

StarDist: U-Net based CNNs on cell boundary detection of electron microscopy

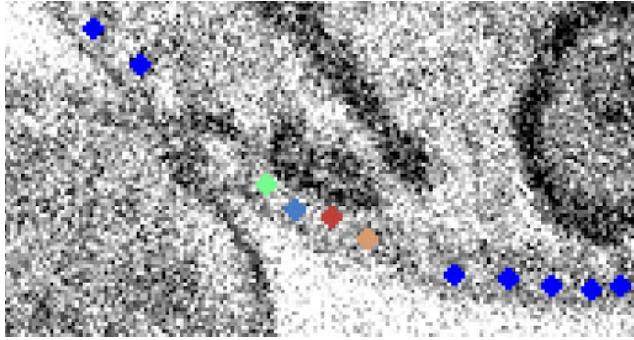


Figure 1. A zoomed-in crop of a cross-segment of the parasite. Blue dots represent microtubules, the other four colored dots represent MtQs.

Abstract

In this project, we explored applying *stardist*, a UNet-based CNN framework specifically developed for cell boundary segmentation, to a task that's common to biological research: The detection and differentiation of cell structures, in this case microtubules and MtQs. We experimented with pretrained models as well as our own models and employed extra preprocessing to make our model more robust. In future studies, we wish to further diversify our data sources as well as pre-segment images for cell count to improve our model performance.

1. Introduction

The de Graffenried lab studies the variety of cellular morphologies that protozoan parasites exhibit to survive within hosts, with the goal of targeting pathways that impact morphology for drug design to treat many neglected tropical diseases. One of these parasites is *Trypanosoma brucei*. Electron microscopy images of cross-segments of *T. brucei* reveal general microtubules and microtubule quartets (MtQs) that form the structure of the parasite, important for better understanding how they undergo morphological changes [6].

Currently these images are painstakingly annotated by hand. Each image takes around thirty minutes to annotate. An automated pipeline to annotate the images would greatly assist both the de Graffenried lab and other labs looking

into similar parasites with imaging data. This is a difficult problem due to the presence of many other features in the image, the low resolution, and the lack of similar datasets to train models on.

Our group planned to use *stardist* [5] as well as a self written CNN on an electron microscopy image dataset provided by the de Graffenried lab. Our goal is to detect and annotate microtubules that form the structure of the parasite, and then track their appearance and disappearance across cross-sections of the parasite. Figure 2 shows an example of what these images look like unannotated and annotated.

The problem that we are solving is difficult because microtubules and MtQs are not instantly differentiable in the electron microscopy images and the training data we have only consists of crude hand-labelled dots. However, by solving this problem we hope to liberate fellow researchers from the toil of having to hand-label microtubules and pave the way to more sophisticated cell detection/segmentation models.

2. Related Work

Our project is theoretically based on the foundational work of Ronneberger et al [4] where they designed and implemented U-Net, a CNN specifically for cell segmentation characterised by its downsampling and upsampling pipeline as well as its skip connections. Huang et al [2] and Oktay et al [3] also detailed the technicalities of more modern U-Net variations. Schmidt et al [5] and Weigert et al [7] were integral in our understanding of how *stardist*, the library we work with, built on U-Net, the additional pipelines and preprocessing as well as performance optimization techniques.

3. Method

Stardist was chosen due to its design for blob-like images, which these microtubules fall into, its good performance on such blobs, and its ease of use, good documentation, and good support on the image.sc forums.

Stardist itself is based on UNet [1], with an additional component of non-maximal suppression and a generated segmentation pixel mask to predict object center probabilities. In brief, UNet is a CNN that downsamples the image to generate an overview of the images pixels, then upsamples to correlate spatial information learned from the downsampling.

