An Automatic Method for Stomatal Pore Detection and Measurement in Microscope Images of Plant Leaf Based on a Convolutional Neural Network Model

Wenlong Song, Junyu Li, Kexin Li, Jingxu Chen and Jianping Huang *

School of Mechanical and Electrical Engineering, Northeast Forestry University, No.26 Hexing Road, Xiangfang District, Harbin 150040, China; wlsong139@126.com (W.S.); ljy9950@163.com (J.L.); dillonlkx@163.com (K.L.); cjxulucky@163.com (J.C.)

* Correspondence: jphuang@nefu.edu.cn; Tel.: +86-185-4585-8294

Received: 15 July 2020; Accepted: 27 August 2020; Published: 1 September 2020

Abstract: Stomata are microscopic pores on the plant epidermis that regulate the water content and CO₂ levels in leaves. Thus, they play an important role in plant growth and development. Currently, most of the common methods for the measurement of pore anatomy parameters involve manual measurement or semi-automatic analysis technology, which makes it difficult to achieve high-throughput and automated processing. This paper presents a method for the automatic segmentation and parameter calculation of stomatal pores in microscope images of plant leaves based on deep convolutional neural networks. The proposed method uses a type of convolutional neural network model (Mask R-CNN (region-based convolutional neural network)) to obtain the contour coordinates of the pore regions in microscope images of leaves. The anatomy parameters of pores are then obtained by ellipse fitting technology, and the quantitative analysis of pore parameters is implemented. Stomatal microscope image datasets for black poplar leaves were obtained using a large depth-of-field microscope observation system, the VHX-2000, from Keyence Corporation. The images used in the training, validation, and test sets were taken randomly from the datasets (562, 188, and 188 images, respectively). After 10-fold cross validation, the 188 test images were found to contain an average of 2278 pores (pore widths smaller than 0.34 µm (1.65 pixels) were considered to be closed stomata), and an average of 2201 pores were detected by our network with a detection accuracy of 96.6%, and the intersection of union (IoU) of the pores was 0.82. The segmentation results of 2201 stomatal pores of black poplar leaves showed that the average measurement accuracies of the (a) pore length, (b) pore width, (c) area, (d) eccentricity, and (e) degree of stomatal opening, with a ratio of width-to-maximum length of a stomatal pore, were (a) 94.66%, (b) 93.54%, (c) 90.73%, (d) 99.09%, and (e) 92.95%, respectively. The proposed stomatal pore detection and measurement method based on the Mask R-CNN can automatically measure the anatomy parameters of pores in plants, thus helping researchers to obtain accurate stomatal pore information for leaves in an efficient and simple way.

Keywords: stomatal detection; pore measurement; deep convolutional neural network; black poplar

1. Introduction

Stomata control the fluxes in carbon dioxide and water vapor levels across a leaf [1–4]. The rest of the epidermis is covered by an impervious cuticle, with limited possibilities for gas exchange [5]. The gas exchange capacity depends on stomatal density (i.e., stomatal number per unit area), stomatal size (length and width), and pore dimensions (length and aperture) [6–9]. Pore area is dynamically adjusted by changes in pore aperture, since pore length is rather rigid during the opening and closure...
of stomata [10,11]. Active pore aperture adjustments in response to internal (e.g., water status) and external (e.g., environmental) factors are physiologically regulated [12–15]. Instead, pore length and stomatal density are anatomical features that are set during leaf elongation [10,16]. These stomatal features have been the focus of a wide range of studies [6–8,17,18]. For instance, they are often used to estimate stomatal conductance (gs) based on the equation by Brown and Escombe [19] or modified versions of this equation [17,20]. In this way, for instance, reconstructions of gs over geological time scales were made possible based on the stomatal traits of fossilized leaves [8]. Estimations of gs from stomatal anatomy features are also important when the effect of changing a single feature on gs is investigated [21,22].

At present, most methods for the measurement of stomatal pores involve manual measurement from images using image processing software, such as ImageJ [23]. This type of method requires researchers to manually label points of interest, such as boundaries, lengths, and widths, in a pore. The disadvantages of this method are that, first, it requires manual intervention, and second, due to the huge amount of data required, only some of the data points can be used to build the model, thus making the results inaccurate. This has led to many automatic measure methods being proposed in an attempt to handle the large data volumes associated with leaf pores and the high accuracy requirements of data analysis.

A method for measuring the parameters of stomatal pore anatomy was first proposed by Omasa and Onoe [24]. This method used Fourier transform and unsharpened mask techniques to remove the noise of original images and calculate the length and width of sunflower pores through edge detection. The disadvantages of this method are that it requires a large amount of computation and is only suitable for single-porosity images. Laga et al. [25] proposed an automatic method. They firstly detected stomata using template matching technology and then extracted stomatal aperture by binary segmentation. However, this method is reliant on a template for each plant species. Liu et al. [26] used maximum stable external regions (MSERs) to detect and measure grape stomata. This is a semi-automatic method because it requires the user to properly select ellipses to fit different stomata. An automatic method for the pore measurement of grape varieties was proposed by Jayakody et al. [27]. This method is based on machine learning theory and uses histogram of oriented gradients (HOG) characteristics to construct a cascade target detector to detect pores and then calculate various associated parameters through binary image segmentation and skeleton technology. Though the method is fully automated, it requires the analyzed microscope image to contain rich background features. Toda et al. [28] firstly used a HOG feature to detect stomata followed by a convolutional neural network (CNN) to classify clipped individual pores as open or closed; finally, they used binary image segmentation to complete automatic pore measurement. As the algorithm needs to manually define the range of parameters (such as area, solids, spindle length, and centroid coordinates), it is not able to identify pores when their size or shape is not within the predefined parameter range. Bhugra et al. [29] proposed a method based on a deep learning neural network to identify pores in SEM images and performed a method of stomatal pore segmentation. This method uses a single shot multi-box detector (SSD) to detect the distribution of stomata using a super-resolution convolutional neural network (SRCNN) to improve the resolution of cropped local images; finally, it uses a fully convolutional network (FCN) to segment the pores. A method for stomatal pore segmentation based on the Chan–Vese (CV) model was presented by Li et al. [30]. The author first used a faster region-based convolutional neural network (Faster R-CNN) to detect the stomatal location and then cropped the detected stomata, thus constructing single stomatal pictures and segmented these single stomata using a CV model, but the method requires manually adjusting the parameters of the CV model, such as the quality of image and the stomatal shape, according to the image being processed. Furthermore, this method can only handle pores with a larger stomatal aperture because it is more difficult to measure pores with a smaller degree of aperture.

