Master of Science in Mechanical Engineering

Self−supervised denoising of SEM images for enhanced metrology and inspection

계측 및 검사 향상을 위한 자기지도학습 기반 SEM 이미지 노이즈 제거

February 2023

Graduate School of Mechanical Engineering
Seoul National University

Jiwon Kang

Self−supervised denoising of SEM images for enhanced metrology and inspection

Advisor Do−Nyun Kim

Submitting a master's thesis of
Mechanical Engineering

October 2022

Graduate School of Mechanical Engineering
Seoul National University

Jiwon Kang

Confirming the master's thesis written by
Jiwon Kang

December 2022

Chair	Yoon−Young Kim	(Signature)
Vice Chair	Do−Nyun Kim	(Signature)
Examiner	Byeng−Dong Youn	(Signature)

Abstract

As the semiconductor manufacturing process complexifies and shrinks in size, metrology and inspection is growing to prominence. The main difficulty is that surface properties are biased and contour geometries are deteriorated due to the innate noise in CD-SEM images. In order to eliminate the noise, multiple frames are averaged or traditional denoising methods are applied. However, since these solutions damage the specimen or show unsatisfying results, recent researches have utilized the power of deep learning in SEM image denoising. Despite the fact that deep learning-based methods show superior performance, they still require simulator or abundant ground truth clean images which are costly or even inaccessible in most real-world cases. Lately, few attempts have been made to devise self-supervised methods which does not require clean images however, they still lack in quality. In this research, SEM noise is analyzed in the 'grain size' point of view to demonstrate the failure of previous methods. Moreover, we propose a novel self-supervised image denoising method that shows superior performance in SEM datasets. This method will be further modified and combined with iterative training procedure for enhanced automation.

Keyword : Self-supervised, Denoising, SEM, Metrology, Inspection
Student Number : 2021-27823

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1. Semiconductor Metrology & Inspection

With the rapid increase of the demand of high-resolution integrated chips, semiconductor manufacturing has gained attention more than ever. However, since semiconductor manufacturing consists of hundreds of steps and each requires a nano-scale degree of precision, quality control is the utmost challenge. Metrology and Inspection is the key to maintaining quality throughout the whole process and appropriate utilization of these techniques can ensure a certain yield to be maintained and detect anomalies to avoid wasting time and resources on the subsequent steps.

The main challenge of metrology is to measure the surface properties of lithographic line space patterns which includes critical dimension (CD), line edge roughness (LER) and line width roughness (LWR). CD is the width of the line space patterns and LER/LWR are the variations of CD values and they can be directly exploited as metrology criteria. [1] For instance, it is known that when LER/LWR values go over a certain limit, circuit performance and production yield decrease. [2] On the other hand, the main challenge of inspection is to characterize defects in patterns. To successfully locate and classify defects, with the emergence of design-based metrology (DBM) [3], extracting accurate CD contour geometry is

indispensable. [4] It is also known that contour extraction is crucial in quality control and device characterization. [5]

Metrology and Inspection is mostly performed based on top-down CD-SEM images of the line space pattern to be evaluated. However, CD-SEM images inherit stochastic noise and this noise increases LER/LWR and hinders the accurate extraction of contour geometry. [6] In order to prevent these consequences, traditional methods remove the noise by averaging multiple frames of images for the same pattern. However as averaged frame numbers increase, the damage to the patterns increase due to the EB irradiation. [7]

1.2. Image Denoising

Image Denoising is a process of restoring an unknown clean image by removing noise from an observed noisy image. Image noise is a random variation of pixel values that resides in a high frequency component of an image and differentiating this from other high frequency structural information is the central objective of image denoising. Since this target is an ill-posed problem, various approaches have been devised to obtain a good estimation. Broadly, these approaches can be categorized into three methods: spatial domain filtering, variational denoising and transform domain filtering. [8]

Spatial domain filtering is one of the elementary denoising methods which applies certain neighborhood operations that acts as

a low−pass filter. This filtering can be subdivided into linear and non−linear operation. Linear filtering includes mean filtering and gaussian filtering. [9] On the other hand, non−linear filtering includes median filtering [10] and bilateral filtering. [11] Variational denoising is another denoising method that functions in the spatial domain. This method removes the noise by optimizing certain energy function based on the assumption that prior knowledge of the clean image is known. For instance, total variational (TV) denoising [12] assumes that clean image has low image gradient and non−local means (NLM) denoising [13] utilizes the fact that similar patches exist extensively in an image. Contrastively, transform domain filtering applies certain denoising procedure after transforming an image into another domain. Transformations are mostly Fourier transform, however algorithms like Block matching and 3D filtering (BM3D) [14] that utilize wavelet transform [15] are showing the most superior results.

Although traditional methods have achieved promising results, they suffer from some serious shortcomings. They require overlapped optimization process for the test phase, manual tuning of parameters and excessive amount of computation time. [16] Recently, these issues have been alleviated with the emergence of deep learning in image denoising. Deep learning is a subclass of machine learning that brought a paradigm shift from traditional model−based approach towards data−based approach using the power of neural networks. Many network architectures were proposed since [17] first introduced deep learning in the field of image denoising. DnCNN [18] utilized residual learning and batch normalization, NBNet [19]

adopted a non-local subspace attention module and FFDNet [20] made an additional use of noise level map to outperform traditional methods.

Evaluation of image denoising can be performed both qualitatively and quantitatively. For quantitative evaluation, peak signal to noise ratio (PSNR) and structural similarity index measure (SSIM) are most widely used. [21] PSNR measures the L2 distance between clean and corrupted images and offers a simple and clear meanings, however cannot guarantee a perceptual quality. SSIM has been developed to mitigate this problem by adopting the concept of structure, luminance and contrast.

1.3. Self-supervised Image Denoising

Supervised deep learning methods show extraordinary performance while consuming small amount of computation time however, they come with a serious drawback. They require perfectly clean ground truth images which are very costly or even impossible to achieve in most of the real-world cases. To address this issue, numerous researches have proposed self-supervised image denoising techniques which only stand in need for corrupted images itself. Noise2Noise [22] first came up with this concept and showed that based on L2 loss statistics, pairs of noisy images are enough. Here, these noisy pairs each are different corrupted version of the same unknown clean image.

Gathering noisy image pairs are also highly demanding therefore, methods using only single noisy images are developed. Most widely used idea is the blind-spot concept. [23]-[27] Noise2Void (N2V) [23] utilizes the blind-spot network which has a receptive field masked at the center while Noise2Self (N2S) [24] randomly masks a noisy image to generate blind-spot image. Both methods work based on the assumption that signal component in the masked pixel contain statistical dependencies with nearby signal values whereas noise component in the masked pixel is conditionally independent from the nearby pixels. Another group of methods split a noisy image into two sub-images in order to provide fake supervision to the model. Neighbor2Neighbor (Ne2Ne) [28] introduces a random neighbor sub-sampling while Noise2Fast (N2F) further simplifies the sub-sampling to propose checkerboard down-sampling. These methods assume the conditional independence between two sub-sampled images which is equal to the assumptions of blind-spot methods. Some other methods achieve the fake supervision via using the true noise distribution. Noisier2Noise (Nr2N) [29] generates a fake noise from the true noise distribution and adds them to noisy images to create noisier images. Then the noisy & noisier image pairs work as a training set for a model to be trained in a supervised manner. Similarly, Recorrupted2Recorrupted (R2R) [30] constructs a pair of noisy images from a single noisy image via matrix transformation. While Nr2N and R2R require the noise distribution to be known, GAN2GAN [31] directly learns the distribution with the help of wGAN. [32]

1.4. SEM Image Denoising

For precise metrology and inspection, proper image denoising should be performed in prior. Traditionally, noisy filters such as gaussian filter [33], [34] or median filter have been applied. Also, some developed an image simulation method [35], [36] to obtain clean images and others used model-based approach [6], [37]-[40] to directly estimate unbiased LER, LWR values. Recently, [7], [41] utilized power spectral density (PSD) to remove noise components in the high frequency region. Although these methods showed valid results, it is shown that noise filters have big impact on LER, LWR values [7] and PSD fails in some images. [42]

In order to alleviate these issues, deep learning has also been adopted in SEM image denoising. [43] introduced SEMNet & EdgeNet and [44] utilized conditional GAN to estimate unbiased LER, LWR accurately. However, these methods depend upon a simulator to generate fake samples. Since such simulators are costly and inaccessible most of the times, various methods are exploiting the virtue of self-supervised learning. Noise2Void has been adopted in few researches [5], [45] however still doesn't show sufficient image quality. [46] combined Noise2Noise with a simulation to propose uMLIQUE framework, however it still need a supervision from a simulator. [47] designed Denoise2Next, Denoise2Best training, yet require images of multiple frames which can damage the SEM patterns.

1.5. Noise Modeling and Generation

For image denoising purposes, extensive researches have been conducted on modeling noise, estimating the noise parameters and generating fake noise samples that mimic the real noise distribution. Conventionally, these processes have been conducted by using wavelet transform [48], [49], contourlet transform [50], MLE [51], PCA [52], [53] and variational bayes [54]. Recently, deep learning-based methods are surpassing these methods and showing remarkable progress in performance and efficiency. These advances include, MLP [55], GAN [56], conditional GAN [57], Dual Adversarial Network [58], Camera Encoding Network [59], Contrastive learning and so on. Despite their performance, the need of clean ground truth images hinders them from real-world applications. FBI-Denoiser [60] overcame this difficulty by combining Poisson-gaussian estimation network (PGE-Net) and generalized Anscombe transformation (GAT). PGE-Net first estimates the parameters of Poisson-gaussian noise and then using these estimated parameters, GAT transforms the noisy image into a unit gaussian noisy image. Since unit gaussian noisy images can be easily denoised with pretrained models, denoising can be performed without any clean images. In spite of the merit, FBI-Denoiser still suffers from high computation of GAT.

Abundant studies have been conducted on noise modeling and generation however, there have been limited research on SEM noise

modeling and generation. SEM noise is mostly approximated as Poisson-gaussian noise [61] while some research [62] modeled SEM noise as a cascade of five steps.

Chapter 2. Motivation

2.1. Self-supervised Image Denoising revisited

Clean images are expensive and sometimes inaccessible in many real-world situations. To overcome this complication, various self-supervised methods have been proposed. Here, these deep learning based self-supervised methods will be briefly compared with traditional ones. In details, supervised method (N2C), N2S, Ne2Ne, and Nr2N will represent deep learning methods while NLM, TV and BM3D will cover for traditional ones. They will be evaluated on both synthetic and real-world datasets to show the motivation of this research.

2.2. Public Dataset

For comparison, well known denoising benchmarks, BSD100, Kodak and Set12 will be used. Each of the dataset will be corrupted with gaussian or poisson noise with two level of intensities. These corrupted versions will be denoised by various methods and be compared with clean images to calculate PSNR and SSIM values.

Table 1, 2 and Figure 1 show the result of the experiment. In all of the cases, the baseline, N2C shows the best performance. Among the gaussian noise cases, self-supervised methods outperform traditional ones and N2S even shows comparable result

to the baseline. However, in the poisson noise cases, N2S and Ne2Ne show very poor performance while Nr2N is even inapplicable since they require clean images to define the noise model. This phenomenon can be explained with the assumptions made by N2S, Ne2Ne, and Nr2N. N2S and Ne2Ne works only for noise that are independent from the clean image structures while Nr2N expects the noise to be additive. Since these conditions only apply to gaussian noise, self-supervised methods fail in poisson noise.

2.3. SEM Dataset

Here, SEM image denoising will be evaluated on three different datasets, SEM1, SEM2 and SEM3. Each is comprised of 1-frame/16-frame image pairs, where 16-frame images are acquired by averaging 16 1-frame images for the same pattern. 16-frame images will work as ground truth clean images to calculate PSNR values. SSIM values will not be compared since the structure of the image degrades while averaging which can lead to misleading numbers.

As it can be seen in Table 3 and Figure 2-4, the results are similar to the poisson noise case in public datasets. N2S and Ne2Ne cannot remove the noise successfully while Nr2N cannot be applied since the noise model is unknown.

2.4. SEM Noise Analysis

Let's assume a simple line−space model as in Figure 5. The image can be divided into three parts and each part roughly is a flat space with a uniform pixel value corrupted with SEM noise which can be simply modeled as poisson−gaussian. However, with the virtue of Equation 1, the poisson noise component for each part can be approximated as gaussian with different standard deviation. Since the combination of two gaussian distributions becomes a single gaussian, SEM noise can be simplified as gaussian with different standard deviations in different flat regions.

$$Poisson(\lambda) \quad \sim \quad Normal(\mu = \lambda, \sigma^2 = \lambda) \qquad (1)$$

If the SEM noise can be seen as gaussian, why do N2S or Ne2Ne fail contrary to the cases in public dataset? This can be elucidated with the 'Grain size' concept. Grain size is the average diameter of particles and it is fixed to 1 in synthetic noise like gaussian or poisson noise whereas bigger than 1 in real SEM noise as it can be seen in Figure 6. When the grain size of the noise is bigger than 1, the principles of N2S or Ne2Ne fail to work.

As briefly mentioned in Chapter 1, N2S randomly masks few pixels and trains the model to estimate the hidden pixel values by looking at the pixels nearby. The trained model can successfully estimate the clean target by utilizing the fact that image structure

information is shared between neighboring region while noise components are fully independent pixel by pixel. However, as the grain size grows bigger than 1, noise components between adjacent pixels tend to share higher dependencies. Similarly, Ne2Ne partitions an image into two to generate fake supervision. This method can also be seen as masking in N2S and can be explained in the same way.

To summarize, SEM noise can be seen as coarse gaussian noise with multiple standard deviation. The coarse condition is violated by N2S and Ne2Ne however, is unaffected by Nr2N and R2R. The 'additive noise' assumption works no matter how big the grain size is, while the only problem is that the noise model should be established. In chapter 3, a modified Nr2N will be proposed which does not require the understanding of the noise model.
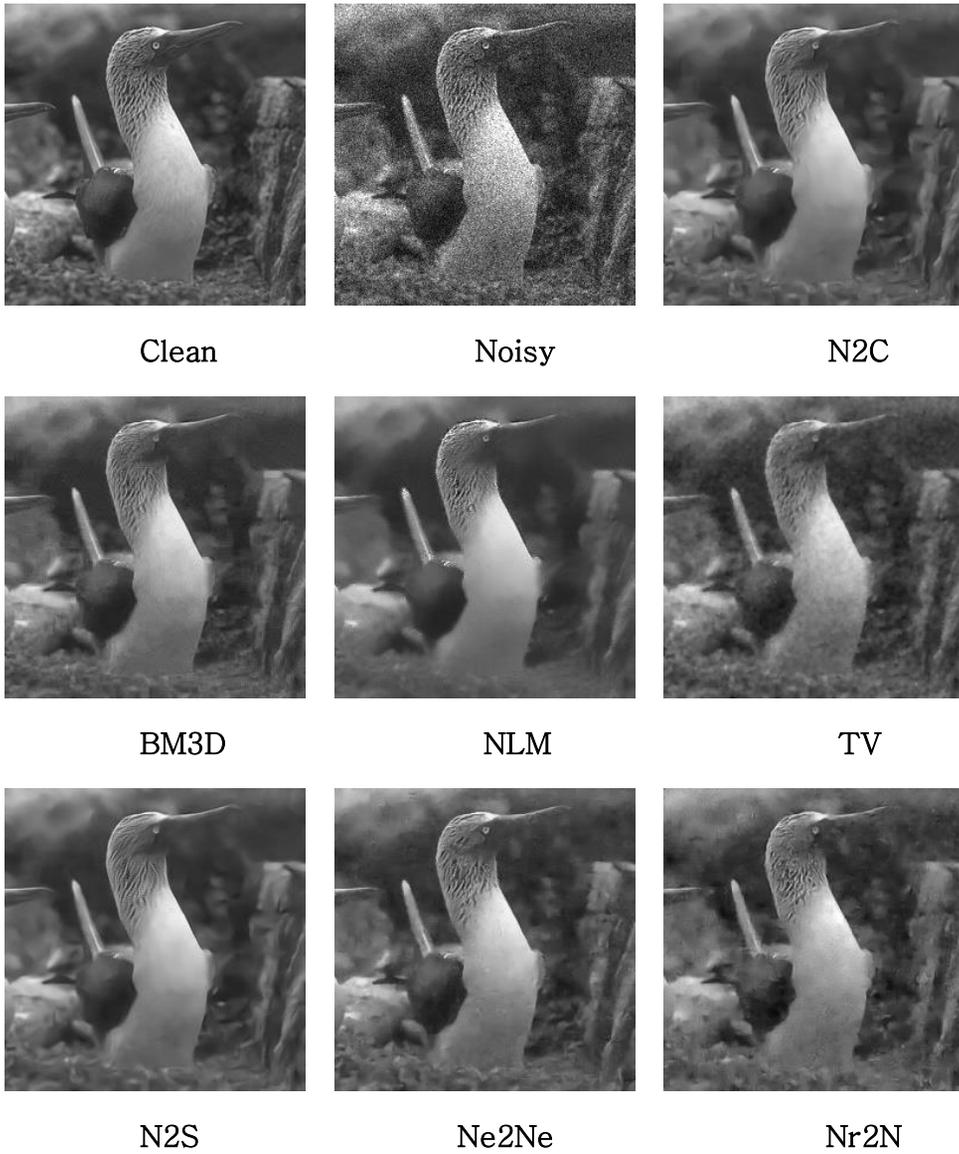
Figure 1. Qualitative evaluation on BSD100 dataset. Clean image is corrupted with gaussian noise ($\sigma=25$).