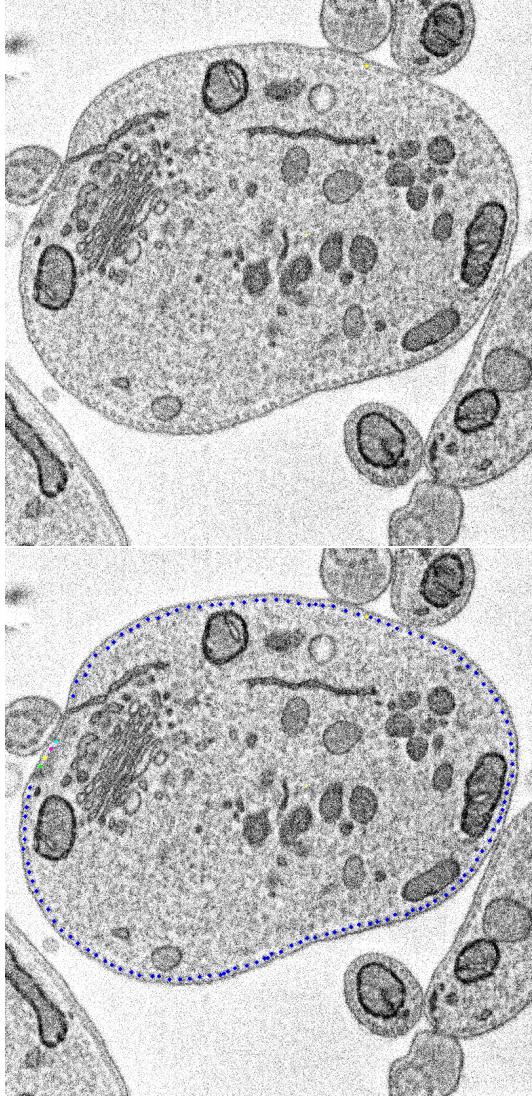


Figure 2. Unannotated cross-section and annotated cross-section. Blue dots represent microtubules, other colored dots represent a particular region called the MtQ that is also of interest.

The skip connections make sure that earlier downsampling information is joined to later upsampling information and improves accuracy.

It does this by finding all polygons that fit onto the number of rays (a user-provided parameter), keeping the polygon with highest probability out of all polygons that overlap, and then removing all the rest with an intersection over union greater than a pre-selected threshold. See 4 for more information.

To create a more effective dataset, our team worked on several preprocessing pipelines including image alignment, image random cropping and image jittering (see 9, 8, 6). These pipelines decreased the extent to which we might

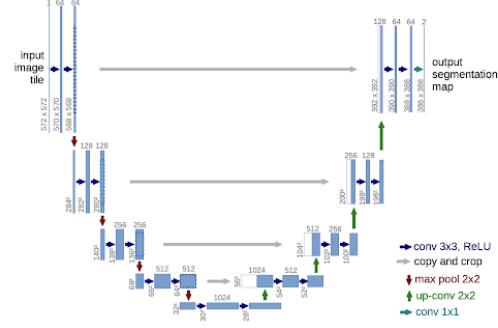


Figure 3. Diagram of how unet works. The grey arrows represent the skip connections.

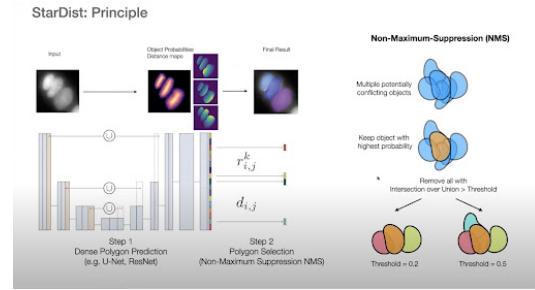


Figure 4. Diagram of how stardist works. On the right, there are multiple potentially conflicting objects. Out of the one with the highest probability, all that overlap greater than the threshold are removed.



Figure 5. Stardist pretrained model sample cell segmentation: Apparent underperformance without knowledge of the number of cells

overfit the training data. We also explored other models like *2D_versatile_fluo* (see 5), a pretrained model for generating object segmentation masks. However, the pretrained model assumes that the training image contains multiple cells, while object segmentation masks cannot differentiate between microtubules and MtQs, suggesting more informative datasets and labelling information in future studies.

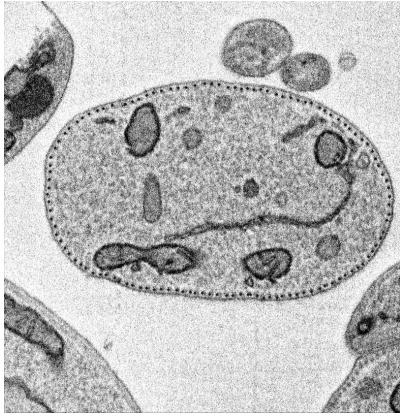


Figure 6. Aligned image example

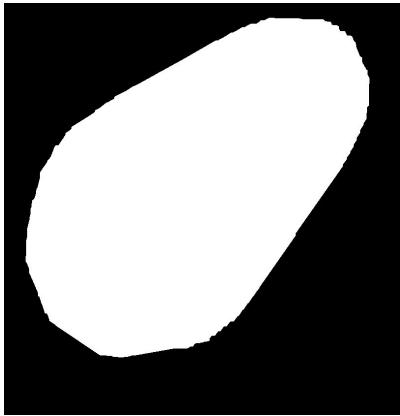


Figure 7. Example of generated object segmentation bit mask:
White pixels represent the segmented cell body

