A Flexible Deep Learning Based Approach for SEM Image Denoising

Jun Chen*, Xinheng Jiang, Keisuke Goto, Takashi Tsutsumi, Yasutaka Toyoda
Hitachi High-tech Corp., Harumi Triton Square Office Tower X 31F 1-8-10, Harumi, Chuo-ku, Tokyo, 104-6031, Japan

ABSTRACT

In the field of semiconductor manufacturing, Scanning Electron Microscope (SEM) is employed for critical dimension (CD) measurements, overlay measurements, and defect inspections to ensure the quality and reliability of semiconductor devices. Nevertheless, SEM images inherently carry a significant level of noise, leading to inaccurate metrology and false defect inspections. Therefore, it is crucial to develop denoising techniques. One widely used method is frame averaging, which reduces cumulative noise by averaging multiple scans. While increasing scan times enhances SEM image quality, it comes with drawbacks such as surface charging, pattern shrinkage, and reduced throughput. Deep learning (DL) techniques, including supervised and unsupervised approaches, have shown remarkable progress in the field of SEM image denoising. However, supervised methods are notably affected by phenomena such as pattern shrinkage and surface charging, which occur during the capture of reference images. On the other hand, unsupervised methods are typically more effective with lower noise levels. In this paper, we introduced a flexible DL method for denoising SEM images that operates without the requirement for paired data. To demonstrate its effectiveness, we analyzed and evaluated its performance in two metrology tasks. Experimental results validated the efficacy of our method in reducing noise, demonstrating its applicability to both ADI and AEI.

Keywords: Image denoising, deep learning, metrology, critical dimension, SEM

1. INTRODUCTION

Various types of Scanning Electron Microscopes (SEMs) are used in semiconductor fabrication plant to ensure semiconductor device quality and reliability during their production lifecycle. Critical Dimension (CD) SEM is specifically designed for high-precision metrology at the nanometer scale. It is employed to measure the dimensions of features on semiconductor wafers, such as linewidths or space widths, to ensure that the fabricated features satisfy the desired specifications. Review SEM is utilized for inspection purposes, focusing on examining the overall quality of semiconductor devices and identifying defects. It captures high-resolution images for comprehensive visual inspection, defect analysis, and general assessment of the manufacturing process. As the circuit pattern becomes smaller and smaller, the number of measurement and inspection points for monitoring manufacturing processes significantly increases. Consequently, engineers are motivated to enhance the throughput and accuracy of metrology and inspection. However, SEM images inherently contain a significant amount of noise, leading to inaccurate metrology and false defect inspections.

SEM image denoising aims to enhance the quality of a noisy SEM image by reducing noise while preserving essential features. Traditional approaches, such as frame averaging or BM3D [1], operate under the assumption that the noise in an SEM image is either additive or that the noise distribution remains consistent across similar circuit patterns. Deep learning (DL) techniques, including both supervised [2-4] and unsupervised approaches [5-8], train denoisers to minimize the differences between convolutional neural network (CNN) outputs and reference images. It should be noted that although an explicit reference image is absent in unsupervised methods, they typically synthesize a reference internally based on certain assumptions. Recent reports have indicated that supervised DL techniques outperform traditional and unsupervised DL approaches in terms of execution speed and precision. However, challenges such as reducing surface charging or pattern shrinkage during reference image capture remain. In this paper, we present a flexible DL approach designed to denoise SEM images. Compared to supervised DL methods, it does not require paired low-quality and clean reference images, thus ensuring robustness against issues such as surface charging and pattern shrinkage. In contrast to unsupervised DL methods, it exhibits a more effective noise reduction performance owing to the involvement of unpaired clean references.

Metrology, Inspection, and Process Control XXXVIII, edited by Matthew J. Sendelbach, Nivea G. Schuch, Proc. of SPIE Vol. 12955, 129550B · © 2024 SPIE
0277-786X · doi: 10.1117/12.3011135

The remainder of this paper is organized as follows: Section 2 reviews related work, while Section 3 outlines the conceptual framework of our proposed method. Section 4 verifies our hypotheses and presents quantitative evaluation results for two metrology tasks.

