New Phytologist

Methods

StomaAI: an efficient and user-friendly tool for measurement of stomatal pores and density using deep computer vision

Na Sai[1,2]*, James Paul Bockman[3,4]*, Hao Chen[3,4], Nathan Watson-Haigh[5,6], Bo Xu[1,2], Xueying Feng[1,2], Adriane Piechatzek[1,2], Chunhua Shen[3,4,7] and Matthew Gilliham[1,2]

[1]Plant Transport and Signalling Lab, ARC Centre of Excellence in Plant Energy Biology, Waite Research Institute, Glen Osmond, SA 5064, Australia; [2]School of Agriculture, Food and Wine, University of Adelaide, Adelaide, SA 5064, Australia; [3]The Australian Institute for Machine Learning, Adelaide, SA 5005, Australia; [4]School of Computer Science, University of Adelaide, Adelaide, SA 5005, Australia; [5]South Australian Genomics Centre, SAHMRI, Adelaide, SA 5000, Australia; [6]Australian Genome Research Facility, Victorian Comprehensive Cancer Centre, Melbourne, Vic. 3000, Australia; [7]Present address: Zhejiang University, Zhejiang 310058, China

Authors for correspondence:
Matthew Gilliham
Email: matthew.gilliham@adelaide.edu.au

Chunhua Shen
Email: chunhua@me.com

Received: *31 October 2022*
Accepted: *23 December 2022*

New Phytologist (2023) **238:** 904–915
doi: 10.1111/nph.18765

Key words: applied deep learning, computer vision, convolutional neural network, phenotyping, stomata.

Summary

- Using microscopy to investigate stomatal behaviour is common in plant physiology research. Manual inspection and measurement of stomatal pore features is low throughput, relies upon expert knowledge to record stomatal features accurately, requires significant researcher time and investment, and can represent a significant bottleneck to research pipelines.
- To alleviate this, we introduce StomaAI (SAI): a reliable, user-friendly and adaptable tool for stomatal pore and density measurements via the application of deep computer vision, which has been initially calibrated and deployed for the model plant Arabidopsis (dicot) and the crop plant barley (monocot grass).
- SAI is capable of producing measurements consistent with human experts and successfully reproduced conclusions of published datasets.
- SAI boosts the number of images that can be evaluated in a fraction of the time, so can obtain a more accurate representation of stomatal traits than is routine through manual measurement. An online demonstration of SAI is hosted at https://sai.aiml.team, and the full local application is publicly available for free on GitHub through https://github.com/xdynames/sai-app.

Introduction

Stomata, derived from the Greek word for *mouth*, are small pores in the epidermal surface of plant aerial organs (leaves and stems). Dicot plants most commonly have a pair of kidney-shape guard cells surrounding each stomatal pore (Fig. 1). In monocot grasses however, such as barley or maize, the stomatal pore is surrounded by dumbbell-shaped guard cells, which are flanked in turn by a pair of subsidiary cells (Cai *et al.*, 2017; Gray *et al.*, 2020). Stomatal pores play a critical role in plant physiology by limiting the diffusion of carbon dioxide (CO_2) into leaves, impacting the rate of photosynthesis – the process that releases oxygen as a by-product and produces the carbohydrates that fuel plant metabolic functions, growth and development. At the same time, water vapour released via stomatal pores enables water transport through plants (Rizhsky *et al.*, 2004; Sinha, 2004). Some plants survive exposure to excessive heat by keeping stomata open and cooling leaves through the evaporation of water. Conversely, stomata close during drought to prevent water loss (Rizhsky *et al.*, 2004). Stomata also respond to stimuli over diel cycles, such as light and dark, and a multitude of other signals to optimise CO_2 gain and water loss (Roelfsema & Hedrich, 2005; Shimazaki *et al.*, 2007). As a consequence, stomatal aperture regulation during daily light and dark cycles, or in response to environmental stresses, directly impacts plant growth, development and survival (Hetherington & Woodward, 2003; Shimazaki *et al.*, 2007; McLachlan *et al.*, 2014).

Due to the importance of stomata, investigating stomatal regulation has become a common task for biologists studying plant signalling pathways and stress perception (Shimazaki *et al.*, 2007; Kim *et al.*, 2010; Ye & Murata, 2016). To study stomatal traits (i.e. size or density) and behaviour, researchers commonly use microscopy (Chater *et al.*, 2015; Eisele *et al.*, 2016; Xu *et al.*, 2021), but this method is not straightforward. Morphological differences in stomata of different species (Fig. 1) and variable image quality make accurate stomatal density and pore measurement a task that requires experience and training. Traditionally, stomatal pore measurement requires manual inspection

*These authors contributed equally to this work.

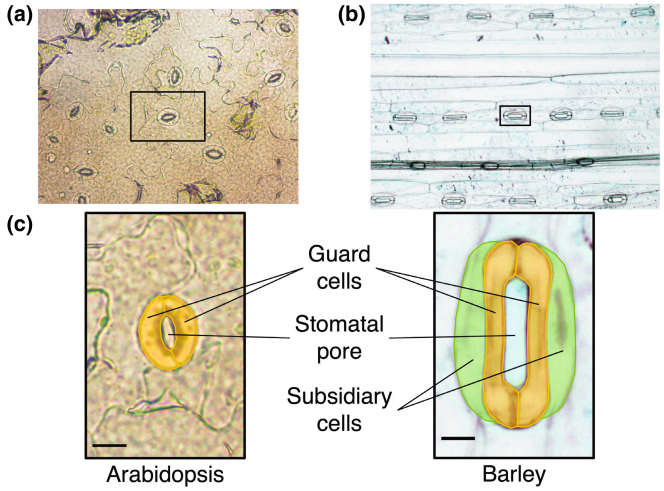

Fig. 1 Stomata of *Arabidopsis thaliana* and barley (*Hordeum vulgare*). Representative captured Arabidopsis (a) and barley (b) epidermis used as inputs to StomaAI. Arabidopsis images were captured at 2592 × 1944 pixel resolution with 10 × 20 magnification, and barley images were captured at 2880 × 2048 pixel resolution with 10 × 10 magnification. Components of Arabidopsis and barley stomata (c) that are rotated and magnified from original image (a) and (b) (indicated by black bounding box) are highlighted in pseudo colour and labelled; bars, 10 μm.

to identify and measure relevant features (i.e. stomatal counts, pore area and aperture). Hundreds of stomatal images need to be analysed to gain sufficient statistical power to support a biological conclusion, a time-consuming and laborious task. Although manual counting and measurement can be aided by image processing software such as FIJI-IMAGEJ (Schindelin *et al.*, 2012), manually tuned parameters are required to produce acceptable performance (Cheng *et al.*, 2014). Thus, an automated stomatal counting and pore measurement system is highly desirable and, with recent advances in computer vision, such a system has become possible to build.

Microscopy-aided imaging presents a uniquely controlled environment in which to apply modern computer vision techniques. Images can be captured in high resolution via calibrated optics, reducing systematic noise. Plant anatomy also enforces regularity in pattern, appearance and orientation (in monocot grasses); these factors remove several of the common *Achilles' heels* of applied vision systems. Previous attempts have been made to quantify stomatal density, pore width and area using traditional computer vision techniques (Omasa & Onoe, 1984; Laga *et al.*, 2014; Vialet-Chabrand & Brendel, 2014; Duarte *et al.*, 2017; Bourdais *et al.*, 2019; Sakoda *et al.*, 2019). Although these methods demonstrate efficacy on their respective tasks, they rely on handcrafted and/or multistage processes. The use of Convolutional Neural Networks (CNNs) to detect stomatal attributes has recently increased in popularity (Jayakody *et al.*, 2017; Saponaro *et al.*, 2017; Toda *et al.*, 2018; Bhugra *et al.*, 2018, 2019; Falk *et al.*, 2019; Fetter *et al.*, 2019; Li *et al.*, 2019; Meeus *et al.*, 2020; Gibbs *et al.*, 2021). Convolutional Neural Networks enable a series of pertinent operations to be learnt from examples – acting as a data-driven approximation of a sequence of computer vision operations.

