Team name:

The Cellulites

Group Members:
Spencer Vore (svore3@gatech.edu)
Madhuri Vihani (mvihani3@gatech.edu)
Joe Laniado (jlaniado3@gatech.edu)

Project Title:
Using transfer learning on multilabel image classification of tissue images.

Project summary (4-5+ sentences). Fill in your problem and background/motivation (why do you want to solve it? Why is it interesting?). This should provide some detail (do not just say "I'll be working on object detection")

The goal of this project is to quantify the number of cells from microscopic images by utilizing instance segmentation. Getting the number of cells is a fundamental step in many biomedical studies, and through automation with computer vision, we can help alleviate the researchers of this monotonous procedure. Solving this problem can also be used as a building block for other types of instance segmentation. Cell segmentation is typically an easier task for object detection, but a more challenging task for instance segmentation, because cells are more homogenous in their features, like size, color, and texture. It is advantageous to enhance edge detection in cytoplasm cells that touch each other, because it is common for two different cells to be confused as one.[1]

What you will do (Approach, 4-5+ sentences) - Be specific about what you will implement and what existing code you will use. Describe what you plan to implement or the experiments you might try, etc. Again, provide sufficient information describing exactly what you will do. One of the key things to note is that just downloading code and running it on a dataset is not sufficient for a description or a project! Some thorough implementation, analysis, theory, etc. must be done for the project.

To solve this problem, we plan to use the Tissue Net Dataset and the DSB2018 Dataset (additional details later). Using additional Python libraries, such as OpenCV, our approach will consist of the following tasks:

1) We will generate a binary segmentation of the cells based on the Tissue Net dataset.
2) We will generate a multiclass segmentation involving the watershed method where the background of the image will have a pixel value of 0, edges will have a pixel value of 1,

1. [1] Al-Kofahi, Y., Zaltsman, A., Graves, R. *et al.* A deep learning-based algorithm for 2-D cell segmentation in microscopy images. *BMC Bioinformatics* **19**, 365 (2018). https://doi.org/10.1186/s12859-018-2375-z

and the center of the cell will have a pixel value of 2. We will be exploring how the thickness of the edges in the ground truth might affect the outcome of the model.

3) The next step will be to try to recreate the methodology used by Bai (2017)[2] for instance-segmentation. More specifically, the Deep Watershed Transform Network and the Direction Network by utilizing the Euclidean distance transform method.
4) We will transfer the weights of binary segmentation to the multiclass segmentation (Watershed).

Future research might include pursuing the following concepts, if possible:

5) If the edges of the cells are blurred and cannot be differentiated, then we can also analyze the vector field of the microscopic images. If the angle between two neighboring pixels is greater than 90 degrees, then we can say that these two pixels belong to two different cells.
6) A 1:1 mapping of nuclear cells to cytoplasm cells. The number of nuclear cells should match the number of cytoplasmic cells.

Resources / Related Work & Papers (4-5+ sentences). What is the state of art for this problem? Note that it is perfectly fine for this project to implement approaches that already exist. This part should show you have done some research into what approaches exist.

Our approach combines preexisting research and applies it in a new context related to cell segmentation. Much of our approach will be recreating existing state-of-the-art methodologies from many different sources. Some of them are listed below:

1. *Greenwald, Noah F., et al. "Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning."*[3]
    a. This paper represents the original paper that created the Tissue Net dataset to do research in cell segmentation. They also created a deep learning-enabled segmentation algorithm that performs nuclear and whole-cell segmentation in tissue imaging data called Mesmer.
2. *M. Bai and R. Urtasun, "Deep watershed transform for instance segmentation"*

[2] Bai, M., Urtasun, R. Deep watershed transform for instance segmentation. ArXiv:1611.08303v2 (2017). https://doi.org/10.48550/arXiv.1611.08303

[3] *Greenwald, Noah F., et al. "Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning." Nature biotechnology 40.4 (2022): 555-565.*

    a. Proposes a new end-to-end approach, for instance segmentation using convolutional neural networks that challenges a lot of state-of-the-art modern methodologies. It also implements aspects of the watershed transform and combines it with deep learning to accomplish their objectives. This is the paper that we will be trying to replicate.

3. *Lempitsky, Victor, and Andrew Zisserman. "Learning to count objects in images."*[4]
    a. A supervised learning approach on how to count objects in images, a cell counting example is mentioned on the paper that can be applicable to our implementation.
4. *"Cellpose: a generalist algorithm for cellular segmentation"*[5]
    a. A deep learning-based segmentation method, which can precisely segment cells from a wide range of image types and does not require model retraining or parameter adjustments.
5. *M. Majurski et al., "Cell image segmentation using generative adversarial networks transfer learning and augmentations"*
    a. Addresses the problem of segmenting cell contours by using convolutional neural networks and compares the accuracy gains of the results using GANs.

A lot of other methodologies currently exist that have an approach at tackling the problem of cell segmentation, but these are some of the most relevant to our project's objective.

Datasets (Provide a link to the dataset). This is crucial! Deep learning is data-driven, so what datasets you use is crucial. One of the key things is to make sure you do not try to create and especially annotate your own data! Otherwise, the project will be taken over by this.

As stated before in the previous sections, the two main datasets we will be working with are: DSB018 and Tissue Net. Below are some links to access samples of both:

1. Tissues Net: https://datasets.deepcell.org/
2. DSB2018 Kaggle: https://www.kaggle.com/competitions/data-science-bowl-2018/overview

Tissue Net was created by Greenwald et. Al[3] for their cell segmentation experiments. It contains over 1 million paired annotations covering whole cells and nuclei from nine different

---

[4] *Lempitsky, Victor, and Andrew Zisserman. "Learning to count objects in images." Advances in neural information processing systems 23 (2010).* https://proceedings.neurips.cc/paper/2010/file/fe73f687e5bc5280214e0486b273a5f9-Paper.pdf

[5] 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

organs imaged using six different platforms. The Kaggle dataset also contains cell imaging that were used in 2018 for a competition. Both datasets are freely accessible.

References:

1. Al-Kofahi, Y., Zaltsman, A., Graves, R. *et al.* A deep learning-based algorithm for 2-D cell segmentation in microscopy images. *BMC Bioinformatics* **19**, 365 (2018). https://doi.org/10.1186/s12859-018-2375-z
2. Bai, M., Urtasun, R. Deep watershed transform for instance segmentation. ArXiv:1611.08303v2 (2017). https://doi.org/10.48550/arXiv.1611.08303
3. *Greenwald, Noah F., et al. "Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning." Nature biotechnology 40.4 (2022): 555-565.*
4. *Lempitsky, Victor, and Andrew Zisserman. "Learning to count objects in images." Advances in neural information processing systems 23 (2010).* https://proceedings.neurips.cc/paper/2010/file/fe73f687e5bc5280214e0486b273a5f9-Paper.pdf
5. 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