

Innovative use of imaging techniques within plant science

Edited by

Lisbeth Garbrecht Thygesen, Gregorio Egea and Alexander Bucksch

Published in

Frontiers in Plant Science



FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-83250-952-4
DOI 10.3389/978-2-83250-952-4

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Innovative use of imaging techniques within plant science

Topic editors

Lisbeth Garbrecht Thygesen — University of Copenhagen, Denmark

Gregorio Egea — University of Seville, Spain

Alexander Bucksch — University of Georgia, United States

Citation

Thygesen, L. G., Egea, G., Bucksch, A., eds. (2022). *Innovative use of imaging techniques within plant science*. Lausanne: Frontiers Media SA.

doi: 10.3389/978-2-83250-952-4

Table of contents

- 05 **Editorial: Innovative use of imaging techniques within plant science**
Gregorio Egea, Alexander Bucksch and Lisbeth G. Thygesen
- 08 **Phenotyping Flowering in Canola (*Brassica napus* L.) and Estimating Seed Yield Using an Unmanned Aerial Vehicle-Based Imagery**
Ti Zhang, Sally Vail, Hema S. N. Duddu, Isobel A. P. Parkin, Xulin Guo, Eric N. Johnson and Steven J. Shirtliffe
- 23 **Creating High-Resolution Microscopic Cross-Section Images of Hardwood Species Using Generative Adversarial Networks**
Dercilio Junior Verly Lopes, Gustavo Fardin Monti, Greg W. Burgreen, Jordão Cabral Moulin, Gabrielly dos Santos Bobadilha, Edward D. Entsminger and Ramon Ferreira Oliveira
- 33 **Darkfield and Fluorescence Macrovision of a Series of Large Images to Assess Anatomical and Chemical Tissue Variability in Whole Cross-Sections of Maize Stems**
Marie Berger, Marie-Françoise Devaux, David Legland, Cécile Barron, Benoit Delord and Fabienne Guillon
- 55 **Using Unmanned Aerial Vehicle-Based Multispectral Image Data to Monitor the Growth of Intercropping Crops in Tea Plantation**
Yujie Shi, Yuan Gao, Yu Wang, Danni Luo, Sizhou Chen, Zhaotang Ding and Kai Fan
- 69 **Detection and localization of citrus fruit based on improved You Only Look Once v5s and binocular vision in the orchard**
Chaojun Hou, Xiaodi Zhang, Yu Tang, Jiajun Zhuang, Zhiping Tan, Huasheng Huang, Weilin Chen, Sheng Wei, Yong He and Shaoming Luo
- 85 **TomatoDet: Anchor-free detector for tomato detection**
Guoxu Liu, Zengtian Hou, Hongtao Liu, Jun Liu, Wenjie Zhao and Kun Li
- 99 **The menace of saffron adulteration: Low-cost rapid identification of fake look-alike saffron using Foldscope and machine learning technology**
Amjad M. Husaini, Syed Anam Ul Haq, Asma Shabir, Amir B. Wani and Muneer A. Dedmari
- 115 **Raman micro-spectroscopy of two types of acetylated Norway spruce wood at controlled relative humidity**
Andrea Ponzecchi, Emil E. Thybring, Ramūnas Digaitis, Maria Fredriksson, Sara Piqueras Solsona and Lisbeth Garbrecht Thygesen

- 124 **Two new methods for severity assessment of wheat stripe rust caused by *Puccinia striiformis* f. sp. *tritici***
Qian Jiang, Hongli Wang and Haiguang Wang
- 142 **Phenotyping Fusarium head blight through seed morphology characteristics using RGB imaging**
Fernanda Leiva, Mustafa Zakieh, Marwan Alamrani, Rishap Dhakal, Tina Henriksson, Pawan Kumar Singh and Aakash Chawade



OPEN ACCESS

EDITED AND REVIEWED BY
Roger Deal,
Emory University, United States

*CORRESPONDENCE

Gregorio Egea
gegea@us.es

SPECIALTY SECTION

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

RECEIVED 24 October 2022

ACCEPTED 02 November 2022

PUBLISHED 22 November 2022

CITATION

Egea G, Bucksch A and Thygesen LG
(2022) Editorial: Innovative use
of imaging techniques within
plant science.
Front. Plant Sci. 13:1079022.
doi: 10.3389/fpls.2022.1079022

COPYRIGHT

© 2022 Egea, Bucksch and Thygesen.
This is an open-access article
distributed under the terms of the
Creative Commons Attribution License
(CC BY). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Editorial: Innovative use of imaging techniques within plant science

Gregorio Egea ^{1*}, Alexander Bucksch ^{2,3,4}
and Lisbeth G. Thygesen ⁵

¹Area of Agroforestry Engineering, School of Agricultural Engineering, University of Seville, Seville, Spain, ²Department of Plant Biology, University of Georgia, Athens, GA, United States, ³Warnell School of Forestry and Natural Resources, University of Georgia, Athens, GA, United States, ⁴Institute of Bioinformatics, University of Georgia, Athens, GA, United States, ⁵Department of Geosciences and Natural Resource Management, University of Copenhagen, Frederiksberg, Denmark

KEYWORDS

imaging, non-destructive measurements, unmanned aerial vehicle (UAV), microscopy, microspectroscopy, artificial intelligence

Editorial on the Research Topic

Innovative use of imaging techniques within plant science

Several examples in the history of biology show how technological advances have facilitated fundamental discoveries in biology. The development and application of imaging techniques in plant sciences represent such an example that is currently unfolding. By using image analysis, spatially resolved information can be obtained that allows new questions in the field to be explored. Furthermore, when applied for example in crop monitoring, quality control or management, these techniques allow objective real-time decisions to be made, often based on non-destructive measurements and a reduction in time and labor that could also translate into cost savings.

This Research Topic brings together research papers that demonstrate how image-based techniques can help solve actual problems in the world of plant sciences. Generally, the presented papers offer image-based solutions to assess plant disease status, predict and detect grain and fruit yield, and analyze wood samples for their species and quality. These general application areas were achieved with a range of imaging instruments from the microscopy level to airborne image collection with unmanned aerial vehicles (UAV).

Zhang et al. (2021) tackle the long-standing and laborious yield prediction problem to precisely quantify yellowness in canola flowers. In doing so, they propose a UAV method to effectively estimate yield in Canola (*Brassica napus* L.) from airborne imagery. Their remote-sensing solution is to define a normalized yellowness vegetation index (NDYI) that demonstrated high predictive performance for seed yield.

Using similar technology, Shi et al. (2022) propose the use of UAV-based multispectral imagery and machine learning (ML) models for aboveground biomass

(AGB) and leaf area index (LAI) estimation of two intercropping species (mung bean and red bean) in tea plantations. Five ML algorithms were evaluated based on the vegetation indices derived from the UAV multispectral images as well as the actual AGB and LAI data. Their results show that two models (Support Vector Machine and Back Propagation Neural Network) outperformed the AGB and LAI prediction of red bean and mung bean as compared to other ML models.

Crop disease detection using image-based techniques is also a field that experiences growth due to the positive impact crop productivity and greater environmental and economic sustainability of agriculture. In this sense, [Jiang et al. \(2022\)](#) have conducted a study aimed at assessing the severity of wheat stripe rust using a low-cost approach based to evaluate images of infected leaves obtained by smartphones. This approach may represent a compromise between the sometimes-subjective visual disease assessment and symptoms assessment using costly devices such as multi- and hyper- spectral cameras. Along, [Leiva et al. \(2022\)](#) compared the performance of two low-cost image-based methods for predicting Fusarium Head Blight (FHB) infection in winter wheat seeds. The two analysis methods use RGB images of wheat seeds to provide various morphological traits of the seed, which were used to predict FHB using multiple regression models.

The development of robots for automatic fruit harvesting is a growing discipline due to the increasing costs of manual harvesting and the difficulty of finding skilled labor. Accurate and robust detection of fruits under natural conditions is crucial for the success of automatic fruit harvesting with robots. In this line of work, [Hou et al. \(2022\)](#) have developed a methodology based on the use of binocular cameras and deep learning to improve both citrus fruit detection and 3D localization under natural lighting conditions in commercial citrus orchards. To this end, an improved version of the YOLO v5s model is proposed for citrus detection, Cr-Cb chromatic mapping together with Otsu threshold algorithm and morphology processing are used to extract citrus shape, and a geometric image model for 3D citrus localization. [Liu et al. \(2022\)](#) present another work aimed at improving the automatic detection of fruits under natural conditions using deep learning models. In their case study, the authors have developed an anchor-free detector based on the CenterNet architecture that outperforms other tomato detection methods.

Another innovative application of the use of image-based techniques is that developed by [Husaini et al. \(2022\)](#) for the detection of fraudulent saffron. Saffron adulteration is a major problem, because Saffron is an expensive spice that is normally used as hand-picked dried flower stigmas. As a technological advancement, the authors have successfully tested two new methods for detecting adulterated saffron, one based on the use of a low-cost optical microscope (Foldscope) in combination with a chemical staining technique for visual identification of fake saffron samples, and another based on deep learning to automatically classify images taken with Foldscope and a smartphone.

[Berger et al. \(2021\)](#) report on a study in which image data obtained using darkfield and fluorescence microscopy was used to quantify the histology in cross sections of whole maize stems. This information was used for phenotyping different maize lines. The method developed makes it possible to assess unusually large cross sections, i.e., in the cm range. It is possible to quantify plant anatomy and autofluorescence after excitation with ultraviolet and/or visible light.

Determining the wood species or genus of timber and wooden artefacts based on light microscopy is important when controlling wood trade, especially to protect endangered tree species. However, wood identification is a skill that requires training and expertise, which means that far less wood is controlled than one could wish for from a conservation viewpoint. Adding to the challenge is the limited availability of microscopy images from known species in species-rich forests. [Lopes et al. \(2022\)](#) describe an exciting first step towards addressing this problem. Their approach involves neural networks to generate artificial images based on microscopy images of known species. In a second step, the method increases the number of images available per species to train neural networks to be able to identify the wood species in microscopy images of unknown species.

The article by [Ponzecchi et al. \(2022\)](#) describes a study where chemically modified wood was studied using Raman microspectroscopy. The novelty of this article lies in the development and test of a miniature climate chamber that makes it possible to adjust the relative humidity of microtomed sample sections mounted below a normal coverslip while they are presented to the instrument. In addition to the advantage of securing a well-defined and adjustable relative humidity, the setup has the advantage of being compatible with immersion objectives.

Together, the articles of this Research Topic illustrate the many useful applications that are currently being explored within this active field of research and development.

Author contributions

All authors listed above have made substantial contributions to the work. All authors have approved the final manuscript for publication.

Acknowledgments

We thank all authors and reviewers for their contributions to this Research Topic and for the support of the editorial office.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.



Phenotyping Flowering in Canola (*Brassica napus* L.) and Estimating Seed Yield Using an Unmanned Aerial Vehicle-Based Imagery

Ti Zhang¹, Sally Vail², Hema S. N. Duddu¹, Isobel A. P. Parkin², Xulin Guo³, Eric N. Johnson¹ and Steven J. Shirliffe^{1*}

¹ Department of Plant Sciences, College of Agriculture and Bioresources, University of Saskatchewan, Saskatoon, SK, Canada, ² Saskatoon Research and Development Center, Agriculture and Agri-Food Canada, Saskatoon, SK, Canada, ³ Department of Geography and Planning, University of Saskatchewan, Saskatoon, SK, Canada

OPEN ACCESS

Edited by:

Alexander Bucksch,
University of Georgia, United States

Reviewed by:

Kanichiro Matsumura,
Tokyo University of Agriculture, Japan
Muthusamy Ramakrishnan,
Nanjing Forestry University, China

*Correspondence:

Steven J. Shirliffe
steve.shirliffe@usask.ca

Specialty section:

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

Received: 26 March 2021

Accepted: 21 May 2021

Published: 17 June 2021

Citation:

Zhang T, Vail S, Duddu HSN,
Parkin IAP, Guo X, Johnson EN and
Shirliffe SJ (2021) Phenotyping
Flowering in Canola (*Brassica napus*
L.) and Estimating Seed Yield Using
an Unmanned Aerial Vehicle-Based
Imagery. *Front. Plant Sci.* 12:686332.
doi: 10.3389/fpls.2021.686332

Phenotyping crop performance is critical for line selection and variety development in plant breeding. Canola (*Brassica napus* L.) flowers, the bright yellow flowers, indeterminately increase over a protracted period. Flower production of canola plays an important role in yield determination. Yellowness of canola petals may be a critical reflectance signal and a good predictor of pod number and, therefore, seed yield. However, quantifying flowering based on traditional visual scales is subjective, time-consuming, and labor-consuming. Recent developments in phenotyping technologies using Unmanned Aerial Vehicles (UAVs) make it possible to effectively capture crop information and to predict crop yield via imagery. Our objectives were to investigate the application of vegetation indices in estimating canola flower numbers and to develop a descriptive model of canola seed yield. Fifty-six diverse *Brassica* genotypes, including 53 *B. napus* lines, two *Brassica carinata* lines, and a *Brassica juncea* variety, were grown near Saskatoon, SK, Canada from 2016 to 2018 and near Melfort and Scott, SK, Canada in 2017. Aerial imagery with geometric and radiometric corrections was collected through the flowering stage using a UAV mounted with a multispectral camera. We found that the normalized difference yellowness index (NDYI) was a useful vegetation index for representing canola yellowness, which is related to canola flowering intensity during the full flowering stage. However, the flowering pixel number estimated by the thresholding method improved the ability of NDYI to detect yellow flowers with coefficient of determination (R^2) ranging from 0.54 to 0.95. Moreover, compared with using a single image date, the NDYI-based flowering pixel numbers integrated over time covers more growth information and can be a good predictor of pod number and thus, canola yield with R^2 up to 0.42. These results indicate that NDYI-based flowering pixel numbers can perform well in estimating flowering intensity. Integrated flowering intensity extracted from imagery over time can be a potential phenotype associated with canola seed yield.

Keywords: canola, flowering, seed yield, multispectral camera, remote sensing

INTRODUCTION

Canola (*Brassica napus* L.) is the predominant oilseed crop grown in Canada (Clayton et al., 2000). Canada has the largest area of canola production in the world (Statistics Canada, 2018). With the growing global demand for canola, Canada needs to maintain and improve canola yield and seed quality to meet the market demands. Yield components of canola consist of the number of pods, the seeds per pod, and the weight per seed (Tayo and Morgan, 1975; McGregor, 1981; Diepenbrock, 2000; Ivanovska et al., 2007; Faraji, 2012). Among these components, pod number retained at maturity is the most important factor as it is influenced most by environmental constraints (Tayo and Morgan, 1975; McGregor, 1981; Diepenbrock, 2000; Ivanovska et al., 2007; Faraji, 2012; Gan et al., 2016; Kirkegaard et al., 2018). The flowering stage in canola is important for yield estimation as flowers produced in the first 2–3 weeks from anthesis contribute to 75% of the pods at maturation (Tayo and Morgan, 1975). Additionally, the flowering period can last from 2 to 6 weeks, which is a major portion of the crop growth cycle (Gan et al., 2016; Kirkegaard et al., 2018). Thus, flower production is one of the most important factors in determining final seed yield (Tayo and Morgan, 1975; Diepenbrock, 2000; Faraji et al., 2008; Faraji, 2012; Fang et al., 2016; Gong et al., 2018; Kirkegaard et al., 2018; Zhang and Flottmann, 2018).

During the plant breeding process, field-based phenotyping plays an important role in evaluating plant performance. It contributes to the selection of ideal genotypes that are high-yielding by associating genotype with the corresponding phenotype (Montes et al., 2007; Sankaran et al., 2015). To select better canola lines and eventually develop better varieties, breeders need to assess many distinct lines grown in multiple environments to detect interactions between genotype and environment (White et al., 2012; Araus and Cairns, 2014).

The quantification of flowering intensity based on traditional visual scales is subjective, labor-consuming, and is often destructive (Sulik and Long, 2015; Fang et al., 2016; Wan et al., 2018). Although ground-based platforms such as Greenseeker, Crop Circle, or time-lapse RGB imaging can provide adequate spectral data, these platforms still require a prohibitive amount of time and labor (Xu et al., 2018; Hassan et al., 2019). Additionally, data collection using these ground-based platforms may cause soil compaction and crop canopy damage (Xu et al., 2018). Therefore, it is necessary to develop an objective, non-destructive, and efficient method to estimate flower numbers. With this, one can model seed yield by assessing real-time radiometric data of the crop canopy, which has the potential to accelerate breeding methods for yield improvement. Current improvements in aerial-based platforms and sensors equipped on aerial platforms make it possible to effectively collect phenotypes via analyzing digital imagery (Kim et al., 2019). Unmanned aerial vehicles (UAVs) equipped with various sensors can quickly provide large quantities of field data enabling plant breeders to efficiently detect traits of numerous plots in large-scale field trials (Kefauver et al., 2017).

Spectral reflectance of the crop canopy is strongly correlated with morphological and physiological traits. Leaf composition

and molecular structure can affect the reflectance of the crop; thus, ratios or differences of different bands in the visual light, near IR (NIR), and shortwave IR wavelengths (i.e., vegetation indices) can be a tool to characterize plant traits (Sankaran et al., 2015; Wójtowicz et al., 2016). Previous studies have shown that multispectral reflectance profiles of visible bands (i.e., blue, green, and red) and NIR bands could estimate canopy features, such as nitrogen use efficiency (Kefauver et al., 2017; Prey et al., 2020), leaf area index (Tunca et al., 2018; Blancon et al., 2019), and flower numbers (Guo et al., 2015; Sulik and Long, 2015, 2016; Carl et al., 2017; Gong et al., 2018; Wan et al., 2018; Xu et al., 2018). These plant traits investigated remotely have the potential to improve yield estimates. Flower numbers, as an important factor in determining crop yield, have exhibited close correlations with optical properties in various crops, such as rice (Guo et al., 2015), cotton (Xu et al., 2018), and canola (Sulik and Long, 2015, 2016; Gong et al., 2018; Wan et al., 2018). Guo et al. (2015) applied a machine learning model, the support vector machine, for flowering quantification using RGB images in rice, which resulted in a good correlation between the actual rice flowering panicles and identified flowering (correlation coefficients ranging from 0.64 to 0.82) (Guo et al., 2015). In canola, there are three different canopy morphologies during the growing season, namely, the vegetative phase (green canopy dominated by leaves), the flowering phase (yellow canopy dominated by the yellowness of flower petals), and the mature phase (green or brown canopy because of pods and branches) (Sulik and Long, 2016). During the flowering phase, the yellowness of canola petals is due to carotenoid absorption of blue and reflectance of a mixture of green and red wavelengths (Sulik and Long, 2015, 2016), but the yellow color has little impact on red edge and NIR reflectance unlike a green vegetative canopy (Shen et al., 2009; Migdall et al., 2010; Sulik and Long, 2015, 2016). Thus, the contributed red light decreases the normalized difference vegetation index (NDVI) values (Equation 1) and adversely impact the ability of NDVI to monitor crop growth condition and estimate yield during the flowering phase (Shen et al., 2009, 2010; Sulik and Long, 2015, 2016). Sulik and Long (2015) found that the ratio of green and blue was strongly correlated with the actual flower numbers with a coefficient of determination (R^2) of 0.87, and they proposed the plot-level normalized difference yellowness index (NDYI) (Equation 2) could be a potential yield predictor ($R^2 = 0.76$) (Sulik and Long, 2016). d'Andrimont et al. (2020) and Han et al. (2021) reported that NDYI successfully captured canola yellowness and detected the peak flowering dates using Sentinel-2 imagery. Fang et al. (2016) found that reflectance at 550 nm was the most sensitive band to estimate flowering coverage with an estimation error below 6% when compared with wavelengths at 490, 670, 720, 800, and 900 nm. Wan et al. (2018) and Gong et al. (2018) reported that combining vegetation index and image classification methods (i.e., k-means clustering method by CIE L^*a^*b space and pixel-level spectral mixture analysis) improved the accuracy of flower numbers and yield estimation in canola with R^2 values of 0.89 and 0.75, respectively.

