RAPID ASSESSMENT OF STOMATA DENSITY USING DIGITAL MICROSCOPY AND MACHINE LEARNING APPROACH

Kunwei Sun

Bioresource Engineering

Department of Agriculture and Environment Science

McGill University, Montreal, Quebec

August 2024

A thesis submitted in partial fulfillment of the requirements of a master's degree

Abstract

The density of stomata per area in leaves of different cultivars, including beans, is vital to understanding several physiological phenomena found in growing plants. Traditional stomata detection methods (*e.g.*, the nail polish method) are labor-intensive and require significant training. A user-friendly platform was developed using a digital microscopy-based machine vision system to detect and count stomata per unit area of fresh leaves. Over 119 bean genotypes were used to compare nail polish and fresh leaf images collected from 10 random leaves representing each genotype. Our platform features a two-module model: a segmentation module using a transformer mechanism and a detection module based on YOLOv5, enhanced with a Vision Transformer (ViT) for improved accuracy. We randomly chose 246 images for the segmentation module as the dataset, utilizing 231 images for training, 10 for validation, and 17 for testing. We randomly chose 211 images for the detection module, utilizing 148 images for training, 42 for validation, and 21 for testing. Compared to different machine learning methods, the stomata detection module developed here improved upon the Mask RCNN, YOLOv5, and traditional convolution neural network (CNN) methods' performances by 16%, 11.2%, and 22.6% mAP (mean Average Precision), respectively. Additionally, our segmentation module outperformed YOLOv5 by 5% mAP. The nail polish image presents 95% mAP, and the fresh leaf microscopy image presents 93% mAP. System validation yielded a simple linear regression with a coefficient of determination (R^2) of 0.87. This platform holds promise for more robust inspection, facilitating better genotype analysis and potentially boosting bean production.

Résumé

L’amélioration de la production du haricot est cruciale pour l’approvisionnement alimentaire humain, et il est essentiel de comprendre la relation entre les génotypes de haricots et les caractéristiques des stomates. Les méthodes traditionnelles de détection des stomates, telles que la méthode du vernis à ongles, demandent beaucoup de travail et nécessitent une formation importante. Cette présentation rend compte du développement d'une plate-forme conviviale pour détecter et compter les stomates par unité de surface à l'aide d'un système de vision industrielle basé sur la microscopie numérique effectuant des mesures de feuilles fraîches. Plus de 119 génotypes de fève soya ont été utilisés pour comparer des images de vernis à ongles et de feuilles fraîches collectées pour 10 feuilles aléatoires représentant chaque génotype. Notre plateforme comprend un modèle à deux modules : un module de segmentation utilisant un mécanisme de transformateur et un module de détection basé sur YOLOv5 amélioré avec un Vision Transformer (ViT) pour une précision améliorée. Nous avons choisi au hasard 246 images pour le module de segmentation comme ensemble de données, en utilisant 231 images pour l'entraînement, 10 pour la validation et 17 pour les tests. Nous avons choisi au hasard 211 images pour le module de détection, en utilisant 148 images pour l'entraînement, 42 pour la validation et 21 pour les tests. En comparant différentes méthodes d'apprentissage automatique, la détection des stomates a présenté une amélioration de 16 % de la mAP (précision moyenne moyenne) par rapport à la méthode Mask RCNN, une amélioration de 11.2 % de la mAP par rapport à YOLOv5 et une amélioration de 22.6 % de la mAP en relation avec le réseau neuronal à

convolution traditionnel (CNN). De plus, notre module de segmentation a surpassé YOLOv5 de 5 % mAP. En effet, l'image du vernis à ongles présente 95% de mAP et l'image Fresh leaf microscopy présente 93% de mAP. La validation du système a donné une régression linéaire simple avec un coefficient de détermination (R^2) égal à 0.87. Cette plateforme est prometteuse pour une inspection robuste, facilitant une meilleure analyse du génotype et potentiellement stimulant la production de fèves soya.

Authorship and Manuscript

The contributions of the individuals involved are as follows:

1) Kunwei Sun: Lead in the design and execution of experiments, along with collection, compilation, analysis, and interpretation of the data communicated in this thesis

2) Prof. Viacheslav Adamchuk: Research supervisor; provided invaluable guidance throughout the experimental design process, oversaw the execution of experiments, and contributed to the comprehensive review of the thesis.

3) Prof. Valerio Hoyos-Villegas: Offered insightful guidance on experimental design and data collection.

4) Henry Cordoba Novoa: Offered optical microscope and Dino-lite guidance.

5) Ryan Smith and Neha Paserkar: Contributed to data collection.

6) Kyle Geddes and John Lan: Contributed to the data collection process.

Acknowledgments

As I complete my master's degree at McGill University, I am grateful for the remarkable individuals who have shaped my academic and personal journey. Their unwavering support, guidance, and friendship have been instrumental in my success, and I would like to express my deepest appreciation to every one of them.

First and foremost, I extend my heartfelt gratitude to my exceptional supervisor, Prof. Viacheslav Adamchuk. You have provided me with the opportunity to study at McGill and do research in machine vision and supported me in my studies and daily life. Your unwavering support, mentorship, and sage advice have been invaluable in shaping my academic and personal growth. I am deeply grateful for your trust in my abilities and dedication to my success.

I would also like to sincerely thank my committee member, Dr. Valerio Hoyos-Villegas. Your expertise, valuable insights, and commitment to academic excellence have greatly enriched my research. I am truly grateful for your involvement in my academic journey.

A special thank you goes to Herny Cordoba Novoa for your support and assistance whenever I encountered challenges during my experiments and for teaching me the nail polish method. Your guidance, expertise, and willingness to lend a helping hand have been invaluable in overcoming obstacles and ensuring the smooth progress of my research. I am deeply grateful for your contributions to my academic journey.

With respect to collecting data samples, I would like to thank Ryan Smith and Neha Paserkar; I extend my sincere appreciation for your collaboration, dedication, and expertise. Our collective efforts have made significant contributions to our field, and I am grateful for the knowledge and insights I gained through our collaboration. Your commitment to excellence and valuable contributions have been instrumental in the success of our work.

I would also like to express my gratitude to my best friends Kyle Geddes and John Lan. Your encouragement and friendship have been a source of strength throughout my journey, helping me process the samples before they were too dry to be useful. Your presence brought joy, laughter, and a sense of belonging to my life. I am truly grateful for the memories we created and for your support.

Furthermore, I would like to thank my parents for their understanding throughout this journey. Thank you for tolerating the physical distance, for always being there for me, and for providing emotional support and guidance.

Finally, I would like to thank Ruilin Shan and Joel Harms for supporting the project and having a grand vision for what is possible, especially in data visibility. I eagerly embrace the future with a grateful heart and a smile.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

FIGURE LIST

TABLE LIST

Abbreviations

ViT: Vision of Transformer

mAP: mean Average Precision

ANOVA: Analysis of Variance

S&D model: Segmentation and Detection (S&D) model

SD: Standard Deviation

SE: Standard Error

R^2: the coefficient of determination

OPT: Optical Microscope Test

NPT: Nail Polish Test

FLT: Fresh Leaf Test

1. INTRODUCTION

1.1 Thesis motivation

Having been cultivated for centuries, the bean is an annual legume that originated in China (Intelligence, 2017). World bean production is predicted to increase by 2.2% annually to 371.3 Tg (terrogram=$10^{12}g$) by 2030. (Masuda & Goldsmith, 2009). To achieve high production, many different genotypes have been developed (Fehr et al., 2003). Various studies have explored the selection and characteristics of bean genotypes for production. Finoto identified specific genotypes with high oil and protein content and those with the highest plant and first pod heights. (Finoto et al., 2021). Palmer highlighted the importance of bean accessions in bean improvement programs, emphasizing the role of sexual hybridization and selection in cultivar development (Palmer & Hymowitz, 2016). Collectively, these studies contributed to the understanding and improvement of bean genotypes for production.

Stomata play an important role in plant growth and influence plant production. The closer to the top leaf, the greater the number of stomata per unit area (Lleras, 1977). Also, during the mature period, signaling from mature to developing leaves predetermines the potential of the developing leaf to maximize its photosynthetic potential during the early stages of development in the enclosed bud (Woodward et al., 2002). The placement of stomata may be coordinated with internal features in some organs (Nadeau & Sack, 2002).

Stomatal formation and patterning are regulated by the frequency and placement of asymmetric divisions (Le et al., 2014).

This thesis focused on investigating the number of stomata from different genotypes in beans to improve their production and quality and to reduce loss during transportation.

1.2 Research problem

Stomata detection has been widely studied, mainly by the leaf surface cast method. (Brewer, 1992). However, as this method requires extra training and experienced operators, developing new techniques to quantify stomata from beans is essential. In this project, we sought to develop an innovative process that required less training and fewer human resources yet was robust enough to be utilized in the real world. This would allow us to gain a more comprehensive understanding of how genotype influences stomata structure and numbers. The project will help establish and improve the production of beans destined for both wholesale and retail markets; it has the potential to greatly benefit the Canadian agricultural sector by maintaining the visual and nutritional quality of beans. The bean industry faces challenges from climate change, with a warmer climate projected to benefit production but also bring increased precipitation deficits, particularly in the Canadian Prairies (Qian et al., 2022). New bean genotypes with greater or lesser stomatal numbers and altered patterns can accommodate these significant changes.

1.3 Hypotheses

The fresh leaf microscopy method yields the same results as the nail polish method. Our model demonstrates a greater ability to calculate stomatal numbers than other models. Additionally, our platform is easily portable and user-friendly.

1.4 Advantages of the S&D (Segmentation & Detection) model

Our method achieves high precision in the completion of its tasks, with the main advantages of detecting stomata, segmenting the blurred portion of stomata images, offering robustness under multiple conditions, and being easy to use in a wide range of settings.

Thus, our approach does not require specific and expensive equipment/hardware and extra training, making it a suitable alternative to assist producers in making decisions about legume crops. Further advantages of our methods include monitoring microscope areas and automated calculation, both of which provide valuable information so that the producer can improve production planning to minimize costs and maximize productivity.

1.5 Project objective

The objective of this research was to develop and evaluate a new method, the fresh leaf microscopy method, in comparison with the traditional nail polish method (traditional method) in terms of stomata numbers detected per unit area (μm^2) and determine if significant differences exist between the two methods.

2. LITERATURE REVIEW

2.1 Stomata detection methods

Stomata detection and analysis are essential for understanding plant physiology and health, which is critical in plant biology and environmental research. Over the years, various methods have been developed to automate this process, ranging from traditional image-processing techniques to advanced machine-learning approaches. Early efforts primarily utilized morphological operations and template matching, which, while foundational, relied heavily on manual counting of stomatal pores per unit area. This manual approach, although informative, has been critiqued for its inefficiency as well as its potential for inaccuracy (Bhaiswar et al., 2016).

Infrared thermography emerged as an alternative method, effectively monitoring stomatal closure in grapevines (Jones et al., 2002). However, this technique has its challenges. Calibration drift in low-cost thermal imagers necessitates corrective procedures, and the presence of non-leaf material within the canopy complicates temperature analysis, highlighting the need for reference surfaces. Additionally, recent explorations into using shaded leaves rather than sunlit ones for detecting stomatal closure have introduced new approaches with advantages and limitations. For instance, the temperature of reference surfaces within the canopy can be influenced by the water status of the canopy itself, thus requiring further investigation.

To address the labor-intensive nature of manual stomatal detection, semi-automatic methods, such as those based on CARTA, have been proposed, particularly for specific plant species. (Higaki et al., 2014). These advancements have enabled higher throughput in stomatal analysis, resulting in accelerating research in plant biology and environmental science. Nevertheless, the need for further automation led to the development of fully automatic methods grounded in machine vision. These methods focus on detecting and counting stomata. (Duarte et al., 2017), with some approaches extending to measure stomatal features like pore dimensions (Jayakody et al., 2017b).

Despite these advancements, challenges remain. For instance, Jayakody's methods are susceptible to the appearance of stomata in training images, which limits their general applicability. Their techniques primarily generate bounding boxes around detected stomata, necessitating additional image-processing steps to analyze stomatal morphology fully. To address the practical needs of field research, mobile applications have been developed to enable real-time detection of stomata in situ (Liu et al., 2016). These mobile solutions represent a significant step forward, though continuous refinement and validation are required to ensure their reliability across diverse environmental conditions.

Recently, with the development of deep learning, more and more methods based on deep learning have been published. Several studies have developed convolutional neural network models to accurately identify and count stomata across various plant species. Some methods achieve precision and recall rates above 90% (Bhugra et al., 2018; Ren et al., 2021) and can

measure stomatal traits such as pore length, width, and area with more than 90% accuracy (Song et al., 2020). Researchers have also developed user-friendly tools like Deep Stomata (Toda et al., 2018) and StomataCounter (Karl C. Fetter et al., 2019) to make these techniques accessible to plant biologists. These automated approaches significantly increase the speed and throughput of stomatal analysis compared to manual methods, enabling large-scale studies of plant responses to environmental stresses (Liang et al., 2021; Millstead et al., 2020). Research by Jayakody et al. and Meeus utilized Cascade Object Detection (COD) algorithms with Histogram of Oriented Gradients (HOG) and Haar-like features to detect stomata (Jayakody et al., 2020; Meeus et al., 2019). However, this method is only partially generalizable to some plant species, and the performance of unseen test data was lower than that of training and validation sets.

