# Deep Learning for Cell Segmentation

*Grace Y. Kim, Will Song, Lucia Zheng*

Stanford
Department of Computer Science

## Background/Introduction

- Light microscopy is a cheap and accessible imaging technique to visualize and quantitatively capture cellular features at a large scale.
- The ability to accurately perform image segmentation on cell microscopy images could enable scientists to study a wide array of biological phenomena and move towards precision medicine treatments.
- Deep learning models have achieve state-of-the-art results for image segmentation tasks in many domains but has had limited growth in cell segmentation tasks due to the limited size of available annotated cellular image data (is an expensive and involved process).

## Problem Statement

Despite the advent of LIVECell, a novel, large annotated dataset of cell cultures, many challenges in the development of high-quality and flexible cell segmentation remain. We seek to iterate in this space through experimentation with models trained for semantic vs. instance segmentation, models of varying levels of capacity, transfer learning, and synthetic data generation.

## Dataset

**LIVECell (Label-free In Vitro image Examples of Cells)**

- Dataset of 5,000 manually annotated and expert-validated fluorescent microscopy images of 2D cell culture (1,686,352 individual cells)
- Eight cell types in the dataset: A172, BT-474, BV-2, Huh7, MCF7, SH-SY5Y, SkBr3, SK-OV-3
- Annotations available in Microsoft MS COCO object detection format

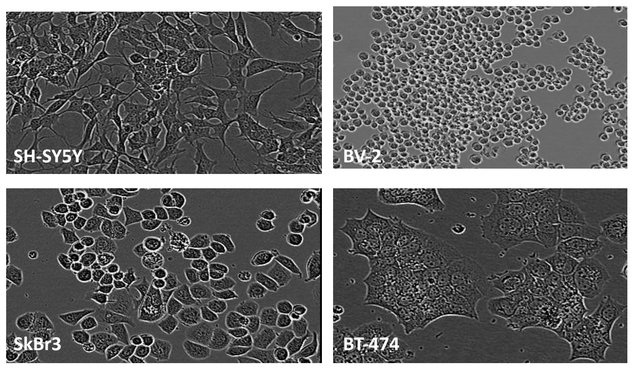
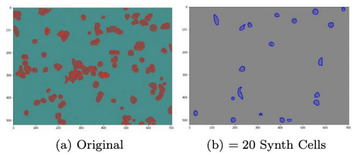


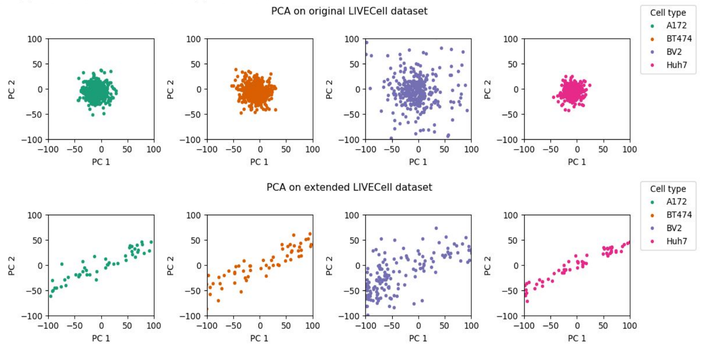

Figure 1. Demonstrates some of the unique cell morphologies in the dataset.

## Methods

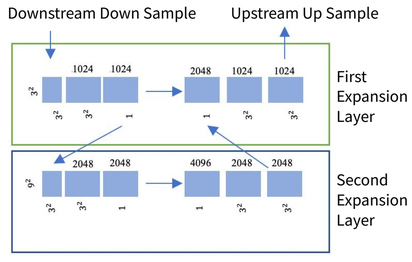


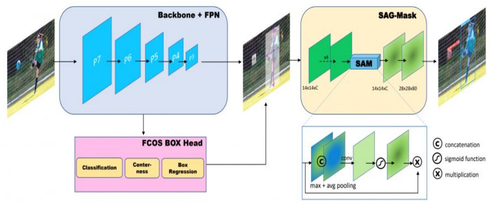

**Synthetic Data Generation**
Extract cells of the same type from each image and vary cell location and density. Mimics differing confluence levels as the cell culture grows. Extended dataset by 5x.

**PCA Analysis** below shows how data extension affected the distribution of data on two principal components identified from the original dataset.




**U-Net -** Commonly used structure for semantic segmentation on biomedical data
**U-Net Expanded -** Add Up and Down layer with 1024 channels at the bottom of the U-Net
**U-Net Expanded Ensemble -** Add 1024 and 2048 channel Up and Down layers and train one model per cell class



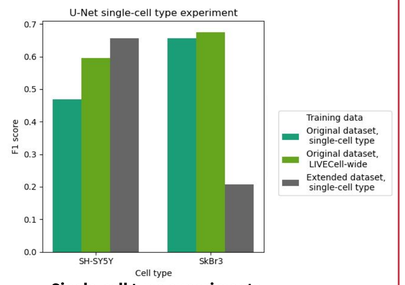

**VoVNet2 backbone** was selected for its efficiency and performance. Pre-trained on MS-COCO 2017 dataset.
**V-99** - The largest VovNet2 backbone
**V-19-SlimDW** - The smallest VoVNet2 backbone (thinner FPN with half the channel size)

**CenterMask** - Uses a Feature Pyramid Network (FPN) to segment objects at different scales. Fully convolutional one stage (FCOS) object detector for anchor free bounding box determination. Spatial Attention-Guided Mask with Spatial Attention module for instance segmentation within each bounding box.

