



OPEN

## Automated stomata detection in oil palm with convolutional neural network

Qi Bin Kwong<sup>1✉</sup>, Yick Ching Wong<sup>1</sup>, Phei Ling Lee<sup>2</sup>, Muhammad Syafiq Sahaini<sup>2</sup>, Yee Thung Kon<sup>1</sup>, Harikrishna Kulaveerasingam<sup>1</sup> & David Ross Appleton<sup>1</sup>

Stomatal density is an important trait for breeding selection of drought tolerant oil palms; however, its measurement is extremely tedious. To accelerate this process, we developed an automated system. Leaf samples from 128 palms ranging from nursery (1 years old), juvenile (2–3 years old) and mature (> 10 years old) were collected to build an oil palm specific stomata detection model. Micrographs were split into tiles, then used to train a stomata object detection convolutional neural network model through transfer learning. The detection model was then tested on leaf samples acquired from three independent oil palm populations of young seedlings (A), juveniles (B) and productive adults (C). The detection accuracy, measured in precision and recall, was 98.00% and 99.50% for set A, 99.70% and 97.65% for set B, and 99.55% and 99.62% for set C, respectively. The detection model was cross-applied to another set of adult palms using stomata images taken with a different microscope and under different conditions (D), resulting in precision and recall accuracy of 99.72% and 96.88%, respectively. This indicates that the model built generalized well, in addition has high transferability. With the completion of this detection model, stomatal density measurement can be accelerated. This in turn will accelerate the breeding selection for drought tolerance.

The key to sustainable agriculture and managing the impact of climate change is selective breeding. Through selective breeding, the global increase in demand for food can be achieved through improvement in yield per unit area, thereby not requiring new agricultural land. Particularly in the case of oil palm, this means a halt to deforestation, therefore conserving the habitats for biodiversity. In the past, selective breeding has been performed using phenotypic selection or limited marker assisted selection. The advancement in genomic sciences has allowed for high throughput, whole genomic marker discovery and deployment for selective breeding. Whole genomic marker-based selection is known as genomic prediction or selection<sup>1</sup>, which has seen a lot of success from cattle to maize breeding<sup>2,3</sup>. Genomic selection is usually conducted by regressing the trait values against the genotypes, resulting in a model that can be used to predict for individuals with known genetic information but unknown trait performance. Since the conceptual introduction of it in oil palm<sup>4</sup>, genomic selection has since been assessed and applied in different oil palm breeding programs<sup>5,6</sup>.

Even though high throughput genotyping has been made possible<sup>7,8</sup>, the phenotyping process usually still relies on manual measurement and recording. This, together with the fact that breeders are usually interested in multiple traits, creates a phenomenon known as phenotyping bottleneck<sup>9</sup>. Time taken for phenotyping essentially limits progress in plant breeding programs, including drought tolerance<sup>10</sup>. Therefore, it is crucial to automate phenotyping processes in order to accelerate selective breeding. The first step towards automating the entire phenotyping process often lies in image analytics and machine learning. Essentially, once an image of a plant has been captured through a device, machine learning can be employed to automate the measurement, counting or classification of the objects in the image. Applications developed based on this include Deep Plant Phenomics, which is capable of leaf counting and mutant classification<sup>11</sup>.

One of the key traits of interest for oil palm is drought tolerance, particularly related to minimizing climate change impact. There are many methods to define drought tolerance, including taking a range of leaf measurements<sup>12</sup>. The most direct measurement of drought tolerance is based on water-use efficiency and yield under water-limiting conditions<sup>13</sup>. Another alternative is to measure the stomatal density. Biologically, stomata are an important “gate” for both gas exchange (CO<sub>2</sub>) of the leaf during photosynthesis and water vapor during transpiration<sup>14</sup>. Previous studies have shown that photosynthesis and efficiency of water usage in plants were

<sup>1</sup>Sime Darby Plantation Technology Centre Sdn Bhd, Serdang, Selangor Darul Ehsan, Malaysia. <sup>2</sup>Department of Biology and Biotechnology, Faculty of Science and Technology, National University of Malaysia, Bangi, Selangor Darul Ehsan, Malaysia. ✉email: kwong.qi.bin@simedarbyplantation.com

affected by the responsiveness, speed and size of stomata<sup>15,16</sup>. Low stomatal density has been reported to increase drought tolerance in many plant species, including barley<sup>17</sup>, rice<sup>18</sup>, wheat<sup>19</sup> and arabidopsis<sup>20</sup>. In addition, regulation capacity of stomatal traits including size and density were also identified as key traits to select for drought tolerance<sup>21</sup>. With that, the focus of this study was to develop an automated method of detecting and counting stomatal density using convolutional neural network (CNN).

## Methodology

**Sample collection.** A total of 128 leaf samples were obtained from advanced commercial oil palms, collected from Sime Darby Plantation nurseries and estates in Carey Island and Banting, Selangor, Malaysia. Among these 128 samples, 42 were collected from nursery seedlings (1 years old (yo)), 43 were from juvenile palms (2–3 yo) planted in estates, and 43 were collected from adult palms (> 10 yo) in commercial estates. For seedlings, bifurcated leaf number 3 was collected, whereas middle leaflets from frond number 17 were collected for juvenile and adult palms.