Although several studies have detected canola flowering number and predicted yield, most of these field experiments were

TABLE 1 | Summary of canola trials and data collection (imagery acquisition and manual flower count) at Saskatoon, SK, Canada from 2016 to 2018 and at Melfort and Scott, SK, Canada in 2017.

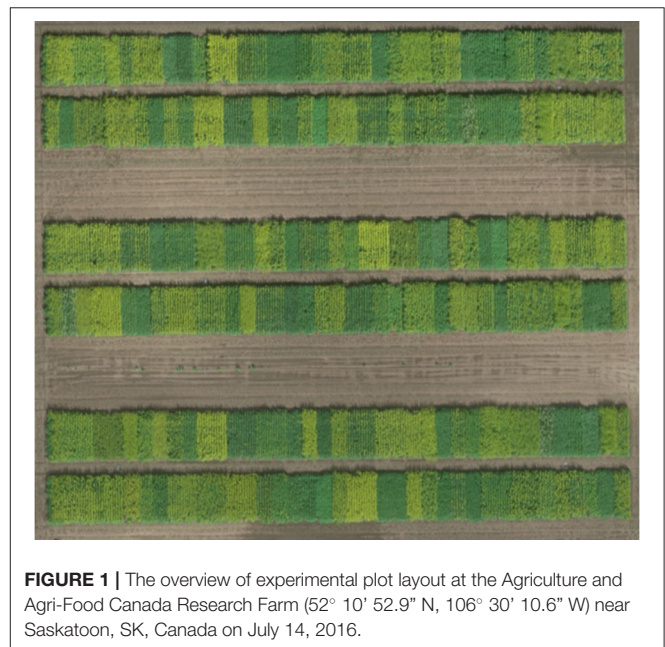
Site	Year	Seeding date	Number of lines/cultivars	Flight altitude (m)	Image acquisition dates	Manual flower count dates
Saskatoon	2016	May 27	56	20	July 14; 19; 26 August 06	July 15; 22; 29 August 05
	2017	May 28	56	20	July 07; 11; 15; 19; 22; 26 August 01; 09; 16; 22	July 01; 18; 25 August 01
	2018	May 21	56	25	June 28 July 06; 09; 16; 20; 24; 27; 30 August 03; 07	July 10; 17; 24; 31
Melfort	2017	May 18	16	15	July 05; 13; 20; 26	July 05; 20; 26
Scott	2017	June 20	16	20	August 09; 16; 29	August 09; 16

conducted with relatively few canola lines and environments, which may neglect the effect of genotype and environmental fluctuations on yellowness of flower (Ohmiya, 2011) and petal size (Jiang and Becker, 2003). In addition, yield estimation models used in those studies were based on only one image date (Sulik and Long, 2016; Gong et al., 2018), which ignores the effect of time and duration of flowering (Tayo and Morgan, 1975). Thus, the reflectance data of flowering throughout the flowering period may provide a better estimate of crop yield. Therefore, the objectives of this study were to use UAV multispectral data to detect flowers within a wide range of canola lines and to estimate seed yield in canola using time series imagery collected during the flowering period.

MATERIALS AND METHODS

Experimental Sites and Plant Materials

The experiment was conducted at the Agriculture and Agri-Food Canada Research Farm near Saskatoon, SK, Canada from 2016 to 2018 (52° 10' 52.9" N, 106° 30' 10.6" W in 2016; 52° 10' 59.3" N, 106° 30' 53.7" W in 2017; and 52° 10' 57.7" N, 106° 30' 01.4" W in 2018), and near Melfort (52° 49' 9.6" N and 104° 35' 46.9" W) and Scott (52° 21' 55.3" N and 108° 52' 32.6" W), SK, Canada in 2017 (Table 1). The soil texture at Saskatoon was a clay loam with a pH of 7.3 and an organic matter content of 5.5%. The field plots were set up in a randomized incomplete block design (rectangular lattice design) with three replications (Figure 1). A rectangular lattice design was used to reduce spatial variability within each block. Individual plot size was 6.0 m long × 1.2 m wide in 2016 and 2018 and 6.0 m long × 1.5 m wide in 2017. Fifty-six genotypes (Saskatoon Research and Development Center, Agriculture and Agri-Food Canada), including 53 diverse *B. napus* lines, two *B. carinata* lines, and a *B. juncea* variety, were selected and planted. Fifty of the diverse lines were used as founders to develop Nested Association Mapping (NAM) population by developing population from crossing to a common reference line (Parkin et al., 2017). This panel, which represents diverse germplasm resources and the historical basis of canola breeding programs, differs in geographic origin, pedigree, phenotypes, and genotype (Parkin et al., 2017). Seeding occurred on May 27, 28, and 21 in 2016, 2017, and 2018, respectively,

**FIGURE 1** | The overview of experimental plot layout at the Agriculture and Agri-Food Canada Research Farm (52° 10' 52.9" N, 106° 30' 10.6" W) near Saskatoon, SK, Canada on July 14, 2016.

at a seeding rate of 108 seeds m⁻² (Table 1). Out of 56 lines, 16 were selected and planted twice in two adjacent but separate plots as double plots. The criteria of line selection for the double plots were based on contrasting seed quality (i.e., seed color, acid detergent lignin, seed glucosinolates, and seed erucic acid) and similarity in flowering timing. The reason for setting double plots was to preserve one plot for imaging without any subsamples being removed. The 16 *B. napus* lines planted in double plots were YN04-C1213, NAM-0, 5, 13, 14, 17, 23, 30, 37, 32, 43, 46, 48, 72, 76, and 79.

The selected 16 *B. napus* lines were planted in a randomized complete block design with three replications at the Melfort and Scott locations in 2017. All lines were planted in 5 m long × 1.2 m wide plots at Melfort and in 5 m long × 1 m wide plots at Scott. Canola was seeded on May 18 at Melfort and June 20 at Scott at a desired seeding rate of 108 seeds m⁻² (Table 1). Edge® (ethalfuralin) was applied as a pre-emergence herbicide at a rate

TABLE 2 | Basic specifications for the multispectral camera (RedEdge) equipped on the unmanned aerial vehicle (UAV) platforms.

GSD ^a (cm per pixel) (per band)	Flight altitude (m)	Sensor resolution per band (MP) ^b	Focal length (mm)	Full width at half maximum (nm)	Peak wavelength (nm)
1.02	15	1.2 ^c	5.5	Blue: 465–485	Blue: 475
1.36	20			Green: 550–570	Green: 560
1.70	25			Red: 663–673	Red: 668
				Red edge: 712–722	Red edge: 717
				NIR: 820–860	NIR: 840

^aGSD: ground sampling distance.

^bMP: megapixel.

^cimage resolution: 1.2 MP = 1,280 × 960 pixels.

of 19.1 kg ha⁻¹ to control weeds. Any weeds not controlled by the herbicides were removed by hand.

Image Acquisition Platform and Sensor

The UAV used in this study was a Draganflyer X4-P model in 2016 and 2017 (DraganFly Inc., Saskatoon, SK, Canada). It is a rotary-wing platform with a maximum payload of 800 g. It can semiautomatically depart and land based on GPS navigation mode and optional Surveyor software. Flight mission was planned in Surveyor software (DraganFly Inc., Saskatoon, SK, Canada) by importing ground coordinates of the field boundaries. The other rotary-wing platform was a Draganflyer Commander model (DraganFly, Inc., Saskatoon, SK, Canada), used in 2018, which differs from the X4-P model in its maximum payload capacity (1,000 g).

A multi-spectral camera (RedEdge, MicaSense Inc., Seattle, WA, United States) was used to acquire images (12-bit image) with an image resolution of 1.2 megapixels (1,280 × 960 pixels) for each of five spectral bands (blue: 475 ± 10 nm; green: 560 ± 10 nm; red: 668 ± 5 nm; red edge: 717 ± 5 nm; and near-infrared: 840 ± 20 nm) (Table 2). The focal length of the camera is 5.5 mm and the ground sampling distance at 15, 20, and 25 m above ground level was 1.02, 1.36, and 1.70 cm per pixel, respectively (Table 2). Images of a MicaSense reflectance panel (RedEdge, MicaSense Inc., Seattle, WA, United States) were taken before and after each UAV flight for radiometric calibration. To geo-reference aerial images, six ground control points (GCPs) were distributed across the experimental area during the whole crop season in 2016 at Saskatoon. The size of the GCPs was 60 × 60 cm, which were geolocated by Trimble GeoExplorer 2008 GPS (Trimble Inc., Westminster, CO, United States). GCPs were manually placed at the same location when phenotyping canola by UAV, which provided an overlay of images taken from various dates and reduced workload by using the same geolocation information for each GCP. For the four locations in 2017 and 2018, GCPs were permanently mounted within guard plots to avoid manually carrying GCPs to the field.

UAV Flight Schedule

The UAV, equipped with a multispectral camera, captured the images of the fields taken weekly during the flowering stage at

Saskatoon in 2016 and at Melfort and Scott in 2017 (Table 1). The imagery was collected semiweekly in 2017 and 2018 at Saskatoon for the duration of canola flowering (Table 1). For the Saskatoon location, although weather conditions such as rain, clouds, and heavy wind limited the flight schedule, image timing interval was achieved as close to 7 days in 2016 and to 4 days in 2017 and 2018. For the Melfort and Scott locations in 2017, image collection was carried out at a 7-day interval.

Image Process and Data Extraction Image Pre-process

Multispectral images were processed, stitched, and calibrated in Pix4Dmapper Pro (Pix4D Inc., San Francisco, CA, United States). Individual images were aligned based on common points from the overlapped images to generate a geo-referenced image that matched the overflowed study area. Geometric calibration was done by importing the geo-location of GCPs to reduce geometric distortion problems of the camera. A system coordinate, World Geodetic System 1984, was applied to generate geo-referenced images. The images of the MicaSense reflectance panel were used in the radiometric calibration to enhance spectral consistency between different flight dates. Then, the five generated reflectance maps were exported and used for further analysis.

Vegetation Index Calculation, Thresholding, and Integration of Flowering Progress

ArcGIS software version 10.4.1 (ESRI Canada, Toronto, ON, Canada) was applied for plot segmentation, vegetation indices calculation, and thresholding. In this study, the middle three rows for each plot were segmented using polygon shapes with assigned plot numbers. The polygon shapes were generated using the “Create Feature” tool. Vegetation index maps were derived *via* calculation of the reflectance maps using the “Rater calculator” tool. Commonly used vegetation indices, NDVI (Rouse et al., 1974), NDYI (Sulik and Long, 2016), green normalized difference vegetation index (GNDVI) (Gitelson et al., 1996), and normalized difference red edge index (NDRE) (Gitelson and Merzlyak, 1997), were calculated as following equations to compare with the actual flower

number counts:

$$NDVI = \left(\frac{R_{NIR} - R_{red}}{R_{NIR} + R_{red}} \right) \quad (1)$$

$$NDYI = \left(\frac{R_{green} - R_{blue}}{R_{green} + R_{blue}} \right) \quad (2)$$

$$GNDVI = \left(\frac{R_{NIR} - R_{green}}{R_{NIR} + R_{green}} \right) \quad (3)$$

$$NDRE = \left(\frac{R_{NIR} - R_{rededge}}{R_{NIR} + R_{rededge}} \right) \quad (4)$$

where R_{NIR} , R_{red} , R_{green} , R_{blue} , and $R_{rededge}$ are the reflectance values at bands centered on 840, 668, 560, 475, and 717 nm, respectively (Table 2). NDVI is the most commonly used vegetation index to identify crop growth conditions and yield estimation (Rouse et al., 1974). NDYI has previously shown a strong correlation with seed yield (Sulik and Long, 2016). GNDVI (Gitelson et al., 1996) and NDRE (Gitelson and Merzlyak, 1997) are related to photosynthesis and have been reported in previous research.

Canola flowers and leaf organs co-existed within each plot during flowering; thus, the “Conditional Function” [Con (index map > threshold value, index map, “”)] in the “Raster Calculator” tool was used to separate flowering pixels from non-flowering pixels by applying threshold values on vegetation index maps. Threshold values were manually determined by comparing the composited RGB images with calculated index maps so that most flowering pixels could be selected and segmented. All pixels in the index map that have values larger than the threshold values were kept in a threshold index map, otherwise, pixels were discarded. Then, the “Zonal Statistics” tool was used to extract the summary statistics of the threshold index map, which included the number of flowering pixels per plot.

This study involved 56 diverse lines with a high flowering density gradient. It is difficult to determine which image date is proper for yield estimation. For this reason, the area under the flowering progress curve (AUFPC) was used to calculate the integration of flowering progress during the flowering season using the following equation:

$$AUFPC = \left(\frac{F_1 + F_2}{2} - F_1 \right) (t_2 - t_1) + \left(\frac{F_2 + F_3}{2} - F_1 \right) (t_3 - t_2) + \dots + \left(\frac{F_{n-1} + F_n}{2} - F_1 \right) (t_n - t_{n-1}) \quad (5)$$

where F_1 , F_2 , F_3 , F_{n-1} , and F_n represent the flowering pixel numbers at each image date and t_1 , t_2 , t_3 , t_{n-1} , and t_n represent Julian date at each image timing. The AUFPC is an adjusted integration equation based on the area under the disease progress curve (AUDPC), which is used in general in pathology studies for estimating the effect of disease progression on crop yield (Jeger and Viljanen-Rollinson, 2001; Simko and Piepho, 2012). Compared with AUDPC, the advantage of the AUFPC is providing a baseline for each line to adjust flowering progress, which can reduce the effect of diverse initial flowering

pixel numbers of each line on the calculated area. The AUFPC equation converted several flowering pixel numbers at a series of image timings into a single value for reporting. The larger the AUFPC value is, the further the flowering had progressed. Figure 2 displays an example of flowering progress over time for a line (NAM-23). Seven data points on the curve line represent NDYI-based pixel numbers for each image date. Pictures under the seven points are corresponding threshold index maps. Then, the area under the curve line was calculated using the AUFPC equation (Equation 5) for NAM-23. The same mathematical method was used to calculate flowering progress for all other lines across 5 site years.

Ground Reference Data/Field Data Collection

The first row of each plot was manually sampled to quantify flowering. Canola flowering typically starts in early July and ends in early August. Flower numbers on the main stem and branches of randomly selected plants were counted at a 7-day interval from July to August. Grain yield was straight combined by a small plot combine harvester when the crop was mature and dry. This occurred multiple times due to differing maturity dates of the *B. napus* lines. To reduce the edge effect, the middle four rows of each plot were harvested. All harvested seeds were air-dried to 10% seed moisture. Final yields were weighed after seed cleaning.

Statistical Analysis

The PROC LATTICE procedure of SAS version 9.4 (SAS Institute, Cary, NC, United States) was used to analyze the data. The LATTICE procedure reduced variations within blocks. After data adjustment, PROC REG in SAS version 9.4 was used as the statistical tool to investigate the simple linear regressions between ground reference data and imagery. Scatterplots of variables were observed to determine whether data could be combined for analysis. In the case where data could not be combined, data were analyzed within site years.

RESULTS AND DISCUSSION

Regression Between Flowering Pixel Number and Actual Flower Numbers

These initial results showed that GNDVI and NDRE did not demonstrate significant correlations with the actual flower count ($P > 0.05$, data not shown). Meanwhile, regression results showed that NDYI had greater coefficients of determination (R^2) than NDVI with actual flower numbers within 3 years of study. An increased red light from the yellow petals can reduce NDVI values and affect its ability to detect canola growth conditions. In addition, there was no strong relationship between plot-level NDYI and actual flower numbers in 2016 (data not shown). Noise from soil background and green vegetation within a plot at the early flowering stage may have resulted in these weak relationships. For this reason, we used NDYI maps to extract flowering pixels and remove non-flowering pixels by the thresholding method. We detected and counted flowering pixels when pixel values were greater than NDYI-based threshold levels.

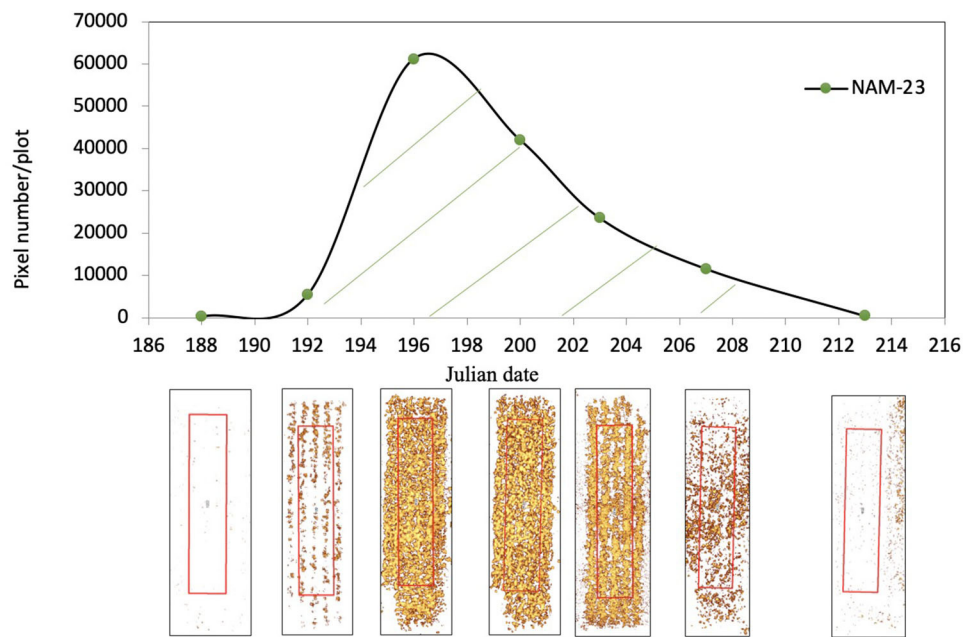


FIGURE 2 | The growth pattern of flowering progress for a *Brassica napus* line (NAM-23) during the flowering stage at the Agriculture and Agri-Food Canada Research Farm (52° 10' 59.3" N, 106° 30' 53.7" W) near Saskatoon, SK, Canada in 2017. The x-axis is the imagery acquisition date (Julian date) in 2017. The y-axis is normalized difference yellowness index (NDYI)-based pixel number per plot. A solid curve line is the flowering progress trend of NAM-23. Seven points on the progress curve line represent NDYI-based pixel number per plot at seven imagery acquisition dates. Seven pictures under each point are corresponding false-color images after thresholding with flowers highlighted in yellow. The region of interest was highlighted in red.