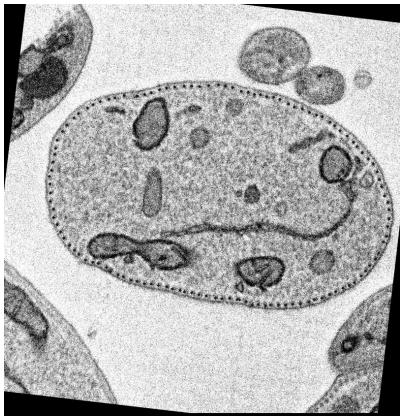


Figure 8. Jittered image example

4. Results

Our TA suggested that we run stardist on a single image and overfit it at first to see if stardist would be suitable for our

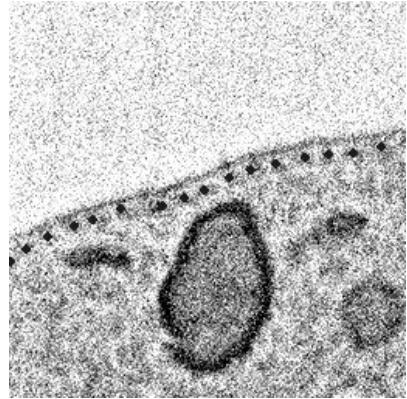


Figure 9. Cropped image with cropping region randomly generated

problem. It appeared that it did quite well on the single image [10](#). The measures used were precision, recall, accuracy, etc. over Intersection over Union (IoU) thresholds, where 0 indicates that any overlap of even 1 pixel of prediction and true label is considered correct, and 1 indicates that the prediction and true label must overlap exactly to be considered correct.

In the end we were able to run the model on the images we had as well, however here the results were lackluster. [11](#). It seems that the model performs well when there's only one cell in the image and fails when there is an unknown number of cells. We did not get around to applying the multi-class version of stardist to the results as the base version was not sufficiently accurate, but we would expect further difficulties with differentiating between microtubules and MTQs there as we suspect there isn't enough information in the single electron microscopy for the model to fully learn the difference between them.

4.1. Technical Discussion

We had a lot of difficulty getting accuracy into the higher single digits. This could decidedly be an issue in our methodology and could be improved. However, part of this is due to how we are calculating accuracy: that being, the combined accuracy of every single labeling of a microtubule and MTQ where 100% accuracy is only calculated where there is pixel-perfect overlap of the computer's prediction and the true label. As the professor suggested, in future explorations we will experiment with other accuracy measures like distance to correct labels rather than intersection over union.

5. Conclusion

As seen from the results, we were able to generate sufficient labeling on some cell images while failing to do so in other images. The accuracy we obtained did not do a sufficient job effectively representing whether the project achieved its functionality. Different accuracy measures are

