

Stomata segmentation using deep learning

Miguel Alonso, Angela Casado-García, Jónathan Heras
Department of Mathematics and Computer Science
University of La Rioja, Spain

{miguel.alonsoa, angela.casado, jonathan.heras}@unirioja.es

Abstract

Stomata are pores in the epidermal tissue of leaf plants formed by specialised cells called guard cells, which regulate the stomatal opening. Stomata facilitate gas exchange, being pivotal in the regulation of processes such as photosynthesis and transpiration. The analysis of the number and behaviour of stomata is a task carried out by studying microscopic images; and, nowadays, this task is mainly conducted manually, or using programs that can count and determine the position of stomata but are not able to determine their morphology. In this paper, we have conducted a study of 10 deep learning algorithms to segment stomata from several species. The model that achieves the best Dice score, with a value of 96.06%, is obtained with the DeepLabV3+ algorithm, whereas the model that provides the best trade-off between inference time and Dice score was trained using the ContextNet architecture. This is a first step towards improving the measurements provided by stomata analysis tools, that will in turn help plant biologists to advance their understanding of dynamics in plants.

1. Introduction

Stomata (singular “stoma”) are pores on a plant leaf that allow the exchange of gases, mainly CO₂ and water vapor, between the atmosphere and the plant. Stomata respond to changes in the environment and are key regulators of the photosynthesis and transpiration of plants, and thus their productivity and water use efficiency. Because of their critical nature, scientists have studied the number, density, size, and behaviour of stomata to understand plant physiology against stress [1, 7], to improve crop management programs [1, 9], to model CO₂ dynamics in the atmosphere, and to predict future carbon and water cycles [7].

In order to analyse stomata, plant biologists take microscopic images of leaves, and manually measure characteristics such as stomata density, individual stomata opening, and morphological traits like the size and shape of the stomata guard cells (a pair of cells that regulate the opening

and closing of the stomatal pore) using programs like ImageJ [15] that have little to no automatization [2, 5]. Those measurements are repeated over hundreds of images from different plant species and growth conditions. This is a tedious, error-prone, time-consuming and subjective task due to the large number of stomata in each image.

Researchers have tried to automatise this task since the 1980s [11], and we can currently find several open-source applications like StomataCounter [6], DeepStoma [16], or LabelStoma [3] that help biologists in the task of counting and measuring stomata. However, those software tools output either the number and position of stoma, or bounding boxes around the detected stomata. Hence, additional processing steps are required to study stomata morphology; and, namely to measure the stomata boundary. This task can be framed as a semantic segmentation problem where pixels of a given image must be classified as either background or as stoma.

Nowadays, semantic segmentation tasks, as most computer vision problems, are tackled by using deep learning techniques [4, 14]; and, this is also the approach followed in this work. Namely, we present a thorough study of deep learning models to segment a stoma from the bounding box that contains it. In this way, our models can be integrated with the aforementioned applications to improve their functionality; and, therefore, help plant biologists in their studies to understand the processes associated with plant gas exchange, and associated with carbon and water cycles. The code, models and datasets associated with this work are available at <https://github.com/mialona/Stomatal-segmentation>.

2. Materials and methods

In our experiments, we have employed the training datasets of StomataCounter [6] and LabelStoma [3]. Such a combined dataset consists of 1055 images from 67 species and was annotated with bounding boxes that indicate the position of the stomata. From the annotated images, we cropped each stoma obtaining 25214 images, see Figure 1. Since manually annotating such an amount of images is

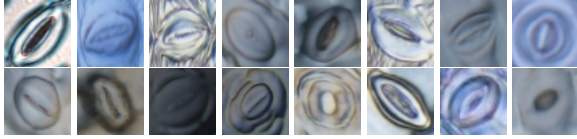


Figure 1. Stomata samples from different species

a time-consuming a tedious task, we implemented several image processing algorithms (based on the combination of morphological operations and thresholding methods) to obtain a mask with the boundary of the stomata. The output produced by the algorithms was manually validated, and we chose the best annotation for each image. This process produced a total of 15686 annotated images, and in addition, we manually annotated 2404 extra images. It is worth noting that none of the image processing algorithms was able to produce a correct annotation for all the stomata; so, we tested several deep segmentation algorithms to tackle this task.

In order to construct the deep segmentation models from the annotated dataset, we employed 14472 (80%) images for training, and the remaining 3618 (20%), for testing — in this split, images from the same species only belong to either the training or the testing set. From the training dataset, we fine-tuned several deep-learning segmentation algorithms. Namely, we have trained 10 architectures: BiSeNet [21], CGNet [19], ContextNet [13], DeepLabV3+ [4], DenseAPP [20], FPENet [10], HRNet-Seg [17], LEDNet [18], OCNNet [22] and U-Net [14]. All the architectures were trained with the libraries PyTorch [12] and FastAI [8] and using the GPU provided by Google Colaboratory. In order to set the learning rate for the different architectures, we employed the procedure presented in [8]; and, we applied early stopping in all the architectures to avoid overfitting using a validation set taken from the training set. As a result of the training process, 10 models were produced that can be used for inference by providing them a stoma image.

3. Results

The trained models were evaluated on the testing set using the Dice score and the Jaccard index as evaluation metrics [14], the results are presented in Table 1. It is worth noting that all the deep segmentation models achieve a Dice score over 90%, and that the best model is obtained using the DeepLabV3+ architecture. We also include in the comparison the best image processing algorithm that was initially employed to annotate the images, but it only achieved a Dice score of 84.42%, far from the results obtained by the deep segmentation models.

