

## **Supplementary Information for**

### **A deep convolutional neural network approach to single-particle recognition in cryo-electron microscopy**

**Yanan Zhu<sup>1</sup>, Qi Ouyang<sup>1,2,3</sup>, Youdong Mao<sup>1,2,4\*</sup>**

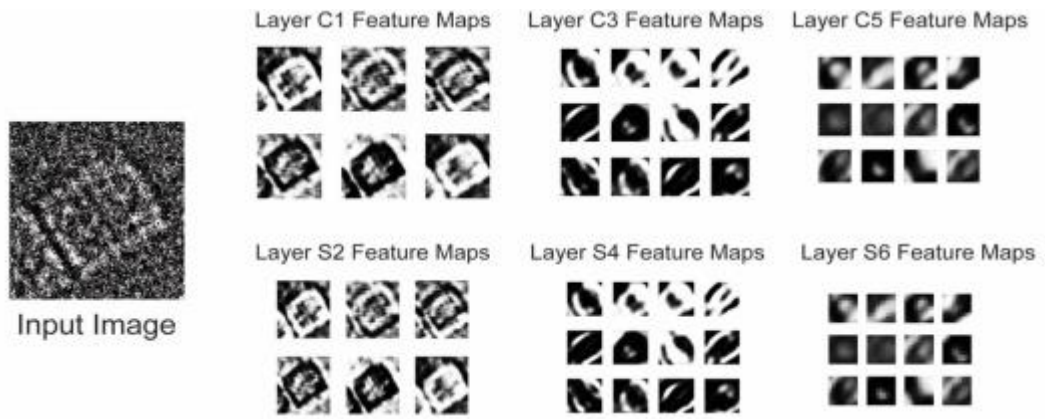
<sup>1</sup>Center for Quantitative Biology, <sup>2</sup>State Key Laboratory for Artificial Microstructure and Mesoscopic Physics, Institute of Condensed Matter Physics, School of Physics,

<sup>3</sup>Peking-Tsinghua Center for Life Sciences at Academy for Advanced

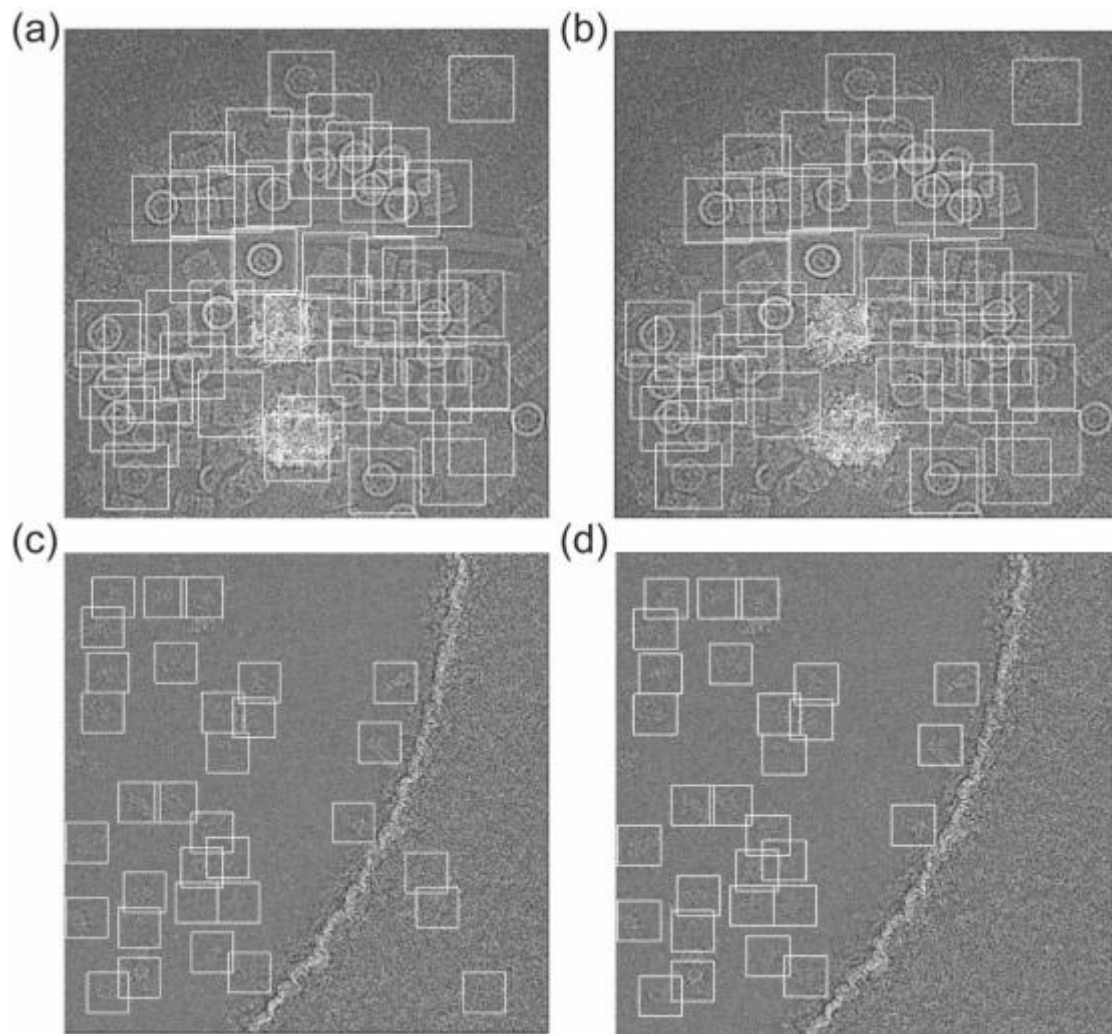
Interdisciplinary Studies, Peking University, Beijing 100871, China. <sup>4</sup>Intel Parallel

