

Cell Nuclei and Lipid Droplets Quantification in Stimulated Raman Images

Adiel Felsen^{*1}, Yuhao Yuan^{*2}, Nikolas Burzynski¹, David Reitano¹, Zhibo Wang², Khalid A. Sethi³, Faike Lu², and Kenneth Chiu^{1†}

Department of {¹Computer Science, ²Biomedical Engineering},

Binghamton University, State University of New York, Binghamton, NY 13902, USA

³Department of Neurosurgery, United Health Services, Johnson City, NY 13790, USA

^{*}These authors contributed equally

[†]kchiu@binghamton.edu

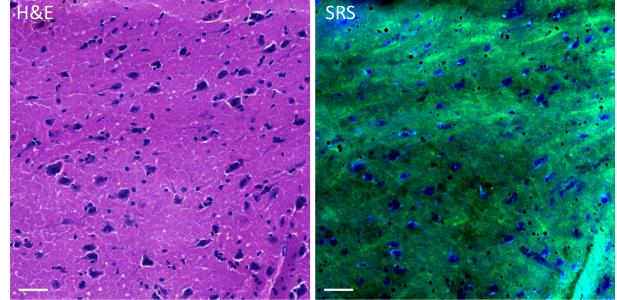
Abstract—Stimulated Raman scattering (SRS) microscopy is a stain-free, laser-scanning imaging technology that allows for rapid chemical imaging at high-resolution. When developing cancer diagnosis based on cellular and tissue pathology, it is critical to understand the number, size, and density of the cell nuclei, as well as other metabolic features, such as the lipid droplets. In our research, we compare the U-Net and Mask R-CNN convolutional neural network architectures to segment cell nuclei from SRS images of cultured cancer cells. We also use a modified version of U-Net to identify the centroids of nuclei and lipid droplets. Combining these centroids with a segmentation, we can generate a Voronoi diagram to estimate the size of each nucleus and lipid droplet. Future work will focus on applying these methods to identify and segment various cellular structure in both cells and human cancer tissues with SRS imaging.

I. INTRODUCTION

Hematoxylin and eosin (H&E) staining is the standard pathological approach for clinical cancer diagnosis [1]. During brain surgery, tissue specimens are sent to the pathology division for rapid diagnosis. The tissue is frozen, sliced and stained with the H&E method to highlight the tissue morphological structures, especially the cell nuclei. The turn around time of this procedure is about 30 minutes, and is only performed a limited number of times during surgery, which limits its extensive applications for intraoperative surgical guidance.

Stimulated Raman scattering (SRS) [2] is an emerging approach to pathological imaging. SRS is based on Raman scattering, which reveals chemical information through light-matter interaction. Two major Raman shifts, 2854 cm^{-1} (CH_2 stretching) and 2930 cm^{-1} (CH_3 stretching) are usually used to image lipids and proteins in cells or tissue, respectively. Cell nuclei are highlighted by weighting, subtracting and overlaying the two channels [3]. However, this simple approach often leads to unreliable results, with unrelated structures identified as cell nuclei. Even after significant processing, lipid/protein two-color SRS imaging fails to achieve the same level of clarity as H&E staining, hindering its further development for rapid cancer diagnosis.

Quantification and segmentation of cell parts are vital for cancer diagnosis based on pathological imaging. Specifically,



Source: [4]

Fig. 1: Comparison of stain-free SRS image and chemically-stained H&E image of a human brain tumor frozen tissue section. Pseudocolor green: lipids; blue: proteins. Scale bar: $50\text{ }\mu\text{m}$.

the size, distribution, and density of the cell nuclei provide the most important information for diagnostic decision-making. Additionally, the distribution of other cell structures such as the intracellular lipid droplets are also of significance for cancer biology studies. Due to the high clarity of H&E stained cells, simple approaches such as thresholding are sufficient to determine these statistics. However, these techniques cannot be directly used on SRS images, which exhibit overlapping and weak chemical contrasts.