**Image capture and stomatal characterization.** The bottom surfaces of leaf samples were cleaned with distilled water and air-dried. A thin layer of nail varnish was applied at the middle part of the oil palm leaf (approximately 1 cm × 3 cm in size). After air-drying, a layer of transparent tape was applied firmly on top of the dried nail varnish. The layer of cellofoam tape was then peeled with the nail varnish stuck together and transferred to a clean slide, generating the stomata imprint<sup>22</sup>. Micrographs of size 1800 μm × 1400 μm were acquired with a 10× objective lens using the RGB-mode illumination on an EVOS FL Auto (Thermo Scientific, US) microscope.

Stomatal density and size were recorded and measured manually from the micrographs. Density was counted as the number of stomata objects in 1 mm<sup>2</sup>. The stomatal size in this study was defined as the horizontal length measured from one stoma end to the other. The average stomatal size was estimated based on the size of 150 randomly selected stomata across all images for that particular age group. Statistical *t*-test was carried out to determine significant differences for stomatal density and size measured between each age group.

**Preprocessing.** The acquired 128 micrographs were of resolution 2048 × 1536. Each image was cropped into 12 independent 516 × 512 tiles, resulting in 1536 tiles. Of these tiles, 51 from young seedlings, 171 from juvenile palms, and 88 from adult palms were manually selected to build the CNN stomata detection model. The tiles were manually annotated for stomata objects using labelImg software<sup>23</sup>. Overall, 20,809 stomata objects were annotated.

The acquired tiles were manually inspected, and 193 tiles of the training set were subjected to augmentation using an in-house Python script. Blurring of the images was done using both “bilateralFilter” and “GaussianBlur” functions from OpenCV library<sup>24</sup>. Brightness of the selected tiles were also altered (brightened and darkened) by a factor of 2. After augmentation, there were 49,837 stomata objects.

**Model building.** This study was carried out using Google Cloud Platform’s Compute Engine, with the specification of Tesla T4 GPU, 128 GB RAM and 24 CPU. The stomata detection model was trained using tensorflow<sup>25</sup> library’s object detection API<sup>26</sup>. MobileNet Version 1 (MobileNetV1) is a light weight deep neural network suitable for mobile vision applications, based on a streamlined architecture with depth-wise (dw) separable convolutions<sup>27</sup>. Table 1 summarized the MobileNetV1 architecture. Single Shot Multibox Detector (SSD) is a deep network based accurate object detector that does not resample pixels or features while forming bounding boxes<sup>28</sup>. Instead of the original VGG-16 architecture<sup>29</sup>, the MobileNetV1-based Single Shot Multibox Detector (SSD) (Fig. 1) model trained on COCO dataset<sup>30</sup> was used in this study. The model was retrained to detect stomata through transfer learning, with the initial weights being MobileNetV1’s pretrained weights.

80% of the tiles were used for the model training set, while 20% were used as a validation set. From the 49,837 stomata objects, 39,877 were used for model training, while 9960 were used for validation. During training, L2 regularization was carried out to prevent overfitting. RMSprop and momentum optimizers were also used during training to reduce training time required. After a few rounds of manual running and adjustments, the initial learning rate was set at 0.003, and the decay factor was set at 0.95, decay steps at 100. Number of classes was set at 1 and batch size was set 20. The detection score threshold was set at 0.2.

The training progress was monitored through the use of Tensorboard<sup>25</sup>. The overall loss function (*L*) used was the combination of classification loss and the weighted sum of the localization loss, which is represented as follow:

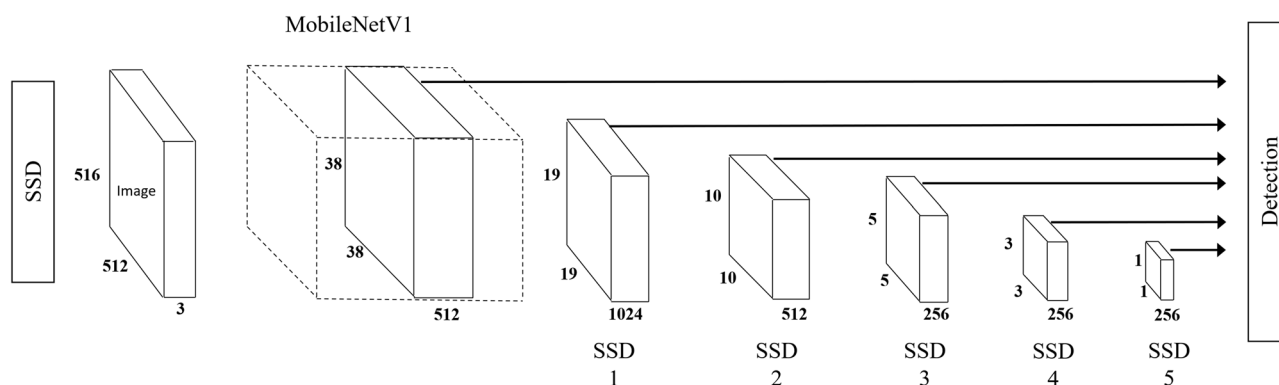
$$L = \frac{1}{N}(L_c + \alpha L_l),$$

where *L<sub>c</sub>*, the classification loss is essentially the softmax loss over the different classes<sup>28</sup>, *L<sub>l</sub>*, the localization loss, is the smooth L1 loss<sup>31</sup> between the predicted bounding box and the ground truth box, *N* is the number of matched default boxes<sup>28</sup> and *α* is the weight factor. The validation metric used to measure the performance of the model was mean average precision (mAP). The model was only saved/updated if it was performing better than the previous checkpoints. Early stopping was employed to prevent model overfitting.