16−Frame	1−Frame	N2C
BM3D	NLM	TV
N2S	Ne2Ne	Nr2N
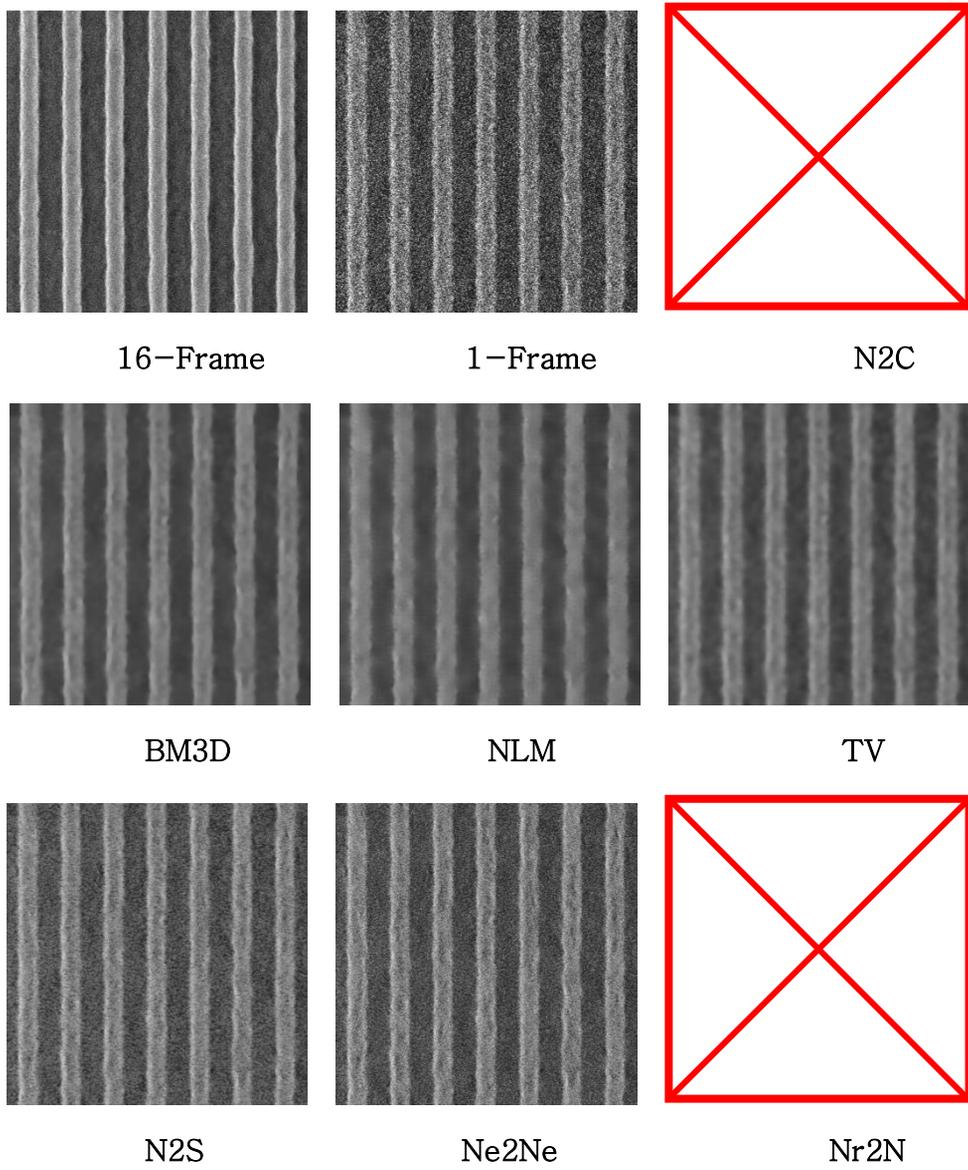
Figure 2. Qualitative evaluation on SEM1 dataset.

16-Frame   1-Frame   N2C

BM3D   NLM   TV

N2S   Ne2Ne   Nr2N
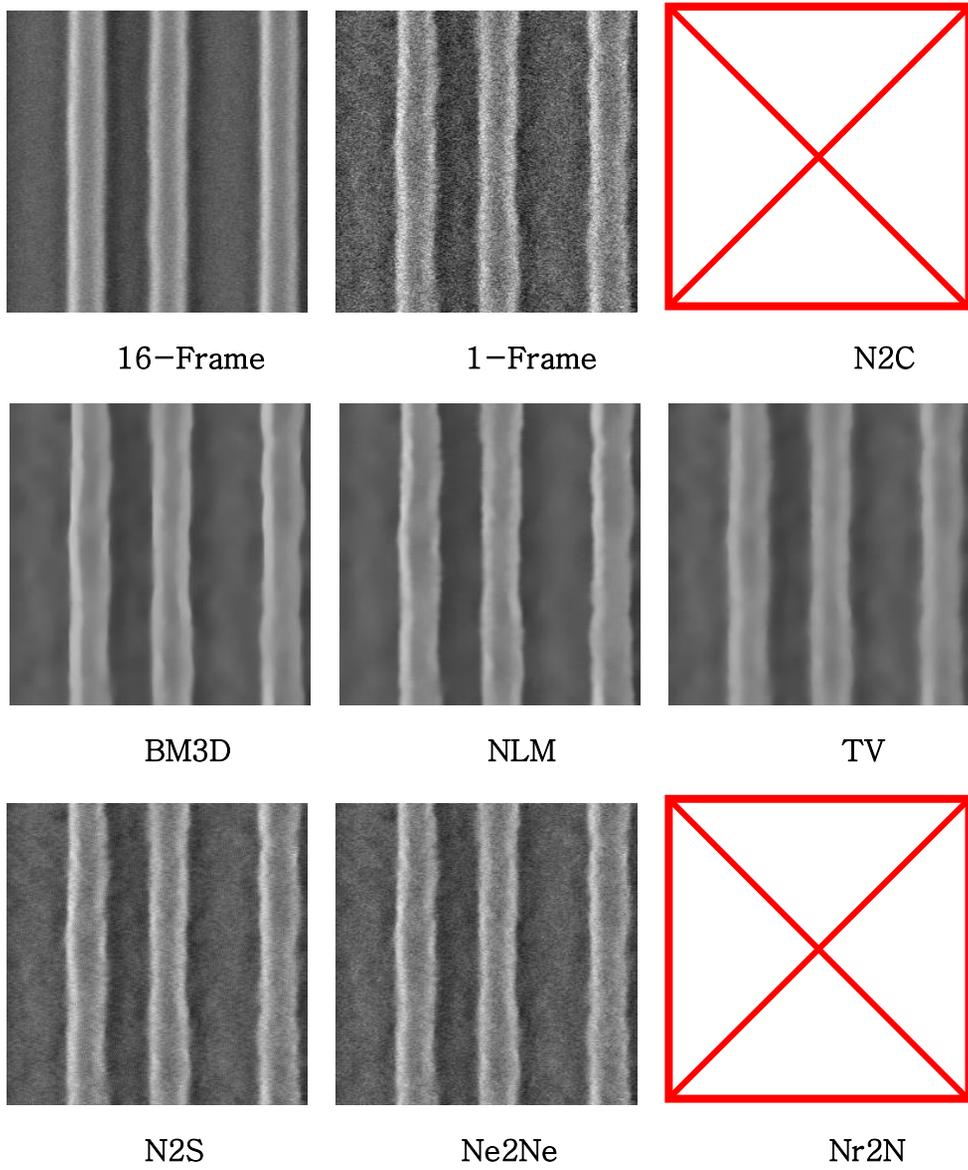
Figure 3. Qualitative evaluation on SEM2 dataset.

16−Frame   1−Frame   N2C

BM3D   NLM   TV

N2S   Ne2Ne   Nr2N
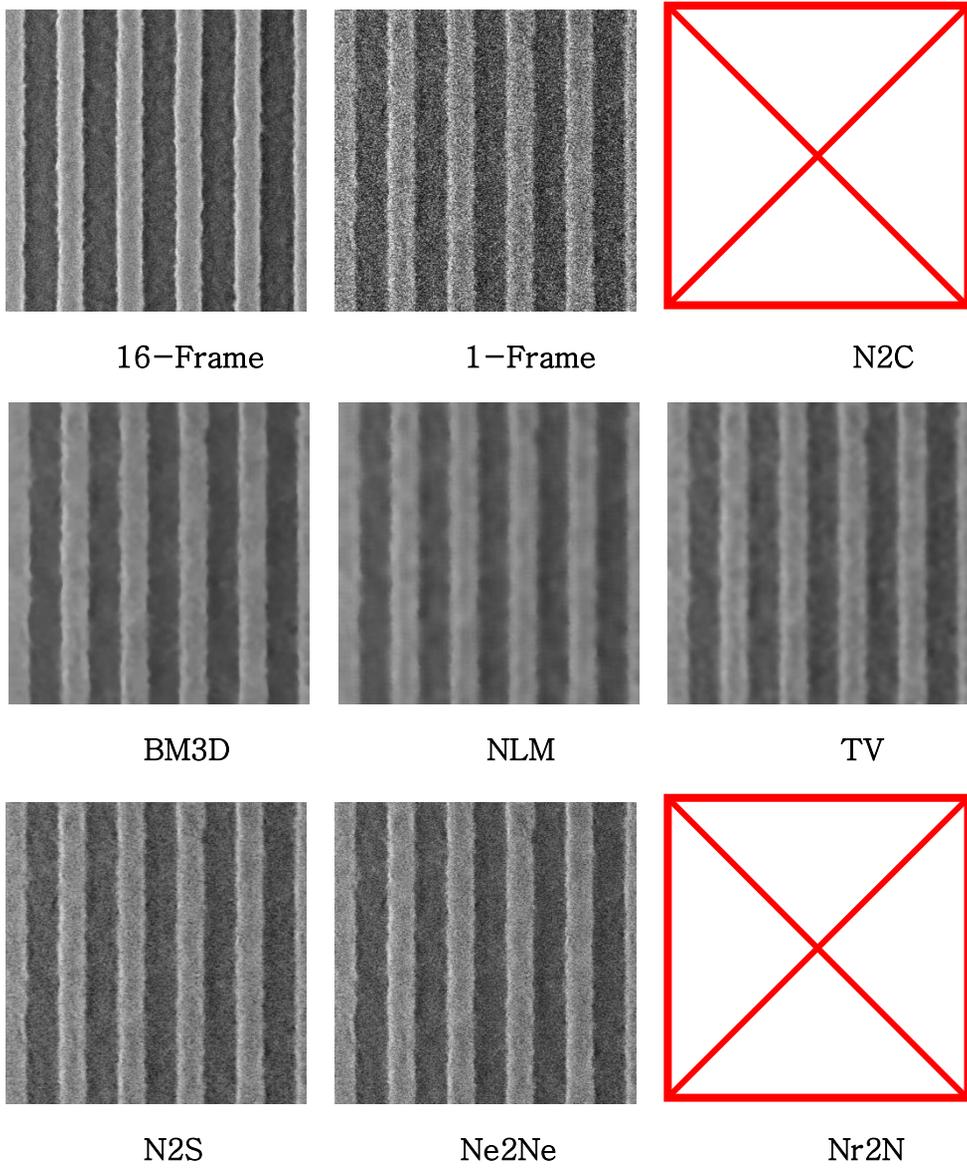
Figure 4. Qualitative evaluation on SEM3 dataset.
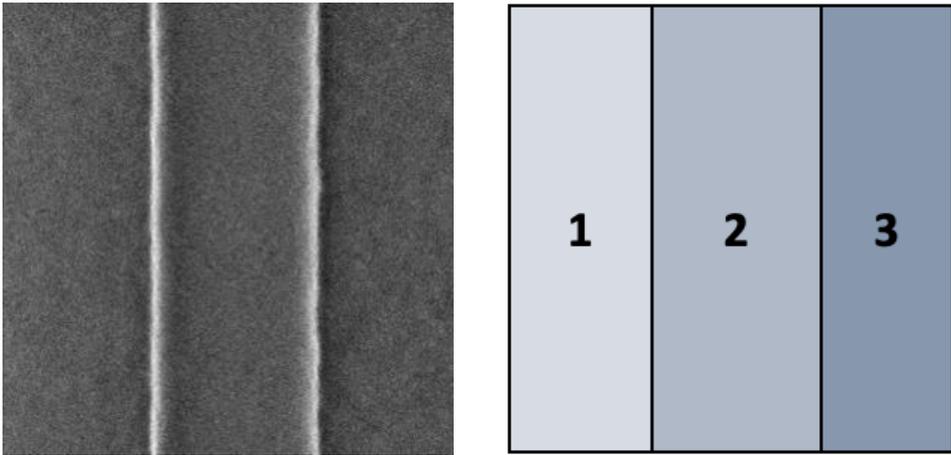
Figure 5. Simple line-space pattern and its diagram
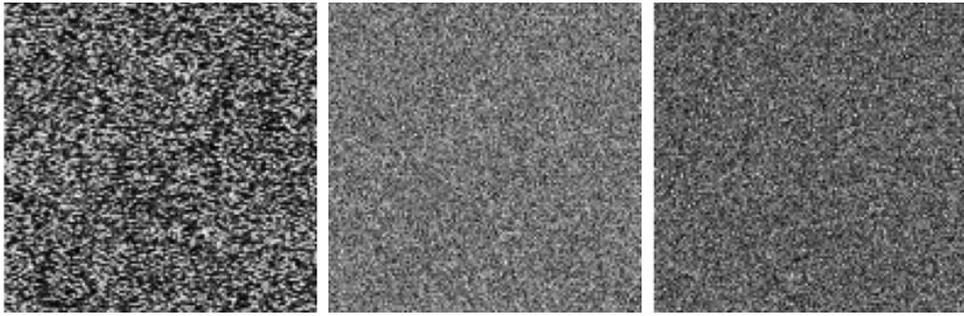
SEM noise.   Gaussian noise   Poisson noise

Figure 6. Various noise with same mean pixel value.

Method	BSD100	Kodak	Set12
Gaussian (std=25)			
N2C [18]	28.03/0.801	29.14/0.804	29.39/0.847
NLM [13]	26.33/0.727	27.46/0.737	27.61/0.792
TV [12]	26.16/0.726	27.20/0.743	27.24/0.790
BM3D [14]	27.48/0.775	<u>28.64</u>/0.776	<u>29.08</u>/0.820
N2S [24]	**27.75/0.790**	**28.94/0.798**	**29.09/0.843**
Ne2Ne [28]	<u>27.63/0.783</u>	28.57/<u>0.779</u>	28.90/<u>0.822</u>
Nr2N [29]	25.90/0.702	26.91/0.707	27.01/0.766
Gaussian (std=50)			
N2C [18]	25.27/0.684	26.42/0.703	26.45/0.765
NLM [13]	23.23/0.566	24.19/0.569	24.01/0.635
TV [12]	23.31/0.591	24.36/0.609	24.06/0.663
BM3D [14]	24.03/0.610	25.13/0.610	25.29/0.678
N2S [24]	**25.09/0.670**	**26.25/0.695**	**26.26/0.758**
Ne2Ne [28]	<u>24.97/0.662</u>	<u>25.96/0.673</u>	<u>25.96/0.732</u>
Nr2N [29]	23.32/0.553	24.24/0.565	23.92/0.619

Table 1. Quantitative comparison (PSNR(dB)/SSIM) for Gaussian noise. The highest values are marked bold while the second highest are underlined. N2C, which is trained with clean images, is excluded.