2.2 Transformer Mechanism

Transformers have significantly revolutionized machine learning, particularly in natural language processing (NLP) tasks (Ghojogh & Ghodsi, 2020). The pioneering Transformer model was first introduced by Vaswani et al. (2017), marking a shift in how sequence data is processed. At the core of the Transformer's architecture is the self-attention mechanism. This breakthrough enabled the model to weigh the relevance of each word in a sequence relative to others, effectively capturing intricate contextual dependencies. This approach contrasts sharply with traditional recurrent neural networks (RNNs), which process inputs sequentially and thus, struggle with long-range dependencies. By processing entire

sequences in parallel, transformers leverage positional encoding to preserve word order information, which is crucial for understanding context. Multi-head attention mechanisms within the architecture allow the model to capture diverse relationships between words while feed-forward neural networks further refine these representations. Enhancements like residual connections and layer normalization contribute to the model's robust learning capabilities, enabling efficient training and the construction of deep, hierarchical representations. As a result, transformers have become a powerful tool for a broad range of applications, extending beyond NLP to areas such as vision and video processing (Zhang et al., 2023). Recent research has explored their ability to mimic complex cognitive functions, such as frontostriatal gating in human working memory (Traylor et al., 2024) and depth-bounded symbolic reasoning (Brinkmann et al., 2024). Moreover, ongoing advancements include competitive ensembles of independent mechanisms (Lamb et al., 2021) and approaches to mitigate computational complexity (Chen et al., 2024). The versatility and efficacy of Transformers led to their successful application in various domains, including text summarization (Sanjabi, 2018), emotion detection, and chemical image recognition(Zhang et al., 2023), with models like BERT and GPT further pushing the boundaries of what these architectures can achieve (Gong, 2022). Despite these advancements, research continues to focus on enhancing the efficiency and performance of Transformers, ensuring their ongoing evolution in machine learning.

2.3 Different kinds of detection models in detecting stomata

A well-known and traditional technique for detecting stomata in plants is *in situ* measurement using a microscope (Collins, 2007). In this method, a tool called ImageJ is applied to annotate the stomata from microscope images and then calculate them. Researchers must manually see and annotate all the images, which could be more effective and accurate.

Several kinds of detection models have been utilized in stomata detection: (*i*) traditional methods that focus more on machine vision and signal processing, or (ii) machine learning methods that focus more on utilizing machine learning methods for detection.

Traditional methods focus on developing automated methods for detecting and analyzing stomata in plant microscopy images. Oliveira presented a technique based on morphological operations (Oliveira et al., 2014). In this method, they built a platform that automatically detects and counts stomata in microscopic images of plant epidermis. To detect these automatically, they used morphological operations. However, this method requires manual counting of the number of stomata, which is difficult and time-consuming, especially when many samples must be analyzed. Moreover, this means that the process depends on expert experience.

A method based on template matching developed by Laga et al. (2014) (Laga et al., 2014), requires individuals to use template matching to detect individual stomata cells in

microscopic images of plant leaves and then performs local analysis within the detected stomata regions to measure morphological and structural features (e.g., stomata opening length and width and guard cell size). Although this method detects stomata successfully, it needs to detect the contrast between the stomata cell region and the surrounding background, which may limit the accuracy of stomata detection and measurement. Advanced image processing techniques like wavelet spot detection (Duarte et al., 2017) can detect and count stomata using wavelet analysis and spot detection in the CIELab color space, then segment stomata images using the Watershed Transform, with the initially detected spots as markers. However, this approach proved time-intensive and could have shown better results in complex background stomata images.

A method based on level set segmentation proposed by Li et al. (2019) (Li et al., 2019), used level set theory and image processing technology, which does not require prior information about the stomata or plant type. Then, they accurately measured morphological features like major/minor axes, area, eccentricity, and opening degree. Moreover, it outperformed existing threshold and skeletonization-based methods in terms of versatility and accuracy. However, this method could have been better at analyzing tilted stomata images.

Recently, machine-learning-based methods have been proposed. One of the most popular methods is based on a convolutional neural network (CNN). A convolutional neural network model (Mask R-CNN) showing potential in instance segmentation and detection

could automatically detect and measure stomatal pore parameters in microscope images of plant leaves (Song et al., 2020). They used a ResNet50 to extract the feature and added a pyramid network to enhance the result which a region proposal module was implemented to suggest candidate regions, while the RoIAlign operation was utilized to normalize region sizes and facilitate segmentation mask generation. However, this method showed poor results for small stomatal apertures, blurry images, impurities in pores, and it did not correctly identify non-stomatal pores. This method requires a high image quality and could be more robust. A simplified method for automated stomata detection and analysis using an imaging technique and a convolutional neural network (Millstead et al., 2020) involved developing an end-to-end solution for computerized stomata analysis, using a CNN for stomata detection and introducing a novel method for estimating stomatal pore boundaries and for calculating pore areas. However, this model's efficiency depends on sample collection and imaging techniques. Also, the CNN-based stomata detection approach produced some false positives that could be due to limitations in the sampling technique or the high level of detail in the slide scanner images. Accordingly, this method may require further training samples to eliminate these. Moreover, the pore estimation algorithm had some incorrect estimations, with 26.27% (409 out of 1,557) images having the wrong result. A new clearing method and a deep learning-based detection model for automated stomatal analysis of bean leaves (Sultana et al., 2021b), added the essential step of incubating leaf samples in absolute ethanol for 1 hour as a fixation and dehydration step, then washing the samples with cold tap water. Finally, they transferred the samples into a clearing agent

consisting of a 1:1 (v/v) mixture of 95% ethanol and 6-14% sodium hypochlorite (NaOCl) for 1 hour to make the leaves transparent. These methods offer better results when compared with other methods. Still, the limitation of this method is the high quality or high resolution of the stomatal images needed to overcome detection failures. If the experimental setup changes, this stomatal identification model can be quickly retrained with little labeling effort. Moreover, the study used only one bean cultivar (Cheongja 3) to develop the new clearing method. In summary, further research is needed to apply the new clearing method to study stomatal traits like size, space, index, and other plant species beyond beans. An improved YOLO v5 deep learning model with an attention mechanism that achieves high precision and recall for automated detection of plant stomata was developed by Zhang et al. (2021) (Zhang et al., 2021). The YOLO v5 object detection model was used to introduce an attention mechanism (SE module) to the backbone network to improve precision and recall. This improved the loss function to avoid problems with selecting the best prediction box. They evaluated the improved model on a corn (*Zea mays* L.) leaf stomata dataset, achieving high precision (94.8%) and recall (98.7%); however, this model did not produce good results when applied to our dataset.

2.4 ViT Mechanism

ViT mechanism represents a significant shift in the approach to visual recognition tasks by adapting the Transformer architecture, originally designed for natural language processing, to process image data. ViT treats an image as a sequence of fixed-size patches,

similar to the way words are treated in text sequences. These image patches are flattened and linearly embedded before being fed into a standard Transformer encoder. Positional embeddings are added to retain spatial information within the patches. Unlike convolutional neural networks (CNNs), which apply convolutional filters to extract local features, ViT leverages the self-attention mechanism to model global relationships between patches, allowing for the capture of both local and global dependencies. This approach has been shown to outperform traditional CNNs on large-scale datasets, especially when pre-trained on extensive datasets such as JFT-300M. The simplicity of the ViT architecture, combined with its ability to leverage Transformer-based pretraining, has spurred a wave of research exploring the use of Transformers in various computer vision tasks, leading to the development of numerous ViT variants aimed at improving performance on smaller datasets, enhancing efficiency, and extending its applicability across different vision-based challenges (Dosovitskiy et al., 2020).

Nowadays, ViT has been recognized as a revolution in computer vision tasks by leveraging attention mechanisms to capture long-range dependencies in images (Fu, 2022). The core components of ViTs include patch division, token selection, position encoding, and attention calculation (Zhou et al., 2024). Recent research has focused on enhancing these mechanisms to improve performance and efficiency. For instance, dynamic mapping re-attention mechanisms have been proposed to assign attention weights adaptively (Labbaf Khaniki et al., 2023), while shift operations have been explored as simple alternatives to attention (Wang et al., 2022). ViTs have demonstrated remarkable robustness to occlusions,

perturbations, and domain shifts (Naseer et al., 2021). Visualization tools like EL-VIT have been developed to aid in understanding ViT operations (Zhou et al., 2023). Ongoing research aims to further enhance ViT efficiency through redesigned attention mechanisms (Heidari et al., 2024), highlighting the continued evolution of this powerful architecture in computer vision applications.

2.5 Bean

Bean, a crucial crop with significant economic and nutritional value as a major source of protein and oil, has become one of the most important crops worldwide (Specht et al., 2014). Bean products are excellent replacements for meat and dairy products, specifically catering to vegetarians (Dwevedi & Kayastha, 2011). While its production is expected to increase due to rising demand (Hartman et al., 2011), challenges to its production include biotic and abiotic constraints and the need for improved yield (Board & Kahlon, 2011). Moreover, the use of genetically modified beans is a contentious issue (Dwevedi & Kayastha, 2011). While the US, Brazil, and Argentina are the leading producers (Taheripour & Tyner, 2018), there is the potential for increased production in other countries (Dourado et al., 2011).

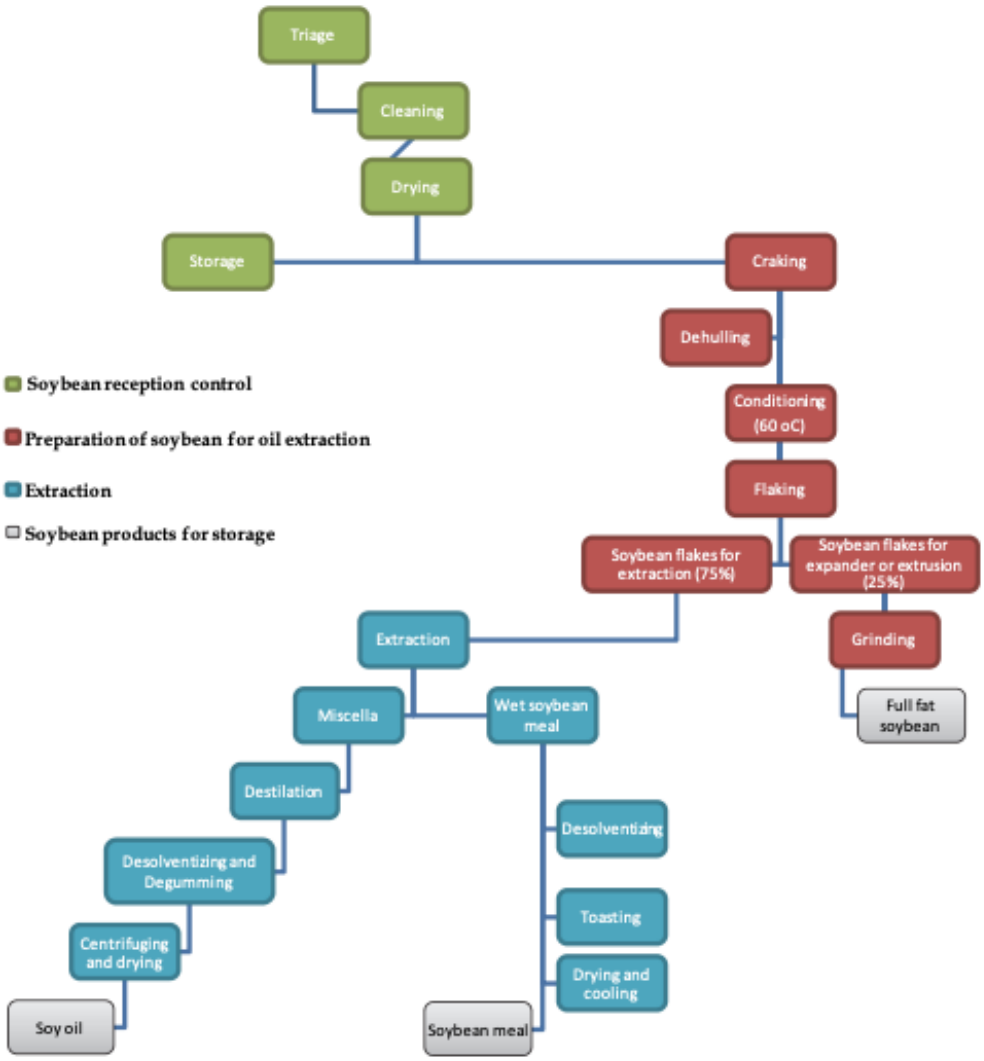

Figure 1. Flowchart of bean processing

In Canada, bean production has seen significant growth over the past few decades, particularly in Ontario, with a concomitant focus on developing short-season cultivars (Cober & Voldeng, 2012). However, the industry faces challenges from climate change, with a warmer climate projected to benefit production and increase precipitation deficits, particularly in the Canadian Prairies (Qian et al., 2022). The need for heat- and drought-tolerant cultivars is required to adapt to these changes. Additionally, the industry has to

contend with viruses, such as the first report of bean pod mottle virus (*Comovirus* sp.) in Canadian bean plants (Michelutti et al., 2002). Despite these challenges, the industry continues to innovate, with ongoing efforts to improve soybean yield through breeding programs (Yoosefzadeh-Najafabadi & Rajcan, 2022).