More recently, Mask Regions with Convolutional Neural Network features (Mask R-CNN) has been used to perform identification and localisation of stomata. This method involves the entire stomatal complex being detected, encircled by a polygon with its orientation and stomatal complex area captured, inferring axis length (Jayakody *et al.*, 2021; Xie *et al.*, 2021) or stomatal density (Bheemanahalli *et al.*, 2021). The algorithms were successfully used across different species with varying image quality (Jayakody *et al.*, 2021). Many studies only estimate stomatal counts for density calculations (Vialet-Chabrand & Brendel, 2014; Duarte *et al.*, 2017; Falk *et al.*, 2019; Fetter *et al.*, 2019; Sakoda *et al.*, 2019; Meeus *et al.*, 2020; Bheemanahalli *et al.*, 2021). Of these methods, Fetter *et al.* (2019) is the only study to provide a user-friendly application for researchers, named 'Stomata Counter'.

Fewer studies focus on stomatal pore measurements, with semi-automated methods that require handcrafted feature extractors or manual post-processing following model inference (Laga *et al.*, 2014; Jayakody *et al.*, 2017; Bhugra *et al.*, 2018, 2019; Li *et al.*, 2019). Ellipse fitting is the common solution used among these studies, for which the fitted ellipse's area, minor axis and major axis are used to estimate pore area, width and length, respectively (Jayakody *et al.*, 2017; Bhugra *et al.*, 2018, 2019; Li *et al.*, 2019; Liang *et al.*, 2021). However, the fitting method is restricted to stomata with an ovular-shaped pore (e.g. Arabidopsis stomata), so other pore shapes (e.g. barley stomata) cannot be represented correctly, resulting in under- or overestimation of pore features (Fig. 1c). None of these studies offers a user-friendly automated stomatal pore measurement tool.

Here, we present StomaAI (SAI) as an accessible and automated tool that allows stomatal counting (for stomatal density estimation) and pore measurement of microscope images. Precise pore measurement is the core novelty of SAI, which is capable of measuring pore area, length, width (i.e. aperture) and width/length ratio. We demonstrate that measurements obtained using SAI are comparable to those taken by human experts, providing assurance of prediction reliability. This key comparison has not been provided by contemporary studies that normally use traditional computer vision evaluation criteria such as F1 score or average precision (AP) to evaluate machine performance. Due to differences in stomatal morphology, SAI includes two class-specific models: a dicot model trained with Arabidopsis data, and a monocot cereal model trained with barley data. We demonstrate that with *c.* 150 annotated images containing *c.* 1700 stomata, SAI can be trained to count and measure pores of different plant species. An online demonstration of our software, where model inference can be viewed, is hosted at https://sai.aiml.team. To measure user-acquired samples, we provide a local version of SAI that can be accessed via https://github.com/xdynames/sai-app. This article contains our experimental materials and the information required to use SAI in the Methods; information to implement and retrain the tool in the Supporting Information; performance of the tool in the Results (and Supporting Information); and the significance of SAI in the discussion.

New Phytologist

Materials and Methods

Plant material, data annotation and modelling

Arabidopsis thaliana ecotype Col-0 and barley (*Hordeum vulgare*, cv Barke) fresh epidermal peels were used as plant materials for image capture. Arabidopsis images were captured using Axiophot Pol Photomicroscope (Carl Zeiss) with 10 × 20 magnification at 2592 × 1944 resolution. A Nikon DS-Fi3 digital camera with a Nikon Diaphot 200 inverted microscope was used to capture barley images with 10 × 10 magnification at 2880 × 2048 resolution.

As mentioned above, researchers currently inspect each captured microscope image to count and/or measure relevant structures manually. Pore measurement depends upon the pore's opening status. The area of an open stoma can be measured by drawing a polygon that encloses its mouth. Pore width and length can either be directly measured or estimated by applying fitting methods to the aforementioned polygon. Lengths of closed pores are acquired by selecting points that mark the beginning and end of their tightly shut mouth.

For SAI, we re-created this process with pore feature annotation using RectLabel (v.3.03.8, https://rectlabel.com). In a given annotation, two labels were ascribed to each stoma – a bounding box delineating a single stoma and its status (open or closed; Fig. 2), and stomatal pore features were recorded as a polygon for open stomata or as a line for closed stomata. In an open stoma, a series of points were labelled such that they form a polygon. This polygon enclosed the area formed between guard cells as a result of the pore opening. Stomatal pore area was derived from the polygon area. Two of the points were positioned where the two guard cells intersect at each end of the pore (where the polygon was longest). These points were extracted and used as the open pore's length/keypoint annotation. However, when a stoma was closed, the pore's length and associated keypoints were derived from the beginning and the end of the line formed by the guard cells pushing against one another (see Fig. 2, indicated by arrows). All information was organised for compatibility with Microsoft Common Objects in Context (MS-COCO), a widely used computer vision benchmark (Lin *et al.*, 2014). Summary statistics of the created database of microscopy images are presented in Table 1.

When a model emulates a researcher performing these measurements, it must locate target structures within a sample, comment on their state and gather relevant measurements. We have reformulated each one of these tasks into a computer vision task. A researcher's initial identification (counting) of stomata and their opening status is re-framed as *object detection*, which consists of drawing boxes around salient objects and predicting that object's semantics. Drawing polygons indicating stomatal pore opening becomes *segmentation*, which highlights regions of interest within images. Selecting end points for a stomatal pore is analogous to *keypoint detection*, which reduces visual features of interest to a defining pixel. Each of these tasks has a library of possible models capable of solving them individually, but candidate models that can solve object detection, segmentation and keypoint detection in tandem are much rarer. The Mask R-CNN is one such model and represents an incremental change atop an already established series of deep-learning architectures (Girshick *et al.*, 2014; Girshick, 2015; He *et al.*, 2017; Ren *et al.*, 2017). Through the use of specialised predictive heads, this iteration comes armed with the requisite predictive powers for our physiological needs.

Deep-learning models were built using DETECTRON 2 (Wu *et al.*, 2019), an open-source framework sitting on top of PYTORCH (Paszke *et al.*, 2019). Both packages were created by Facebook's Artificial Intelligence Research division (FAIR). Adaptions were made to FAIR's Mask R-CNN model to better suit stomatal pore measurement, specifically, increasing the resolution of prediction heads responsible for segmentation and keypoint detection. Mean average precision (mAP), as defined in the MS-COCO Challenge 2017, was used to evaluate and compare models on all tasks (Lin *et al.*, 2014). Justification and verification of model design choices and training regimes are presented in Supporting Information Methods S1–S8. Pseudocode for the tool's core measurement and processing loop is provided in

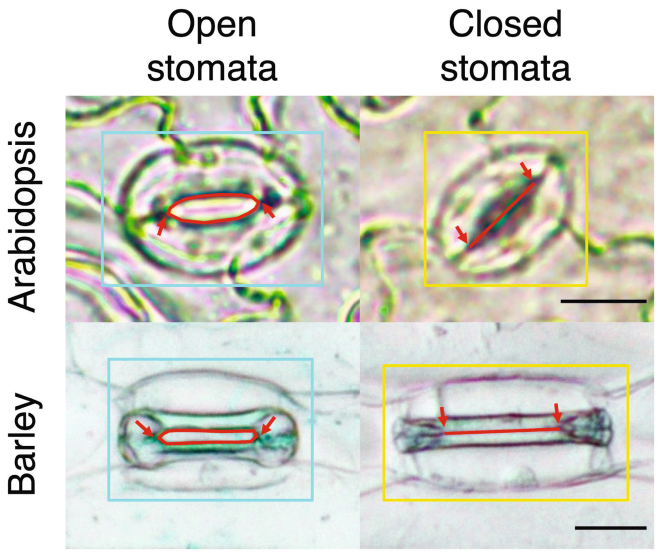

Fig. 2 Annotation examples of Arabidopsis and barley stomata. Bounding boxes contain a single Arabidopsis or barley stoma (i.e. a pair of guard cells and a pair of subsidiary cells from barley) and its opening status is determined in different labels (open stoma in cyan, closed stoma in yellow). The red polygon and line define the stomatal pore. Red arrows indicated the position of points that were extracted from the polygon or line as stomatal pore length annotation. Bars: 10 μm in Arabidopsis; 20 μm in barley.

