# Convolutional Models for Biomedical Image Segmentation

## Kaggle 2018 Data Science Bowl

*Alexander Haigh (haighal), Frits van Paasschen (fritsvp), and Joseph Murphy (murphyjm)*

## Problem

- Cell identification in biomedical images is essential to drug development
- Data is inherently inhomogeneous and labeling by hand is costly in time and funding
- **Goal:** Deliver robust model for labeling cell nuclei across wide variety of cell types, microscopy methods
- **Dataset:** 603 train images, 67 dev, 65 test using multiple microscopy techniques and ; provided by Kaggle. Train and dev images have associated masks, submit test prediction to Kaggle for evaluation

## Model Architecture

- UNet, introduced by Ronneberg et al. 2015
- Modifications to their model
  - 256 x 256 input image (instead of 572 x 572)
  - "Same" convolutions result in 256 x 256 output, no cropping
  - Sigmoid activation with BCE Loss rather than softmax classification
  - Modified final layers
  - Post-processing: global map → individual masks

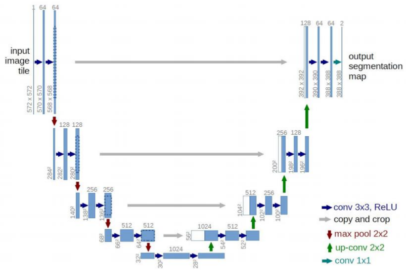



**Figure 1: U-Net Architecture,** as introduced by Ronneberg et al. The U-Net is a fully convolutional NN that predicts a global segmentation map over the entire image (outputs a binary 0 or 1 for each pixel, with 1 corresponding to part of a cell). The contracting path collects high-level semantic information about the image at each step; these features are then upsampled and combined with lower-level information in the expansion path.

## Data Augmentation

- Convert to black and white, normalize to [0, 1]
- Invert B&W versions of 3-channel images
- Resize input to 256 x 256
- Random piecewise affine transformations
- Merge individual cell masks into one mask for training

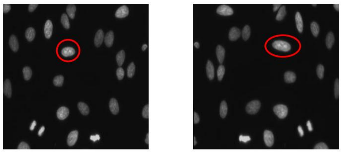


**Figure 2: Sample Augmented Training Example.** Left, raw microscope image. Right, image after a horizontal flip and a piecewise affine transformation. Red circle highlights cell warps.

## Postprocessing and Prediction

- Otsu Thresholding: prob. distribution → global mask
- Split global mask into local masks by BFS

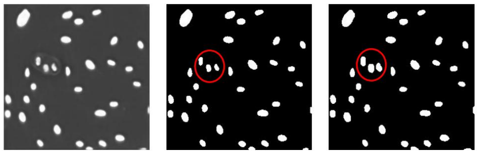
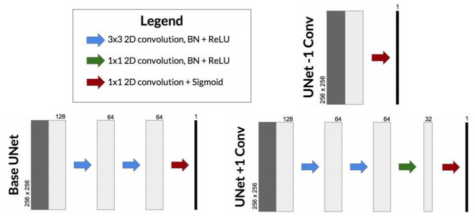


**Figure 3: Sample Prediction. Left,** output of the UNet (float values in the range [0, 1]). **Middle,** predicted mask after applying Otsu Thresholding, which binarizes the 0-1 heatmap by finding the threshold that minimizes the weighted sum of intra-class variance. **Right,** ground-truth labeled segmentation mask. Sample difference in red.

## Final Layer Modifications




## Evaluation Metric

- Kaggle defines evaluation metric, LB
- For each image, average cell-wise precision over various IoU thresholds for predicted and actual objects:

$$\frac{1}{|thresholds|}\sum_t \frac{TP(t)}{TP(t)+FP(t)+FN(t)}$$

- LB metric averages the average precision of entire dataset
- ALso measure average pixel-wise IoU, precision, and accuracy over the global mask

## Results

| Model | Dev | | | | Test |
|---|---|---|---|---|---|
| | Accuracy | Precision | IoU | LB | LB |
| Base UNet | 0.860 | 0.478 | 0.804 | **0.342** | **.277** |
| UNet +1 Conv | 0.931 | 0.451 | 0.786 | 0.337 | .265 |
| UNet -1 Conv | 0.913 | 0.460 | 0.797 | 0.321 | .274 |

## Error Analysis and Next Steps

- Artifacts in images - conjoined masks, leaked information
  - Waterfall thresholding, hysterisis, binary opening in postprocessing, no resizing at test time
  - Weighted Loss: force algorithm to learn border pixels
- Heterogeneous input: Different in-convolution for 3-channel images
- New Architecture: Mask-RCNN to predict individual masks directly

**Figure 4: Sample Error. Left,** raw microscope image. **Middle,** predicted mask. **Right,** ground-truth mask label. Errors include conjoined masks as well as masks that leak information as a result of imaging quality; this includes finger-like projections, "holes" in masks, and split masks. Example errors circled in red.