Computing Center for Structural Biology, Dana-Farber Cancer Institute, Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA 02115, USA.

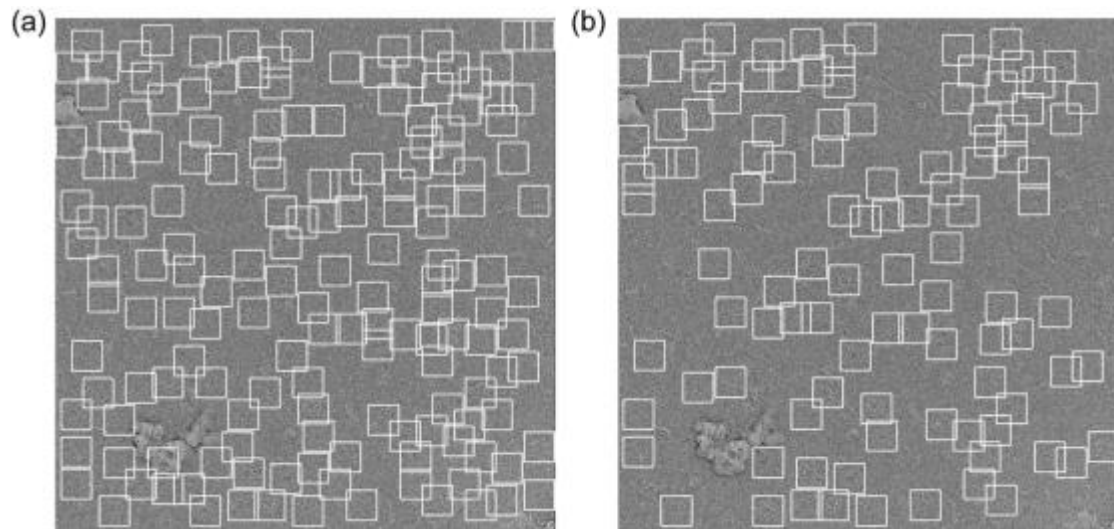
\*Corresponding author. Tel: +1 617 632 4358. Fax: +1 617 632 4338. E-mail address: youdong\_mao@dfci.harvard.edu (Y. M.).