**Application.** Three independent datasets, consisting of 55 leaf samples from 1 yo seedlings (set A), 135 leaf samples from 2 to 3 yo palms (set B) and 100 samples from > 10 yo palms (set C) were collected and used as the application/test set. Similarly, the micrographs were split into tiles. In this case, however, they were split into 20 overlapping (120 pixels overlapping horizontally and 40 pixels vertically) tiles, resulting in 5800 tiles. The stomata objects on these tiles were detected using the model built previously with detection score threshold set at

Type/stride	Filter shape	Input size
Conv/s2	$3 \times 3 \times 3 \times 32$	$224 \times 224 \times 3$
Conv dw/s1	$3 \times 3 \times 32$ dw	$112 \times 112 \times 32$
Conv/s1	$1 \times 1 \times 32 \times 64$	$112 \times 112 \times 32$
Conv dw/s2	$3 \times 3 \times 64$ dw	$112 \times 112 \times 64$
Conv/s1	$1 \times 1 \times 64 \times 128$	$56 \times 56 \times 64$
Conv dw/s1	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv/s1	$1 \times 1 \times 128 \times 128$	$56 \times 56 \times 128$
Conv dw/s2	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv/s1	$1 \times 1 \times 128 \times 256$	$28 \times 28 \times 128$
Conv dw/s1	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv/s1	$1 \times 1 \times 256 \times 256$	$28 \times 28 \times 256$
Conv dw/s2	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv/s1	$1 \times 1 \times 256 \times 512$	$14 \times 14 \times 256$
$5 \times$	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
	$1 \times 1 \times 512 \times 512$	$14 \times 14 \times 512$
Conv dw/s2	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
Conv/s1	$1 \times 1 \times 512 \times 1024$	$7 \times 7 \times 512$
Conv dw/s2	$3 \times 3 \times 1024$ dw	$7 \times 7 \times 1024$
Conv/s1	$1 \times 1 \times 1024 \times 1024$	$7 \times 7 \times 1024$
Avg Pool/s1	Pool $7 \times 7$	$7 \times 7 \times 1024$
FC /s1	$1024 \times 1000$	$1 \times 1 \times 1024$
Softmax /s1	Classifier	$1 \times 1 \times 1000$

**Table 1.** MobileNet body architecture<sup>27</sup>.



**Figure 1.** Diagram showing SSD<sup>28</sup> using MobileNetV1<sup>27</sup> as backbone.

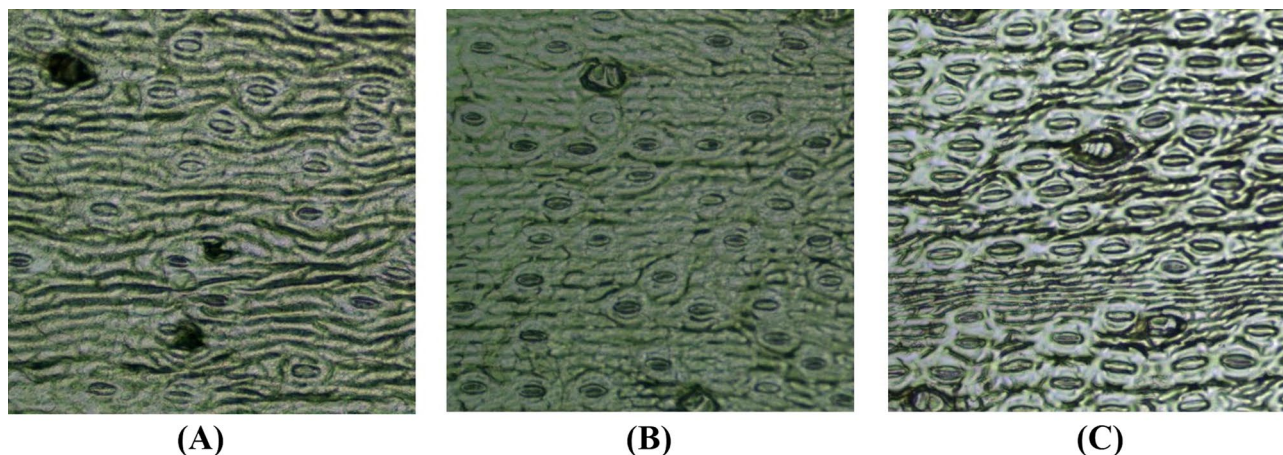
0.2. After prediction, the bounding boxes generated were overlaid on the full-size micrographs using a Python script. Non-maximum suppression was used to solve the issue of overlapping bounding boxes caused by splitting, with the threshold set as 0.01. The acquired results were manually inspected and validated. Accuracy of the prediction was measured in both precision and recall. Precision was defined as true positives divided by total detected cases in the image, whereas recall was defined as true positives divided by total stomata in the image. As this dataset was not annotated, both precision and recall were calculated manually. In this case, correct detections were defined as bounding boxes overlapping with the stomata objects.

**Challenge/limitation.** Another 20 independent leaves were sampled from palms sharing the same genetic background as Set C. However, in this case the micrographs were captured under the  $4\times$  objective lens of an ECLIPSE Ci-L Nikon, Japan microscope. Stomata images of size  $1589 \mu\text{m} \times 1192 \mu\text{m}$  were captured from this step. The model was cross applied onto the current dataset, hereby referred to as Set D.