2.6 Different genotypes and their influence on beans

With the development of gene editing technology, the genetic map of beans has become increasingly clear, with many more traits being discovered.

A range of studies have highlighted the significant influence of different genotypes on the performance of bean plants. Studies have found that common bean (*Phaseolus vulgaris* L.) genotypes respond differently to soil fertility and plant density (Haag et al., 1978), and research continues on how genotype influences production and quality (Hampton et al., 1997). Recent research has emphasized the role of genotypes in determining the iron and zinc content of common beans and their adaptability to different environments. (Nwadike et al., 2014; Philipo et al., 2020). However, this research cannot be extended because of differences in soil quality in different regions. Also, some research focused on the faba bean (*Vicia faba* L.) and identified genotypes more tolerant to drought stress and stability across different agroecological zones (Siddiqui et al., 2015). What's more, the research across three different agroecologies in Tanzania shows the influence of environment, genotype, and the genotype-environment interaction on common bean seed yield stability and adaptability of common bean landraces, lines, and improved varieties was

investigated (Philipo et al., 2021); However, this research was only representative of Tanzania which means the study has geographical limitations. Similarly, the interaction between genotypes, *Rhizobium* strains, and the environment was found to significantly affect common bean nodulation and productivity and its response to drought stress (Argaw & Muleta, 2017; Dastneshan et al., 2019). These studies collectively underscore the importance of genotype selection in bean cultivation.

There are also a range of studies that have explored the genetic diversity and traits of common bean genotypes. Ekincialp and Şensoy (2018) found high genetic diversity among common bean genotypes, with distinct clusters based on origin, growth, and seed color; however, they did not present data on possible cross-pollination between genotypes due to the presence of wild bees in the study area. Cross-pollination led to some genotypes not clustering with the main Andean and Mesoamerican gene pools. Meanwhile, Nkhata et al. (2021) identified markers associated with bean fly resistance and other agro-morphological traits, while Basavaraja et al. (2020) highlighted the potential of specific genotypes for breeding programs. Rooting genotype was also emphasized in determining growth and yield (Wagara et al., 2011; White & Castillo, 1992) with Wagara et al. (2011) also identified potential sources of resistance to angular leaf spot. However, the genotypes tested were not resistant to all of the pathogens, indicating a high level of complexity in the pathogen population, and the resistant or moderately resistant genotypes were small-seeded types, which are less popular commercially. Additionally, Yan explored the impact of soil conditions on bean genotypes and discovered differences in phosphorus (P) efficiency (Yan

et al., 1995). However, Yan's research did not explore the mechanisms of phosphorus (P) acquisition related to root traits, nor did it address the efficiency of P use, indicating a need for further studies in these areas. While Afzal identified salt-tolerant genotypes (Afzal et al., 2022), lack of a standard growing environment makes replicability and consistent results across laboratories challenging. Previous studies have treated salinity tolerance as a single trait and used visual scoring, even though it is a polygenic trait with variability in agricultural soil can result in difficulty in performing field evaluations, increasing the coefficient of variation and leading breeders away from their objectives. Fageria et al. (2012) further highlighted the importance of soil acidity tolerance in bean genotypes (Fageria et al., 2012). These studies underscore the significance of genetic diversity and specific traits in common bean genotypes.

2.7 Environment influences the expression of the bean's genotype

The expression of bean genotype (*i.e.*, phenotype) is significantly influenced by the environment, particularly in yield and yield quality (Sözen et al., 2017). This interaction between genotype and environment can overestimate the influence of genetic parameters, affecting the selection of genotypes (Coimbra et al., 1999). However, Coimbra did not consider the interaction between genotype and environment, so it overestimated the prediction of genetic parameters like variance and heritability. Accordingly, direct selection led to superior genetic gains compared to indirect selection methods. The genotype-environment interaction was inconsistent across different environment pairs and the

genotype-environment interaction was highest in the first environment, which limited the ability to extrapolate findings. The environment plays a major role in expressing various characteristics, including growth, yield, and yield components (Nwadike et al., 2014; Papastylianou et al., 2021). However, Nwadike ignored the need for further multi-location trials over multiple seasons to select the desired genotype properly. However, the interaction between genotype and environment can be small, making it easier to select superior genotypes (Pereira et al., 2015). The effectiveness of *Rhizobium* inoculation and the productivity of common bean genotypes are also influenced by this interaction (Argaw & Muleta, 2017). The *Rhizobium* isolates tested did not perform consistently across all locations and genotypes, indicating a limitation in the generalizability of the results. In addition, the results suggest the need to develop specific *Rhizobium* strains for different locations, implying that the tested isolates were insufficient for common bean production in the study locations. Furthermore, the environment can affect common bean genotypes' iron and zinc concentrations (Nchimbi–Msolla & Tryphone, 2010). Conducted at only two locations, the study had limitations: a randomized complete block design with 3 replications was used for 20 genotypes, and data analysis was limited to ANOVA, Duncan's multiple range test, and correlation techniques. Lastly, while Cirimwami et al. (2015) noted that the environment influenced the adaptability of biofortified bean genotypes, they ignored climatic, soil, and technical constraints that lead to low and variable bean production and significant genotype-environment interactions that reduce the effectiveness of selecting the best-performing genotypes.

2.8 Structure of stomata

Stomata are essential structures in land plants, consisting of guard cells surrounding a pore, that regulate gas exchange and water loss (Merced & Renzaglia, 2017; Willmer, 1993). In bryophytes, stomata are found only on sporangia, while in vascular plants, they occur on leaves and stems (Merced & Renzaglia, 2017). Grass stomata have a unique four-celled structure with dumbbell-shaped guard cells flanked by subsidiary cells, allowing rapid responses to environmental cues (Cai et al., 2017). The development and patterning of stomata involve complex molecular mechanisms, including transcriptional networks and peptide signaling (McKown & Bergmann, 2018). The mechanical properties of guard cell walls, particularly their anisotropy, play a crucial role in stomatal opening (Marom et al., 2017). Stomatal structure and function have evolved across plant lineages, from mosses to advanced grasses, with variations in anatomy, ultrastructure, and physiological responses (Lucas & Renzaglia, 2002; Serna & Fenoll, 2000; Willmer, 1993).

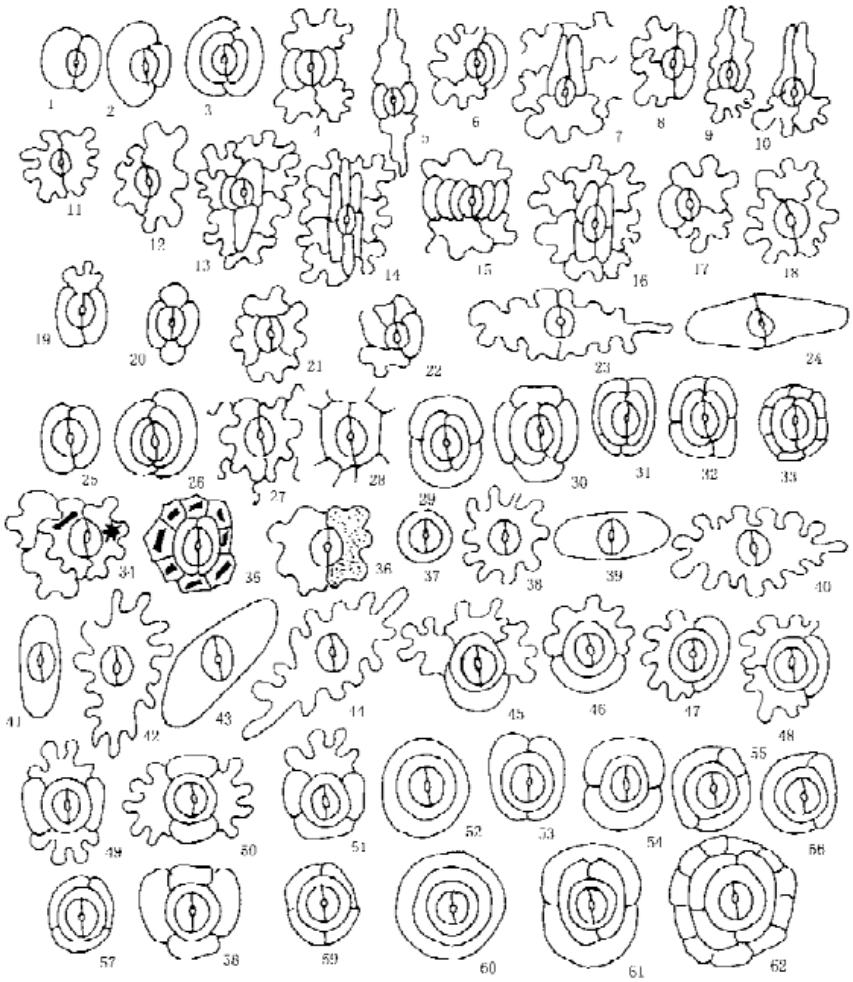

Figure 2. Different type of stomata (Prabhakar, 2004)

Diagrammatic representation of stomatal types based on subsidiary cell arrangement. Numbers 1–24 depict variations of paracytic stomata with distinct guard cells and subsidiary cells. Numbers 25–26 represent stomata with distinct subsidiaries, while 27–28 lack clear subsidiary differentiation. Remaining numbers illustrate pericytic stomata with various orientations and arrangements of subsidiary cells (e.g., transverse, parallel, or oblique).

2.9 Function of stomata

Stomata are epithelial openings in plant leaves and stems, which are structures unique to the plant epidermis. They facilitate the exchange of gases, allowing for the entry of CO_2 and O_2, which are essential for photosynthesis and respiration (Levitt, 1976; Roberts, 1990). Stomata are generally regulated by movement through the opening and closing action of guard cells and play an important role in plant physiology (Willmer & Fricker, 1996). It is the main outlet for water vapor to be discharged from the plant body to the outside during transpiration. It is the channels for photosynthesis and respiration to exchange gases with the outside world. What's more, stomata also regulate water loss through transpiration, closing to conserve water when necessary (Levitt, 1976; Roberts, 1990). The control of stomatal function is influenced by guard cell metabolism and osmoregulatory pathways (Lawson & Matthews, 2020). There remains a need for better understanding of the coordination between stomatal function, guard cell metabolism, and mesophyll photosynthesis. Thus, future work to produce new transgenic and nontransgenic lines with alterations in key genes involved in leaf osmoregulation is needed. The largely unknown role of guard cell chloroplasts and photosynthesis in stomatal function requires further research.

2.10 Mechanics of stomata opening and closure

The opening and closing of stomata in plants is controlled by guard cells. Each opening and closing of the stomata affect plant transpiration, photosynthesis, and respiration. When the stomata are open, the plant's transpiration increases, which, by cooling the leaf, protects

the plant from the sun's heat. It also expels CO_2, facilitating the plant's carbon assimilation and photosynthesis. (Levitt, 1976; Roberts, 1990). While most plants' stomata have two guard cells, grasses have a unique four-celled "graminoid" morphology that allows for faster stomatal movements (Nunes et al., 2019). The stomatal aperture is controlled by changes in guard cell osmotic pressure and turgor (Araújo et al., 2011). However, the molecular hierarchy and signal transduction pathways involved in stomatal movement are not fully understood and require further experimentation and analysis. Additional components of the signal transduction pathways and effectors involved in stomatal responses must be identified through further genetic, molecular, and biochemical analyses. A more comprehensive understanding of the biochemical mechanisms and environmental cues underlying stomatal movement is needed. Environmental factors like light, CO_2 concentration, and plant hormones influence stomatal behavior (Lawson & Matthews, 2020). Stomatal function, guard cell metabolism, and mesophyll photosynthesis still need further understanding. To do so, investigators need to generate new transgenic and non-transgenic lines with alterations in key genes involved in osmoregulation. These different lines allow researchers to investigate the specific role of genes involved in osmoregulation and how they impact guard cell function. The largely unknown role of guard cell chloroplasts and photosynthesis in stomatal function requires further research. Climate change plays an important role in stomata function and is expected to impact stomatal development and function in the future, with significant consequences for global water and carbon cycles (Matthews & Lawson, 2019).

Recent research identified key genetic and molecular mechanisms in stomatal development (Endo & Torii, 2019). Genetic modification of plants is restricted in many countries due to concerns about environmental disturbance. Some plant species are difficult to genetically engineer given their resistance to *Agrobacterium*, low regeneration capacity, and incomplete genome information. Exogenous manipulation of stomatal development using peptides is limited by the high cost of peptide synthesis, preventing large-scale agricultural applications. Understanding these processes may improve crop water use efficiency and productivity, which is crucial for sustainable agriculture in the face of changing environmental conditions (Endo & Torii, 2019; Nunes et al., 2019).

2.11 Regression analysis model

Regression analysis is a statistical method to model relationships between variables and make predictions (Yu et al., 2019). The regression model can be divided into general and Ridge models. The general regression model consists of systematic and irregular components, with the former explained by predictor variables and the latter representing unexplained variations (Glaser, 2004). The general regression model encompasses various models, including linear, polynomial, logistic, and Poisson regression, each suited for different data types and research questions (Gupta et al., 2017). However, regression models can suffer from underfitting and overfitting when dealing with complex datasets, which means Ridge and Lasso regression are considered the best models to address these issues. In contrast, Ridge regression is practical when data suffers from multicollinearity (high

correlation between independent variables) (Gupta et al., 2017). Lasso regression differs from Ridge regression by using absolute values instead of squares in the penalty function. Ridge and Lasso models are more widespread than traditional ones (Constantin, 2017). Regression analysis is widely applied in empirical research, particularly in actuarial science, finance, and higher education (Djehiche, 2011). To improve a model's prediction reliability, one of the most common methods is to use quantitative data and validate models according to least squares method assumptions (Bethea & Rhinehart, 2019).

