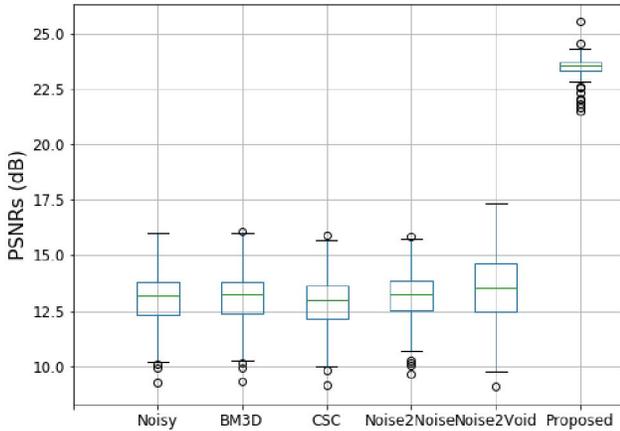
Case 2 (TEM_{ZB} and Charge noise): Switching from Gaussian to Charge noise (case 2, Fig. 3b) maintained an additive artifact-interference strategy similar to that in case 1. Here, however, the proposed method was the only approach capable of consistently achieving PSNRs > 22 dB. Fig. 5 shows that the Charge noise caused by electron beam damage was effectively removed through the presence of cyclic loss in the proposed model. This type of noise has blob-like artifacts on top of a Gaussian noise pattern scattered across acquired images. It is not easily defined by mathematical formulas, making it very difficult to design hand-crafted filters to account for it. Therefore, the pro-



(a) Case 1



(b) Case 2



(c) Case 4

Figure 3: PSNRs (in dB) of related methods compared to the ground truth images.

posed approach’s approximation of the forward and inverse noise models by neural networks conveys certain benefits, such as the inclusion of electron beam charge damage location and size information as conditions that can be incorporated by the learning process.

Case 3 (SEM_{ZB} and Charge noise): The model trained with TEM_{ZB} data and Charge noise in case 2 was blindly deployed on unseen SEM_{ZB} images. Fig. 6 shows a side-by-side comparison of the results for one representative example across all tested methods. By visual inspection alone, we determined that the proposed method was able to effectively remove blob-like Charge noise and increase apparent signal-to-noise in the acquired SEM images because its discriminators were trained using noise-free TEM images. Subtle differences between the proposed method and Noise2Noise are also visible. The proposed method reconstructed an output image with sharper edges and clearer background, which could be due to its separate discriminators for noise-free and noise pattern images.

Case 4 (TEM_{DR5} and Film noise): Multiplicative noise was used to assess reconstruction performance in case 4. As shown in Fig. 3c, the proposed method substantially outperformed other approaches. Fig. 7 reinforces this impression in a visual example, which shows that the proposed method produces results that deviate less from the ground truth than the others. This case differed from previous cases in that, despite still having variable image content and noise behaviors (i.e. TEM_{DR5} and TEM_{PPC} and Film noise), all its input images arose from same modality (TEM). The relative success of the proposed model suggests that the noise-like electron-absorption artifacts included in Film noise are not easily estimated by the noise assumptions underlying the other models. However, cyclically reconstructing the noise-altered and noise-free images apparently maintains the DC components (i.e. low frequency intensities) of the data as long as enough training samples are available.

Case 5 (TEM_{PPC} and Film noise): The model trained with TEM_{DR5} and Film noise in case 4 was blindly deployed on unseen TEM_{PPC} images. Fig. 8 shows the results of the proposed method and others. Here, the TEM_{PPC} image data is noticeably darker due to suppression of image contrast by Film noise. By visual inspection only, our method appeared to improve the contrast closer to that of noise-free TEM images while maintaining high-frequency information better than other tested methods. This production of less blurred results is particularly important for EM applications, in which identification of fine structures such as synaptic vesicles and clefts is a key goal.