Threshold values were 0.59, 0.52, and 0.45 in 2016, 2017, and 2018, respectively.

Across 5 site years, the R^2 values between flowering pixel numbers and actual flower numbers ranged from 0.54 to 0.95 during flowering duration (Figures 3–7). There were significant relationships between flowering pixel numbers and actual flower numbers in 2016 at Saskatoon (Figure 3). Not surprisingly, the early flowering stage (July 15) had the strongest regression relationship with actual flower numbers with an R^2 of 0.85 (Figure 3A). Developing flowers were on the upper part of a plant at the early flowering stage so sensors could easily detect these early-blooming flowers. Whereas, the late flowering stages (August 05) showed the weakest regression ($R^2 = 0.54$) (Figure 3D), which may be a result of the lower sensitivity of NDYI to differentiate yellow flowers and dark green pods. Dark green pods impart more green reflectance, which can make NDYI less sensitive to yellow flowers, as yellow is a composite color of green and red (Yates and Steven, 1987; Sulik and Long, 2015, 2016). Additionally, the potential reason why it had the smallest R^2 value is that many flowers growing on the lower branches adversely affected the ability of the sensor to detect the late-developing flowers.

The Saskatoon location in 2017 and 2018 had similar regression patterns between flowering pixel numbers and actual flower numbers (Figures 4, 5). There were very strong relationships at the early flowering stages (July 10, 2017 and July 17, 2018) (Figures 4, 5). Similar to 2016, the relationships became weaker with the late flowering stages (August 01, 2017 and July

31, 2018) (Figures 4, 5). Although the late flowering stages had weaker regressions compared with the early flowering timing, the regressions at the peak flowering dates (July 25, 2017 and July 24, 2018) were relatively strong (Figures 4, 5).

For the Melfort location in 2017, the first image date (July 05) had the weakest regression ($R^2 = 0.71$) (Figure 6A). Variability from subsampling plants can be a potential reason for decreased regressions at the early flowering stage. However, the peak flowering time (July 20) and late flowering stage (July 26) showed strong relationships with the value of R^2 of up to 0.91 (Figures 6B,C). The potential reason why this site year had a greater R^2 at the late flowering stage is that flight altitude (15 m) at Melfort in 2017 was lower than the other site years (Table 1). The high resolution may have increased the ability of the sensor to detect flowers growing lower in the canopy. Although the flight altitude was relatively low compared with other locations, there was no significant canopy movement due to the UAV platform. The seeding date at Scott was June 22, 2017. Flowering started relatively late with a shorter duration compared with other site years. There was no imagery collected at the end of the flowering stage, and thus, those relationships are unknown. At Scott, the R^2 values for the regressions between flowering pixel numbers and actual flower numbers followed similar patterns as the Saskatoon location. The early flowering stage (August 09) and the peak flowering time (August 16) had strong relationships (Figures 7A,B).

In this study, we used a zero-intercept linear regression model in the regression analysis as there was no flowering

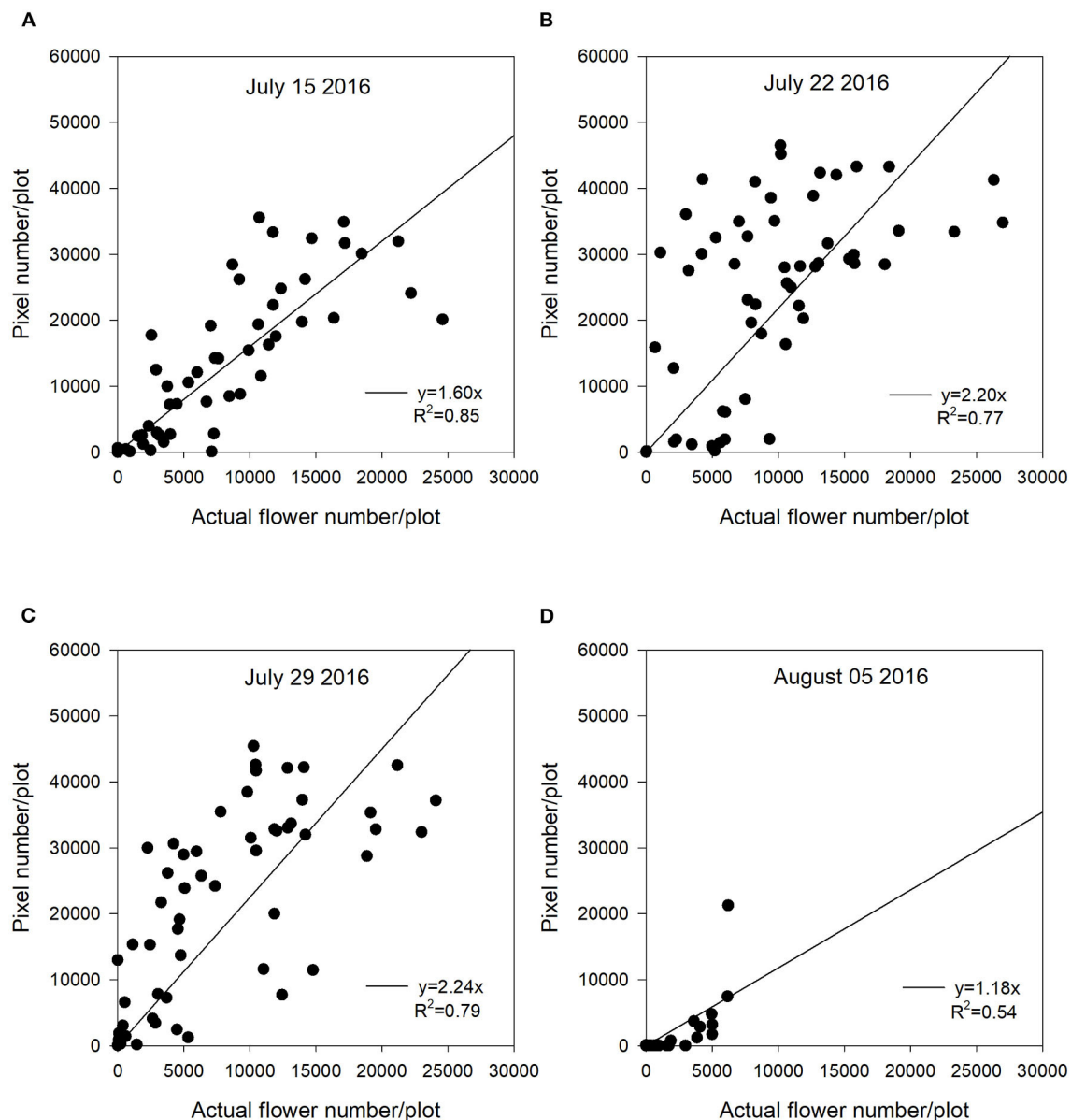


FIGURE 3 | The relationship between actual flower numbers per plot and pixel numbers extracted from aerial images during the flowering stage at Saskatoon, SK, Canada in 2016. Actual flower numbers per plot were manually measured. Pixel number per plot was detected by the thresholding method. **(A)** Regression equation for July 15, 2016: $y = 1.60x$, $R^2 = 0.85$. **(B)** Regression equation for July 22, 2016: $y = 2.20x$, $R^2 = 0.77$. **(C)** Regression equation for July 29, 2016: $y = 2.24x$, $R^2 = 0.79$. **(D)** Regression equation for August 05, 2016: $y = 1.18x$, $R^2 = 0.54$.

pixel prior to the commencement of flowering. Furthermore, the fitted intercept values were close to zero in most cases. For the Saskatoon location over 3 years, slopes were relatively consistent at the early flowering stages (Figures 3–5). Slope values became smaller with the delayed flowering stage. There was a smaller slope value at the late flowering stage (slope = 1.18) compared with the peak flowering time (slope = 2.20) at Saskatoon in 2016. The Saskatoon location in 2017 and 2018 had similar patterns (Figures 4, 5). The Melfort location had similar patterns with a smaller slope at the late flowering stage (Figure 6), but the

slope of the first image date (slope = 3.70) was greater than the other image dates. This indicated that early flowering imagery overestimated the actual flower numbers. Experimental plots at this location showed non-uniform flowering with fewer flowers at the front of each plot, which may be caused by the edge effect. Thus, manual flower count based on subsampling plants at the front row of a plot may not accurately represent the average flower numbers. In 2017, at Scott, slopes were consistent at the early and the peak flowering times (Figure 7). The slope values at this location were smaller than the other site years. A

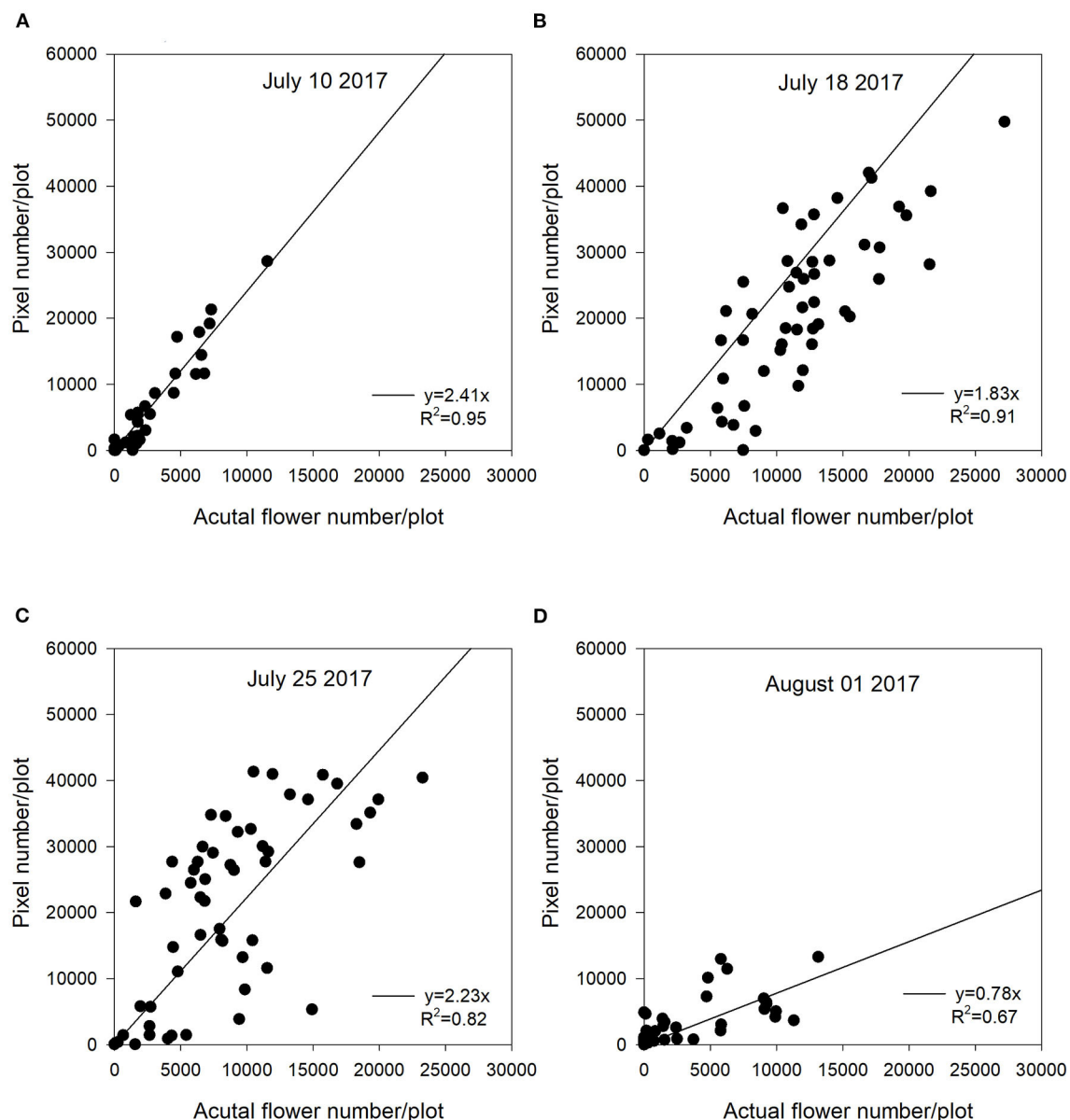


FIGURE 4 | The relationship between actual flower numbers per plot and pixel numbers extracted from aerial images during the flowering stage at Saskatoon, SK, Canada in 2017. Actual flower numbers per plot were manually measured. Pixel number per plot was detected by the thresholding method. **(A)** Regression equation for July 10, 2017: $y = 2.41x$, $R^2 = 0.95$. **(B)** Regression equation for July 18, 2017: $y = 1.83x$, $R^2 = 0.91$. **(C)** Regression equation for July 25, 2017: $y = 2.23x$, $R^2 = 0.82$. **(D)** Regression equation for August 01, 2017: $y = 0.78x$, $R^2 = 0.67$.

potential reason for this underestimation of flower numbers is that the plots had a more condensed canopy and there were more branches at this site year than other site years (data not shown) due to poor emergence percentage. Thus, for the Scott location, there were more flowers produced on the lower branches which could not be detected by the sensor. As mentioned above, there was no available data collected at the end of flowering; thus, the relationship at this stage is unknown.

In general, although the linear regression slopes varied across site years, the high R^2 values indicated that the flowering pixel

numbers extracted from the threshold NDVI map performed well to predict actual flower numbers at the early and peak flowering stages in canola (R^2 up to 0.95). These results were consistent with that reported by Sulik and Long (2015), wherein the ratio of blue and green strongly correlated with the yellow flowers in canola with a significant R^2 value of 0.87 at the full flowering stage. Wan et al. (2018) reported good estimation for the flowering number of canola using the k-means clustering method based on the CIE L^*a^*b space model during the full flowering period. Xu et al. (2018) found that white cotton flowers

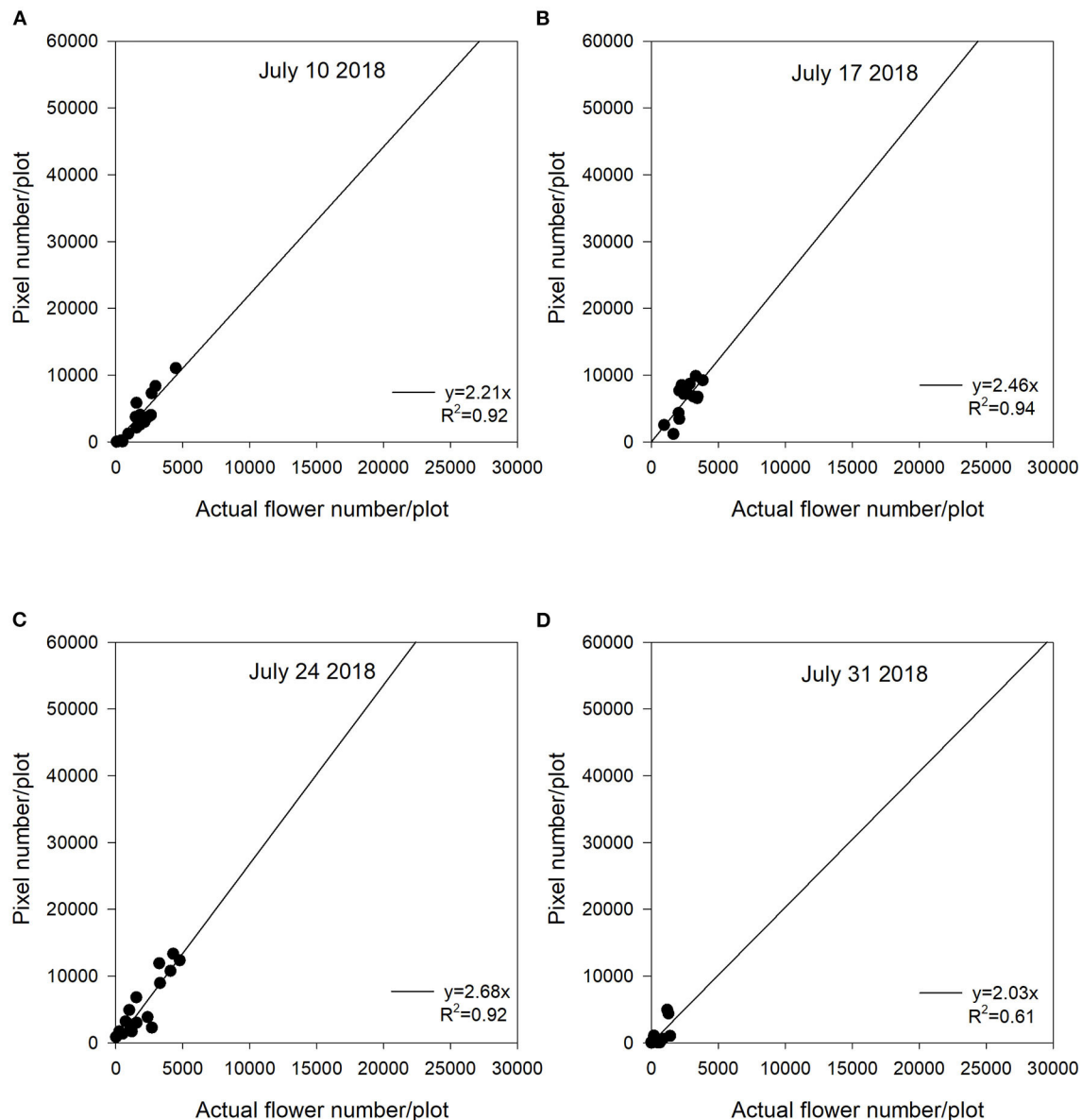


FIGURE 5 | The relationship between actual flower numbers per plot and pixel numbers extracted from aerial images during the flowering stage at Saskatoon, SK, Canada in 2018. Actual flower numbers per plot were manually measured. Pixel number per plot was detected by the thresholding method. **(A)** Regression equation for July 10, 2018: $y = 2.21x$, $R^2 = 0.92$. **(B)** Regression equation for July 17, 2018: $y = 2.46x$, $R^2 = 0.94$. **(C)** Regression equation for July 24, 2018: $y = 2.68x$, $R^2 = 0.92$. **(D)** Regression equation for July 31, 2018: $y = 2.03x$, $R^2 = 0.61$.

had higher prediction accuracy at the early flowering stage. The lower classification accuracy at the later growth stage may have resulted from coverage of leaves which increased misclassified non-flowers when using a convolutional neural network (Xu et al., 2018). They recommended that using one raw image might solve this issue, as more cotton flowers would be detected from different perspectives. Moreover, the early flowering stages across 5 site years showed greater slope values, as most flowers at this early stage were visible and had less overlap. In contrast, flowers growing on lower branches were likely to be underestimated at the late flowering stages. Subsampling variability may make the

actual flower count non-representative for a plot, which may reduce the accuracy of flower estimation.

Yield Estimation Using Integrated Flowering Accumulation During Flowering Period

Flowering pixel numbers derived from the threshold NDVI map were able to estimate actual flower numbers across 5 experimental site years. Initially, we did regression analysis between yield and flowering pixel numbers at each image date.