Table 1 Summary metrics for stomatal pore dataset used in model training and evaluation.

	Images		Stoma instance (Open/Closed)	
	Train	Validation	Train	Validation
Arabidopsis	200	42	974/293	235/55
Barley	150	33	1000/692	268/89

Method S9. Code used to train models, associated model weights and relevant image data can be found at GitFront: https://gitfront.io/r/jpb/u6BtFFMkNGCv/SAI-training/. The local version of SAI is available at https://github.com/xdynames/sai-app. Detailed user installation instructions and user guide are available in the Notes S1 and S2.

Average-Human/Machine Test

To determine whether SAI predictions were consistent with human measurement, an Average-Human/Machine Test was designed. In total, 35 microscopy images (15 of barley and 20 of Arabidopsis) were collated as a test dataset (details summarised in Table 2). Four plant stomatal morphology experts manually measured stomata (Arabidopsis: $n = 149$, barley: $n = 175$) in the 35 images. For stomatal counts, the *human-level* reference was the number that the majority of experts identified; where there was no majority count, a mean was used. For other measurement attributes, mean values derived from the four human experts were used as the *human-level* reference. Participants used the data annotation schema as described above.

To compare SAI and human expert measurements, stomatal counts and density were matched on a per-image basis, while stomatal pore width, length, area and width/length ratio were matched per stoma. Differences between SAI/human experts and *human-level* reference measurements were visualised with scatter plots and quantified by relative error (RE). Error in stomatal counting was evaluated using the difference between SAI's and the *human-level* reference's count. To evaluate the agreement between SAI, human experts and the *human-level* reference, the concordance correlation coefficient (CCC; ranging from −1 to 1) was applied (Lin, 1989). Furthermore, for each measurer, the total number of stomata counted and the means of the other measurement attributes were compared using one-way ANOVA with a Tukey's Honest Significance Difference (HSD) test.

SAI in practice

To justify whether SAI can be used as a replacement for traditional stomatal pore measurement methods, we compared a previously published conventionally measured dataset with SAI measurements to ascertain whether the same scientific conclusions using both measurement techniques were aligned. Manually measured image datasets of both Arabidopsis and barley were obtained from Xu *et al.* (2021). Two different experimental designs were selected to evaluate SAI's real-world performance: the 25 μM abscisic acid (ABA) ± 2 mM γ-aminobutyric acid (GABA) for Arabidopsis; and dark-to-light transition ±1 mM GABA for barley (figs 3b, 5g from Xu *et al.*, 2021). Reference measurements for the two datasets were made by different researchers: Arabidopsis measurements were taken by human expert 2, and barley samples were measured by human expert 4. Measurements produced by SAI were subjected to the same statistical tests used in Xu *et al.* (2021) to examine whether SAI provides results consistent with human experts.

Table 2 Summary of Average-Human/Machine Test dataset.

		Stomata	
	Images	Open	Closed
Arabidopsis	20	129	20
Barley	15	109	66

Inference time assay

The efficiency of SAI was tested on a range of commonly available computer hardware using the same set of 35 sample images used in the above Average-Human/Machine Test (details summarised in Table 2). A series of processors were tested on the same image set of Arabidopsis and barley at the native resolution of images for each species. For all tests, no user-defined filter was applied. Time to process each image was recorded and used to estimate the average model inference time with post-processing with sample standard deviation for each processor (Arabidopsis: $n = 20$, barley: $n = 15$). These measures were then used to compare throughputs.

Overview of SAI workflow

The inference and post-processing pipeline used in SAI is simplified and illustrated in Fig. 3. Images of stomata from fresh epidermal peels captured using a bright field optical microscope are required for SAI. A 10 × 20 magnification at 2592 × 1944 resolution for Arabidopsis images and 10 × 10 magnification at 2880 × 2048 resolution for barley are recommended for best results. Captured microscope images may require format conversion if images are not in a supported format (png, bmp, jpg, jpe, jp2 and tiff).

At its core, SAI uses the deep neural network Mask-R-CNN to generate proposals for pore measurements in two main stages. Initially, an image is parsed by the region proposal network and smaller image areas are identified as regions of interest. In our case, regions being selected should contain a single stoma. Once the region proposals are generated, each one is passed to four different prediction heads, each responsible for a different type of measurement. One head produces a bounding box that further localises the stomatal complex; another predicts whether the stomatal pore is open or closed; a third attempts to mask the open region of the pore; and the final head estimates the location of two keypoints denoting the beginning and end of the stomatal pore.

Many of these predictions can be duplicated, incomplete or contain errors and thus require treatment. SAI has a series of post-processing steps to increase measurement validity. Removing duplicate detections is done using nonmaximum suppression. This technique finds predicted bounding boxes with significant overlap – an intersection over union (IoU) of > 0.2 – and discards the measurements for the complex with the lowest confidence score. The confidence score is an arbitrary scalar between 0 and 1 output by the network, which indicates how strong the

New Phytologist

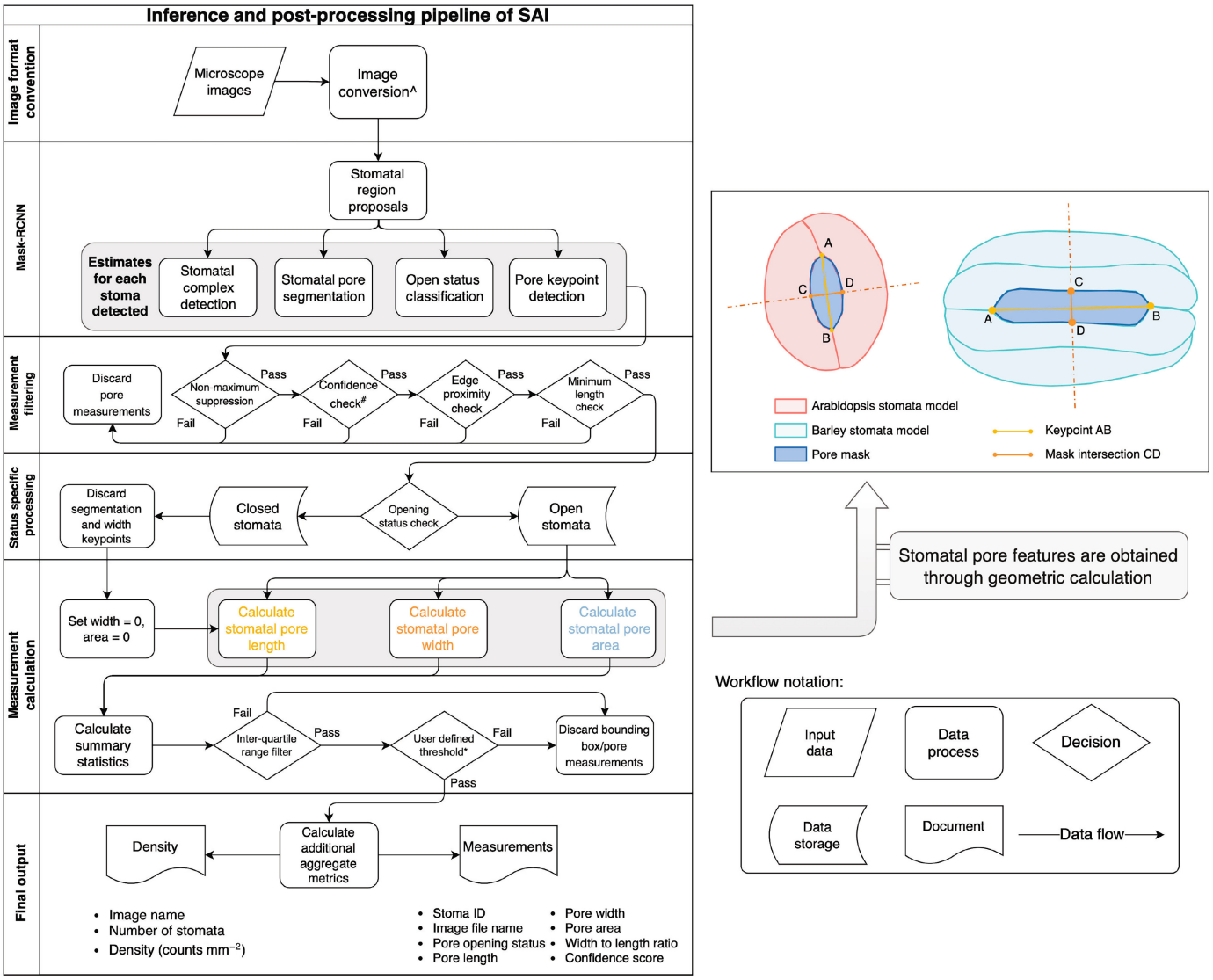

Fig. 3 The inference and post-processing pipeline of StomaAI (SAI). The image-to-measurement process is divided into several stages. At the measurement calculation stage, stomatal pore area was calculated by the number of pixels within pore mask (blue); length was calculated through the geometric distance between predicted keypoints (AB), and width was obtained by finding the intersection points of pore mask and mid-perpendicular line of AB (i.e. points C and D). Arrow indicate the direction of data flow. ^Supported image format: 'png', 'bmp', 'jpeg', 'jpg, 'jpe', 'jp2', 'tiff', 'tif'. #Internal confidence threshold = 0.6. *User defined threshold including immature stomata filtering (via stomate pore length) and confidence threshold.

