

Deep Learning Based Segmentation of Nuclei from Fluorescence Microscopy Images

Prabhakar R. Gudla^{1,3}, George Zaki⁴, Sigal Shachar^{2,3}, Tom Misteli^{2,3} and Gianluca Pegoraro^{1,3*}

¹. High-throughput Imaging Facility (HiTIF),² Cell Biology of Genomes Group,

³. National Cancer Institute, National Institutes of Health, Bethesda, MD, USA.

⁴. Biomedical Informatics and Data Science Directorate, Frederick National Lab for Cancer Research, Frederick, MD, USA.

* Corresponding author: gianluca.pegoraro@nih.gov

Segmentation of cellular objects is a necessary step in many image processing pipelines of fluorescent microscopy images. As an example, High-Throughput Imaging (HTI) assays often require accurate detection of nuclei and other subcellular compartments to quantify the biological effects of RNAi or chemical reagents [1]. Deep Convolution Neural Networks models (D-CNNs) have surpassed traditional image processing approaches for biological object detection [2], are mostly non-parametric, and require few and simple image pre-processing steps. However, successful training of these models generally requires large numbers (hundreds to thousands) of representative images along with their human expert-generated ground-truth annotations. This hinders the rapid implementation of D-CNN models for a wide range of HTI assays.

Here we describe a novel image analysis workflow based on D-CNNs and transfer learning for the segmentation of fluorescently labelled mammalian cell nuclei. The workflow (Fig. 1A) uses few heavily augmented images for training of D-CNNs, generation of preliminary ground-truth annotations by traditional image processing methods or by pre-trained D-CNNs, a rapid preliminary ground-truth correction step by expert annotators, and two state-of-the-art D-CNNs: U-Net [2] and Mask R-CNN [3]. U-Net was converted to an instance segmentation architecture by formulating the classification task as a regression task for predicting nucleus edges and internal seeds, which were then used as inputs to a seeded watershed algorithm [4]. Mask R-CNN was used without any modifications [5].

Both D-CNNs were trained using 7 fluorescence microscopy images (1278x1078 pixels), containing on average 95 nuclei per image, from a single well of a multi-well plate containing DAPI-stained human MCF10A cells (Fig. 1B). The preliminary ground truth annotations for these images were generated using traditional image segmentation approaches [2], followed by expert-driven correction of any errors. The input fluorescence microscopy images were first heavily augmented, and random patches (256x256 pixels) of these augmented images were then used to train the U-Net and Mask R-CNN models. The D-CNNs trained on the random patches of MCF10A cells images were subsequently used in inference mode to generate preliminary ground-truth annotations for test microscopy images of another well of MCF10A cells (7 images), and for two additional cell lines: U2OS (1 image) and HCT-116 (1 image). The nuclear morphology of the U2OS and HCT-116 cells is substantially different from MCF10A cells. These preliminary segmentation results were then manually verified and corrected (Fig. 1B).

As seen in Fig. 1B, this analysis workflow can segment individual nuclei in both training (MCF10A) and in test images (U2OS and HCT-116), even when the nuclei are tightly packed and touching, such as in the case of U2OS and HCT-116 cells. This is remarkable because of their difference in nuclear morphology when compared to MCF10A cells, and also because the test image datasets were acquired at a different

magnification (20X) from the training image dataset (60X) (Fig. 1B). Quantitative evaluation of D-CNN's performance on images of MCF10A, U2OS and HCT-116 cells, as measured by the average precision (AP) at different Intersection over Union (IoU) thresholds for both models (Fig. 1C), indicates that Mask R-CNN results were more stable at higher IoU thresholds when compared to U-Net, and that the source of the preliminary ground truth annotations (i.e. either from U-Net or Mask R-CNN) for a specific test cell line has a strong influence on the performance of each D-CNN (Fig. 1B and 1C).