2. RELATED WORKS

2.1.1 Traditional approaches

Frame averaging serves as the primary noise reduction method, working on the assumption that the noise in a SEM image is both additive and random. It reduces random fluctuations above and below the actual intensity by calculating the arithmetic mean of the intensity values for each pixel position. Theoretically, noise magnitude decreases with the square root of the number of samples for averaging, making it advisable to increase the number of scans to synthesize a high-quality SEM image. As the number of scans for averaging increased, there was a significant reduction in the noise evident in the images. Figure 1 illustrates a comparison of the SEM images averaged over 1, 4 and 64 scans to demonstrate the effectiveness of frame averaging. However, while enhancing image quality, increasing scan times extends acquisition time, slowing SEM throughput, and raising the risk of pattern shrinkage. BM3D [1] is another traditional method used to reduce additive Gaussian noise. It operates by segmenting the image into overlapping blocks and identifying similar blocks to implement collaborative filtering. Subsequently, adaptive noise reduction is performed using Wiener filtering. Despite its effectiveness, BM3D is sensitive to parameters and tends to blur the pattern boundaries.

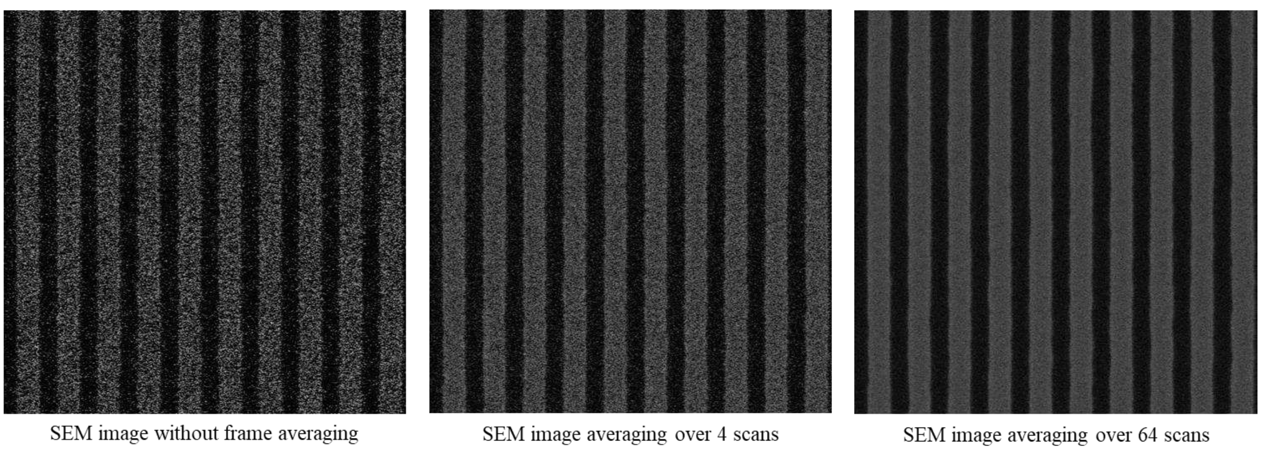

Figure 1. Comparison among SEM images averaged over 1, 4 and 64 scans.

2.1.2 Supervised deep learning approaches

Supervised DL denoising methods generally involves two steps. Initially, a CNN model [2-4] is trained using pairs of low- and high-quality SEM images. Subsequently, a well-trained model is used to convert an input low-quality SEM image into a clean image. When the high-quality counterpart is ideal, the model can effectively estimate the noise distribution during training and reduce the noise during inference. However, obtaining ideal images is challenging in most scenarios, as they are typically captured by averaging dozens or hundreds of scans, introducing issues such as surface charging and pattern shrinkage. Figure 2 shows an example of pattern shrinkage, where the linewidth significantly decreased as the number of scans increased. A model trained on such images learns both noise reduction and pattern shrinkage, consequently altering the features of the input image, particularly the linewidth.

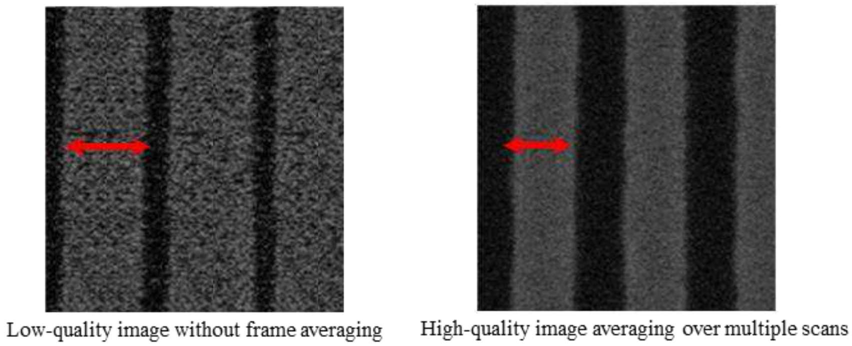

Figure 2. Example of pattern shrinkage.