**Statement of consent.** This study on oil palm complies with relevant institutional, national, international guidelines and legislation. All samples collected are maintained and belong to Sime Darby Plantation R&D, Malaysia.

Set	Palm age (yo)	Count mean (/mm <sup>2</sup> )	Count sd (/mm <sup>2</sup> )	Size mean (μm)	Size sd (μm)
A	1	107 <sup>a</sup>	12	31.82 <sup>a</sup>	2.64
B	2–3	200 <sup>a</sup>	18	34.41 <sup>a,β</sup>	2.67
C	> 10	208	16	43.01 <sup>β</sup>	3.51

**Table 2.** Stomatal count and size across different palm age groups. The stomatal counts and standard deviations (sd) were rounded to its' nearest integer. Significant differences at p-value < 0.05 were represented as α for A versus B, and β for B versus C.



**Figure 2.** Representative stomata tiles for three palms' developmental stages collected in this study. (A) 1 yo palm seedling, (B) 2–3 yo juvenile palm, (C) > 10 yo adult palm.

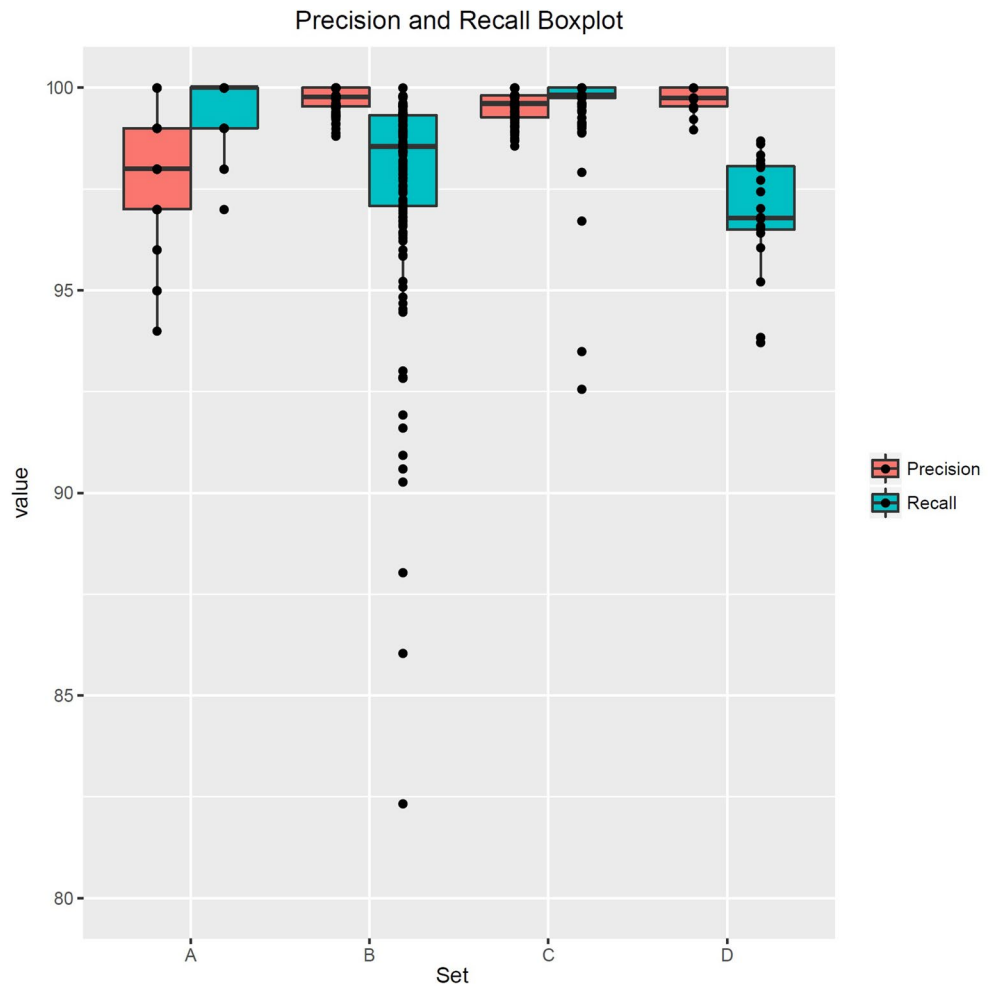
## Results

**Stomatal characterization across different palm developmental stages.** The average stomatal count in an image was  $107 \pm 12$  (mean  $\pm$  standard deviation) for the 1 yo palms (Set A) and  $200 \pm 18$  for 2–3 yo palms (Set B) and  $208 \pm 16$  for > 10 yo palms (Set C). Besides having the highest stomatal count, the stomatal size of Set C was also the largest, at  $43.01 \pm 3.51$  μm, as compared to Set A and Set B ( $31.82 \pm 2.64$  μm and  $34.41 \pm 2.67$  μm, respectively) (Table 2). Example stomata tiles for all three stages can be found in Fig. 2.