## Experiments & Analysis

| | CenterMask-V-19-SlimDW (L) | | | CenterMask-V-99 (L) | | |
|---|---|---|---|---|---|---|
| Split | F1 | AP | AFNR | F1 | AP | AFNR |
| LIVECell | 0.823 | 34.7 | 56.2 | 0.95 | 48.5 | 44.8 |
| A172 | 0.92 | 20.5 | 61.2 | 0.95 | 39.4 | 49.7 |
| BT-474 | 0.82 | 27.9 | 55.5 | 0.86 | 45.6 | 43.3 |
| BV-2 | 0.813 | 46.7 | 47.1 | 0.86 | 53.3 | 40.2 |
| Huh7 | 0.81 | 31.3 | 45.9 | 0.90 | 54.7 | 33.1 |
| MCF7 | 0.88 | 25.2 | 65.7 | 0.91 | 40.1 | 52.4 |
| SH-SY5Y | 0.80 | 15.9 | 75.2 | 0.84 | 27.7 | 63.2 |
| SkBr3 | 0.9025 | 61.5 | 33.2 | 0.93 | 66.5 | 28.6 |
| SK-OV-3 | 0.91 | 27.7 | 55.8 | 0.94 | 54.4 | 37.5 |

| Split | U-Net (L) F1 | U-Net (EL) F1 | Deep U-Net (L) F1 | Deep U-Net Ensemble (EL) F1 |
|---|---|---|---|---|
| LIVECell | 0.5873 | 0.2591 | 0.5832 | 0.6052 |
| A172 | 0.4391 | 0.3229 | 0.4326 | 0.4496* |
| BT-474 | 0.7100 | 0.0123 | 0.7127 | 0.6184 |
| BV-2 | 0.7146 | 0.7953 | 0.7266 | 0.7999 |
| Huh7 | 0.6401 | 0.0324 | 0.6358 | 0.6798 |
| MCF7 | 0.6374 | 0.1543 | 0.6293 | 0.6304 |
| SH-SY5Y | 0.5960 | 0.0959 | 0.5825 | 0.6104* |
| SkBr3 | 0.6751 | 0.4360 | 0.6759 | 0.7128 |
| SK-OV-3 | 0.2857 | 0.2245 | 0.2700 | 0.3405* |

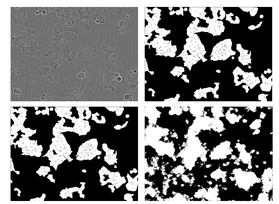



**LIVECell-wide experiments**

- Immortalized cell lines (proliferate indefinitely -> high cell density): BV-2, SH-SY5Y, Huh7. This attribute may be related to why BV-2 & SH-SY5Y were the few cell lines that benefited from training on the dataset extended with our synthetic data generation method.
- Data extension not as helpful - may have produced data that was too out of distribution, may be difficult for U-Net with limited capacity to learn meaningful patterns on diverse data
- Greater model capacity and transfer learning improved performance, especially for irregularly shaped cell lines such as SH-SY5Y and A172 and less so for regular cell lines such as BV-2 and SkBr3

**Single-cell type experiments**

- Original LIVECell paper: single-cell type training worse than LIVECell-wide training for all cell types
- Our hypothesis: single-cell type training bottlenecked by single-cell type training data size / diversity, extending the dataset with synthetic data generation could help
- Exp. results: training on the extended dataset improved single-cell type training above LIVECell-wide and single-cell type training on the original dataset for SH-SY5Y, not for SkBr3, likely because of:
(1) SH-SY5Y's unique morphology
(2) The distribution of cell densities in our synthetic images matched the true distribution of cell densities of immortalized cell lines more closely



*Top Left Clockwise* - Original image, Ground Truth Mask, Deep U-Net Mask, CenterMask-V-19-SlimDW Mask

Instance segmentation training objective + transfer learning produced crisper binary segmentation masks than U-Net models explicitly trained for semantic segmentation

## Conclusions & Future Work

- Applications of CenterMask and U-Net work very well to segment individual cells in cell microscopy, especially transfer learning from other tasks.
- Our U-Net models performed well for a model that can fit on 1 T4 GPU.
- Further exploration of automated generation of biologically relevant synthetic data.
- Exploring characterization of time as a latent parameter to the evolving cell geometries across phases (such as using a VAE or Transformer).