Literature Summary Statement

The measurement of stomatal density has the potential to become a vital tool in assessing soybean production in agricultural fields, as different soybean genotypes exhibit varying stomatal densities. The fresh leaf microscope method allows for direct, real-time measurement of stomatal density, offering a novel approach that has not been extensively explored in the literature. While numerous studies have investigated stomatal detection, few have specifically addressed soybean genotypes, color images, or fresh leaf microscopy. Existing methods tend to focus on other plant species, making them less applicable and transferable to soybean datasets. In situ measurements of stomatal density are critical for improving our understanding of plant physiology in real-world conditions, and our fresh leaf microscopy method represents an important step in this direction. To the best of our knowledge, no previous research has reported similar methodologies or results specific to fresh leaf stomatal measurements in soybeans.

3. MATERIALS AND METHODS

3.1 Dataset and data pre-processing

The leaves of different bean genotypes were sampled from December 2023 to June 2024 from plants grown in the Raymond Greenhouse and Emile A. Lods Agronomy Research Centre, McGill University, Montreal, Canada (N 45° 25' 34", W 73° 55' 43"). Ten leaf samples were taken for each genotype, and each sample was divided into two symmetrical parts, one for the nail polish method and the other for the fresh leaf microscopy method. Each sample was repeated on three occasions.

Soybean was the variety of bean used in this study, and it has different genotypes and shows different traits. The light duration is 16 hours, fertilizer, and water once each month. We used the first unifoliate leaves for the experiment because these leaves are the newest and freshest ones. In the phenological stage of growth, the leaves are taken from the topmost fully expanded leaves. The sampled area is located on the leaves' lower surface.

3.2 Data augmentation

Data augmentation is a worldwide method in computer vision that aims to increase the number of images in a dataset. We processed the data using numerous data augmentation methods (e.g., rotation, shear, horizontal and vertical).

3.3 Nail polish method

The cells in plant leaves are not distributed in a single layer but in overlapping layers (Ciha & Brun, 1975). Accordingly, it is hard to observe all layers if one is using an optical microscope (Larkin et al., 1997). Also, some results suggest that overlapping layers will cause harm to the optical microscope (Kalve et al., 2014).

The nail polish method (Wu & Zhao, 2017), aims to fix the status of stomatal guard cells instantly and provides clear, stable, and almost permanent slides of epidermal impressions for measurement of stomatal aperture (Figure. 3). Compared with the leaf surface cast method (Brewer, 1992), this method is cheaper, requires less experience, and is easier to use in research. The nail polish approach does not require specific equipment or countless repetitions to collect stomata image data. It can be considered an alternative to microscopic data-collecting methods.

Nail Polish method includes the following key steps:

(i) Apply nail polish (Essie Clear Nail Polish, Longwear Gel Nail Polish Gel Couture, Essie, Clear) to the part of the leaf surface to be sampled.

(ii) Wait for the nail polish to dry and use scotch tape to sample. This step separates the overlapping layers of cells into a single layer.

(iii) Transfer the sample on the scotch tape (Amazon Basics Matte Finish Tape, BOPP film, Amazon Basics) to the slide.

(iv) Mark the sample number.

(v) Observing and counting by optical microscopy and machine learning model

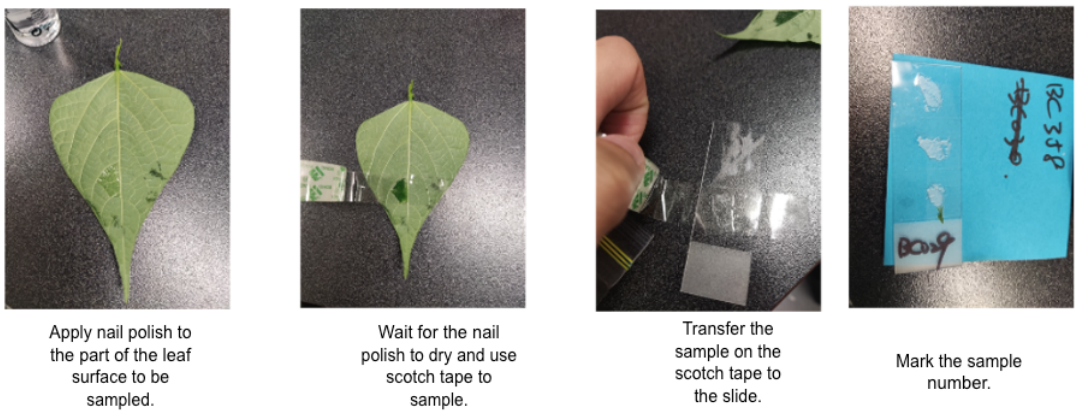

Figure 3. Steps of Nail polish method

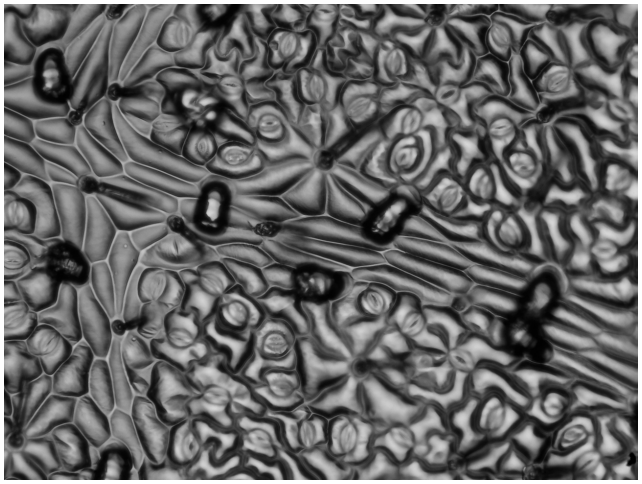

Figure 4. Nail polish image

3.5 Fresh leaf microscopy method

While chemicals emanating from nail polish can be harmful to some people, using the fresh leaf microscopy method, one need only cut leaves into slides and utilize a Dino-lite

microscope to observe (Figure. 5). Thus, our fresh leaf microscopy method approach does not require extra training and protection for the experimenter to collect stomata image data, making it a good alternative to improve data collection.

Our method uses a digital microscope, such as Dino-Lite (AnMo Electronics Corporation , AM4113ZT, New Taipei City, Taipei, Taiwan) (*Amazon.ca: Electronics*) The AM4113ZT model offers adjustable magnification levels of up to 220x and is equipped with integrated LED lighting, providing clear and detailed images for both live observation and digital capture. Their compact design and USB connectivity make them highly convenient for fieldwork and small-scale laboratory settings. With advanced features like measurement software and polarization options, fresh leaf microscopy microscopes have become essential in various research fields, offering a reliable and cost-effective alternative to traditional optical microscopes.

The fresh leaf microscopy method was implemented using the following steps:

(i) Cut the sample from the lower surface of the leaves where it needs to be sampled

(ii) Transfer it to the slide

(iii) Observing and counting by Dino-lite microscope and machine learning model

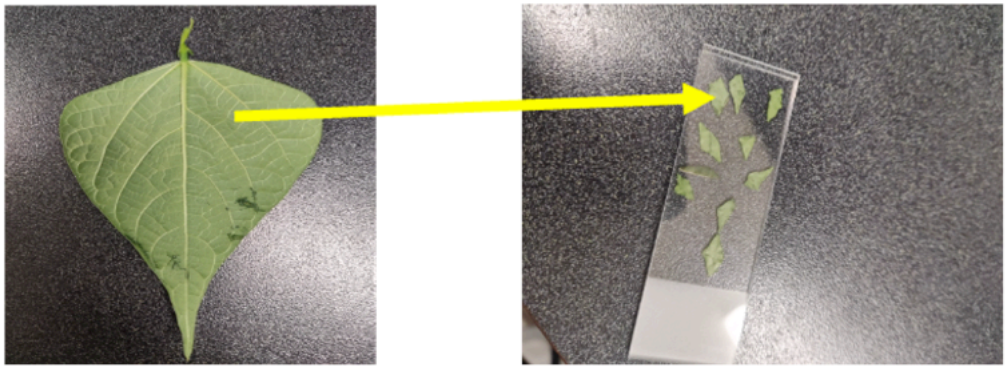

Figure 5. Steps of Fresh leaf microscopy method

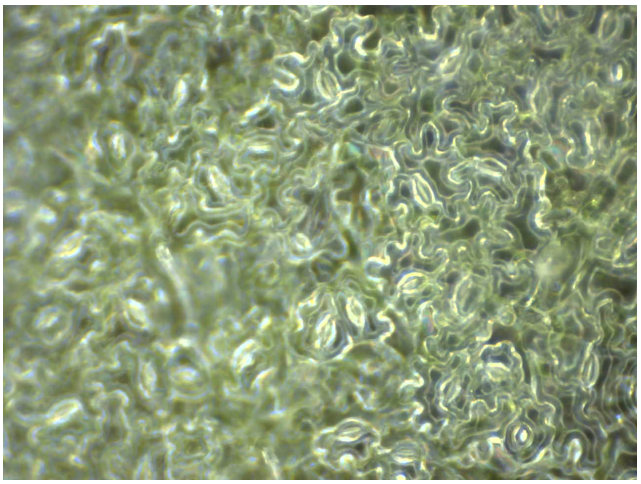

Figure 6. Fresh leaf microscopy image

3.6 S&D (Segmentation and Detection) model

Our S&D model can be divided into 2 parts, the segmentation module, and the detection module. Each module is linked to a calculating module to obtain the result (stomata number per unit area) (Figure. 7).

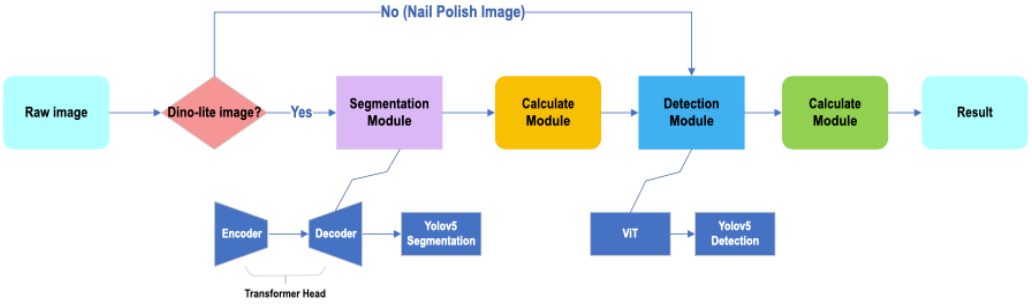

Figure 7. Pipeline of model

i. Segmentation Module Architecture

To improve its performance, our segmentation module was comprised of two parts: 1) a YOLO-based segmentation model and 2) a transformer head before YOLOv5 segmentation.

- **Transformer head**

Vaswani (2017) developed the first Transformer model (Vaswani et al., 2017). With a transformer head, a neural network can extract features more easily than traditional neural networks. A multi-scale strategy utilizes the higher-resolution maps inside our transformer decoder. Following the approach of (Jain et al., 2023), it would prove more straightforward to build a transformer head based on MaskRCNN, as demonstrated in our model architecture. In our model, we feed the object queries (Q) and the multi-scale outputs from the pixel decoder (F_i) as inputs. We use the features with different resolution values of the original image alternatively to update Q using a masked cross-attention (CA) operation, followed by self-attention (SA) and finally a feed-forward network (FFN). We perform talternate

operations L times inside the transformer decoder. The final query outputs from the transformer decoder are mapped to a $(K + 1)$ dimensional space for class predictions, where K denotes the number of classes and an extra $+1$ for the no-object predictions. To obtain the final masks, we decode F_i with the help of an enum operation between Q and F_{i-1}. During inference, we follow the same postprocessing technique as (Cheng et al., 2022) to obtain the final panoptic, semantic, and instance segmentation predictions.

YOLOv5 segmentation model

The YOLOv5 module was applied for segmentation. We borrowed a transformer head to improve our segmentation result. We adopted mean average precision (mAP) as the indicator of detection accuracy:

$$Precision\ (mAP) = \frac{TP}{TP + FP} \tag{1}$$

where TP is the true positives, which means a true segment result, FP indicates the false positives, which means a false segment result.