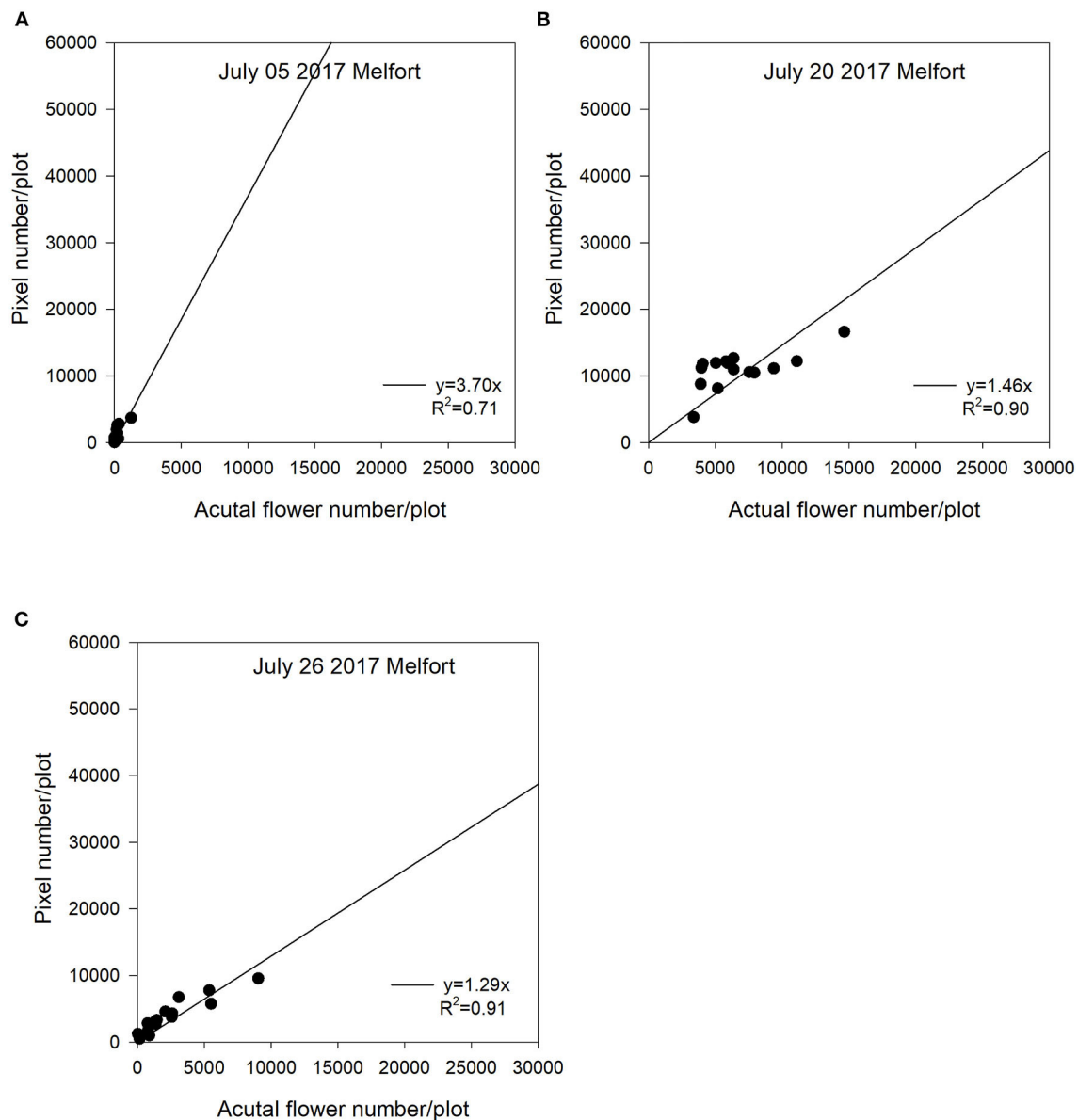


FIGURE 6 | The relationship between actual flower numbers per plot and pixel numbers extracted from aerial images during the flowering stage at Melfort, SK, Canada in 2017. Actual flower numbers per plot were manually measured. Pixel number per plot was detected by the thresholding method. **(A)** Regression equation for July 05, 2017: $y = 3.70x$, $R^2 = 0.71$. **(B)** Regression equation for July 20, 2017: $y = 1.46x$, $R^2 = 0.90$. **(C)** Regression equation for July 26, 2017: $y = 1.29x$, $R^2 = 0.91$.

Among the 5 site years, in most cases, there were no significant relationships until the middle of July when most varieties started blooming (Table 3). In addition, it is difficult to determine a single well-defined image time for crop yield estimation because of various environmental fluctuations and various flowering timings in large-scale breeding programs, especially involving many diverse lines. Furthermore, we may miss important flowering progress information if yield estimation is only based on the imagery from a single date (Haynes and Weingartner, 2004; Gan et al., 2016). Although flower formation at the later stage may contribute less than early timing points, they may

still have the potential to increase final grain yield. Therefore, integrating all aspects of the entire flowering duration using AUFPC can reflect flowering accumulation progress and improve the accuracy of crop yield estimation.

We found significant relationships between integrated flower accumulation and yield during the flowering period (Figures 8, 9). In 2016, at Saskatoon, integrated flower accumulation had a moderate relationship with yield ($R^2 = 0.12$, $P < 0.05$) (Figure 8A). We calculated the flower accumulation progress by integrating the flowering pixel numbers over four image dates at a 7-day interval, which missed

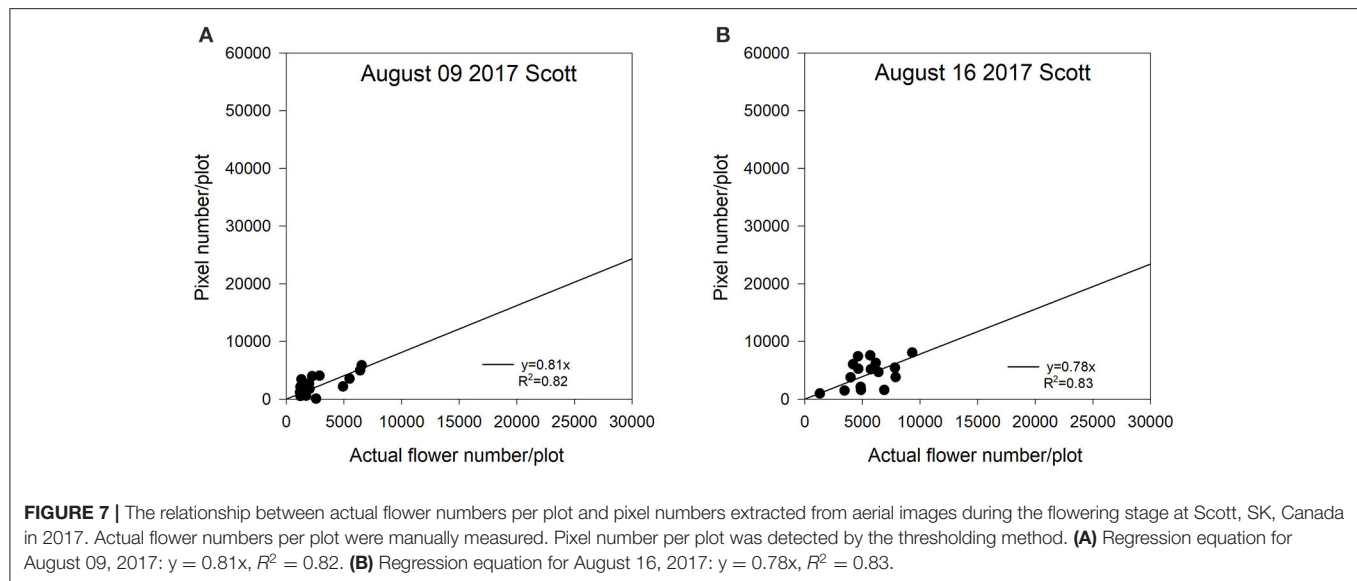


TABLE 3 | The coefficient of determination (R^2) between flowering pixel numbers from a single image date and yield at Saskatoon, SK, Canada from 2016 to 2018 and at Melfort and Scott, SK, Canada in 2017.

Site	2016	R^2	2017	R^2	2018	R^2
Saskatoon	July 14	0.04	July 07	0.02	June 28	<0.01
	July 19	<0.01	July 11	<0.01	July 06	0.02
	July 26	0.02	July 15	0.04	July 09	0.06
	August 06	0.04	July 19	0.29***	July 16	0.36***
			July 22	0.33***	July 20	0.22***
			July 26	0.06	July 24	0.07
			August 01	0.06	July 27	<0.01
			August 09	0.02	July 30	0.03
			August 16	0.05	August 03	0.03
			August 22	0.05	August 07	0.02
Melfort			July 05	<0.01		
			July 13	0.23*		
			July 20	0.02		
			July 26	0.14		
Scott			August 09	0.46**		
			August 16	0.32*		
			August 29	0.01		

*Significant at the 0.05 probability level.

**Significant at the 0.01 probability level.

***Significant at the 0.001 probability level.

the starting point of the flowering period. There was no adequate imagery data for the entire flowering period, so it may be the reason for the low accuracy of yield estimation. In both 2017 and 2018 at Saskatoon, we collected imagery semiweekly (Table 1). For the 2 site years, the relationships between integrated flower accumulation and seed yield were relatively stronger compared to the 1st experimental year ($R^2 = 0.30$, $P < 0.05$ in 2017; $R^2 = 0.34$, $P < 0.05$ in 2018) (Figures 8B,C). At the Melfort and Scott locations in 2017, there were more consistent and stronger regressions (Figure 9) using the integration of flowering progress, when compared with a single image date (Table 3).

In general, compared with using a single image, applying the integration of flowering progress to estimate yield includes more information to provide consistent accuracy (Figures 8, 9). Although the R^2 values for yield estimation are not very high, our results still demonstrate potential ability of AUFPC to predict yield, especially for crops producing bright flowers (e.g., canola and cotton) under different environmental conditions.

Several studies have reported similar results (Sulik and Long, 2016; Gong et al., 2018; Xu et al., 2018; Hassan et al., 2019). Sulik and Long (2016) reported that the plot-level NDVI values during flowering had high accordance with field yield observations (R^2

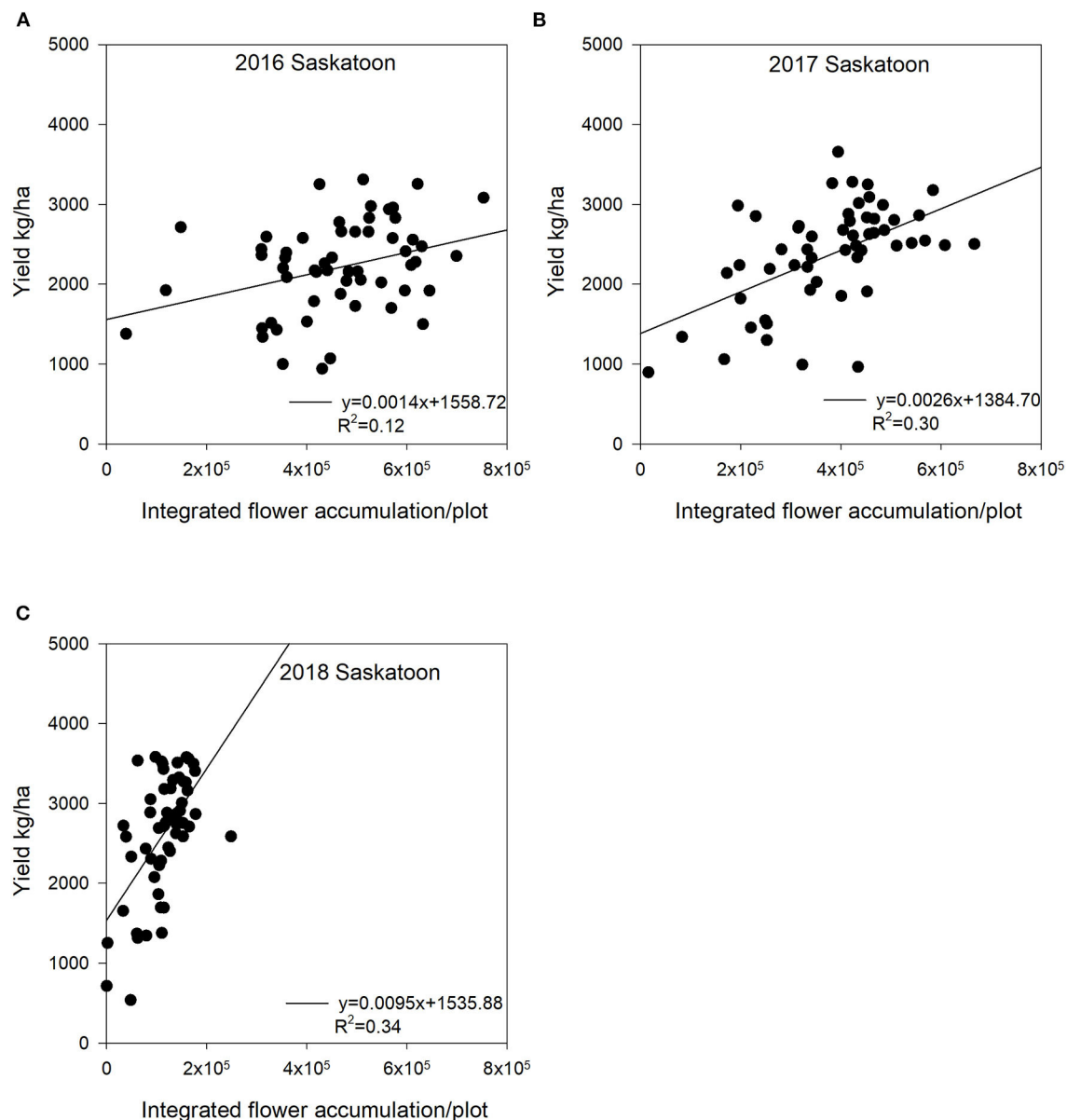
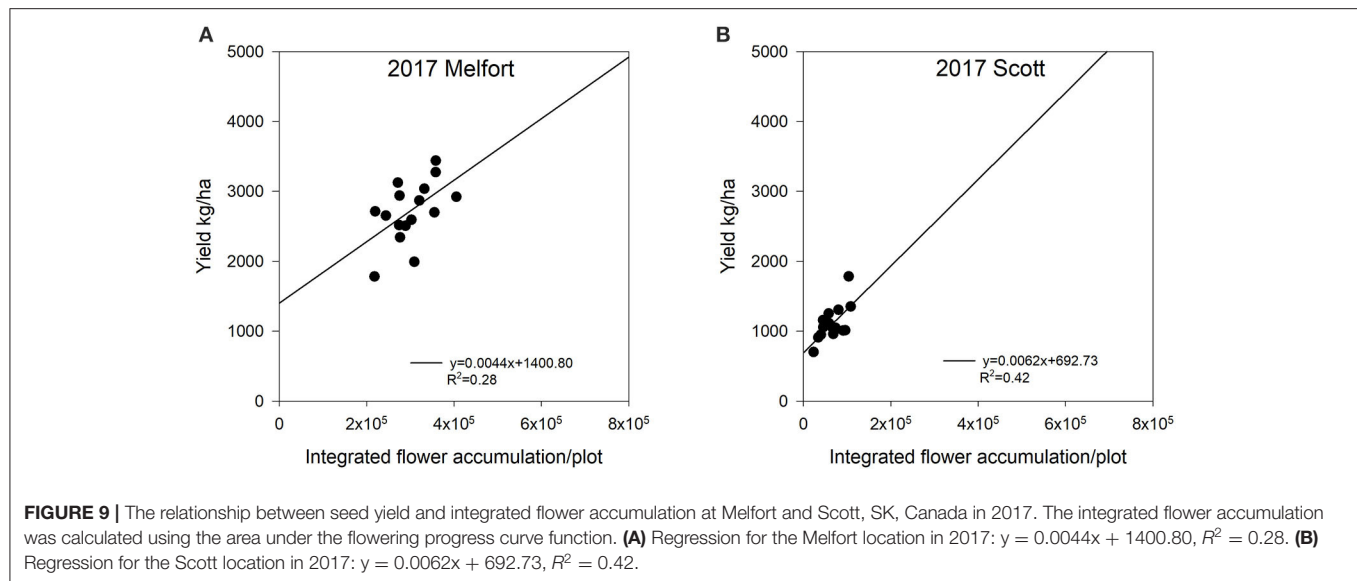


FIGURE 8 | The relationship between seed yield and integrated flower accumulation at Saskatoon, SK, Canada from 2016 to 2018. The integrated flower accumulation was calculated using the area under the flowering progress curve function. **(A)** Regression for the Saskatoon location in 2016: $y = 0.0014x + 1558.72$, $R^2 = 0.12$. **(B)** Regression for the Saskatoon location in 2017: $y = 0.0026x + 1384.70$, $R^2 = 0.30$. **(C)** Regression for the Saskatoon location in 2018: $y = 0.0095x + 1535.88$, $R^2 = 0.34$.

$= 0.72$), which showed a better correlation with seed yield than NDVI at the peak flowering time in canola. Gong et al. (2018) found that NDVI multiplied by leaf-related canopy fraction had the strongest relationship with canola yield with low estimation errors (coefficient of variation $< 13\%$) at the early flowering stages. Some research also investigated yield estimation using canopy reflectance data in other crops including cotton and wheat (Xu et al., 2018; Hassan et al., 2019). Xu et al. (2018) reported that the estimated cotton flower numbers derived from aerial images using a convolutional neural network significantly

correlated with cotton yield ($R^2 = 0.36$). Hassan et al. (2019) reported that UAV-based NDVI measured at the grain filling stage could be a promising tool for wheat yield prediction with R^2 ranging from 0.83 to 0.89 in field conditions.

Our regression results had smaller R^2 values compared with the previous studies. This is probably associated with many diverse lines (i.e., 56 diverse lines) estimated in this study. Most previous research only planted one or few lines. The stability of pigments in rapeseed flowers for each line may change under different developmental stages (Ohmiya, 2011). These factors can



impact yellow to some degree (Ohmiya, 2011). The inconsistent yellowness may explain that the more varieties included in regression analysis, the less model variation could be explained by integrated flower accumulation. Furthermore, flowering pixels extracted based on threshold values may not be highly consistent over the flowering stage, as each threshold value was determined manually. In addition, canola yield components include plant density, pod number per plant, seed number per pod, and seed weight. Although pod number per plant is highly correlated with seed yield (Tayo and Morgan, 1975; McGregor, 1981; Ivanovska et al., 2007), only 45% of flowers produce pods (McGregor, 1981). Seed weight per pod and thousand seed weight also significantly correlated with seed yield (Ivanovska et al., 2007). The simple regression analysis of flower numbers could not fully explain yield variation. Additional yield components considered in the yield estimation model would improve the accuracy of seed yield estimation. Finally, flower abortion and poor pod formation can happen under drought and heat stress during the crop season (Faraji et al., 2008). Flowering progress only reveals part of crop growth stages, so some varieties even with high AUFPC may end up with low yield under stress, which may result in a weaker relationship between integrated flower accumulation and seed yield. Combining UAV-based reflectance data at both flowering and pod stages may enhance yield estimation accuracy.

Usually, breeding programs need to assess a large number of varieties or breeding lines across multiple environmental conditions. Therefore, from a practical perspective, these results revealed a more realistic yield estimation trend for large-scale breeding programs. Moreover, most previous research used one image date or selected the largest reflectance index value for each plot across all sampling dates to estimate crop yield. In fact, it is difficult to determine the best image date for yield estimation using multiple crop varieties grown in differing environmental conditions. Fluctuating environments can influence flowering progress; therefore, integrated flower accumulation is a

promising and predictable variable in the descriptive yield model.