In this paper, we propose an automatic high-throughput method based on the mask region-based convolutional neural network (Mask R-CNN) [31] to acquire parameters of stomatal pore anatomy. The Mask R-CNN, based on the Faster R-CNN [32], is an instance segmentation model.


When a new stomatal image is used as input, this model processes the image by not only providing a bounding box around the object but also providing a prediction about the category of each pixel. This method was applied to the detection of stomata based on microscope images of plant leaves, and the results were evaluated visually and in terms of several quantitative indices.

2. Materials and Methods

2.1. Data Acquisition

In this paper, we used one-year-old black poplar (*Populus nigra*) growing in its natural environment as the experimental plant. Black poplar is often used as a protective forest species and plays an important role in environmental protection. It also serves as a habitat for many animals, and its seeds can be eaten by finches. Therefore, black poplar plays an important role in ecology and the protection of the environment.

The Keyence VHX-2000 microscope observation system with a large depth-of-field [33] was used to obtain stomatal microscope images of fully focused black poplar leaves, as shown in Figure 1.

![Figure 1](image_url)  
**Figure 1.** The Keyence VHX-2000 large depth-of-field microscope observation system.

The resolution of obtained stomatal pores images was 1600 × 1200, and the magnification was 1000x. A self-made lifting device was used to simultaneously obtain leaves from each position, including the bottom, middle, and top leaves of different trees, as well as the top, middle, and base of a single leaf. To arrive at a relatively robust conclusion, the task of collecting images of different parts of poplar leaves continued for three months (July–October). Poplar leaves need to be fixed with a self-made leaf holder to make them flatter during the collection process. Examples of the collected images are shown in Figure 2.
The second step used the trained Mask R-CNN to detect and segment the pore region in the test set, parameters detected by the model, including the length and width of the pore and the pore area, to generate a binary map with the same size as that of the input image for each pore region. The validation and test sets were used to optimize the network hyperparameters. Details of the hyperparameter adjustment process are found in Section 2.4. The second step used the trained Mask R-CNN to detect and segment the pore region in the test set, parameters detected by the model, including the length and width of the pore and the pore area, to generate a binary map with the same size as that of the input image for each pore region.

2.2. Methods

The aim of this study was to acquire a fully automated method for stomatal pore segmentation and the quantitative measurement of plants. The overall flow of the proposed method is shown in Figure 3. The input was a stomatal microscope image. The output was the stomatal pore anatomy parameters detected by the model, including the length and width of the pore and the pore area, the elliptical eccentricity based on the pore [26], and the degree of stomatal opening—for a detailed introduction of anatomy parameters, refer to Section 2.2.2. The method consisted of 3 steps: (1) make a network training set, validation set, and test set to obtain a network model; (2) use the trained network model to obtain the mask (segmentation results) contour coordinates of each pore in the test image; and (3) obtain the parameters of the pores based on the least squares fitting of ellipses according to the contour coordinates of the mask and stomatal pore measurement model [34].

First, all acquired black poplar microscope images were randomly divided into a training set, a validation set, and a test set. The labeling tool LabelMe [35] was used to perform manual labeling under the guidance of a botanist. Examples of labeled results are shown in Figure 4. Each microscope image included about 10–20 open stomata. The experimental network was achieved based on the model of Matterport [36] according to the actual requirements of the experiment (for a detailed introduction to the Mask R-CNN, refer to Section 2.2.1). The validation and test sets were used to optimize the network hyperparameters. Details of the hyperparameter adjustment process are found in Section 2.4. The second step used the trained Mask R-CNN to detect and segment the pore region in the test set, as well as to generate a binary map with the same size as that of the input image for each pore region.
To measure the pore anatomy parameters, we cropped the binary map to create a small map for use in the stomatal pore measurement model, thus ensuring that the pore was located in the center. The mask area on the cropped binary map is represented by the white region in Figure 3. The measurement model performed boundary extraction (the red line in Figure 3) on the mask area to obtain the contour coordinates of the mask area. The third step was ellipse fitting based on least squares [37]. The ellipse fitting (the blue line in Figure 3) of the mask area was performed according to the contour coordinates of the mask region, and the length, width, area, eccentricity, and stomatal aperture of the pores were the output (for a detailed introduction to pore measurement, refer to Section 2.2.2).

Figure 3. Model flowchart of the proposed method. (RPN: region proposal network; FC: fully connected layers; and FCN: fully convolutional network).

