Detection and Calculation of Stomatal Density Using YOLOv5: A Study of High-Yielding Patchouli Varieties

Arie Qur'ania^{1,2}, Yeni Herdiyeni^{1,*}, Wisnu Ananta Kusuma¹, Aji Hamim Wigena³, and Sri Suhesti^{4,5}

¹ Computer Science Department, IPB University, Bogor, Indonesia

²Computer Science Department, Pakuan University, Bogor, Indonesia

⁴ Plantation Instruments Standardization Agency, Bogor, Indonesia

⁵ Indonesian Center for Estate Crops Research and Development, Bogor, Indonesia

Email: ariequrania@apps.ipb.ac.id (A.Q.); yeni.herdiyeni@apps.ipb.ac.id (Y.H.); ananta@apps.ipb.ac.id (W.A.K.);

aji_hw@apps.ipb.ac.id (A.H.W.); hesti.khrisnawijaya@gmail.com (S.S.)

*Corresponding author

Abstract-Stomatal density influences plant photosynthesis, transpiration, and secondary production like fruit and oil. It could serve as a selection criterion for developing plant varieties. The genetic diversity of patchouli is still relatively limited owing to a lack of flowering and fruiting; therefore, genetic variability is also limited. One approach to overcoming this problem is to collect plants from specific regions, called accessions, to identify potential varieties capable of producing abundant and high-quality patchouli oil. Parameters such as stomatal density were evaluated during this process. Conventional manual calculations have inherent drawbacks, including time constraints, low precision, and susceptibility to bias. Therefore, automated methods are essential for stomatal detection models and counting calculations based on deep learning. The dataset consisted of 100 and 400 microscopy images split at a ratio of 8:2 for the training and testing data, respectively. A stomata detection model using YOLOv5 achieved precision, recall, and F1-Score of 0.88 each. The accuracy of the stomata calculation on the test data was 97%. This result demonstrates the ability of the model to calculate the stomatal density in microscopy images.

Keywords—deep learning, patchouli, stomata detection, stomatal density, YOLOv5

I. INTRODUCTION

Patchouli is an herbaceous plant or shrub, and a rare wood that provides a pleasant scent for medicines, perfumes, and cosmetics. Patchouli produces essential oils, known as patchouli oil. The main compound responsible for the aroma is patchoulol or patchouli alcohol. Patchouli can be used as a spice, eaten as functional food or prescribed as a drug. The chemical compound patchoulol causes patchouli to react as a drug in the treatment of human diseases. Patchouli is useful as an influenza antiviral, antidepressant, lung-protective, brain-protective, and antibacterial agent [1]. However, the genetic diversity of patchouli remains relatively low. This is because patchouli does not flower or bear fruit; therefore, there are no natural or artificial crossings. Plant propagation through stem cutting and tissue culture results in a lack of genetic variability [2, 3].

Efforts have been made to increase the genetic diversity of patchouli by exploring various production centres and other regions. Plants for patchouli research were collected from specific sites (accessions). The next step was to continue the selection of accessions based on morphological characteristics such as leaf shape and stem colour. Accessions were observed, characterised, evaluated, and selected to obtain patchouli accessions with the potential to produce high-quality products [4]. The parameters observed during accession selection included the number of stomata on the leaves based on microscopic observations.

Stomatal density directly affects the photosynthesis and transpiration of the plants, which in turn affects the secondary production of the plants, such as flowers, fruits, and oil [5]. Stomatal density has potential as a selection criterion for varietal development in tea plants [6] and for the breeding of drought-tolerant oil palms [7]. Studies on stomatal density and patchouli oil glands have been conducted to determine the internal structure (anatomy) of patchouli [8, 9]. The analysis of patchouli stomatal density can be used for plant breeding in the selection process of new patchouli varieties based on ideotype breeding (character traits for selection). Ideotype breeding is a plant breeding method that focuses on the selection and development of plants with specific characteristics [10].

The traditional method for obtaining stomatal density is manual calculation, which is time-consuming and has low accuracy [11]. This occurs because of the weak and limited image quality [12]. To address this problem, an intelligent stomatal counting model based on object detection is needed, which will help researchers calculate the amount and density of stomata. Research on stomatal detection has

³ Statistic Department, IPB University, Bogor, Indonesia

Manuscript received December 27, 2023; revised February 17, 2024; accepted March 21, 2024; published July 5, 2024.

been conducted using template-matching techniques on microscopy images of wheat and barley leaves that have almost the same stomatal size [13]. Six templates were utilised to capture variations in stomatal shape. In image analysis, template matching compares an input image with a predefined reference image to identify and locate objects or patterns. Stomatal detection depends on the template that has been specifically chosen, so it cannot capture the various shapes and positions of the stomata, leading to inaccurate detection. Low contrast values between stomatal cells and backgrounds result in low accuracy in distinguishing stomata from other objects, and template matching requires manual segmentation of multiple training images to build template datasets, which can be time-consuming and biased.

Generally, research on stomata in microscopy images has been conducted using image-processing approaches with segmentation, thresholding, and morphological operations within the framework of small-scale data. Image-processing research has focused on machine learning and computational processes that can recognise increasingly diverse object patterns. Computer vision is widely associated with image processing and machine learning, and it is used to predict or detect objects in a broader range of studies [14]. Stomata detection research has continued to grow with the development of machine learning and deep learning algorithms for object detection.

Computer vision has developed in various fields, such as agriculture, to classify varieties of chickpeas [15], wheat [16], and breeding programmes [17], and to classify varieties for the nursery plant industry [18]. Using dermoscopic images, Deep Convolutional Neural Network (DCNN) models in the medical field, which contain various artefacts such as hair, gel bubbles, and blood vessels, which pose a challenge in the skin lesion classification process, are used to help diagnose skin diseases [19]. The DCNN was applied for the identification of human digestive tract abnormalities using endoscopic images. The problems of interclass similarity in gastrointestinal abnormalities and the presence of artefacts in the images became a challenge in the identification task [20]. Similar to medical images, microscopy images of patchouli stomata contain artefacts such as trichomes and epidermis, which pose a challenge in the detection of stomata overlapping with trichomes. Microscopy images are downscaled to train the model without losing details on the data content or knowledge of stomatal object recognition. YOLOv5 is part of the CNN method that uses the Convolution + Batchnorm + SiLU (CBS) module.