5. Conclusion

In this paper, an asymmetrically cyclic adversarial network was proposed for performing denoising tasks with

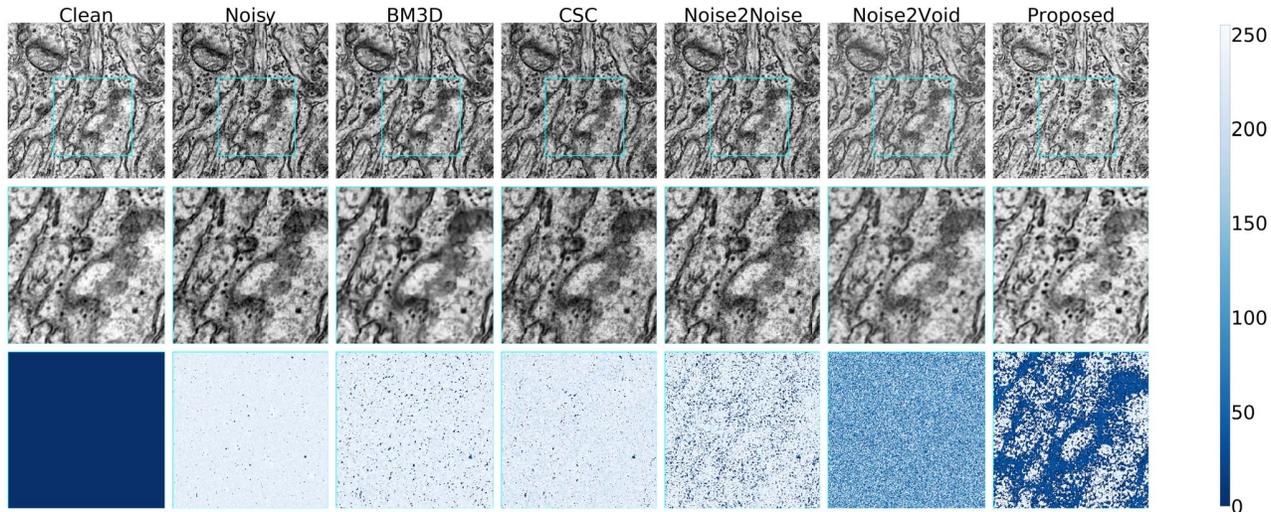


Figure 4: Comparison across methods for a TEM_{ZB} dataset example in case 1. Color bar indicates absolute pixel intensity deviation from ground truth (Clean).