Figure 4. An example of manually labeled results: (a) original stomatal image, (b) mask overlaid on the original image, and (c) the labeled mask image (red is the mask area of the pore and black is the background).

2.2.1. Architecture of the Model

In this work, a Mask R-CNN based on deep learning was used to detect and segment the stomatal pores of a microscope image of black poplar. The Mask R-CNN is a two-stage instance segmentation algorithm (where the first stage is a region proposal that judges whether a positive object exists or not and the second stage predicts the category of the object that is output from the first stage and describes their masks) and is composed of the following four modules, whose network architecture is shown in Figure 5.
(1) Feature Extraction: The Mask R-CNN is composed of a ResNet50 network and a feature pyramid network (FPN) that have been turned into a top-down multiscale feature extraction network. The use of a multiscale feature extraction network is superior to the use of the ResNet50 network in isolation as an extraction network. The FPN transforms the feature map (C1–C5) extracted by the residual network into different scale feature maps by up-sampling (P2–P5) and max-pooling (P6 subsamples from P5 with a stride of 2, only used in the region proposal network). This architecture makes the network pay attention to both detailed information and semantic information in the images. In the process of feature extraction of the neural network, the shallower the layer in the network, the more detailed the information in the image; however, the semantic information of the image is ignored. On the contrary, the deeper the network layer, the poorer the detailed information becomes—but the semantic information of this image improves. The FPN uses the method of cascade links with different scale feature maps by sharing information, acquiring an improved feature extraction result.

(2) Region Proposal Network (RPN): This network proposes a probable target region by the branches of sub-classification and sub-regression. Anchors are used in this process. In one image, a different size of the target means it cannot be predicted with a cell of the same size, so anchors with different aspect ratios are generated by the network at each pixel, and a feature map generates multiple anchors. Anchors slide into the feature map and generate the candidate region features, which are extracted as low-dimensional features and then sent to the sub-fully connection layer in a process called boundary box regression; additionally, the classification layers acquire the proposed target.

Figure 5. The network architecture of the Mask R-CNN (region-based convolutional neural network) (C1–C5 are the five stages of the feature extraction of Resnet-50. FCN: fully convolutional network; RPN: region proposal network; ROI: region of interest; bbox: bounding box; bbox_pred: the prediction of bounding box; and cls_prob: the probability of class).
(3) RoIAlign: In the head architecture of the Mask R-CNN, the input size is fixed to 7 × 7. However, the region of interest (ROI) size for the RPN of each target is different, so RoIAlign is required for size normalization. The difference between the RoIAlign and RoIPooling structure used by the Faster R-CNN is that the former uses bilinear interpolation in pixel processing and not the rounding operation used by the latter. This improves the accuracy of mask generation in segmentation tasks.

(4) Mask-Generated: The Mask R-CNN adds an FCN on the basis of the Faster R-CNN to generate a mask for the target, and several branches (box regression, classification, and mask-generated) operate in parallel. After the feature map passes the RPN, the ROI is sent to the mask-generated branch.

The total loss function is defined as:

\[ L_{loss} = l_{cls} + l_{reg} + l_{mask} \]  

where \( L_{cls} \) is the SoftMax loss, \( L_{reg} \) is the smooth \( L_1 \) loss, and \( L_{mask} \) is the binary cross-entropy loss. \( L_{reg} \) is given by:

\[ L_{reg} = smooth_{L1}(t - t^*) \]  

where the smooth function is defined as \( smooth_{L1}(x) = \begin{cases} 0.5x^2, & |x| \leq 1 \\ 0.5|x| - 0.5, & \text{otherwise} \end{cases} \), \( t \) represents the box coordinates generated by the prediction, and \( t^* \) represents the manually labeled box coordinates.

The \( L_{mask} \) is defined as:

\[ L_{mask}(s^*, s) = -(s^* \log(s) + (1 - s^*) \log(1 - s)) \]  

where \( s \) is the true binary mask by manual label and \( s^* \) is the predicted binary mask.

2.2.2. Stomatal Pore Measurement

The pore measurement algorithm consists of mask contour coordinate extraction, ellipse fitting, and parameter calculation. The pore parameter measurement process is shown in Figure 6.

![Figure 6](image)

Figure 6. The pore parameter measurement process.

The mask generation branch of the network generates a binary image with the same size as the input image for each stomatal pore in each test image. Each pixel in the image is used as a coordinate point. After calculation, 4.8 pixels = 1 μm. The value of each point is either false or true, where false means that the area in which the pixel is located is not the mask area of the pores and true means that the pixel is in the area of the pores in the mask. Each picture generates \( n \) binary maps through the mask generation branch, where \( n \) is the number of detected pores in the picture. The stomatal pore measurement model extracts the boundary coordinates of the mask in each binary image and crops the pore region from these binary images, and—through the least squares ellipse fitting technique [34,37]—the values of the pore length and width [24] can be obtained and recorded as \( 2a \) and \( 2b \), as shown in Figure 6. In order to further analyze the physiological characteristics of the stomatal pores, the anatomical parameters of pores can also be obtained, including area, eccentricity, and the degree of pore opening.
The pore area represents the size of the channel for gas and moisture exchange \cite{9}, and it determines the stomatal conductance to CO\textsubscript{2} and H\textsubscript{2}O \cite{17}, which is defined as the product of the pore length (2a), pore width (2b), and \pi. The pore area is:

\[
\text{Area} = \pi ab.
\]  
(4)

Stomatal aperture adjusts depending on the prevailing ambient environmental conditions \cite{38}, which is the ratio of the pore width (2b) to the pore length (2a). The stomatal aperture is defined as:

\[
\text{Stomatal aperture} = \frac{b}{a}
\]  
(5)

The eccentricity of the pores after being fit into an ellipse can also be used to calculate the stomatal aperture \cite{26}. The eccentricity is determined as follows:

\[
\text{eccentricity} = \sqrt{1 - \left(\frac{b}{a}\right)^2}
\]  
(6)

2.3. Evaluation Indices

In addition to a visual assessment, several quantitative indices were calculated to evaluate the performance of the proposed method in comparison with the predicted and ground truth, and these included precision, recall, the intersection of union (IoU), and pore measurement accuracy.