Since 2014, deep learning-based object detection has evolved. It is primarily divided into two parts: a singlestage and a two-stage detector. The single-stage detector technique comprises You Look Only Once (YOLO), single-shot multi-box detector (SSD), Retina-Net, CornerNet, and CenterNet. Object detection techniques that include two-stage detectors include Region-Convolutional Neural Networks (R-CNNs), Spatial Pyramid Pooling Networks (SPPNNs), fast R-CNNs, faster R-CNNs, and Feature Pyramid Networks (FPNs) [21].

YOLO was the first single-stage detector to apply a single neural network to an entire image in deep learning. YOLO divides an image into multiple regions and simultaneously predicts bounding boxes and probabilities from each region. YOLO is trained to understand images and immediately optimizes recognition performance. The YOLO algorithm was designed to detect objects in images by predicting the bounding box and class probabilities directly from the image in a single stage. This approach differs from two-stage object detection algorithms, which use a two-stage proposal process for region classification [21].

Stomata detection research was conducted on a small dataset (183 soybean leaf images), resulting in high accuracy and precision. Some stomata in the images were visually blurred but were still analysed if they were recognisable as stomata by the human eye. The object detection model using YOLO demonstrated exceedingly high average accuracies for YOLOv3, YOLOv4, and YOLOv5 (94%, 98%, and 99%, respectively) [22]. This indicates that the YOLO model can successfully recognise stomatal patterns in the test images without requiring a large number of labelled images.

Researchers encounter three main challenges when working with microscopic images: size, quantity, and the time required for annotation [23]. The YOLO model utilises smaller images without losing important details, and the trained model can detect objects in larger images. However, YOLO has the disadvantage of reduced accuracy in detecting small objects. An improved YOLO has been developed to overcome this problem. YOLOv5 is open-source software that enables researchers to develop the YOLOv5 architecture by adding an attention mechanism module. Research using YOLOv5 encompasses the detection of stomata in broad bean leaves with a precision value of 93% [24]. Other research on nonstomata objects include the detection of mummy berry disease with 96% accuracy [25] and fish detection with 95% accuracy [26].

Stomatal detection studies commonly use datasets from fruit crops, nuts, and food crops such as wheat and maize. However, there is a lack of stomata datasets for aromatic plants such as patchouli in the Cuticle Database (https://cuticledb.eesi.psu.edu), especially for superior patchouli varieties. Therefore, we need patchouli stomata datasets for training and testing object detection models. The analysis of microscopy images of stomata in plants with diverse leaf structures is challenging [27]. Some images may be of consistent quality and ease of analysis, but obstructed objects and unclear guard cells hinder successful detection. Stomatal density, which represents the number of stomata per field of view, was critical for analysis. Current calculations require manual labelling using software such as ImageJ, which is a cumbersome and time-consuming process with limited accuracy. Biologists often select microscopy images with visible stomata for ease of observation and counting, overlooking images with poor quality and noise [27].

Patchouli has trichomes (leaf hairs) on its leaves, so the resulting microscopy images sometimes overlap between stomata and trichomes. The patchouli epidermis is so thin that the stomata sometimes appear blurred. Overlapping and blurring of microscopy images is the main challenge in patchouli stomata detection research. The microscopy image dataset is divided into datasets without noise and those with noise. YOLOv5 has several advantages to support more applications, including being lightweight, fast, and able to run on mobile platforms [28]. The algorithm used for stomatal detection is YOLOv5, which includes YOLOv5s, YOLOv5m, and YOLOv5l, which were combined in this study as a baseline for detecting stomata in superior varieties of patchouli.

The objectives of this study were 1) to integrate a deep learning algorithm for automatic detection of stomata in microscopy images of patchouli and 2) to count stomata to analyse stomatal density as a parameter that affects production and essential oil levels in patchouli. The results of this study can be used to conduct stomatal density analysis to contribute to plant breeding (proliferation) in the selection of new patchouli varieties based on ideotype breeding (characteristics for selection) and patchouli secondary parameters.

A. Problem-Solving Approach

Stomata regulate two physiological functions, namely, photosynthesis and transpiration [29]. Variations in stomatal density suggest that this trait can be improved through well-designed breeding strategies [6]. Ideotype breeding is a plant breeding method used to select and develop plants with specific characteristics. Ideotype breeding involves the selection of desirable traits to create an ideal plant [30]. However, ideotype breeding aims to produce plants with better yield potential and pest and disease resistance.

Traditional stomata counting relies on visual observations by researchers. Stomata types are diverse and scattered in many images of different sizes, causing manual stomata calculations to be easily missed, timeconsuming, and costly. To address this problem, a stomatal detection model and automatic stomata count calculations based on computer vision are needed.

B. State of the Art and Novelty

Many studies on stomatal detection have been conducted using plant datasets at the plant family and species levels, and this study proposes patchouli varieties. Related studies are presented in Table I.

