

## Deep-Learning Based Autofocus Score Prediction of Scanning Electron Microscope

Huisoo Kim<sup>1</sup>, Moohyun Oh<sup>1</sup>, Heerang Lee<sup>1</sup>, Jonggyu Jang<sup>2</sup>, Myeung Un Kim<sup>2</sup>, Hyun Jong Yang<sup>1,2,\*</sup>, Michael Ryoo<sup>1</sup>, and Junhee Lee<sup>3</sup>.

<sup>1</sup>. Egovid Inc., UNIST-gil 50, Ulsan 44919, Korea.

<sup>2</sup>. Ulsan National Institute of Science and Technology, UNIST-gil 50, Ulsan 44919, Korea.

<sup>3</sup>. Coxem Co. Ltd, Techno 2-ro 199, Yuseong-gu, Daejeon 34025, Korea.

\* corresponding author

Although many autofocus (AF) algorithms have been proposed and compared in the literature [1], [2], none of the existing algorithms can work perfectly for images of a scanning electron microscope (SEM) in practice. This is because a simple mathematical scalar metric cannot perfectly capture the quality of images, especially for a variety of SEM samples, hardware specifications, measurement environment, and etc. In addition, a simple scalar AF metric cannot govern a variety of control parameters such as sharpness, contrast, and brightness. At the era of the 4th industrial revolution, the ultimate goal would be a fully autonomous machine controlling the SEM parameters to get the best image just as a SEM specialist does. To take the first step to the automatic machine-controlled SEM development, we propose a supervised learning framework that automatically assesses the quality of sample images as if a SEM specialist does. Specifically, we develop a deep learning computer software that uses an input of a sample image and current control parameters such as brightness, contrast, and focus to automatically score the quality of the sample image. To evaluate how accurately the proposed deep learning software can score, we define a mean squared error loss (MSE)  $L$  as  $L = \frac{1}{n} \sum_{i=1}^n (y_i - \underline{y}_i)^2$ , where  $y_i$  is the score that a SEM specialist gives and  $\underline{y}_i$  is the score predicted by the proposed software. The following neural network architectures are used for deep learning: i) parallelized convolutional neural network (CNN) and fully-connected neural network (FCNN) (PACF) as in Fig. 1, ii) VGG [3], and iii) ResNet [4]. Gray-scale images with a sample type of a grid or tin ball are used. The original image resolution is 640x480-pixel; however, we augment the amount of data by cropping the images to 224x224 pixels, flipping them vertically, and rotating them randomly. As a result, the total number of images is 2134. For supervised learning, the deep learning network should be trained with known inputs and outputs. Thus, 1493 images out of the 2134 images are used for training, and the remaining 641 images are used for testing. Control parameters such as brightness, contrast, and focus of each image are set differently. In addition, each image is scored by a SEM specialist as an overall score on the quality of the image.

Experiments are implemented under the following environment: GeForce GTX 1080M, CUDA 10.0, Python 3.6.7, Pytorch 1.0.0, and OpenCV 4.0.1. Table 1 shows the performance comparison for the average test runtime of each image and the Root of MSE (RMSE) between the existing AF algorithms and deep-learning based score prediction networks for 100 epoch and 0.00015 learning rate. For comparison, the values of the AF functions are normalized to scores by  $Score_{normalized} = AF_{original} \times (\frac{Score_{max} - Score_{min}}{AF_{max} - AF_{min}})$ . The proposed networks outperform the existing AF functions with respect to both of the test runtime and RMSE. Fig. 2 shows the examples of the scores from the specialist, one of the existing AF algorithms based on absolute variance, and one of the proposed networks, PACF, respectively. Even in the distorted images, our proposed software predicts the score much more accurately than the conventional AF algorithms.

To sum up, the deep-learning based score prediction approach for an automatic SEM system is proven to be superior to the existing AF algorithms even in the test runtime as well as RMSE. Based

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2018-0-0095,Development of Joint Electrical/Mechanical Drone Beamforming based on Target Detection and Precise Attitude Control), the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2019-2017-0-01635) supervised by the IITP, and the Technology development Program(S2557960) funded by

on the proposed automatic score prediction algorithm, a fully autonomous SEM control software can be developed by using the existing deep learning frameworks, such as reinforcement learning, which will be discussed in our future work.