2.1.3 Unsupervised deep learning approaches

Noise2Void (N2V) [5-6] improves the quality of low-quality SEM images without the need for clean reference images. It operates in a self-supervised manner, learning the underlying structure and characteristics of noise through statistical analysis of pixel intensity within a neighborhood. This approach is particularly useful for tasks, such as microscopic image analysis, in which obtaining clean reference images is impractical. However, its performance is significantly influenced by the noise level due to the utilization of neighborhood data. If noise level exceeds a certain threshold, the inferred images may not show a significant difference compared to the input images. Noise2Noise (N2N) [7-8] is another type of unsupervised DL method that trains a DL model to learn mapping from noisy images to clean images using pairs of noisy images, without explicitly requiring clean images. It capitalizes on the inherent similarity among all noisy images to estimate noise distribution and perform noise reduction. However, its effectiveness may be affected by the inconsistency of noise in the training images. Similar to supervised DL approaches, capturing redundant noisy images from the same scene proves challenging in most scenarios.

3. ALGORITHM

Figure 3 provides an overview of our proposed algorithm, comprising three phases: (1) pattern-structure-aware noise simulation and image generation, (2) model training, and (3) inference. Phases (1) and (2) operate offline on the server-side, whereas inference can be performed in real-time on the SEM to improve the throughput of metrology and inspection. In our setup, a high-quality SEM image averaged over dozens or hundreds of scans, served as a clean reference. A low-quality SEM image is used for noise reduction. Importantly, the proposed algorithm does not depend on paired low- and high-quality SEM images, allowing them to be captured from different wafers or dies.

3.1.1 Pattern-structure-aware noise simulation and image generation

While a pre-designed additive noise algorithm [8] generates noisy images by applying Gaussian noise to a clean image, assuming a consistent noise distribution across the entire image, our analysis reveals variations in noise distribution among circuit patterns. Consequently, the application of global Gaussian noise to the entire image may distort the pattern boundaries and reduce the accuracy of metrology and inspection. To address this issue, we introduce a pattern-structure-aware noise simulation module. This module uses a low-quality image as a reference and degrades the input high-quality SEM image to match the noise level of the reference considering the circuit structures.

3.1.2 Model training

In this phase, we employ supervised learning to train a CNN denoiser that learns the mapping from a simulated low-quality image to its original high-quality version. As depicted in Figure 3(2) and (3), differences in the linewidths between the training and inference images may exist. If these variations in the inference images are not sufficiently covered by the training dataset, the prediction may become unreliable. To address this concern, we employ data augmentation techniques, such as zooming in, zooming out, flipping, and rotation, to enhance the diversity and quantity of the training images. This approach improves generalization and robustness of the denoiser.

3.1.3 Inference

In the inference phase, a well-trained denoiser is applied to make predictions for new, unseen, and low-quality SEM images. In our setup, we utilized Open Visual Inference and Neural Network Optimization (OpenVINO [9]) to accelerate the inference process.