Based on previous research and studies on stomatal detection, the best accuracy was achieved using the YOLOv5 algorithm. Based on the strength of YOLOv5, the YOLOv5 algorithm was used to detect and calculate stomatal density in patchouli. Patchouli has trichomes, or leaf hairs, so the stomata overlap with the trichomes. This condition creates noisy microscopy images. This is a challenge and a novelty in research related to stomatal detection in patchouli microscopy images.

| TABLE I. | RESEARCH ON STOMATA (| (2017 - 2023) |) |
|----------|-----------------------|---------------|---|
|----------|-----------------------|---------------|---|

| Methods | Object | Accuracy |
|---|---------------------------------------|----------|
| Cascade Object Detection (COD) [27] | Grape | 74% |
| Deep Convolutional Neural Networks (DCNN) [31] | Cuticle database | 94% |
| YOLOv3 [23] | Common beans, barley, and soybeans | 91% |
| Deep learning (VGG19) [32] | Herbarium | 94% |
| Single-Shot Detector (SSD) [33] | Quinoa (whole grain) | 94% |
| Mask Region-CNN (Mask R- CNN) [34] | Gymnosperms | 90% |
| Faster R-CNN [35] | Wheat | 92% |
| DCNN [36] | Maize | 94% |
| SSD [7] | Oil palm | 96% |
| YOLOv5 [22] | Soybean | 98% |
| YOLOv5 [24] | Broad bean | 93% |

II. MATERIALS AND METHODS

The research method of the object detection-based smart count model for calculating patchouli stomatal density is described in the research flowchart shown in Fig. 1.



Fig. 1. Research flow chart of the patchouli stomata smart count model.

The research method was divided into three stages: the first stage of data collection and preprocessing, the second stage of building a stomatal detection model using the YOLOv5 technique, and the third stage of analysing stomatal density parameters. Below is a description of each activity.

A. Data Collection and Preprocessing Stages

The Research Center for Spice and Medicinal Plants (BALITTRO) released two patchouli varieties, Sidikalang and Patchoulina2, from which the data for this study were obtained. Both varieties were selected because they produce high-quality patchouli oil. In addition, Sidikalang is tolerant to bacterial wilt, while Patchoulina2 is resistant to bacterial wilt. Patchouli propagated from stem cuttings was grown in the experimental garden of the Plantation Instrument Standardization Agency in West Java, Indonesia. The process of stomatal microscopy image acquisition is illustrated in Fig. 2.



Fig. 2. Stomata image acquisition process.

Microscopy images were obtained by observing the stomata of each variety. The fifth leaf from the tip was then collected from different stems. Leaves with a central width of one centimetre were picked for observation. Stomata were prepared via the whole-mount method. Using an Olympus BX53 microscope at 200× magnification, we counted the number of stomata for observation.

The whole-mount method determines the morphology of the epidermis and the stomata. The materials used were nitric acid (HNO₃), distilled water (aquades), glycerine, and safranin. The working method was to prepare a solution of nitric acid and water in a ratio of 1:3 in a glass container, which was then heated to boiling for 2-4 min along with leaf samples that had been cut to a size of 1×1 cm to peel off the epidermis. The epidermis was removed and placed in a petri dish filled with water. Next, the epidermis was removed (placed on a glass object) and observed. After cleaning with water to remove dirt, safranin was applied until evenly distributed, and then the samples were rinsed with water until clean. Subsequently, glycerine was sprinkled, and the sample was covered with a cover glass. Nail polish was placed on the edge of the cover glass to prevent air from entering the object of preparation (the research specimen), which was ready for observation.

B. Patchouli Stomata Dataset

Patchouli has characteristics that distinguish it from other plants, such as the presence of numerous trichomes or leaf hairs that cover the stomata. The epidermis is thin; therefore, the guard cells, or pores, of the stomata are opaque. On this basis, the acquisition of microscopy images of patchouli was divided into two parts: 1) images with clearly visible stomata (without noise) and 2) images with trichomes (leaf hairs) or blurred stomata with images with added contrast and brightness (images with noise). In the hope that guard cells and pore stomata can be observed and detected, the contrast and brightness enhancement of the blurred stomata can be observed via microscopic imaging.

The dataset used in this study comprised two datasets: a dataset consisting of 100 images and 400 microscopy images of stomata. The image size was 1920×1080 pixels. The microscopy image dataset was divided into an 80% training dataset and a 20% test dataset. A total of 10,840 stomatal objects from 400 microscopy images of patchouli were manually labelled with bounding boxes using the LabelMe software tool. During the labelling process, bounding boxes were marked on all stomatal positions and conditions, as well as the contrast and brightness settings of the images, to obtain clearer stomata. Some stomata appeared unclear with random stomatal positions, such as vertical, horizontal, and oblique positions. An example of manually labelled stomatal objects in a microscopy image of patchouli at 200× magnification is shown in Fig. 3. This study focused on images with noise, and it is challenging for the model to correctly detect stomata. The dataset contains noisy images, as shown in Fig. 4.



Fig. 3. Manual labelling of stomata with bounding boxes.



Fig. 4. Microscopy image with noise.

C. Stomata Detection Model Using the YOLOv5 Algorithm

Stomata image data were analysed using the YOLOv5 architecture [19], which consists of 24 layers, including layers in the backbone (5 CBS, 4 C3, and 1 SPPF) and layers in the neck (4 concat, 2 upsampling, 4 CBS, and 4

C3). The training model was validated to assess its performance before the testing. The evaluation involved a comparison between the manual calculations performed by experts and the results obtained from the stomata calculation model using the YOLOv5 architecture. YOLOv5 follows the consistent design philosophy of the YOLO series in its algorithm. It is structured into three basic components: the backbone, the neck, and the head. The input layer processes the image detection and sends it to the backbone for feature extraction, resulting in feature maps of different sizes. Subsequently, the feature fusion network (neck) combines these features to generate three feature maps, namely, P3, P4, and P5, with dimensions of 80×80, 40×40, and 20×20, respectively, in YOLOv5. These maps are used to detect large, medium, and small objects within the image in the head section [37]. The training process using the YOLOv5 architecture is illustrated in Fig. 5.



Fig. 5. Stomata detection model with YOLOv5.