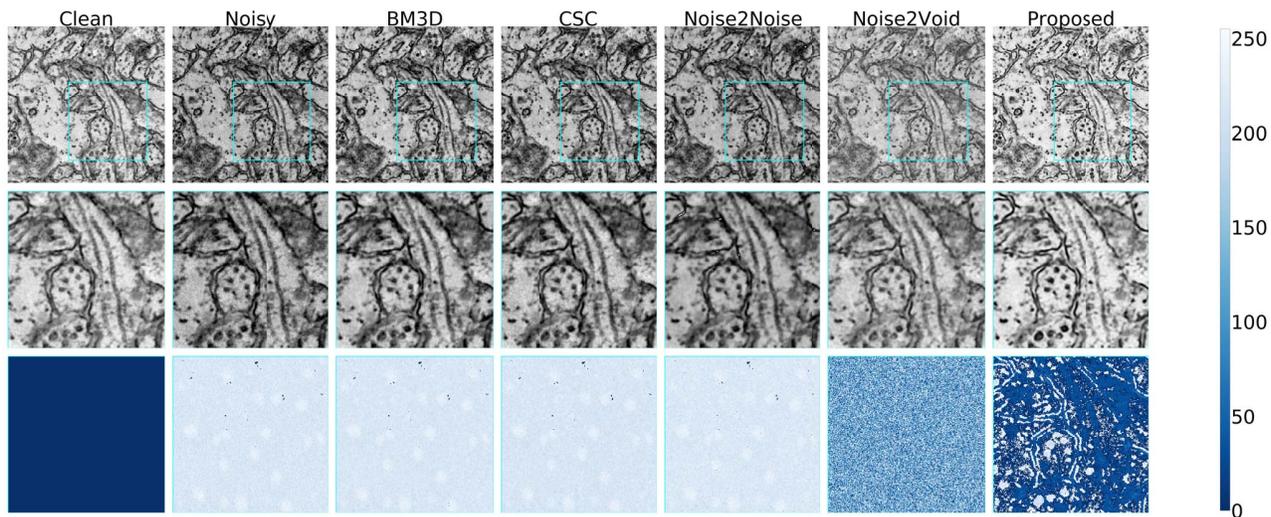


Figure 5: Comparison across methods for a TEM_{ZB} dataset example in case 2. Color bar indicates absolute pixel intensity deviation from ground truth (Clean).

a focus on EM image analysis. Unlike other common deep learning methods that have been trained with paired data in which the target instances are typically noise-free images, our work focused on empirical observations. We showed that the proposed method effectively estimates noise patterns (associated with films or charge damage) from selected regions of acquired image data for use as priors alongside generative adversarial loss, thereby making it possible to create noise-corrupted images, reconstruct original noise patterns, and produce nearly noise-free denoised outputs. We demonstrated that the proposed solution outperforms other state-of-the-art deep learning-based approaches in denoising unseen images that are corrupted by a variety of noise patterns irrespective

of how well the noise can be modeled with with a prior mathematical description. Thanks to the nature of one-time feedforward deployment on trained neural networks, the proposed method has the potential to be applied rapidly to the massive numbers of images generated by high-throughput EM imaging workflows.

Acknowledgement

This work was supported by the National Research Foundation of Korea via the Ministry of Science and ICT (NRF-2017R1D1A1A09000841 and NRF-2019M3E5D2A01063819); by the Leon Levy Foundation; and by NIH grants (R21NS085320 and RF1MH114047).

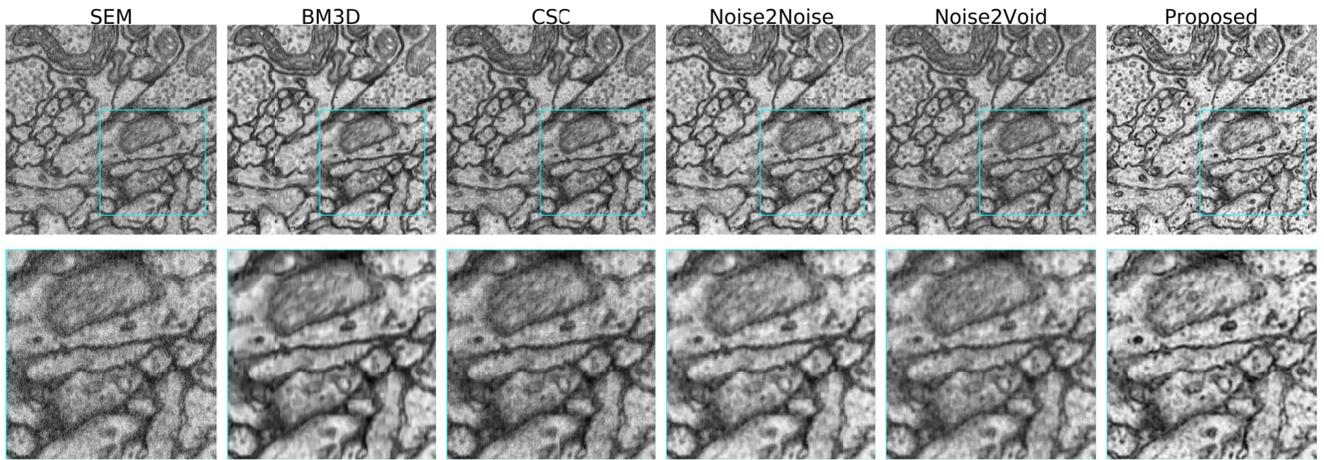


Figure 6: Comparison across methods for a SEM_{ZB} dataset example in case 3.