**Model and application.** The mean average precision (mAP) of the detection model was 96% (with intersection over union (IoU) threshold set at 0.5) when training was stopped. At this point, the overall loss value was 1.34. The model built was then applied on independent test sets (A, B, C and D). Figure 3 shows the boxplot summarizing precision and recall values acquired for all four test sets. For the image with the highest stomatal density (Supplementary Fig. S1), both precision and recall calculated were 99.83%. As for the image with the lowest stomatal density (Supplementary Fig. S2), precision was 98.91% and recall was 99.45%. The average precision acquired when testing the model on set A was  $98.00 \pm 1.40\%$  (mean  $\pm$  standard deviation), with the corresponding recall being  $99.50 \pm 0.70\%$ . Comparatively, the average precision acquired for set B was  $99.70 \pm 0.29\%$  and the recall was  $97.65 \pm 2.76\%$ . The image with the lowest recall value (82%) can be found in Supplementary Fig. S3. As for set C, the model achieved average precision of  $99.55 \pm 0.37\%$ , and recall of  $99.62 \pm 1.05\%$ . The cross application of the model onto set D dataset exhibited precision of  $99.72 \pm 0.29\%$  and a slight drop in recall of  $96.88 \pm 1.40\%$ . The worst performing image is included as Supplementary Fig. S4. In addition, examples of raw and annotated tiles are shown in Fig. 4, including successful application on an unfocused image (Fig. 4D). Figure 5 shows an example of a fully annotated stomata image using the built model.

## Discussion

To the best of our knowledge, this is the first CNN-based oil palm stomata detection model reported. In other plant species such as grapevine<sup>32</sup> and oak<sup>33</sup>, stomata detection and morphological feature estimation have been automated through cascade object detector utilizing feature extractors such as histogram of oriented gradients. In oak, the estimated precision and recall were 95% and 85%, respectively. As for grapevine, the precision and recall were 92% and 79%, respectively. In these methods, parameters associated with these feature extractors were manually defined. Though the precision values acquired from these studies were very high, the recall values were lower by comparison. Compared to these methods that use specific feature engineering methods, a more recent methodology of automating stomata detection using multiple feature extraction techniques and learning methods was introduced in maize<sup>34</sup>, achieving 97.1% detection accuracy (measured in precision). This study also highlighted the use of deep learning features for stomata detection application. StomataCounter<sup>35</sup> was one