Precision and recall are widely used to evaluate the performance of object detection methods. Precision and recall are defined as follows:

\[
\text{precision} = \frac{TP}{TP + FP}
\]  
(7)

\[
\text{recall} = \frac{TP}{TP + FN}
\]  
(8)

where True Positive (TP) means that the pores are correctly identified by the model from the defined pore region, False Positive (FP) means the background is misidentified as pores, and False Negative (FN) means the pores are misidentified as the background.

The IoU index can be used to compare similarities and differences between finite sample sets. The greater the IoU coefficient, the higher the similarity between the segmentation result and the corresponding ground truth.

The IoU index is defined as follows:

\[
\text{IoU} = \frac{S_{\text{pred}} \cdot S_{\text{gt}}}{(S_{\text{pred}} + S_{\text{gt}}) - S_{\text{pred}} \cdot S_{\text{gt}}}
\]  
(9)

where \(S_{\text{pred}}\) is the detection result location and \(S_{\text{gt}}\) is the real object location.

We used the relative error to verify the effect of stomatal pore measurement. In this paper, the relative errors of the parameters of the true mask of each pore in the image and the fitting ellipses of the segmentation mask were obtained. The relative error formula for each pore is as follows:

\[
\text{relative_error} = \frac{g_{\text{gt}} \text{parameter} - p_{\text{pred}} \text{parameter}}{g_{\text{gt}} \text{parameter}}
\]  
(10)

where \(g_{\text{gt}} \text{parameter}\) represents the anatomical parameters of pore areas marked by ground truth after ellipse fitting (the pore length and width, area, eccentricity, and stomatal aperture) and \(p_{\text{pred}} \text{parameter}\) represents the anatomical parameters of the mask obtained by our method of segmentation after ellipse fitting.
2.4. Model Parameters and Operating Environment

In order for the original network to train our self-made dataset, we adjusted the output channel of the fully connected layer of the network from 81 to 2 (for the two categories of pore and background) prior to network training. The weights generated by the Mask R-CNN training on the balloon dataset were used as the initial weights for training [36]. Compared with the training result using the weights downloaded with ImageNet, these weights resulted in a faster convergence of the model parameters with an increased accuracy of results. In order to make the network perform optimally, we adjusted the hyperparameters, including the training epoch, learning rate, size of the image, and batch size of the network, according to the characteristics of the data. During the training process, we determined the number of training epochs according to the change trend of the validation set accuracy and adjusted other hyperparameters, including the learning rate, size of image, and batch size, according to the test accuracy of the model on the test set. In the proposed method, the epochs for network training were (1) 40, (2) 120, and (3) 160, which were the numbers of train epochs for the (1) head architecture, (2) depth 4+ of ResNet50, and (3) the entire network, respectively, and there were 100 steps in each epoch. The batch size was set as 1; the size of the image was resized to $1024 \times 1024$; the learning rate was (1) 0.001, (2) 0.001, and (3) 0.0001 (the number of train epochs for the (1) head architecture, (2) depth 4+ of Resnet, and (3) entire network, respectively); the weight decay rate was 0.0001, the learning momentum was 0.9, and the detected confidence was 90%. The hyperparameter settings are shown in Table 1.

| Parameter            | Value               |
|----------------------|---------------------|
| Learning Rate        | 0.001, 0.001, 0.0001|
| Learning momentum    | 0.9                 |
| Weight decay         | 0.0001              |
| Epochs               | 40, 120, 160        |
| Steps per epoch      | 100                 |
| Batch size           | 1                   |
| GradientClipNorm     | 5.0                 |

In this study, the experiments were performed using the deep learning platforms Keras (2.0.8) and Tensorflow (1.10.0) in Python 3.5. The hardware support for the experiment was a Nvidia GeForce RTX2080ti. The operating system was Windows 10 with an Intel (R) Core (TM) i7-9700k, 3.6 GHz CPU, and 16 GB memory.

3. Results

3.1. Pore Detection and Segmentation

Examples of the detection and segmentation results of the proposed method are shown in Figure 7. As can be seen in Figure 7, the proposed method achieved a good performance regarding the detection and segmentation of stomatal pores. We randomly divided the training and test sets into 10 subdatasets, used nine subdatasets to train the model, and used the rest of the subdataset for testing. Through the statistical analysis of 188 test images and after 10-fold cross validation, an average of 2201 stomata were detected from an average of 2278 stomata marked by the proposed method. In order to quantitatively analyze the validity and feasibility of the proposed model, the areas corresponding to stomatal pores were manually labeled using the LabelMe [35] marking tool under the guidance of botanical experts, and the obtained results were used as the ground truth (GT). The annotation results were compared with the measured results obtained by the proposed method. The average precision was 96.72%, the average recall rate was 96.87%, and the average IoU of the pore was 0.82. The average time to process an image was about 912 ms. The complete code for the project can be accessed at https://github.com/lijunyu159/stomatal_pore_measurement-MaskRCNN (accessed on 15 July 2020).
pore was 0.82. The average time to process an image was about 912 ms. The complete code for the project can be accessed at https://github.com/lijunyu159/stomatal_pore_measurement (accessed on 15 July 2020).