In recent years, significant advances have been made in the field of image segmentation. Semantic segmentation partitions an image on a pixel-by-pixel level, while instance segmentation identifies objects at an instance level. In this paper we focus on two particular networks: U-Net [5] and Mask R-CNN [6]. U-Net is an image-to-image technique designed for semantic segmentation of biomedical images. Detectron2 [7] is a popular implementation of the Mask R-CNN architecture, which performs instance segmentation. Both approaches use convolutional models to tackle segmentation problems. In this paper, we compare the use of these networks for the segmentation and quantification of cell nuclei and lipid droplets in SRS images of cultured cancer cells.

II. METHODOLOGY

A. Segmentation with U-Net

U-Net and variations of it have been used previously for cell nuclei segmentation in stained pathology images [8] [9]. We attempt to use this model for the stain-free SRS images. As input to the model, we gathered 150 1024×1024 SRS images with two channels: a lipid channel using a 2854 cm⁻¹ Raman shift and a protein channel using a 2930 cm⁻¹ shift. Note, while the cells generally sit in the same plane, there is some depth to these cultures which can lead to overlapping cells. Our labels were generated via thresholding of the two-photon fluorescence images of cell nuclei stained with Hoechst 33342. Since fluorescence imaging has a slightly larger depth visibility than SRS, there are some cells visible in the fluorescence that are not visible in the SRS images. These incorrect labels were removed before training and testing.

Our data was divided into training and testing sets. The training set contains 120 (80%) randomly selected images, while our testing set contains the remaining 30 images. For training, each image was divided evenly into four 512×512 segments. Additionally, random transformations were performed on the training data. This includes random horizontal and vertical flipping, as well as a 10% brightness, saturation, and contrast jitter. Jitter is generated using the default PyTorch [10] color jitter transformation.

For the testing data, 92 pixels of each edge of each image are mirrored. Then, each image is divided into 9 512×512 segments. After passing through our network, these segments are compiled together, with each segment consisting of 1/9 of the final compilation. This relatively complex process is necessary due to the size reduction that occurs during U-Net's convolutional steps. For a 512×512 input image, only the center 328×328 pixels are labeled. A similar process is described in the original U-Net paper to account for this shrinking issue.

While U-Net is capable of segmenting cell nuclei, it cannot distinguish between touching or overlapping nuclei. We experimented altering our labels by creating a 1-pixel wide black line between any connected labels. However, U-Net still failed to reliably create a separation between touching nuclei.

B. Mask R-CNN to Identify Overlapping Nuclei

We next used Mask R-CNN to account for these overlapping nuclei. The labels for Mask R-CNN were similar to those for U-Net, with a notable difference: any overlapping labels were manually circled using ImageJ and separated into different image stack layers. Then, each distinct label was outlined by a polygon, which we used as our final label for Detectron2. Unlike U-Net, a single pixel can be associated with more than a single nucleus.

We implemented our Mask R-CNN instance segmentation algorithm using Detectron2 from Facebook Research. We used the same dataset, and separated our training and testing sets in the exact same way. Each 1024×1024 image is fed into the network as-is, without any subdivision or mirroring. Our

results show that Mask R-CNN was fully capable of accurately labeling overlapping nuclei. However, the manual labeling of overlapping object was labor intensive and it is impractical to use this network on the micro lipid droplets.

C. U-Net for Centroid Detection

In an attempt to rival the accuracy of Mask R-CNN using U-Net, we focused on centroid detection. Given the centroid of each object, we used a Voronoi tessellation to separate touching objects. A Voronoi tessellation partitions a plane into regions based on distance to a set of objects.

Our inspiration for using U-Net for centroid detection was a project originally designed to count bees leaving their hive [11]. We use the same U-Net structure as we had for nuclei segmentation, but provide new labels. Instead of labeling the entire area of a nucleus, we just label the centroid of each nucleus. Our final label is a binary image, with a single pixel to represent the centroid of each object.