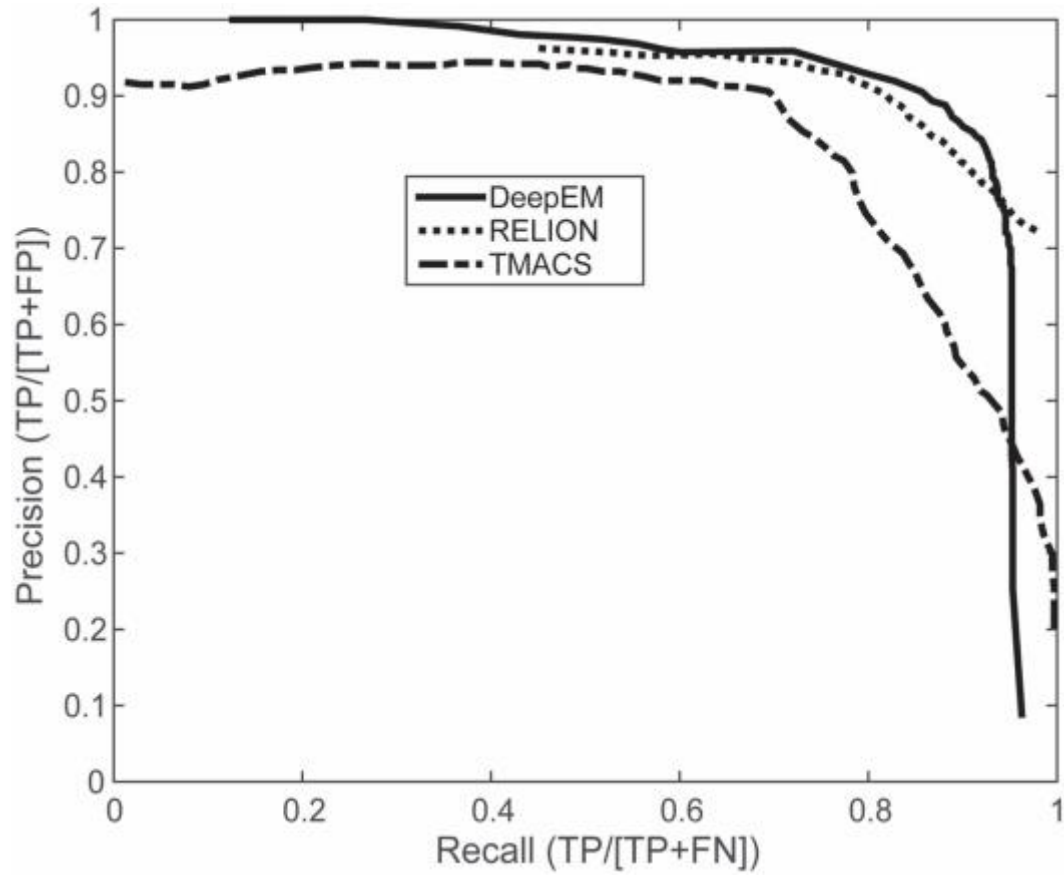
**Supplementary Figure 1.** The feature maps of convolutional and subsampling layers from a typical particle image of KLH learned by our CNN.



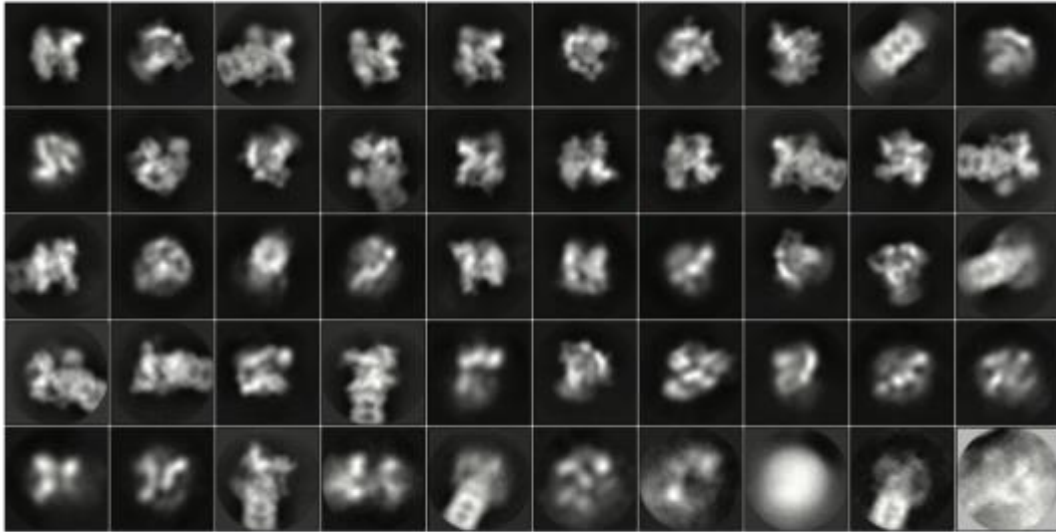
**Supplementary Figure 2.** (a) and (b) show the comparison of the results between before and after additional selection using standard deviation of the KLH dataset, respectively. (c) and (d) show the comparison of the results between before and after additional selection using standard deviation of the 19S, respectively.