References:

[1] A. Santos *et al.*, Journal of Microscopy **188** (1997) p. 264-272.  
 [2] Y. Yao *et al.*, SPIE (2006) p. 1-12.  
 [3] K. Simonyan and A. Zisserman, ICLR (2015) p. 1-14.  
 [4] K. He *et al.*, IEEE CVPR (2016) p. 770-778.

Autofocus Function	Time (s)	RMSE	Neural Network	Time (s)	RMSE
<b>Absolute Variance (ACM Variance)</b>	0.0713	1.5374	<b>PACF</b>	0.0050	1.2256
<b>Squared Variance</b>	0.0758	2.2610	<b>VGG11</b>	0.0061	1.4400
<b>Absolute Gradient</b>	0.0270	2.1937	<b>VGG13</b>	0.0076	1.2035
<b>Squared Gradient</b>	0.0351	2.9930	<b>VGG19</b>	0.0100	1.3250
<b>Correlation (Vollath's F4)</b>	0.0522	2.5133	<b>ResNet50</b>	0.0110	0.8329
<b>Entropy</b>	0.0820	2.3036	<b>ResNet101</b>	0.0174	0.8049

Table 1. Performance of the existing AF algorithms and deep-learning based score prediction networks

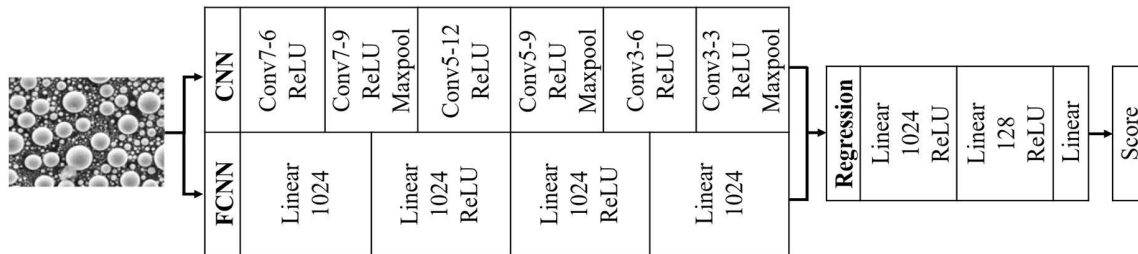


Figure 1. PACF (Parallelized CNN + FCNN) Model

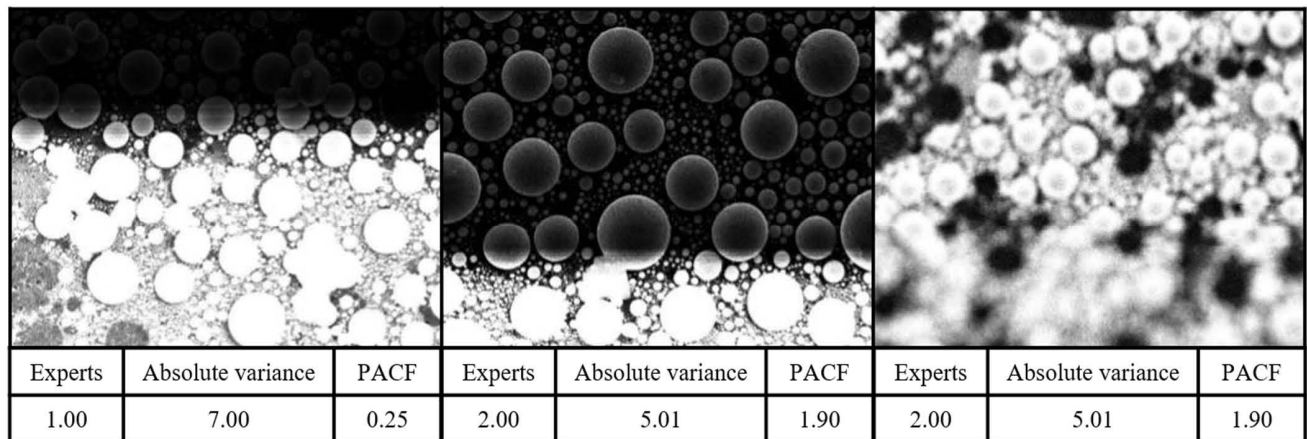


Figure 2. Score comparison of distorted tin ball images