Loss is determined by cross-entropy loss (CEL) between the output of the network and the label. Due to the much more drastic class imbalance positive labels are given significantly higher weight in the loss function. However, with only single-pixel labels, small errors leads to a drastic increase in the loss. This incentivizes the network to label large regions around each centroid to minimize the number of false negatives. This proved to be problematic, since these large regions occasionally overlap, so individual nuclei could not be reliably identified.

To coerce the network to shrink its predictions, we created new labels which mark the distance of each pixel to the nearest centroid. Any positive pixel contributes to the loss based on its distance from a centroid. The final loss for some output pixel o with labels c and d is:

$$\text{CEL}(o, c) + \text{ReLU}(o) \cdot d$$

Segmentation of the lipid droplets proved to be a more challenging task due to their small size. A single lipid droplet could be represented by only a few pixels in the original SRS image. To account for this issue, we artificially increased the resolution of our images using bilinear interpolation. Following this minor change, the accuracy of our network drastically increased. This demonstrates that the network has the capability to count smaller structures.

III. RESULTS

Figure 2 shows the cell nuclei segmentation results using both U-Net and Mask R-CNN. We found that semantic segmentation with U-Net could not distinguish the two adjacent cell nuclei, while instance segmentation with Mask R-CNN could distinguish the overlapping nuclei.

Figure 3 shows examples of nuclei and lipid counting with out modified U-Net algorithm. Both are images from a testing dataset, which the network has never seen. Note that the lipid and nuclei counting models were trained on different datasets.

Figure 4 shows a Voronoi diagram generated by a combination of U-Net nuclei counting and segmentation. In the center

of the image, two overlapping nuclei are properly separated by this algorithm. A few errors are visible. On the top left, a single nucleus is mistakenly divided into two parts. On the bottom middle, two nuclei are not identified and are not separated.

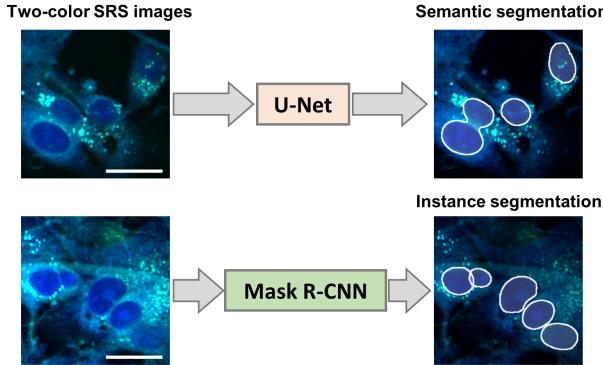


Fig. 2: Cell nuclei segmentation results using U-Net (above) and Mask R-CNN (below), respectively. Scale bar, 20 μ m

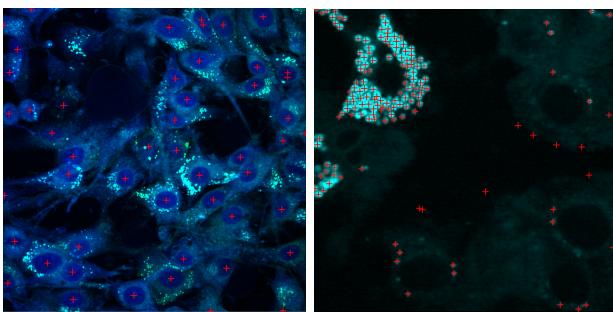


Fig. 3: Cell nuclei (left) and lipid droplets (right) counting using U-Net with the modified loss function. Red crosses indicate the centers detected by the network.

