# Robust Deep-learning Based Autofocus Score Prediction for Scanning Electron Microscope

Hyun Jong Yang[1,2], Moohyun Oh[2], Jonggyu Jang[1], Hyeonsu Lyu[1] and Junhee Lee[3]

[3]Ulsan National Institute of Science and Technology, Ulsan, Ulsan-gwangyoksi, Republic of Korea, [2]EgoVid Inc., Ulsan, Ulsan-gwangyoksi, Republic of Korea, [4]Coxem, Daejeon, Taejon-jikhalsi, Republic of Korea

**Introduction:** Deep-learning has come into its own in analyzing scanning electron microscope (SEM) images [1]. As the image analysis part is becoming autonomous with the aid of deep-learning, the main challenge in the use of SEMs is to get high quality images by deftly controlling SEM parameters such as brightness, contrast, focus, and etc. While the other studies attempt to use deep-learning as a classifier or analyzer for images, our previous work [2] has shown that deep-learning can be potentially used as a controller of SEM parameters. To automate the SEM control, the most crucial part is to accurately evaluate the quality of input SEM images as if SEM experts do, because existing mathematical autofocus (AF) metrics cannot capture the image quality by scrutinizing both regional features and the entire image for a variety of types of samples. The deep-learning based algorithm proposed in [2] has shown to mimic SEM experts better in judging the quality of SEM images than the existing AF algorithms [3], [4], [5], which ultimately results in a full automation of SEM control.

**Contribution Summary:** In fact, the quality of SEM images becomes more susceptible to the focus control under higher magnification. The previous work [2] does not consider this fact in designing the deep-learning algorithm, and hence suffers from performance degradation in the presence of magnification variation.

Therefore, we design a new deep neural network (DNN) architecture and data collection criteria to cope with all possible control parameters such as brightness, contrast, focus, and *magnification*. The proposed scheme shows more robust performance than the previous work [2] for a variety of magnification setups.

**New Labeling and Dataset Construction:** The dataset used in [2] is comprised of SEM images with labels of integer scores ranging from 0 to 7. The labeling of the dataset was done by several highly-qualified SEM experts. However, this dataset includes different amounts of images for each magnification, which may result in inconsistent deep-learning performance for different magnification settings. In the new dataset construction, we collect the same number of images for each of the five magnification settings: (500, 1000, 2000, 5000, and 10000). For labeling, images are given 1-point when the sample begins to be recognizable from an unrecognizable initial setting. The images perfectly focused are scored 10-point. In addition, we add noise, white, and black images as 0-point images in the new dataset. Then, the new dataset is labeled by the range of [0, 10], which results in finer granularity in scoring than in [2].

**Deep Neural Network Design:** In our new DNN, magnification is also treated as an important input along with SEM images. The proposed neural network model is composed of ResNet18 and three fully connected neural network (FCN) layers with the rectified linear unit (ReLU) activation function. As shown in Figure 1, the input image passes through the ResNet18 layer, and then the output is concatenated with the magnification of the image. This concatenated tensor passes through the FCN layers to predict the score of the input image. To avoid biased result and unstable weight update of the DNN, the magnification values of (500, 1000, 2000, 5000, and 10000) are normalized into the range of [0, 1].

**Experiment Settings:** The experimental setups are the same as those in [2]. That is, tinball and grid samples are used in both training and test, and the score prediction error is measured by root mean square error (RMSE). The stochastic gradient descent (SGD) optimizer is used for updating our DNN with the following hyper parameters: maximum epoch of 100, batch size of 1, learning rate of 2e-4, and gradient momentum of 0.9. Our new dataset consists of 30,272 images in total, amongst which 27,235 images are used for training and 3,037 images are used to evaluate the final model in test. The images are augmented by 320x240 sized random cropping and vertical/horizontal flipping. In the test, only 320x240 sized center-cropping is applied to input images. Experiments are implemented under the following environment: NVIDIA RTX 2080Ti, CUDA 10.0, Python 3.6.8, Pytorch 1.1.0, and OpenCV 4.1.0.