Method	BSD100	Kodak	Set12
Poisson (peak=25)			
N2C [18]	26.78/0.760	27.93/0.769	28.01/0.813
NLM [13]	24.49/0.645	25.57/0.659	25.56/0.701
TV [12]	25.04/0.690	25.83/0.679	25.80/0.706
BM3D [14]	**25.52/0.698**	**26.69/0.706**	**26.90/0.729**
N2S [24]	18.84/0.474	20.19/0.440	21.39/0.471
Ne2Ne [28]	19.09/0.456	19.62/0.398	19.84/0.418
Nr2N [29]	x	x	x
Poisson (peak=50)			
N2C [18]	28.28/0.817	29.38/0.819	29.49/0.852
NLM [13]	25.87/0.701	27.01/0.717	27.24/0.771
TV [12]	26.09/0.730	27.00/0.732	27.11/**0.777**
BM3D [14]	**27.17/0.767**	**28.35/0.771**	**28.41**/0.774
N2S [24]	22.23/0.594	23.37/0.569	24.38/0.599
Ne2Ne [28]	21.67/0.578	22.19/0.511	22.59/0.525
Nr2N [29]	x	x	x

Table 2. Quantitative comparison (PSNR(dB)/SSIM) for Poisson noise. The highest values are marked bold while the second highest are underlined. N2C, which is trained with clean images, is excluded.

Method	SEM1	SEM2	SEM3
NLM [13]	17.93	<u>27.02</u>	17.41
TV [12]	<u>18.00</u>	26.77	**17.42**
BM3D [14]	**18.00**	**27.03**	<u>17.42</u>
N2S [24]	15.69	24.13	15.48
Ne2Ne [28]	15.89	24.23	15.61
Nr2N [29]	x	x	x

Table 3. Quantitative comparison (PSNR(dB)) for SEM datasets. For each dataset, the highest values are marked bold while the second highest are underlined.

Chapter 3. Proposed Method

3.1. Noisier2Noise revisited

We consider a situation where a deep neural network has to be trained to perform image denoising with only noisy images available. Let χ, A be the distribution of unknown clean images and noise. Then noisy image can be formulated as $Y \equiv X + N$ where $X \sim \chi$ and $N \sim A$. Since the noise distribution A is known, additional synthetic noise $M \sim A$ can be extracted.

In order to imitate supervised training strategy, Nr2N first generates noisier image $Z \equiv Y + M = X + N + M$ for each noisy image. Then the neural network will be trained to predict Y given Z using pixel−wise L_2 loss function. This training is an optimization process for the parameter θ of the network f and can be formulated as follows:

$$\min_{\theta} E_Z[||f(Z; \theta) - Y||_2] \tag{2}$$

Due to the statistical nature of L_2 loss function, the optimal prediction in equation 2 would be the conditional mean of Y given Z, $E[Y|Z]$. This value can be further decomposed as in equation 3 using the fact that $Y \equiv X + N$ and $E[M|Z] = E[N|Z]$. The second term can be

justified since M and N are independent while originating from the same distribution.

$$2E[Y|Z] = 2(E[X|Z] + E[N|Z])$$
$$= E[X|Z] + (E[X|Z] + E[N|Z] + E[M|Z])$$
$$= E[X|Z] + E[X + N + M|Z]$$
$$= E[X|Z] + Z \tag{3}$$

Equation 3, in other words, can be interpreted as $E[X|Z] = 2E[Y|Z] - Z$. This tells us that doubling the output of noisier2noise model and subtracting noisier value would be equivalent to the output of a model trained with clean & noisier image pairs.

3.2. Patch-based Noisier2Noise Framework

Nr2N enables the training without clean images when the noise distribution is given. However, since the noise distribution of SEM images is inaccessible, Nr2N cannot be applied in SEM image denoising.

One distinct characteristic of SEM images is that most of the areas are flat. Being flat in other words, can be interpreted as having uniform image structure component. Thus, when the flat area is subtracted with its mean value, only pure noise component will exist. Therefore, by extracting the noise component in the flat region, noise distribution of SEM image dataset can be approximated. By virtue of this discovery, a modified Nr2N which we call 'Patch-based Nr2N'

will be proposed. As illustrated in Figure 6, our method can be divided into three steps.

For data pre−processing, pure noise patches will be gathered from the dataset. We utilize the smooth patch extraction method in GAN2GAN which is based on the 2D discrete wavelet transform (DWT). This extraction involves a single parameter λ which determines the purity of extracted patches. Bigger λ indicates that extracted patches will contain larger quantity of image structure components. Since the size of extracted patches are smaller than the size of SEM images, extracted patches will be tiled to match the size. Finally, tiled patches will be subtracted by its mean to comprise approximated noise distribution.

In the training stage, a pure noise patch will be randomly chosen from the distribution and added to a noisy image to form a noisier version of its own. Prepared noisier, noisy image pairs will be then used to train the deep neural network in a supervised manner.

For inference, as mentioned in equation 3, output of the network will be doubled and then subtracted with noisier to form prediction. Moreover, the inference trick in R2R will be adopted to boost the performance further. The trick is to generate multiple noisier input for a single noisy image by adding different noise patches. With the multiple inputs, multiple inferences can be made which then can be averaged to further reduce the variance of the leftover noise. As it can be seen in Table 4, this inference trick remarkably improves the denoising performance. Furthermore, when

the trick is combined with the power of batch－wise training, the computation burden from multiple inferences can be alleviated.

3.3. SEM dataset revisited

Our methods were evaluated and compared both quantitatively and qualitatively. In case of SEM3 dataset, out methods were inapplicable since most of the images were densely filled with patterns and sufficient number of noise patches couldn’t be extracted.

As in table 4, simple Nr2N shows comparable performance to BM3D while other self－supervised methods fail to remove the noise. Moreover, Nr2N with the inference trick almost shows same performance to BM3D. However, when it comes to the visual quality as illustrated in Figure 7 and 8, our method shows outstanding results, removing noise almost perfectly while BM3D still consisting of unneglectable level of noise.

The disagreement between quantitative and qualitative evaluation is due to the incorrect establishment of the ground truth for PSNR calculation. For the calculation, 16－frame images are assumed as the ground truth. However, since 16－frame images still carry certain level of noise, the calculated PSNR cannot fully represent the denoising performance. Thus, additional metrics will be adopted in the following chapter.
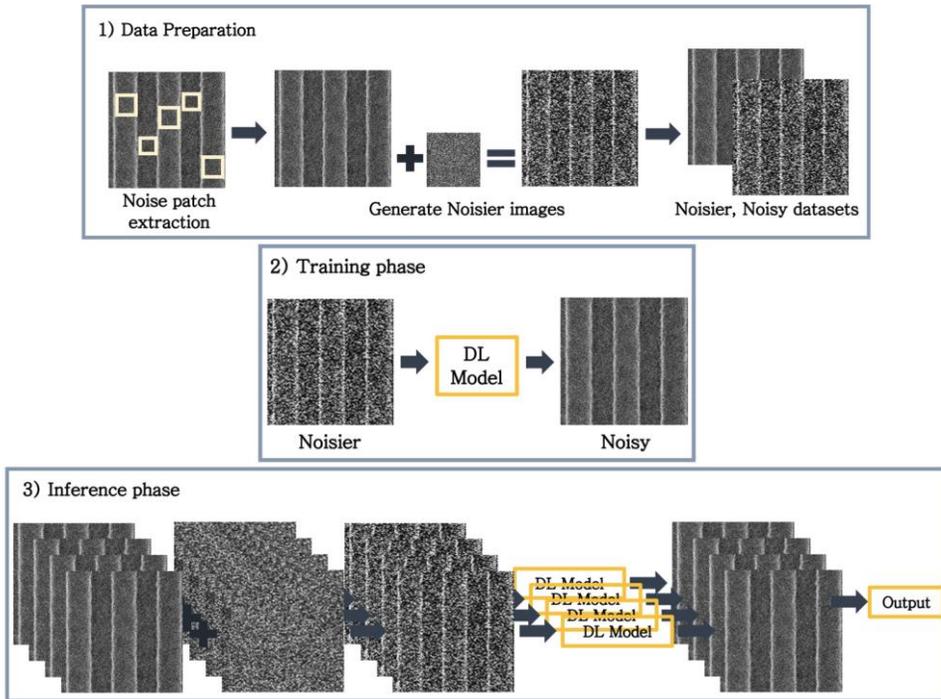
Figure 7. A diagram of Patch-based Noisier2Noise process.

16-Frame	1-Frame	N2S
BM3D	Nr2N	Nr2N + R2R
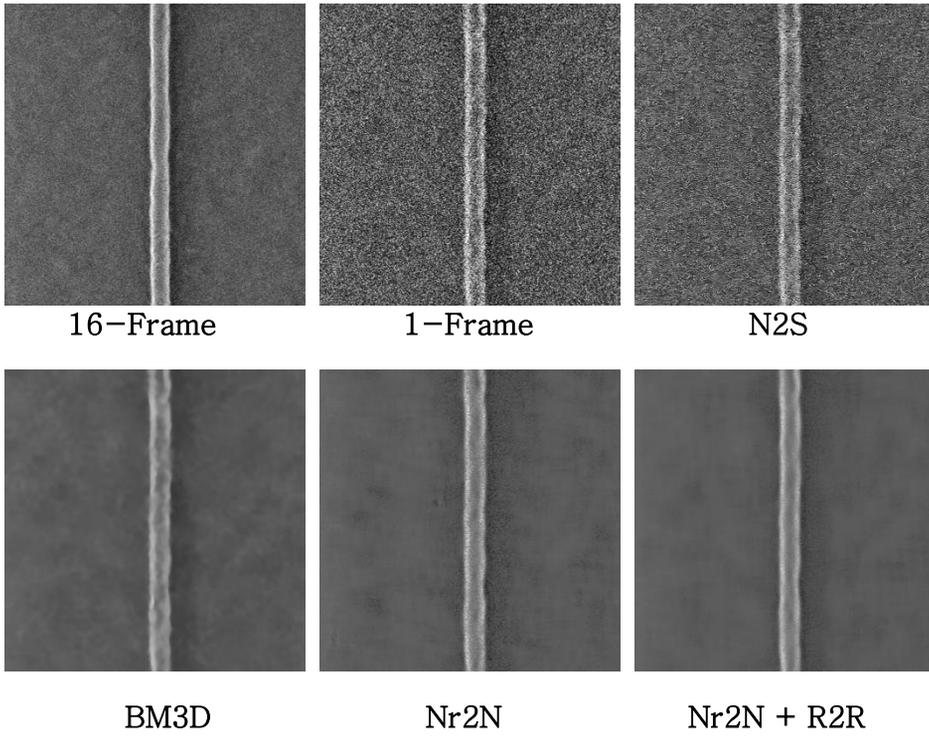
Figure 8. Qualitative comparison on SEM1 dataset

16-Frame   **1-Frame**   **N2S**

BM3D   **Nr2N**   **Nr2N + R2R**
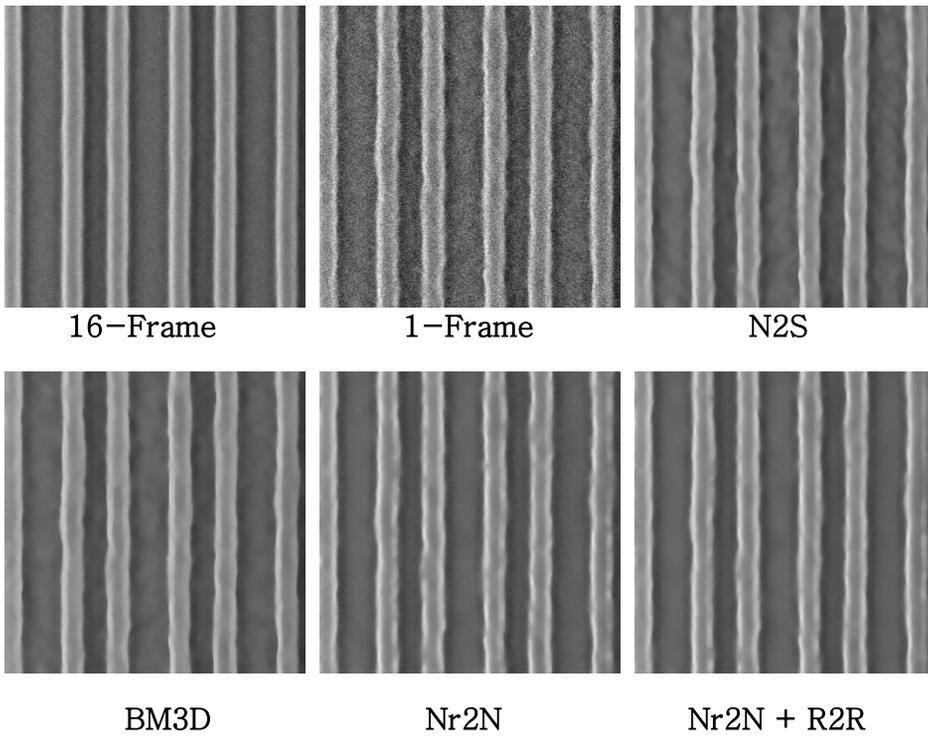
Figure 9. Qualitative comparison on SEM2 dataset

Method	SEM1	SEM2	SEM3
BM3D [14]	18.00	27.03	17.42
N2S [24]	15.69	24.13	15.48
Nr2N [29]	17.55	26.60	x
Nr2N + R2R (5)	17.96	26.85	x

Table 4. Quantitative comparison (PSNR(dB)) for SEM datasets. For each dataset, the highest values are marked bold while the second highest are underlined. '+R2R' means that inference trick has been added while the number inside the parenthesis corresponds to the number of averaged inferences.

Chapter 4. Improvements

4.1. Modeling-based Noisier2Noise

Despite the high performance of patch-based Nr2N, this method still cannot be utilized in real-world due to the patch extraction step. This step consumes extra time and cannot be applied in some datasets as mentioned in part 3.3. Furthermore, this step hinders the automation of the whole process since it requires troublesome manual tuning of parameter λ. As in Table 5, small shift of the parameter leads to drastic performance change which increases the necessity for precise control of the parameter for each dataset.

In order to deal with these circumstances, we propose 'Modeling-based Nr2N'. Instead of extracting noise patches to approximate the true noise distribution, a certain noise model will represent the distribution. Then the parameters of the noise model will be estimated via a pretrained network named 'Noise Parameter Estimation Network' or NPE-Net. With the predefined noise model and the estimated noise parameters, synthetic noise can be generated for the Nr2N training procedure.

In order to train a network like NPE-Net that can estimate noise parameters, ground truth clean images or computationally heavy process like GAT [63] is required as seen in previous studies [59], [60], [64]. However, these constraints can be omitted by

directly utilizing the flat regions in SEM images for noise parameter estimation. Since these flat regions have uniform pixel value, the ground truth clean images for training NPE-Net can be replaced with simple uniform images. Thus, dataset for training NPE-Net will be prepared as follows. First flat images with random intensity will be generated. Then the images will be corrupted with the predefined noise model and randomly generated noise parameter vectors. Finally, these images and their corresponding noise parameter vectors will be paired to form the training dataset.

Table 6 shows the result of a simple experiment. For each dataset, large number of noise patches have been extracted and put into pretrained NPE-Net to estimate the noise parameters. The distribution of the parameter in each dataset was highly homogeneous having small value of standard deviation. To the extreme degree, this implies that a single noise patch can represent the whole dataset for noise parameter estimation. Thus, with the help of pretrained NPE-Net, modeling-based Nr2N can mimic the process of patch-based Nr2N without the burdensome patch extraction step.

4.2. Iterative Noisier2Noise Training

Modeling-based Nr2N greatly simplifies the process of patch-based Nr2N, however falls behind in performance. This is due to the imprecise modeling of SEM noise, since gaussian noise or poisson noise is insufficient as describe in chapter 2.4. This can be

solved by incorporating the concept of 'Iterative training' [31] in our modeling-based Nr2N to form 'Iterative Nr2N' method.

As illustrated in Figure 9, Iterative Nr2N follows the concept of the modeling-based Nr2N concept. The main difference is that the noise patch and the training images will also be denoised as the test images. These denoised patch and training dataset will be reused as the dataset for the training and inference in the subsequent iteration.

The qualitative evaluation of iterative Nr2N method is illustrated in figure 10-12. The first iteration, which is equal to the simple modeling based Nr2N, shows superior image quality over BM3D, yet fails to reach the quality of patch-based Nr2N. However, from the second iteration, the prediction shows comparable image quality to the prediction of patch-based Nr2N. Moreover, in the quantitative comparison as in table 7, as the iteration continues, the PSNR of iterative Nr2N converges to the PSNR of patch-based method.

