Automatic Cell Detection for Plant-Based Xylem Tissue Applications

Sean Wu
Seaver College
Pepperdine University
Malibu, CA 90265
Sean.wu@pepperdine.edu

Timothy Chen
Seaver College
Pepperdine University
Malibu, CA 90265
Timothy.chen5@pepperdine.edu

***Abstract*—The quantification of the cells in a plant's xylem tissue is a task that can provide valuable information about the plant's anatomy, development, and immune system. However, this measurement process is usually undertaken in a manual fashion, where botanists will record and count thousands of cell elements in each cross section. In this study we automated this evaluation of xylem cross sectional images by implementing a iterative Deep Neural Network object detection algorithm that counts and regionalizes every cell type. An extention of this object detection is the measurement (diameter, width) of the vessels, which is completed by measuring pixel ratio of the cell type in respect to the size of the image.**

I. Introduction

The analysis of a plant's xylem tissue produced by a light micrograph imaging technique can give botanists crucial insights to a plant's hydraulic efficiency, internal strength, and a variety of other traits that are essential for analyzing a plant's overall health. The xylem tissue of a plant is composed of three unique cell types that we are interested in detecting.

- Fiber: An elongated cell type which is primarily responsible for the stem or root of a plants structure. The cell type that is most commonly found in our dataset.
- Parenchyma: A shorter cell type with darker features that is responsible for the temporary transport of starch material and water. Visually, the parenchyma are usually in column formations.
- Vessel: A cell type that features a large radius which is responsible for the transport of water. Typically the largest cell types in the images, with a water-like substance inside.

Although each cell has its distinct characteristics, it can be difficult to classify and count each cell manually. Plant biologists have historically used fluorescent dyes to color individual cell walls for cell identification, a tedious process.

To accelerate identification and quantification of the fiber, parenchyma, and vessel, we propose a deep-learning object detection algorithm to automate this process. The algorithm would count and visually border internal cells, allowing botanists to determine the ratios of each cell in an accurate and efficient manner. This information can be extrapolated to comment on the relationship between cell ratio and several plant properties such as strength and water retention.

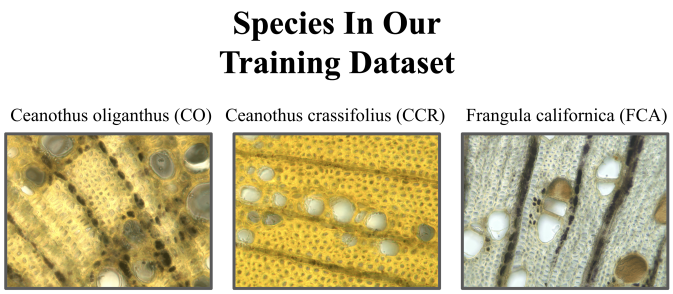

Fig. 1. Depiction of the variety of species that we used to diversify and generalize our model.

We want to quantify the vessels because it has been shown that the number of vessels has a direct correlation to its wood anatomy and development. [6] Additionally, the number of vessels is also linked to its resistance to vessel embolism, which is essentially a clot or air bubble in the cell type. The quantification of fibers provides important information about the embolism resistance and mechanical support to the stem, and potentionally has a correlation to vessel cavitation resistance [8]. Finally, the botanists are interested in the number of parenchyma in a given cross section, so they can analyze the state of the pathogen defense system of the plant. [7]

Our models utilize multiple machine learning techniques for object detection, including some popular networks such as YOLOv5 [1]. We experiment with the compatibility of our dataset with YOLOv5's trained COCO dataset model [2] as well as how well it can generalize with external data. Furthermore, to quantify the vessels that are detected by our network, we can simply measure the pixel-number relationship between the bounded region and the normal image width and height.