Since the final aim of this project is to incorporate one of the analysed models into a pipeline that extracts the boundaries of hundreds of stomata images, it is also important to

Algorithm	Dice score	Jaccard index
Image processing	84.42	76.20
BiSeNet	95.98	92.26
CGNet	95.91	92.13
ContextNet	95.33	91.08
DeepLabV3+	96.06	92.42
DenseAPP	95.91	92.14
FPENet	95.54	91.46
HRNet	95.95	92.21
LEDNet	91.40	84.16
OCNet	95.87	92.06
UNet	94.07	88.80

Table 1. Results obtained by the segmentation models. In bold the best result

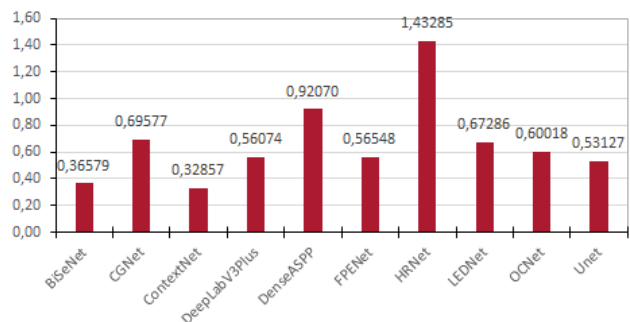


Figure 2. Inference time (in seconds) of the trained segmentation models on a set of 100 images

study the inference time of each model, see Figure 2. Using this metric, the best model is provided by the ContextNet architecture that takes 0.32 seconds for processing a set of 100 images. This is also the model that provides the best trade-off between inference time and Dice score.

4. Conclusions and further work

In this paper, we have presented a study of several deep learning architectures for segmenting stomata. The model that provides the best trade-off between inference time and Dice score is obtained with the ContextNet architecture. Such a model generalises properly to images from species that were not included in the training set and has a reasonable inference time. As further work, it remains the task of incorporating the best model to a software tool that facilitates its usage. This will also require the implementation of a mechanism that allows users to validate the generated segmentations and modify them if needed. The suggested approach requires the combination of a detection algorithm with our segmentation model; so, it will be also worth exploring the usage of a model that can directly segment all the stomata of an image instead of processing them individually.

References

- [1] B. Buttery, C. S. Tan, R. I. Buzzell, et al. Stomatal numbers of soybean and response to water stress. *Plant and Soil*, 149(2):283–288, 1993.
- [2] R.S. Caine, X. Yin, J. Sloan, et al. Rice with reduced stomatal density conserves water and has improved drought tolerance under future climate conditions. *New Phytologist Trust*, 221:371–384, 2019.
- [3] Á. Casado-García et al. LabelStoma: A Tool for Stomata Detection based on the YOLO algorithm. *Computers and Electronics in Agriculture*, 178:105751, 2020.
- [4] L. C. Chen, Y. Zhu, G. Papandreou, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Computer Vision – ECCV 2018*, pages 833–851. Springer International Publishing, 2018.
- [5] C. Dutton, H. Hōrak, Christopher Hepworth, et al. Bacterial infection systemically suppresses stomatal density. *Plant, Cell & Environment*, 42(8):2411–2421, 2019.
- [6] K. C. Fetter, S. Eberhardt, R. S. Barclay, et al. Stomata-Counter: a neural network for automatic stomata identification and counting. *New Phytologist*, 223(3):1671–1681, 2019.
- [7] A. M. Hetherington and F. I. Woodward. The role of stomata in sensing and driving environmental change. *Nature*, 424(6951):901–908, 2003.
- [8] J. Howard and R. Thomas. Practical deep learning for coders. <https://course.fast.ai/>, 2019.
- [9] J. Hughes, C. Hepworth, C. Dutton, et al. Reducing Stomatal Density in Barley Improves Drought Tolerance without Impacting on Yield. *Plant Physiology*, 174(2):776–787, 2017.
- [10] M. Liu and H. Yin. Feature pyramid encoding network for real-time semantic segmentation. In *30th British Machine Vision Conference*, page 1031, 2019.
- [11] K. Omasa and M. Onoe. Measurement of Stomatal Aperture by Digital Image Processing. *Plant and Cell Physiology*, 25(8):1379–1388, 1984.
- [12] A. Paszke, S. Gross, F. Massa, et al. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019.
- [13] R. Poudel et al. Contextnet: Exploring context and detail for semantic segmentation in real-time. In *29th British Machine Vision Conference*, page 146, 2018.
- [14] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, volume 9351 of *Lecture Notes in Computer Science*, pages 234–241, 2015.
- [15] C. T. Rueden, J. Schindelin, M. C. Hiner, et al. ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics*, 18:529, 2017.
- [16] Y. Toda, S. Toh, G. Bourdais, et al. DeepStomata: Facial Recognition Technology for Automated Stomatal Aperture Measurement. *bioRxiv*, 2018.
- [17] J. Wang, K. Sun, T. Cheng, et al. Deep high-resolution representation learning for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020.
- [18] Yu Wang et al. Lednet: A lightweight encoder-decoder network for real-time semantic segmentation. *arXiv preprint arXiv:1905.02423*, 2019.
- [19] Tianyi Wu, Sheng Tang, Rui Zhang, Juan Cao, and Yongdong Zhang. Cgnet: A light-weight context guided network for semantic segmentation. *IEEE Transactions on Image Processing*, 30:1169–1179, 2020.
- [20] M. Yang et al. Denseaspp for semantic segmentation in street scenes. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, 2018.
- [21] Changqian Yu, Jingbo Wang, Chao Peng, Changxin Gao, Gang Yu, and Nong Sang. Bisenet: Bilateral segmentation network for real-time semantic segmentation. In *Computer Vision – ECCV 2018*, pages 334–349, Cham, 2018. Springer International Publishing.
- [22] Yuhui Yuan, Xilin Chen, and Jingdong Wang. Object-contextual representations for semantic segmentation. 2020.