CONCLUSIONS

In this study, we proposed a simple and effective approach to estimate relative flower numbers and model seed yield based on the integrated flowering pixel. This study results showed that flowering pixel numbers estimated by the thresholding method regressed strongly with manual flower count during the flowering stage with an R^2 value of up to 0.95, indicating that flowering pixel numbers can be used as a good indicator of flowering intensity in the field. Additionally, the integrating flowering progress from consecutive images *via* AUFPC math function was more consistently and strongly related to yield compared with using a single image date because integrated flowering pixel over time utilizes more growth information. Therefore, the integrated flower accumulation can be a good indicator for yield estimation. These tools do not require extra coding or strong computer science background, can be used for calculating thresholding and vegetation indices, and is a convenient tool for agronomists and breeders. Future studies need to consider and test a multivariate model including multiple vegetation indices related to other yield components and more reflectance information from the pod stage to improve yield estimation accuracy.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

SS and SV designed the field experiments. SV and IP provided and prepared the plant materials. SS contributed to funding

acquisition. TZ, HD, and SS collected the images and performed imagery pre-processing. TZ conducted ground data collection, performed the statistical analysis, and wrote the manuscript under the supervision of SS and with contributions from SV, HD, EJ, and XG. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by a Canada First Research Excellence Fund to the Plant Phenotyping and Imaging Research

Center (P2IRC), Global Institute for Food Security (GIFS) at the University of Saskatchewan, SK, Canada, and Chinese Scholarship Council (201206300150).

ACKNOWLEDGMENTS

We would like to thank the technical staff at the Department of Plant Sciences in the University of Saskatchewan for their assistance in field data and imagery collection. We would also like to thank the technician staff at the Agriculture and Agri-Food Canada for their help in field management.

REFERENCES

- Araus, J. L., and Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci.* 19, 52–61. doi: 10.1016/j.tplants.2013.09.008
- Blanc, J., Dutartre, D., Tixier, M. H., Weiss, M., Comar, A., Praud, S., et al. (2019). A high-throughput model-assisted method for phenotyping maize green leaf area index dynamics using unmanned aerial vehicle imagery. *Front. Plant Sci.* 10:685. doi: 10.3389/fpls.2019.00685
- Carl, C., Landgraf, D., der Maaten-Theunissen, V., Biber, P., and Pretzsch, H. (2017). *Robinia pseudoacacia* L. flower analyzed by using an unmanned aerial vehicle (UAV). *Remote Sens.* 9:1091. doi: 10.3390/rs9111091
- Clayton, G., Turkington, K., Harker, N., O'Donovan, J., and Johnston, A. (2000). High yielding canola production. *Better Crop.* 84:26–27. Available online at: [http://www.ipni.net/publication/bettercrops.nsf/0/405DA9F381C7D59F85257980081FDE7/\\$FILE/Better%20Crops%202000-1%20p26.pdf](http://www.ipni.net/publication/bettercrops.nsf/0/405DA9F381C7D59F85257980081FDE7/$FILE/Better%20Crops%202000-1%20p26.pdf) (accessed January 05, 2020).
- d'Andrimont, R., Taymans, M., Lemoine, G., Ceglar, A., Yordanov, M., and van der Velde, M. (2020). Detecting flowering phenology in oil seed rape parcels with Sentinel-1 and-2 time series. *Remote Sens. Environ.* 239:1–14. doi: 10.1016/j.rse.2020.111660
- Diepenbrock, W. (2000). Yield analysis of winter oilseed rape (*Brassica napus* L.): a review. *Field Crops Res.* 67, 35–49. doi: 10.1016/S0378-4290(00)00082-4
- Fang, S., Tang, W., Peng, Y., Gong, Y., Dai, C., Chai, R., et al. (2016). Remote estimation of vegetation fraction and flower fraction in oilseed rape with unmanned aerial vehicle data. *Remote Sens.* 8: 416. doi: 10.3390/rs8050416
- Faraji, A. (2012). Flower formation and pod/flower ratio in canola (*Brassica napus* L.) affected by assimilates supply around flowering. *Int. J. Plant Product.* 4, 271–280. doi: 10.22069/ijpp.2012.710
- Faraji, A., Latifi, N., Soltani, A., and Rad, A. H. S. (2008). Effect of high temperature stress and supplemental irrigation on flower and pod formation in two canola (*Brassica napus* L.) cultivars at Mediterranean climate. *Asian J. Plant Sci.* 7, 343–351. doi: 10.3923/ajps.2008.343.351
- Gan, Y. T., Harker, K. N., Kutcher, H. R., Gulden, R. H., Irvine, B., May, W. E., et al. (2016). Canola seed yield and phenological responses to plant density. *Can. J. Plant Sci.* 96, 151–159. doi: 10.1139/cjps-2015-0093
- Gitelson, A. A., and Merzlyak, M. N. (1997). Remote estimation of chlorophyll content in higher plant leaves. *Int. J. Remote Sens.* 18, 2691–2697. doi: 10.1080/014311697217558
- Gitelson, A. A., Merzlyak, M. N., and Lichtenthaler, H. K. (1996). Detection of red edge position and chlorophyll content by reflectance measurements near 700 nm. *J. Plant Physiol.* 148, 501–508. doi: 10.1016/S0176-1617(96)80285-9
- Gong, Y., Duan, B., Fang, S., Zhu, R., Wu, X., Ma, Y., et al. (2018). Remote estimation of rapeseed yield with unmanned aerial vehicle (UAV) imaging and spectral mixture analysis. *Plant Methods* 14, 1–14. doi: 10.1186/s13007-018-0338-z
- Guo, W., Fukatsu, T., and Ninomiya, S. (2015). Automated characterization of flowering dynamics in rice using field-acquired time-series RGB images. *Plant Methods* 11, 1–15. doi: 10.1186/s13007-015-0047-9
- Han, J., Zhang, Z., and Cao, J. (2021). Developing a new method to identify flowering dynamics of rapeseed using landsat 8 and sentinel-1/2. *Remote Sens.* 13:105. doi: 10.3390/rs13010105
- Hassan, M. A., Yang, M., Rasheed, A., Yang, G., Reynolds, M., Xia, X., et al. (2019). A rapid monitoring of NDVI across the wheat growth cycle for grain yield prediction using a multi-spectral UAV platform. *Plant Sci.* 282, 95–103. doi: 10.1016/j.plantsci.2018.10.022
- Haynes, K. G., and Weingartner, D. P. (2004). The use of area under the disease progress curve to assess resistance to late blight in potato germplasm. *Am. J. Potato Res.* 81, 137–141. doi: 10.1007/BF02853611
- Ivanovska, S., Stojkovski, C., Dimov, Z., Marjanović-Jeromela, A., Jankulovska, M., and Jankuloski, L. (2007). Interrelationship between yield and yield related traits of spring canola (*Brassica napus* L.) genotypes. *Genetika* 39, 325–332. doi: 10.2298/GENSRO703325I
- Jeger, M. J., and Viljanen-Rollinson, S. L. H. (2001). The use of the area under the disease-progress curve (AUDPC) to assess quantitative disease resistance in crop cultivars. *Theoret. Appl. Gen.* 102, 32–40. doi: 10.1007/s001220051615
- Jiang, L., and Becker, H. C. (2003). Inheritance of apetalous flowers in a mutant of oilseed rape. *Crop Sci.* 43, 508–510. doi: 10.2135/cropsci2003.0508
- Kefauver, S. C., Vicente, R., Vergara-Díaz, O., Fernandez-Gallego, J. A., Kerfal, S., Lopez, A., et al. (2017). Comparative UAV and field phenotyping to assess yield and nitrogen use efficiency in hybrid and conventional barley. *Front. Plant Sci.* 8:1733. doi: 10.3389/fpls.2017.01733
- Kim, J., Kim, S., Ju, C., and Son, H. I. (2019). Unmanned aerial vehicles in agriculture: a review of perspective of platform, control, and applications. *IEEE Access* 7, 105100–105115. doi: 10.1109/ACCESS.2019.2932119
- Kirkegaard, J. A., Lilley, J. M., Brill, R. D., Ware, A. H., and Walela, C. K. (2018). The critical period for yield and quality determination in canola (*Brassica napus* L.). *Field Crops Res.* 222, 180–188. doi: 10.1016/j.fcr.2018.03.018
- McGregor, D. I. (1981). Pattern of flower and pod development in rapeseed. *Can. J. Plant Sci.* 61, 275–282. doi: 10.4141/cjps81-040
- Migdall, S., Ohl, N., and Bach, H. (2010). "Parameterisation of the land surface reflectance model SLC for winter rape using spaceborne hyperspectral CHRIS data," in *Hyperspectral Workshop*. (Vol. ESA SP-683) (Frascati).
- Montes, J. M., Melchinger, A. E., and Reif, J. C. (2007). Novel throughput phenotyping platforms in plant genetic studies. *Trends Plant Sci.* 12, 433–436. doi: 10.1016/j.tplants.2007.08.006
- Ohmiya, A. (2011). Diversity of carotenoid composition in flower petals. *Jap. Agri. Res. Quart.* 45, 163–171. doi: 10.6090/jarq.45.163
- Parkin, I., Vail, S., and Robinson, S. (2017). *Development of a Germplasm Resource to Dissect Complex Traits in Brassica napus*, ADF Project #20110155 Final Report. Available online at: https://www.saskcanola.com/quadrant/System/research/reports/Parkin.ADF20110155.Long_Report-1.pdf (accessed May 1, 2021).
- Prey, L., Hu, Y., and Schmidhalter, U. (2020). High-throughput field phenotyping traits of grain yield formation and nitrogen use efficiency: optimizing the selection of vegetation indices and growth stages. *Front. in Plant Sci.* 10:1672. doi: 10.3389/fpls.2019.01672
- Rouse, J. W., Haas, R. H., Schell, J. A., Deering, D. W., and Harlan, J. C. (1974). *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*. NASA/GSFC Type III Final Report. Greenbelt, MD, 371.

- Sankaran, S., Khot, L. R., Espinoza, C. Z., Jarolmasjed, S., Sathuvalli, V. R., Vandemark, G. J., et al. (2015). Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: a review. *Eur. J. Agron.* 70, 112–123. doi: 10.1016/j.eja.2015.07.004
- Shen, M., Chen, J., Zhu, X., and Tang, Y. (2009). Yellow flowers can decrease NDVI and EVI values: evidence from a field experiment in an alpine meadow. *Can. J. Remote Sens.* 35, 99–106. doi: 10.5589/m09-003
- Shen, M., Chen, J., Zhu, X., Tang, Y., and Chen, X. (2010). Do flowers affect biomass estimate accuracy from NDVI and EVI? *Int. J. Remote Sens.* 31, 2139–2149. doi: 10.1080/01431160903578812
- Simko, I., and Piepho, H. P. (2012). The area under the disease progress stairs: calculation, advantage, and application. *Phytopathology* 102, 381–389. doi: 10.1094/PHYTO-07-11-0216
- Statistics Canada (2018). *Seeding Decisions Harvest Opportunities for Canadian Farmers*. Available online at: <https://www150.statcan.gc.ca/n1/pub/95-640-x/2016001/article/14813-eng.htm> (accessed May 23, 2018).
- Sulik, J. J., and Long, D. S. (2015). Spectral indices for yellow canola flowers. *Int. J. Remote Sens.* 36, 2751–2765. doi: 10.1080/01431161.2015.1047994
- Sulik, J. J., and Long, D. S. (2016). Spectral considerations for modeling yield of canola. *Remote Sens. Environ.* 184, 161–174. doi: 10.1016/j.rse.2016.06.016
- Tayo, T. O., and Morgan, D. G. (1975). Quantitative analysis of the growth, development and distribution of flowers and pods in oil seed rape (*Brassica napus* L.). *J. Agri. Sci.* 85, 103–110. doi: 10.1017/S0021859600053466
- Tunca, E., Köksal, E. S., Çetin, S., Ekiz, N. M., and Balde, H. (2018). Yield and leaf area index estimations for sunflower plants using unmanned aerial vehicle images. *Environ. Monitor. Assess.* 190, 1–12. doi: 10.1007/s10661-018-7064-x
- Wan, L., Li, Y., Cen, H., Zhu, J., Yin, W., Wu, W., et al. (2018). Combining UAV-based vegetation indices and image classification to estimate flower number in oilseed rape. *Remote Sens.* 10:1484. doi: 10.3390/rs10091484
- White, J. W., Andrade-Sanchez, P., Gore, M. A., Bronson, K. F., Coffelt, T. A., Conley, M. M., et al. (2012). Field-based phenomics for plant genetics research. *Field Crops Res.* 133, 101–112. doi: 10.1016/j.fcr.2012.04.003
- Wójtowicz, M., Wójtowicz, A., and Piekarczyk, J. (2016). Application of remote sensing methods in agriculture. *Commun. Biomet. Crop Sci.* 11, 31–50.
- Xu, R., Li, C., Paterson, A. H., Jiang, Y., Sun, S., and Robertson, J. S. (2018). Aerial images and convolutional neural network for cotton bloom detection. *Front. Plant Sci.* 8:2235. doi: 10.3389/fpls.2017.02235
- Yates, D. J., and Steven, M. D. (1987). Reflexion and absorption of solar radiation by flowering canopies of oil-seed rape (*Brassica napus* L.). *J. Agri. Sci.* 109, 495–502. doi: 10.1017/S0021859600081703
- Zhang, H., and Flottmann, S. (2018). Source-sink manipulations indicate seed yield in canola is limited by source availability. *Eur. J. Agron.* 96, 70–76. doi: 10.1016/j.eja.2018.03.005

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Zhang, Vail, Duddu, Parkin, Guo, Johnson and Shirliffe. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Creating High-Resolution Microscopic Cross-Section Images of Hardwood Species Using Generative Adversarial Networks

Dercilio Junior Verly Lopes^{1*}, Gustavo Fardin Monti², Greg W. Burgreen³, Jordão Cabral Moulin⁴, Gabrielly dos Santos Bobadilha¹, Edward D. Entsminger¹ and Ramon Ferreira Oliveira¹

¹ Department of Sustainable Bioproducts, Forest and Wildlife Research Center, Mississippi State University, Starkville, MS, United States, ² Universidade Federal do Espírito Santo, Centro Universitario do Norte do Espírito Santo, São Mateus, Brazil, ³ Center for Advanced Vehicular Systems, Mississippi State University, Starkville, MS, United States, ⁴ Departamento de Ciências Florestais e da Madeira, Universidade Federal do Espírito Santo, Jerônimo Monteiro, Brazil

OPEN ACCESS

Edited by:

Lisbeth Garbrecht Thygesen,
University of Copenhagen, Denmark

Reviewed by:

Gabriele Ehmcke,
Thünen Institute of Wood Research,
Germany

Natalie M. Clark,
Iowa State University, United States

*Correspondence:

Dercilio Junior Verly Lopes
dvl23@msstate.edu

Specialty section:

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

Received: 17 August 2021

Accepted: 20 September 2021

Published: 13 October 2021

Citation:

Lopes DJV, Monti GF,
Burgreen GW, Moulin JC,
dos Santos Bobadilha G,
Entsminger ED and Oliveira RF (2021)
Creating High-Resolution Microscopic
Cross-Section Images of Hardwood
Species Using Generative Adversarial
Networks.
Front. Plant Sci. 12:760139.
doi: 10.3389/fpls.2021.760139

Microscopic wood identification plays a critical role in many economically important areas in wood science. Historically, producing and curating relevant and representative microscopic cross-section images of wood species is limited to highly experienced and trained anatomists. This manuscript demonstrates the feasibility of generating synthetic microscopic cross-sections of hardwood species. We leveraged a publicly available dataset of 119 hardwood species to train a style-based generative adversarial network (GAN). The proposed GAN generated anatomically accurate cross-section images with remarkable fidelity to actual data. Quantitative metrics corroborated the capacity of the generative model in capturing complex wood structure by resulting in a Fréchet inception distance score of 17.38. Image diversity was calculated using the Structural Similarity Index Measure (SSIM). The SSIM results confirmed that the GAN approach can successfully synthesize diverse images. To confirm the usefulness and realism of the GAN generated images, eight professional wood anatomists in two experience levels participated in a visual Turing test and correctly identified fake and actual images at rates of 48.3 and 43.7%, respectively, with no statistical difference when compared to random guess. The generative model can synthesize realistic, diverse, and meaningful high-resolution microscope cross-section images that are virtually indistinguishable from real images. Furthermore, the framework presented may be suitable for improving current deep learning models, helping understand potential breeding between species, and may be used as an educational tool.

Keywords: wood anatomy, machine learning, artificial intelligence, wood image transformation, microscopic images, StyleGAN

INTRODUCTION

Transverse microscopic cross-sections of wood species have long been used for forensic wood identification, for analysis of critically important properties such as permeability and treatability with chemical agents, and to gain an understanding of the functioning of the tree (Zhang and Cai, 2006; Esteves and Pereira, 2008; Martins et al., 2013; Leggate et al., 2020; Lengowski et al., 2020;

Ślupianek et al., 2021). Microscopic capture of various anatomical features is accomplished in the lab by preparing individual thin slices of wood samples through standard stringent procedures that include several manually intensive steps: softening, cutting, clearing, staining, dehydrating, and mounting of the thin wood sections (Jansen et al., 1998).

Historically, creating and curating large datasets of microscopic wood images has been cumbersome with only a handful of datasets available to the public for research and development. The dataset produced by Martins et al. (2013) is perhaps the most used dataset for benchmarking several different wood identification approaches. The art of wood identification using such datasets is limited to only highly trained and experienced wood anatomists, due to the complexity of the wood structure within species and among a multitude of different species. Moreover, the number of senior wood anatomists with broad taxonomic expertise is declining (Lens et al., 2020). These limitations have set the stage for new artificial intelligence/machine-learning (AI/ML) technologies to make significant advances into the wood identification process.

Currently, deep learning in the form of convolutional neural networks (CNN) and optimization algorithms is beginning to revolutionize wood identification services. In fact, this technology is matching or surpassing expert wood anatomists in both macroscopic and microscopic image recognition and is being increasingly proposed as an adjunct to human wood identification decision-making (Hafemann et al., 2014; Lens et al., 2020; Lopes et al., 2020, 2021; Olschofsky and Köhl, 2020; de Geus et al., 2021; Fabijańska et al., 2021). The growth of computer-based wood identification and many other recognition tasks is facilitated by recent advancements in computational power, especially with graphical processing units (GPUs), which have enabled the widespread use of supervised machine-learning.

The AI/ML approaches have a rich potential within wood science and technology. For example, computer vision approaches could help identify and protect forests in the future (Lens et al., 2020). In this case, the expansion of computer vision-based wood identification would heavily depend on either establishing traditional extensive collaborations across wood science organizations as explained by Hwang and Sugiyama (2021) or through the development and application of artificial intelligence solutions that are novel, economically relevant, innovative, and stakeholder-engaged.