response of its classification head was to the region. SAI internally filters all measurements from detected complexes with a confidence score < 0.6 as this allows the majority of false positives to be removed while retaining valid detections for counting and pore measurement. Occasionally, images will contain stomata that are cut off by the image's framing; detections found close to the edge have their associated pore measurements discarded, but those whose bounding box is > 50% of the average size for that image are retained for counting. Any stomata whose pore length measurement is extremely small (< 5 pixels) will also have their measurements discarded, but the bounding box retained for counting.

After this initial filtering, additional measurements are computed for each stoma. All detections will have their length estimated from keypoint detections. Depending on their opening status, detections are treated differently. If a stoma is classified as closed, the pore's width and area measurements are set to zero. Open stomata have their pore areas and widths calculated from their mask and keypoint detections, respectively. Stomatal pore width calculation is done by finding the intersection points of the stomatal pore mask and the mid-line normal to the length keypoints (Fig. 3, right upper panel). Collectively, measurements from all stomata across all images are used to inform anomaly removal via interquartile range (IQR) filtering. All samples whose pore length is < 2 IQRs below the median are excluded, as are samples whose bounding box width or height are 2 IQRs above or below the median value (bounding box IQR calculation excludes stomata on image edges). Samples that pass all these criteria are then subjected to two user-defined filters, where users can adjust the confidence score threshold to > 0.6 and/or exclude

samples below a minimum length to discard immature stomata. The minimum length filter is not applied to the results of the counting task.

After user-defined filtering, the remaining samples have their width/length ratios and stomatal density calculated, and output as two comma-separated values (CSV) files for further analysis. For pore measurements, each stoma has a row that includes an identifier for the pore (*id*), the name of the image it was extracted from (*image name*), whether the pore is open or closed (*class*), its measured pore *length*, *width*, *area*, *width*/*length* ratio and the stomatal pore's associated confidence score (*confidence*). For counting, each image has a separate row that contains the image name (*image name*), *number of stomata* and estimated *density*. This information can be saved by the users in an image format visualising the masks and measurements overlaid onto the original images to allow users to verify SAI's predictions. Bounding boxes indicate counted stomata, while polygon and lines mask over the stomatal pore indicate pore measurements. Example of output CSV and measurement visualisation are illustrated in the Fig. S1.

Results

Average-Human/Machine Test

Beyond assessing performance using traditional metrics (details in Methods S2), we show that SAI produces measurements equivalent to *human-level* performance. To compare independent human operators (multiple plant physiology researchers) with SAI, we applied an Average-Human/Machine Test. Stomatal counts, density, pore width, length, area and width/length ratio were measured by SAI (Figs 4, S2) and human experts (Fig. S3). Stomatal counts and pore width measurements obtained from SAI, when plotted against *human-level* reference, generally align with $y = x$; indicating that number and pore width measurements are consistent with the *human-level* reference (Fig. 4). Stomata incorrectly identified as open or closed can be identified along the *x*- or *y*-axes (Fig. 4b).

Overall, SAI achieved a CCC in Arabidopsis and barley of 0.969 and 0.992 in stomata counting (Fig. 4a), and 0.891 and 0.984 for stomatal pore width (Fig. 4b), respectively. Concordance correlation coefficients in stomatal density (Fig. S2a), and pore length, area and width/length ratio (Fig. S2b) were 0.971, 0.909, 0.928 and 0.865 for Arabidopsis, and 0.978, 0.965, 0.985 and 0.985 for barley, respectively. Considering that any open Arabidopsis stomata with a pore width of < 1 μm will have a minimal impact on transpiration, excluding them from the *human-level* reference increases the SAI CCC to 0.916. Human experts show average CCCs of 0.949 and 0.985 in Arabidopsis and barley stomata counting (Fig. S3a), and 0.945 and 0.985 in Arabidopsis and barley stomatal pore width (Fig. S3c), respectively. Relative error in stomatal density and pore measurements from SAI were distributed in a similar pattern to those from human measurements (Figs 4b, S3b–f). Judging stomatal pore width and area for near-closed stomata is more difficult than for open stomata, which creates a skew in the RE histogram where errors are more frequently observed in small measurements. Estimation of stomatal pore length was not affected by stomatal opening status, so REs were spread evenly between under- and overestimation.

The total number of counted stomata, means of stomatal density and pore measurements were calculated and compared from each measurement source (SAI and human experts; Fig. 5; Tables S1, S2). No significant differences were found for stomatal density (Fig. 5a) or stomatal pore features (Fig. 5b) across either Arabidopsis or barley samples when SAI was compared with the *human-level* reference. Additionally, SAI exhibits no significant difference from individual human experts, except in the case of expert 2 for Arabidopsis length measurements (Fig. 5b). In fact, human expert 2 provided the most conservative results for Arabidopsis, with significantly smaller measurements for pore area, length or width than human experts 1 and 3, and pore length measurements than SAI.

Application on real-world physiological experiments

SAI was used to measure two sets of published physiological experiments. The original images used for Xu *et al.* (2021) were processed with SAI. Traditionally, researchers will exercise their discretion by consciously measuring only stomata they deem mature. SAI measures indiscriminately. However, we were able to emulate this human practice via filtering detections by excluding stomata with pore length < 2 μm in Arabidopsis and < 16 μm in barley. SAI and the original manual measurements from Xu *et al.* (2021) were compared using one-way ANOVA with Tukey's HSD (Fig. 6). Scientific conclusions drawn from the statistical tests were consistent between SAI and the original researchers: Arabidopsis stomata close in response to 25 μM ABA ±2 mM GABA, and light-induced barley stomatal opening was inhibited by the presence of 1 mM GABA.

The mean and distribution of stomatal pore widths obtained by SAI in each treatment group were compared with the original published measurements. SAI detected 66.4% and 91.4% of the manually measured data reported by Xu *et al.* (2021) in Arabidopsis and barley, respectively. The better detection performance on barley stomata is likely due to the regular stomatal orientation and distribution on a barley leaf, and the higher image quality of barley data (see Fig. 1b). For both barley and Arabidopsis, distributions of stomatal pore width measurements produced by SAI were similar to those produced via manual inspection (Fig. S4). Human expert 2, who was identified as the most conservative measurer in our Average-Human/Machine Test, was responsible for measuring Arabidopsis stomata in the Xu *et al.* (2021) report, so the lower mean value for published Arabidopsis stomatal pore widths was expected (Figs 5, 6).