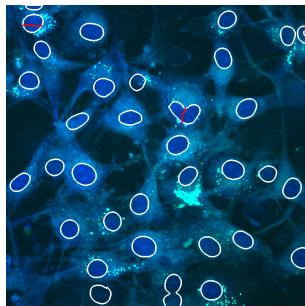


Fig. 4: Voronoi diagram partitioning segmented nuclei using the U-Net nuclei counting algorithm to denote distinct objects. Red lines show each partition performed by this algorithm.

IV. DISCUSSION

Overall, we found that Mask R-CNN provides accurate and true-to-reality results. It was capable of accurately labeling overlapping objects. However, the manual labeling required is more labor-intensive than that for U-Net. Also, it may be unfeasible to implement Mask R-CNN for smaller structures, like lipid droplets.

U-Net segmentation combined with U-Net centroid detection is a more flexible approach, requiring only a moderate amount of manual labeling. This approach allows for a reasonable estimation of the size, count, and distribution of cell nuclei and lipid droplets.

We did notice a limitation with our centroid detection technique. Due to the way our loss function determines cost, two disconnected regions labeling the same object are often allowed to exist. As of now, labeling two disconnected pixels 1 away from centroid has same loss as labeling two connected pixels the same distance away. This can lead to "double counting", which interferes with our final cell statistics. This is clear in figure 4, where a single nucleus is mistakenly identified as two separate nuclei. We have considered punishing the network more severely when it labels pixels that are disconnected from a centroid. However, we have yet to find an effective and differentiable method to implement this in our loss function.

V. CONCLUSION

We will continue to refine our approach to cell nuclei segmentation and lipid droplets counting. We hope to greatly reduce the "double counting" issue we encountered. In the future we plan on applying this technique to other cellular structures. Additionally, we will also apply these networks to SRS images of human cancer tissues.

REFERENCES

- [1] A. H. Fischer, K. A. Jacobson, J. Rose, and R. Zeller, "Hematoxylin and eosin staining of tissue and cell sections," *Cold spring harbor protocols*, vol. 2008, no. 5, pp. pdb-prot4986, 2008.
- [2] C. W. Freudiger, W. Min, B. G. Saar, S. Lu, G. R. Holtom, C. He, J. C. Tsai, J. X. Kang, and X. S. Xie, "Label-free biomedical imaging with high sensitivity by stimulated raman scattering microscopy," *Science*, vol. 322, no. 5909, pp. 1857–1861, 2008.
- [3] F.-K. Lu, D. Calligaris, O. I. Olubiyi, I. Norton, W. Yang, S. Santagata, X. S. Xie, A. J. Golby, and N. Y. Agar, "Label-free neurosurgical pathology with stimulated raman imaging," *Cancer research*, vol. 76, no. 12, pp. 3451–3462, 2016.
- [4] F.-K. Lu, S. Santagata, X. S. Xie, A. J. Golby, and N. Y. R. Agar, "Label-Free Neurosurgical Pathology with Stimulated Raman Imaging: A Dataset," 2016. [Online]. Available: <https://doi.org/10.7910/DVN/EZW4EK>
- [5] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," 2015.
- [6] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," 2018.
- [7] Y. Wu, A. Kirillov, F. Massa, W.-Y. Lo, and R. Girshick, "Detectron2," <https://github.com/facebookresearch/detectron2>, 2019.
- [8] Z. Zeng, W. Xie, Y. Zhang, and Y. Lu, "Ric-unet: An improved neural network based on unet for nuclei segmentation in histology images," *Ieee Access*, vol. 7, pp. 21 420–21 428, 2019.
- [9] F. Long, "Microscopy cell nuclei segmentation with enhanced u-net," *BMC bioinformatics*, vol. 21, no. 1, pp. 1–12, 2020.
- [10] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems* 32, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, Eds. Curran Associates, Inc., 2019, pp. 8024–8035. [Online]. Available: <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>
- [11] M. Kelcey, "counting bees on a rasp pi with a conv net," May 2018. [Online]. Available: <http://matpalm.com/blog/countingbees/>