Successful applications of deep learning for wood identification are based on supervised learning algorithms that critically depend on labeled data for training purposes (Hwang and Sugiyama, 2021). For example, Martins et al. (2013); Filho et al. (2014), and Hafemann et al. (2014) applied deep CNN models on macroscopic and microscopic images by manually labeling the forest wood species. Their custom deep learning-based model achieved 96.0 and 97.0% accuracies on the macroscopic and microscopic datasets, respectively. Similarly, Fabijańska et al. (2021) automatically identified 14 European tree species using a residual convolutional encoder network in a sliding window with 99.0% accuracy. Collecting large sets of labeled training data constitutes a non-trivial bottleneck

in AI/ML workflows. However, AI/ML has the potential to artificially synthesize the requisite labeled data, which we will explore in this manuscript.

Generative adversarial networks (GANs) are special types of deep learning where two neural networks are trained simultaneously, with the generator Network G, focusing on image generation from feedback given by a discriminator Network D, that is designed to determine whether a given input data is from an actual dataset or is synthetically generated (fake) by G (Yi et al., 2019). The GANs can achieve state-of-the-art synthetic generation of remarkably realistic images using CNN in an unsupervised manner. The GANs have been successfully applied in many fields including medical analysis, satellite imagery, computational fluid dynamics, and precision agriculture (Goodfellow et al., 2014; Nie et al., 2018; Wu et al., 2020; Pang et al., 2021).

Given the ability to use deep learning to synthesize images from multiple domains, we herein seek to explore the utility of GANs to map and generate labeled microscopic images on a large number of hardwood species. Therefore, the purpose of this manuscript is fourfold: (1) to demonstrate the feasibility of image synthesis in the field of wood anatomy; (2) to quantitatively and qualitatively assess the quality of generated images; (3) to present synthetically generated images to experts in the field through a visual Turing test (VTT); and (4) to raise awareness of the potential of deep learning techniques for steering the forestry and forest and wood products industry toward transformative directions.

To our knowledge, no study has been conducted using GANs to synthesize and critically evaluate microscopic cross-sectional images of hardwood species or in wood anatomy in general. This study seeks to demonstrate proof-of-concept technical and computational feasibility of performing image domain transformation to better equip wood anatomists and to introduce the wood science and technology communities to a novel AI/ML-based approach.

MATERIALS AND METHODS

Transverse Microscopic Hardwoods Section Dataset

This study was conducted using a publicly accessible transverse section of microscopic hardwood species dataset obtained from the Xylarium Digital Database (XDD) for Wood Information Science and Education – Kyoto University Research Information. This database was created, curated, processed, and labeled by Sugiyama et al. (2020). It was created in an effort to expand research and development in the area of wood anatomy and wood identification. The methods for obtaining the cross-section of the wood species are thoroughly described in the series of manuscripts published by the XDD research team in Hwang et al. (2018, 2020a,b) and Kobayashi et al. (2019). **Figure 1** shows eight different woody species present in the dataset.

Observing **Figure 1**, it is evident the diversity in anatomical structure with clear growth rings distinction, latewood and earlywood transitions, parenchyma cells, arrangement of

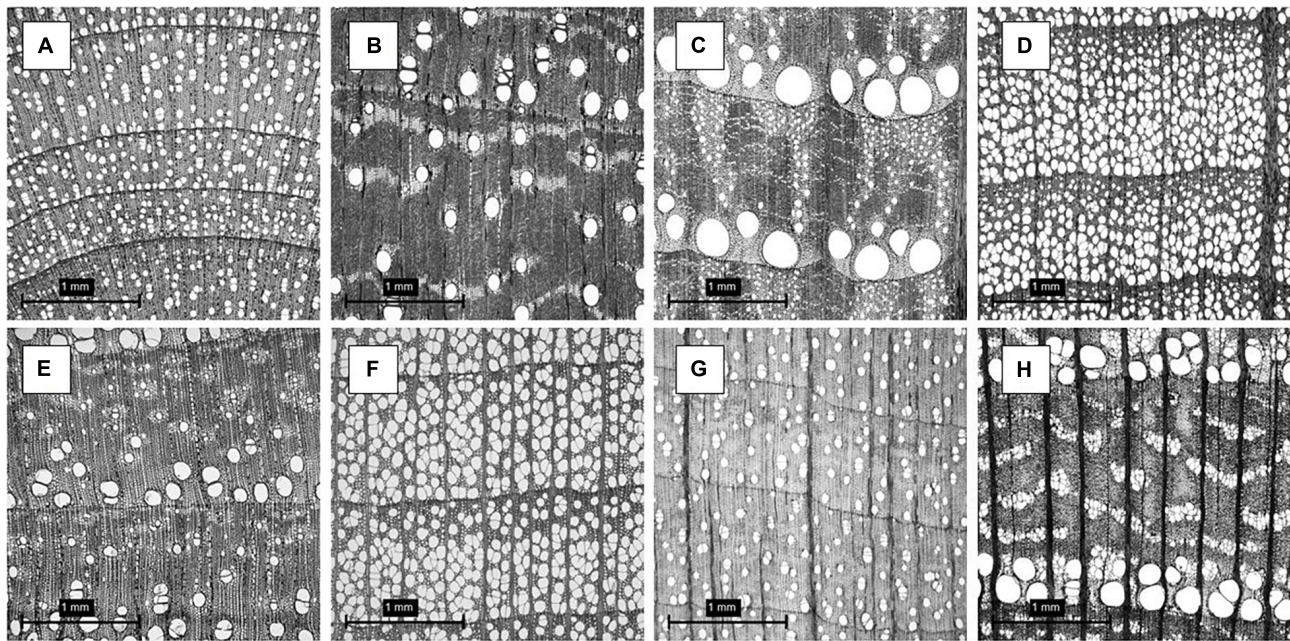


FIGURE 1 | (A) Betulaceae - *Alnus firma*; (B) Cannabaceae - *Aphananthe aspera*; (C) Fagaceae - *Quercus crispula*; (D) Fagaceae - *Fagus japonica*; (E) Lauraceae - *Sassafras tzumu*; (F) Magnoliaceae - *Liriodendron tulipifera*; (G) Sapindaceae - *Acer distylum*; (H) Ulmaceae - *Ulmus laciniata*. Refer to the dataset for full dataset description.

parenchyma cells, fiber, vessel elements, pores and pores arrangement, multiple porosity classifications (ring, semi-ring, semi-diffuse, and diffuse porous), pore structure, and rays. These features are examples of key anatomical elements for hardwood identification. The full list of features and their terminology can be seen in Wheeler et al., 1989. The XDD dataset contained 7,051 images from 33 genus, 119 species, and 540 individuals at a resolution of $2.96 \mu\text{m}/\text{pixel}$ in a compressed HDF5 file at a grayscale resolution of $900 \text{ pixels} \times 900 \text{ pixels}$ in JPEG (Joint Photographic Expert Group) format. The full description of the wood species can be seen in the **Supplementary Material 1**.

Custom Training the Generative Adversarial Network

We leveraged the style-based generative adversarial network, henceforth StyleGAN model developed by Karras et al. (2019) to generate realistic microscopic cross-section images of hardwood species. The StyleGAN includes the progressive increase of resolution by adding layers to the network as described in Karras et al. (2018) with a series of later modifications described in Karras et al. (2019). The main reason for choosing StyleGAN was that it achieves state-of-the-art in human face transformations with extraordinary levels of detail. Similarly, to human faces, wood is a biological material with high-level attributes and stochastic variation in its structure, which requires an AI/ML framework that generates small and subtle intricacies of wood anatomy such as fibers, cells, pores shapes, pore arrangements, and rays, etc.

As the original image size was $900 \text{ pixels} \times 900 \text{ pixels}$, we resized the images to be $512 \text{ pixels} \times 512 \text{ pixels}$ without further image processing. In this implementation, the StyleGAN progressively increased image size from 4^2 pixels to 512^2 pixels. We used 5,650 images for training. A latent vector of dimension 512 was used. The batch size decreased from 256 to 4 as training progressed. The adaptive momentum estimator (Adam) (Kingma and Ba, 2015) optimizer was used for training. The learning rate for the discriminator and generator were initially set to 0.0015 up to the resolution of 128^2 pixels and slowly increased to 0.02 and 0.03 for resolutions of 256^2 and 512^2 pixels, respectively. The training setup doubled the image resolution when 600,000 images were shown to the discriminator. Training finished when the model had seen 7.5 million synthesized images. The Wasserstein GAN-gradient penalty (WGAN-GP) loss developed by Gulrajani et al. (2017) with modifications included by Karras et al. (2019) was used. Throughout the training session, the model serialized checkpoints for later inference by using a script for image generation. The training took approximately 10 days. The computational resources used for this study included a workstation powered by $4 \times$ NVIDIA GeForce RTX 2080Ti graphics processing units (GPU) with 11 GB of memory each and an Intel Core i9-9920K with a central processing unit (CPU) with 128 GB of memory.

Quantitative Analysis of Generative Adversarial Network Images

There is no unified and universal metric to compare and evaluate generative adversarial networks (Borji, 2019). In the case of wood

anatomy, the quantitative measure of GANs is limited or even non-existent. This work, to the best of our knowledge, is the first study to present GAN metrics in the domain of wood anatomy. For GAN metrics, we relied on the Fréchet inception distance (FID) by Heusel et al. (2018) and the Structural Similarity Index Measure (SSIM) by Hore and Ziou (2010) to assess the realism and diversity of the images generated by the StyleGAN.

The FID score is a metric that measures the maximum Gaussian entropy distribution for given mean and covariance. The difference of two Gaussians is then measured by Eq. 1:

$$FID = \|\mu_r - \mu_g\|^2 + TR \left(C_r + C_g - 2(C_r C_g)^{\frac{1}{2}} \right) \quad (1)$$

where, μ_r and μ_g and C_r and C_g are the mean and covariance of real and generated images.

The lower FID score means higher accuracy in synthetically generating microscopic cross-sectional images. The FID score enables a quantifiable anatomical comparison between a ground-truth image and a GAN generated image with respect to the fidelity of generated images.

The SSIM is a quality metric used to measure the similarity between two images. It is considered to be correlated with the quality perception of the human visual system (HVS) (Hore and Ziou, 2010). The SSIM is designed by modeling any image distortion as a combination of three factors, namely loss of correlation, luminance, and contrast distortions. The SSIM was defined by Eq. 2:

$$SSIM(f, g) = l(f, g) c(f, g) s(f, g) \quad (2)$$

where,

$$l(f, g) = \frac{2\mu_f \mu_g + C_1}{\mu_f^2 + \mu_g^2 + C_1} \quad (3)$$

$$c(f, g) = \frac{2\sigma_f \sigma_g + C_2}{\sigma_f^2 + \sigma_g^2 + C_2} \quad (4)$$

$$s(f, g) = \frac{\sigma_{fg} + C_3}{\sigma_f \sigma_g + C_3} \quad (5)$$

Equations 3–5, respectively, refer to the luminance comparison function that measures the closeness of two images mean luminance (μ_f and μ_g); the contrast comparison function, which calculates the closeness of the contrast of the two images by the standard deviation (σ_f and σ_g); and the structure comparison function that measures the correlation coefficient between the two images, f and g . The σ_{fg} argument is the covariance between f and g . A value of zero (0) means no correlation between images, and a value of one (1) means that $f = g$ (Hore and Ziou, 2010).

Visual Turing Test

To compare between actual and generated microscopic cross-section images of hardwood species, we used a VTT based on Park et al. (2021) and Chuquicusma et al. (2018). Our VTT experiments were conducted by a group of eight wood anatomy experts divided into two levels of expertise for analysis of microscopic wood images, namely, four intermediate wood

anatomy experts [more than 1 and less than 5 years of experience (Group I)], and four advanced wood anatomy experts [more than 5 years of experience (Group II)].

The wood anatomists were blinded to each other's evaluations of experiments and were not shown real or generated images prior to the experiments. The VTT contained 60 distinct 512² images (30 actual images and 30 generated images). We randomly selected the images from the actual dataset, such that a minimum of three images were selected from each family. To avoid any bias, the generated image data were automatically generated by the StyleGAN. Furthermore, these images were not individually selected by our group.

The experts were given two choices to classify the fidelity of the images, namely, actual image or generated image. A website (Google Forms) was created to upload the images in a random manner. The link for the website can be seen in the GitHub.¹ The visuals evaluated did not contain any information about the wood species and only the microscopic cross-section of hardwood species was presented.

In this experiment, the experts were not informed how many of the images were real or not real. The non-disclosed ratio allowed the evaluation of three important metrics: (1) number of incorrectly identified actual images (a high number represents how real the generated images look), (2) number of corrected identified real images (a high number represents how accurately the experts recognized salient anatomical features), and (3) a confusion metric that represents how effective our results were to confuse experts in identifying actual versus generated images.

The mean sensitivity, specificity, and accuracy of the eight expert VTT evaluations were calculated by Eqs 6–8.

$$\text{Sensitivity} = \frac{\text{True positive}}{(\text{True positive} + \text{False negative})} \quad (6)$$

$$\text{Specificity} = \frac{\text{True negative}}{(\text{True negative} + \text{False positive})} \quad (7)$$

$$\text{Accuracy} = \frac{(\text{True positive} + \text{True negative})}{\text{Number of observations}} \quad (8)$$

A statistical *t*-test was used to compare the means of the experts' evaluations across the experiment. The scientific computing Python package Pauli et al. (2020) was used for the statistical analyses with the significance level set at $p \leq 0.05$.

RESULTS AND DISCUSSION

Feasibility of StyleGAN Generative Adversarial Network Training

The first goal of the study was to demonstrate the feasibility of training StyleGAN from scratch to generate realistic microscopic cross-section images of hardwood species. We found concomitant training improvement of the model up to approximately 7.5 million images seen by the discriminator,

¹<https://github.com/LignumResearch/stylewood-model-usage>

which corresponded to training at the final resolution of 512^2 pixels. **Figure 2** illustrates the progress of image generation as the resolution increased during training from 4^2 to 512^2 pixels. Initially, at 4^2 pixels resolution, the generated images were pure abstract noise with concomitant progress in image quality with remarkable realism obtained at resolution of 512^2 pixels. The StyleGAN trained as expected and was found to generate visually acceptable synthetic cross-section images of hardwood species.

Qualitative Analysis of Generated Images

Artificial intelligence and deep learning frameworks are revolutionizing interpretation, identification, and decision-making in wood species recognition. As data quantity and quality are critical to train deep learning-based image recognition systems, the proposed method herein should be useful to assist the computer vision wood identification community by providing realistic and meaningful microscopic images of cross-section of hardwood species. Using trained StyleGAN model, examples of the random generation of synthetic microscopic cross-section hardwood species are shown in **Figure 3**.

Qualitatively, a remarkable variety of anatomical elements was generated by the trained generative adversarial StyleGAN network. The StyleGAN was capable of synthesizing high detail levels of the earlywood and latewood bands and growth ring transitions; ray width, height, and arrangement of apotracheal and paratracheal parenchyma cells; porosity such as ring-porous, semi-ring, semi-diffuse, and diffuse porous; and vessels with different arrangements and diameters were produced and recognized. Such detailed anatomical elements are what enable wood anatomists to scientifically identify wood species. Correct wood identification promotes reliable utilization of wood in various forms as in flooring, structural elements, plywood, particleboard, cross-laminated timber (CLT), various engineered wood products, and many other structural applications. **Figure 4** illustrates the learned anatomical elements by the generative model in detail. **Figures 4A–D** should be carefully analyzed as these species do not exist, although may look similar to actual data. They were created using the StyleGAN generator,

which allows control over various aspects of the image. They represent the capability of the proposed network in generating realistic and meaningful microscopic cross-section images of hardwood species.

In **Figure 4A**, there is a visible transition between earlywood and latewood growth ring bands, parenchyma and fiber cells noticeable, and uniseriate rays that are clearly seen. Different earlywood and latewood pores, pore arrangement, rays, initial earlywood band, and a few nested pores or pore clusters are also identifiable. There are also numerous solitary pores. No tyloses or mineral deposits can be seen in the vessel elements. The presence of paratracheal vasicentric, paratracheal aliform, and/or paratracheal confluent longitudinal parenchyma cells was not identifiable. Possible semi-ring-porous wood with clear separation between earlywood and latewood pores in **Figure 4A**.

In **Figure 4B**, the growth ring bands are visible, parenchyma and fiber cells noticeable, and uniseriate rays are clearly seen. There are numerous pore multiples that occur throughout, where two or more pores are connected to another pore. The radially arranged series of pore multiples or closely arranged solitary pores are visible as pore chains. These characteristics along with no clear separation between earlywood and latewood pores, small vessel element pore sizes, uniform pore size, and evenly distribution of the pores make this a possible diffuse-porous wood in **Figure 4B**.

In **Figure 4C**, shows a visible transition between earlywood and latewood growth ring bands, parenchyma and tracheids cells noticeable, and uniseriate rays are clearly seen in this cross-sectional view. Different earlywood and latewood pores, broad rays and pore arrangement, and initial earlywood band are observable. The pores are arranged in irregular concentric bands that are tangential in the earlywood are wavy bands (ulmiform pore arrangement). A few nested pores or pore clusters are also identifiable. Few tyloses can be seen in the vessel elements as well. A few solitary pores that do not touch any other pores are clearly seen. Possible ring-porous wood with clear separation between earlywood and latewood pores in **Figure 4C**.

In **Figure 4D**, the growth ring bands are visible, parenchyma cells noticeable, and uniseriate rays are evident. The growth ring

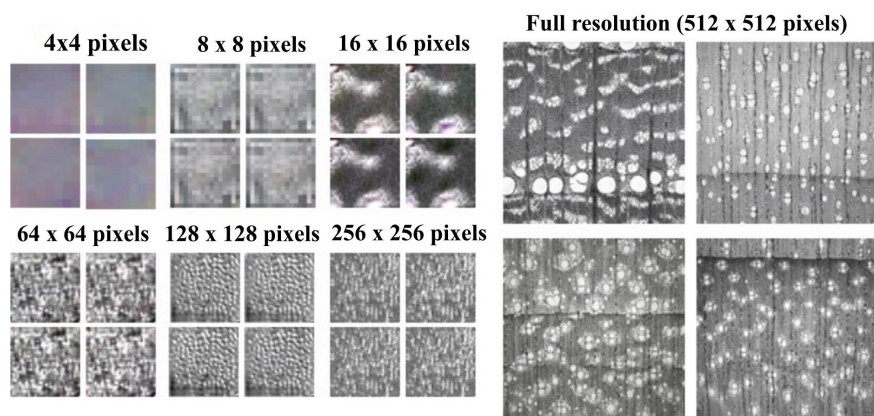


FIGURE 2 | Overview of StyleGAN training using progressively increased image resolution from 4×4 pixels to 512×512 pixels.

