Segmenting single cells from blood/tissue sample is critical to identify diseased cells to arrive at early diagnosis. On the other side, the behaviors of single cells have brought more attention in biological researches. For example, the lineage tracking of single embryo cells is vital to understand the mechanisms in early stage embryo developments; the shape analysis of mammalian cells during Epithelial-Mesenchymal transition is key to dissect the transformation of how normal cells become cancer cells. Till now, identifying cells in images (cell segmentation) has been challenging due to the high density and imbalanced illumination of the sample. Recently, machine learning and deep learning has become the instrumental tool for computer vision tasks [1][2]. In this work, we proposed a highly efficient and effective neural network model that has accurate predictions on both cell segmentation and cell count.

**Dataset**

**ZEISS LSM 700 confocal microscope Dataset**

- 46X objective on MCF-10A human breast cancer cell line
- The cell nucleus was stained by DAPI and illuminated by laser with wavelength 405nm
- Training dataset includes 484 images with ground-truth binary mask, and 8 images with the corresponding masks for validation
- The size of image is of 488 × 488, and each image contains around 100 cells.
- The ground-truth labels are computed using thresholding and other segmentation methods. Segmentation errors are manually corrected image-by-image. The ground truth coordinates for bounding boxes are computed based on the ground truth labels.

**Data augmentation**

- Augmented data with cropping, adding noise, rotation, mirror

**Contribute**

- Capturing each single cell in cell clusters, we integrated Region Proposal Networks (RPN) to detect cells with bounding boxes.
- Predict the cell segmentations by integrating U-Net with pre-computed weight maps.
- Integrate Watershed algorithm with cell centroids (predicted by RPN) and cell segmentations (predicted by U-Net) as inputs to improve the boundary predictions of densely touching cells.
- Extensive quantitative evaluations illustrated the superiority of the proposed method compared with state-of-the-art segmentation models.

**Instance-U-Net**

- To improve the cell segmentation performance, we bring the objectness insights into original U-Net.
- The aforementioned feature extractor network as encoder of U-Net is improved in integrating with instance segmentation scores.
- Two cost functions are implemented: binary cross-entropy (BCE) loss function for segmentation and dice loss.
- The BCE loss is denoted as:
  \[ L_{BCE} = -\sum w(x) \log(p_{y|x}(x)) \]
  \( w(x) \) is the weight map. \( \log(p_{y|x}(x)) \) is the pixel-wise cross-entropy sum over all pixels.
- The weight map is computed for each ground-truth segmentation mask to urge the network to learn the small separation borders between touching cells.
- \( d_{ij}(s) \) denote the distance to the border of the nearest cell second nearest cell. In practice we set \( w_{ij} = 10 \) and \( \sigma = 5 \).
- The so-called Dice loss is denoted as:
  \[ L_{dice} = 1 - 2 \frac{\sum p_{i}(x)p_{j}(x)}{\sum p_{i}(x) + \sum p_{j}(x) + \varepsilon} \]
  \( p_{i}(x) \) and \( p_{j}(x) \) are the binary labels from prediction and ground truth labels. \( \varepsilon \) is a small factor to prevent division by zero, in practice we set \( \varepsilon = 10^{-10} \).