Figure 7. (a)–(d) are examples of the segmentation results (the color mask represents the pore area).

3.2. Pore Measurement

The black poplar stomatal pore parameters of the ground truth value and the corresponding predicted values were calculated, and the corresponding relative error of measurement was acquired. The error results (the average value after 10-fold cross validation) are shown in Table 2, and a scatterplot of the true values against the predicted values is shown in Figure 8.

Table 2. Anatomy parameter calculation results for stomatal pores.

| Number of Stomatal Pore | Average Pore Length Accuracy (%) | Average Pore Width Accuracy (%) | Average Area Accuracy (%) | Average Eccentricity Accuracy (%) | Average the Degree of Stomatal Opening Accuracy (%) |
|-------------------------|---------------------------------|---------------------------------|---------------------------|----------------------------------|-----------------------------------------------|
| 2201                    | 94.66                           | 93.54                           | 90.73                     | 99.09                            | 92.95                                         |

Figure 7. (a)–(d) are examples of the segmentation results (the color mask represents the pore area).

3.2. Pore Measurement

The black poplar stomatal pore parameters of the ground truth value and the corresponding predicted values were calculated, and the corresponding relative error of measurement was acquired. The error results (the average value after 10-fold cross validation) are shown in Table 2, and a scatterplot of the true values against the predicted values is shown in Figure 8.

Table 2. Anatomy parameter calculation results for stomatal pores.

| Number of Stomatal Pore | Average Pore Length Accuracy (%) | Average Pore Width Accuracy (%) | Average Area Accuracy (%) | Average Eccentricity Accuracy (%) | Average the Degree of Stomatal Opening Accuracy (%) |
|-------------------------|---------------------------------|---------------------------------|---------------------------|----------------------------------|-----------------------------------------------|
| 2201                    | 94.66                           | 93.54                           | 90.73                     | 99.09                            | 92.95                                         |
Figure 8. The scatterplot of the automatically quantified stomatal apertures versus the manually quantified apertures.

The calculation process is as follows: (1) Each manually labeled stomatal pore corresponding to the mask generated by the model is found according to the maximized IoU; (2) through the ‘Stomatal Pore Measurement’ section and the formulas proposed in Section 2.3, evaluation indices, the anatomical parameters of the manually labeled pore region, the anatomical parameters of the mask generated by the model are obtained; and (3) the relative errors between them can be calculated the average relative error for each of the pore anatomical parameters is summed separately for each respective pore and divided by the total number of pores.

In addition, the relationship between the stomatal aperture and the measurement error was analyzed. Based on 10-fold cross validation, the number of pores with a stomatal aperture larger than 40% was an average of 469, that larger than 30% and less than 40% was an average of 954, that larger than 20% and less than 30% was an average of 646, and that larger than 10% and less than 20% was an average of 91. The relationship between the degree of stomatal opening and the average measurement error is shown in Figure 9. It can be seen that the accuracy of measurement results was proportional to the degree of stomatal opening.

Figure 9. Relationship between stomata aperture and measurement error.
3.3. Algorithm Comparison

The proposed method was compared with Li’s method [30] in terms of measurement accuracy and time. Li’s method used a Faster R-CNN and CV model to detect and segment pores, and the anatomical parameters of pores were obtained with an ellipse fitting technique. The disadvantages of this method are as follows: (1) Only one single stoma can be used at once during stomatal pore segmentation, resulting in a long overall process time (about 1.58 s per stoma); (2) the CV model needs to be manually adjusted in the process of segmenting pores; and (3) Li’s method is incapable of fitting an incomplete pore at the boundary of an image, two different examples of original images as shown in Figure 10a (top row and bottom row), and corresponding failed segmentation examples as shown in Figure 10b. Our method improves upon the drawbacks of Li’s method. First, we only used a network model, the Mask R-CNN, to achieve pore positioning and segmentation, resulting in a great improvement in measurement speed. Second, the neural network adjusted the parameters according to the stomatal pore characteristics of the training set to obtain a network model with a good degree of generalization, thus eliminating the need to adjust the parameters twice. Third, the proposed method could segment and perfectly fit the pore at the boundary of the image (top row and bottom row of Figure 10a), the corresponding segmentation and ellipse fitting results as shown in Figure 10c,d, respectively.

![Figure 10. Comparison of different segmentation methods. Top row: the condition one. Bottom row: the condition two. (a) The original image, (b) an example of failed segmentation using Li’s methods, (c) segmentation results of the proposed method, and (d) the ellipse fitting results of the proposed method.](image)

In this experiment, we used Li’s method to test our data. The comparison results are given in Table 3.

| Methods                      | Li’s   | Proposed |
|------------------------------|--------|----------|
| Average pore length error    | 16.8%  | 5.3%     |
| Average pore width error     | 19.3%  | 6.5%     |
| Average area error           | 37.2%  | 9.27%    |
| Average eccentricity error   | 1.5%   | 0.91%    |
| Average stomatal aperture error | 13%   | 7.05%    |

In Table 3, it can be seen that the errors for the proposed method were clearly smaller than those of Li’s method.