4.3. Further Evaluation

Since image quality and PSNR are insufficient for thorough evaluation, additional evaluation will be conducted. In figure 13-15, two 1-frame samples are denoised with iterative Nr2N method and each pair are compared with their 2D profiles. In table 8, LER and LWR values of 1-frame, 16-frame and images denoised by BM3D and iterative Nr2N are compared. 1-frame image shows the highest

values while image denoised with iterative Nr2N shows the lowest values.　Since LER and LWR are the variance of profile variations, iterative Nr2N showing the lowest values supports the superiority of iterative Nr2N. Furthermore, the 2D profiles show us that iterative Nr2N removes the unwanted variations in the profile while maintaining the main structure.
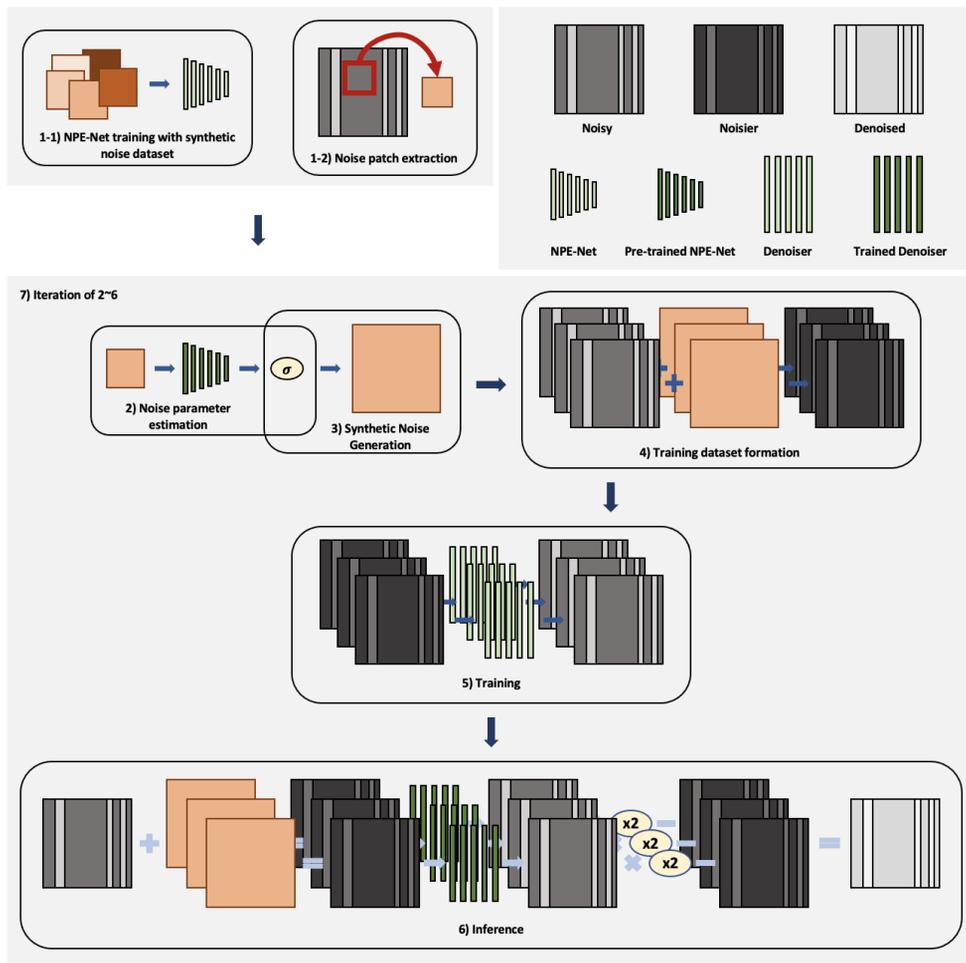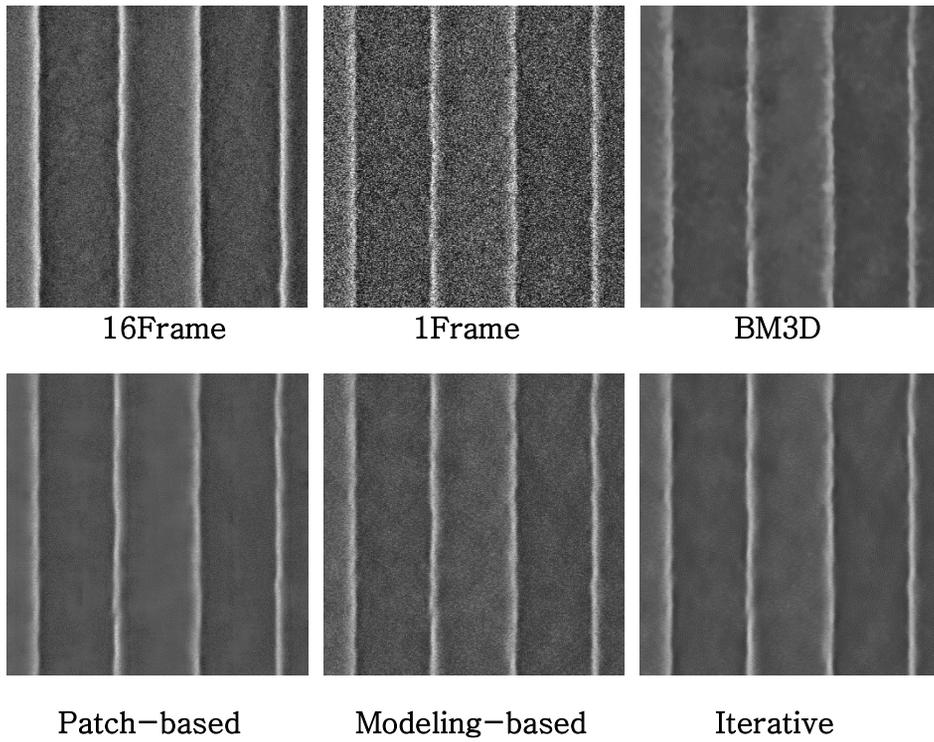
Figure 10. A diagram of the whole 'Iterative Nr2N' process.

Figure 11. Qualitative comparison on SEM1 dataset.

16Frame	1Frame	BM3D
Patch−based	Modeling−based	Iterative
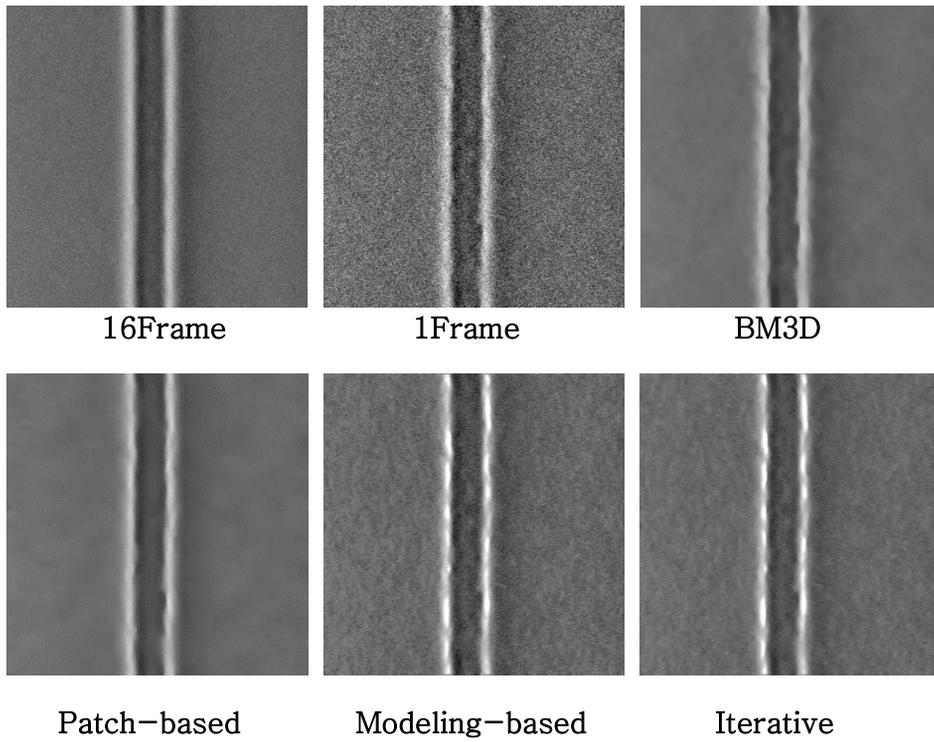
Figure 12. Qualitative comparison on SEM2 dataset

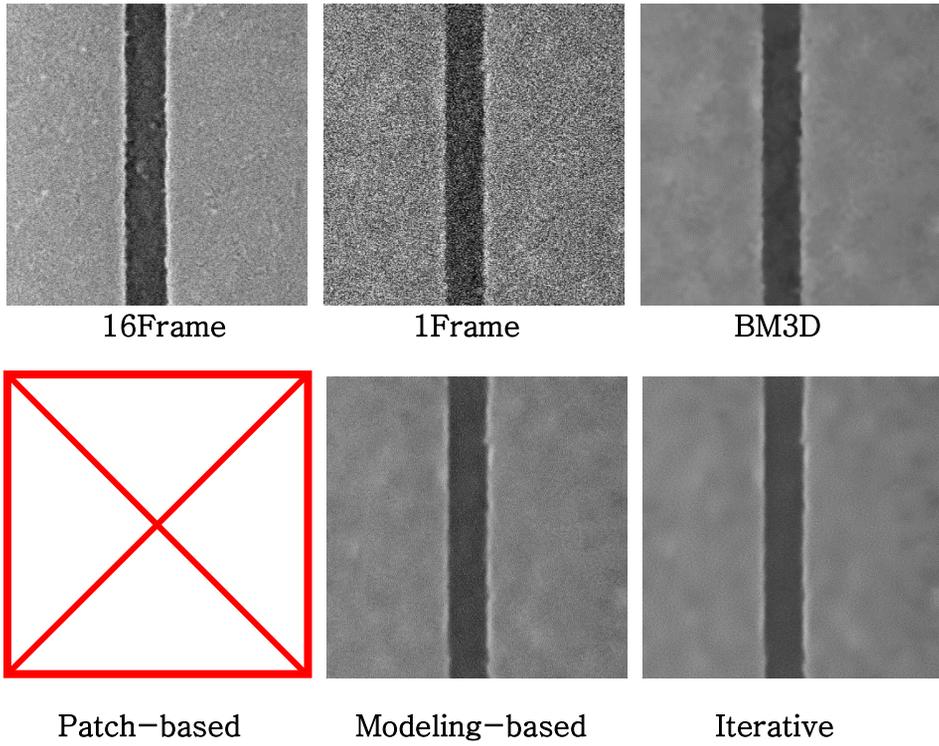| 16Frame | 1Frame | BM3D |
| Patch-based | Modeling-based | Iterative |

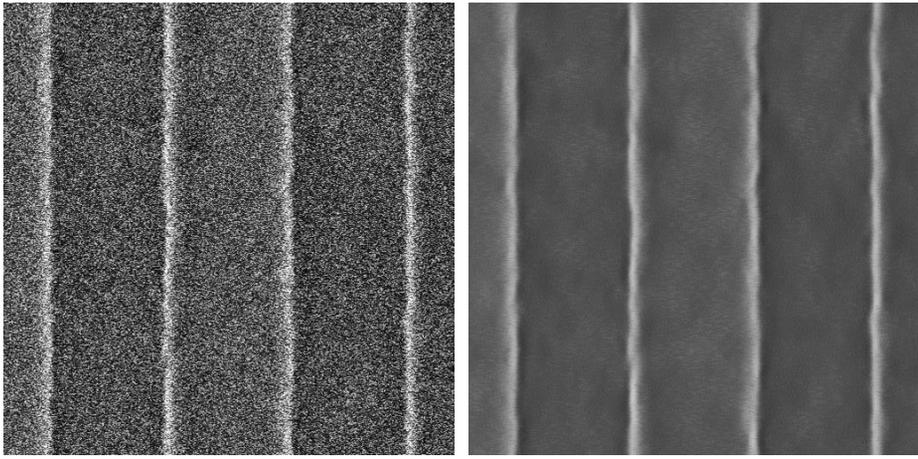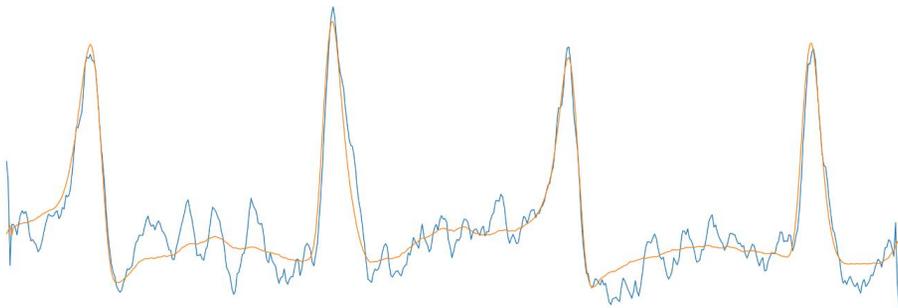**Figure 13.** Qualitative comparison on SEM3 dataset

1Frame Iterative

Figure 14. Noisy image, Denoised image and 2D profile comparison on SEM1 dataset
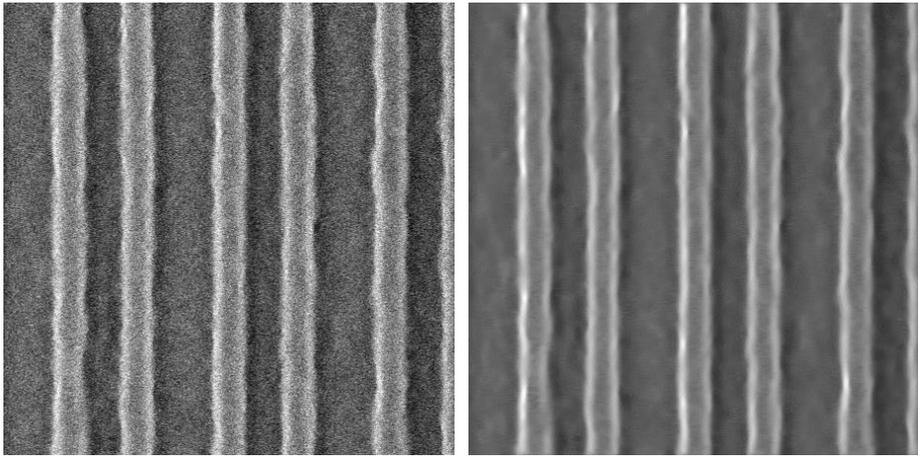
1Frame   Iterative
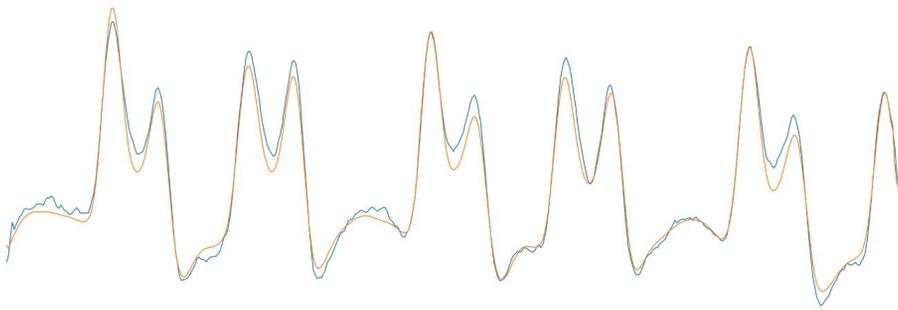
Figure 15. Noisy image, Denoised image and 2D profile comparison on SEM2 dataset

1Frame   Iterative
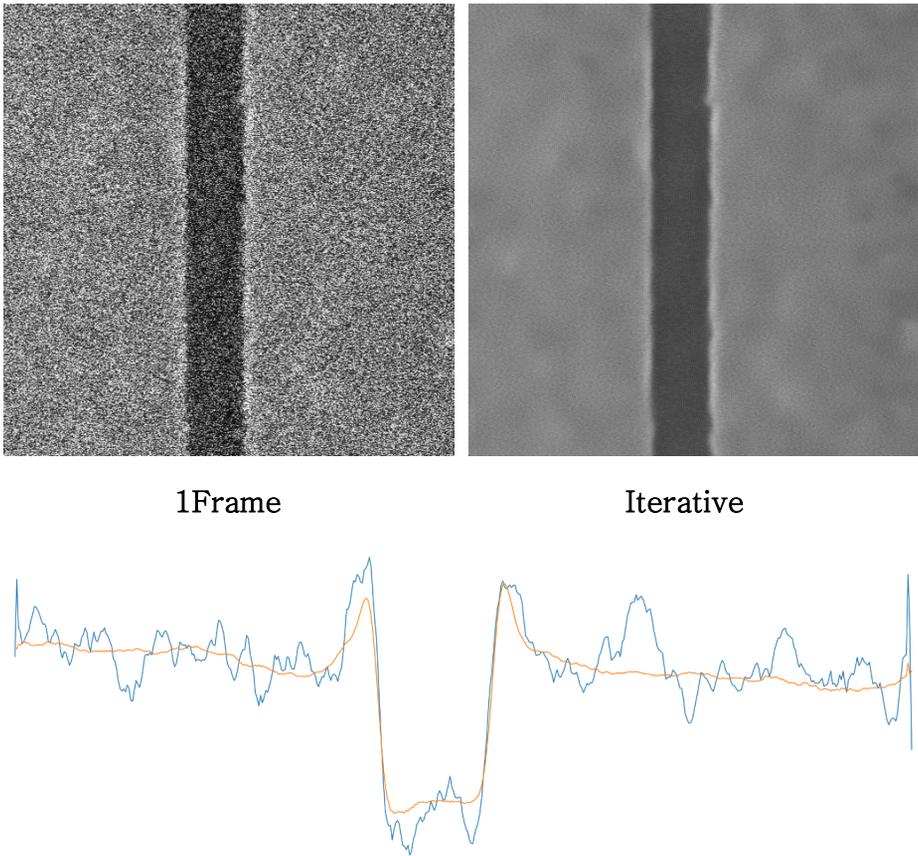
Figure 16. Noisy image, Denoised image and 2D profile comparison on SEM3 dataset

Parameter λ	SEM1
Single	13.25
0.19	17.82
0.23	17.92
0.28	16.79

Parameter λ	SEM2
Single	22.22
0.20	26.83
0.30	24.51
0.60	20.31

Table 5. Denoising performance (PSNR(dB)) for SEM datasets with different parameter values. Single corresponds to an extreme case where only one patch is extracted.