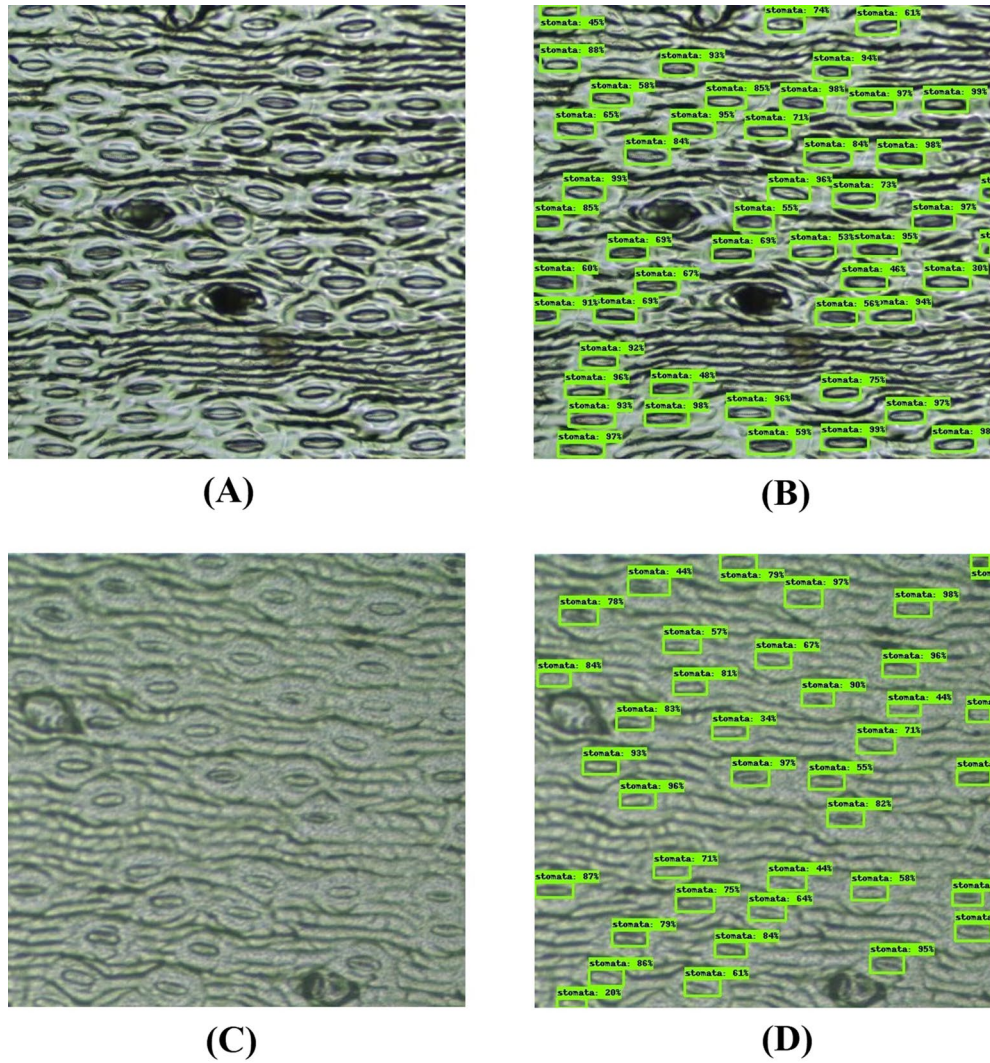


**Figure 3.** Precision and recall boxplot for four test sets (A, B, C and D).

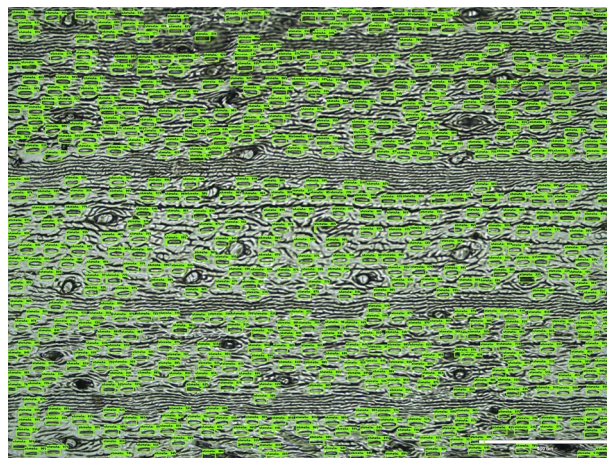
of the earliest efforts in building CNN-based stomata detection model. The model built reached precision of 99% and recall of 93%, indicating that CNN performed better than other methods. It was also noted in the study that model accuracy would decrease significantly when cross-applied on images from another species.

Our current study focused on developing an automated stomata detection model specifically for oil palm. For this purpose, we have selected oil palm samples from different developmental stages or age groups, ranging from young seedling to adult. One of the concerns with only using training dataset images from a single age group was that the resulting detection model might not generalize well across other age groups. From the stomatal characterization performed, we observed significant difference in stomatal density between seedling and juvenile palms. Although the stomatal density of adult palms was the highest, it was not statistically different from juvenile palms. On the other hand, stomatal size measurements showed significant differences across all age groups. It is noted, however, that the size difference observed between juvenile and seedling was far lesser when compared to the difference between adult and juvenile. This indicates that the juvenile palms' (age 2–3 yo) stomatal profile captures the transitional stages of stomata from seedling to adult. By combining these images with images acquired from both seedling and adult palms, the resulting model would be capable of detecting stomata across all age groups in oil palm. Thus, not only can this model be used for yield-related studies in adult palms, it can also be used early screening of drought tolerant palms in nursery.

In general, the accuracy acquired in this study was similar to a previous publication on stomata detection using CNN<sup>35</sup>, with the acquired recall values reported here being slightly higher. This is probably attributed to the preset condition that our model was both trained and expected to only be used for oil palm stomata micrographs, as compared to a general model<sup>35</sup> that is expected to work across different species. The high recall and precision values indicated that the model seldom miss stomata, nor misclassify non-stomata objects. In fact, compared to other non-CNN methods<sup>32,33</sup>, CNN-based detection models achieved far better recall values, indicating that they generalized better. This is probably because parameters involved in CNN were determined through training and not manually decided. In addition, CNN has a deeper architecture which provides exponentially more expressive capability, and its' hierarchical feature representation enables multilevel representation from pixel to high-level features<sup>36</sup>. Although CNN models generalize better, the required condition to build CNN model is that the number of samples needs to be large, which was a limitation faced in previous grapevine and oak



**Figure 4.** Representative tiles before and after automated detection using developed stomata model. The detected stomata were labelled in green bounding boxes together with a detection score on top. (A) Clear tile. (B) Stomata detected on clear tile. (C) Unfocused tile. (D) Stomata detected on unfocused tile.



**Figure 5.** Stomata detection on entire microscopic image. The detected stomata were labelled in green bounding boxes together with a detection score on top.

publications<sup>32,33</sup>. As our final goal is to develop a mobile application that enables field stomatal count phenotyping, the MobileNet CNN architecture, which is known to be light weight was selected. Other CNN architectures, such as VGGNet<sup>29</sup> were not assessed, which can be seen as a limitation of this study.

With large volume of data comes the problem of data quality. Two problems faced when working on micrographs in this study were inconsistent brightness and unfocused images, which were predominantly faced during set B application (example shown in Fig. 4C). This is potentially caused by uneven stomata imprint or non-optimal microscope brightness and contrast settings. In moving towards high-throughput phenotyping using micrographs, this represents one of the practical issues that needs to be solved. Another option is to develop a robust detection model. Instead of discarding these low-quality images, the CNN model was also trained on them as well. In fact, to increase the number of low-quality images, some of the high-quality images were artificially altered to be unfocused and darkened/lightened in a process known as augmentation. The augmentation process also increases the number of annotated images for training, which is often regarded as a taxing and expensive process in developing an object detection model<sup>37,38</sup>. With this, the resulting model was also capable of detecting stomata on most low-quality images (example shown in Fig. 4D). Another method to reduce manual annotation used in this study was only annotating selected tiles from a full-size image instead of annotating every single one. With tiles within the same image sharing similar background properties, it was rather redundant to annotate all of them. Instead, representing every training image with a few tiles allows for building of a more robust model with lesser effort.

Although the resulting oil palm stomata detection model performed well, it was not without limitation. In rare cases, we observed low recall but high precision, which indicates failure of the model in detecting certain stomata objects. Upon inspection of these images (Supplementary Figs. S3, S4), it was found that the undetected stomata fall within the image regions that had a combination of unfocused stomata objects, low-contrast or noisy background. Further improvements to the model can be made through training on images with similar properties. In addition, the stomata objects that were found near the edges of the images were sometimes undetected. A potential solution that can be implemented is to subset the test image by ignoring the last few pixels both vertically and horizontally<sup>33,35</sup>.