The segmentation model's loss is given as:

$$loss = \lambda_{coord} \sqrt{\sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{ij}^{obj} [(x_i - \hat{x_i})^2 + (y_i - \hat{y_i})^2]} \tag{2}$$

where S^2 is the number of grids into which the image is divided, B is the number of bounding boxes predicted per grid, λ_{coord} is a weighting factor that controls the importance of the coordinate loss relative to other components of the total loss. It's used to balance the impact of different parts of the loss function during training, I_{ij}^{obj} is the indicator function that is 1 if the j^{th} bounding box in grid and i is responsible for the prediction of the object, and 0 otherwise, (x_i, y_i) are the coordinates of the center of the predicted bounding box in the i^{th} grid, $(\widehat{x_i}, \widehat{y_i})$ are the coordinates of the center of the ground truth bounding box in the i^{th} grid, $[(x_i - \widehat{x_i})^2 + (y_i - \widehat{y_i})^2]$ is the loss term which calculates the squared error between the predicted and ground truth coordinates.

ii. Calculate module architecture

Our calculation module was designed to calculate the clear part of the raw image. The calculate blurred area Algorithm (Alg.1) followed the steps below:

Algorithm 1. Calculate blurred area

```
1  Procedure Blurred calculate(original image, mask image)
2      for i in original image
3          if i in mask image
4              blured_pixel++
5          else                                              (Alg.1)
6              clear_pixel++
7      result=blurred_pixel/(clear_pixel+blurred_pixel)
8  return result
```

iii. Detection Module Architecture

YOLOv5 detection model

The YOLO algorithm can achieve rapid detection of targets (Redmon et al., 2016), detecting the object of interest by dividing images into grids, and processing grids with convolutional layers. This method can achieve real-time detection. However, in our project, the YOLO showed poor accuracy. This was attributable to the fact that without a detailed grid division, there tended to be several targets in the same grid (Comba et al., 2018; Palaniswami et al., 2011).

ViT model

To improve the performance of YOLOv5, we proposed the Vision Transformer (ViT) layer (Dosovitskiy et al., 2020) for feature extraction. ViT is a deep learning architecture that has achieved state-of-the-art performance on various computer vision tasks, including image classification, object detection, and semantic segmentation. This layer partitions an input image into non-overlapping patches of a fixed size. It projects each patch into a lower-dimensional embedding space to produce a sequence of patch embeddings. The multi-head self-attention mechanism then computes a set of attention weights for each embedding, by evaluating its similarity to other embeddings in the sequence. By capturing long-range dependencies in the image, ViT enhances the detection of objects that span multiple regions, thereby improving detection accuracy. In summary, ViT uses a combination of patch

embeddings and attention mechanisms to effectively analyze images and identify complex objects that may be distributed across different regions.

The loss function (Dong et al., 2022) can be defined as:

$$loss = loss_{obj} + loss_{cls} + loss_{box} \tag{3}$$

The confidence loss $loss_{obj}$ can be defined as:

$$\begin{aligned} loss_{obj} = & \sum_{i=0}^{s^2} \sum_{j=0}^{B} I_{ij}^{obj} (C_i \log(C_i) + (1 - C_i) \log (C_i)) \\ & - \lambda_{nobj} \sum_{i=0}^{s^2} \sum_{j=0}^{B} I_{ij}^{obj} (C_i \log(C_i) + (1 - C_i) \log (C_i)) \end{aligned} \tag{4}$$

where $C_i \; is$ confidence score of the i^{th} grid, representing the probability that an object is present.

λ_{nobj} is the hyperparameter that balances the loss between the confidence scores of objects and non-objects.

The classification loss, $loss_{cls}$, can then be defined as:

$$loss_{cls} = \sum_{i=0}^{s^2} \sum_{j=0}^{B} I_{ij}^{obj} (P_i \log(P_i) + (1 - P_i) \log (P_i)) \quad (5)$$

where P_i is the probability distribution over the classes for the i^{th} grid.

The GLoU loss function (Rezatofighi et al., 2019) is used to express the position loss of the target box and the prediction box, $loss_{box}$.

$$loss_{box} = L_{GIoU} = 1 - (IoU - \frac{|C - (A \cup B)|}{|C|}) \quad (6)$$

where L_{GIoU} is the generalized Intersection over Union loss, which measures the difference between the predicted bounding box and the ground truth bounding box

IoU is the intersection over the Union, the area of overlap between the predicted bounding box and the ground truth, divided by the area of their union.

C is the smallest enclosing box that contains both the predicted bounding box A and the ground truth bounding box B.

iv. Calculate module architecture

Our calculation module was designed to calculate the clear part in the raw image. The Calculate Stomata per unit area algorithm, which generates a result of $\frac{Stomata\ number}{unit\ area}$ follows the steps (Alg.2).

The calculate module followed the format.

Algorithm 2. Calculate Stomata per unit area

```
1   Procedure Stomata per unit area calculate(stomata number, unit area)
2       Result=stomata number/unit area
3   return result
```

(Alg.2)

So, the total loss score in the S&D model can be calculated as follows:

$$loss = \lambda_1 loss_{Segementation} + \lambda_2 loss_{Detection} \quad (7)$$

Where $loss_{Segementation}$ equal the $loss$ from format.2, and $loss_{Detection}$ equal the $loss$ from format.3. The λ_1 and λ_2 based on the definition, we define that if this is in the Fresh leaf microscopy images, we utilized the Segmentation module and Detection module so the λ_1=λ_2=0.5, while in the Nail polish methods, we didn't utilize the segmentation module, so the λ_1=0 while λ_2=1.

v. Regression model

Model Definition

To investigate the relationship between the stomata number per area measured using two different methods, we developed a linear regression model. In this model, the dependent

variable Y represents the stomata number per area obtained using the opt method, while the independent variable X represents the stomata number per area measured by the Dinolite method. The linear regression model can be expressed as:

$$Y = \beta_0 + \beta_1 X + \in \quad (8)$$

Where the Y denotes the stomata number per area as measured by the opt method,

X denotes the stomata number per area as measured by the Dinolite method,

β_0 is the intercept of the regression line,

β_1 is the slope of the regression line,

$\in$ represents the error term, capturing the deviation of the observed values from the predicted values.

Based on our data we chose The Least Squares Method to estimate the parameters of the regression model. This method is widely used to minimize the sum of the squared differences between the observed values and the values predicted by the model. The approach provides estimates of the regression coefficients β_0 and β_1 that best fit the observed data.

Data Preparation

The data was organized into two columns for the regression analysis: one column for the stomata number per area obtained using the opt method (Y), and another for the stomata number per area obtained using the Dinolite method (X).

Parameter of model

The SD (standard deviation) was calculated as:

$$SD = \sqrt{\frac{\sum(X_i - \mu)^2}{n}} \tag{9}$$

where n means the sum number of the genotype.

X_i means the genotype result.

μ means the average in all the genotype results.

The SE (standard error) was calculated as:

$$SE = \frac{\sqrt{SD}}{n} \tag{10}$$

Also, the residuals can be defined as $e_i = y_i - f_i$ (forming a vector e),

where y_i is the observed value,

f_i is the predicted value from the regression model.

So that the SS_{res} was calculated as:

$$SS_{res} = \sum_i (y_i - f_i)^2 , \quad (11)$$

So that the SS_{tot} was calculated as:

$$SS_{tot} = \sum_i (y_i - \mu)^2 , \quad (12)$$

So that the R^2 was calculated as:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} , \quad (13)$$

Network training and evaluation

The tool LabelMe was used for image annotation. The dataset is randomly split into training validation and test groups with a rate of 7:2:1, which means our segmentation module dataset includes 123 images. With data augmentation, we utilized horizontal and vertical images as data augmentation methods, so each image has been augmented into 2 images, for a total of 246 images. 231 images were utilized for training, 10 images for validating, and 17 images for testing.

Our detection module dataset includes 211 images, 148 images were utilized for training, 42 images for validating and 21 images were employed for testing. The network training and validation were conducted on the Windows 11 operating system, and all the tasks from the training were carried out on a GeForce RTX 3080Ti graphics card.

The experiment database comprises a total of 1446 images derived from 119 distinct genotypes, with each genotype represented by at least 10 repeated images. To ensure comprehensive representation, one or two images were randomly selected from each genotype, resulting in a complete dataset without missing genotypes.

The batch size in the segmentation module is 16, the epochs we chose were 100, and the image size had been selected with 640 ((640,640,3), which means 640 pixel×640 pixel×3 channels (Red, Blue, Green)) The batch size in the detection module is 4, epochs were 100, and the image size was selected with 640 ((640,640,3), which means 640 pixel×640 pixel×3 channels (Red, Blue, Green)).

Our code can be found in: https://github.com/empersun/YOLOv5forstomatadetection

3.7 Experiment setup

After we finished training our S&D model, we randomly chose 1329 Fresh leaf microscopy images, including 130 genotypes, to validate our Fresh leaf microscopy method and 1329 nail polish images, including the same 130 genotypes, to validate our Nail Polish method. Each image we collected from the same leaves in symmetrical parts.

4. RESULTS AND DISCUSSIONS

4.1 Segmentation model result

Model name	Trans-YOLOv5	YOLOv5	Mask-RCNN
mAP	89.4%	78.2%	73.4%

Table 1: Segmentation result

We evaluated the segmentation module's performance by computing its accuracy rates (mAP). The mAP increases from 78.2% to 89.4% with the transformer head. Figure 8 shows the segmentation result between our Trans-YOLOv5 model and Mask RCNN and YOLOv5 model.

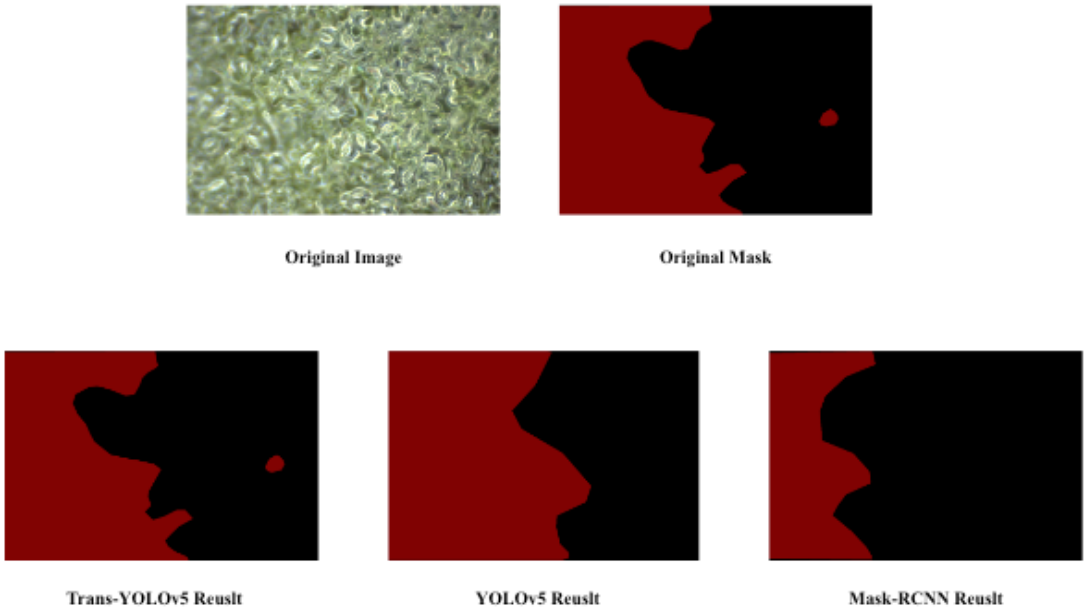

Figure 8. Segmentation Results: Segmentation result compares between Trans-YOLOv5, YOLOv5, and Mask-RCNN.

The red mask means the blurred part and the black mask shows the cleared part. Our model significantly includes more details, making segmentation stomata images more accurate.

We can observe that the segmentation results produced by our Improved YOLOv5 model provide more information and exhibit lower loss compared to the previous versions.

In this study, the segmentation model is designed to partition the image into two distinct regions: clear and unclear. The primary objective is to calculate the result based on the clear region using a detection model. A threshold of 40% is applied to determine the usability of the image. Specifically, an image is classified as unusable if the clear region comprises less than 40% of the total area.

4.2 Detection model result

Model Name	Trans-YOLOv5	YOLOv5	CNN
mAP	97.7%	92.7%	75.1%

Table 2. Detection module result

We evaluated the performance of the detection module by computing its mAP. With the ViT head, the mAP increases from 75.1% to 97.7%

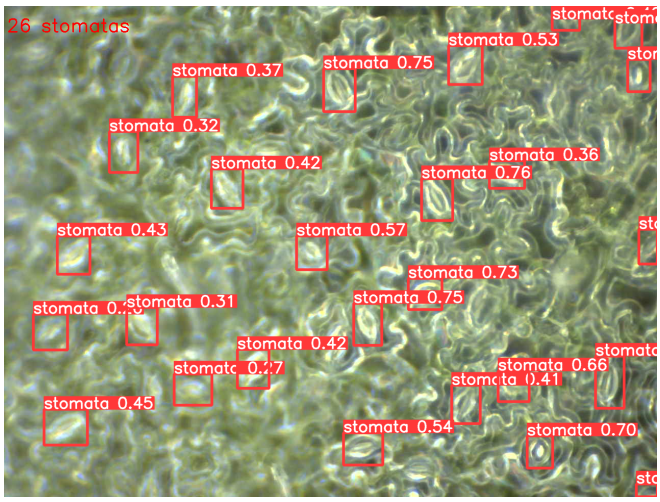

Figure 9. Fresh leaf microscopy detection result

We can observe that the detection results in the Fresh leaf microscopy image which was produced by our Improved YOLOv5 model presented almost all the stomata and ignored some seems-stomata.