	Gaussian (std)		Poisson (peak)	
	mean	std	mean	std
SEM1	0.2128	0.0078	0.0314	0.0032
SEM2	0.0770	0.0050	0.2590	0.0439
SEM3	0.2613	0.0030	0.0277	0.0008

Table 6. Mean and standard deviation of the estimated noise parameters from SEM datasets. Gaussian noise model case and Poisson noise model case are considered.

Method	SEM1	SEM2	SEM3
BM3D [14]	18.00	27.03	17.42
Patch-based Nr2N	17.96	26.85	x
Modeling-based Nr2N	16.96	25.62	16.16
Iterative Nr2N	18.00	26.37	17.28

Table 7. Quantitative comparison (PSNR(dB)) for SEM datasets.

	LER	LWR
1−Frame	5.122	7.220
16−Frame	2.310	3.232
BM3D	2.229	3.044
Iterative Nr2N	1.865	2.585

Table 8. LER, LWR comparison on SEM1 dataset

Chapter 5. Experimental Details

5.1. Training and Evaluation

In the case of public datasets, training have been conducted with 1000 grayscale ImageNet images cropped with the size of 256 while evaluation have been performed on grayscale BSD100, Kodak and Set12 cropped with the size of 256. For the traditional methods, hyper-parameters with the highest PSNR values have been chosen for comparison. For deep learning-based methods, hyper-parameters and experimental details were chosen based on the papers and official codes for each method.

Two neural networks have been used for training and evaluation. DnCNN have been used for the main denoising network. For the noise parameter estimation network, the last layer of the PGE-Net in FBI-Denoiser have been replaced with global average pooling layer in order to output the estimated parameters directly.

5.2. Extra Results

In the following pages, extra results are displayed. First, extra qualitative results on both public datasets and SEM datasets are shown. Then contour extraction results on SEM datasets are shown. The extraction is conducted with canny algorithm with the optimal parameter chosen experimentally.
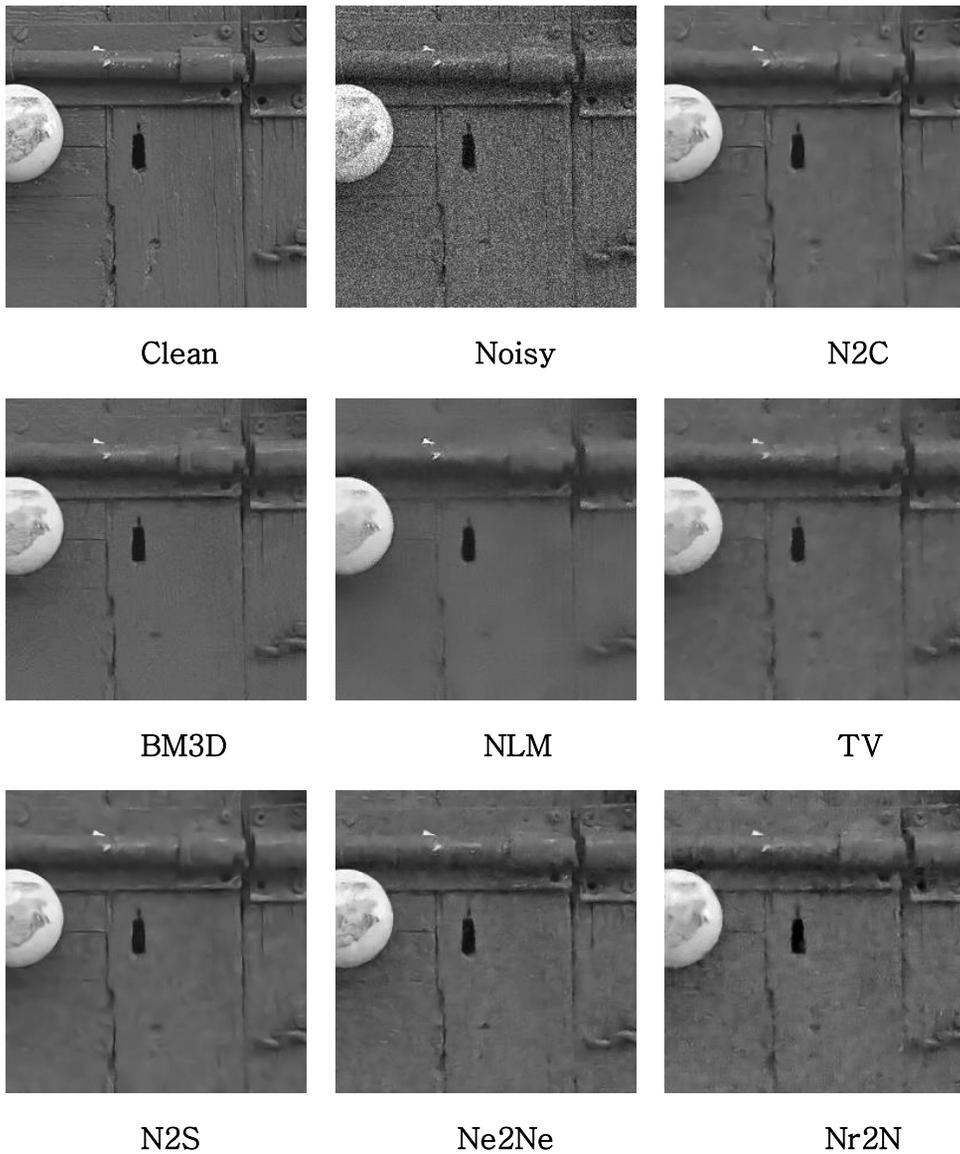
Figure 17. Qualitative evaluation on Kodak dataset. Clean image is corrupted by gaussian noise ($\sigma=25$).
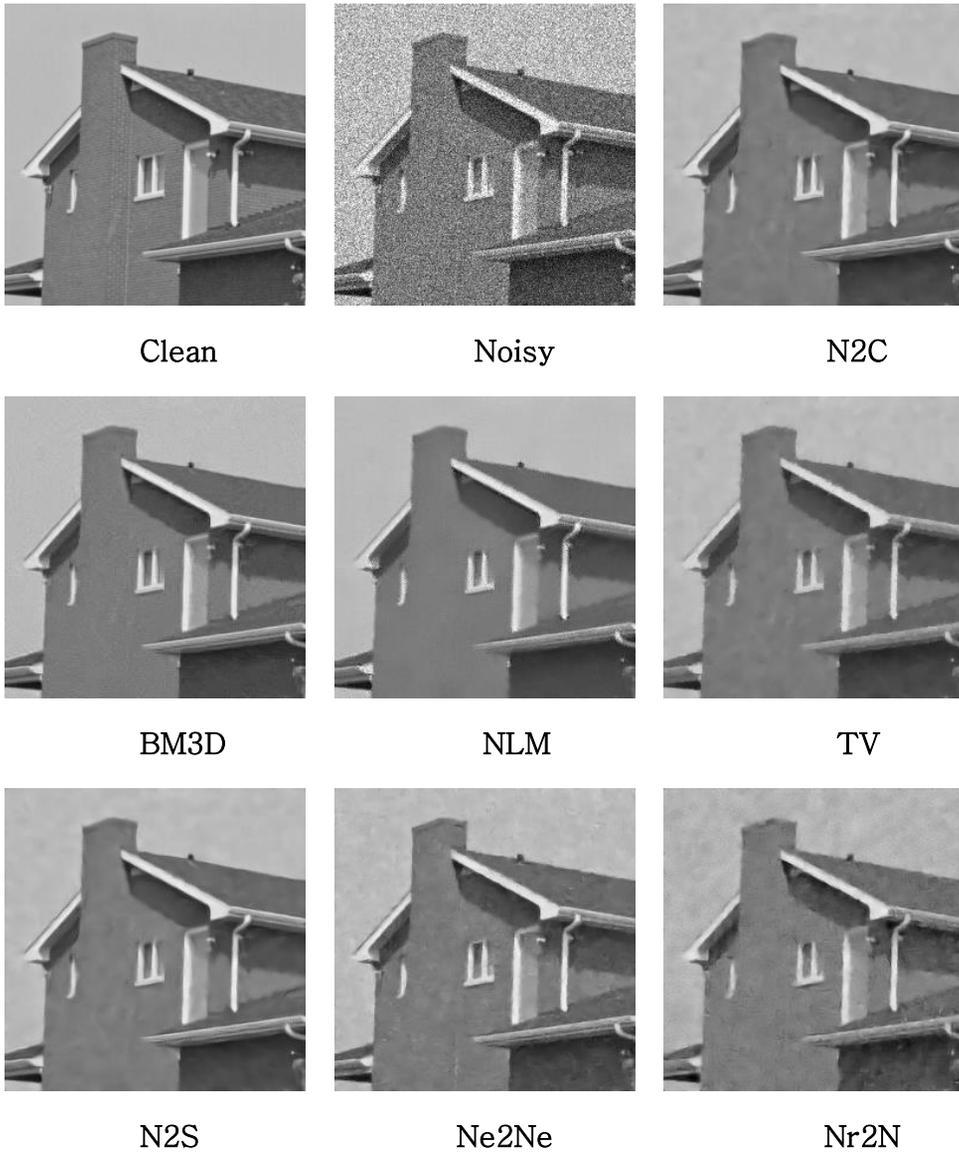
Figure 18. Qualitative evaluation on Set12 dataset. Clean image is corrupted by gaussian noise (σ=25).

Clean	Noisy	N2C
BM3D	NLM	TV
N2S	Ne2Ne	Nr2N

Figure 19. Qualitative evaluation on BSD100 dataset. Clean image is corrupted by gaussian noise (σ=50).

Figure 20. Qualitative evaluation on Kodak dataset. Clean image is corrupted by gaussian noise (σ=50).

Figure 21. Qualitative evaluation on Set12 dataset. Clean image is corrupted by gaussian noise (σ=50).

Figure 22. Qualitative evaluation on BSD100 dataset. Clean image is corrupted by poisson noise $(\alpha=25)$.

Figure 23. Qualitative evaluation on Kodak dataset. Clean image is corrupted by poisson noise ($\alpha=25$).

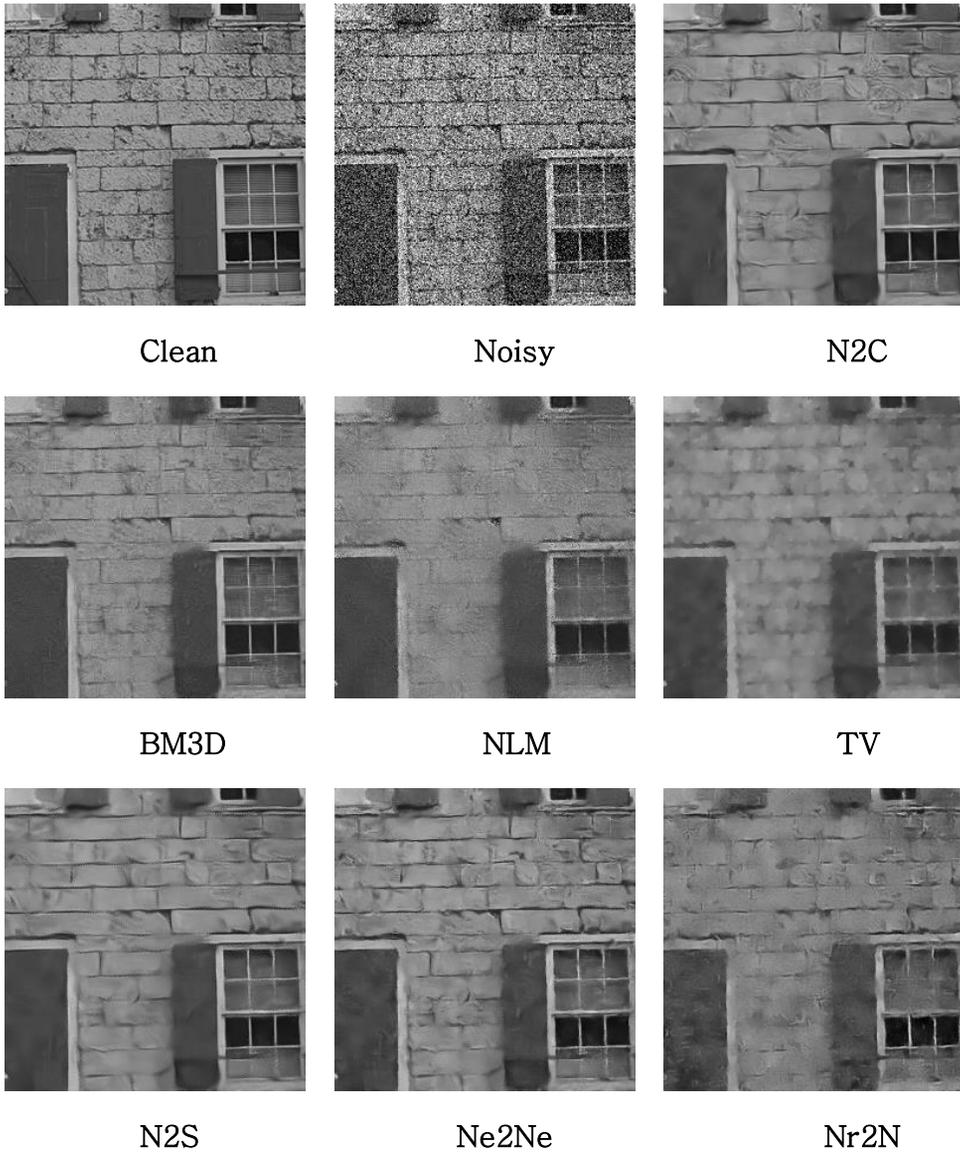|		
Clean	Noisy	N2C
BM3D	NLM	TV
N2S	Ne2Ne	Nr2N
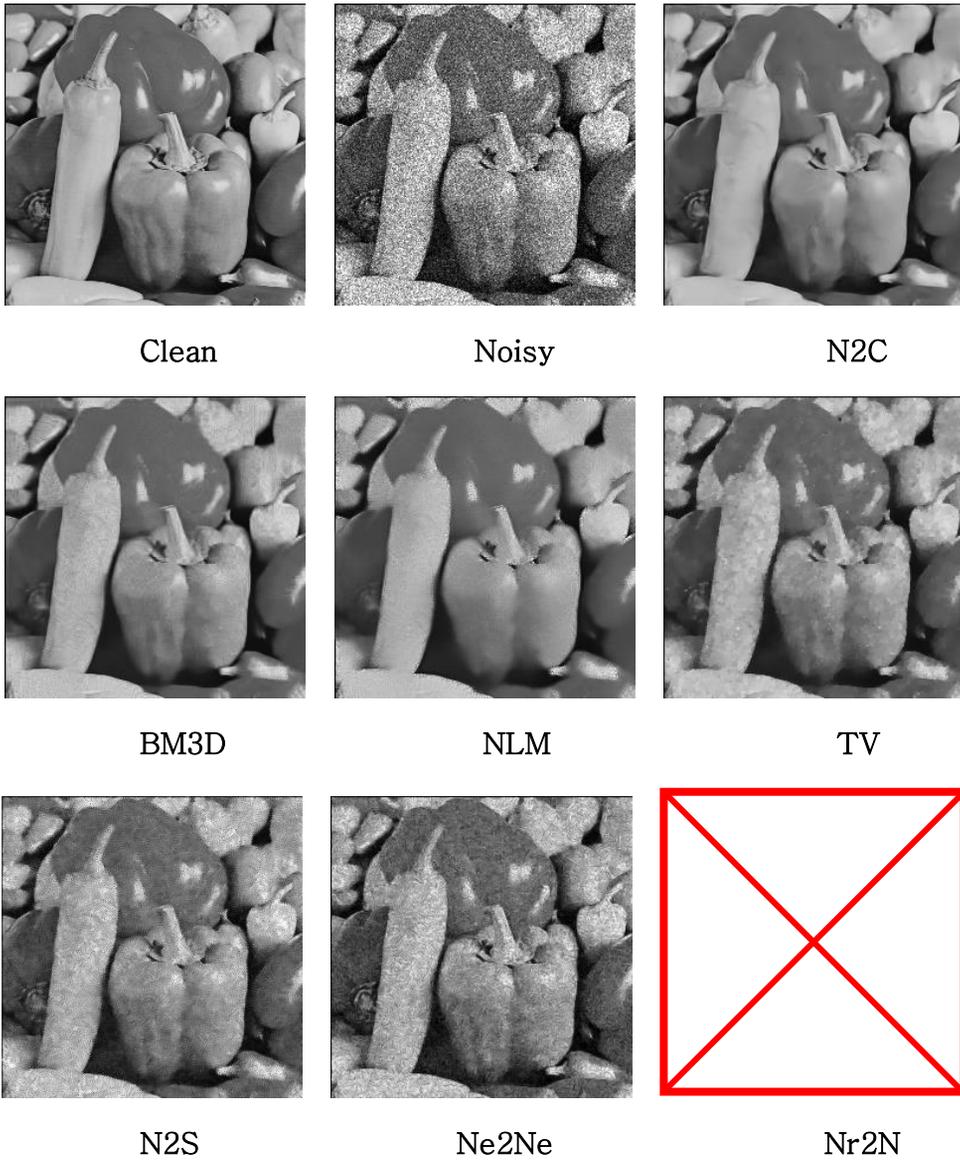
Figure 24. Qualitative evaluation on Set12 dataset. Clean image is corrupted by poisson noise ($\alpha=25$).