3.4. Model Generalization Ability

To test the generalization ability of the proposed model, we tested the pores of two tree species: ginkgo and poplar. The corresponding datasets can be found in [39].
The images of the poplar were divided into (1) 60, (2) 20, and (3) 20 image sets, and those of the ginkgo were divided into (1) 55, (2) 18, and (3) 18 image sets ((1) training set, (2) validation set, and (3) test set). We only used the test dataset of the poplar and ginkgo to test the performance of the trained model based on the black poplar dataset without fine-tuning. The process of fine-tuning was on the basis of the model trained with black poplar. The training sets for poplar and ginkgo were used to train the model with transfer learning.

A comparison of the segmentation results of the model without and with fine-tuning for poplar is shown in Figure 11.

![Figure 11. Comparison of the segmentation results of the model without and with fine tuning for poplar: (a) original image, (b) ground truth labeled manually, (c) segmentation results without fine tuning, and (d) segmentation results with fine tuning.](image-url)
The model did not show a good generalization ability without fine-tuning because it lacked diversity in sample features since we only used black poplar as the training dataset. However, the model performed well with fine-tuning when using a small dataset. Table 4 shows the improvement of the generalization ability of the model with fine-tuning through transfer learning. After the model was retrained with a small dataset of a different tree species, the detection precision and recall rates were greatly improved, thus indicating that this model has some generalization ability.

Table 4. The difference between the precision rate and the recall rate of the model without and with fine-tuning.

| Dataset | Without Fine-Tuning | With Fine-Tuning |
|---------|---------------------|-----------------|
|         | Precision | Recall | Precision | Recall |
| Ginkgo  | 64.6%     | 32.4%   | 84.7%     | 69%    |
| Poplar  | 12.4%     | 7.2%    | 76.5%     | 80%    |
Table 5 shows that the model performed well in parameter measurement with fine-tuning.

| Dataset  | Relative Error (%) |
|----------|--------------------|
|          | Area   | Pore Length | Pore Width | Eccentricity | Stomatal Aperture |
| Ginkgo   | 13.65  | 7.5        | 10.83      | 0.97         | 13.65            |
| Poplar   | 19.7   | 7.79       | 14.1       | 1.72         | 11.69            |

### 4. Discussion

The results of the experiment demonstrated that the proposed method could obtain a higher segmentation accuracy for most stomatal pores and does not require non-uniform illumination correction for leaf microscope images. For black poplar, the pore detection accuracy of the proposed method was 96.72% and the recall rate was 96.87%. Examples of the stomata that our method failed to detect are shown in Figure 13. Reasons for this failure include that the degree of stomatal aperture was too small, the improper operation of the microscope resulted in fuzzy images, and the presence of impurities such as trichomes in the pores.

![Figure 13. The case of failed prediction in the model.](image)

In the first example, the stomatal aperture was too small, as shown in the first row in Figure 13. After measurement, the minimum threshold at which the model can be segmented was 1.86 pixels, about 0.38 μm. Secondly, stomatal pores were blurred due to the improper operation of the microscope, which prevented the successful detection of the pores, as shown in the second row in Figure 13. Lastly, because the dataset was composed of living stomata, there were many impurities such as trichomes that could be found in the pores of stomata, thus preventing their successful detection by the algorithm, as shown in the third row in Figure 13.

In this work, when the pores were manually labeled, we made a label rule that the pores at the edge of the image had to be exposed by more than 50% before being considered pores. Otherwise, they were regarded as part of the background of the image. As shown in Figure 14, when we manually conducted labeling (whether it was for the training, validation, or test sets), we considered these pores as part of the background, and they remained unlabeled. When the proposed model was used for testing, similar to the first picture in the last line of Figure 14, they were detected as pores, such that the number of pores detected by the algorithm was slightly greater than the number of manually labeled pores and the precision rate of the algorithm was lower than expected.
In addition, the proposed method incorrectly identified non-stomatal pores as pores due to the influence of guard cells and leaf background colors in the microscope images of living plant leaves, as shown in Figure 15, thus reducing the precision rate of the proposed method.

Figure 14. The case of error prediction in the model (the black area is the automatic filling of the image by the algorithm).

Figure 15. Cases of algorithm misjudgment (the red mask in the figure represents the pore area predicted by the model).

5. Conclusions

In this work, an automatic method for the segmentation and measurement of plant stomatal pores based on the Mask R-CNN model was proposed. The method consists of three parts: segmentation based on the Mask R-CNN model, the extraction of contour coordinates for pore region segmentation results for ellipse fitting, and the calculation of pore anatomy parameters. After 10-fold cross validation by segmenting and measuring an average of 2201 pores, the average measurement accuracies of 10 results of the (1) pore length, (2) pore width, (3) area, (4) eccentricity, and (5) stomatal aperture were (1) 94.66%, (2) 93.54%, (3) 90.73%, (4) 99.09%, and (5) 92.95%, respectively. The experimental results showed that the proposed method can provide more accurate stomatal pore anatomy parameters in comparison with state-of-the-art stomatal segmentation methods. After fine-tuning with small datasets in the optimized model, the proposed method also demonstrated good performance when applied to other species, which could reduce the labeling workload of researchers to some extent. In future work, we will apply the proposed method to the stomata of more plant species and further improve the generalization ability of the proposed model.

