Network Optimization with Limited Bioimage Data for Robust Semantic Segmentation

Jihyeon Je[1], Alice Lucas PhD[2], David Stirling PhD[2], Beth Cimini PhD[2]

[1]Department of Biomedical Engineering, Duke University,[2]Cimini Lab, Broad Institute of MIT and Harvard

Motivation

- Bioimage segmentation is a common but challenging task, especially in dealing with complex biological properties due to closely packed structures, high variation in shapes and sizes, and complex morphology.
- While deep learning provides an opportunity to tackle previously intractable bioimage analysis tasks, large amounts of annotated data is generally required to train a reliable network.
- However, curation of a sizable and diverse training dataset is labor intensive and challenging, which limits the development of high accuracy segmentation models on such datasets.
- Approach: We present a robust training pipeline that utilizes
    1. Data augmentation and data regularization
    2. Feature extraction and representation learning
    3. Optimized network structure to efficiently utilize limited data

Dataset

- A 3D image stack of the C. elegans germline composed of 1024x1024 pixel images of 0.207 μm x 0.207 μm pixels.
- High variation in cell sizes and shapes.
- Ground truth annotations were created using 3dmod.

Background

- Conventional image analysis tools along with existing deep learning tools (CellPose[1], StarDist[2], NucleAlzer[3]) largely failed in performing segmentation tasks on a high-variation, sparse, and limited dataset, though we achieved moderate success with ilastik + multi-state CellProfiler[4] segmentation.

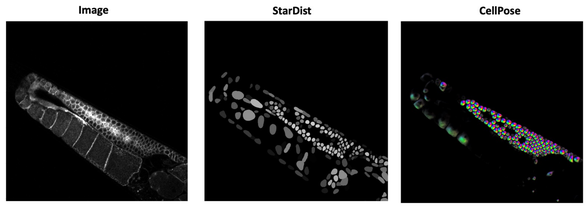

Figure 1. StarDist and CellPose segmentation results

Methods

- A single representative stack was chosen and 27 2D slices from the stack was split into 512x512 pixel tiles. Each tile was then divided into *training* (80%) and *validation* (20%) sets based on a stratified Monte Carlo sampling method.
- In order to account for the small dataset, a series of image augmentations were performed. In addition, L2 regularization was applied to prevent the model from overfitting.
- A deep learning model using a modified UNet framework utilizing a pre-trained MobileNetV2 (CNN)[5] model as the encoder and a Pix2Pix model (cGAN)[6] as the decoder was developed to segment cell, cell boundaries, and background.
- Encoder robustness was boosted via transfer learning with ImageNet. The model was trained to convergence for 65 epochs and the top 54 layers of the model were trained in isolation for greater performance.

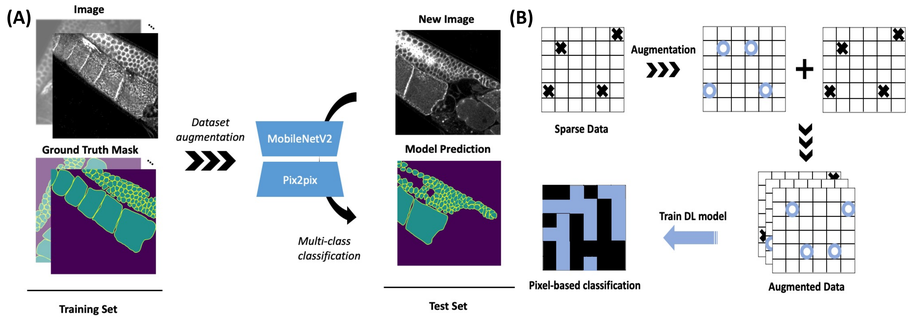

Figure 2. Deep learning framework to optimize segmentation performance with limited data. **(A)** Proposed workflow. **(B)** Simplified augmentation framework.

Results

Effective representation learning with limited labeled data is possible

- Model achieved 0.985 accuracy in training set and 0.973 accuracy in validation set.
- The model with the lowest validation loss was tested on an unseen, independent set of 2D slices from a separate stack.
- Model performance was evaluated on visual and performance metrics and confirmed that the model is capable of multi-class segmentation even on a highly sparse, unequally labeled, limited dataset.
- CellProfiler pipeline was utilized to post-process the model prediction results

Results, cont'd

3D analysis allows quantitative measurement extraction

- 2D segmentations were combined across multiple planes and re-segmented into 3D objects. The reconstructed 3D stack is visualized utilizing a napari[7] plugin.
- The pipeline can identify the total number of cell compartments and extract 3D volumetric dimensionalities from 2D segmentation results.

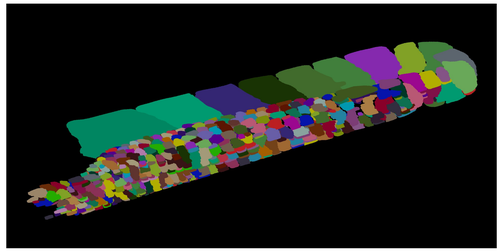

Figure 3. 3D reconstruction of cellular compartments

Conclusions

- We applied data augmentation strategies to express class invariances along with feature extraction techniques and optimized the network structure to effectively utilize extremely small dataset.
- We demonstrated that deep learning with adequate framework can be effectively used to tackle challenging computer vision analysis tasks, especially when training data is sparse.

Future Directions

- Incorporate learning representations to reweight weak supervision and to enforce structured prior knowledge
- Determine the optimal training set size necessary to achieve the desired performance

Acknowledgements

Michael Werner PhD[3], Jack Linehan[3], Amy Shaub Maddox PhD[3]

[3]Department of Biology, UNC-Chapel Hill

This work was supported by the Center for Open Bioimage Analysis (NIH P41 GM135019)

[1] Stringer, C., Wang, T., Michaelos, M. *et al.* Cellpose: a generalist algorithm for cellular segmentation. *Nat Methods* **18,** 100–106 (2021). https://doi.org/10.1038/s41592-020-01018-x.
[2] Schmidt, Uwe et al. "Cell Detection with Star-convex Polygons." *ArXiv* abs/1806.03535 (2018): n. pag.
[3] Lucas, Alice M., et al. "Open-Source Deep-Learning Software for Bioimage Segmentation." *Molecular Biology of the Cell*, edited by Diane Lidke, vol. 32, no. 9, Apr. 2021, pp. 823–29. *DOI.org (Crossref)*, https://doi.org/10.1091/mbc.E20-10-0660.
[4] McQuin C, Goodman A, Chernyshev V, Kamentsky L, Cimini BA, Karhohs KW, Doan M, Ding L, Rafelski SM, Thirstrup D, Wiegraebe W, Singh S, Becker T, Caicedo JC, Carpenter AE (2018). CellProfiler 3.0: Next-generation image processing for biology. PLoS Biol. 16(7):e2005970 / doi. PMID: 29969450 (Research article)
[5] Sandler, Mark, et al. "MobileNetV2: Inverted Residuals and Linear Bottlenecks." *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, 2018, pp. 4510–20. *DOI.org (Crossref)*, https://doi.org/10.1109/CVPR.2018.00474.
[6] Isola, Phillip, et al. "Image-to-Image Translation with Conditional Adversarial Networks." *ArXiv:1611.07004 [Cs]*, Nov. 2018. *arXiv.org*, http://arxiv.org/abs/1611.07004.
[7] napari contributors (2019). napari: a multi-dimensional image viewer for python. doi:10.5281/zenodo.3555620