A. Related Works

Cell detection is an intriguing task that biologists have been particularly invested in over the past few years, making breakthroughs in cancer cell detection of fish [5], the quantification of a plant's fitness levels through image segmentation-(R-CNN) architectures [2], and white blood cell segmentation [4]. These studies exhibit innovative applications for computer

vision and imaging which we can build on. In terms of plant cell research, there have been various studies on methods to segment plant tissues and their interior organs. However, few studies have been conducted on the object detection of a plant's parenchyma, fiber, and vessel- more specifically in the California fern species. In our previous works, we explored possible ways to classify the fiber, parenchyma, and vessels from manually cropped images using traditional transfer-learned neural networks. Nonetheless, the process of classifying a cropped cell-type still gives little information of the state and health of the plant as a whole, as a cropped cell image is needed for input. Our ambition with this research project is to add an additional detection phase to the plant cell classification.

II. Methods

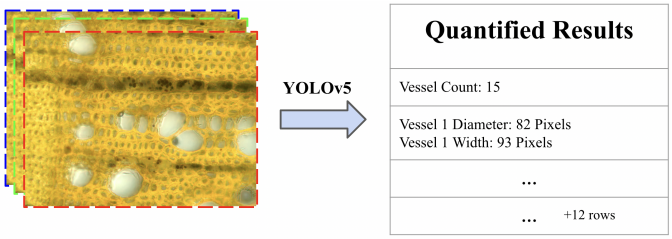

Fig. 2. Illustration of the output that we hope our model can produce after learning from the training set.

In this study we explore possible techniques to quantify and detect the fibers, parenchyma, and vessels in different plant cell cross sections. We plan on experimenting on how our dataset performs on popular supervised object detection algorithms such as YOLOv5. The purpose of this study is to not only quantify the different cell types in xylem tissue, however we would also like to explore the possibility of semi-self supervised learning to label a vast amount of data for us, while only being trained on a small sample of cross sectional data.

A large downside of the object detection methodology that we are conducting is the lack of negative sampling in our dataset. As illustrated in Figure 1, almost every pixel of the cross section is consisting on either a fiber, parenchyma, or vessel, which can potentially make it difficult for an object detection system to learn.

A. Dataset Acquisition

A popular imaging technique to obtain xylem tissue data is a light microscopy technique where each stem or root sample is converted to cross-sections using a sledge microtome. The dataset that is pertinent to our project are different stem and root species from a myriad of chaparral shrub species gathered from the Santa Monica Mountains which contains the Ceanothus crassifolius (CCR), Ceanothus oliganthus (CO), and Frangula californica (FCA) shrub species. The cross section that were obtained by previously mentioned techniques were all stained by I_2KI then magnified by through a range of 100-200x zoom. The reason we are so interested with the California shrub species is because analyzing the xylem tissue cells of these plants can give us crucial information such as how they survive six month long droughts every year.

B. Treble Model Collection

Performing a object detection task on three different classes that are so spacially close together may hinder the network from learning the correct bounding box placements. This is due to the lack of negative sampling, and it is difficult for the model to learn what is in the background and what is not a cell type. To combat this, we trained three separate models, one to detect the vessels, another to detect the parenchyma, and finally one to detect the fibers. By doing this, we are in a way separating the tasks of object detection and classification, allowing for these models to focus on one cell type at a time, and ascertain lots of background and negative samples for each image.

In practice, having three models also allows the user to choose a specific task to be performed by the model, for example, if the user only wanted to measure the width and height of vessels, they could do so without being overwhelmed by the number of bounding boxes for the fiber and parenchyma.

C. Iterative Computerized Labeling

One of the greatest challenges of training this network is the labeling of the training data. Whilst there may only be ten to twenty vessels per cross section, the number of parenchyma and fibers can span to the hundreds or thousands. To expedite the process of labeling the parenchyma and fibers, we implemented an iterative labeling algorithm from a subset of the cross sections. First, we fully label some high quality cross sections ourselves, making sure to include every fiber or parenchyma. Once we have this subset of labeled data, we then train a object detection network to perform these tasks at a certain accuracy. We then deploy that model on every unlabeled cross section and choose the best machine-labeled images to be added to our training dataset. This process is iterated over and over until we have a full training dataset that are psuedo-labeled.