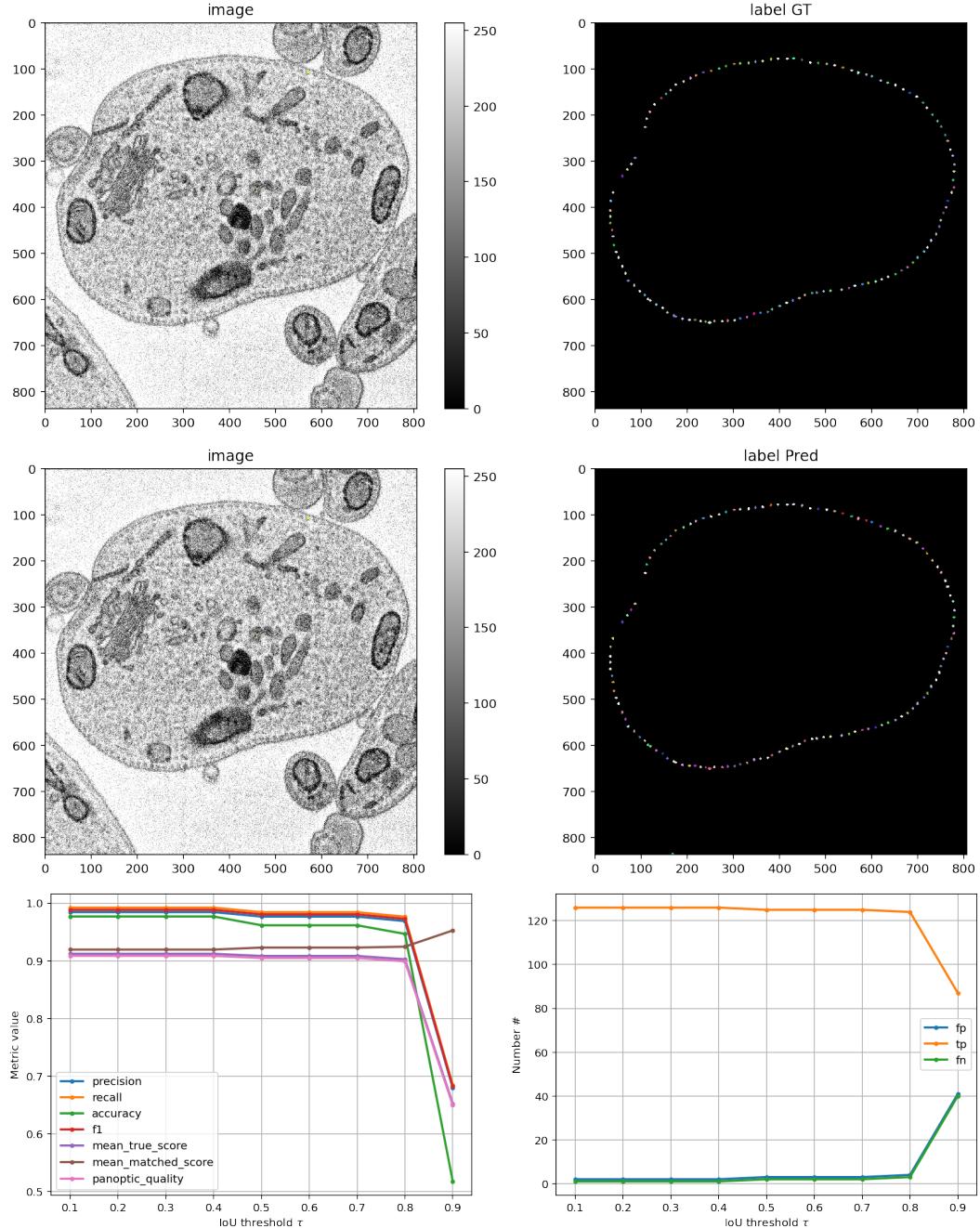


Figure 10. *Top:* Correct labels for single image. *Middle:* Predicted labels for model trained on that single image. *Bottom:* Accuracy results and false positives/true positives/false negatives numer across different IoU thresholds.

required to have a more precise analysis on the results, and while some predictions are accurate enough for lab use, the consistency leaves much to be desired. Regardless, more improvements need to be made so that the project can convincingly provide labeling that does not require much human effort.

References

- [1] Thorsten Falk, Dominic Mai, Robert Bensch, Özgün Çiçek, Ahmed Abdulkadir, Yassine Marrakchi, Anton Böhm, Jan Deubner, Zoe Jäckel, Katharina Seiwald, Alexander Dovzhenko, Olaf Tietz, Cristina Dal Bosco, Sean Walsh, Deniz Saltukoglu, Tuan Leng Tay, Marco Prinz, Klaus Palme, Matias Simons, Ilka Diester, Thomas Brox, and Olaf Ronneberger. U-net: deep learning for cell counting, detection, and mor-