YOLOv5 performs object detection at multiple stages. In the initial stage, images collected in the model are processed through the deep layers of the artificial neural network. These layers are responsible for extracting features from an image and learning the patterns of the image presented. This model can identify objects in an image by learning patterns. In the next step, the model generates predictions for each cell in the grid. This prediction consists of a bounding box that indicates the object's location in the image, along with a class label specifying the type of object that was detected. After obtaining predictions for each cell, the next step is to combine and filter the predictions [38]. This process utilises the Non-maximum Suppression (NMS) technique outlined in Algorithm 1 to eliminate redundant or overlapping detections. The highest-scoring bounding boxes are retained while overlapping bounding boxes with lower scores are removed [39].

| Algorithm 1. Non-Maximum Suppression (NMS) | | | | |
|---|--|--|--|--|
| Require: Set of predicted bounding boxes B, confidence | | | | |
| scores S, IoU threshold τ , confidence threshold T | | | | |
| Ensure: Set of filtered bounding boxes F | | | | |
| $1: F \leftarrow \emptyset$ | | | | |
| 2: Filter the boxes: $B \leftarrow \{b \in B \mid S(b) \ge T\}$ | | | | |
| 3: Sort the boxes B by their confidence scores in descending | | | | |
| order | | | | |
| 4: while $B \neq \emptyset$ do | | | | |
| 5: Select the box b with the highest confidence score | | | | |
| 6: Add b to the set of final boxes F: $F \leftarrow F \cup \{b\}$ | | | | |
| 7: Remove b from the set of boxes B: $B \leftarrow B - \{b\}$ | | | | |
| 8: for all remaining boxes r in B do | | | | |
| 9: Calculate the IoU between b and r: you \leftarrow IoU(b, r) | | | | |
| 10: if you $\geq \tau$ then | | | | |
| 11: Remove r from the set of boxes B: $B \leftarrow B - \{r\}$ | | | | |
| 12: end if | | | | |
| 13: end for | | | | |
| 14: end while | | | | |

In YOLOv5, the loss function is determined by evaluating several metrics to quantify the disparities between the predicted outputs of the model and the actual labels. The two primary components of the loss calculation are the localisation loss (box loss) and the confidence loss. The localisation loss is a combination of the coordinate loss, as depicted in Eq. (1), and the size loss, as shown in Eq. (2). The confidence loss is a combination of the objectness loss, which assesses the probability of an object being present in the proposed region, as shown in Eq. (3). The absence of an object is calculated using Eq. (4), and the class loss is calculated using Eq. (5) [40].

$$\lambda_{coord} \ \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{ij}^{obj} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] \ (1)$$

$$\lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{ij}^{obj} \left[(\sqrt{w_i} - \sqrt{\widehat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\widehat{h}_i})^2 \right]$$
(2)

$$\sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{ij}^{obj} (C_i - \hat{C}_i)^2$$
(3)

$$\lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{noobj} (C_i - \hat{C}_i)^2$$
(4)

$$\sum_{i=0}^{S^2} 1_{ij}^{obj} \sum_{c \in classes} (p_i(c) - \hat{p}_i(c))^2 \tag{5}$$

where 1_{ij}^{obj} is an object and 1_{ij}^{noobj} is no object. The ground-truth coordinates are (x_i, y_i, w_i, h_i) , whereas the prediction coordinates are $(\hat{x}_i, \hat{y}_i, \hat{w}_i, \hat{h}_i)$. C_i is the ground truth bounding box confidence value, while \hat{C}_i is the prediction confidence value and intersection over union (IoU). The value p_i (c) is the ground truth and $\hat{p}_i(c)$ prediction of the class.

D. Evaluation of the Stomatal Detection Model

Model evaluation was conducted to measure model performance using a confusion matrix, which is a table describing the number of bounding boxes correctly and incorrectly identified by the model, including True Positives (TPs), True Negatives (TNs), False Positives (FPs), and False Negatives (FNs).

- The Average Precision (AP) is a metric used to measure the accuracy of stomatal detection. The AP measures the difference between the bounding box generated by the overlap model and the ground-truth bounding box. The AP was calculated for each object class and then averaged. The threshold used was AP50 (IoU = 0.5).
- The mean Average Precision (mAP) is the average of the AP values for all the object classes. This provides an overview of the performance of the model for stomatal detection using Eq. (6).

$$mAP = \frac{\sum AP}{n} \tag{6}$$

• Precision and recall are the metrics used to measure the accuracy and recall of a model, respectively. Precision measures whether the bounding box results generated by the model are correct, while recall measures the performance of the model in detecting all actual stomata objects. The formula for calculating precision is given by Eq. (7), and recall is given by Eq. (8).

$$Precision = \frac{TP}{TP+FP}$$
(7)

$$Recall = \frac{TP}{TP + FN}$$
(8)

• The F1–Score is a combination of accuracy and recall, providing the overall picture of the model's performance. The formula for calculating the F1–Score is shown in Eq. (9).

$$F1-Score = \frac{2 \times precision \times recall}{precision + recall}$$
(9)

These metrics were used to measure the performance of the YOLOv5 model in terms of stomatal detection. The evaluation was performed by comparing the model prediction with the ground truth and measuring the extent to which the model prediction matched the actual object in the stomatal image. The higher the F1–Score is, the better the performance of the model in detecting the stomata. The next step after the model has successfully detected and counted stomata is to analyse the stomatal density using the formula shown in Eq. (10).

$$Stomata \ density = \frac{\sum stomata}{area \ of \ field \ of \ view}$$
(10)

To measure the density of stomata, the field of view was used at $200 \times$ magnification, and the field of view area was measured using the following formula:

Area of the field of view = length \times width

 $= 661.52 \ \mu m \times 370.52 \ \mu m$

 $= 0.66152 \text{ mm} \times 0.37052 \text{ mm}$

 $= 0.2451063904 \text{ mm}^2 = \text{simplified to } 0.245 \text{ mm}^2$

III. RESULTS AND DISCUSSION

The model was trained using 200 epochs with a randomly selected training dataset to determine the accuracy of stomata detection in the test data. A comparative study was conducted by running the YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv4-CSP models. A comparison was made with YOLOv4-CSP because it is an improved version of YOLOv4 that achieves the optimal balance between speed and accuracy [41]. Tests were conducted on a dataset comprising 100 and 400 patchouli microscopy images to compare the mAP@0.5, precision, and recall values, as shown in Table II.