D. Pre-Processing and Data Augmentation

Before applying any machine learning algorithms to create a object detection system, we must first pre-process and clean our data. A plant botanist utilized popular data labeling website Roboflow to manually annotate 89 cross sectional images of cross sections from various California shrub species. We partitioned the dataset into a (71,9,9) train validation split. In each cross sectional image, every vessel was given a manual bounding box, and n parenchyma and m fibers were labeled depending on which sub-model we were training.

To prevent overfitting and to generalize our model to unseen test data, a mosiac dataloader was used in YOLOv5, which takes a few images as input, and randomly shuffles the order, while applying data augmentation to that whole sequence. The scaling, color space, and resize transforms were also

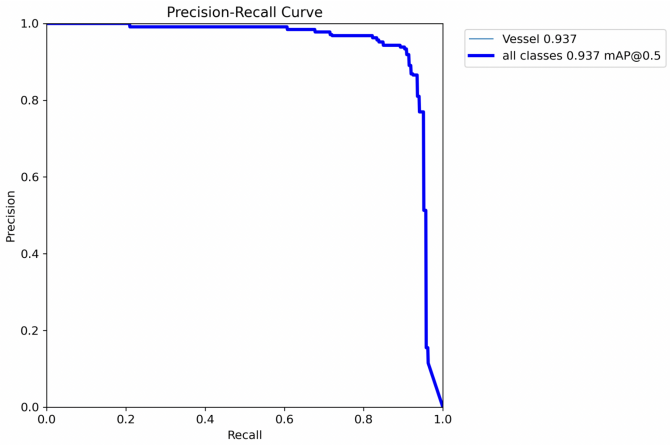

Fig. 3. Results of YOLOv5 precision recall curve after training on 150 epochs

implemented synonymously within all bounding boxes of a cross section.

E. Supervised Approaches

For our supervised baseline approach, we have implemented the PyTorch YOLOv5 model, which is one of the most recent and popular open-sourced models for object detection. We expect that training the YOLOv5 model on our 70 labeled cross sections will yield results that can detect the vessels well, because the plant's vessels have a visually distinct shape and color.

Upon successful computation of the bounding boxes for every cell type, we saved the best weights of each network and on embedding models into a user friendly web application in which researchers can simply upload a single cross sectional image and receive the quantifiable information in seconds. Additionally, we encourage those who do use the online platform to allow us to use these external cross sections to create a database so that more robust and generalizable models can be trained.

III. Results

Our research consisted on training three separate pipelines of separate tasks. The first detection model is to to detect the vessels in the xylem cross sections, which have shown to represent distinct shape, size and color. For this task, we manually labeled every vessel in the cross sections manually, and trained the YOLOv5 network for 150 epochs to reach a global minimum for the validation loss. The fiber and parenchyma detection tasks also had their respective training pipelines, where a semi-self supervised approach is taken. Each model is trained three times, with an increased training sub-set each iteration. The semi-self supervised system starts off with 6 labeled images with two images in the validation and test set. On the next iteration, we choose 24 cross sections to be in the training set with 4 images in the validation and testing set. Finally on the last iteration, we label every cell in the cross sections to train a final model on the whole dataset.

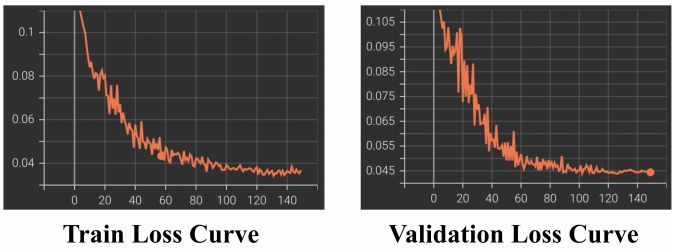

Fig. 4. Visualization of train and validation loss curves for vessel detection

Due to the fact that we are lacking an actual ground truth label for the vessel width and height, the results from the evaluation are a measurement of the relative size of the cell type compared to the cross section as a whole.

A. Evaluation Metrics

To evaluate our three networks' performances, we utilized the F1 confidence score and area under the precision-recall curve. These metrics are both dependent on the precision and recall of the model's outputs. The precision is calculated by the true positives divided by the true positives + false positives, and the recall is computed by the true positives divided by the true positives + false negatives. In our case, we try to maximize the F1 confidence score because we would like to minimize the false positive and negative rates, while keeping a balance between the precision and recall scores (harmonic mean).

