SEMNet: Deep Synthetic Training for Unprecedented SEM Image Denoising and Super-Resolution

Junghun Cha[1], Ph.D Doohyun Cho[1], Ph.D SeungJae Lee[1]
[1]D.notitia Inc., Seoul, Republic of Korea

The Scanning Electron Microscope (SEM), an essential tool in electron microscopy, excels in delivering detailed visuals by directing an electron beam over a specimen's surface. Its widespread use in material science, biology, and various industrial fields stems from its ability to provide high-resolution, in-depth focus, and three-dimensional imagery.

Despite these advantages, acquiring high-quality SEM images can be challenging. This is due to inherent resolution limitations and noise factors that impair image clarity. Addressing these issues, our study introduces a method to enhance SEM image quality through advanced deep-learning algorithms, focusing on both denoising and super-resolution (SR) to refine image details and clarity effectively.

We introduce SEMNet, a Convolutional Neural Network (CNN) specifically designed for SEM image restoration with an upscale factor of 2. Inspired by the successes of the RRDBNet [1] in real-world image restoration, our model innovatively applies this architecture to SEM imagery, overcoming typical denoising and SR challenges. Fig. 1 illustrates our SEMNet.

SEMNet's core consists of 6 Residual-in-Residual Dense Blocks (RRDBs), a sophisticated structure pivotal for meticulously refining image details. Within each RRDB, we embed 3 Residual Dense Blocks (RDBs) characterized by their densely connected convolutional layers (denoted as Conv in Fig. 1). This design ensures a seamless flow of data and gradients throughout the network, significantly boosting feature extraction performance. The RDBs each consist of 5 convolutional layers paired with LeakyReLU activation functions (denoted as LeakyReLU in Fig. 1), having a negative slope coefficient of 0.2. Overall, SEMNet comprises 94 convolutional layers and 19 LeakyReLU activation functions, considering the additional convolutional layers outside the RRDBs as well. The network concludes with an up-sampling convolutional layer (denoted as Upsample in Fig. 1) that utilizes nearest neighbor interpolation to achieve SR, enhancing the quality and resolution of SEM images.

To effectively tackle the challenges posed by diverse noise types and low-resolution issues in SEM images, we devise a degradation modeling strategy specialized to producing synthesized low-quality (LQ) images from high-quality (HQ) counterparts. Specifically, our methodology concentrates on

simulating two primary noise types frequently observed in SEM images: Gaussian noise, with a mean of 0 and a standard deviation ranging from 2 to 25, and Poisson noise, with a lambda value (λ) between 10^2 and 10^4. Moreover, we integrated additional degradations into our datasets, such as Gaussian blurring effects and up-downscaling processes. This comprehensive approach is designed to bolster the robustness of our SEM image restoration model by broadening its exposure to a variety of noise and other degradation scenarios. Note that it serves as a form of data augmentation, dedicately aimed at enhancing image restoration performance.

We trained the model for 5.44×10^5 iterations using Adam optimizer, with a batch size of 4. The learning rate starts at 0.0001 and is halved when the iteration reaches a multiple of 2.0×10^5. The training utilizes the following loss functions with respective weights: L1 loss (1.0), Histogram loss [2] (0.1), and Total Variation loss (0.01). The L1 loss effectively minimizes absolute pixel differences, playing a crucial role in image restoration. Histogram loss targets the reduction of pixel distribution disparities, ensuring the restored image faithfully mirrors the statistical characteristics of the original. Total Variation loss focuses on diminishing noise and maintaining fine details, like edges, enhancing the overall quality of restored images. Our model is trained with 1300 real SEM images and 1300 images sampled from Flickr2K [3], which is the public dataset that has high-resolution real-world images from multiple categories such as nature, cityscape, human, animal, etc.

To quantitatively evaluate the performance of SEM image SR, we employ two metrics: Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM). PSNR measures the pixel-level accuracy between the reconstructed and the original images, while SSIM assesses the perceptual quality by considering changes in luminance, contrast, and structural similarity. For both metrics, higher values indicate better performance. A testing set of 25 SEM images was used to calculate both metrics.

Compared to the classical restoration using bicubic interpolation (for SR) and Gaussian blurring (for denoising), SEMNet achieved an increase of 0.212dB in average PSNR and 0.1379 in average SSIM. These results suggest that deep learning-based image restoration methods surpass classical image processing techniques in enhancing the quality of SEM images. Fig.2 shows the qualitative results of classical restoration and the SEMNet. The original sizes of the LQ and HQ images are 640×480 and 1280×960, respectively. Our SEMNet demonstrates superior visual performance compared to the classical restoration method.

In conclusion, we introduce SEMNet, a new deep-learning model for restoring LQ SEM images to HQ ones. With a diverse training dataset, SEMNet effectively removes multiple types of noises and upscales the image

resolution through detailed degradation modeling and specialized loss functions. It surpasses classical restoration methods in improving SEM image quality, both quantitatively and qualitatively. Our work will contribute to the field of microscopy image restoration by introducing innovative deep-learning architecture designs and training strategies, specifically tailored to handle highly complex and diverse degradations in SEM images.

Graphic:

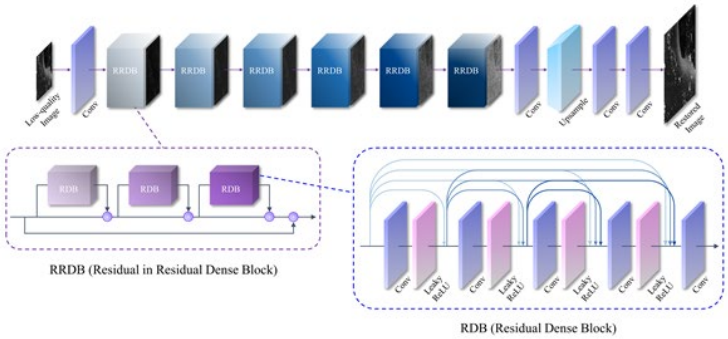

Fig 1. Overview of our SEMNet. It is consisted of 6 Residual-in-Residual Dense Blocks (RRDBs). Within each RRDB, we embed 3 Residual Dense Blocks (RDBs) characterized by their densely connected convolutional layers (denoted as Conv). The RDBs each consist of 5 convolutional layers paired with LeakyReLU activation functions (denoted as LeakyReLU). Overall, SEMNet comprises 94 convolutional layers and 19 LeakyReLU activation functions, considering the additional convolutional layers outside the RRDBs as well. Also, the network concludes with an up-sampling convolutional layer.

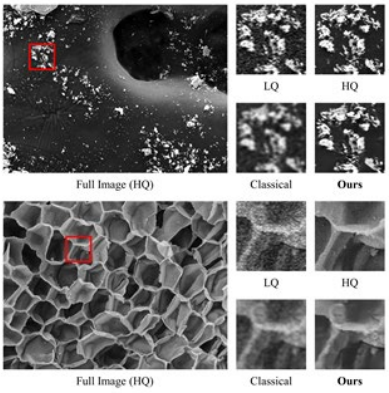

Fig 2. Qualitative results of our proposed methods. Our SEMNet demonstrates superior visual performance compared to the classical restoration method.

Keywords:

SEM, Image Restoration, Deep Learning

Reference:

[1] Wang et al., "Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data", ECCVW, 2018
[2] Afifi et al., "HistoGAN: Controlling Colors of GAN-Generated and Real Images via Color Histograms", CVPR, 2021
[3] Lim et al., "Enhanced Deep Residual Networks for Single Image Super-Resolution", CVPRW, 2017