Author Contributions: J.L. constructed the stomatal pore detection, segmentation, and measurement model; carried out the comparison experiment; concluded the results; and drafted the manuscript. K.L. conducted the comparison experiments and contributed to writing the paper. J.H. supplied the main goals of the project and contributed to writing the paper. W.S. identified the main goals of the project and supervised the project. J.C. acquired the microscope images for the research. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China (NO. 31470714, 61701105) and the Fundamental Research Funds for the Central Universities (FRFCU, NO. 2572016CB03), and the project was funded by China Postdoctoral Science Foundation (NO. 2017M610199).

Acknowledgments: We thank the National Natural Science Foundation of China, the Fundamental Research Funds for the central Universities, and the Project funded by China Postdoctoral Science Foundation for funding the research and Kexin Li for help with supplying the black poplar.

Conflicts of Interest: The authors declare no conflict of interest.
References

1. Fanourakis, D.; Bouranis, D.; Giday, H.; Carvalho, D.R.; Nejad, A.R.; Ottosen, C.O. Improving stomatal functioning at elevated growth air humidity: A review. *J. Plant Physiol.* 2016, 207, 51–60. [CrossRef] [PubMed]

2. Fanourakis, D.; Aliniaieifard, S.; Sellin, A.; Giday, H.; Körner, O.; Nejad, A.R.; Delis, C.; Bouranis, D.; Koubouris, G.; Kambourakis, E.; et al. Stomatal behavior following mid- or long-term exposure to high relative air humidity: A review. *Plant Physiol. Biochem.* 2020, 153, 92–105. [CrossRef]

3. Fanourakis, D.; Nikoloudakis, N.; Pappi, P.; Markakis, E.; Doupis, G.; Charova, S.N.; Delis, C.; Tsaniklidis, G. The role of proteases in determining stomatal development and tuning pore aperture: A review. *Plants* 2020, 9, 340. [CrossRef] [PubMed]

4. Hetherington, A.M.; Woodward, F.I. The role of stomata in sensing and driving environmental change. *Nature* 2003, 424, 901–908. [CrossRef] [PubMed]

5. Fanourakis, D.; Hyldgaard, B.; Giday, H.; Aulik, I.; Bouranis, D.; Körner, O.; Ottosen, C.-O. Stomatal anatomy and closing ability is affected by supplementary light intensity in rose (Rosa hybrida L.). *Hortic. Sci.* 2019, 46, 81–89. [CrossRef]

6. Giday, H.; Kjaer, K.H.; Fanourakis, D.; Ottosen, C.O. Smaller stomata require less severe leaf drying to close: A case study in rosa hybrida. *J. Plant Physiol.* 2013, 170, 1309–1316. [CrossRef]

7. Carvalho, D.R.A.; Fanourakis, D.; Correia, M.J.; Monteiro, J.A.; Araújo-Alves, J.P.L.; Vasconcelos, M.W.; Almeida, D.P.F.; Heuvelink, E.; Carvalho, S.M.P. Root-to-shoot aba signaling does not contribute to genotypic variation in stomatal functioning induced by high relative air humidity. *Environ. Exp. Bot.* 2016, 123, 13–21. [CrossRef]

8. Franks, P.J.; Beerling, D.J. Maximum leaf conductance driven by co2 effects on stomatal size and density over geologic time. *Proc. Natl. Acad. Sci. USA* 2009, 106, 10343–10347. [CrossRef]

9. Zhu, J.; Yu, Q.; Xu, C.; Li, J.; Qin, G. Rapid estimation of stomatal density and stomatal area of plant leaves based on object-oriented classification and its ecological trade-off strategy analysis. *Forests* 2018, 9, 616. [CrossRef]

10. Fanourakis, D.; Giday, H.; Milla, R.; Pieruschka, R.; Kjaer, K.H.; Bolger, M.; Vasilevski, A.; Nunes-Nesi, A.; Fiorani, F.; Ottosen, C.O. Port size regulates operating stomatal conductance, while stomatal densities drive the partitioning of conductance between leaf sides. *Ann. Bot.* 2015, 115, 555–565. [CrossRef]

11. Fanourakis, D.; Heuvelink, E.; Carvalho, S.M.P. Spatial heterogeneity in stomatal features during leaf elongation: An analysis using rosa hybrida. *Funct. Plant Biol.* 2015, 42, 737–745. [CrossRef] [PubMed]

12. Fanourakis, D.; Giday, H.; Hyldgaard, B.; Bouranis, D.; Körner, O.; Ottosen, C.O. Low air humidity during cultivation promotes stomatal closure ability in rose. *Eur. J. Hortic. Sci.* 2019, 84, 245–252. [CrossRef]

13. Fanourakis, D.; Bouranis, D.; Tsaniklidis, G.; Nejad, A.R.; Ottosen, C.-O.; Woltering, E.J. Genotypic and phenotypic differences in fresh weight partitioning of cut rose stems: Implications for water loss. *Acta Physiol. Plant.* 2020, 42, 1–10. [CrossRef]

14. Fanourakis, D.; Giday, H.; Li, T.; Kambourakis, E.; Lisogixagis, E.K.; Papadimitriou, M.; Stratidakis, A.; Bouranis, D.; Fiorani, F.; Heuvelink, E.; et al. Antitranspirant compounds alleviate the mild-desiccation-induced reduction of vase life in cut roses. *Postharvest Biol. Technol.* 2016, 117, 110–117. [CrossRef]

15. Zhang, L.; Liu, L.; Zhao, H.; Jiang, Z.; Cai, J. Differences in near isohydric and anisohydric behavior of contrasting poplar hybrids (i-101 (populus alba l.) × 84k (populus alba l. × populus glandulosa uyeki)) under drought-rehydration treatments. *Forests* 2020, 11, 402. [CrossRef]