Altogether, these results show that, by combining input augmentation and transfer learning, D-CNNs can be trained with a very limited number of objects to accurately segment nuclei for a variety of cell lines with different nuclei morphologies, and that they can be used to generate preliminary ground-truth annotations for previously unseen datasets, which can then be further quickly corrected by human experts. These advancements will facilitate the adoption of D-CNN workflows for the analysis of large and diverse image datasets generated in HTI assays [6].

References:

- [1] G Pegoraro and T Misteli, *Trends Genet* **33** (2017), p. 604.
- [2] T Falk et al., *Nat Methods* **16** (2019), p. 67.
- [3] K He et al., arXiv e-prints [Online], (2015).
- [4] F Meyer, *Comp Imag Vis* **2** (1994), p. 77.
- [5] W Abdulla, [Online], (2017), https://github.com/matterport/Mask_RCNN.
- [6] This research was supported by funding from the Center for Cancer Research as part of the Intramural Research Program at NCI/NIH. The authors acknowledge the CBIIT Server Team and the High-Performance Computing Group at NCI/NIH for computational support. Prabhakar R. Gudla and George Zaki contributed equally.

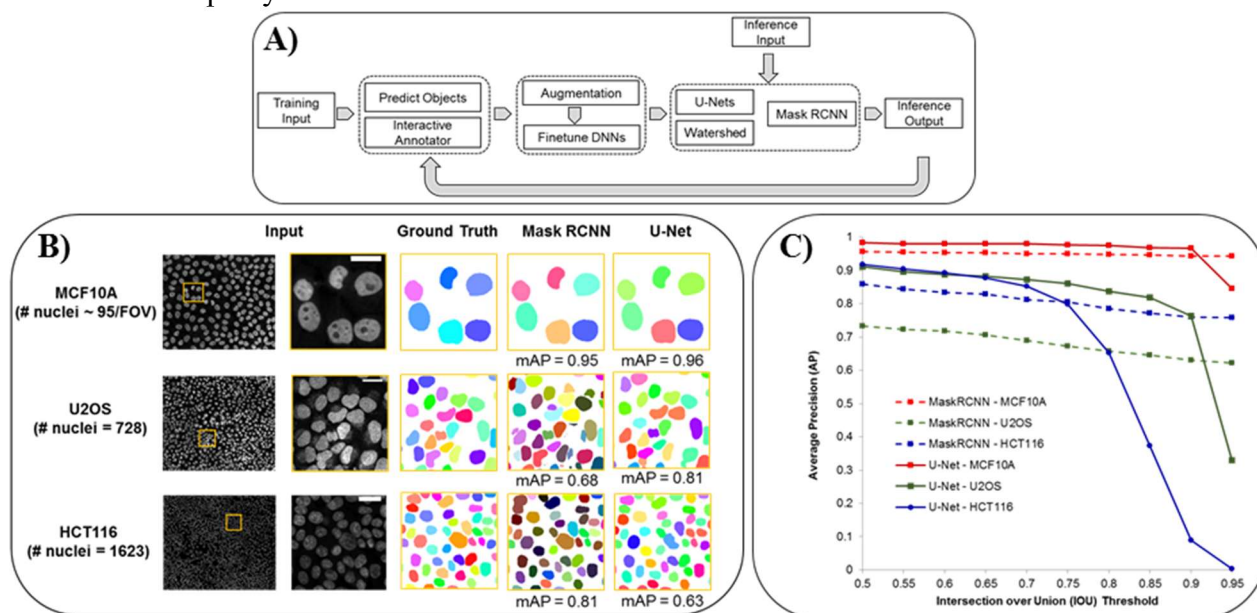


Figure 1. A) Schematics of the D-CNN(s) based workflow for nuclei segmentation from fluorescence microscopy images. B) Representative images (greyscale) of human MCF10A, U2OS and HCT116 nuclei stained with the DAPI DNA dye. Pseudo-colored images of ground truth annotations and inference results using the U-Net and Mask-RCNN architectures, respectively, on the greyscale images along with the mean average precision (mAP). Scale bars: 15 μm (MCF10A), 25 μm (U2OS), and 25 μm (HCT-116). C) AP curves for the D-CNN models for different cell lines.