Inference time assay

Fig. 7 shows the average time required to process a microscope image (module inference time and post-processing time) on a range of commonly available hardware. Model inference time is predominantly limited by image resolution and computation speed. Due to this, barley data (2880 × 2048 resolution)

New Phytologist

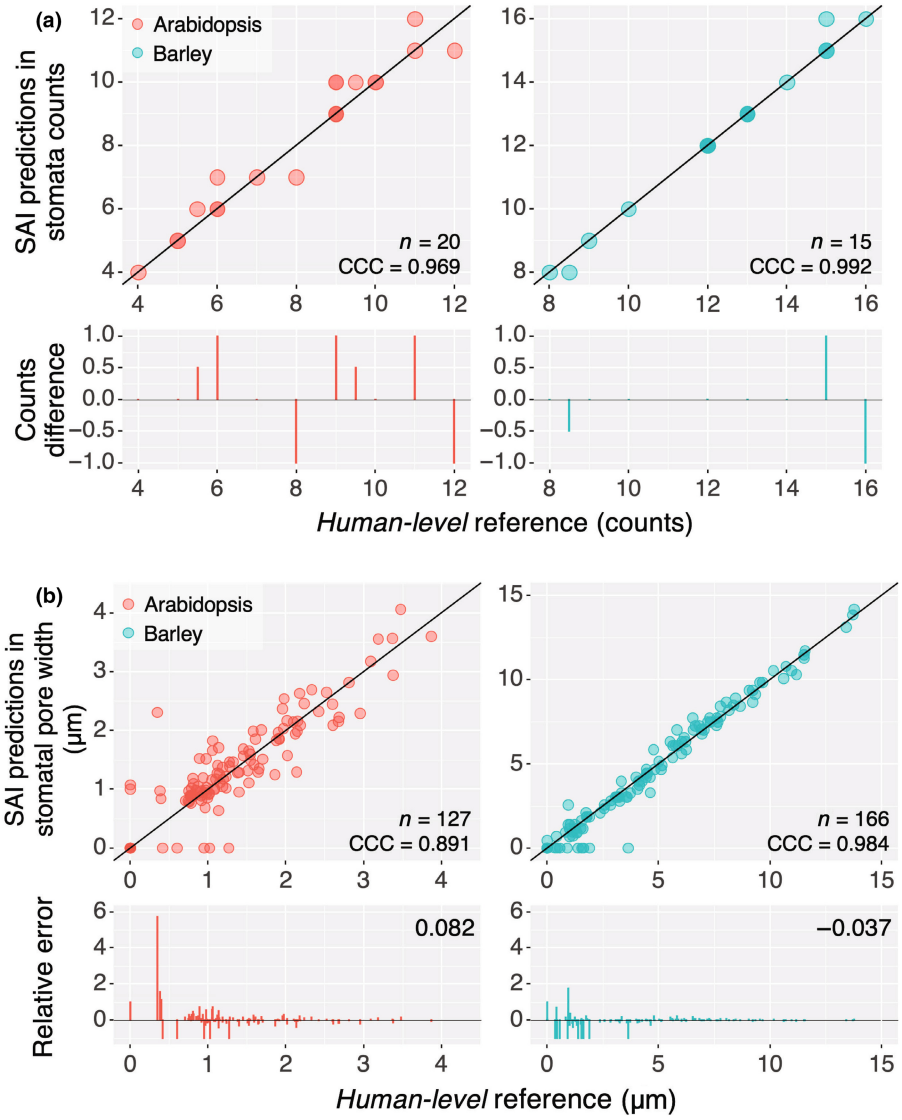

Fig. 4 StomaAI (SAI) prediction vs average *human-level* reference set in Arabidopsis and barley stomatal counting (a) and pore width (b). Stomatal counts and pore measurements from four human experts were collected and converted to a *human-level* reference. In the upper panel, SAI predictions were compared against the reference and the concordance correlation coefficient (CCC, ranging from −1 to 1) was used to determine accuracy performance. The black diagonal line represents $y = x$ and CCC is a measure of dispersion for the points from this line. The corresponding count difference and relative error to *human-level* reference is presented in the lower panel.

generally took longer to process than Arabidopsis data (2592 × 1944 resolution) on the same processor.

Discussion

SAI provides a new way to analyse stomata, one of the most studied plant cell types. Our tool allows stomatal counting and measurement of a series of stomatal pore features including stomatal density and pore opening status, width, length and area to be extracted from one set of images simultaneously. SAI generally produced better predictions on barley than Arabidopsis reflected in the CCC values obtained (Figs 4, S2), largely due to image quality used in this study, and leaf epidermis morphological structure (Fig. 1). Although stomata themselves have a relatively uniform structure, the random distribution and orientation of Arabidopsis stomata (Fig. 1a) make the measurement task more challenging than for barley, whose stomata are aligned in parallel rows with fixed orientations (Fig. 1b). Using criteria outlined in Jayakody *et al.* (2021) to assess image quality, we found that Arabidopsis samples were rated as medium quality, whereas barley samples are considered high quality. The observed disparity in measurement quality supports the claim that image quality has a major impact on model performance (Jayakody *et al.*, 2021).

Ellipse fitting is a commonly used method in automated stomatal pore measurement systems to estimate pore area, width and length (Jayakody *et al.*, 2017; Bhugra *et al.*, 2018, 2019; Li *et al.*, 2019; Liang *et al.*, 2021). However, ellipse fitting is limited to stomatal pores that have ovular shapes, such as dicot stomata delineated by kidney-shaped guard cells (Fig. 1c). Plants like barley, which have stomatal pores delineated by dumbbell-shaped guard cells, have pores that resemble a coin slot, which cannot be represented accurately by an ellipse, leading to under- or overestimation in derived measurements (Fig. 1c). SAI uses direct mask segmentation of the stomatal pores, which can flexibly represent any pore shape and calculate pore area from the masked pixel area. Moreover, the efficacy of ellipse fitting is positively correlated with the extent of stomatal pore opening (Li *et al.*, 2019), so it cannot be used effectively for stomatal pore assays under