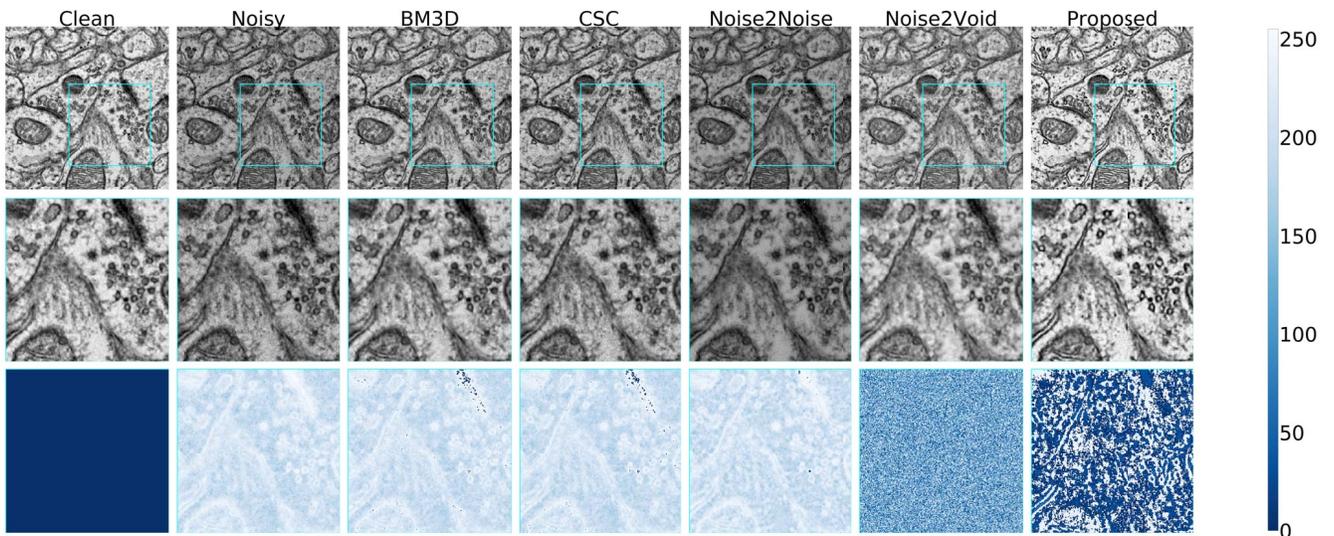


Figure 7: Comparison across methods for a TEM_{DR5} dataset example in case 4. Color bar indicates absolute pixel intensity deviation from ground truth (Clean).