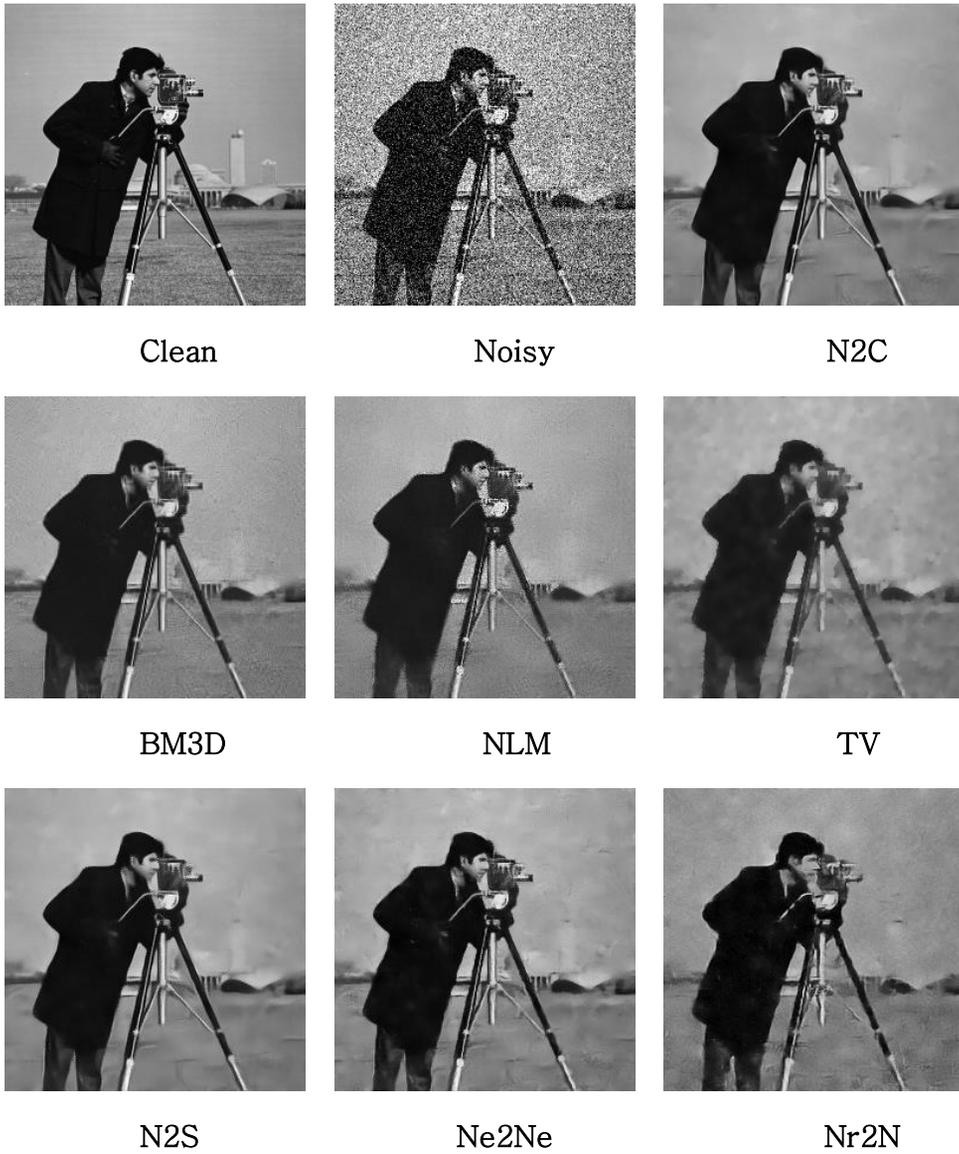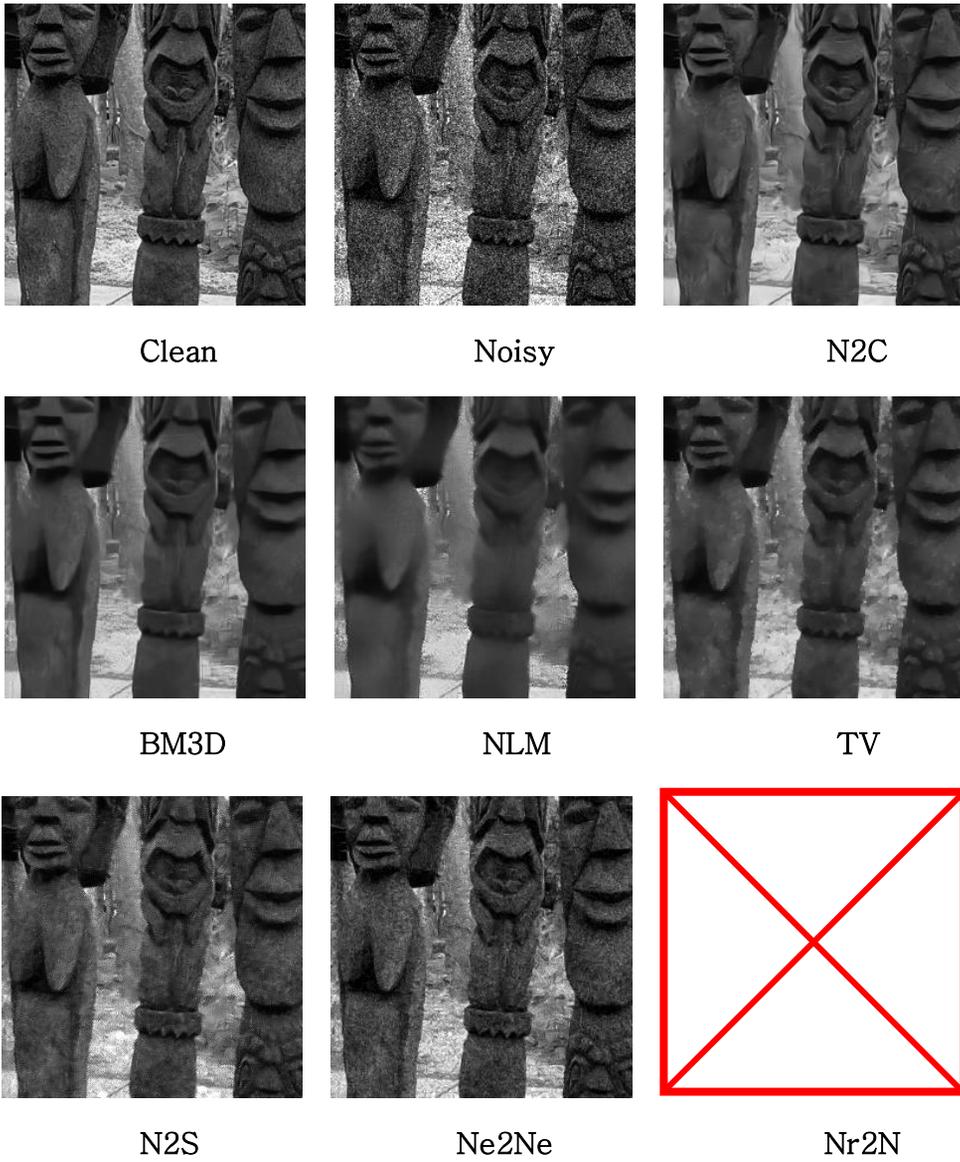|		
Clean	Noisy	N2C
BM3D	NLM	TV
N2S	Ne2Ne	Nr2N

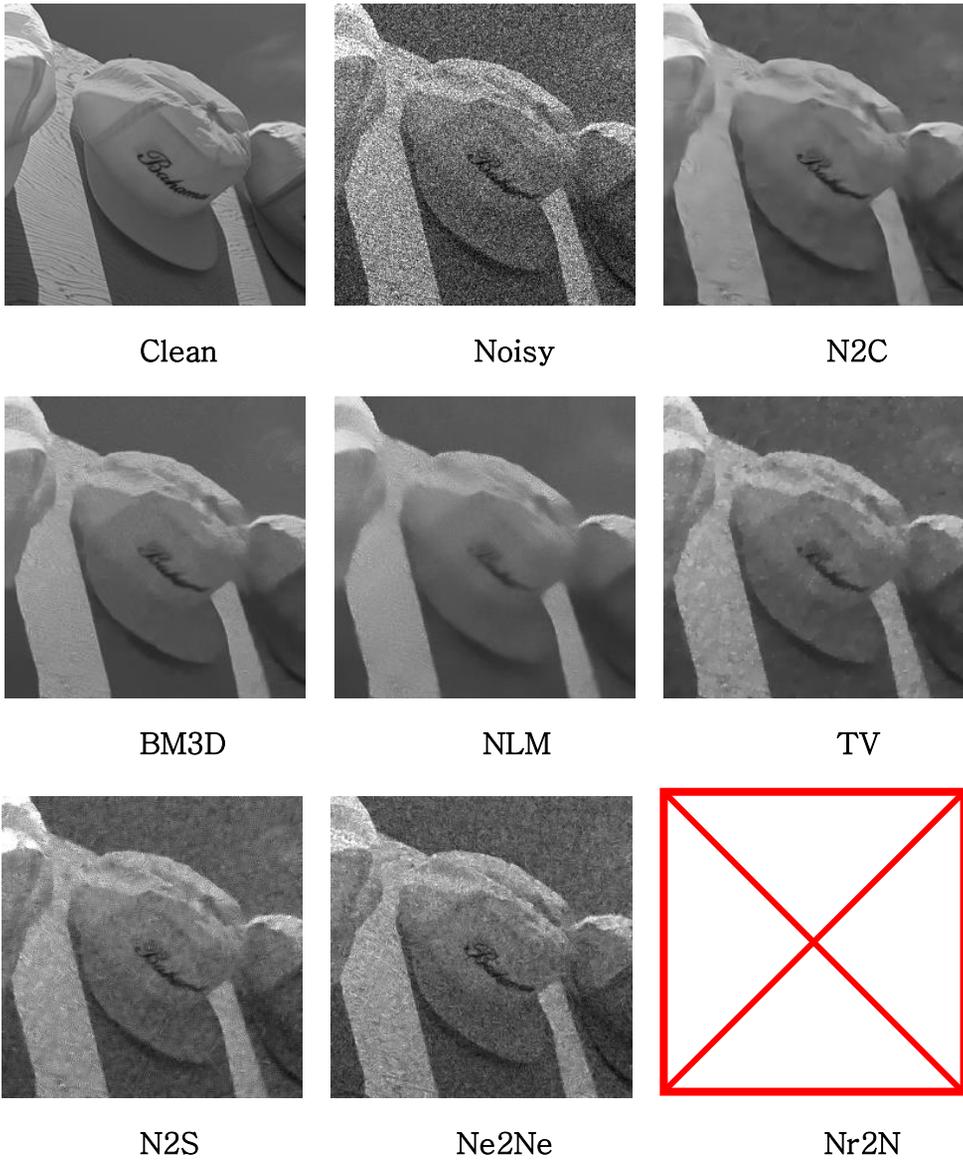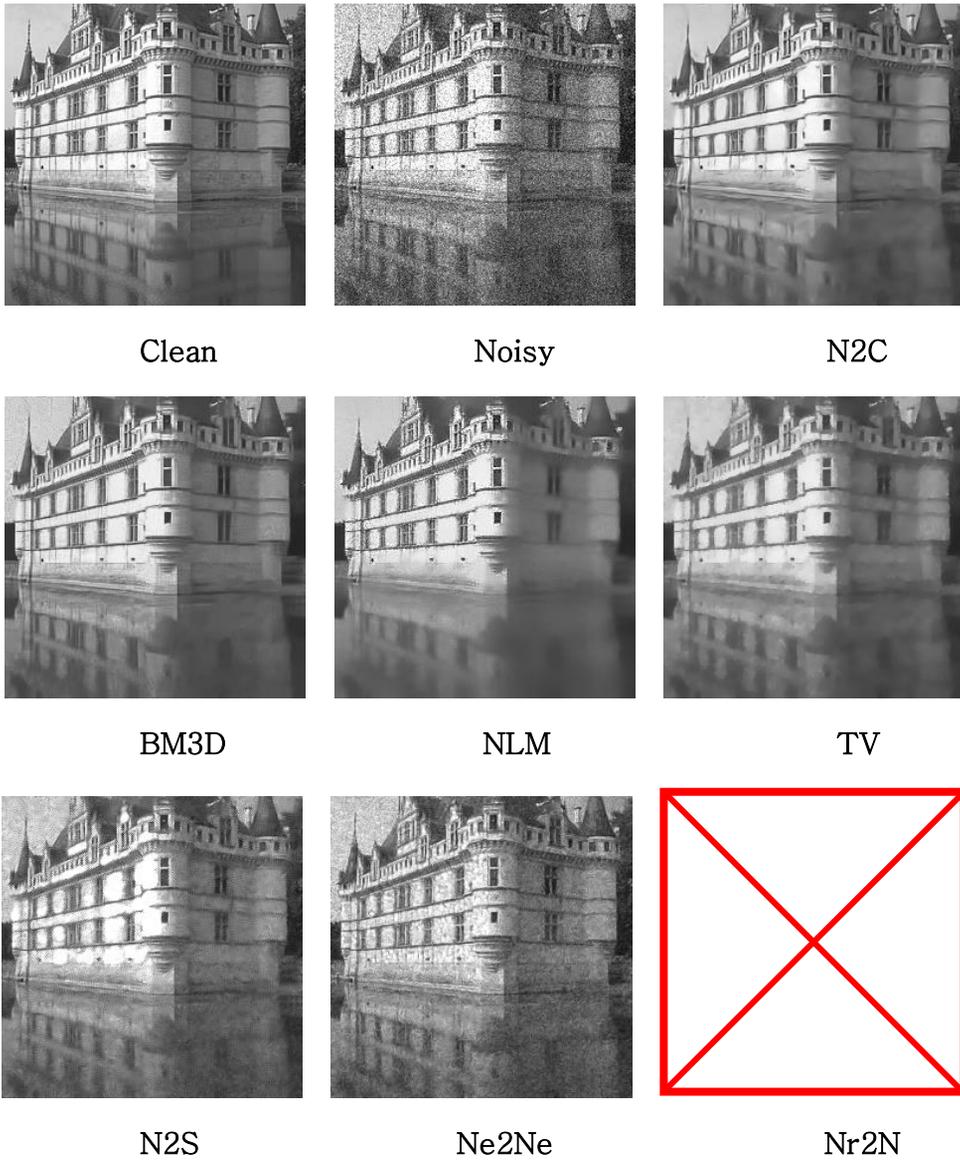**Figure 25.** Qualitative evaluation on BSD100 dataset. Clean image is corrupted by poisson noise (α=50).
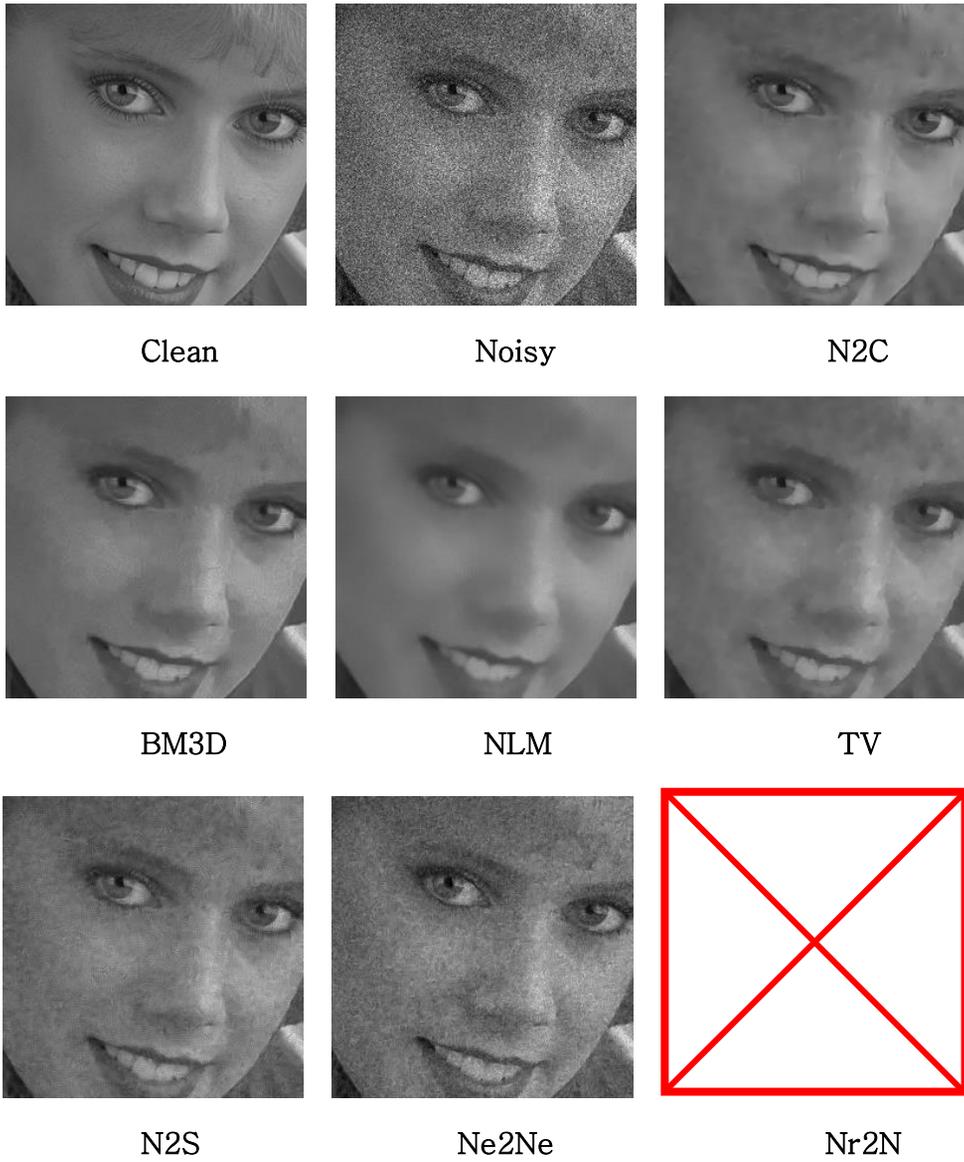
Figure 26. Qualitative evaluation on Kodak dataset. Clean image is corrupted by poisson noise $(\alpha=50)$.
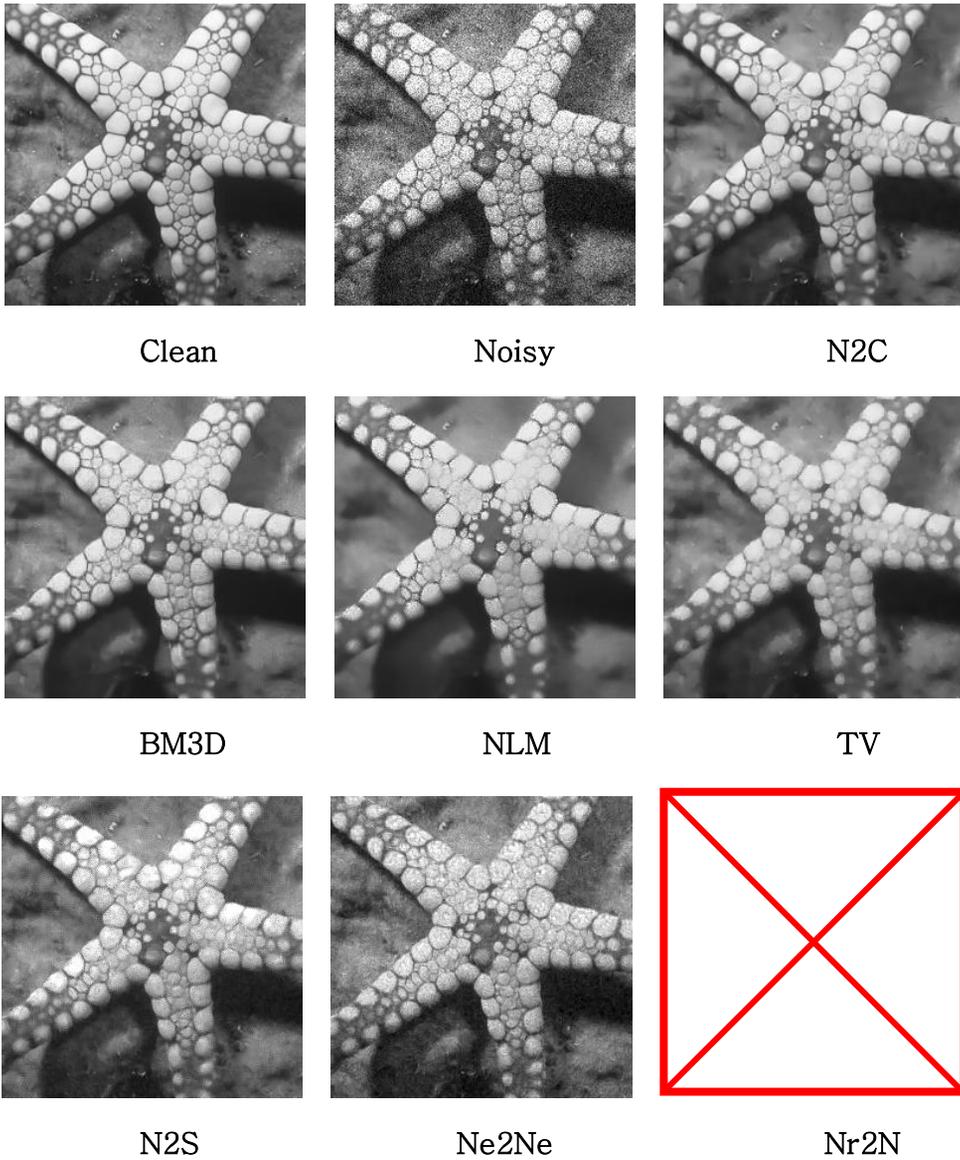
Figure 27. Qualitative evaluation on Set12 dataset. Clean image is corrupted by poisson noise ($\alpha=50$).