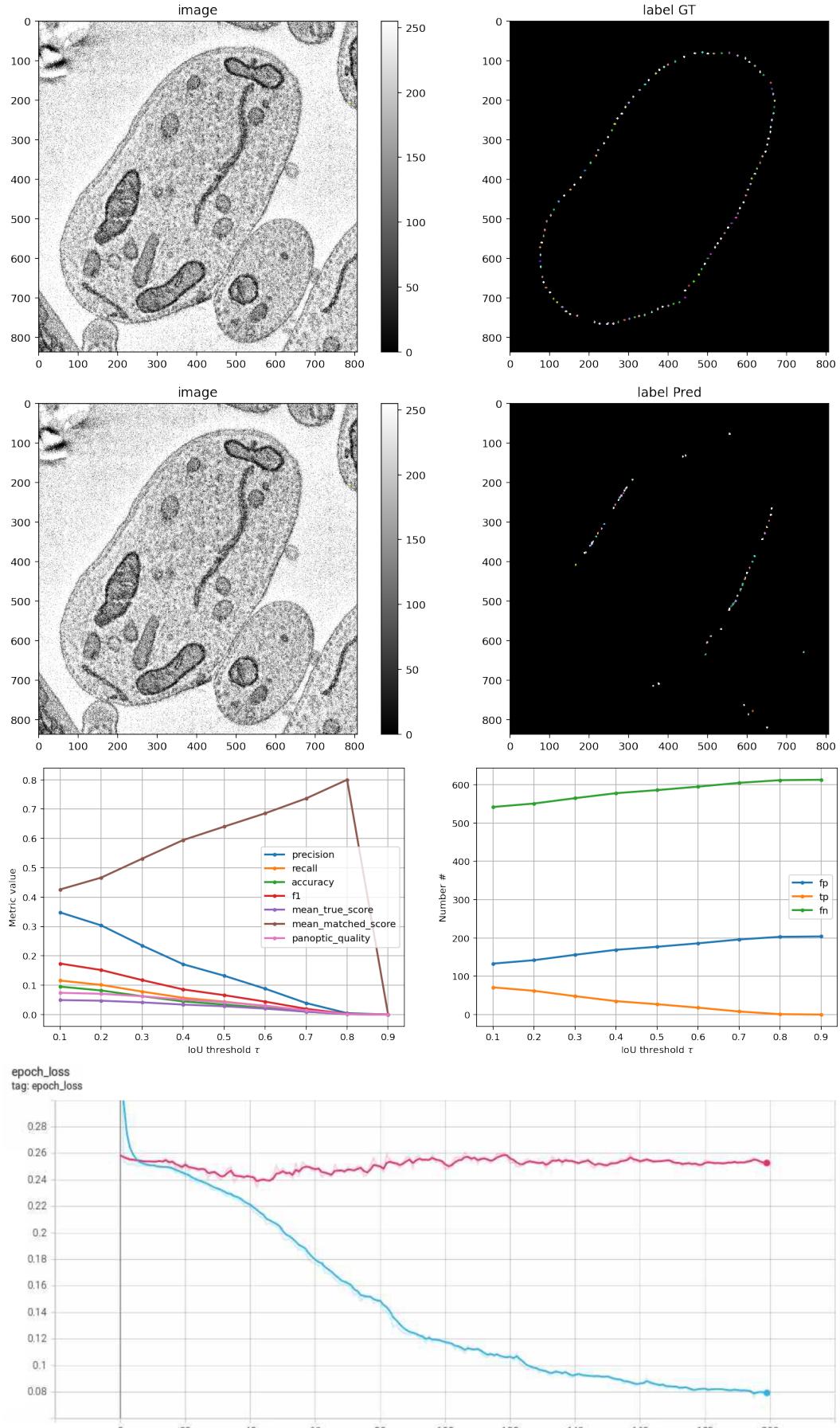


Figure 11. *Top:* Correct labels for example image. *Second:* Predicted labels for example image. *Third:* Accuracy measures and false/true/positives for range of IoU thresholds. *Fourth:* Graph where red is validation loss, blue is training loss, across 200 epochs.

- photometry. *Nature Methods*, 16, 2019. 1
- [2] Huimin Huang, Lanfen Lin, Ruofeng Tong, Hongjie Hu, Qiaowei Zhang, Yutaro Iwamoto, Xianhua Han, Yen-Wei Chen, and Jian Wu. Unet 3+: A full-scale connected unet for medical image segmentation. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1055–1059. IEEE, 2020. 1
 - [3] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Matthias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y Hammerla, Bernhard Kainz, et al. Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*, 2018. 1
 - [4] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III* 18, pages 234–241. Springer, 2015. 1
 - [5] Uwe Schmidt, Martin Weigert, Coleman Broaddus, and Gene Myers. Cell Detection with Star-convex Polygons. jun 2018. 1
 - [6] Amy N. Sinclair and Christopher L. de Graffenreid. More than microtubules: The structure and function of the subpellicular array in trypanosomatids, 2019. 1
 - [7] Martin Weigert, Uwe Schmidt, Robert Haase, Ko Sugawara, and Gene Myers. Star-convex polyhedra for 3d object detection and segmentation in microscopy. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pages 3666–3673, 2020. 1