Optimization of Mesmer, A Deep Learning Model for Cytoplasmic Segmentation Using LIVECell Dataset

Stanford CS231n Final Project Presentation

Liang Yi, Wei Da and Zhaoqiang Bai
June 4, 2022
Contents

- Background/Introduction
- Problem statement
- Dataset
- Methods
- Experiments & Analysis
- Conclusions & Future Work
- References
Background/Introduction

Cell Segmentation

- Cell segmentation plays a crucial role in understanding, diagnosing, and treating diseases.

- It remains challenging to accurately segment densely packed cells in cell membrane images using deep learning-based cell segmentation methods. Two critical challenges -
  
  1. Establish a scalable approach for generating large volumes of pixel-level training data in tissue images
  2. Set up an integrated deep learning pipeline that utilizes these data to achieve human-level performance.

- LIVECell and Mesmer are cutting-edge solutions to the above two challenges, respectively
Problem Statement

Objectives

- We apply Mesmer onto the LIVECell dataset in an attempt to contribute to the cell segmentation solutions
  
  › Combine the two pieces of the state of the arts and explore whether and how we can optimize the Mesmer algorithm so that it works well on the Livecell dataset.
  
  › Tuning backbone models, hyper-parameters, semantic transformations, among other knobs
  
  › Gather understandings of the best practices
Datasets

LIVECell

- We trained the Mesmer model using the LIVECell dataset.
  - 1.6 million cells
  - 5,239 high-quality, manually annotated and expert-validated images
  - 8 cell types: A172, BT-474, BV-2, Huh7, MCF7, SH-SY5Y, SkBr3 and SK-OV-3
  - diverse set of cell morphologies and culture densities.

- File structure organized in 2 subdirectories named “images” and “annotations”.
  - Image - “train val” and “test”
  - Annotations - the whole, single cell types and various size-split

- Data Processing and Augmentation
  - Processed by DeepCell API - SemanticDataGenerator
  - Data Augmentation
  - Data Normalization
Methods

Mesmer Network
- Mesmer is essentially a Feature Pyramid Network, consisting of:
  - A bottom-up pathway of a usual ConvNet backbone (could be ResNet, DenseNet, EfficientNet, etc.)
  - A top-down pathway of a feature pyramid
  - A lateral connection coupling the two above

- Two semantic heads for pixel-level predictions
  - “Pixel-wise transform” - predicts whether each pixel belongs to the cell interior, cell boundary or background
  - “Inner-distance transform” captures the distance of each pixel inside a cell to that cell’s centroid

- Watershed and post processing
  - Convert continuous predictions to discrete label images
Experiments & Analysis

Semantic Segmentation

- Combinations of the semantic transformations were tested to optimize the Mesmer trained by LIVECell dataset.
  - The combination of “inner-distance” with “pixel-wise” is significantly better than ones without “inner-distance”.
  - In agreement with the practices in Mesmer trained by TissueNet; Confirm the importance of the features captured by the distance of each pixel inside a cell to that cell’s centroid, which helps classify cell interior, cell boundary or background.
  - In addition Using “outer-distance” and/or “foreground/background” would not help Mesmer model training by the one-channel only, non-strained transparent cell images like those from LIVECell.
Experiments & Analysis

Cell Types
- We aim to evaluate the sensitivity of Mesmer model to the training that uses the images of each single cell types from LIVECell datasets.
- Mesmer trained by the images of A172 performed the best amongst all the cell types, and the one for BV2 was the worst.
- It might be helpful to generate more aggressive image augmentations from LIVECell so that Mesmer’s generalizability will be increased to cover more diversified cell morphology.
Experiments & Analysis

Backbone alternatives

- We attempted ResNet50, DenseNet201 and EfficientNetb7 as the backbone models

- DenseNet201 shows the best performance

- It might be due to the fact that the cell images share a lot of low-level features. The DenseNet is designed to strengthen feature propagation and encourage feature reuse which should be suitable in this case.

- The trend of the learning curve also suggest under-fitting and prediction accuracy could be improved by more extensive training (more epochs).
Conclusion & Future work

Conclusion

- This study used small sample dataset of LIVECell to train and optimize Mesmer. Our main conclusions are -
  
  - Combination of “inner-distance” and “pixel-wise” semantic heads leads to optimal cytoplasmic image segmentation.
  
  - Models trained by each single cell type images show various performances, with the one by A172 being the best.
  
  - DenseNet201 as Mesmer backbone provides the best result, outperforming the ResNet50 as used in the original Mesmer paper.

Future work

- More extensive search to identify the best combination of the hyperparameters.
- Test how other cutting-edge models, like Mask RNN + UNet, compares with Mesmer, after being trained and optimized using LIVECell datasets.
Reference

Key reference

LIVECell


Mesmer


Refer to our final project reports for other referenced literature