TABLE II. COMPARISON OF THE TWO DATASETS

| Model | Dataset image | mAP@0.5 | Precision | Recall |
|------------|------------------|---------|-----------|--------|
| YOLOv4_CSP | 100 | 0.874 | 0.94 | 0.71 |
| YOLOv5s | 100 | 0.921 | 0.88 | 0.88 |
| YOLOv5m | 100 | 0.918 | 0.89 | 0.86 |
| YOLOv51 | 100 | 0.916 | 0.89 | 0.86 |
| YOLOv4_CSP | 400 | 0.845 | 0.87 | 0.88 |
| YOLOv5s | 400 | 0.918 | 0.88 | 0.87 |
| YOLOv5m | 400 | 0.913 | 0.88 | 0.88 |
| YOLOv51 | 400 | 0.913 | 0.88 | 0.87 |

YOLOv5s is the lightest variant, providing computational advantages with higher mAP@0.5 results compared to YOLOv4-CSP and the heavier YOLOv5 variants YOLOv5m and YOLOv5l, which require more computations. YOLOv4-CSP tended to achieve higher precision but lower recall values on small datasets, while YOLOv5s tended to be more consistent. YOLOv5s is a compact model designed to improve speed, making it suitable for deployment on devices with limited computational resources. YOLOv5 has undergone numerous developments owing to its user-friendly nature, which has led to its widespread adoption in many research studies. Research on YOLOv5s includes the development of an attention mechanism for real-time fruit detection and counting [42], as well as the detection and counting of citrus fruits in plantations [43].

Based on the training experiment, the precision and recall values were both 0.88 and the F1–Score was also 0.88 for dataset 100. The precision and recall for the 100

and 400 datasets are shown in Fig. 6. The model performance results indicate that the constructed model effectively and accurately detected stomata.



Fig. 6. Evaluation metrics of stomatal detection models, (a). Dataset 100 with YOLOv5s (orange) and YOLOv4-CSP (light blue), (b). Dataset 400 with YOLOv5s (dark blue) and YOLOv4-CSP (red)

Another metric used to measure the model's accuracy is the mAP@0.5, which resulted in model accuracies of 0.923 and 0.920 for 100 and 400 values, respectively, based on the precision-recall (PR) curve shown in Fig. 7. The PR results indicate that the model can predict accurately, as both the precision and recall values remain high (above 80%).



Fig. 7. Stomata detection model accuracy (dataset 400).



Fig. 8. Stomata that overlap trichomes and guard cells appear unclear.

The validation dataset showed a stomatal detection error rate of 0.08. This was mainly due to stomata overlapping with trichomes or unclear stomatal guard cells, as shown in Fig. 8. The yellow arrows indicate stomata overlapping with trichomes, while the green arrows indicate unclear guard cells. These results in stomata cannot be detected by the model.

The blue box indicates that the confidence score for identifying the stomata should be greater than 0.50. This gives a True Positive (TP) value that is used to determine the stomatal density, as shown in Eq. (10). A confidence value of less than 0.50 will indicate that the stomata are non-stomata, as indicated by the green bounding box in Fig. 9. This results in a False Negative (FN), which will not be included in the calculation of stomatal density.



Fig. 9. The green bounding boxes for stomata with confidence scores values less than 0.50.

Another factor that affects the inaccuracy of calculations when analysing stomatal density is the occurrence of False Positives (FPs), which occur when non-stomata are mistakenly identified as stomata. The possibility of FP occurs because non-stomatal objects have similar features to stomata. Further research is needed to develop techniques to enhance the model for detecting the unique features of stomata. An example of stomatal detection at 200× magnification from a microscopy image, resulting in TP, FP, and FN, is shown in Fig. 10.

The model detects the base of the trichome as the stomata because of its similar shape to a stomatal pore, resulting in an FP that is included in the calculation of stomatal density. This phenomenon occurs in microscopy images of objects other than stomata, such as trichome objects or leaf hairs, as found in patchouli. Further research is needed to address the shortcomings of the model in recognising the distinctive characteristics of stomata. Attention techniques can be incorporated into the featureextraction part of the YOLOv5 backbone. This should enhance the accuracy and precision of the model by lowering the FPs and FNs.



Fig. 10. Stomata that produce TP (blue boxes) were detected. FP (white arrow) and FN (black arrow).

The stomatal density is determined by the number of stomata per field of view. An inaccurate reduction in the number of stomata affects the density of stomata in microscopy images. To improve the accuracy of the stomata detection model, it is necessary to employ techniques that reduce False Positives (FPs) and increase recall by minimising False Negatives (FNs), thus ensuring precise stomatal density calculations. The loss-testing function metric was utilised to evaluate the performance of the stomata detection model by comparing its results with the training data from the training datasets. The validation loss function was also utilised to assess the model's performance on the validation datasets. The results obtained during training showed a box loss of 0.031, an object loss of 0.128, and a class loss of 0. This occurred during both training and validation because only one object, the stomata, was labelled and detected. During validation, the results showed a box loss value of 0.035 and an object loss of 0.187. The loss function results indicate that the model can effectively predict stomata and is optimally suited for training and validation. Therefore, the model was neither overfit nor underfit. The loss function metrics for datasets 100 and 400 are presented in the form of graphs in Figs. 11 and 12. The loss functions for all the stomatal detection models are listed in Table III.

| Model | Dataset image | Box loss (train) | Object loss (train) | Box loss (validation) | Object loss (validation) |
|------------|---------------|---------------------|------------------------|--------------------------|-----------------------------|
| YOLOv4_CSP | 100 | 0.024 | 0.101 | 0.034 | 0.204 |
| YOLOv5s | 100 | 0.031 | 0.128 | 0.035 | 0.187 |
| YOLOv5m | 100 | 0.027 | 0.105 | 0.035 | 0.196 |
| YOLOv51 | 100 | 0.033 | 0.123 | 0.037 | 0.174 |
| YOLOv4_CSP | 400 | 0.019 | 0.094 | 0.032 | 0.235 |
| YOLOv5s | 400 | 0.027 | 0.145 | 0.032 | 0.200 |
| YOLOv5m | 400 | 0.022 | 0.101 | 0.032 | 0.216 |
| YOLOv51 | 400 | 0.026 | 0.115 | 0.034 | 0.195 |

TABLE III. LOSS FUNCTION OF THE MODEL



Fig. 11. Loss function of the model (dataset 100).