16Frame | 1Frame | BM3D

Patch−based | Modeling−based | Iterative
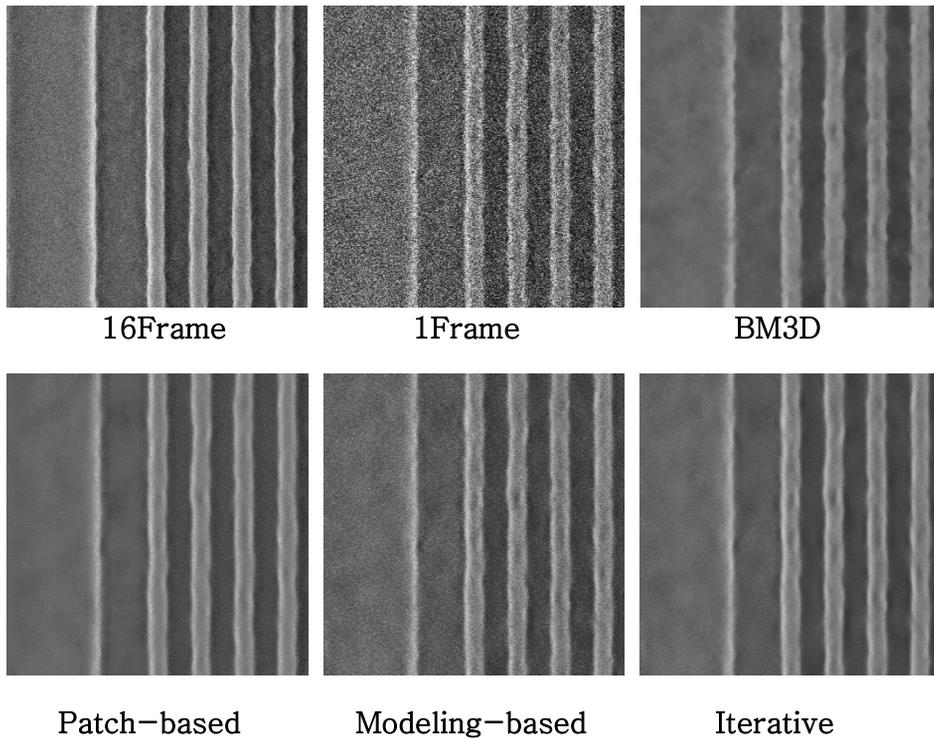
Figure 28. Qualitative evaluation on SEM1 dataset.

| 16Frame | 1Frame | BM3D |

| Patch-based | Modeling-based | Iterative |
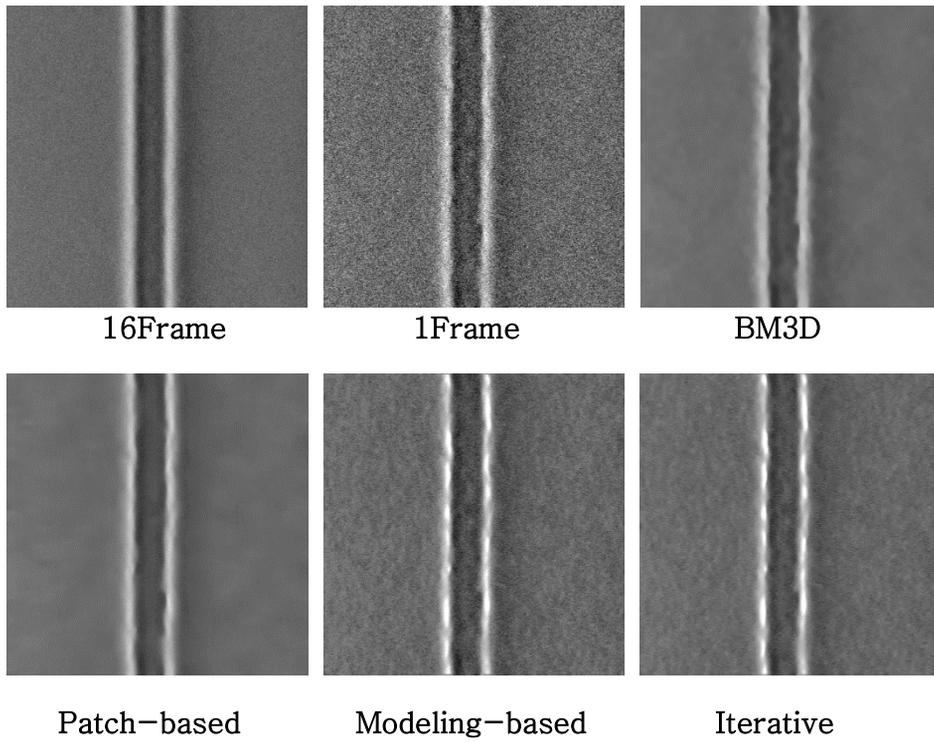
Figure 29. Qualitative evaluation on SEM2 dataset.

| 16Frame | 1Frame | BM3D |
| Patch-based | Modeling-based | Iterative |
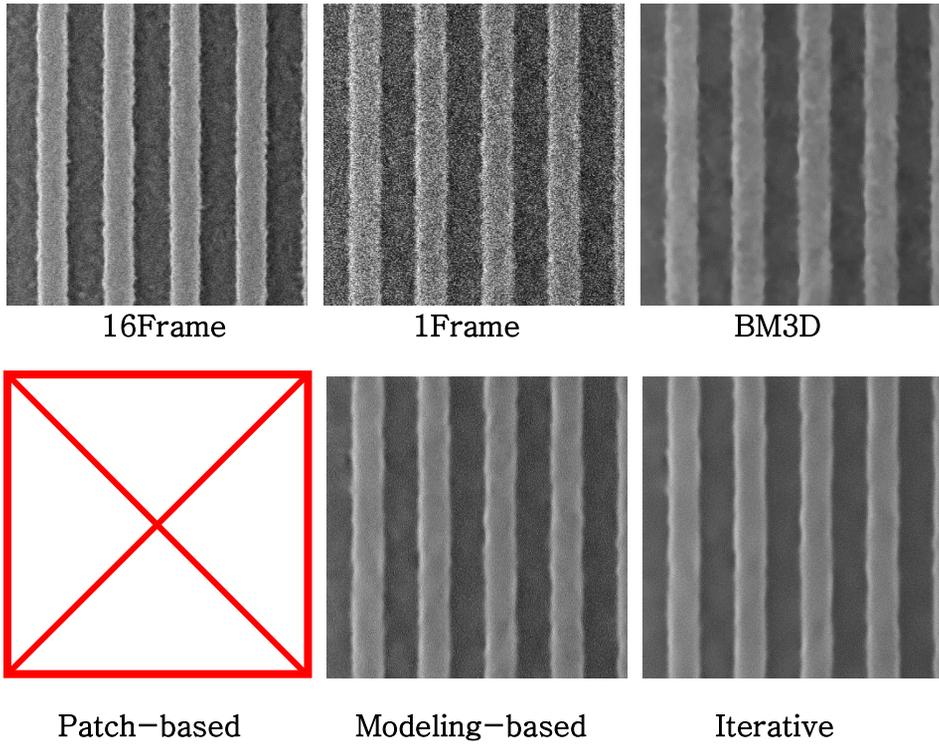
Figure 30. Qualitative evaluation on SEM3 dataset.
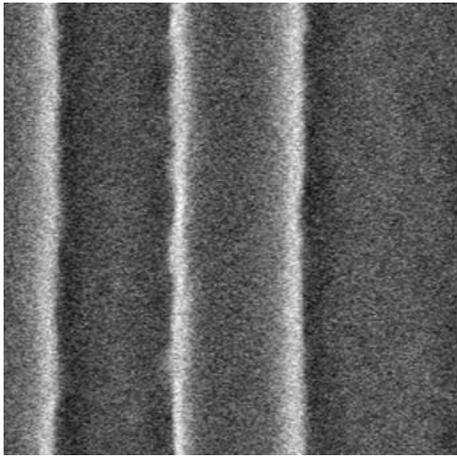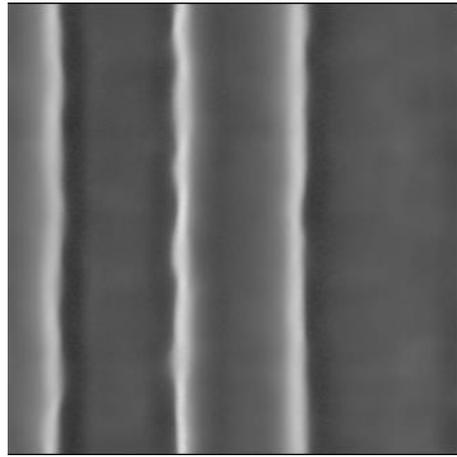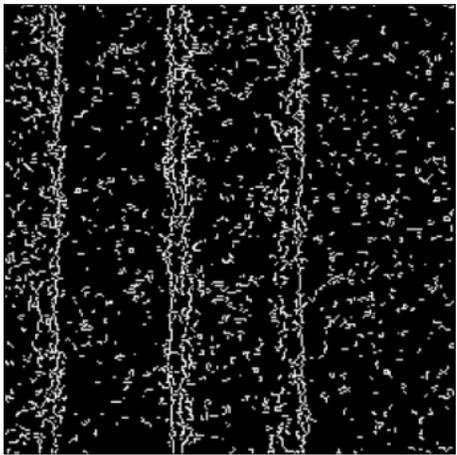
1Frame   Iterative
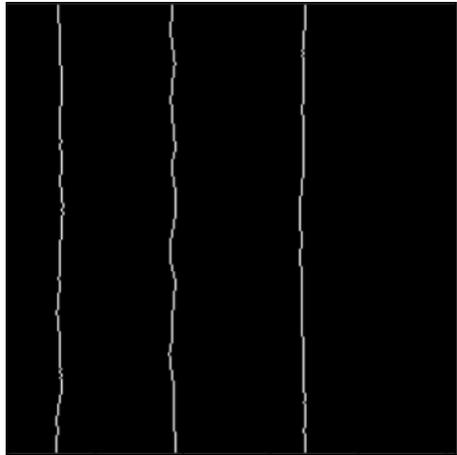
1Frame（Contour）   Iterative（Contour）

Figure 31. Contour extraction on SEM dataset with Canny.

