

For segmentation purposes, four different datasets were used

1. BBBC039: This dataset is a derivative from BBBC022. This data set has a total of 200 fields of view of nuclei captured with fluorescence microscopy using the Hoechst stain. The collection has around 23,000 single nuclei manually annotated to establish a ground truth collection for segmentation evaluation. A small fraction of this dataset overlaps with Kaggle Data Bowl 2018 competition. The images have been split in three subsets: training, validation and test. A metadata package contains files with the names of the images in each of these subsets.
Website: <https://data.broadinstitute.org/bbbc/BBBC039/>
2. Data Science Bowl 2018: This dataset contains a large number of segmented nuclei images. The images were acquired under a variety of conditions and vary in the cell type, magnification, and imaging modality (brightfield vs. fluorescence). For each cell image, ground truth for each nuclei is provided.
Website: <https://www.kaggle.com/c/data-science-bowl-2018/data>
3. ISBI 2017: (from authors): In this dataset, each patient has 3 to 7 annotated 512x512 samples from their original WSI. It is a pixel wise annotation, i.e. each pixel value above 0 is considered as the label of a corresponding nucleus. We have annotated a considerable amount of cells, including normal epithelial and myoepithelial breast cells (localized in ducts and lobules), invasive carcinomatous cells, fibroblasts, endothelial cells, adipocytes, macrophages and inflammatory cells (lymphocytes and plasmocytes). In total, we have 33 images with a total of 2754 annotated cells, the maximum number of cells in one sample is 293 and the minimum number of cells in one sample is 5.
Website: <https://peterjacknaylor.github.io/PeterJackNaylor.github.io/2017/01/15/Isbi/>
Paper: Nuclei Segmentation in Histopathology Images Using Deep Neural Networks
4. MoNuSeg: Training data containing 30 images and around 22,000 nuclear boundary annotations has been released to the public previously as a dataset article in *IEEE Transactions on Medical imaging* in 2017.
Website: <https://monuseg.grand-challenge.org/Data/>