One of the key criteria when building a predictive model is that the training dataset must be representative of the test/application set by having similar distribution<sup>37</sup>. In practice, however, it is impractical to assume that images acquired under different microscope and different conditions will share the same underlying distribution. As such, the detection model was tested on another set of images (D) acquired from a different microscope under different conditions such as lighting and image size. Set D achieved the similar precision with B and C at 99.72%, indicating that the model was able to detect stomata in the images from different microscope with low false positives. Despite the slight drop in recall to 96.88%, overall the model performed rather well in this challenge test. In other words, the model can be cross-applied to micrographs generated using other microscopes. Yet, the slight drop in recall also indicates that the detection model needs to be continuously be improved using images acquired under different conditions.

With our main goal being to develop drought tolerant yet productive palms, commercial populations developed from key breeding programs were selected for this study. Stomata detection from micrographs was identified as one of the key time-consuming steps. To ease this phenotyping process, this stomata detection CNN model was developed. The resulting model showed promising result, particularly on adult palms. From our experience, manual stomata counting of a single full image can take up to 5 min, and automating stomata counting on a large dataset can save months of manpower. Without the limitation in phenotype analysis, the result acquired using this model can be used for genome-wide association study or quantitative trait locus analysis, which identifies key genes controlling drought tolerance and requires a large dataset of palms to be used. Also, high-throughput phenotyping methods will accelerate genomic selection<sup>5,6</sup>, thereby reducing time required for breeding and selection. As an initial step towards automating stomata phenotyping in oil palm, future models will need to not only detect stomata, but also segment out and measure their sizes and openings. The model can then be incorporated into a portable device for field phenotyping. Given the high accuracy observed, other CNN models can be developed in the future for many other commercially useful traits and for precision agriculture, for example automatic palm detection from drone/satellite images and disease classification in estates or nurseries.

## Conclusion

This paper describes the development of the first oil palm CNN-based stomata detection model using leaves collected from palms of different ages. The resulting model showed very high accuracy (99.08% precision and 98.92% recall) when tested on three different age groups. The model also demonstrated high transferability, achieving high accuracy (99.72% precision and 96.88% recall) when cross-applied onto micrographs acquired under different conditions.

Received: 2 March 2020; Accepted: 9 July 2021

Published online: 26 July 2021

## References

1. Meuwissen, T. H., Hayes, B. J. & Goddard, M. E. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**, 1819–1829 (2001).
2. Hayes, B. J., Bowman, P. J., Chamberlain, A. J. & Goddard, M. E. Invited review: Genomic selection in dairy cattle: Progress and challenges. *J. Dairy Sci.* **92**, 433–443 (2009).
3. Bruce, W. B., Edmeades, G. O. & Barker, T. C. Molecular and physiological approaches to maize improvement for drought tolerance. *J. Exp. Bot.* **53**, 13–25 (2002).