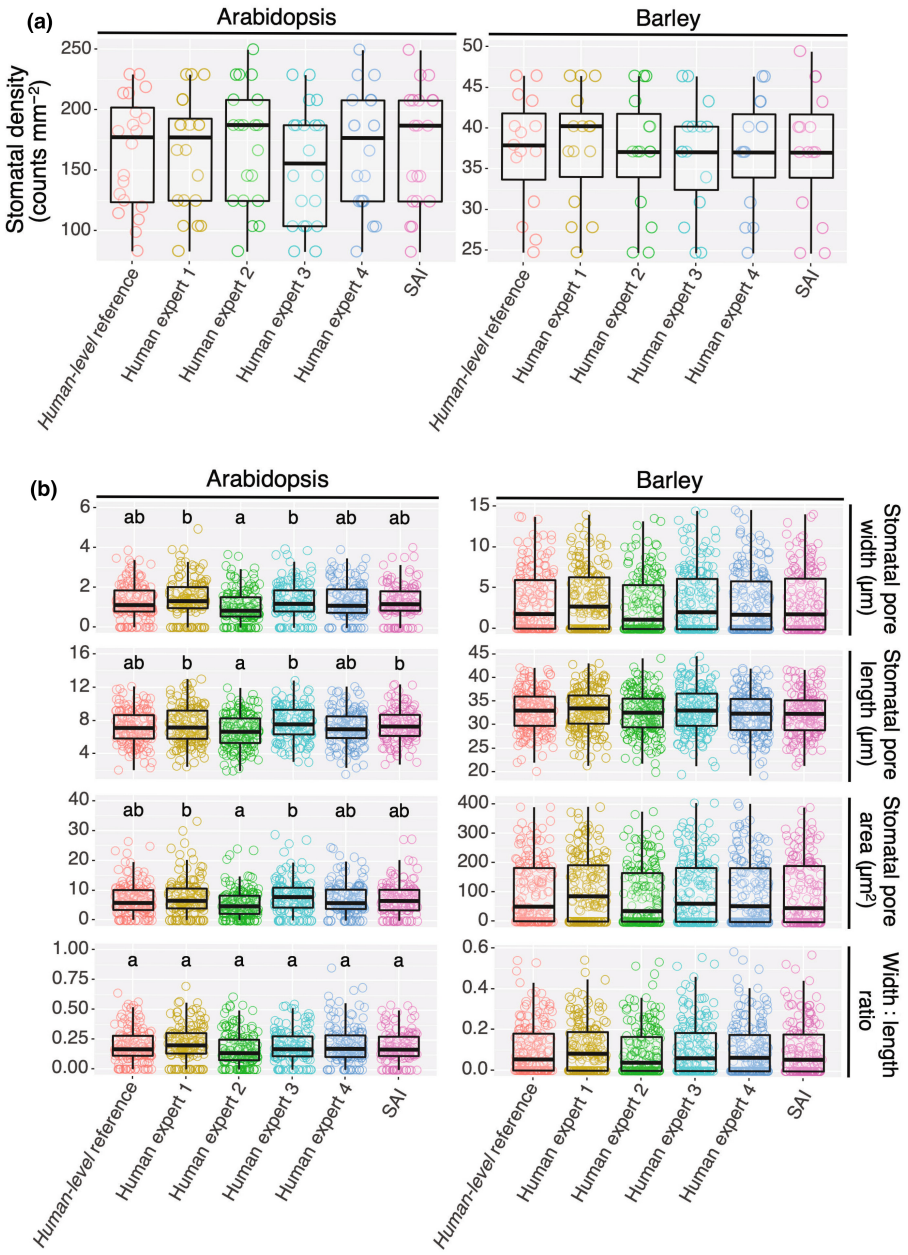

Fig. 5 Measurement comparison of stomatal density (a) and pore measurements (b) in width, length, area and width/length ratio of Arabidopsis and barley. Stomatal density from individual image and stomatal pore measurement were visualised in box and whiskers plots. Whiskers plot represents minimum and maximum values, and box plot represents first quartile, median and third quartile. Four human experts on stomatal morphology, *human-level* reference (the average of human measurements) and StomaAI (SAI) presented with one-way ANOVA with Tukey's Honest Significance Difference test (density: $n = 20$ in Arabidopsis, $n = 15$ in barley; pore measurements: $n > 120$/measurer in Arabidopsis, $n > 160$/measurer in barley). No differences found between source of measurements in stomatal density and all pore measurement attributes barley, a and b represent groups without significant difference in Arabidopsis, $P \leq 0.05$ between group.

experimental conditions that measure stomata that are partially or completely closed (e.g. stomata exposed to ABA, high CO_2, H_2O_2, or darkness; Zhang *et al.*, 2001; Chater *et al.*, 2015; Xu *et al.*, 2021). To solve this issue, SAI classifies stomata before measuring, so that open and closed stomata are managed separately and accurately reported (Fig. 3, status-specific processing).

We have determined that SAI achieves *human-level* performance and produces consistent conclusions with human researchers (Figs 4–6). Beyond this, SAI has many advantages compared with manual measurement. SAI produces stable and reproducible measurements, unlike rater's bias created by individual humans as observed in our Average-Human/Machine Test (Figs 5, S3). SAI's consistent predictions guarantees that counting and measurements are reproducible from the same set of samples. SAI can also enable researchers to rapidly verify conclusions from existing reports without weeks of human effort. Provided that the samples from which a biologist draws their conclusions are available, SAI can produce a set of measurements within minutes, ready for statistical analysis and assessment (Fig. 7).

SAI detects fewer stomata than experts for pore measurement but not counting (Tables S1, S2). The tasks of stomatal counting and pore measurement have different purposes. Stomatal counting aims to identify each stoma in an image, while pore measurement requires analysis of fine details of the stomatal pore. Human experts can use their experience to measure stomatal pores that are blurry, occluded, truncated or unresolved, whereas SAI will ascribe a low level of confidence to such samples and discards them below a minimum value of confidence (0.6) to prevent false positives (incorrect prediction of stomatal presence). SAI prioritises measurement precision over quantity, discarding

New Phytologist

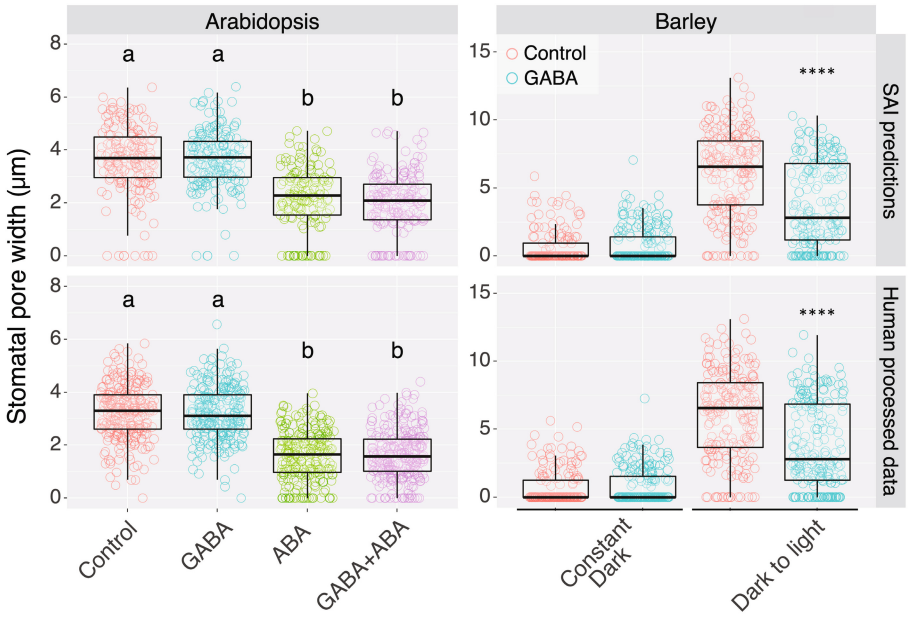

Fig. 6 StomaAI (SAI) predicted measurements are consistent with outcomes obtained by human researchers. Stomatal pore width in SAI predictions and human measurements collected from original microscope images with treatment under 25 μM ABA with 2 mM GABA (Arabidopsis) or 1 mM GABA (barley) during a dark-to-light transition (Xu *et al.*, 2021). Boxes show first quartile, median and third quartile; whiskers extend to minimum and maximum values. Data were tested using one-way ANOVA followed by Tukey's Honest Significance Difference ($n > 140$/group in Arabidopsis, $n > 150$/group in barley). Letters indicate significant differences at $P \leq 0.0001$; ****, $P \leq 0.0001$.

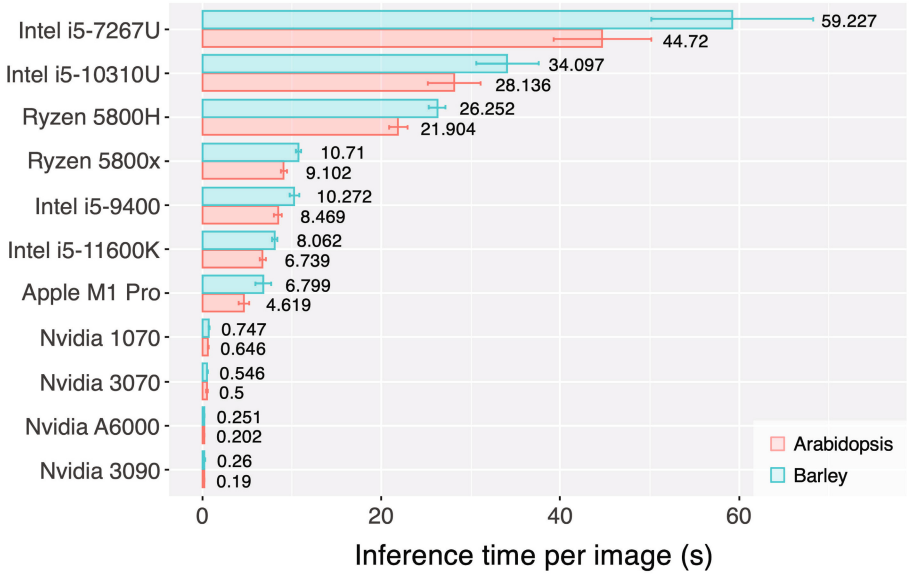

Fig. 7 Time elapsed for measurement from start to finish was measured when processing a single microscope image using StomaAI Image processing time (model inference time and post-processing time) were recorded. All processors are tested on the same image set of Arabidopsis and barley at the native resolution of images from each species. For all tests, the confidence threshold was set to 0.6. Bars shown mean ± SD of inference time per sample (Arabidopsis: $n = 20$, barley: $n = 15$). All processors were on desktop computers, except Intel i5-7267U (MacBook Pro 2017), Intel i5-10310U, Ryzen 5800H, Apple M1 Pro (MacBook Pro 2022) and Nvidia 3070, which were on a laptop.

blurry, occluded or unresolved stomata, or those that appear to be truncated by an image edge, although these latter may be retained for counting. While pore measurements that could be salvaged by experts are often discarded by SAI, it is important to note that the extraction of pore measurements from every stoma in an image becomes less important due to SAI's high throughput and ability to measure orders of magnitude more stomatal pores across many more images in a fraction of the time. Future enhancements to SAI could incorporate the ability to measure stomatal pore apertures from images that cover the same area but at a different focal plane to incorporate the majority of currently discarded stomata.