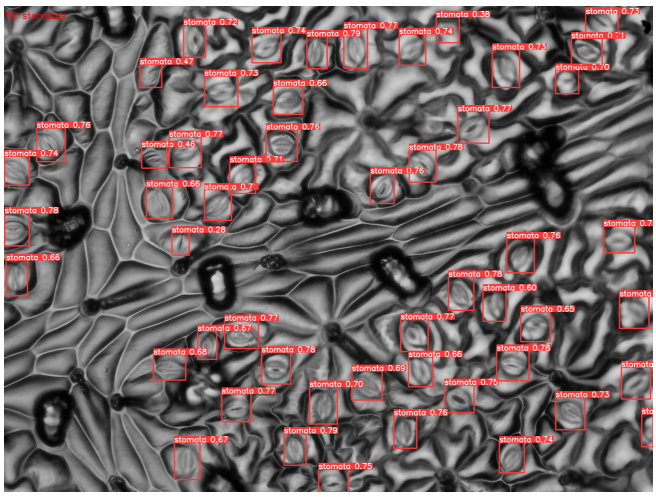

Figure 10. Nail polish detection result

The visual results show that our model can accurately identify and represent stomata placement and structure, even when the stomata may be slightly occluded. The boxes detected must be primarily located in the center of the stomata and not cover all pixels from the background or seems-stomata. This is important because the models are trained to detect the true stomata. Additionally (Figures 9 and 10) with fresh leaf microscopy and stomata results, our method successfully detects stomata.

Figure.11 presents a scatter plot illustrating the linear relationship between two variables. Both the x-axis and y-axis represent the number of stomata in the image. Each data point is represented by a dot, with error bars indicating the measurement errors. The dashed line in the plot represents the linear regression fit, described by the equation y=0.99x+0.09. In general, if the range of R^2 between $0.5 \leq R^2 \leq 1$ we can consider the data as regression (Allen, 2004). The coefficient of determination, R^2, is 0.99, indicating a strong positive correlation between the two variables. The standard error (SE) in our model is 0.40. The SE

from regression is consistent with measurement variability, supporting the reliability of the manual counting method detection results alongside the S&D model method detection result. Thus, these two data can be considered significant, which means the S&D model result and manual counting result are the same result.

The statistical analysis involved in validating the reproducibility of stomata per image measurements used two methods: Manual and S&D model. Levene's test ($p = 0.99$) demonstrated no significant difference in variances, supporting the reproducibility across methods. An F-test ($p = 0.49$) compared variances between the methods and confirmed they were statistically comparable.

These findings indicate that the S&D model and manual measurement methods produce consistent and reproducible results in stomata density measurements. The non-significant Levene's test result suggests that both methods maintain comparable precision across their measurements. Similarly, the F-test result indicates that there is no significant difference in the variances between the two methods, further validating their consistency.

The statistical comparability of these methods implies that the S&D model can be confidently employed as an alternative to manual measurement. This is particularly beneficial for large-scale studies where manual measurement may be time-consuming. Given that the S&D model performs equally well in terms of variability, it presents an efficient and reliable tool for stomata analysis. Future studies can rely on these findings to choose the S&D model for rapid analysis without compromising on accuracy.

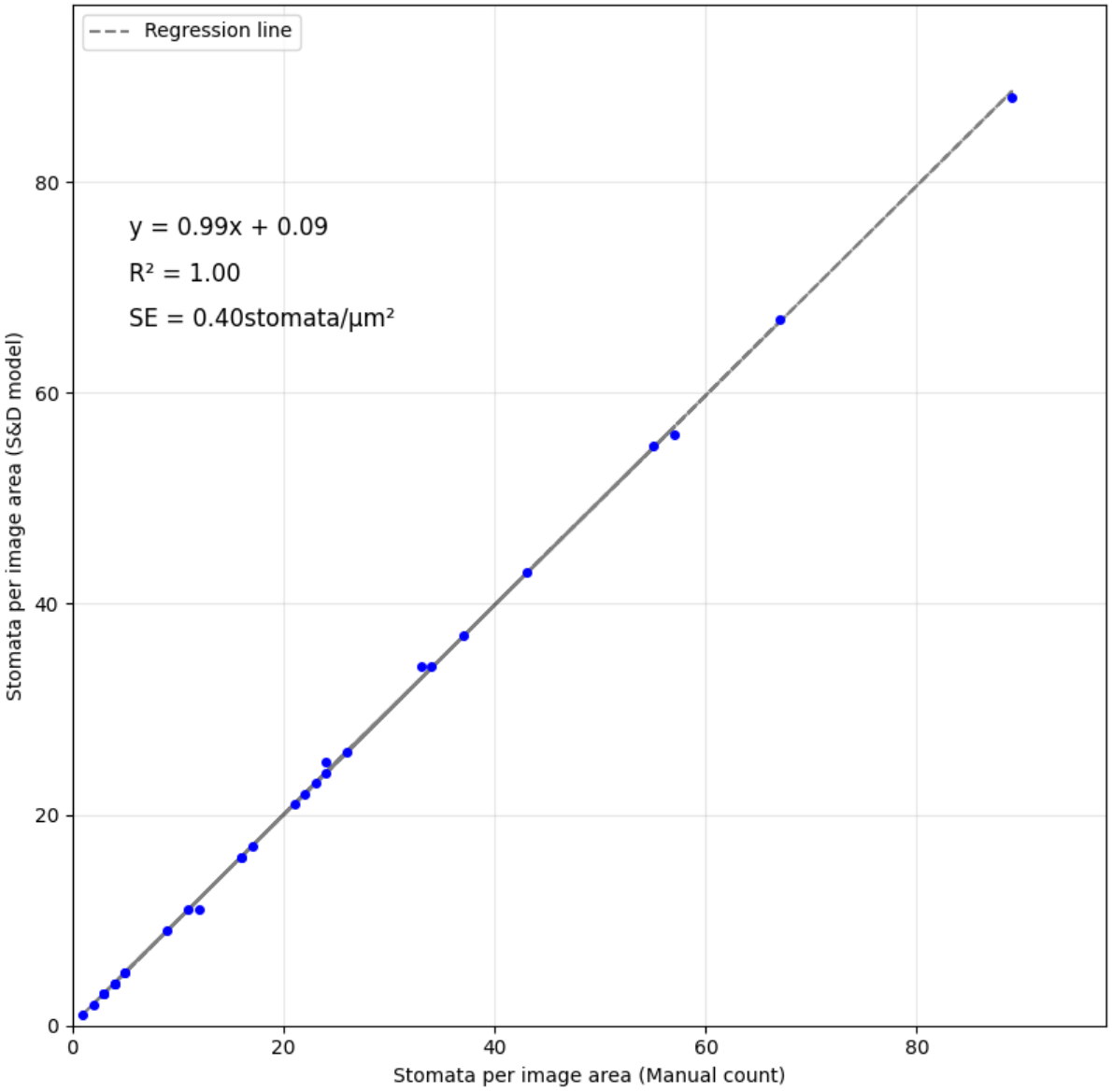

Figure 11. Regression analysis of detection result between S&D model and manual result

4.3 Regression model result

In this study, we employed a dual-microscopy approach to optimize the accuracy of leaf surface imaging. Our primary imaging tool, the Fresh leaf microscopy digital microscope, provides illumination from a lens-mounted light source. However, when used to observe slides, this setup leads to undesirable reflections that interfere with image clarity. Conversely,

traditional optical microscopes, which illuminate from below the sample, are well-suited for raw sample observations but face challenges with overlapping leaf cell structures. These overlaps can obscure key features such as stomatal pores, compromising image interpretation.

The Fresh leaf microscopy zoom is between 839.7 to 850.1, which means the raw image area is between 161702 to 161791μm^2, and our optical microscope area is 284375μm^2.

Figure.12 presents a scatter plot illustrating the linear relationship between two variables. Based on the provided dataset and regression analysis, the results indicate a strong linear relationship between stomata density measurements obtained from the nail-polish method and the Dino-lite method. The high values of the multiple R (0.93) and R-squared (0.87) suggest that the nail-polish method can reliably predict stomata density as measured by Dino-lite with approximately 87% of the variance in Dino-lite measurements being explained by the nail-polish method.

The adjusted R-squared value (0.87) further confirms the model's robustness, indicating that the relationship holds consistently across the dataset. Additionally, the standard error of 29.00 suggests a relatively low dispersion of residuals, implying accurate prediction by the regression model.

The ANOVA results show a highly significant regression model with an F-statistic of 787.11 ($p < 0.001$), providing strong evidence that the relationship between the two methods

is statistically significant. The coefficient for the nail-polish method is 0.81 ($p < 0.001$), indicating a strong positive relationship between the two methods.

In conclusion, the analysis demonstrates that the two methods are highly comparable, with statistically significant linear correlations, minimal variance, and strong reproducibility, making both techniques viable for stomata measurement depending on the practical constraints of the research. However, it is important to note that the Dino-lite method detects approximately 10% fewer stomata compared to the nail-polish method, as indicated by the regression analysis. Despite this slight underestimation, the Dino-lite method remains a reliable alternative due to its strong correlation with the nail-polish method and ease of use in non-invasive measurements.

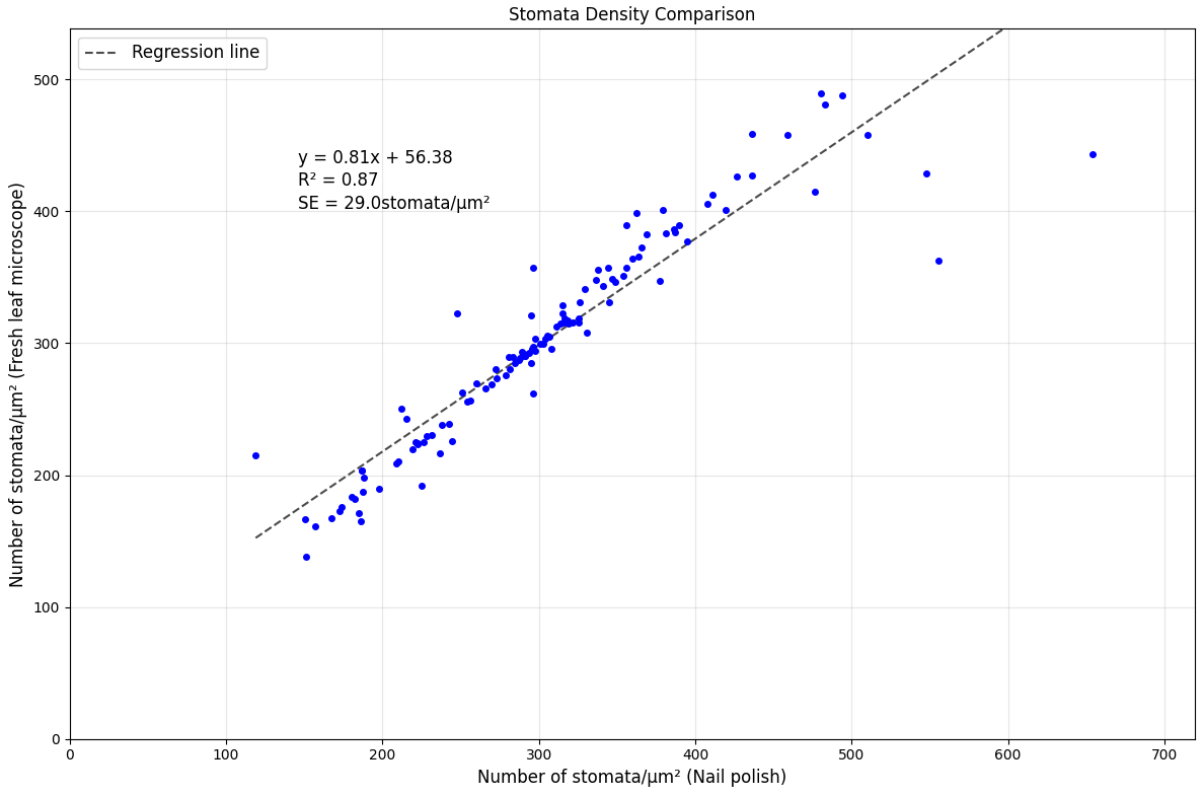

Figure 12. Regression analysis result of nail polish method and Fresh leaf microscopy method

4.4 Discussion

The S&D model demonstrated superior accuracy in object detection tasks, achieving 89.4% accuracy compared to Mask-RCNN and 97.7% accuracy compared to CNN. This shows the model's advanced integration of transformer techniques with the YOLOv5 framework, leading to enhanced precision in detecting and segmenting objects. The high accuracy performance of the S&D model suggests that it is highly applicable for practical object detection and segmentation tasks, making it a valuable tool for real-world applications. Its effectiveness in various contexts highlights its potential for broad deployment in complex environments. However, the model still needs more data, including different periods of stomata, to optimize to reduce potential sources of error and variability in object detection to enhance model performance further.

The regression model, analyzing the relationship between stomata density measurements from the opt and fresh leaf microscope methods, achieved the R^2 value of 0.87. This indicates that the Fresh leaf microscopy measurements explain approximately 87% of the variation in the stomata number per area measured by the optical microscope method. The SE of $29.0 stomata/\mu m^2$ suggest that while the model is accurate, inherent variability must be addressed. The strong correlation between the opt and Fresh leaf microscopy methods indicates that the regression model helps estimate stomata density. This model can be applied in field settings where the Fresh leaf microscopy method is used as an alternative to the opt method, especially when high precision is required. Additional data will be needed

to calibrate and validate the Fresh leaf microscopy method against the opt method, which is necessary to ensure its reliability and accuracy in various settings. and determine the reason for SD changes (e.g., which geno-pair, what kind of DNA fold will have a significant influence). Investigate sources of variability indicated by the high SD to improve model precision and reliability. Understanding and mitigating these factors will enhance the accuracy of stomata density measurements.