Similarly, the area under the precision-recall curve is also dependent on the relationship between the precision and recall, where we large area under the curve means that the precision and recall are maximized, meaning there is a low false positive rate whilst also having a low false negative.

B. Vessel Detection

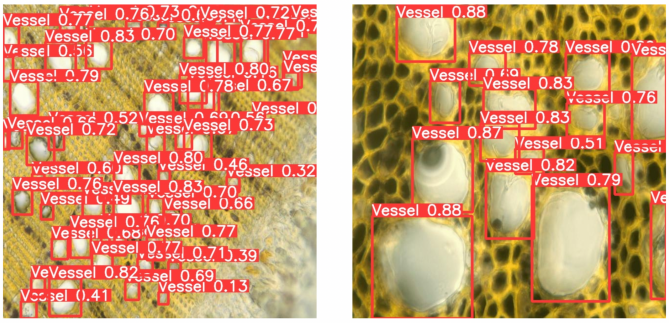

Fig. 5. Illustration of our trained model detecting all vessels

Learning to detect the vessels in these cross sections has proven to be a relatively easy task for the network. It is important to note that YOLOv5 utilizes a GlobU loss function which is a generalized intersection over the union loss, specially made for bounding box detection problems. As illustrated in Figure 4, the validation loss decreased synonymously with the training loss, inferring that our model is generalizing well on

unseen data, and that the loss has converged to a minima. Figure 3 depicts the metrics of the vessel detection model, where the area under the precision recall curve is large, and a high F1-confidence curve is achieved. A visualization of the networks testing set can be seen in=Figure 5, where virtually every vessel is properly detected with a high confidence. It is also clear that the model is generalizing well on cross sections of different scaled microscopes.

C. Fiber Detection

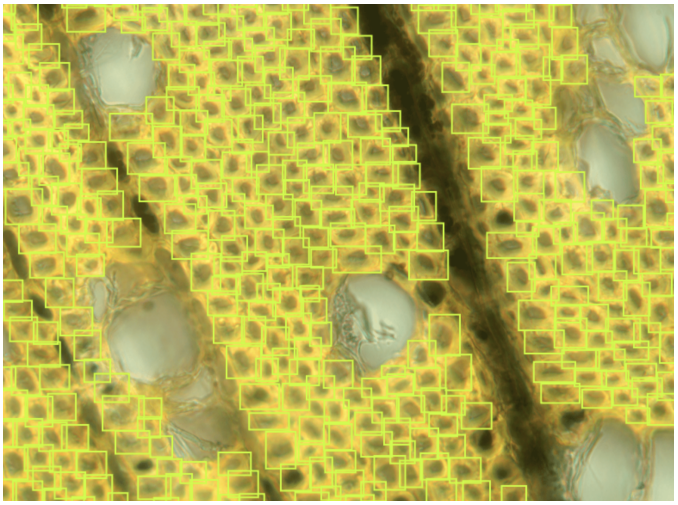

Fig. 6. Illustration of the psuedo-ground truth bounding boxes that the first training iteration provided.

The fiber detection algorithm is trained 3 separate times, each on a different training sized subset. We first train a YOLOv5 algorithm on 10 cross sections, with 1 image in the validation set and 2 images in the test set. This first training resulted in a precision of 0.71 and recall of 0.62. With this preliminary network, we then ran inference on all of the unlabled data samples, and chose a specific confidence threshold eg. 0.5 to automatically annotate the images. This process can be seen in Figure 6, where the errors that arose in the pseudo labelling process were manually fixed by hand. The second iteration consisted of a training set of 31 images, with 5 images in the validation and test set, and consisted of a precision score of 0.80 and a recall of 75.8. We then used the second trained model to fully annotate all of the cross sections, and retrained a YOLOv5 from scratch for 200 epochs, where the model stopped training if there was no improvement for 100 epochs. The results exemplify that the quantification of fibers is not an easy task, where the area under the precision-recall curve is 0.61 and the F1 confidence is 0.51. The confusion matrix of task task can be seen in Figure 7 where the network is experiencing difficult discernment between the fibers and non-fibers.