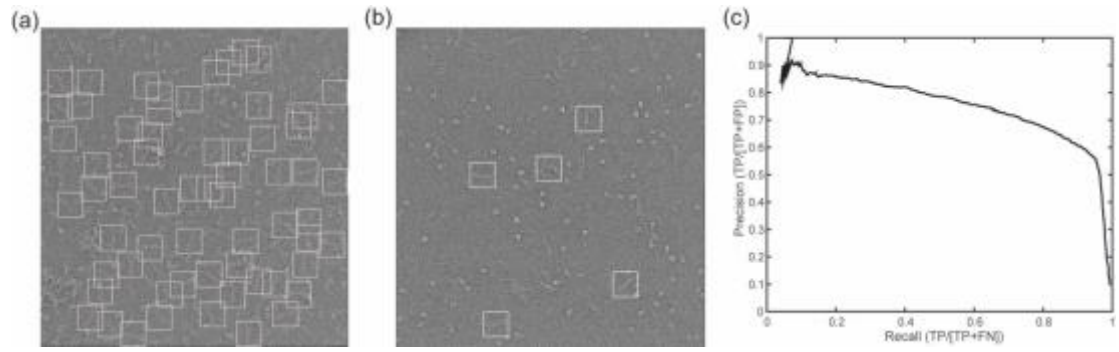
**Supplementary Figure 3.** (a) and (b) show the comparison of the results before and after optimization of the training dataset, respectively.



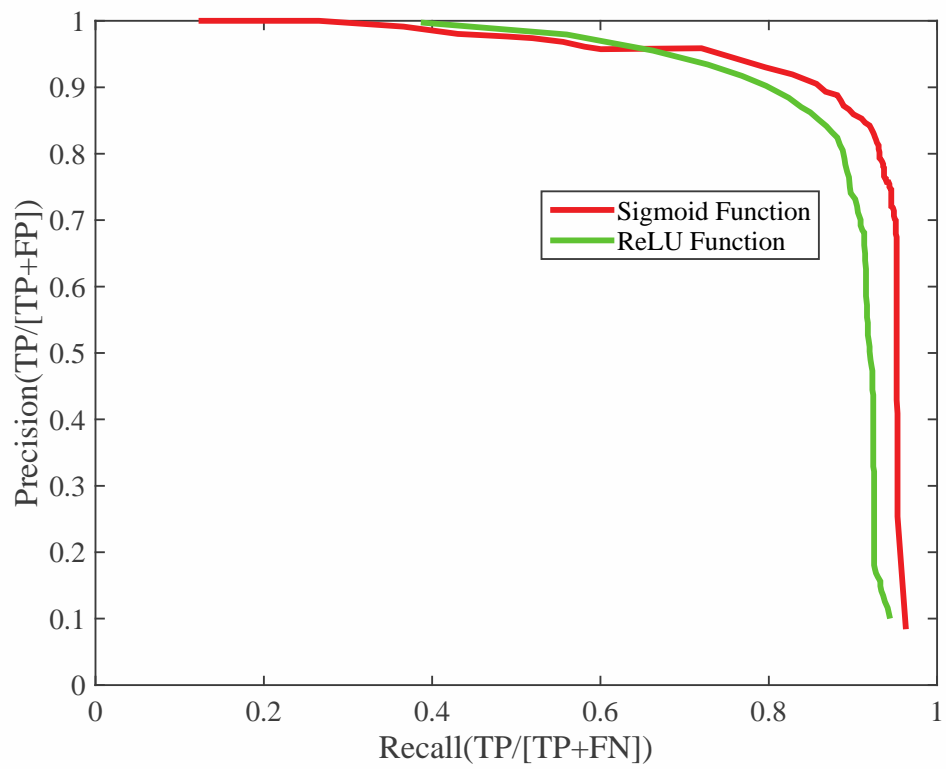
**Supplementary Figure 4.** Comparison of the DeepEM with TMACS and RELION in the KLH dataset. The curves of TMACS and RELION are obtained from the published data directly.



**Supplementary Figure 5.** The reference free 2D classification of 19S proteasomes recognized by DeepEM.



**Supplementary Figure 6.** The results in the recognition of the side view of 26S proteasome by DeepEM.



**Supplementary Figure 7.** The comparison results of different activation functions tested on the KLH dataset.