FINE TUNED MODELS FOR OOCYTE SEGMENTATION IN STAINED MOUSE OVARIAN SECTIONS

Alexandre Cunha[1], *Máté Borsos*[2], *Sirisha Gudavalli*[2], *Viviana Gradinaru*[2]

[1]Center for Image Analysis, Beckman Institute
[2]Division of Biology and Biological Engineering
California Institute of Technology, Pasadena, California, USA

1. ABSTRACT

One of the first steps during mammalian ovary development is the formation of dormant primordial ovarian follicles. These consist of a single layer of granulosa cells surrounding each oocyte. Later these follicles start to grow as granulosa cells proliferate and oocyte size increases in diameter, see Fig.1. To study this process we require accurate counts and sizes of follicles, similar to [1]. We are currently using oocyte size as a proxy to determine follicle growth stage and we rely on automatic segmentation of oocytes present in images of stained sections of mice ovary to provide us quantitative data to follow growth.

We adopted Cellpose [2] as our preferred segmentation tool as it shows excellent performance segmenting round structures and it can be readily used for training new models. Its pretrained models cyto, cyto2, and others are insufficient to capture only oocytes of interest as they add many false positives and masks with inferior boundary accuracy, even after adjusting threshold parameters (see red contours in Fig.1G crops). We fine tuned these pretrained models creating new models capable of segmenting oocytes of any size and containing nuclei. Our assumption when mandating nuclei presence is that sections of oocytes containing nuclei are larger and better represent real growth size, assuming an oocyte is roughly a sphere with its nucleus at its center. Periodic acid-Schiff (PAS) stained ovarian slices were acquired at $0.3776\mu m$ pixel resolution totaling 55 training images containing approximately 2,050 oocytes.

An expert biologist (second author) used our Collaborative Segmentation web tool, Fig.1F, to annotate with dots and scribbles true positive oocytes in over-segmented images. We used those to perform two rounds of weakly supervised (false negatives not corrected) fine tuning, each round assuming different shuffles of training and test images. Annotations amounted to 499 ground truth oocytes ($\sim$ 25% of total). After training with 2,000 (first round) and 1,000 (second round) epochs, we selected best models based on a convex combination of training $\mathcal{L}_r$ and test loss $\mathcal{L}_t$ values reported by Cellpose: selected models have low values of $\mathcal{L} = 2(\mathcal{L}_r + 1.5\mathcal{L}_t)/5$ in the range $\mathcal{L} \in [0.045, 0.066]$ typically with $\mathcal{L}_t < \mathcal{L}_r$. Our new models segmented well test images avoiding nucleus free regions and capturing oocytes within primordial and growing follicles with median diameters, respectively, of $18\mu m$ and $42\mu m$, within the expected $[10, 60]\mu m$ diameter range. We plan to further extend our models to segment full follicles and combine results from our best models in a segmentation fusion strategy [3].

We thank financial support from the Beckman Institute at Caltech to the Center for Image Analysis (AC) and from Bill and Melinda Gates Foundation (MB,VG). Corresponding author: cunha@caltech.edu

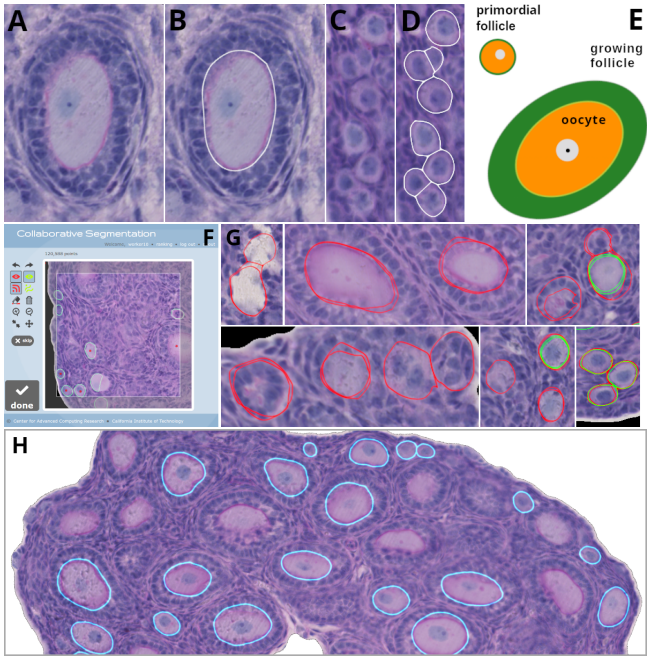

Fig. 1. A growing follicle (A,E) originates from a small primordial follicle (C,E) and it continuosly expands in size. Oocytes within follicles are surrounded by granulosa cells (green region in cartoon E) which proliferate from a single layer of cells. Measuring segmented oocytes (B,D) allows us to approximately determine follicle growth stage. Pretrained Cellpose models cyto and cyto2 over segment (red contours in G) and they are not adequate to segment only oocytes of interest, those of any size exclusively containing a nucleus (H). Our new models segment well test images identifying all oocytes and avoiding nucleus free oocytes (H, and green contours in G).

2. REFERENCES

[1] Carolina Lliberos, Seng H Liew, Pirooz Zareie, Nicole L La Gruta, Ashley Mansell, and Karla Hutt, "Evaluation of inflammation and follicle depletion during ovarian ageing in mice," *Nature Scientific Reports*, vol. 11, no. 1, pp. 278, 2021.

[2] Marius Pachitariu and Carsen Stringer, "Cellpose 2.0: how to train your own model," *Nature Methods*, vol. 19, no. 12, pp. 1634–1641, 2022.

[3] Carlos HC Pena, Tsang Ing Ren, Pedro D Marrero Fernandez, Fidel A Guerrero-Peña, and Alexandre Cunha, "An ensemble learning method for segmentation fusion," in *2022 IEEE International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2022, pp. 1–6.