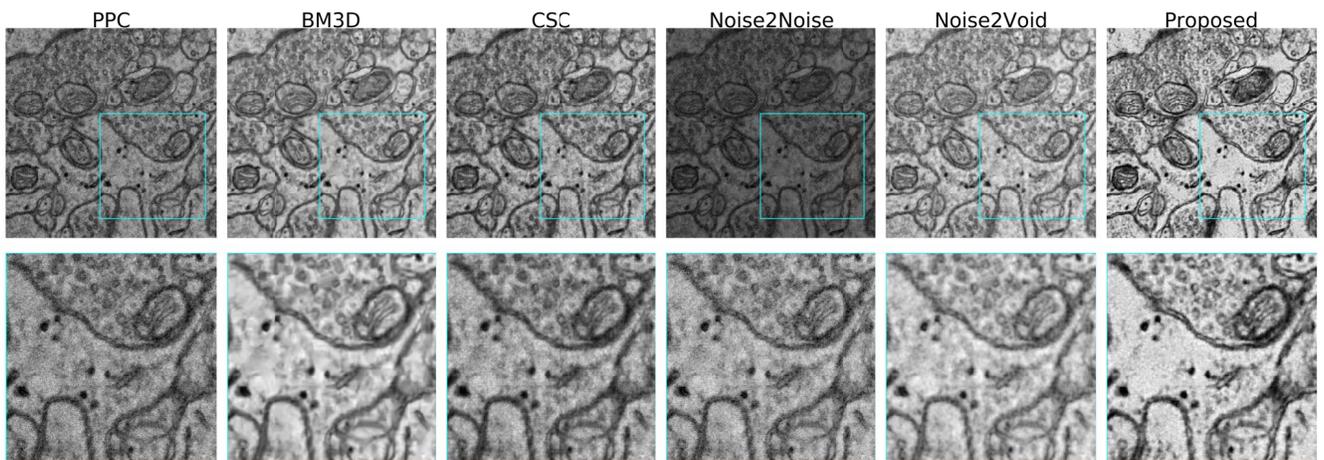


Figure 8: Comparison across methods for a TEM_{PPC} dataset example in case 5.

References

- [1] D. D. Bock, W.-C. A. Lee, A. M. Kerlin, M. L. Andermann, G. Hood, A. W. Wetzel, S. Yurgenson, E. R. Soucy, H. S. Kim, and R. C. Reid. Network anatomy and in vivo physiology of visual cortical neurons. *Nature*, 471(7337):177–182, Mar. 2011. 1
- [2] A. Buades, B. Coll, and J. Morel. A non-local algorithm for image denoising. In *Proceedings of CVPR*, volume 2, pages 60–65 vol. 2, June 2005. 2
- [3] D. Carrera, G. Boracchi, A. Foi, and B. Wohlberg. Sparse Overcomplete Denoising: Aggregation Versus Global Optimization. *IEEE Signal Processing Letters*, 24(10):1468–1472, Oct. 2017. 2, 5
- [4] A. Chambolle. An Algorithm for Total Variation Minimization and Applications. *Journal of Mathematical Imaging and Vision*, 20(1):89–97, Jan. 2004. 2
- [5] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095, Aug. 2007. 2, 5
- [6] M. Elad and M. Aharon. Image Denoising Via Sparse and Redundant Representations Over Learned Dictionaries. *IEEE Transactions on Image Processing*, 15(12):3736–3745, Dec. 2006. 2
- [7] R. Girshick. Fast R-CNN. In *Proceedings of IEEE ICCV*, pages 1440–1448, 2015. 2
- [8] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative Adversarial Nets. In *Proceeding of NIPS*, pages 2672–2680, 2014. 2
- [9] B. J. Graham, D. G. C. Hildebrand, A. T. Kuan, J. T. Maniates-Selvin, L. A. Thomas, B. L. Shanny, and W.-C. A. Lee. High-throughput transmission electron microscopy with automated serial sectioning. *bioRxiv*, 2019. 1, 3
- [10] B. J. Graham, D. G. C. Hildebrand, and W.-C. A. Lee. Gridtape imaging stage. Google Patents, WO2018089578A1, 2018. 1, 3
- [11] Y. Han and J. C. Ye. Framing u-net via deep convolutional framelets: Application to sparse-view ct. *IEEE Transactions on Medical Imaging*, 37(6):1418–1429, June 2018. 2
- [12] K. J. Hayworth, J. L. Morgan, R. Schalek, D. R. Berger, D. G. C. Hildebrand, and J. W. Lichtman. Imaging ATUM ultrathin section libraries with WaferMapper: a multi-scale approach to EM reconstruction of neural circuits. *Frontiers in Neural Circuits*, 8, June 2014. 1
- [13] K. He, G. Gkioxari, P. Dollár, and R. Girshick. Mask R-CNN. In *Proceedings of IEEE ICCV*, pages 2961–2969, 2017. 2
- [14] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of IEEE CVPR*, pages 770–778, 2016. 2
- [15] M. Helmstaedter. Cellular-resolution connectomics: Challenges of dense neural circuit reconstruction. *Nature Methods*, 2013. 1
- [16] D. G. C. Hildebrand, M. Cicconet, R. M. Torres, W. Choi, T. M. Quan, J. Moon, A. W. Wetzel, A. Scott Champion, B. J. Graham, O. Randlett, G. S. Plummer, R. Portugues, I. H. Bianco, S. Saalfeld, A. D. Baden, K. Lillaney, R. Burns, J. T. Vogelstein, A. F. Schier, W.-C. A. Lee, W.-K. Jeong, J. W. Lichtman, and F. Engert. Whole-brain serial-section electron microscopy in larval zebrafish. *Nature*, 545(7654):345–349, May 2017. 3
- [17] D. G. C. Hildebrand, B. J. Graham, and W.-C. A. Lee. Gridtape for fast nanoscale imaging. Google Patents, WO2017184621A1, 2017. 1, 3
- [18] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger. Densely connected convolutional networks. In *Proceedings of IEEE CVPR*, pages 2261–2269, 2017. 2
- [19] P. Isola, J. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. In *Proceeding of CVPR*, pages 5967–5976, July 2017. 2
- [20] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. 2014. 5
- [21] A. Krizhevsky, I. Sutskever, and G. E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks. In *Proceedings of NIPS*, pages 1097–1105, 2012. 2
- [22] A. Krull, T. O. Buchholz, and F. Jug. Noise2void - Learning Denoising from Single Noisy Images. *arXiv preprint arXiv:1811.10980*, 2018. 2, 3
- [23] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, May 2015. 2
- [24] W.-C. A. Lee, V. Bonin, M. Reed, B. J. Graham, G. Hood, K. Glattfelder, and R. C. Reid. Anatomy and function of an excitatory network in the visual cortex. *Nature*, 532(7599):370–374, Apr. 2016. 1
- [25] 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 ICML*, pages 2965–2974. PMLR, 10–15 Jul 2018. 2, 3, 5
- [26] J. W. Lichtman and W. Denk. The Big and the Small: Challenges of Imaging the Brains Circuits. *Science*, 334(6056):618–623, Nov. 2011. 1
- [27] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of IEEE CVPR*, pages 3431–3440, 2015. 2
- [28] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, July 1990. 2
- [29] T. M. Quan, T. Nguyen-Duc, and W. Jeong. Compressed Sensing MRI Reconstruction Using a Generative Adversarial Network With a Cyclic Loss. *IEEE Transactions on Medical Imaging*, 37(6):1488–1497, June 2018. 2
- [30] J. Roels, J. Aelterman, H. Luong, S. Lippens, A. Pizurica, Y. Saeys, and W. Philips. An overview of state-of-the-art image restoration in electron microscopy. *Journal of Microscopy*, 271(3):239–254. 2
- [31] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *Proceeding of MICCAI*, pages 234–241, 2015. 2
- [32] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, Nov. 1992. 2