The strong correlation between the nail polish method and the Fresh leaf microscopy method indicates that the Fresh leaf microscopy method is a reliable approach for estimating stomata density. While the model shows promise for practical applications, addressing variability and validating the model against other methods will further enhance its reliability.

In conclusion, the strengths of the S&D model offer a robust framework for advancing stomata detection and calculation. Future research should focus on validating these findings, optimizing performance, and exploring integrated approaches to fully leverage these models' advantages in practical applications.

Several previous studies have explored automated methods for stomata detection, such as AlexNet method (Jayakody et al., 2017a), MobileNet (Kwong et al., 2021), StomataCounter (Karl C Fetter et al., 2019) and plantprofile (Sultana et al., 2021a). However, AlexNet exhibited significant limitations in processing speed, requiring 30 to 40 minutes to analyze a single image, whereas our fresh leaf microscopy method delivers results within 3 minutes. When we applied MobileNet to our database, the model's accuracy decreased

substantially due to its cross-application on images from different species, a challenge similarly faced by StomataCounter. Although PlantProfile achieved the highest accuracy of 94.5% on nail polish images, its performance was adversely affected by the presence of color in fresh leaf microscope images. Also, these automated methods struggled primarily with effectively eliminating noise and adequately capturing the complexity of stomatal structures. As a result, they often misidentified non-stomatal structures as stomata or failed to detect actual stomata altogether.

In comparison to previously published automated stomata detection methods, our approach demonstrates a high correlation ($R^2 = 0.87$) between Dino-lite and Nail-polish techniques, with comparable variance based on Levene's (p=0.49) and F-tests (p=0.06). While automated methods may boast higher throughput, our technique ensures robust reproducibility without significant variance, offering a balance between accuracy and method validity. Moreover, the consistency of the regression SE (1.68 stomata/μm²) and SD (13.27 stomata/μm²) further supports the precision of manual measurements when compared to advanced automated models.

5. CONCLUSIONS AND RECOMMENDATIONS

5.1 General conclusion

This study aims to develop an innovative platform utilizing Fresh leaf microscopy technology and comparing its performance with traditional nail polish methods. We propose a novel model to segment and detect stomata, calculating their number per unit area. Our segmentation module integrates a transformer head, trained to optimize segmentation accuracy by minimizing loss points. Similarly, our detection module incorporates a Vision Transformer (ViT) mechanism, trained to enhance feature extraction and detection precision. The proposed stomata detection model offers improved accuracy, robustness against noise, and real-time detection capabilities for live legume crops. Experimental results demonstrate that our model effectively calculates stomata density from in-field images. Although this paper primarily addresses stomata calculation, the proposed model could be extended to other applications, such as "seem-stomata calculation" or other parts of live legume crops. Future work will involve testing the model on these extended applications as relevant data becomes available.

5.2 Suggestion for future work

In this experiment, the fresh leaf microscopy method demonstrated its potential to become one of the most convenient ways for stomata detection. Recent studies have used live plants from 130 genotypes to make slides to take stomata images. Further research should focus on exploring and understanding other genotypes and determining if they would influence our regression analysis model. Future research should examine how genotype influences stomata structure. Also, we invested leaves that are less than $3*3mm^2$ because the stomata in these leaves are not mature. We could develop our model further to find the unmatured stomata from small leaves and compare different genotypes' influence on stomata structures. More bean genotypes and high spectrum analysis may need to be included to provide a better understanding of the effects of genes on stomata numbers and quality. In addition, the different treatments in this experiment were done under room temperature storage. The shelf-life quality of bean leaves deteriorated within 2 to 3 days; therefore, it would be beneficial to test the effects of temperature on stomata structure and number of beans under different temperature conditions to determine how we can increase bean production.

REFERENCES

Afzal, M., Alghamdi, S. S., Migdadi, H. H., El-Harty, E. H., & Al-Faifi, S. A. (2022). Agronomical and Physiological Responses of Faba Bean Genotypes to Salt Stress. *Agriculture*.

Allen, M. P. (2004). *Understanding regression analysis*. Springer Science & Business Media.

Araújo, W. L., Fernie, A. R., & Nunes-Nesi, A. (2011). Control of stomatal aperture. *Plant Signaling & Behavior*, *6*, 1305 - 1311.

Argaw, A., & Muleta, D. B. (2017). Effect of genotypes-Rhizobium-environment interaction on nodulation and productivity of common bean (Phaseolus vulgaris L.) in eastern Ethiopia. *Environmental Systems Research*, *6*, 1-16.

Bethea, R. M., & Rhinehart, R. R. (2019). Regression Analysis. *Applied Engineering Statistics*.

Bhaiswar, N., Dixit, D. V. V., & Student, P. G. (2016). A Review: Methods of Automatic Stomata Detection and Counting Through Microscopic Images of a Leaf.

Bhugra, S., Mishra, D., Anupama, A., Chaudhury, S., Lall, B., Chugh, A., & Chinnusamy, V. (2018). Deep Convolutional Neural Networks Based Framework for Estimation of Stomata Density and Structure from Microscopic Images. ECCV Workshops,

Board, J. E., & Kahlon, C. S. (2011). Soybean yield formation: what controls it and how it can be improved. *Soybean physiology and biochemistry*, 1-36.

Brewer, C. A. (1992). Responses by stomata on leaves to microenvironmental conditions. *Tested studies for laboratory teaching*, *13*, 67-75.

Brinkmann, J., Sheshadri, A., Levoso, V., Swoboda, P., & Bartelt, C. (2024). A Mechanistic Analysis of a Transformer Trained on a Symbolic Multi-Step Reasoning Task. *ArXiv*, *abs/2402.11917*.

Cai, S., Papanatsiou, M., Blatt, M. R., & Chen, Z. (2017). Speedy Grass Stomata: Emerging Molecular and Evolutionary Features. *Molecular plant*, *10 7*, 912-914.

Chen, Y., Pu, H., & Qu, Y. (2024). An analysis of attention mechanisms and its variance in transformer. *Applied and Computational Engineering*.

Cheng, B., Misra, I., Schwing, A. G., Kirillov, A., & Girdhar, R. (2022). Masked-attention mask transformer for universal image segmentation. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition,

Ciha, A., & Brun, W. (1975). Stomatal size and frequency in soybeans 1. *Crop Science*, *15*(3), 309-313.

Cober, E. R., & Voldeng, H. D. (2012). A retrospective look at short-season soybean cultivar development in Ontario. *Canadian Journal of Plant Science*, *92*(7), 1239-1243.

Coimbra, J. L. M., Guidolin, A. F., Carvalho, F. I. F. d., Coimbra, S. M. M., & Hemp, S. (1999). Reflexos da interação genótipo X ambiente e suas implicações nos ganhos de seleção em genótipos de feijão (Phaseolus vulgaris L.). *Ciencia Rural*, *29*, 433-439.

Collins, T. J. (2007). ImageJ for microscopy. *Biotechniques*, *43*(S1), S25-S30.

Comba, L., Biglia, A., Aimonino, D. R., & Gay, P. (2018). Unsupervised detection of vineyards by 3D point-cloud UAV photogrammetry for precision agriculture. *Computers and electronics in agriculture*, *155*, 84-95.

Constantin, C. P. (2017). Using the Regression Model in multivariate data analysis.

Dastneshan, S., Bihamta, M. R., Abbasi, A., & Sabokdast, M. (2019). The Effect of Different Levels of Drought Stress on some Physiological Traits and Chlorophyll Fluorescence of Bean Genotypes (Phaseolus Vulgaris L.). *Journal of Crop Breeding*.

Djehiche, B. (2011). Regression modeling with actuarial and financial applications by Edward W. Frees. *Scandinavian Actuarial Journal*, *2011*, 319 - 319.

Dong, X., Yan, S., & Duan, C. (2022). A lightweight vehicles detection network model based on YOLOv5. *Engineering Applications of Artificial Intelligence*, *113*, 104914.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., & Gelly, S. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.

Dourado, L. R. B., Pascoal, L. A. F., Sakomura, N. K., Costa, F. G. P., & Biagiotti, D. (2011). Soybeans (Glycine max) and soybean products in poultry and swine nutrition. *Recent trends for enhancing the diversity and quality of soybean products*, 175-190.

Duarte, K. T. N., Carvalho, M. A. G., & Martins, P. S. (2017). Segmenting High-quality Digital Images of Stomata using the Wavelet Spot Detection and the Watershed Transform. VISIGRAPP,

Dwevedi, A., & Kayastha, A. M. (2011). Soybean: a multifaceted legume with enormous economic capabilities. *Soybean—biochemistry, chemistry and physiology. InTech Publisher, Janeza Trdine, Croatia*, 165-188.

Endo, H., & Torii, K. U. (2019). Stomatal Development and Perspectives toward Agricultural Improvement. *Cold Spring Harbor perspectives in biology*, *11 5*.

Fageria, N. K., Baligar, V. C., Melo, L. C., & Oliveira, J. P. d. (2012). Differential Soil Acidity Tolerance of Dry Bean Genotypes. *Communications in Soil Science and Plant Analysis*, *43*, 1523 - 1531.

Fehr, W. R., Hoeck, J. A., Johnson, S. L., Murphy, P. A., Nott, J. D., Padilla, G. I., & Welke, G. A. (2003). Genotype and environment influence on protein components of soybean. *Crop Science*, *43*(2), 511-514.

Fetter, K. C., Eberhardt, S., Barclay, R. S., Wing, S., & Keller, S. R. (2019). StomataCounter: a neural network for automatic stomata identification and counting. *New Phytologist*, *223*(3), 1671-1681.

Fetter, K. C., Eberhardt, S., Barclay, R. S., Wing, S. L., & Keller, S. R. (2019). StomataCounter: a neural network for automatic stomata identification and counting. *The New phytologist*.

Finoto, E. L., Soares, M. B. B., Albuquerque, J. d. A. A. d., Monteiro Neto, J. L. L., Maia, S. d. S., Dona, S., Chagas, E. A., Soares-da-Silva, E., & Abanto-Rodriguez, C. (2021). Selection of soybean genotypes for yield, size, and oil and protein contents. *Australian Journal of Crop Science*, *15*(1), 48-50.

Fu, Z. (2022). Vision Transformer: Vit and its Derivatives. *ArXiv*, *abs/2205.11239*.

Ghojogh, B., & Ghodsi, A. (2020). Attention Mechanism, Transformers, BERT, and GPT: Tutorial and Survey.

Glaser, E. (2004). Regression Analysis. *The Missouri Review*, *27*, 140 - 141.

Gong, H. (2022). Implements of Transformer in NLP and DKT. *2022 4th International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM)*, 805-807.

Gupta, A., Sharma, A., & Goel, A. (2017). Review of Regression Analysis Models. *International Journal of Engineering Research and*, *6*.

Haag, W., Adams, M. W., & Wiersma, J. V. (1978). Differential Responses of Dry Bean Genotypes to N and P Fertilization of a Central American Soil 1. *Agronomy Journal*.

Hampton, J. G., Mesquita, F. J. A., & Hill, M. J. (1997). The effect of genotype and plant density on grain yield and cooking quality in common bean (Phaseolus vulgaris L.).

Hartman, G. L., West, E. D., & Herman, T. K. (2011). Crops that feed the World 2. Soybean—worldwide production, use, and constraints caused by pathogens and pests. *Food Security*, *3*, 5-17.

Heidari, M., Azad, R., Kolahi, S. G., Arimond, R. e., Niggemeier, L., Sulaiman, A., Bozorgpour, A., Aghdam, E. K., Kazerouni, A., Hacihaliloglu, I., & Merhof, D. (2024). Enhancing Efficiency in Vision Transformer Networks: Design Techniques and Insights. *ArXiv*, *abs/2403.19882*.

Higaki, T., Kutsuna, N., & Hasezawa, S. (2014). CARTA-based semi-automatic detection of stomatal regions on an Arabidopsis cotyledon surface. *Plant Morphology*, *26*, 9-12.

Intelligence, M. (2017). Soybean market – Growth, trends, and forecast (2020–2025). *https://www.mordorintelligence.com/industry-reports/soybean-market*.

Jain, J., Li, J., Chiu, M. T., Hassani, A., Orlov, N., & Shi, H. (2023). Oneformer: One transformer to rule universal image segmentation. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,

Jayakody, H., Liu, S., Whitty, M., & Petrie, P. (2017a). Microscope image based fully automated stomata detection and pore measurement method for grapevines. *Plant Methods*, *13*, 1-12.

Jayakody, H., Liu, S., Whitty, M. A., & Petrie, P. R. (2017b). Microscope image based fully automated stomata detection and pore measurement method for grapevines. *Plant Methods*, *13*.

Jayakody, H., Petrie, P. R., Boer, H. J. d., & Whitty, M. A. (2020). A generalised approach for high-throughput instance segmentation of stomata in microscope images. *Plant Methods*, *17*.

Jones, H. G., Stoll, M., Santos, T. P., Sousa, C. d., Chaves, M. M., & Grant, O. M. (2002). Use of infrared thermography for monitoring stomatal closure in the field: application to grapevine. *Journal of experimental botany*, *53 378*, 2249-2260.