Chapter 6. Conclusion

We propose patch-based Noisier2Noise which is specifically designed to denoise SEM images without any ground truth clean image. Our method successfully removes the noise by combining patch extraction, self-supervised training and inference trick. The success of our method is further justified with the analysis of SEM noise with the grain size concept. Moreover, we propose modeling-based Noisier2Noise which enables the automation process by utilizing noise parameter network. Furthermore, the performance degradation is alleviated with the help of iterative training framework. Quantitatively, our methods show comparable performance to the traditional methods in PSNR and improved LER, LWR values while showing much faster computation. Qualitatively, our methods show enhanced visual quality and contour extraction results while preserving the image structure.

Bibliography

[1]   N. G. Orji *et al.*, "Metrology for the next generation of semiconductor devices," *Nature Electronics*, vol. 1, no. 10. 2018. doi: 10.1038/s41928-018-0150-9.

[2]   Y. Ban, S. Sundareswaran, R. Panda, and D. Z. Pan, "Electrical impact of line-edge roughness on sub-45nm node standard cell," in *Design for Manufacturability through Design-Process Integration III*, 2009, vol. 7275. doi: 10.1117/12.814355.

[3]   L. Capodieci, "Design-driven metrology: a new paradigm for DFM-enabled process characterization and control: extensibility and limitations," in *Metrology, Inspection, and Process Control for Microlithography XX*, 2006, vol. 6152. doi: 10.1117/12.659181.

[4]   R. Yumiba *et al.*, "Deep learning's impact on contour extraction for design based metrology and design based inspection," 2019. doi: 10.1117/12.2514898.

[5]   B. Dey *et al.*, "Unsupervised machine learning based SEM image denoising for robust contour detection," 2021. doi: 10.1117/12.2600945.

[6]   J. S. Villarrubia and B. D. Bunday, "Unbiased estimation of linewidth roughness," in *Metrology, Inspection, and Process Control for Microlithography XIX*, 2005, vol. 5752. doi: 10.1117/12.599981.

[7]   S. Kobayashi *et al.*, "New method removing SEM image noise to characterize CD and LWR," 2019. doi: 10.1117/12.2514661.

[8]   L. Fan, F. Zhang, H. Fan, and C. Zhang, "Brief review of image denoising techniques," *Visual Computing for Industry, Biomedicine, and Art*, vol. 2, no. 1. 2019. doi: 10.1186/s42492-019-0016-7.

[9]   G. Deng and L. W. Cahill, "Adaptive Gaussian filter for noise reduction and edge detection," in *IEEE Nuclear Science Symposium & Medical Imaging Conference*, 1994, no. pt 3. doi: 10.1109/nssmic.1993.373563.

[10]   R. Yang, L. Yin, M. Gabbouj, J. Astola, and Y. Neuvo, "Optimal weighted median filtering under structural constraints," *IEEE Transactions on Signal Processing*, vol. 43, no. 3, 1995, doi: 10.1109/78.370615.

[11]   C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proceedings of the IEEE International Conference on Computer Vision*, 1998. doi: 10.1109/iccv.1998.710815.

[12]   L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 1-4, 1992, doi: 10.1016/0167-2789(92)90242-F.

[13]   A. Buades, B. Coll, and J.-M. Morel, "Non-Local Means Denoising," *Image Processing On Line*, vol. 1, 2011, doi: 10.5201/ipol.2011.bcm_nlm.

[14]   K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 8, 2007, doi: 10.1109/TIP.2007.901238.

[15]   R. D. da Silva, R. Minetto, W. R. Schwartz, and H. Pedrini, "Adaptive edge-preserving image denoising using wavelet transforms," *Pattern Analysis and Applications*, vol. 16, no. 4, 2013, doi: 10.1007/s10044-012-0266-x.

[16]   C. Tian, L. Fei, W. Zheng, Y. Xu, W. Zuo, and C. W. Lin, "Deep learning on image denoising: An overview," *Neural Networks*, vol. 131. 2020. doi: 10.1016/j.neunet.2020.07.025.

[17]   V. Jain and H. S. Seung, "Natural image denoising with convolutional

networks," in *Advances in Neural Information Processing Systems 21 – Proceedings of the 2008 Conference*, 2009.

[18]   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*, vol. 26, no. 7, 2017, doi: 10.1109/TIP.2017.2662206.

[19]   S. Cheng, Y. Wang, H. Huang, D. Liu, H. Fan, and S. Liu, "NBNet: Noise Basis Learning for Image Denoising with Subspace Projection," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2021. doi: 10.1109/CVPR46437.2021.00486.

[20]   K. Zhang, W. Zuo, and L. Zhang, "FFDNet: Toward a fast and flexible solution for CNN-Based image denoising," *IEEE Transactions on Image Processing*, vol. 27, no. 9, 2018, doi: 10.1109/TIP.2018.2839891.

[21]   Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, 2004, doi: 10.1109/TIP.2003.819861.

[22]   J. Lehtinen *et al.*, "Noise2Noise: Learning image restoration without clean data," in *35th International Conference on Machine Learning, ICML 2018*, 2018, vol. 7.

[23]   A. Krull, T. O. Buchholz, and F. Jug, "Noise2void-Learning denoising from single noisy images," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2019, vol. 2019-June. doi: 10.1109/CVPR.2019.00223.

[24]   J. Batson and L. Royer, "Noise2Seif: Blind denoising by self-supervision," in *36th International Conference on Machine Learning, ICML 2019*, 2019, vol. 2019-June.

[25]   S. Laine, T. Karras, J. Lehtinen, and T. Aila, "High-quality self-supervised deep image denoising," in *Advances in Neural Information Processing Systems*, 2019, vol. 32.

[26]   J. Xu *et al.*, "Noisy-as-Clean: Learning Self-Supervised Denoising From Corrupted Image," *IEEE Transactions on Image Processing*, vol. 29, 2020, doi: 10.1109/tip.2020.3026622.

[27]   W. Khademi, S. Rao, C. Minnerath, G. Hagen, and J. Ventura, "Self-supervised poisson-gaussian denoising," in *Proceedings - 2021 IEEE Winter Conference on Applications of Computer Vision, WACV 2021*, 2021. doi: 10.1109/WACV48630.2021.00218.

[28]   T. Huang, S. Li, X. Jia, H. Lu, and J. Liu, "Neighbor2Neighbor: Self-supervised denoising from single noisy images," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2021. doi: 10.1109/CVPR46437.2021.01454.

[29]   N. Moran, D. Schmidt, Y. Zhong, and P. Coady, "Noisier2Noise: Learning to denoise from unpaired noisy data," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2020. doi: 10.1109/CVPR42600.2020.01208.

[30]   T. Pang, H. Zheng, Y. Quan, and H. Ji, "Recorrupted-to-Recorrupted: Unsupervised Deep Learning for Image Denoising," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2021. doi: 10.1109/CVPR46437.2021.00208.

[31]   G. Kahl, " [ICLR ' 21]gan2gan_generative_noise_learning_for_blind_denoising_with_single_noisy_images," *The Dictionary of Genomics, Transcriptomics and Proteomics*, no. 2018, 2015.

[32] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in *34th International Conference on Machine Learning, ICML 2017*, 2017, vol. 1.

[33] G. P. Patsis, V. Constantoudis, A. Tserepi, E. Gogolides, and G. Grozev, "Quantification of line-edge roughness of photoresists. I. A comparison between off-line and on-line analysis of top-down scanning electron microscopy images," *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures*, vol. 21, no. 3, 2003, doi: 10.1116/1.1570845.

[34] D. Li *et al.*, "Noise filtering for accurate measurement of line edge roughness and critical dimension from SEM images," *Journal of Vacuum Science & Technology B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena*, vol. 34, no. 6, 2016, doi: 10.1116/1.4968184.

[35] C. A. Mack and G. F. Lorusso, "Determining the ultimate resolution of scanning electron microscope-based unbiased roughness measurements. I. Simulating noise," *Journal of Vacuum Science & Technology B*, vol. 37, no. 6, 2019, doi: 10.1116/1.5122758.

[36] T. Kameda, S. Takada, M. Suzuki, T. Yokosuka, S. Borisov, and S. Babin, "SEM image prediction based on modeling of electron-solid interaction," in *Metrology, Inspection, and Process Control for Microlithography XXXI*, 2017, vol. 10145. doi: 10.1117/12.2257661.

[37] A. Yamaguchi, R. Steffen, H. Kawada, and T. Iizumi, "Bias-free measurement of LER/LWR with low damage by CD-SEM," in *Metrology, Inspection, and Process Control for Microlithography XX*, 2006, vol. 6152. doi: 10.1117/12.655496.

[38] G. F. Lorusso, V. Rutigliani, F. van Roey, and C. A. Mack, "Unbiased roughness measurements: Subtracting out SEM effects," *Microelectron Eng*, vol. 190, 2018, doi: 10.1016/j.mee.2018.01.010.

[39] G. F. Lorusso, V. Rutigliani, F. van Roey, and C. A. Mack, "Unbiased roughness measurements: Subtracting out SEM effects, part 2," *Journal of Vacuum Science & Technology B*, vol. 36, no. 6, 2018, doi: 10.1116/1.5046477.

[40] V. Constantoudis and E. Gogolides, "Noise-free estimation of spatial line edge/width roughness parameters," in *Metrology, Inspection, and Process Control for Microlithography XXIII*, 2009, vol. 7272. doi: 10.1117/12.822729.

[41] "Line edge roughness metrology: recent challenges and advances toward more complete and accurate measurements," *Journal of Micro/Nanolithography, MEMS, and MOEMS*, vol. 17, no. 04, 2018, doi: 10.1117/1.jmm.17.4.041014.

[42] E. Giannatou, G. Papavieros, V. Constantoudis, H. Papageorgiou, and E. Gogolides, "Deep learning denoising of SEM images towards noise-reduced LER measurements," *Microelectron Eng*, vol. 216, 2019, doi: 10.1016/j.mee.2019.111051.

[43] N. Chaudhary, S. A. Savari, and S. S. Yeddulapalli, "Line roughness estimation and Poisson denoising in scanning electron microscope images using deep learning," *Journal of Micro/Nanolithography, MEMS, and MOEMS*, vol. 18, no. 02, 2019, doi: 10.1117/1.jmm.18.2.024001.

[44] Y. Midoh and K. Nakamae, "Image quality enhancement of a CD-SEM image using conditional generative adversarial networks," 2019. doi: 10.1117/12.2515152.

[45] B. Dey *et al.*, "SEM image denoising with unsupervised machine learning for better defect inspection and metrology," 2021. doi: 10.1117/12.2584803.

6 4

[46] L. Yu, W. Zhou, L. Pu, and W. Fang, "SEM image quality enhancement: an unsupervised deep learning approach," 2020. doi: 10.1117/12.2552883.

[47] H. Lei, C. Teh, L. Yu, G. Fu, L. Pu, and W. Fang, "Denoising sample−limited SEM images without clean data," 2021. doi: 10.1117/12.2584653.

[48] A. Foi, M. Trimeche, V. Katkovnik, and K. Egiazarian, "Practical Poissonian−Gaussian noise modeling and fitting for single−image raw−data," *IEEE Transactions on Image Processing*, vol. 17, no. 10, 2008, doi: 10.1109/TIP.2008.2001399.

[49] V. A. Pimpalkhute, R. Page, A. Kothari, K. M. Bhurchandi, and V. M. Kamble, "Digital Image Noise Estimation Using DWT Coefficients," *IEEE Transactions on Image Processing*, vol. 30, 2021, doi: 10.1109/TIP.2021.3049961.

[50] S. Lee, M. S. Lee, and M. G. Kang, "Poisson−gaussian noise analysis and estimation for low−dose X−ray images in the NSCT domain," *Sensors (Switzerland)*, vol. 18, no. 4, 2018, doi: 10.3390/s18041019.

[51] D. Meng and F. D. la Torre, "Robust matrix factorization with unknown noise," in *Proceedings of the IEEE International Conference on Computer Vision*, 2013. doi: 10.1109/ICCV.2013.169.

[52] X. Liu, M. Tanaka, and M. Okutomi, "Single−image noise level estimation for blind denoising," *IEEE Transactions on Image Processing*, vol. 22, no. 12, 2013, doi: 10.1109/TIP.2013.2283400.

[53] S. Pyatykh, J. Hesser, and L. Zheng, "Image noise level estimation by principal component analysis," *IEEE Transactions on Image Processing*, vol. 22, no. 2, 2013, doi: 10.1109/TIP.2012.2221728.

[54] Q. Zhao, D. Mengt, Z. Xut, W. Zuo, and L. Zhang, "Robust principal component analysis with complex noise," in *31st International Conference on Machine Learning, ICML 2014*, 2014, vol. 2.

[55] S. Nam, Y. Hwang, Y. Matsushita, and S. J. Kim, "A Holistic Approach to Cross−Channel Image Noise Modeling and Its Application to Image Denoising," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016, vol. 2016−December. doi: 10.1109/CVPR.2016.186.

[56] J. Chen, J. Chen, H. Chao, and M. Yang, "Image Blind Denoising with Generative Adversarial Network Based Noise Modeling," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2018. doi: 10.1109/CVPR.2018.00333.

[57] D. W. Kim, J. R. Chung, and S. W. Jung, "GRDN:Grouped residual dense network for real image denoising and GAN−based real−world noise modeling," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2019, vol. 2019−June. doi: 10.1109/CVPRW.2019.00261.

[58] Z. Yue, Q. Zhao, L. Zhang, and D. Meng, "Dual Adversarial Network: Toward Real−World Noise Removal and Noise Generation," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2020, vol. 12355 LNCS. doi: 10.1007/978−3−030−58607−2_3.

[59] K. C. Chang *et al.*, "Learning Camera−Aware Noise Models," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2020, vol. 12369 LNCS. doi: 10.1007/978−3−030−58586−0_21.

[60] J. Byun, S. Cha, and T. Moon, "FBI−Denoiser: Fast Blind Image Denoiser for Poisson−Gaussian Noise," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2021. doi:

10.1109/CVPR46437.2021.00571.

[61]   P. Cizmar, A. E. Vladár, B. Ming, and M. T. Postek, "Simulated SEM images for resolution measurement," *Scanning*, vol. 30, no. 5, 2008, doi: 10.1002/sca.20120.

[62]   F. Timischl, M. Date, and S. Nemoto, "A statistical model of signal–noise in scanning electron microscopy," *Scanning*, vol. 34, no. 3, 2012, doi: 10.1002/sca.20282.

[63]   J. Byun and T. Moon, "Learning Blind Pixelwise Affine Image Denoiser with Single Noisy Images," *IEEE Signal Process Lett*, vol. 27, 2020, doi: 10.1109/LSP.2020.3002652.

[64]   G. Jang, W. Lee, S. Son, and K. Lee, "C2N: Practical Generative Noise Modeling for Real–World Denoising," in *Proceedings of the IEEE International Conference on Computer Vision*, 2021. doi: 10.1109/ICCV48922.2021.00235.

Abstract in Korean

반도체 공정의 품질 관리를 위해서는 단계 별 정확한 계측 및 검사가 필수적이다. 이러한 과정들은 반도체 패턴의 SEM 이미지를 분석하여 진행되는데, 이미지에 내재되어 있는 노이즈로 인해서 정밀도가 하락하는 문제가 존재한다. 노이즈를 제거하기 위한 다양한 방법 중 딥러닝 기반의 방법들이 가장 좋은 성능을 보이고 있는데, 산업에서는 노이즈가 없는 깨끗한 정답 이미지가 없기 때문에 적용할 수 없다는 한계점이 존재한다. 따라서 정답 이미지 없이 딥러닝 학습을 진행하는 자기지도학습 기반 방식들이 고안되었으나 SEM 이미지에서는 좋은 성능을 보이지 못한다. 따라서 본 논문에서는 먼저 기존의 자기지도학습 기반의 방법들이 실패한 원인을 SEM 노이즈의 입자크기를 통해 설명하고 SEM 이미지의 노이즈를 성공적으로 제거할 수 있는 'Patch-based Noisier2Noise'를 제안한다. 또한 이 방법이 갖는 한계점인 효율성을 개선한 'Modeling-based Noisier2Noise'를 고안하고 'Iterative training'을 접목하여 성능 또한 끌어올렸다. 제안한 방식은 전통적인 방식에 비해서 최대 50배 단축된 시간을 보여주면서 대등한 PSNR과 개선된 LER, LWR 수치를 보여준다. 또한 이미지의 시각적 품질과 2D 단면 윤곽선 측면에서 기존의 방식들에 비해 이미지 구조를 파괴하지 않으면서 더 효과적으로 노이즈를 제거함을 확인할 수 있다.