16. Fanourakis, D.; Hyldgaard, B.; Giday, H.; Bouranis, D.; Körner, O.; Nielsen, K.L.; Ottosen, C.-O. Differential effects of elevated air humidity on stomatal closing ability of kalanchoe blossfeldiana between the c 3 and cam states. *Environ. Exp. Bot.* 2017, 143, 115–124. [CrossRef]

17. Taylor, S.H.; Franks, P.J.; Hulme, S.P.; Spriggs, E.; Christin, P.A.; Edwards, E.J.; Woodward, F.I.; Osborne, C.P. Photosynthetic pathway and ecological adaptation explain stomatal trait diversity amongst grasses. *New Phytol.* 2012, 193, 387–396. [CrossRef]

18. Sørensen, H.K.; Fanourakis, D.; Tsaniklidis, G.; Bouranis, D.; Nejad, A.R.; Ottosen, C.-O. Using artificial lighting based on electricity price without a negative impact on growth, visual quality or stomatal closing response in passiflora. *Sci. Hortic.* 2020, 267, 109354. [CrossRef]

19. Brown, H.T.; Escombe, F. Static diffusion of gases and liquids in relation to the assimilation of carbon and translocation in plants. *Ann. Bot.* 1900, os-14, 537–542. [CrossRef]
20. Franks, P.J.; Drake, P.L.; Beerling, D.J. Plasticity in maximum stomatal conductance constrained by negative correlation between stomatal size and density: An analysis using eucalyptus globulus. *Plant Cell Environ.* **2009**, *32*, 1737–1748. [CrossRef]

21. Roth-Nebelsick, A.; Grein, M.; Utescher, T.; Konrad, W. Stomatal pore length change in leaves of eotrigonobalanus furcinervis (fagaceae) from the late eocene to the latest oligocene and its impact on gas exchange and co2 reconstruction. *Rev. Palaeobot. Palynol.* **2012**, *174*, 106–112. [CrossRef]

22. Fanourakis, D.; Heuvelink, E.; Carvalho, S.M. A comprehensive analysis of the physiological and anatomical components involved in higher water loss rates after leaf development at high humidity. *J. Plant Physiol.* **2013**, *170*, 890–898. [CrossRef]

23. Kuznichov, D.; Zvirin, A.; Honen, Y.; Kimmel, R. Data augmentation for leaf segmentation and counting tasks in rosette plants. *arXiv* **2019**, arXiv:1903.08583.

24. Omasa, K.; Onoe, M. Measurement of stomatal aperture by digital image processing. *Plant Cell Physiol.* **1984**, *25*, 1379–1388. [CrossRef]

25. Laga, H.; Shahinnia, F.; Fleury, D. Image-based plant stomata phenotyping. In Proceedings of the 13th International Conference on Controle, Automation, Robotics and Vision (ICARCV 2014), Marina Bay Sands, Singapore, 10–12 December 2014; pp. 217–240.

26. Liu, S.; Tang, J.; Petrie, P.; Whitty, M. A Fast Method to Measure Stomatal Aperture by Mser on Smart Mobile Phone. In Proceedings of the Applied Industrial Optics: Spectroscopy, Imaging and Metrology 2016, Heidelberg, Germany, 25–28 July 2016.

27. Jayakody, H.; Liu, S.; Whitty, M.; Petrie, P. Microscope image based fully automated stomata detection and pore measurement method for grapevines. *Plant Methods* **2017**, *13*, 94. [CrossRef]

28. Todo, Y.; Toh, S.; Bourdais, G.; Robatzek, S.; Maclean, D.; Kinoshita, T. Deepstomata: Facial recognition technology for automated stomatal aperture measurement. *bioRxiv* **2018**. [CrossRef]

29. Bhugra, S.; Mishra, D.; Anupama, A.; Chaudhury, S.; Lall, B.; Chugh, A.; Chinnusamy, V. Deep convolutional neural networks based framework for estimation of stomata density and structure from microscopic images. In Proceedings of the Computer Vision—ECCV 2018 Workshops, Munich, Germany, 8–14 September 2018; pp. 412–423.

30. Li, K.; Huang, J.; Song, W.; Wang, J.; Lv, S.; Wang, X. Automatic segmentation and measurement methods of living stomata of plants based on the cv model. *Plant Methods* **2019**, *15*, 67. [CrossRef]

31. Wada, K. Labelme: Image Polygonal Annotation with Python. Available online: https://github.com/wkentaro/labelme (accessed on 6 July 2020).

32. Halır, R.; Flusser, J. Numerically stable direct least squares fitting of ellipses. In Proceedings of the 6th International Conference in Central Europe on Computer Graphics and Visualization, Plzen-Bory, Czech Republic, 9–13 February 1998.

33. Liang, Y.K.; Dubos, C.; Dodd, I.C.; Holroyd, G.H.; Hetherington, A.M.; Campbell, M.M. Atmyb61, an r2r3-myb transcription factor controlling stomatal aperture in arabidopsis thaliana. *Curr. Biol.* **2005**, *15*, 1201–1206. [CrossRef]

34. Fetter, K.C.; Eberhardt, S.; Barclay, R.S.; Wang, S.; Keller, S.R. Stomatacounter: A neural network for automatic stomata identification and counting. *New Phytol.* **2019**, *223*, 1671–1681. [CrossRef]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).