**Experiment Results:** We first evaluate the RMSE performance of the proposed DNN in comparison to ResNet50, which is the known best scheme in [2]. Furthermore, to assess the impact of the newly constructed dataset, ResNet50 and the proposed DNN are evaluated with both the old and new datasets. Table 1 shows the RMSE results of the evaluated schemes, which are averaged over different magnification values ranging from 500 to 10,000. As seen from the results, the new dataset brings sizable performance gain for both ResNet50 and the proposed DNN. Moreover, the proposed DNN architecture delivers further RMSE reduction (0.5732 to 0.4548) in case of the new dataset.

Table 2 shows the RMSE performance of the existing AF algorithm [3], ResNet50 [2] with the new dataset, and the proposed DNN with the new dataset for different magnification setups. Although both ResNet50 and the proposed DNN suffer from performance degradation as the magnification increases, the proposed DNN relatively works well even in high magnification.

|  | ResNet18 w/ new dataset | ResNet18 w/ old dataset | ResNet50 w/ new dataset | ResNet50 w/ old dataset |
|---|---|---|---|---|
| RMSE | 0.4548 | 0.7132 | 0.5732 | 0.7407 |

**Table 1**. RMSE of ResNet50 [2] and the proposed DNN with the old and new datasets. The RMSE is averaged over different magnification values ranging from 500 to 10,000.

| Magnification | Absolute Variance | ResNet 50 w/ new dataset | Proposed DNN |
|---|---|---|---|
| 500 | 3.0467 | 0.3791 | 0.3056 |
| 1,000 | 2.9133 | 0.3602 | 0.3152 |

| 2,000  | 2.7270 | 0.5837 | 0.3152 |
|--------|--------|--------|--------|
| 5000   | 2.7924 | 0.5402 | 0.4174 |
| 10,000 | 2.7741 | 0.8749 | 0.6729 |
| Total  | 2.8500 | 0.5732 | 0.4548 |

**Table 2**. RMSE of the absolute variance algorithm [3], ResNet50 [2], and the proposed DNN for different magnification settings.

 Figure 2 shows several examples of score prediction.