Kalve, S., De Vos, D., & Beemster, G. T. (2014). Leaf development: a cellular perspective. *Frontiers in plant science*, *5*, 86936.

Kwong, Q. B., Wong, Y. C., Lee, P. L., Sahaini, M. S., Kon, Y. T., Kulaveerasingam, H., & Appleton, D. R. (2021). Automated stomata detection in oil palm with convolutional neural network. *Scientific reports*, *11*(1), 15210.

Labbaf Khaniki, M. A., Mirzaeibonehkhater, M., & Manthouri, M. (2023). Enhancing Pneumonia Detection using Vision Transformer with Dynamic Mapping Re-Attention Mechanism. *2023 13th International Conference on Computer and Knowledge Engineering (ICCKE)*, 144-149.

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

Lamb, A., He, D., Goyal, A., Ke, G., Liao, C.-F., Ravanelli, M., & Bengio, Y. (2021). Transformers with Competitive Ensembles of Independent Mechanisms. *ArXiv*, *abs/2103.00336*.

Larkin, J. C., Marks, M. D., Nadeau, J., & Sack, F. (1997). Epidermal cell fate and patterning in leaves. *The Plant Cell*, *9*(7), 1109.

Lawson, T., & Matthews, J. S. A. (2020). Guard Cell Metabolism and Stomatal Function. *Annual review of plant biology*.

Le, J., Zou, J., Yang, K., & Wang, M. (2014). Signaling to stomatal initiation and cell division. *Frontiers in plant science*, *5*, 99423.

Levitt, J. (1976). Physiological Basis of Stomatal Response.

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*.

Liang, X., Xu, X., Wang, Z., He, L., Zhang, K., Liang, B., Ye, J., Shi, J., Wu, X., Dai, M., & Yang, W. (2021). StomataScorer: a portable and high-throughput leaf stomata trait scorer combined with deep learning and an improved CV model. *Plant Biotechnology Journal*, *20*, 577 - 591.

Liu, S., Tang, J., Petrie, P. R., & Whitty, M. A. (2016). A Fast Method to Measure Stomatal Aperture by MSER on Smart Mobile Phone.

Lleras, E. (1977). Differences in stomatal number per unit area within the same species under different micro-environmental conditions: a working hypothesis. *Acta Amazonica*, *7*, 473-476.

Lucas, J. R., & Renzaglia, K. S. (2002). Structure and Function of Hornwort Stomata. *Microscopy and Microanalysis*, *8*, 1090 - 1091.

Marom, Z., Shtein, I., & Bar-On, B. (2017). Stomatal Opening: The Role of Cell-Wall Mechanical Anisotropy and Its Analytical Relations to the Bio-composite Characteristics. *Frontiers in plant science*, *8*.

Masuda, T., & Goldsmith, P. D. (2009). World soybean production: area harvested, yield, and long-term projections. *International food and agribusiness management review*, *12*(4), 1-20.

Matthews, J. S. A., & Lawson, T. (2019). Climate Change and Stomatal Physiology. *Annual Plant Reviews online*.

McKown, K. H., & Bergmann, D. C. (2018). Grass stomata. *Current Biology*, *28*, R814-R816.

Meeus, S., Wyffels, F., & Van den Bulcke, J. (2019). From leaf to label: A robust automated workflow for stomata detection. *Ecology and Evolution*, *10*, 9178 - 9191.

Merced, A., & Renzaglia, K. S. (2017). Structure, function and evolution of stomata from a bryological perspective. *Bryophyte Diversity and Evolution*, *39*, 7-20.

Michelutti, R., Tu, J., Hunt, D., Gagnier, D., Anderson, T., Welacky, T., & Tenuta, A. (2002). First report of Bean pod mottle virus in soybean in Canada. *Plant Disease*, *86*(3), 330-330.

Millstead, L., Jayakody, H., Patel, H., Kaura, V., Petrie, P. R., Tomasetig, F., & Whitty, M. A. (2020). Accelerating Automated Stomata Analysis Through Simplified Sample Collection and Imaging Techniques. *Frontiers in plant science*, *11*.

Nadeau, J. A., & Sack, F. D. (2002). Stomatal development in Arabidopsis. *The Arabidopsis book/American Society of Plant Biologists*, *1*.

Naseer, M., Ranasinghe, K., Khan, S. H., Hayat, M., Khan, F. S., & Yang, M.-H. (2021). Intriguing Properties of Vision Transformers. Neural Information Processing Systems,

Nchimbi–Msolla, S., & Tryphone, G. M. (2010). The effects of the environment on iron and zinc concentrations and performance of common bean (Phaseolus vulgaris L.) genotypes. *Asian Journal of Plant Sciences*, *9*, 455-462.

Nunes, T. L. d. S., Zhang, D., & Raissig, M. T. (2019). Form, development and function of grass stomata. *The Plant journal : for cell and molecular biology*.

Nwadike, C., Vange, T., & Ochigbo, A. E. (2014). Genotype - Environment Interaction for Plant Development and Some Yield Components in Common Bean (Phaseolus vulgaris L.) during the 2012 Wet Season. *Journal of Plant breeding and Crop Science*, *2*, 001-008.

Oliveira, M. W. d. S., Casanova, D., Pinheiro, L. F. S., Kolb, R. M., & Bruno, O. M. (2014). Automatic counting of stomata in epidermis microscopic images.

Palaniswami, C., Gopalasundaram, P., & Bhaskaran, A. (2011). Application of GPS and GIS in Sugarcane Agriculture. *Sugar Tech*, *13*, 360-365.

Palmer, R., & Hymowitz, T. (2016). Soybean: Germplasm, breeding, and genetics.

Papastylianou, P., Vlachostergios, D. N., Dordas, C., Tigka, E. L., Papakaloudis, P., Kargiotidou, A., Pratsinakis, E. D., Koskosidis, A., Pankou, C. I., Kousta, A., Mylonas, I., Tani, E., Abraham, E. M., Karatassiou, M., & Kostoula, S. (2021). Genotype X Environment Interaction Analysis of Faba Bean (Vicia faba L.) for Biomass and Seed Yield across Different Environments. *Sustainability*.

Pereira, T. C. V., Schmit, R., Haveroth, E. J., Melo, R. C. d., Coimbra, J. L. M., Guidolin, A. F., & Backes, R. L. (2015). Reflexo da interação genótipo x ambiente sobre o melhoramento genético de feijão. *Ciencia Rural*, *46*, 411-417.

Philipo, M., Ndakidemi, P. A., & Mbega, E. R. (2020). Environmental and genotypes influence on seed iron and zinc levels of landraces and improved varieties of common bean (Phaseolus vulgaris L.) in Tanzania. *Ecological Genetics and Genomics*, *15*, 100056.

Philipo, M., Ndakidemi, P. A., & Mbega, E. R. (2021). Environmentally stable common bean genotypes for production in different agro-ecological zones of Tanzania. *Heliyon*, *7*.

Prabhakar, M. (2004). Structure, delimitation, nomenclature and classification of stomata. *Journal of Integrative Plant Biology*, *46*(2), 242.

Qian, B., Smith, W., Jing, Q., Kim, Y. M., Jégo, G., Grant, B., Duguid, S., Hester, K., & Nelson, A. (2022). Climate conditions in the near-term, mid-term and distant future for growing soybeans in Canada. *Canadian Journal of Plant Science*, *103*(2), 161-174.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. Proceedings of the IEEE conference on computer vision and pattern recognition,

Ren, F., Zhang, Y., Liu, X., Zhang, Y., Liu, Y., & Zhang, F. (2021). Identification of Plant Stomata Based on YOLO v5 Deep Learning Model. *Proceedings of the 2021 5th International Conference on Computer Science and Artificial Intelligence*.

Rezatofighi, H., Tsoi, N., Gwak, J., Sadeghian, A., Reid, I., & Savarese, S. (2019). Generalized intersection over union: A metric and a loss for bounding box regression. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition,

Roberts, B. R. (1990). What Are Stomates and How Do They Work? *Arboriculture & Urban Forestry*.

Sanjabi, N. (2018). Abstractive text summarization with attention-based mechanism.

Serna, L., & Fenoll, C. (2000). Stomatal development and patterning in Arabidopsis leaves. *Physiologia Plantarum*, *109*, 351-358.

Siddiqui, M. H., Al-Khaishany, M. Y., Al-Qutami, M. A., Al-Whaibi, M. H., Grover, A., Ali, H. M., Al-Wahibi, M. S., & Bukhari, N. A. (2015). Response of Different Genotypes of Faba Bean Plant to Drought Stress. *International journal of molecular sciences*, *16*, 10214 - 10227.

Song, W., Li, J., Li, K., Chen, J., & Huang, J. (2020). An Automatic Method for Stomatal Pore Detection and Measurement in Microscope Images of Plant Leaf Based on a Convolutional Neural Network Model. *Forests*.

Sözen, Ö., Karadavut, U., Ozcelik, H., Bozoğlu, H., & Akçura, M. (2017). Genotype x environment interaction of some dry bean (Phaseolus vulgaris L.) genotypes. *Legume Research*, *41*, 189-195.

Specht, J. E., Diers, B. W., Nelson, R. L., de Toledo, J. F. F., Torrion, J. A., & Grassini, P. (2014). Soybean. *Yield gains in major US field crops*, *33*, 311-355.

Sultana, S. N., Park, H., Choi, S. H., Jo, H., Song, J. T., Lee, J.-D., & Kang, Y. J. (2021a). Optimizing the experimental method for stomata-profiling automation of soybean leaves based on deep learning. *Plants*, *10*(12), 2714.

Sultana, S. N., Park, H., Choi, S. H., Jo, H., Song, J. T., Lee, J.-D., & Kang, Y. J. (2021b). Optimizing the Experimental Method for Stomata-Profiling Automation of Soybean Leaves Based on Deep Learning. *Plants*, *10*.

Taheripour, F., & Tyner, W. E. (2018). Impacts of possible Chinese protection on US soybeans.

Traylor, A., Merullo, J., Frank, M. J., & Pavlick, E. (2024). Transformer Mechanisms Mimic Frontostriatal Gating Operations When Trained on Human Working Memory Tasks. *ArXiv*, *abs/2402.08211*.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*.

Wagara, I. N., Mwang'ombe, A. W., Kimenju, J. W., Buruchara, R. A., & Kimani, P. M. (2011). Reaction of selected common bean genotypes to physiological races of phaeoisariopsis griseola occuring in Kenya. *African Crop Science Journal*, *19*, 343-355.

Wang, G., Zhao, Y., Tang, C., Luo, C., & Zeng, W. (2022). When Shift Operation Meets Vision Transformer: An Extremely Simple Alternative to Attention Mechanism. *ArXiv*, *abs/2201.10801*.

White, J. W., & Castillo, J. A. (1992). Evaluation of Diverse Shoot Genotypes on selected Root Genotypes of Common Bean under Soil Water Deficits. *Crop Science*, *32*, 762-765.

Willmer, C., & Fricker, M. (1996). *Stomata* (Vol. 2). Springer Science & Business Media.

Willmer, C. M. (1993). The evolution, structure and functioning of stomata. *Botanical Journal of Scotland*, *46*, 433-445.

Woodward, F., Lake, J., & Quick, W. (2002). Stomatal development and CO 2: ecological consequences. *New Phytologist*, 477-484.

Wu, S., & Zhao, B. (2017). Using clear nail polish to make Arabidopsis epidermal impressions for measuring the change of stomatal aperture size in immune response. *Plant pattern recognition receptors: methods and protocols*, 243-248.

Yan, X., Lynch, J. P., & Beebe, S. E. (1995). Genetic Variation for Phosphorus Efficiency of Common Bean in Contrasting Soil Types: I. Vegetative Response. *Crop Science*, *35*, 1086-1093.

Yoosefzadeh-Najafabadi, M., & Rajcan, I. (2022). Six decades of soybean breeding in Ontario, Canada: a tradition of innovation. *Canadian Journal of Plant Science*, *103*(4), 333-352.

Yu, S., Qi, X., Hu, Y.-s., Zheng, W. J., Wang, Q., & Yao, H. (2019). [Overview of multivariate regression model analysis and application]. *Zhonghua yu fang yi xue za zhi [Chinese journal of preventive medicine]*, *53 3*, 334-336.

Zhang, E. Y., Cheok, A. D., Pan, Z., Cai, J., & Yan, Y. (2023). From turing to transformers: A comprehensive review and tutorial on the evolution and applications of generative transformer models. *Sci*, *5*(4), 46.

Zhang, F., Ren, F., Li, J. Q., & Zhang, X. (2021). Automatic stomata recognition and measurement based on improved YOLO deep learning model and entropy rate superpixel algorithm. *Ecol. Informatics*, *68*, 101521.

Zhou, H., Zhang, R., Lai, P., Guo, C., Wang, Y., Sun, Z., & Li, J. (2023). EL-VIT: Probing Vision Transformer with Interactive Visualization. *2023 IEEE International Conference on Data Mining Workshops (ICDMW)*, 118-127.

Zhou, T., Niu, Y., Lu, H., Peng, C., Guo, Y., & Zhou, H. (2024). Vision transformer: To discover the "four secrets" of image patches. *Inf. Fusion*, *105*, 102248.