Fig. 12. Loss function of the model (dataset 400).

High stomatal density has been shown to influence vulnerability to fungal, bacterial, and pathogenic attacks, as observed in *Arabidopsis thaliana* [44]. High stomatal density determines susceptibility to bacterial infection [45]. During this study, in addition to stomatal detection and counting, a detection and counting model was developed for two varieties of patchouli. Additionally, we calculated the stomatal density using Eq. (10). We compare stomatal detections and the manual calculations for 18 test images of Patchouli varieties in Fig. 13.



Fig. 13. Comparison of the number of stomata in patchouli varieties.

To evaluate the variance between the outcomes of the manual and modelled stomatal counts, we use Eq. (11) to determine the accuracy of the stomatal counts. To determine the linear relationship between the manual calculation by the expert and the automatic calculation by the stomatal detection model, the R^2 coefficient equation for simple linear regression Eq. (12) is used.

$$Counting \ accuracy = 1 - \frac{abs \ (automatic \ count-manual \ count)}{manual \ count}$$
(11)

$$R^{2} = 1 - \frac{\sum_{i} (x_{i} - y_{i})^{2}}{\sum_{i} (x_{i} - \hat{y})^{2}}$$
(12)

In this case, x_i is the manual calculation, y_i is the automatic calculation, and \hat{y} is the average. The test dataset has a calculation accuracy of 0.97 and an R^2 value of 0.94, as shown in Fig. 14. These results demonstrate a strong positive correlation between the manual calculation results and the patchouli microscopy image dataset model. The accuracy of the calculations for high-contrast and noisy images was also high, indicating a strong positive correlation. This shows that the detection model can accurately identify stomata even when they appear blurred or when other objects, such as trichomes or epidermis, are present in the microscopy image. The complete accuracies and R^2 values are listed in Table IV.

TABLE IV. EVALUATION OF STOMATA COUNT

| # | Image | Accuracy | \mathbb{R}^2 |
|---|---------------------|----------|----------------|
| 1 | Image without noise | 99% | 0.98 |
| 2 | Image with noise | 95% | 0.94 |



Fig. 14. Comparison of the stomata calculations.

Stomata counting using the stomata detection model in noisy images, images with blurred or faint stomata, and stomata with a confidence score < 0.5, as summarized in Figs. 8 and 9, remains a challenge and an opportunity for further research. There is a need for a model that can remember stomata even when they are faint, blurred, or overlapping with background objects, such as trichomes. The results of the YOLOv5s model for stomatal detection and stomatal density calculation in noisy images are shown in Fig. 15. The figure illustrates the identification and quantification of stomata in an image containing noise, where there is an increase in the number of stomata detected in dataset 400. In general, there was no significant difference in the average number of stomata between the Sidikalang and Patchoulina2 varieties when manual and model calculations were compared, as shown in Fig. 13. Stomatal density was classified as low (<300/mm²), medium density (300–500/mm²), or high (>500/mm²). The results of the stomatal density calculations using Eq. (10) show that both varieties had low stomatal densities. This finding is consistent with previous research on the impact of high stomatal density on susceptibility to fungal, bacterial, and pathogenic attacks. Sidikalang is tolerant, while Patchoulina2 is resistant to bacterial wilt, a common disease affecting patchouli plants and its production of patchouli oil (patchoulol).









Fig. 15. Example of object detection in an image with noise, (a). Dataset 100, (b). Dataset 400.

IV. CONCLUSION

Many studies related to stomatal detection have been conducted to help researchers measure stomatal quantity and density. Each study was conducted on different leaf specimens, each of which faced difficulty in obtaining microscopy images of the stomata. This study focused on two high-yielding varieties of patchouli: Sidikalang, known for its ability to produce high-quality patchouli oil and tolerance to bacterial wilt disease, and Patchoulina2, chosen for its resistance to bacterial wilt disease, which can affect leaf production to produce high amounts of essential oils. A deep learning approach using the YOLOv5 algorithm was developed to create a stomatal object detection model, which was used to calculate the number and density of stomata in patchouli. The performance of the model, as measured by the evaluation metrics, yielded precision, recall, and F1–Scores of 88% each. The model achieved 97% accuracy for stomatal detection in the test dataset. Efforts are required to improve the ability of this model to detect stomata that overlap with trichomes, guard cells that are not visible, and blurred stomata.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Author A.Q. wrote this paper with direction, guidance, and supervision from Y.H., W.A.K., A.H.W., and S.S., who contributed to the background, literature review, methods, and model evaluation. Y.H. focused on methods, image processing, computer vision, and deep learning algorithms. W.A.K. focused on methods and deep learning algorithms. A.H.W. focused on the methods, statistical calculations, and correlations between parameters. S.S. contributed to the method for acquiring microscopic images of patchouli and the analysis of the model detection results; all authors had approved the final version.

FUNDING

This research was supported by the Directorate General of Higher Education, Research, and Technology, Ministry of Education, Culture, Research, and Technology of Indonesia, which has provided Doctoral Research Grant No. 102/E5/PG.02.00.PL/2023, 19 June 2023. Contract No.18867/IT3.D10/PT.01.02/M/T/2023.

ACKNOWLEDGEMENT

The authors also thank Dra. Endang Hadipoentyanti, MS, an expert in patchouli, for her support and direction.