Compared with manual processes, SAI is exceptionally efficient. From experience, human experts take 2–5 min to process one image depending on its quality, the number of stomata present, the measurements required and stomatal opening status. SAI can produce measurements from a high-resolution image in 6–12 s while running on a mid-range desktop computer's central processing unit (CPU; Fig. 7). When using a graphics processing unit (GPU), SAI further increases the disparity between human and machine measurement. On an Nvidia GTX 1070, SAI is able to process an image in 700 ms. With a higher-level GPU, SAI can process hundreds of images per minute – the equivalent output of 4–9 h of human work. Automatically processing images using SAI makes it trivial to quickly achieve minimum measurement numbers required per treatment group for statistical testing (100–200 stomata are routinely used in physiological research); indeed, SAI enables researchers to increase the statistical power of

their conclusions by easily counting thousands of stomata, a previously unthinkable number. By decoupling human effort from measurement, researchers can also analyse stomata under many more experimental conditions. The law of large numbers will provide more accurate summary statistics and presumably allow detection of smaller variations in stomatal function due to treatments or genetics than at present.

This hassle-free, high-resolution, time-efficient data acquisition assistant has the potential to accelerate research on a cell type that has a major impact on plant physiology. Moreover, SAI's ability to learn how to measure stomata from both barley and Arabidopsis gives us confidence in its ability to do so in other species. Towards this, we provide an additional model with SAI that has been trained on a combined species dataset (available at GitFront https://gitfront.io/r/jpb/u6BtFFMkNGCv/SAI-training). As stomata share some common visual features, this model can be used as a starting point for researchers wishing to use SAI on a new type of plant. A set of measured examples that conform to our annotation format could be used to fine-tune the provided combined species model.

In this study, we only consider using SAI in stomata counting and pore feature measurement. However, given sufficient labelled data, extension to measurement of other relevant cell structures would be possible (e.g. total stomatal size to extrapolate conductance measurements, or measurements of subsidiary cell characteristics). More generally, SAI could be used to measure other structures captured via microscopy (e.g. trichomes or surface lesions caused by pathogens). We encourage researchers to contribute to SAI themselves, or to contact us and collaborate to increase its utility.

StomaAI is a reliable and efficient solution to automate stomatal pore measuring for plant biologists via a user-friendly web app (online demo at https://sai.aiml.team; the full version is available at https://github.com/XDynames/SAI-app). StomaAI is a new tool that can free researchers from labour-intensive, low-throughput measurement tasks, accelerating the speed of physiology-based plant research, regardless of the shape of the stomatal pore.

Acknowledgements

This research was supported by PhD Studentship, Grains Research and Development Corporation (GRDC)/ARC CoE Plant Energy Biology Scholarship (UWA1708-010RTX) to NS and Australian Government Research Training Program (RTP) Scholarship to JPB. MG was supported by the Australian Research Council Centre of Excellence CE140100008 and Discovery Project, Australian Research Council DP210102828. High-performance compute resources used in this work were funded by Linkage Project, Australian Research Council LE190100080. Lockheed Martin Australia supported this research via a research scholarship awarded to JPB. We thank Natalie Betts for manuscript editing. Open access publishing facilitated by The University of Adelaide, as part of the Wiley - The University of Adelaide agreement via the Council of Australian University Librarians. Open access publishing facilitated by The University of Adelaide, as part of the Wiley - The University of Adelaide agreement via the Council of Australian University Librarians.

Competing interests

None declared.

Author contributions

NS, MG and CS conceived the research. NS produced barley images; developed and labelled all model training data annotation; developed and conducted experiments for Average-Human/Machine Test (Figs 4, 5, S2, S3); conducted experiments for 'Application on real-world physiological experiments' in the Results section (Figs 6, S4); provided advice on the development of SAI online demo and offline application; and drafted all figures. JPB translated the measurement problem into a set of machine vision tasks; converted human annotations to machine learning compatible format for Arabidopsis; conducted deep-learning experiments to determine feasibility, model type and hyperparameters (Methods S1–S9); and developed and deployed SAI online demo showcasing and offline application. NS and JPB developed the postprocessing procedures for improving measurement quality and immature stomata filtration and conducted run time benchmarking analysis of SAI on various hardware (Fig. 7). HC set up the experimental framework for Mask-R-CNN and translated human annotations for barley into a machine learning compatible format. NW-H, CS, NS and JPB developed the evaluation system for Average-Human/Machine Test. BX and XF provided Arabidopsis images. BX, XF, AP and NS provided human processed measurements for Average-Human/Machine Test. NS and JPB drafted the original manuscript. NS, JPB, MG, CS, NW-H, BX and HC provided edits and revisions to the final manuscript. NS and JPB contributed equally to this work.

ORCID

James Paul Bockman https://orcid.org/0000-0002-2840-2533
Hao Chen https://orcid.org/0000-0003-4417-614X
Xueying Feng https://orcid.org/0000-0002-9573-6082
Matthew Gilliham https://orcid.org/0000-0003-0666-3078
Adriane Piechatzek https://orcid.org/0000-0002-7958-5771
Na Sai https://orcid.org/0000-0002-9482-6115
Chunhua Shen https://orcid.org/0000-0002-8648-8718
Nathan Watson-Haigh https://orcid.org/0000-0002-7935-6151
Bo Xu https://orcid.org/0000-0002-7583-2384

Data availability

The data that support the findings of this study are openly available on GitFront at https://gitfront.io/r/jpb/u6BtFFMkNGCv/SAI-training.

References

Bheemanahalli R, Wang C, Bashir E, Chiluwal A, Pokharel M, Perumal R, Moghimi N, Ostmeyer T, Caragea D, Jagadish SVK. 2021. Classical phenotyping and deep learning concur on genetic control of stomatal density and area in sorghum. *Plant Physiology* **186**: 1562–1579.

Bhugra S, Mishra D, Anupama A, Chaudhury S, Lall B, Chugh A. 2018. Automatic quantification of stomata for high-throughput plant phenotyping. In: *24th International Conference on Pattern Recognition (ICPR)*. Los Alamitos, CA, USA: IEEE Computer Society, 3904–3910.

Bhugra S, Mishra D, Anupama A, Chaudhury S, Lall B, Chugh A, Chinnusamy V. 2019. Deep convolutional neural networks based framework for estimation of stomata density and structure from microscopic images. In: Leal-Taixé L, Roth S, eds. *The European Conference on Computer Vision (ECCV) Workshops*. Cham, Switzerland: Springer International, 412–423.

Bourdais G, McLachlan DH, Rickett LM, Zhou J, Siwoszek A, Häweker H, Hartley M, Kuhn H, Morris RJ, MacLean D *et al.* 2019. The use of quantitative imaging to investigate regulators of membrane trafficking in Arabidopsis stomatal closure. *Traffic* **20**: 168–180.

Cai S, Papanatsiou M, Blatt MR, Chen Z-H. 2017. Speedy grass stomata: emerging molecular and evolutionary features. *Molecular Plant* **10**: 912–914.

Chater C, Peng K, Movahedi M, Dunn JA, Walker HJ, Liang Y-K, McLachlan DH, Casson S, Isner JC, Wilson I *et al.* 2015. Elevated CO_2-induced responses in stomata require ABA and ABA signaling. *Current Biology* **25**: 2709–2716.

Cheng Y, Cao L, Wang S, Li Y, Wang H, Zhou Y. 2014. Analyses of plant leaf cell size, density and number, as well as trichome number using cell counter plugin. *Bio-Protocol* **4**: e1165.

Duarte KTN, de Carvalho MAG, Martins PS. 2017. Segmenting high-quality digital images of stomata using the wavelet spot detection and the watershed transform. In: *The 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications – Volume 4: VISAPP, (VISIGRAPP 2017)*. Setúbal, Portugal: SciTePress, 540–547.