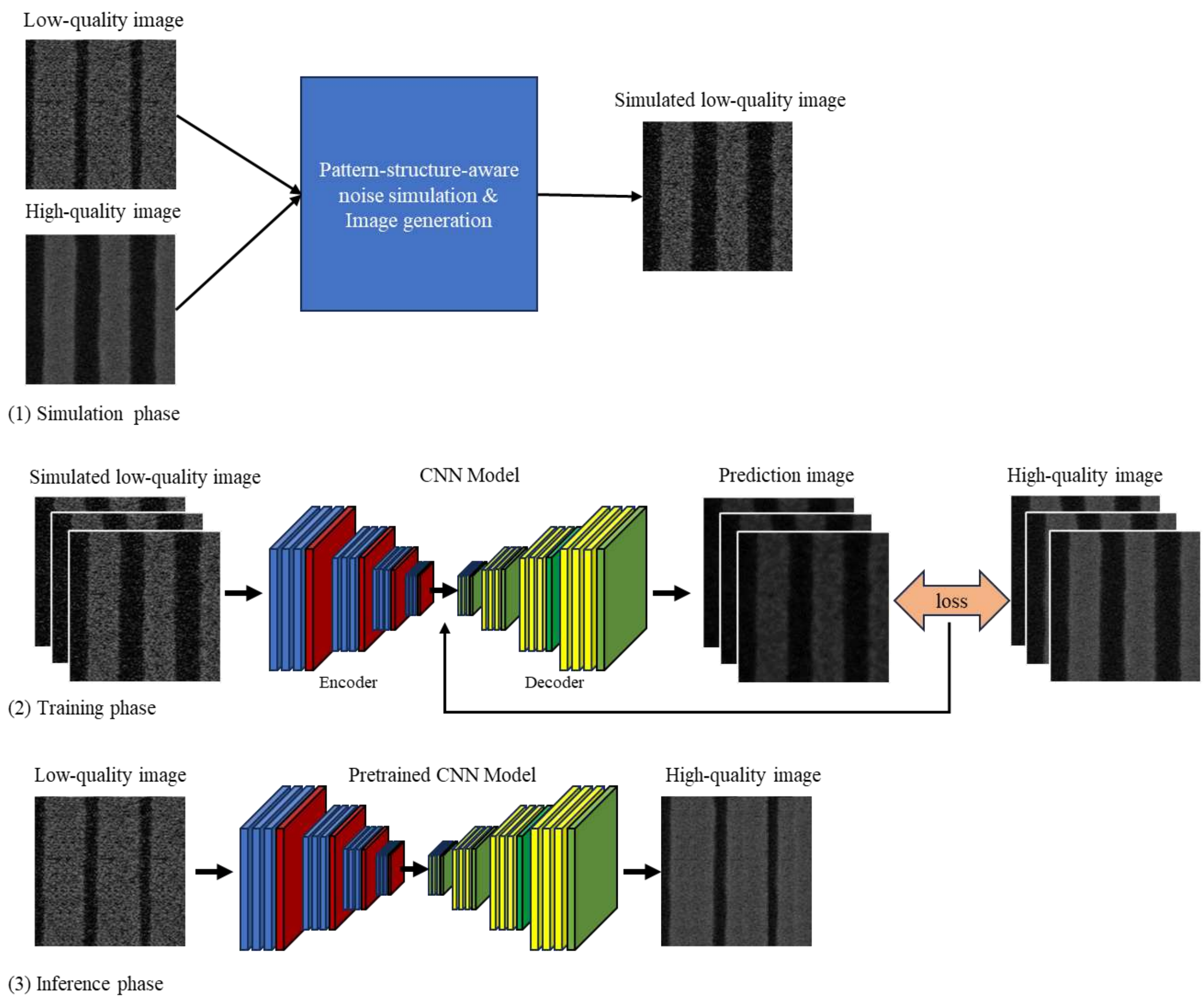

Figure 3. Overview of the proposed method.

4. EXPERIMENTAL RESULTS

In this section, we initially confirm the variation in noise distribution across the different circuit pattern. We then compare the performance of the proposed method with the following two latest approaches in two metrology tasks to demonstrate its effectiveness.

- N2C [2] as a recently developed supervised DL image denoising method that utilizes pairs of noisy and clean images for model training.

- N2V [5] is an unsupervised DL image denoising method trained using only noisy images.

4.1.1 Verification of noise distributions

We utilized a pair of low- and high-quality line/space pattern SEM images to demonstrate the impact of pattern structures on noise levels. The line and space regions, highlighted by red and yellow boxes in Figure 4 (a), were manually extracted from the SEM images. Assuming that the high-quality image represents the ground truth and that the noise follows a Gaussian distribution, we computed the respective noise distributions and fitted them to Gaussian curves, as shown in Figure 4 (b) and (c). The calculated standard deviations of the noise, being 9.1 and 3.1 respectively, indicate that the noise varies with the circuit pattern.