REFERENCES

- [1] H. Lee, J. Lee, D. Smolensky, and S. Lee, "International Immunopharmacology Potential benefits of patchouli alcohol in prevention of human diseases: A mechanistic review," *Int Immunopharmacol*, vol. 89, no. August, p. 107056, 2020. doi: 10.1016/j.intimp.2020.107056
- [2] M. K. Swamy, S. Balasubramanya, and M. Anuradha, "In vitro multiplication of Pogostemon cablin Benth. through direct regeneration," *Afr J Biotechnol*, vol. 9, no. 14, pp. 2069–2075, 2010.
- [3] N. S. Ravindra *et al.*, "Evaluation of somaclonal variation for genetic improvement of patchouli (Pogostemon patchouli), an exclusively vegetatively propagated aromatic plant," *J. Crop Sci Biotechnol*, vol. 15, no. 1, pp. 33–39, Mar. 2012. doi: 10.1007/s12892-011-0068-5
- [4] W. Haryudin and S. Suhesti, "Morphological characteristics yield and quality of five patchouli accession numbers at three agroecological zones," *Informatika Pertanian*, vol. 23, no. 1, pp. 29–34, 2014.
- [5] J. F. Eisele, F. Fabler, P. F. Bürgel, and C. Chaban, "A rapid and simple method for microscopy-based stomata analyses," *PLoS One*, vol. 11, no. 10, October. 2016. doi: 10.1371/journal.pone.0164576
- [6] D. Borthakur, J. Saikia, and S. Roy, "Stomatal density as a selection criterion for developing tea varieties with high physiological efficiency," *Journal of Plant Physiology and Breeding*, vol. 7, no. 1, pp. 121–131, 2017.

- [7] Q. Bin Kwong *et al.*, "Automated stomata detection in oil palm with convolutional neural network," *Sci Rep*, 0123456789, pp. 1–9, 2021, doi: 10.1038/s41598-021-94705-4
- [8] A. M. Senna, M. M. F. Vieira, R. M. Machado-de-Sena, A. O. Bertolin, S. C. Nune, and M. S. Ribeiro, "Photodynamic inactivation of Candida ssp. on denture stomatitis. A clinical trial involving palatal mucosa and prosthesis disinfection," *Photodiagnosis and Photodynamic Therapy*, vol. 22, pp. 212–216, 2018. doi: 10.1016/j.pdpdt.2018.04.008
- [9] M. Zahara, S. Suwarniati, Q. Aini, and M. Muslim, "The effect of organic fertilizers on the leaf morphology and stomatal density of Pogostemon cablin Benth," *J. Nat.*, vol. 21, no. 2, pp. 52–56, 2021. doi: 10.24815/jn.v21i2.20989
- [10] D. G. Tork, N. O. Anderson, D. L. Wyse, and K. J. Betts, "Domestication of perennial flax using an ideotype garden performance," *Agronomy*, vol. 9, pp. 2–16, 2019. doi: 10.3390/agronomy9110707
- [11] S. Jian, C. Zhao, and Y. Zhao, "Based on remote sensing processing technology estimating leaves stomatal density of populus euphratica," in *Proc. 2011 IEEE International Geoscience and Remote Sensing Symposium*, 2011, pp. 547–550. doi: 10.1109/IGARSS.2011.6049186
- [12] J. Wang, H. J. Renninger, Q. Ma, and S. Jin, "StoManager1: An enhanced, automated, and high-throughput tool to measure leaf stomata and guard cell metrics using empirical and theoretical algorithms," arXiv preprint, arXiv:2304.10450, 2024.
- [13] H. Laga and D. Fleury, "Image-based plant stomata phenotyping," in Proc. 13th Int. Conf. Control. Autom. Robot. Vis., 2014, pp. 10– 12.
- [14] V. Wiley and T. Lucas, "Computer vision and image processing: A paper review," *International Journal of Artificial Intelligence Research*, vol. 2, no. 1, pp. 28–36, 2018. doi:10.29099/ijair.v2i1.42
- [15] R. Pourdarbani, S. Sabzi, V. M. G. Amicis, G. G. Mateos, J. M. M. Martínez, and A. R. Canales, "Automatic classification of chickpea varieties using computer vision techniques," *Agronomy*, vol. 9, no. 11, Oct. 2019. doi: 10.3390/agronomy9110672
- [16] K. Laabassi, M. A. Belarbi, S. Mahmoudi, S. A. Mahmoudi, and K. Ferhat, "Wheat varieties identification based on a deep learning approach," *Journal of the Saudi Society of Agricultural Sciences*, vol. 20, no. 5, pp. 281–289, Jul. 2021. doi: 10.1016/j.jssas.2021.02.008
- [17] Y. Xing, P. Lv, H. He, J. Leng, H. Yu, and X. Feng, "Traits expansion and storage of soybean phenotypic data in computer vision-based test," *Front Plant Sci.*, vol. 13, Mar. 2022. doi: 10.3389/fpls.2022.832592
- [18] S. B. Martínez, J. Simo, A. Gras, M. Mestre, R. Boque, and F. Tarres, "Combining computer vision and deep learning to classify varieties of Prunus dulcis for the nursery plant industry," J. *Chemom*, vol. 36, no. 2, Feb. 2022. doi: 10.1002/cem.3388
- [19] I. Iqbal, M. Younus, K. Walayat, M. U. Kakar, and J. Ma, "Automated multi-class classification of skin lesions through deep convolutional neural network with dermoscopic images," *Computerized Medical Imaging and Graphics*, vol. 88, Mar. 2021. doi: 10.1016/j.compmedimag.2020.101843
- [20] I. Iqbal, K. Walayat, M. U. Kakar, and J. Ma, "Automated identification of human gastrointestinal tract abnormalities based on deep convolutional neural network with endoscopic images," *Intelligent Systems with Applications*, vol. 16, Nov. 2022. doi: 10.1016/j.iswa.2022.200149
- [21] Z. Zou, K. Chen, Z. Shi, Y. Guo, and J. Ye, "Object detection in 20 years: A survey," in *Proc. The IEEE*, pp. 1–20, 2023. doi: 10.1109/jproc.2023.3238524
- [22] S. N. Sultana et al., "Optimizing the experimental method for stomata-profiling automation of soybean leaves based on deep learning," *Plants*, vol. 10, no. 12, 2021. doi: 10.3390/plants10122714
- [23] A. C. García et al., "LabelStoma: A tool for stomata detection based on the YOLO algorithm," *Comput. Electron. Agric.*, vol. 178, 2020. doi: 10.1016/j.compag.2020.105751
- [24] X. Li, S. Guo, L. Gong, and Y. Lan, "An automatic plant leaf stoma detection method based on YOLOv5," *IET Image Process.*, vol. 17, no. 1, pp. 67–76, 2023, doi: 10.1049/ipr2.12617.
- [25] E. Y. Obsie, H. Qu, Y. J. Zhang, S. Annis, and F. Drummond, "Yolov5s-CA: An improved Yolov5 based on the attention mechanism for mummy berry disease detection," *Agric.*, vol. 13, no. 1, 2022. doi: 10.3390/agriculture13010078