Eisele JF, Fäßler F, Bürgel PF, Chaban C. 2016. A rapid and simple method for microscopy-based stomata analyses. *PLoS ONE* **11**: e0164576.

Falk T, Mai D, Bensch R, Çiçek Ö, Abdulkadir A, Marrakchi Y, Böhm A, Deubner J, Jäckel Z, Seiwald K *et al.* 2019. U-Net – deep learning for cell counting, detection, and morphometry. *Nature Methods* **16**: 67–70.

Fetter KC, Eberhardt S, Barclay RS, Wing S, Keller SR. 2019. StomataCounter: a neural network for automatic stomata identification and counting. *New Phytologist* **223**: 1671–1681.

Gibbs J, Mcausland L, Robles-Zazueta CA, Murchie E, Burgess A. 2021. A deep learning method for fully automatic stomatal morphometry and maximal conductance estimation. *Frontiers in Plant Science* **12**: 2703.

Girshick R. 2015. Fast R-CNN. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. Los Alamitos, CA, USA: IEEE Computer Society, 1440–1448.

Girshick R, Donahue J, Darrell T, Malik J. 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Los Alamitos, CA, USA: IEEE Computer Society, 580–587.

Gray A, Liu L, Facette M. 2020. Flanking support: how subsidiary cells contribute to stomatal form and function. *Frontiers in Plant Science* **11**: 881.

He K, Gkioxari G, Dollar P, Girshick R. 2017. Mask R-CNN. In: *The IEEE International Conference on Computer Vision (ICCV)*. Los Alamitos, CA, USA: IEEE Computer Society, 2961–2969.

Hetherington AM, Woodward FI. 2003. The role of stomata in sensing and driving environmental change. *Nature* **424**: 901–908.

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

Jayakody H, Petrie P, de Boer HJ, Whitty M. 2021. A generalised approach for high-throughput instance segmentation of stomata in microscope images. *Plant Methods* **17**: 1–13.

Kim TH, Böhmer M, Hu H, Nishimura N, Schroeder JI. 2010. Guard cell signal transduction network: advances in understanding abscisic acid, CO_2, and Ca^{2+} signaling. *Annual Review of Plant Biology* **61**: 561–591.

Laga H, Shahinnia F, Fleury D. 2014. Image-based plant stomata phenotyping. In: *13th International Conference on Control Automation Robotics & Vision (ICARCV)*. IEEE, 217–222.

Li K, Huang J, Song W, Wang J, Lv S, Wang X. 2019. Automatic segmentation and measurement methods of living stomata of plants based on the CV model. *Plant Methods* **15**: 67.

Liang X, Xu X, Wang Z, He L, Zhang K, Liang B, Ye J, Shi J, Wu X, Dai M *et al.* 2021. StomataScorer: a portable and high-throughput leaf stomata trait scorer combined with deep learning and an improved CV model. *Plant Biotechnology Journal* **20**: 1–15.

Lin LI-K. 1989. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* **45**: 255–268.

Lin TY, Maire M, Belongie S, Hays J, Perona P, Ramanan D, Dollár P, Zitnick CL. 2014. Microsoft COCO: common objects in context. In: Fleet D, Pajdla T, Schiele B, Tuytelaars T, eds. *The European Conference on Computer Vision (ECCV)*. Cham, Switzerland: Springer, 740–755.

McLachlan DH, Kopischke M, Robatzek S. 2014. Gate control: guard cell regulation by microbial stress. *New Phytologist* **203**: 1049–1063.

Meeus S, Van den Bulcke J, Wyffels F. 2020. From leaf to label: a robust automated workflow for stomata detection. *Ecology and Evolution* **10**: 9178–9191.

Omasa K, Onoe M. 1984. Measurement of stomatal aperture by digital image processing. *Plant and Cell Physiology* **25**: 1379–1388.

Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, Antiga L *et al.* 2019. PYTORCH: an imperative style, high-performance deep learning library. In: Wallach H, Larochelle H, Beygelzimer A, Alché-Buc F, Fox E, Garnett R, eds. *Advances in neural information processing systems*. Curran Associates, 8026–8037.

Ren S, He K, Girshick R, Sun J. 2017. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **39**: 1137–1149.

Rizhsky L, Liang H, Shuman J, Shulaev V, Davletova S, Mittler R. 2004. When defense pathways collide. The response of Arabidopsis to a combination of drought and heat stress. *Plant Physiology* **134**: 1683–1696.

Roelfsema MRG, Hedrich R. 2005. In the light of stomatal opening: new insights into 'the Watergate'. *New Phytologist* **167**: 665–691.

Sakoda K, Watanabe T, Sukemura S, Kobayashi S, Nagasaki Y, Tanaka Y, Shiraiwa T. 2019. Genetic diversity in stomatal density among soybeans elucidated using high-throughput technique based on an algorithm for object detection. *Scientific Reports* **9**: 7610.

Saponaro P, Treible W, Kolagunda A, Chaya T, Caplan J, Kambhamettu C, Wisser R. 2017. DeepXScope: segmenting microscopy images with a deep neural network. In: *The IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Los Alamitos, CA, USA: IEEE Computer Society, 843–850.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B *et al.* 2012. FIJI: an open-source platform for biological-image analysis. *Nature Methods* **9**: 676–682.

Shimazaki KI, Doi M, Assmann SM, Kinoshita T. 2007. Light regulation of stomatal movement. *Annual Review of Plant Biology* **58**: 219–247.

Sinha RK. 2004. *Modern plant physiology*. Pangbourne, UK: Alpha Science.

Toda Y, Toh S, Bourdais G, Robatzek S, Maclean D, Kinoshita T. 2018. DeepStomata: facial recognition technology for automated stomatal aperture measurement. *bioRxiv*. doi: 10.1101/365098.

Vialet-Chabrand S, Brendel O. 2014. Automatic measurement of stomatal density from microphotographs. *Trees* **28**: 1859–1865.

Wu Y, Kirillov A, Massa F, Lo WY, Girshick R. 2019. *Detectron2*. [WWW document] URL https://github.com/facebookresearch/detectron2 [accessed 16 September 2020].

Xie J, Fernandes SB, Mayfield-Jones D, Erice G, Choi M, Lipka AE, Leakey ADB. 2021. Optical topometry and machine learning to rapidly phenotype stomatal patterning traits for maize QTL mapping. *Plant Physiology* **187**: 1462–1480.

Xu B, Long Y, Feng X, Zhu X, Sai N, Chirkova L, Betts A, Herrmann J, Edwards EJ, Okamoto M *et al.* 2021. GABA signalling modulates stomatal opening to enhance plant water use efficiency and drought resilience. *Nature Communications* **12**: 1952.

Ye W, Murata Y. 2016. Microbe associated molecular pattern signaling in guard cells. *Frontiers in Plant Science* 7: 583.

Zhang X, Zhang L, Dong F, Gao J, Galbraith DW, Song C-P. 2001. Hydrogen peroxide is involved in abscisic acid-induced stomatal closure in *Vicia faba*. *Plant Physiology* **126**: 1438–1448.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Example of exported measurement visualisation and output of stomatal counting/density and pore measurement comma-separated values files.

Fig. S2 StomaAI prediction vs *human-level* reference set for Arabidopsis and barley stomatal density and pore measurements.

Fig. S3 Individual human measurements vs *human-level* reference set for Arabidopsis and barley with corresponding counts difference or relative error of stomatal counts, density, pore width, length, area and width/length ratio.

Fig. S4 Mean of stomatal pore width and measurement distribution illustration.

Methods S1 Default training procedure.

Methods S2 Model performance.

Methods S3 Batch size.

Methods S4 Keypoint head complexity.

Methods S5 Keypoint and mask head pooler resolution.

Methods S6 Image resolution.

Methods S7 Keypoint head complexity at higher resolution.

Methods S8 Final schema.

Methods S9 Pseudocode.

Notes S1 Downloading and installing StomaAI.

Notes S2 Running StomaAI.

Table S1 Summary of counted stomata in number and mean of stomatal density.

Table S2 Summary of measured stomata in number and mean of corresponding stomatal pore feature measurements.

Please note: Wiley is not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.