- [33] S. Soltanayev and S. Y. Chun. Training deep learning based denoisers without ground truth data. In *Proceedings of NIPS*, 2018. 2, 3
- [34] W. F. Tobin, R. I. Wilson, and W.-C. A. Lee. Wiring variations that enable and constrain neural computation in a sensory microcircuit. *eLife*, 6:e24838, May 2017. 1
- [35] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Proceedings of ICCV*, pages 839–846, Jan. 1998. 2
- [36] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155, July 2017. 2
- [37] Z. Zheng, J. S. Lauritzen, E. Perlman, C. G. Robinson, M. Nichols, D. Milkie, O. Torrens, J. Price, C. B. Fisher, N. Sharifi, S. A. Calle-Schuler, L. Kmecova, I. J. Ali, B. Karsh, E. T. Trautman, J. A. Bogovic, P. Hanslovsky, G. S. X. E. Jefferis, M. Kazhdan, K. Khairy, S. Saalfeld, R. D. Fetter, and D. D. Bock. A Complete Electron Microscopy Volume of the Brain of Adult *Drosophila melanogaster*. *Cell*, 174(3):730–743.e22, July 2018. 1
- [38] J. Zhu, T. Park, P. Isola, and A. A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceeding of ICCV*, pages 2242–2251, Oct 2017. 2, 4, 5