- [26] L. Xu et al., "Bass detection model based on improved YOLOv5 in circulating water system," PLoS One, vol. 18, pp. 1–14, 2023. doi: 10.1371/journal.pone.0283671
- [27] H. Jayakody, S. Liu, M. Whitty, and P. Petrie, "Microscope image based fully automated stomata detection and pore measurement method for grapevines," *Plant Methods*, pp. 1–12, 2017. doi:10.1186/s13007-017-0244-9
- [28] C. Dewi, R. C. Chen, Y. C. Zhuang, and H. J. Christanto, "Yolov5 series algorithm for road marking sign identification," *Big Data and Cognitive Computing*, vol. 6, no. 4, Dec. 2022. doi: 10.3390/bdcc6040149
- [29] D. F. Parkhurst, "The adaptive significance of stomatal occurrence on one or both surfaces of leaves," *J. Ecol.*, vol. 66, no. 2, p. 367, 1978. doi: 10.2307/2259142
- [30] Y. Abhilash, A. H. Madakemohekar, M. S. Reddy, P. Yaswanth, and K. Swetha, "Ideotype breeding in maize a review," *Int. J. Curr. Microbiol. Appl. Sci.*, no. 11, pp. 2603–2616, 2020.
- [31] K. C. Fetter, S. Eberhardt, R. S. Barclay, S. Wing, and S. R. Keller, "StomataCounter: A neural network for automatic stomata identification and counting," *New Phytol.*, vol. 223, no. 3, pp. 1671– 1681, 2019. doi: 10.1111/nph.15892
- [32] S. Meeus and B. J. V. Den, "From leaf to label: A robust automated workflow for stomata detection," *Ecology and Evolution*, pp. 9178– 9191, 2020. doi:10.1002/ece3.6571
- [33] A. Razzaq et al., "Stomatal state identification and classification in quinoa microscopic imprints through deep learning," Complexity, vol.1, 2021.
- [34] H. Jayakody, P. Petrie, H. J. Boer, and M. Whitty, "A generalised approach for high-throughput instance segmentation of stomata in microscope images," *Plant Methods*, vol. 17, no. 1, 2021. doi: 10.1186/s13007-021-00727-4
- [35] C. Zhu *et al.*, "A deep learning-based method for automatic assessment of stomatal index in wheat microscopic images of leaf epidermis," *Front. Plant Sci.*, vol. 12, 2021. doi: 10.3389/fpls.2021.716784
- [36] A. H. Aono *et al.* (2021). A stomata classification and detection system in microscope images of maize cultivars. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/jou rnal.pone.0258679
- [37] H. Liu, F. Sun, J. Gu, and L. Deng, "SF-YOLOv5: A lightweight small object detection algorithm based on improved feature fusion mode," *Sensors*, pp. 1–14, 2022.
- [38] X. Xu et al., "Research on machine vision and deep learning based recognition of cotton seedling aphid infestation level," *Frontiers in Plant Science*, pp. 1–14, 2023. doi:10.3389/fpls.2023.1200901
- [39] D. Forsyth, "Non-maximum-suppression," Computer (Long. Beach. Calif), vol. 47, no. 2, pp. 6–7, 2009.
- [40] J. Redmon, D. Santosh, G. Ross, and A. Farhadi, "You only look once: Unified, real-time object detection," ACM Int. Conf. Proceeding Ser., pp. 2–22, 2015. doi:10.1145/3243394.3243692
- [41] C. Y. Wang, A. Bochkovskiy, and H. Y. M. Liao, "Scaled-Yolov4: scaling cross stage partial network," in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Nashville, TN, USA, 20–25 June, 2021, pp. 13024–13033.
- [42] O. M. Lawal, S. Zhu, and K. Cheng, "An improved YOLOv5s model using feature concatenation with attention mechanism for real-time fruit detection and counting," *Front Plant Sci.*, vol. 14, 2023. doi: 10.3389/fpls.2023.1153505
- [43] Y. Lin, W. Hu, Z. Zheng, and J. Xiong, "Citrus identification and counting algorithm based on improved YOLOv5s and deepsort," *Agronomy*, vol. 13, no. 7, Jul. 2023. doi: 10.3390/agronomy13071674
- [44] L. Jordá et al., "ERECTA and BAK1 receptor like kinases interact to regulate immune responses in Arabidopsis," *Front. Plant Sci.*, vol. 7, no. June, pp. 1–15, 2016. doi: 10.3389/fpls.2016.00897
- [45] C. Tateda et al., "The host stomatal density determines resistance to septoria gentianae in Japanese gentian," Mol. Plant-Microbe Interact., vol. 32, no. 4, pp. 428–436, 2019. doi: 10.1094/MPMI-05-18-0114-R

Copyright © 2024 by the authors. This is an open access article distributed under the Creative Commons Attribution License (<u>CC BY-NC-ND 4.0</u>), which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.