Examples of StyleGAN generated microscopic cross section images

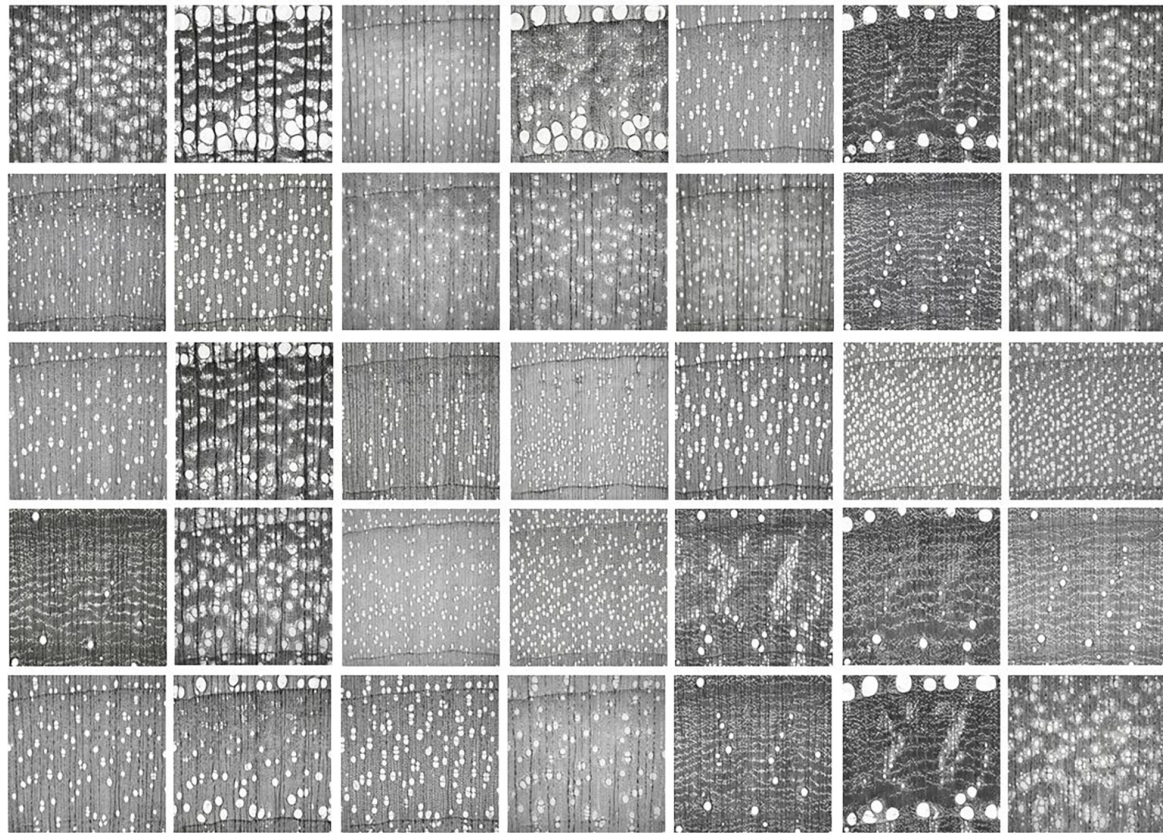


FIGURE 3 | Examples of synthetic cross-section images of hardwood species produced by the StyleGAN.

boundary is clearly delineated by a line of marginal parenchyma as several cells thick of longitudinal parenchyma. There are numerous pore multiples that occur throughout, where two or more pores are connected to another pore. The radially arranged series of pore multiples or closely arranged solitary pores are visible as pore chains. These characteristics along with no clear separation between earlywood and latewood pores transitions, the small vessel element pore sizes, uniform pore size, and evenly distribution of the pores make this a possible diffuse-porous wood in **Figure 4D**.

The potential applicability of generative adversarial in wood science and technology is tremendous. As macroscopic cross-section datasets become publicly available for research and development, especially from tropical species, GANs can be trained to generate unlimited numbers of realistic cross-sections of endangered wood species listed by CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora). The synthetic and meaningful images could then be implemented to train, validate, and test current deep learning wood species recognition models. The methodology of this work could potentially eliminate economic and processing burdens in acquiring images of tropical species for machine-learning purposes. Furthermore, the GANs framework proposed herein is a logical step to increase collaboration among academia, research

laboratories, local, state, and federal agencies, private sector, and the industry.

Another innovative use of the StyleGAN framework demonstrated in this work is to generate anatomical elements of a hybrid from two targeted parental species. The training of GAN on microscopic cross-section images from two parental species would potentially generate a hybrid species. The generated hybrid would then be validated by a real hybrid species. If the generated hybrid possesses relevant and accurate information, this technology could potentially steer a series of new research directions within the wood science and technology field, especially in breeding and genetics for estimating wood permeability, strength, density, and calculating the hydraulic potential of the tree trunk of a species that has not even been planted.

While the StyleGAN implementation appears to be very useful in creating realistic and meaningful microscopic cross-section images for more robust deep learning models and targeted biological engineering, it could also create content to facilitate training and education in wood anatomy. The realistic images could provide personalized interactions based upon an individual's experience and areas of expertise. For students interested in anatomical elements, the GAN

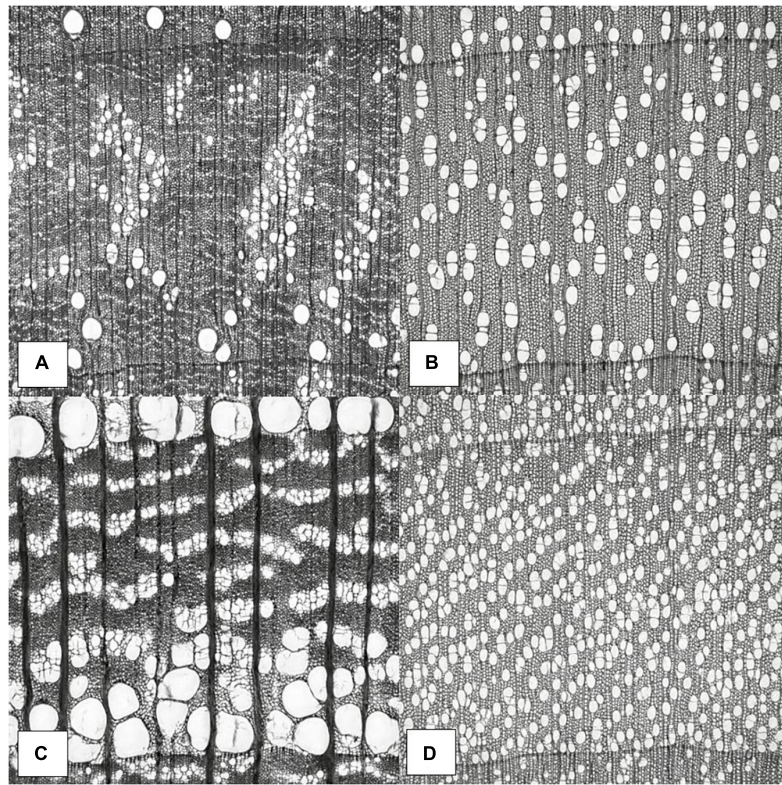


FIGURE 4 | Wood anatomy images generated by StyleGAN. Synthesis of anatomical elements. Potential semi-ring-porous wood (A), diffuse-porous wood (B), ring-porous wood (C), and diffuse-porous wood (D).

frameworks could provide new content that would help in training a new workforce faster and cheaper. In that case, this work has the capabilities of extending the wood anatomy and wood identification body beyond research and development.

Quantitative Analysis of Generated Images

The FID score was calculated on 5,650 images drawn from the generator. The score was calculated by using the Inception-V3-network (Szegedy et al., 2016). The FID scores are reported in **Figure 5**. It was noted that as the model was trained, the StyleGAN model decreased the FID score from 657 points to a final value of 17.38, which indicates more realistic image quality generation at full resolution of 512^2 pixels. The lower FID score of 17.38 means higher similarity between the two distributions, namely, between actual and synthetic data.

In the context of wood anatomy, it is not possible to compare the FID score to prior research or literature because this work is the first known application of generative adversarial for wood cross-sectional synthesis. However, GANs have been extensively used in different non-wood domains with comparatively low FID scores being reported. For instance, FID scores in Karras et al. (2019) were 4.40 for Flickr-Faces-HQ (FFHQ) on human faces, 2.65 for Large-scale Scene Understanding (LSUN) on a

bedroom, and 3.27 for LSUN on car datasets, using an identical model. Conversely, in research by Skandarani et al. (2021), the FID scores were 24.74, 23.72, and 29.06 for cardiac, liver, and diabetic retinopathy datasets, respectively, also using StyleGAN. It is worth mentioning that the datasets used in Karras et al. (2019) were much larger than those in Skandarani et al. (2021) and in this work.

However, FID scores do not completely ensure reliability when evaluating diversity of image data (Borji, 2019). In order to further quantitatively assess the quality of our image synthesis, we calculated the structural similarity index for ground-truth pairs and ground-truth/generated image pairs on 5,650 actual and 5,650 generated images.

The XDD dataset used in this work consisted of hundreds of different species that would bring the SSIM to near zero (0.00) if the images were not correlated. The lower the SSIM, the more structurally different two given images are, which denotes diversity. To that end, the calculated SSIM for ground-truth training data pairs was 0.061 ± 0.015 , which indicates a highly diverse dataset. Generally, collapsed GANs would generate similar images to the training set as explained by Srivastava et al. (2017); Lala et al. (2018), and Thanh-Tung and Tran (2020). In that case, the SSIM for ground-truth and generated images for collapsed GANs would be much higher, tending to approach 1.00. In this study, the calculated SSIM for the comparison between ground-truth and StyleGAN generated images was

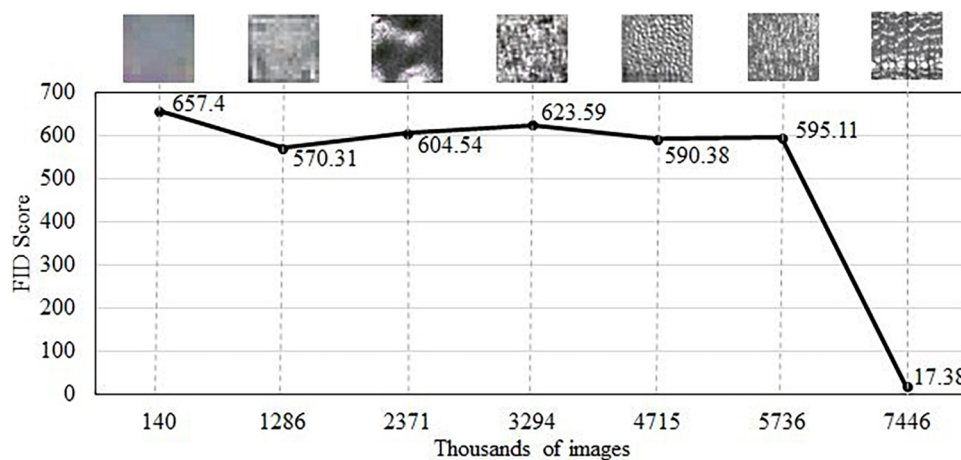


FIGURE 5 | The Fréchet inception distance (FID) score achieved by the StyleGAN generative adversarial network (GAN) on cross-section images of hardwood species. Top images show the evolution of anatomical detail with training.

0.061 \pm 0.026. The intuition is relatively simple. The lower the SSIM, the more diverse the StyleGAN generated pairs seem to be. Likewise, Odena et al. (2017) used the same concept to evaluate the diversity of generated images from the ImageNet dataset. Furthermore, in this work, the StyleGAN model generated images as diverse as the training set, where the orange and blue curves highly overlapped (see **Supplementary Material 2**).

In order to provide a clear understanding about the StyleGAN implemented in this research, we developed a graphic user interface where one can generate images of microscopic hardwood species in a menu-driven and intuitive web application. The goal of this application is to provide knowledge about StyleGAN via user interactions. The application is an open-source framework available at <https://github.com/LignumResearch/stylewood-model-usage>. It is worth noting that the user has the capability of generating unlimited amount of data (images) with this pre-trained model.

Anatomic Validation via Visual Turing Test

Table 1 summarizes the results of the realism assessment of images from the VTT by the eight wood anatomists. The mean accuracy obtained in the entire VTT was statistically lower than the random guessing [221/480 (46.04%) vs 240/480 (50.00%), respectively, $p = 0.018$]. In terms of correctly identifying generated images (specificity), there was no statistical difference between all wood anatomists and random guessing [116/240 (48.33%) vs 120/240 (50.00%), respectively, $p = 0.6717$]. Similarly, there was no statistical difference between all eight wood anatomists and random guessing to correctly identify actual images [105/240 (43.75%) vs 120/240 (50.00%), respectively, $p = 0.064$], despite the trend was in the predicted direction ($p \leq 0.05$).

By analyzing Groups I and II, there was no statistical significant difference between the two groups for accuracy, sensitivity, and specificity, respectively [45.0 vs 47.1% ($p = 0.548$),

40.8 vs 46.6% ($p = 0.317$), and 49.2 vs 47.5% ($p = 0.873$)]. The only actual species captured (100% true positive) by all wood anatomists was *Litsea glutinosa*. Additionally, none of the wood anatomists (100% false negative) captured *Zelkova serrata*, which was also an actual species. The full data regarding the VTT can be obtained in the GitHub.

In summary, results of the VTT indicated that the StyleGAN synthetically generated image fidelity comparable to actual data. The VTT data suggests that the generated images were highly realistic and indistinguishable from real microscopic cross-section images of hardwood species, regardless of the level of expertise in anatomical evaluation.

TABLE 1 | Assessment of the realism of 60 images by the eight professional wood anatomists readers by the visual Turing test (VTT).

Group	Accuracy ^a (%)	Sensitivity ^b (%)	Specificity ^c (%)
Group I^d			
Wood Anatomist 02	50.0	43.3	56.7
Wood Anatomist 04	46.7	40.0	53.3
Wood Anatomist 05	46.7	40.0	53.0
Wood Anatomist 06	45.0	63.0	26.7
Group II^e			
Wood Anatomist 01	41.7	43.3	40.0
Wood Anatomist 03	50.0	40.0	60.0
Wood Anatomist 07	40.0	36.7	43.3
Wood Anatomist 08	48.3	43.3	53.3

^aOverall mean [95% CI (confidence interval)] accuracy 46.1 (42.9–49.1).

^bOverall mean (95% CI) sensitivity 43.7 (36.9–50.5).

^cOverall mean (95% CI) specificity 48.3 (39.1–57.4).

^dGroup I: Wood anatomists with 1–5 years of experience. Mean (95% CI) accuracy 47.1 (43.8–50.4), sensitivity 46.7 (28.8–64.5), and specificity 47.5 (25.3–69.7).

^eGroup II: Wood anatomists with >5 years of experience. Mean (95% CI) accuracy 45.0 (37.2–52.8), sensitivity 40.8 (35.8–45.9), and specificity 49.15 (25.3–69.7).

CONCLUSION

This study shows that StyleGAN can successfully synthesize highly realistic and anatomically meaningful 512^2 microscopic cross-section images of hardwood species that are virtually indistinguishable from real cross-section images. We confirmed the realism and diversity for generated images by calculating the FID score, an SSIM distribution, and a VTT using two groups of professional wood anatomists with different levels of expertise.

We discussed several novel research directions involving wood anatomy and wood identification, StyleGAN, namely, data augmentation for current computer vision-based wood identification, dataset generation for wood species that are listed as threatened, endangered, or critical by CITES, and simulation of breeding between two parental woody species. Along with these applications, the StyleGAN can be used as an educational tool for improving training of a new workforce in wood anatomy and wood identification. It is our ultimate goal to provide AI/ML solutions that are reliable, economically relevant, safe, and robust to better equip the forestry and forest and wood products industries, students, researchers, staff, faculty, and enthusiasts in the field.

Future research will focus of exploring latent space when generating images. It would allow us to explore single attributes of a given species, for example porosity, ray thickness, growth ring, etc. to potentially increase model's generalization. Specifically, this research would increase the meaning and realism of images and enable targeted effects on the generated images. Additionally, GANs can perform multimodal learning that enables image synthesis by feature description.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://repository.kulib.kyoto-u.ac.jp/dspace/handle/2433/250016>.

REFERENCES

- Borji, A. (2019). Pros and cons of GAN evaluation measures. *Comput. Vision Image Underst.* 179, 41–65. doi: 10.1016/j.cviu.2018.10.009
- Chuquicuma, M. J. M., Hussein, S., Burt, J., and Bagci, U. (2018). "How to fool radiologists with generative adversarial networks? A visual turing test for lung cancer diagnosis," in *Proceeding of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)* (IEEE), 240–244. doi: 10.1109/ISBI.2018.8363564
- de Geus, D., Meletis, P., Lu, C., Wen, X., and Dubbelman, G. (2021). "Part-aware panoptic segmentation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5485–5494.
- Esteves, B. M., and Pereira, H. M. (2008). Wood modification by heat treatment: a review. *BioResources* 4, 370–404. doi: 10.15376/biores.4.1.370-404
- Fabijańska, A., Danek, M., and Barniak, J. (2021). Wood species automatic identification from wood core images with a residual convolutional neural

AUTHOR CONTRIBUTIONS

DL, JM, GSB, and EE established the scope of the project. DL and GB collected the dataset. DL and RO provided the context and contributed to the introduction. DL, JM, GSB, GB, GM, and RO prepared the dataset, implemented the machine learning pipeline, and analyzed the data. DL, GB, and EE wrote the manuscript. All authors read and approved the final manuscript.