D. Parenchyma Detection

Our first iteration of the parenchyma detection involved 16 images in the training dataset and 2 images in the test and validation sets. This first iteration resulted in a 0.75 precision and 0.62 recall score. We then applied this model onto the remaining cross sections, where 36 training images were used to annotate the rest of the data samples. This second training model also has a precision of 0.75 and recall of 0.663. A possible explanation for the similarity of precision and recall despite there being more training samples, is that more difficult samples are added to the test and validation set. When training the final labeled dataset from scratch, the area under the precision-recall curve resulted in 0.751.

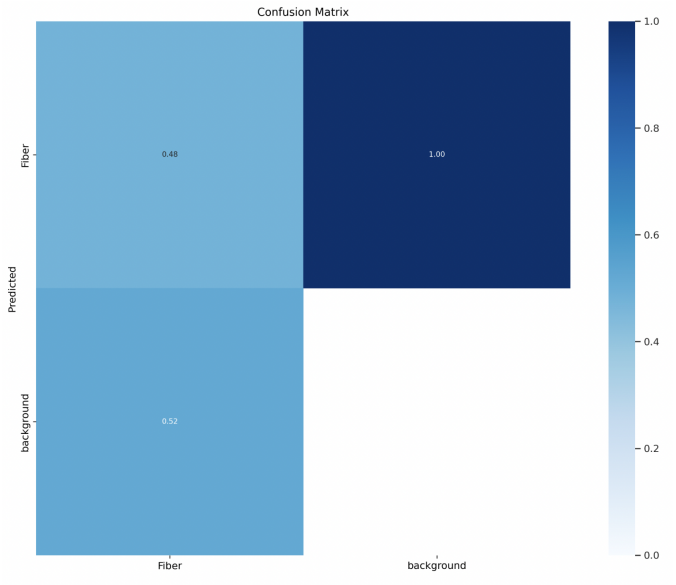

Fig. 7. Confusion matrix of the fiber detection model.

IV. Discussion

In this research project, we explored a possible object detection algorithm (YOLOv5) to confidently detect different cell types within a plant's xylem cross section. We mainly utilized two novel approaches to this detection task. The first approach was to train a separate YOLOv5 instance for each cell type to drastic increase negative sampling, which simplified the learning to really study what is a vessel, and what is in the background. The second methodology that we implemented was a semi self-supervised network, where we accelerated the labeling process by starting off with a small training set, and we iteratively trained a network and evaluated on the data until our data was all labeled.

Some possible issues that could arise in this study, is the manual labeling of the cells. Due to the variability of the xylem staining and imaging process, there are some cross sections that leaves a high room for error in labeling these cell types. For example, if the cross section were stained too extremely, then some fibers could be confused as parenchyma due to having a darker color. Additionally the rows of parenchyma can visibly look like one black strip, making the individual cells almost indistinguishable. Another potential labeling confusion is mistakenly labeling a growing vessel as a fiber, or vice versa.

In these extreme cases, the quantification and measurement of the vessels may be offset. This issue can be visualized in Figure 8, where the growing vessel is mistaken as part of the background cells, rather than being a vessel.

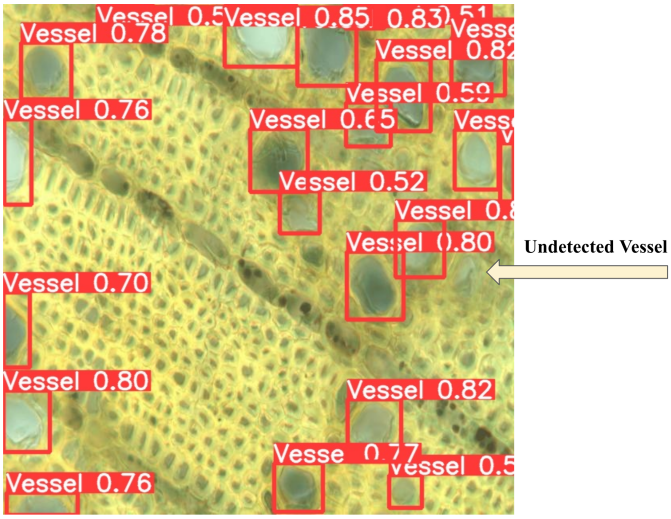

Fig. 8. Example of the network missing a growing vessel (right-middle)