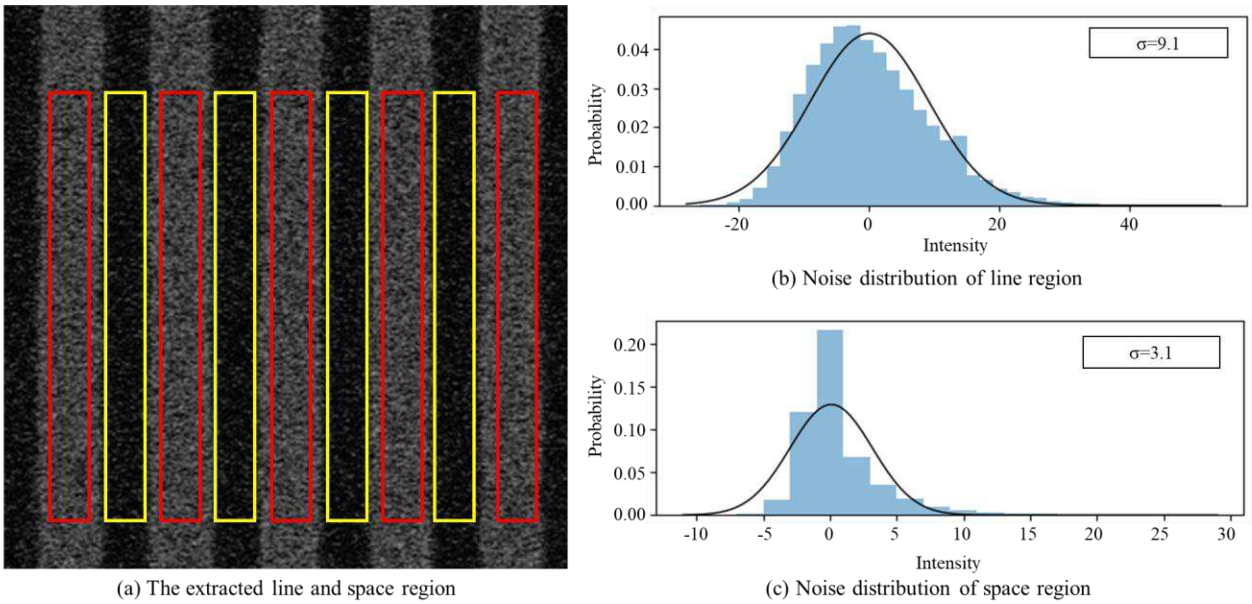

Figure 4. Extracted line and space regions and their noise distributions.

4.1.2 Metrological evaluation

In this experiment, we evaluated the proposed, N2C, and N2V methods using two datasets detailed Table 1. For the first dataset, we synthesized high-quality images by scanning 127 measurement points 64 times, where the initial scans were considered as low-quality images. We allocated 80% of the dataset for model training and the remaining 20% for testing. The second dataset consisted of 129 pairs of high- and low-quality SEM images. The high-quality images, acting as the reference for both model training and quantitative assessment, were the averages of 64 scans, while the low-quality images were synthesized by averaging over 4 scans. From this dataset, we randomly picked up 90 pairs to train the DL models, while the remaining images were reserved for assessing measurements accuracy.

Because the low- and high-quality images were captured from the same positions, they could be directly used to train the N2C model without additional processing. As mentioned previously, the N2V model was trained using only low-frame images, whereas our method utilizes both simulated and original high-quality images for training.

Table 1. Details of evaluation datasets.

	Pattern	Low quality (scan times)	High quality (scan times)	Number of training images (frames)	Number of test images (frames)	Capturing conditions
ADI EUV resist with shrinkage	Line/space Pitch=45 nm	1	64	105	22	Voltage: 800v Current: 8.0pA Magnification: 150k
AEI Non-shrinkage	Line/space Pitch=90 nm	4	64	90	39	Voltage: 800v Current: 8.0pA Magnification: 100k

4.1.3 Denoising for the dataset with pattern shrinkage (ADI)

Figure 5(a) and (b) give examples of low- and high-quality images from the dataset, revealing noticeable pattern shrinkage between the two images. Figure 5(c) presents an example of a simulated low-quality image. The predictions of the three models are shown in Figure 5(d)-(f). Both the proposed method and N2C effectively reduced most noise, whereas N2V did not exhibit effective noise reduction.

Because of the presence of noise, the linewidth computed from low-quality images lacks sufficient accuracy, which makes unsuitable for monitoring the manufacturing process. However, our observations revealed that most of the extracted edges were correct, suggesting that the values can still provide a reasonable estimation of the actual physical width. As depicted by the blue and orange curves in Figure 6, the approximated average linewidths of low-quality and high-quality images are 27.62 nm and 24.71 nm, respectively. This result is consistent with the pattern shrinkage phenomenon, where the pattern width tends to decrease with increased in scanning time. N2C utilizes high-quality SEM images as references during training, enabling it to learn not only noise reduction, but also pattern shrinkage. Consequently, it generated clear images with an average linewidth of 24.87 nm, as shown in Figure 5(d) and orange curve in Figure 6. N2V effectively maintained the linewidth, averaging at 27.95 nm, as indicated by the yellow curve in Figure 6. However, the visual quality of the prediction remained similar to that of the input image, as illustrated in Figure 5(e), indicating that N2V failed to enhance the signal-to-noise ratio (SNR). The result of our proposed method, depicted as a red curve in Figure 6, show an average linewidth of approximately 27.49 nm, closely aligns with that of the low-quality SEM images. Additionally, as shown in Figure 5(f), a significant noise reduction was observed. This experiment confirmed that the proposed method is suitable for after development inspection (ADI).