**Acknowledgement**
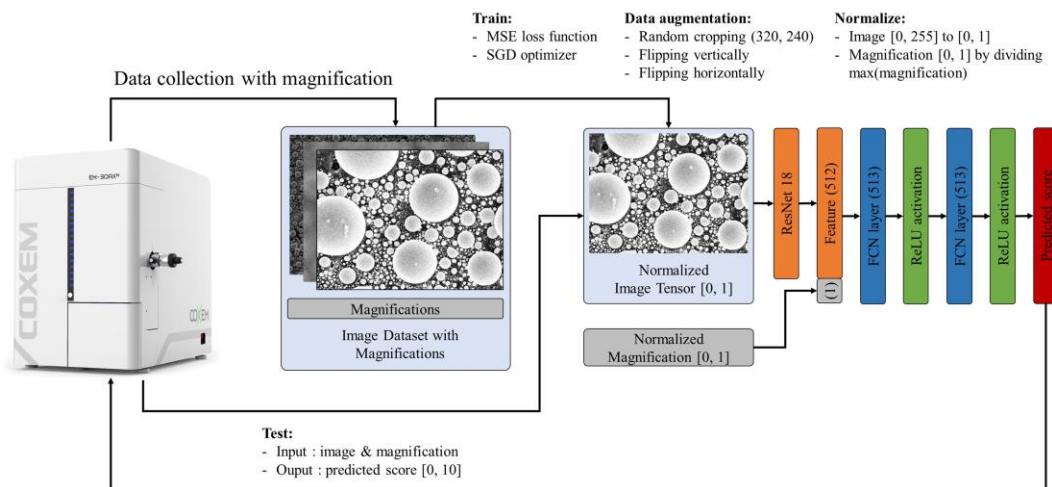
**Figure 1.** Proposed deep-learning-based autofocus score prediction algorithm. Unlike the previous work [2], the magnification is also inserted as an input. The proposed deep-learning architecture includes only a few layers, which results in faster computation than ResNet50 [2].
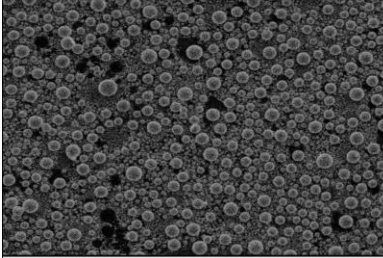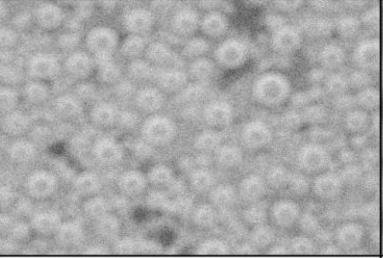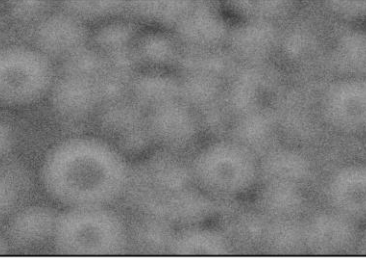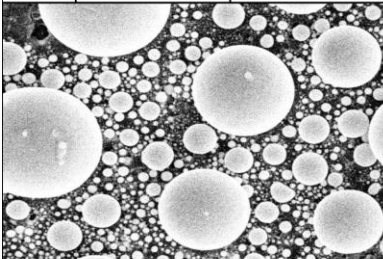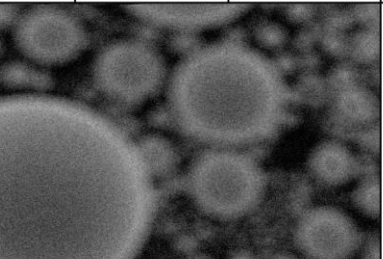
| 500 | | | 1000 | | | 2000 | | |
|---|---|---|---|---|---|---|---|---|
| Expert | ResNet50 | Proposed DNN | Expert | ResNet50 | Proposed DNN | Expert | ResNet50 | Proposed DNN |
| 9.0 | 8.8169 | 8.9995 | 5.0 | 5.5526 | 5.1890 | 5.0 | 3.7179 | 4.4746 |

| 5000 | | | 10000 | | |
|---|---|---|---|---|---|
| Expert | ResNet50 | Proposed DNN | Expert | ResNet50 | Proposed DNN |
| 10.0 | 9.8954 | 9.9533 | 8.0 | 7.3533 | 7.9229 |

**Figure 2.** Score prediction examples of three cases: i) SEM experts' score, ii) score predicted by ResNet50 with the new dataset, and iii) score predicted by the proposed DNN with the new dataset. As seen from the examples, the previous work [2] often fails to accurately predict the AF score, while the proposed DNN shows robust prediction performance for different magnification settings.

References
[1] De Nobili, Cristiano, (2017) Deep Learning for Nanoscience Scanning Electron Microscope Image Recognition *in Master in High Performance Computing (MHPC 2017.)*
[2] H. Kim, (2019) Deep-Learning Based Autofocus Score Prediction of Scanning Electron Microscope *in Microscopy & MicroAnalysis, 2019. (M&M 2019).*
[3] Frans C.A. Groen, (1985) A Comparison of Different Focus Functions for Use in Autofocus Algorithms in *in Journal of Microscopy,* 6 p. 81-91
[4] A. Santos, (1997) Evaluation of Autofocus Functions in Molecular Cytogenetic Analysis *in Journal of Microscopy,* 188 p. 264-272.
[5] Y. Yao, (2006) Granular Computing for Data Mining *in Society of Photographic Instrumentation Engineers* (*SPIE 2006*) p. 1-12.
[6] K. He, (2016) Deep Residual Learning for Image Recognition *in Conference on Computer Vision and Pattern Recognition (IEEE CVPR 2016).*