4. Wong, C. K. & Bernardo, R. Genomewide selection in oil palm: Increasing selection gain per unit time and cost with small populations. *Theor. Appl. Genet.* **116**, 815–824 (2008).
5. Kwong, Q. B. *et al.* Genomic selection in commercial perennial crops: Applicability and improvement in oil palm (*Elaeis guineensis* Jacq.). *Sci. Rep.* **7**, 2872 (2017).
6. Cros, D. *et al.* Genomic selection prediction accuracy in a perennial crop: Case study of oil palm (*Elaeis guineensis* Jacq.). *Theor. Appl. Genet.* **128**, 397–410 (2015).
7. Appleby, N., Edwards, D. & Batley, J. New technologies for ultra-high throughput genotyping in plants. *Methods Mol. Biol.* **513**, 19–39 (2009).
8. Dearlove, A. M. High throughput genotyping technologies. *Brief. Funct. Genomic Proteomic* **1**, 139–150 (2002).
9. Furbank, R. T. & Tester, M. Phenomics—Technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* **16**, 635–644 (2011).
10. Grosskinsky, D. K. *et al.* Phenotyping in the fields: Dissecting the genetics of quantitative traits and digital farming. *New Phytol.* **207**, 950–952 (2015).
11. Ubbens, J. R. & Stavness, I. Deep plant phenomics: A deep learning platform for complex plant phenotyping tasks. *Front. Plant Sci.* **8**, 1190 (2017).
12. Zu, X. F. *et al.* A new method for evaluating the drought tolerance of upland rice cultivars. *Crop J.* **5**, 488–498 (2017).
13. Tuberosa, R. Phenotyping for drought tolerance of crops in the genomics era. *Front. Physiol.* **3**, 347 (2012).
14. Kim, T. H., Bohmer, M., Hu, H., Nishimura, N. & Schroeder, J. I. Guard cell signal transduction network: Advances in understanding abscisic acid, CO<sub>2</sub>, and Ca<sup>2+</sup> signaling. *Annu. Rev. Plant Biol.* **61**, 561–591 (2010).
15. Tanaka, Y., Sugano, S. S., Shimada, T. & Hara-Nishimura, I. Enhancement of leaf photosynthetic capacity through increased stomatal density in *Arabidopsis*. *New Phytol.* **198**, 757–764 (2013).
16. Lawson, T. & Blatt, M. R. Stomatal size, speed, and responsiveness impact on photosynthesis and water use efficiency. *Plant Physiol.* **164**, 1556–1570 (2014).
17. Hughes, J. *et al.* Reducing stomatal density in barley improves drought tolerance without impacting on yield. *Plant Physiol.* **174**, 776–787 (2017).
18. Caine, R. S. *et al.* Rice with reduced stomatal density conserves water and has improved drought tolerance under future climate conditions. *New Phytol.* **221**, 371–384 (2019).
19. Dunn, J. *et al.* Reduced stomatal density in bread wheat leads to increased water-use efficiency. *J. Exp. Bot.* **70**, 4737–4748 (2019).
20. Hepworth, C., Doheny-Adams, T., Hunt, L., Cameron, D. D. & Gray, J. E. Manipulating stomatal density enhances drought tolerance without deleterious effect on nutrient uptake. *New Phytol.* **208**, 336–341 (2015).
21. Xu, Z. & Zhou, G. Responses of leaf stomatal density to water status and its relationship with photosynthesis in a grass. *J. Exp. Bot.* **59**, 3317–3325 (2008).
22. Voleníková, M. & Tichá, I. Insertion profiles in stomatal density and sizes in *Nicotiana tabacum* L. plantlets. *Biol. Plant.* **44**, 161–165 (2001).
23. Tzutalin. LabelImg. Git code. <https://github.com/tzutalin/labelImg> (2015). Accessed 1 July 2018.
24. Bradski, G. The OpenCV Library. *Dr. Dobbs' Journal of Software Tools.* **120**, 122–125 (2000).
25. Abadi, M. *et al.* TensorFlow: Large-scale machine learning on heterogeneous distributed systems. [arXiv:1603.04467v2](https://arxiv.org/abs/1603.04467v2) (2016).
26. Huang, J. *et al.* Speed/accuracy trade-offs for modern convolutional object detectors. [arXiv:1611.10012](https://arxiv.org/abs/1611.10012) (2016).
27. Howard, A. G. *et al.* MobileNets: Efficient convolutional neural networks for mobile vision applications. [arXiv:1704.04861](https://arxiv.org/abs/1704.04861) (2017).
28. Liu, W. *et al.* SSD: Single shot MultiBox detector. In *Computer Vision—ECCV 2016* Vol. 9905 (eds Leibe, B. *et al.*) 21–37 (Springer, 2016).
29. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. [arXiv:1409.1556](https://arxiv.org/abs/1409.1556) (2015).
30. Lin, T. Y. *et al.* Microsoft COCO: Common objects in context. In *Computer Vision—ECCV 2014* (eds Fleet, D. *et al.*) 740–755 (Springer, 2014).
31. Girshick, R. Fast R-CNN. [arXiv:1504.08083](https://arxiv.org/abs/1504.08083) (2015).
32. Jayakody, H., Liu, S., Whitty, M. & Petrie, P. Microscope image based fully automated stomata detection and pore measurement method for grapevines. *Plant Methods* **13**, 94 (2017).
33. Violet-Chabrand, S. & Brendel, O. Automatic measurement of stomatal density from microphotographs. *Trees* **28**, 1859–1865 (2014).
34. Aono, A. H. *et al.* A stomata classification and detection system in microscope images of maize cultivars. *Comput. Electron. Agric.* **9**, 467 (2019).
35. Fetter, K. C., Eberhardt, S., Barclay, R. S., Wing, S. & Keller, S. R. StomataCounter: A neural network for automatic stomata identification and counting. *New Phytol.* **223**, 1671–1681 (2019). [arXiv:1409.1556](https://arxiv.org/abs/1409.1556) (2015).
36. Zhao, Z. Q., Zheng, P., Xu, S. T. & Wu, X. Object detection with deep learning: A review. *IEEE Trans. Neural Netw. Learn. Syst.* **30**, 3212–3232 (2019).
37. Xu, X. Y., Li, G. Q., Xie, G., Ren, J. C. & Xie, X. L. Weakly supervised deep semantic segmentation using CNN and ELM with semantic candidate regions. *Complexity* **2019**, 12 (2019).
38. Bhagat, P. K. & Choudhary, P. Image annotation: Then and now. *Image Vis. Comput.* **80**, 1–23 (2018).

## Acknowledgements

We will like to thank Sharon Khor Mi Bee, Ong Ai Ling, Heng Huey Ying, Song Cheen and Kee Shao Yong for their assistance in annotating stomata objects in the micrographs. Also, we will like to thank Siti Khadijah Mohd Rais, Mohd Nor Azizi Shabudin, Leona Daniela Jeffery Daim and Neoh Bee Keat for samples collection and preparation of stomata imprint in this study.

## Author contributions

Conceptualization, Q.B.K., Y.C.W. and D.R.A.; Methodology, Q.B.K., Y.C.W. and Y.T.K.; Software, Q.B.K. and M.S.S.; Formal Analysis, Q.B.K., Y.C.W., P.L.L. and D.R.A.; Investigation, Q.B.K., Y.C.W., P.L.L. and M.S.S.; Resources, Y.C.W., D.R.A. and H.K.; Data Curation, Q.B.K., Y.T.K. and M.S.S.; Writing—Original Draft, Q.B.K., Y.C.W. and H.K.; Writing—Review & Editing, D.R.A. and Y.T.K.; Visualization, Q.B.K., P.L.L. and M.S.S.; Supervision, H.K. and D.R.A.; Project Administration, Q.B.K. and Y.C.W.; Funding Acquisition, D.R.A. and H.K.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-94705-4>.



**Correspondence** and requests for materials should be addressed to Q.B.K.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021