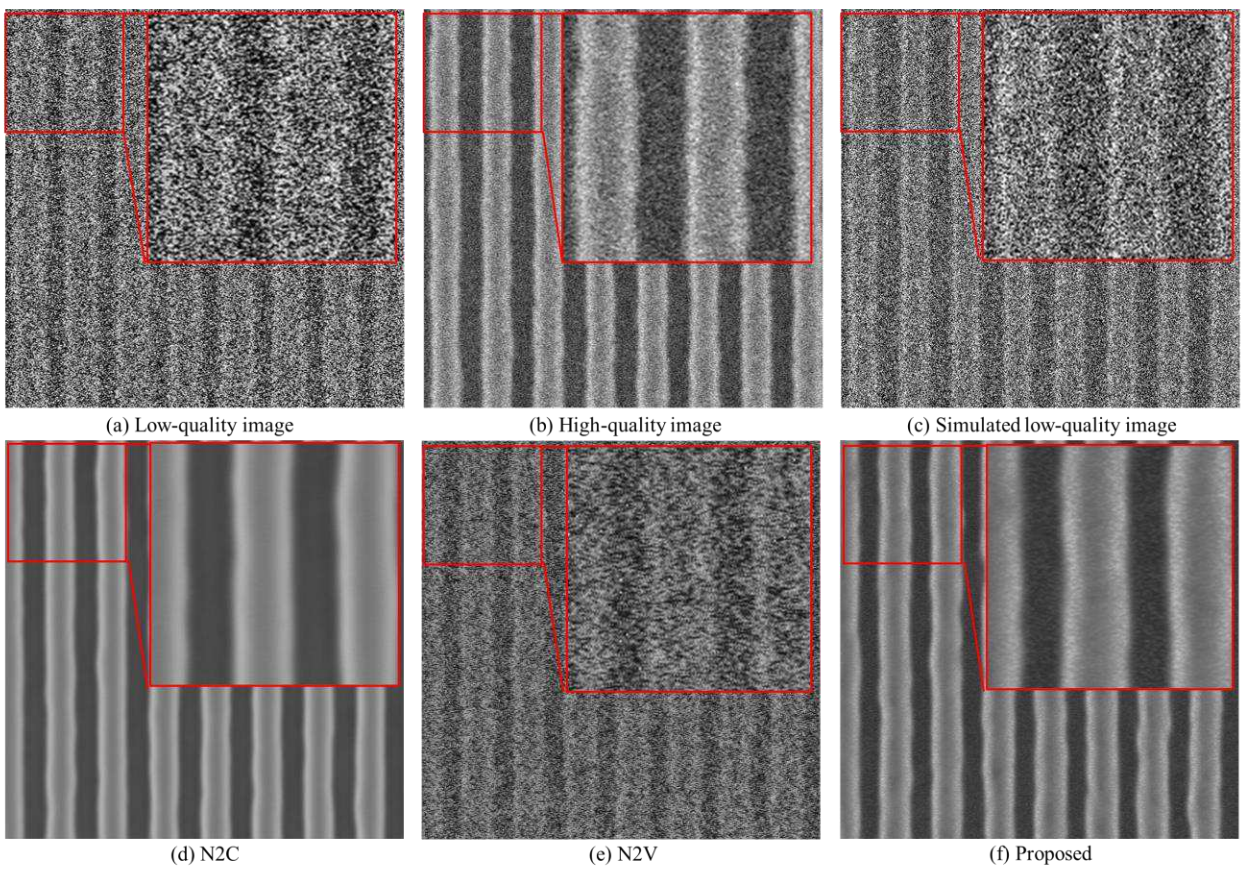

Figure 5. Examples of samples with pattern shrinkage, simulated low-quality images, and predicted results of the three methods.

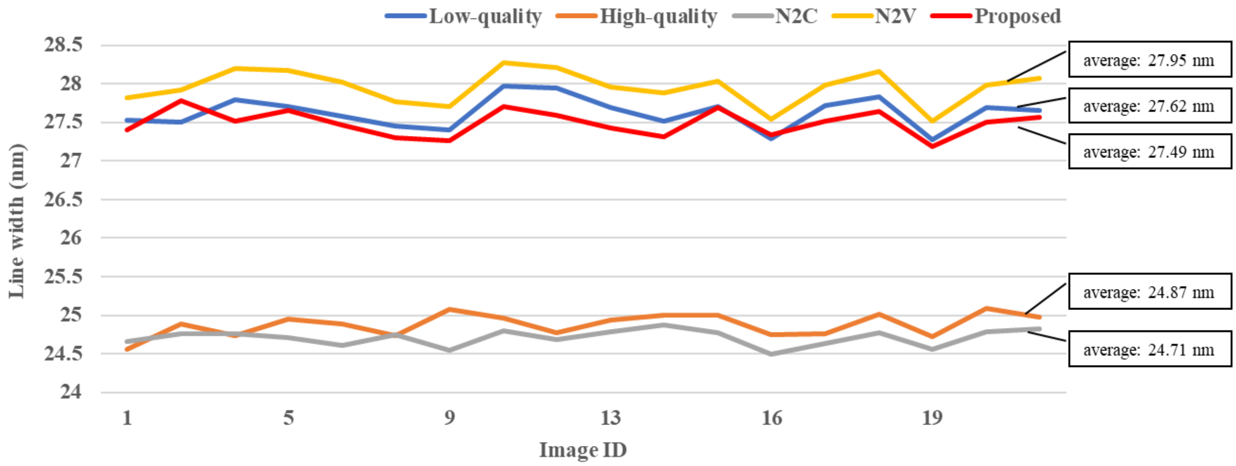

Figure 6. Comparison of linewidths of low-quality images, high-quality images, and predictions of the three methods.