While the F1 confidence and area under the precision-recall curve scores are substantial, there can be improvements to hasten the labeling process and better generalize our network on unseen data. On future datasets, there are two main active learning techniques that we could utilize to improve our models performance. The first method is pool-based sampling method [10], inspired by the active learning application to linear regression [9], where we start off with an initial subset of the training set, and after training that model, we evaluate the examples with the lowest confidence or predicted outcome. Once we know what which instance of cells are difficult to detect, we could then try to only manually label those low scoring cells and retrain our model that way. By doing this, we continuously improve the low performing cells until the model becomes very robust.

Another active learning pipeline that we are considering adding in the future is a stream-based selective sampling method [10], where the network actually has a choice to choose to use the manually labeled ground truth value, or use the pseudo generated labels. While this pipeline may demonstrate a performance boost, it is actually very computationally expensive because the model has to choose whether to utilize a real label for every bounding box, and in our case, that could mean thousands per cross section.

V. Conclusion

We have researched an efficient way to automatically detect every cell type in a plant's xylem tissue cross section. This research project could possibly be integrated into the labs of botanists all around the world free of charge, where our model is used from a web-based server to quantify their respective cell types. A mutually beneficial relationship may even be formed with this deep learning project if biologists continuously providing more training data whilst using the network. We hope that with this research, botanists can spend more time on meaningful research rather than expending hours manually labeling images.

VI. Acknowledgements

We would like to thank R. Brandon Pratt for help with shrub sample collection.

References

[1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[2] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft Coco: Common Objects in Context," Computer Vision – ECCV 2014, pp. 740–755, 2014.

[3] X. Lin et al., "Self-Supervised Leaf Segmentation under Complex Lighting Conditions," Pattern Recognition, vol. 135, p. 109021, Mar. 2023, doi: 10.1016/j.patcog.2022.109021.

[4] X. Zheng, Y. Wang, G. Wang, and J. Liu, "Fast and robust segmentation of white blood cell images by self-supervised learning," Micron, vol. 107, pp. 55–71, Apr. 2018, doi: 10.1016/j.micron.2018.01.010.

[5] C. Albuquerque et al., "Object detection for automatic cancer cell counting in zebrafish xenografts," PLOS ONE, vol. 16, no. 11, p. e0260609, Nov. 2021, doi: 10.1371/journal.pone.0260609.

[6] A. E. Zanne et al., "Angiosperm wood structure: Global patterns in vessel anatomy and their relation to wood density and potential conductivity," American Journal of Botany, vol. 97, no. 2, pp. 207–215, Feb. 2010, doi: 10.3732/ajb.0900178.

[7] H. Morris, C. Brodersen, F. W. M. R. Schwarze, and S. Jansen, "The Parenchyma of Secondary Xylem and Its Critical Role in Tree Defense against Fungal Decay in Relation to the CODIT Model," Frontiers in Plant Science, vol. 7, Nov. 2016, doi: 10.3389/fpls.2016.01665.

[8] A. L. Jacobsen, F. W. Ewers, R. B. Pratt, W. A. Paddock, and S. D. Davis, "Do Xylem Fibers Affect Vessel Cavitation Resistance?," Plant Physiology, vol. 139, no. 1, pp. 546–556, Aug. 2005, doi: 10.1104/pp.104.058404.

[9] M. Sugiyama and S. Nakajima, "Pool-based active learning in approximate linear regression," Machine Learning, vol. 75, no. 3, pp. 249–274, Jan. 2009, doi: 10.1007/s10994-009-5100-3.

[10] B. Settles, "Active Learning Literature Survey," minds.wisconsin.edu, 2009. https://minds.wisconsin.edu/handle/1793/60660