4.1.4 Denoising for the dataset without pattern shrinkage

Examples of low-quality, high-quality, and simulated images are shown in Figure 7(a)-(c), respectively. The predictions of the three methods are presented in Figure 7(d)-(f). Since the noise level in the low-quality image was relatively moderate, it is evident that all three methods effectively reduced most of the noise.

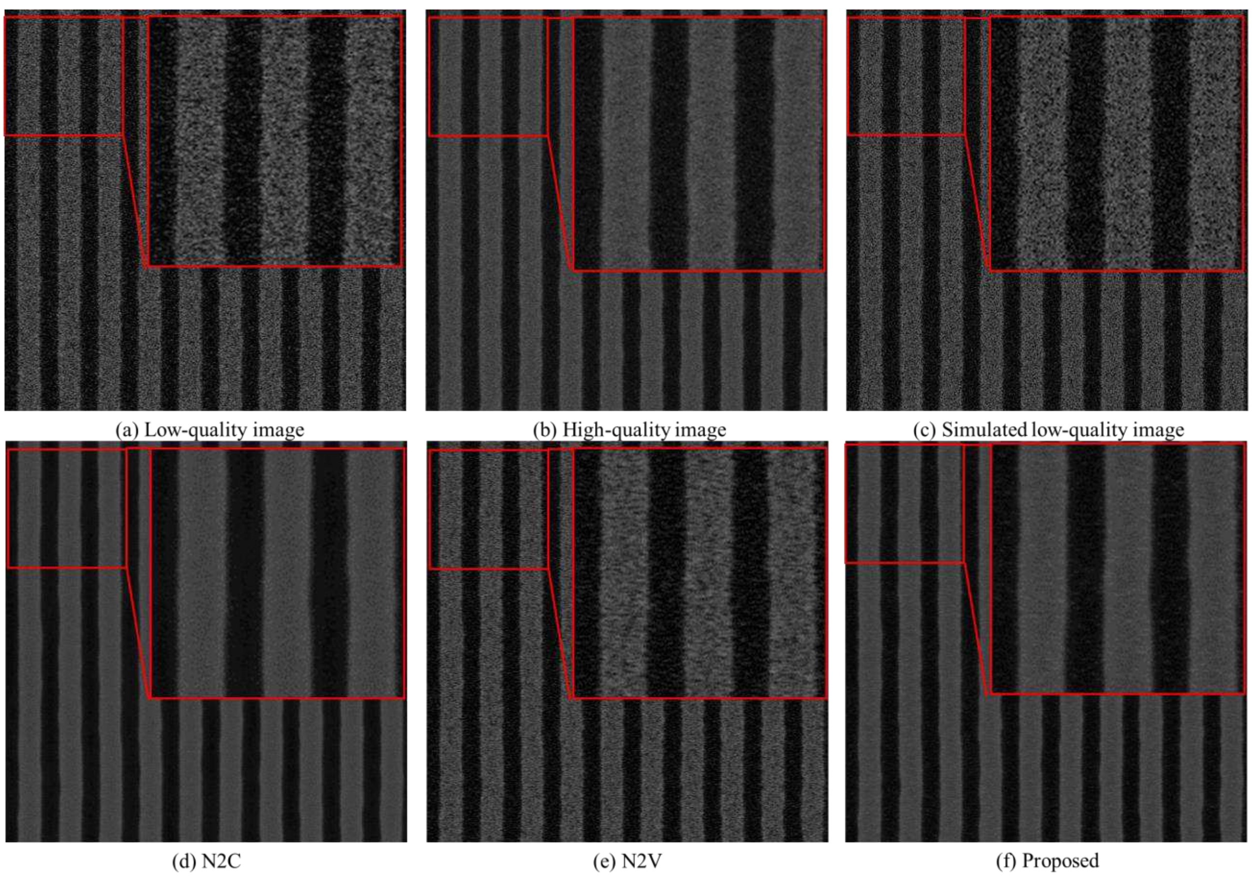

Figure 7. Examples of samples without pattern shrinkage, simulated low-quality images, and predicted results of the three methods.

Statistical measures, including slope, R-squared (R^2), and skewness (skew), were used to assess how closely the linewidths of the predictions aligned with those of the reference image. A higher slope and R^2 (closer to 1) and a lower skew (closer to 0) indicate a better prediction.

As shown in Fig. 8, the proposed method outperforms N2V across all three metrics. Especially, there is a significant improvement in R^2 and skew, indicating that our method can produce more accurate predictions than the N2V method. In comparison with N2C, our method exhibited similar performance across all metrics. This demonstrates that our method, operating under more flexible conditions, can achieve a performance comparable to that of N2C. This experiment proved that the proposed method can also be extended to after etching inspection (AEI).

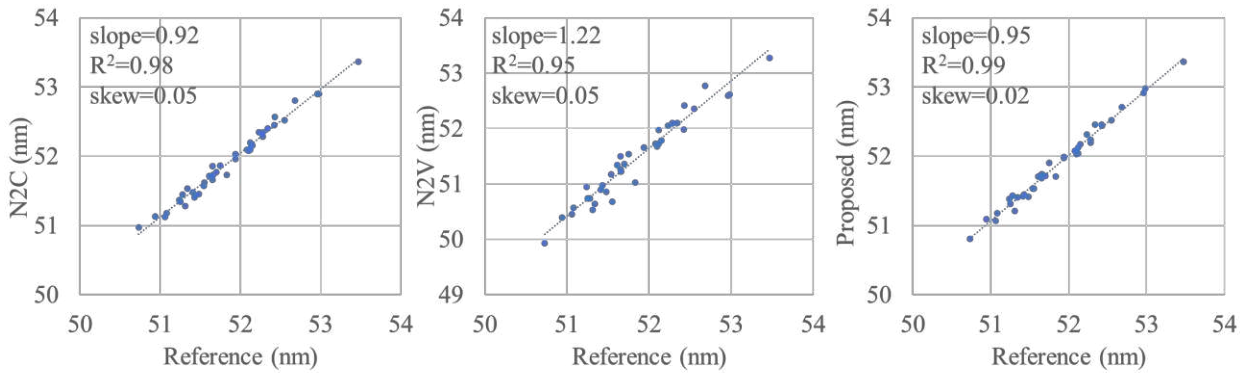

Figure 8. Quantitative evaluation results of N2C, N2V and proposed method.

5. CONCLUSIONS

In this study, we introduce a novel DL-based SEM image denoising method designed for scenario where capturing an ideal clean reference is impractical. In the proposed method, we first take the low-quality image as a reference and simulate a noisy image from a high-quality image using a pattern-structure-aware noise simulation module. Subsequently, a CNN denoiser is trained using these simulated images and the original high-quality images. Experimental results across two CD metrology tasks validate the efficacy of our method in reducing noise while preserving critical features such as linewidth. In the future, we aim to assess its efficiency using various SEM images. Furthermore, we plan to reduce the burden of noise simulations and model training to improve the usability.

REFERENCES

[1] Danielyan A, Katkovnik V, Egiazarian K. BM3D frames and variational image deblurring. IEEE Transactions on image processing, 2011, 21(4): 1715-1728.

[2] T. Okuda, J. Chen, T. Motoyoshi, et al., Comparison between supervised and self-supervised deep learning for SEM image denoising. Metrology, Inspection, and Process Control XXXVII. SPIE, 2023, 12496: 916-924.

[3] Zhang K, Zuo W, Chen Y, et al. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. IEEE transactions on image processing, 2017, 26(7): 3142-3155.

[4] J. Liang, J. Cao, G. Sun, Swinir: Image restoration using swin transformer. Proceedings of the IEEE/CVF international conference on computer vision. 2021: 1833-1844.

[5] A. Krull, T. O. Buchholz, F. Jug, Noise2void-learning denoising from single noisy images. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019: 2129-2137.

[6] B. Dey, S. Halder, K. Khalil, et al., SEM image denoising with unsupervised machine learning for better defect inspection and metrology. Metrology, Inspection, and Process Control for Semiconductor Manufacturing XXXV. SPIE, 2021, 11611: 245-254.

[7] Lehtinen J, Munkberg J, Hasselgren J, et al. Noise2Noise: Learning Image Restoration without Clean Data. International Conference on Machine Learning. PMLR, 2018: 2965-2974.

[8] Lei H, Teh C, Yu L, et al. Denoising sample-limited SEM images without clean data. Metrology, Inspection, and Process Control for Semiconductor Manufacturing XXXV. SPIE, 2021, 11611: 289-299.

[9] OpenVINO Deep Learning Deployment Toolkit, 2019. https://github.com/openvinotoolkit/openvino