FUNDING

DL, EE, GSB, and RO would like to thank the support of the U.S. Department of Agriculture (USDA), Research, Education, and Economics (REE), Agriculture Research Service (ARS), Administrative and Financial Management (AFM), Financial Management and Accounting Division (FMAD), Grants and Agreements Management Branch (GAMB), under Agreement No. 58-0204-9-164 for funding this novel research. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the USDA. The authors would also like to thank Mississippi State University (MSU), Division of Agriculture, Forestry, and Veterinarian Medicine (DAFVM), College of Forest Resources (CFR), Forest and Wildlife Research Center (FWRC), and the Department of Sustainable Bioproducts (DSB) for the extramural grant funding to conduct this research. This publication is a contribution of the Forest and Wildlife Research Center (FWRC) at Mississippi State University. The FWRC publication number is SB 1037. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA. The USDA is an equal opportunity provider and employer. GB, GM, and JM contributed to this work under independent R&D support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.760139/full#supplementary-material>

network. *Comput. Electr. Agric.* 181:105941. doi: 10.1016/j.compag.2020.105941

- Filho, P. L. P., Oliveira, L. S., Nisgoski, S., and Britto, A. S. (2014). Forest species recognition using macroscopic images. *Mach. Vision Appl.* 25, 1019–1031. doi: 10.1007/s00138-014-0592-7
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2014). *Generative Adversarial Nets*. Montreal, QC: Advances in Neural Information Processing Systems.
- Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., and Courville, A. (2017). "Improved training of wasserstein GANs," in *Proceedings of the 31st International Conference on Advances in Neural Information Processing Systems*.
- Hafemann, L. G., Oliveira, L. S., and Cavalin, P. (2014). "Forest species recognition using deep convolutional neural networks," in *Proceeding of the 2014 22nd International Conference on Pattern Recognition (IEEE)*, 1103–1107. doi: 10.1109/ICPR.2014.199

- Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., and Hochreiter, S. (2018). "GANs trained by a two time-scale update rule converge to a local nash equilibrium," in *proceeding of the Advances in Neural Information Processing Systems*.
- Hore, A., and Ziou, D. (2010). "Image quality metrics: PSNR vs. SSIM," in *Proceeding of the 2010 20th International Conference on Pattern Recognition (IEEE)*, 2366–2369. doi: 10.1109/ICPR.2010.579
- Hwang, S. w., and Sugiyama, J. (2021). Computer vision-based wood identification and its expansion and contribution potentials in wood science: a review. *Plant Methods* 17, 1–21. doi: 10.1186/s13007-021-00746-1
- Hwang, S. w., Kobayashi, K. K., and Sugiyama, J. J. (2020b). "Evaluation of a model using local features and a codebook for wood identification," in *Proceeding of the IOP Conference Series: Earth and Environmental Science*.
- Hwang, S. w., Kobayashi, K., and Sugiyama, J. (2020a). Detection and visualization of encoded local features as anatomical predictors in cross-sectional images of lauraceae. *J. Wood Sci.* 66:16. doi: 10.1186/s10086-020-01864-5
- Hwang, S. w., Kobayashi, K., Zhai, S., and Sugiyama, J. (2018). Automated identification of lauraceae by scale-invariant feature transform. *J. Wood Sci.* 64, 69–77. doi: 10.1007/s10086-017-1680-x
- Jansen, S., Kitin, P., De Pauw, H., Idris, M., Beeckman, H., and Smets, E. (1998). Preparation of wood specimens for transmitted light microscopy and scanning electron microscopy. *Belgian J. Bot.* 131, 41–49.
- Karras, T., Aila, T., Laine, S., and Lehtinen, J. (2018). "Progressive growing of GANs for improved quality, stability, and variation," in *Proceeding of the 6th International Conference on Learning Representations, ICLR 2018Conference Track Proceedings*, 1–26.
- Karras, T., Laine, S., and Aila, T. (2019). "A style-based generator architecture for generative adversarial networks," in *Proceeding of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (IEEE)*, 4396–4405. doi: 10.1109/CVPR.2019.00453
- Kingma, D. P., and Ba, J. (2015). "Adam: a method for stochastic optimization," in *Proceeding of the 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*.
- Kobayashi, K., Kegasa, T., Hwang, S. W., and Sugiyama, J. (2019). Anatomical features of fagaceae wood statistically extracted by computer vision approaches: some relationships with evolution. *PLoS One* 14:e0220762. doi: 10.1371/journal.pone.0220762
- Lala, S., Shady, M., Belyaeva, A., and Liu, M. (2018). *Evaluation of Mode Collapse in Generative Adversarial Networks*. Waltham, MA: High Performance Extreme Computing, IEEE, 10.
- Leggatte, W., Shirmohammadi, M., McGavin, R. L., Chandra, K., Knackstedt, M., Knuefing, L., et al. (2020). Influence of wood's anatomical and resin traits on the radial permeability of the hybrid pine (*Pinus Elliottii* x *Pinus Caribaea*) wood in australia. *BioResources* 15, 6851–6873. doi: 10.15376/biores.8.3.6851-6873
- Lengowski, E. C., Bonfatti, E. A., Nisgoski, S., de Muñiz, G. I. B., and Klock, U. (2020). Properties of thermally modified teakwood. *Maderas. Ciencia y Tecnología* 23, 1–16. doi: 10.4067/S0718-221X2021000100410
- Lens, F., Liang, C., Guo, Y., Tang, X., Jahanbanifard, M., da Silva, F. S. C., et al. (2020). Computer-assisted timber identification based on features extracted from microscopic wood sections. *IAWA J.* 41, 660–680. doi: 10.1163/22941932-bja10029
- Lopes, D. V. Jr., Bobadilha, G. d. S., Burgreen, G. W., and Entsminger, E. D. (2021). Identification of north american softwoods via machine-learning. *Can. J. Forest Res.* 51, 1245–1252. doi: 10.1139/cjfr-2020-0416
- Lopes, D. V. Jr., Burgreen, G. W., and Entsminger, E. D. (2020). North american hardwoods identification using machine-learning. *Forests* 11:298. doi: 10.3390/f11030298
- Martins, J., Oliveira, L. S., Nisgoski, S., and Sabourin, R. (2013). A database for automatic classification of forest species. *Mach. Vision Appl.* 24, 567–578. doi: 10.1007/s00138-012-0417-5
- Nie, D., Trullo, R., Lian, J., Wang, L., Petitjean, C., Ruan, S., et al. (2018). Medical image synthesis with deep convolutional adversarial networks. *IEEE Trans. Biomed. Eng.* 65, 2720–2730. doi: 10.1109/TBME.2018.2814538
- Odena, A., Olah, C., and Shlens, J. (2017). "Conditional image synthesis with auxiliary classifier GANs," in *Proceeding of the 34th International Conference on Machine Learning*, 2642–2651.
- Olschofsky, K., and Köhl, M. (2020). Rapid field identification of cites timber species by deep learning. *Trees Forests People* 2:100016. doi: 10.1016/j.tfp.2020.100016
- Pang, Y., Lin, J., Qin, T., and Chen, Z. (2021). Image-to-image translation: methods and applications. *arXiv [preprint]*. arXiv: 2101.08629,
- Park, H. Y., Bae, H. J., Hong, G. S., Kim, M., Yun, J., Park, S., et al. (2021). Realistic high-resolution body computed tomography image synthesis by using progressive growing generative adversarial network: visual turing test. *JMIR Med. Informatics* 9:e23328. doi: 10.2196/23328
- Pauli, V., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., et al. (2020). Scipy 1.0: fundamental algorithms for scientific computing in python. *Nat. Methods* 17, 261–272. doi: 10.1038/s41592-019-0686-2
- Skandarani, Y., Jodoin, P. M., and Lalande, A. (2021). GANs for medical image synthesis: an empirical study. *arXiv [preprint]*. arXiv: 2105.05318,
- Stupianek, A., Dolzblasz, A., and Sokołowska, K. (2021). Xylem parenchyma—role and relevance in wood functioning in trees. *Plants* 10:1247. doi: 10.3390/plants10061247
- Srivastava, A., Russell, C., Valkov, L., Gutmann, M. U., and Sutton, C. (2017). "VeeGAN: reducing mode collapse in gans using implicit variational learning," in *Proceeding of the 31st Conference on Neural Information Processing Systems*, 3310–3320.
- Sugiyama, J., Hwang, S. W., Zhai, S., Kobayashi, K., Kanai, I., and Kanai, K. (2020). *Xylarium Digital Database for Wood Information Science and Education (XDD_016)*. Kyoto: Kyoto University Research Information Repository, doi: 10.14989/XDD_016
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2016). "Rethinking the inception architecture for computer vision," in *proceeding of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (IEEE)*, 2818–2826. doi: 10.1109/CVPR.2016.308
- Thanh-Tung, H., and Tran, T. (2020). "Catastrophic forgetting and mode collapse in GANs," in *proceeding of the 2020 International Joint Conference on Neural Networks*, 1–10.
- Wheeler, E. A., Baas, P., and Gasson, P. E. (1989). "IAWA list of microscopic features for hardwood identification," in *International Association of Wood Anatomists at the National Herbarium of the Netherlands*, Vol. 10, (Leiden), 219–332.
- Wu, H., Liu, X., An, W., Chen, S., and Lyu, H. (2020). A deep learning approach for efficiently and accurately evaluating the flow field of supercritical airfoils. *Comput. Fluids* 198:104393. doi: 10.1016/j.compfluid.2019.104393
- Yi, X., Walia, E., and Babyn, P. (2019). Generative adversarial network in medical imaging: a review. *Med. Image Anal.* 58:101552. doi: 10.1016/j.media.2019.101552
- Zhang, Y., and Cai, L. (2006). Effects of steam explosion on wood appearance and structure of sub-alpine fir. *Wood Sci. Technol.* 40, 427–436. doi: 10.1007/s00226-005-0053-6

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Lopes, Monti, Burgreen, Moulin, dos Santos Bobadilha, Entsminger and Oliveira. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Darkfield and Fluorescence Macrovision of a Series of Large Images to Assess Anatomical and Chemical Tissue Variability in Whole Cross-Sections of Maize Stems

Marie Berger^{1,2}, Marie-Françoise Devaux^{1*}, David Legland^{1,3}, Cécile Barron⁴, Benoit Delord² and Fabienne Guillon¹

¹ UR1268 BIA, INRAE, Nantes, France, ² Limagrain Europe, Saint-Beauzire, France, ³ PROBE Research Infrastructure, BIBS Facility, INRAE, Nantes, France, ⁴ IATE, Univ Montpellier, INRAE, Institut Agro, Montpellier, France

OPEN ACCESS

Edited by:

Lisbeth Garbrecht Thygesen,
University of Copenhagen, Denmark

Reviewed by:

Jozef Mravec,
University of Copenhagen, Denmark

Wei Zeng,
Zhejiang Agriculture and Forestry
University, China

*Correspondence:

Marie-Françoise Devaux
marie-francoise.devaux@inrae.fr

Specialty section:

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

Received: 11 October 2021

Accepted: 26 November 2021

Published: 14 December 2021

Citation:

Berger M, Devaux M-F,
Legland D, Barron C, Delord B and
Guillon F (2021) Darkfield
and Fluorescence Macrovision of a
Series of Large Images to Assess
Anatomical and Chemical Tissue
Variability in Whole Cross-Sections
of Maize Stems.
Front. Plant Sci. 12:792981.
doi: 10.3389/fpls.2021.792981

The proportion and composition of plant tissues in maize stems vary with genotype and agroclimatic factors and may impact the final biomass use. In this manuscript, we propose a quantitative histology approach without any section labelling to estimate the proportion of different tissues in maize stem sections as well as their chemical characteristics. Macroscopic imaging was chosen to observe the entire section of a stem. Darkfield illumination was retained to visualise the whole stem cellular structure. Multispectral autofluorescence images were acquired to detect cell wall phenolic compounds after UV and visible excitations. Image analysis was implemented to extract morphological features and autofluorescence pseudospectra. By assimilating the internode to a cylinder, the relative proportions of tissues in the internode were estimated from their relative areas in the sections. The approach was applied to study a series of 14 maize inbred lines. Considerable variability was revealed among the 14 inbred lines for both anatomical and chemical traits. The most discriminant morphological descriptors were the relative amount of rind and parenchyma tissues together with the density and size of the individual bundles, the area of stem and the parenchyma cell diameter. The rind, as the most lignified tissue, showed strong visible-induced fluorescence which was line-dependant. The relative amount of para-coumaric acid was associated with the UV-induced fluorescence intensity in the rind and in the parenchyma near the rind, while ferulic acid amount was significantly correlated mainly with the parenchyma near the rind. The correlation between lignin and the tissue pseudospectra showed that a global higher amount of lignin resulted in a higher level of lignin fluorescence whatever the tissues. We demonstrated here the potential of darkfield and autofluorescence imaging coupled with image analysis to quantify histology of maize stem and highlight variability between different lines.

Keywords: autofluorescence multispectral imaging, darkfield imaging, quantitative histology, macrovision, maize stem

INTRODUCTION

Maize is a major productive crop worldwide and the most widely used forage crop in dairy cow feeding (Boon et al., 2008; Barros-Rios et al., 2012; Barrière, 2017). In addition, maize stems are considered one of the key lignocellulosic feedstocks to produce biofuels and other value-added chemicals (Carpita and McCann, 2008; Barrière et al., 2009b; Melati et al., 2019). Many of these uses involve efficient degradation of the cell walls either by enzymes or microbes.

The cell wall in maize stems is a complex polymer network of polysaccharides, namely, cellulose, hemicelluloses, and phenolics, as well as other minor compounds, such as pectins and structural proteins (Carpita and McCann, 2008). Cell wall phenolics comprise lignins and hydroxycinnamates and para-coumaric and ferulic acid derivatives. Lignin is a heterogeneous mixture of randomly polymerised phenolic monolignols (Vanholme et al., 2019). In maize stems, the amount of lignin in the cell wall fraction typically ranges between 12 and 20% (Jung and Buxton, 1994; Méchin et al., 2000; Barrière et al., 2009a,b). Para-coumaric acid, which accounts for approximately 1.5–2.5% of the cell wall, is mainly associated with lignin, while ferulic acid, which accounts for approximately 0.7% of the cell wall, is either linked to lignin or ester linked to hemicelluloses (Ralph et al., 1998; Méchin et al., 2000; Jung and Casler, 2006a; Barrière et al., 2009a; Hatfield et al., 2017). Ferulic acid can mediate cross-linkages between hemicelluloses and lignins or between hemicellulosic chains *via* diferulic bridges (Ralph et al., 1998; Hatfield et al., 2017; Terrett and Dupree, 2019). The amount of lignins in the cell wall, their variable structure, and the cross-linkages between cell wall components have been suggested to have a variable depressive effect on cell wall polysaccharide degradation by enzymes or microorganisms (Méchin et al., 2000; Fontaine et al., 2003; Jung and Phillips, 2010; Barrière, 2017; Casler and Jung, 2017; Hatfield et al., 2017; Terrett and Dupree, 2019). This effect is related to the limited access of enzymes or microbes to degradable/fermentable polysaccharides (Meng and Ragauskas, 2014; Melati et al., 2019). However, the main determinant among these factors for recalcitrance is still not clear (McCann and Carpita, 2015; Melati et al., 2019; Zoghalmi and Paës, 2019), which is partly due to the large variability of biomass and assignment of biomass as a bulk material without considering the heterogeneity of plant cell walls according to organs and tissues.

In fact, within a species, the cell wall composition depends on the genotypes and the plant-breeding environment but also on other components, such as the organs, stems, and leaves to the tissues and cell types. For example, a maize stem or internode is composed of different tissues, namely, rind, parenchyma and vascular bundles, whose proportions, morphologies and compositions vary according to the genotype, maturity and agro-climatic conditions (Cone and Engels, 1993; Morrison et al., 1998; Jung and Casler, 2006a,b; Legland et al., 2017; Perrier et al., 2017; El Hage et al., 2018; Zhang et al., 2019). It has been shown that these tissues differ in their fermentation/digestibility yield and rate, which has been related to their cell wall composition (Akin, 1989; Scobbie et al., 1993; Wilson et al., 1993; Wilson and Mertens, 1995; Hatfield et al., 1999; Jung and Casler, 2006b;

Barros-Rios et al., 2012; Ding et al., 2012; Devaux et al., 2018). Several authors have reported that the relative proportion of tissues and lignin distribution within organs can explain the differences in digestibility observed at an equivalent stage of maturity (Akin, 1989; Wilson and Mertens, 1995; Méchin et al., 2005; Barros-Rios et al., 2012).

To study the histological features of plant organs, methods are required to quantify the proportion of tissues and their composition. Microscopic techniques are generally proposed for this purpose. However, these methods are not compatible with large-scale or high-throughput studies. In the case of maize stems, the stem section area can be of 1–2 cm² while the cells diameter can be of approximately 60 µm (Zhang et al., 2013; Legland et al., 2014, 2017); moreover, the experiments often need to be repeated to tackle the biological variability. For a few years, whole stem section imaging has been developed (Zhang et al., 2013, 2019; Legland et al., 2014, 2017, 2020; Heckwolf et al., 2015; Perrier et al., 2017). Images of hand- or microtome-cut stem cross-sections were acquired with either a microscope, a microscope slide scanner or a flatbed scanner. Different modes of illumination (darkfield, brightfield, and epifluorescence, etc.) associated or not with contrast-enhancing methods, such as Fagsta staining (Tolivia and Tolivia, 1987), are implemented to visualise the tissues. Other authors favour 3D imaging and use micro-computed tomography technology for stem imaging (Zhang et al., 2018, 2020, 2021). Optical macrovision systems have the advantage of being relatively inexpensive compared to more sophisticated equipment, such as X-ray tomographs and are well suited for studying histology because they combine a large field of view and good spatial resolution, thus allowing for observations of a whole stem cross-section and differentiation of the different tissues (Legland et al., 2014; Corcel et al., 2016).

Regardless of the image acquisition methods, image analysis is required to identify and quantify morphological features, which are also called anatomical traits. Maize stems include the proportions of tissues, e.g., rind, parenchyma cells, and vascular bundles, and the morphology and density of cells and vascular bundles. Different image analysis workflows have been proposed, which depend on the targets and on the contrast in the images. Most workflows include a tissue segmentation step followed by morphological feature quantification. Heckwolf et al. (2015) developed custom image processing software that utilises a variety of global thresholding and local filtering to extract rind, pith and vascular bundle sizes from stem cross-section scanned images. Legland et al. (2017) proposed a series of morphological filters to identify the rind and vascular bundles in the pith from stem cross-sections after Fagsta staining. Zhang et al. (2018, 2020, 2021) presented an image analysis pipeline to extract micro-phenotypic traits from 3D tomography images that combine threshold-based segmentation and morphological operations.

Once tissues are segmented, it is a straightforward process to measure the rind thickness, pith area, vascular bundle area or vascular bundle size or shape. These descriptors were related to stem lodging (Zhang et al., 2018), developmental stages (Zhang et al., 2020), and water stress (Legland et al., 2017; El Hage et al., 2018) or used to analyse the phenotypic variation between lines (El Hage et al., 2018; Zhang et al., 2021). In addition to tissue

segmentation, Devaux and Legland (2014) proposed applying grey-level granulometry using morphological closings to directly extract cell size distributions from grey-level images. Legland et al. (2020) used the method on maize stem images to compute the parametric maps of cell size.

Chemical imaging techniques are required to reveal the variations in the cell wall composition of tissues or cell types. Specific staining methods or spectral imaging can be used for this purpose, and each technique leads to very different image analyses. Several authors have used Fasca staining to assess the distribution of lignin in maize or sorghum stem tissue according to developmental stages (Zhang et al., 2013, 2019) or in response to water deficit (Legland et al., 2017; Perrier et al., 2017; El Hage et al., 2018). Fasca staining colours lignified tissues in red and non-lignified tissues in blue. In Zhang et al. (2013, 2019), lignification was assessed by the ratio of red to blue intensity. The image analysis workflow made it possible to assess the distribution of lignin within a cross-section by profiling the red/blue intensity ratio from the epidermis to the centre of the cross-section. The image processing workflow was further improved and fully automated by Legland et al. (2017), and it was designed for measuring the amount of blue and red intensities in the parenchyma and the amount of red intensity in the rind. Perrier et al. (2017) also developed a dedicated tool in ImageJ software for analysing Fasca-stained cross-sections from sorghum internodes. The dedicated script allowed quantification of the outer zone area in percentage of internode cross-section area, the percentage of sclerenchyma tissue in the outer zone, the percentage of nonlignified tissue in the central zone of the internode and the density of vascular bundles in the central zone.

Apart from histochemical staining, spectral imaging techniques have been proposed to perform chemical mapping of cell wall variations in wood or plant stems. Microspectroscopy, such as Fourier transform infrared (FT-IR) or Raman microspectroscopy, is very useful to study carbohydrates or phenolic constituents (lignin or hydroxycinnamic acids) (Gierlinger, 2018; Beć et al., 2020). The main drawback is that the techniques are time-consuming, thus allowing for the mapping of only a small region of the sample, which limits the application of these techniques for the comparison of large numbers of samples. To monitor the chemical variation in tissue composition, multispectral fluorescence imaging can be applied (Corcel et al., 2016). Full-field fluorescence macroscopy has a sufficient spatial resolution ($\approx 3 \mu\text{m}$ per pixel), high acquisition speed and large fields of view. Taking advantage of the autofluorescence properties of many plant compounds, fluorescence imaging can be performed with little tissue preparation and, more importantly, without labelling. Fluorescence imaging techniques have two main attributes over other techniques associated with their greater sensitivity and selectivity due to the unique properties of autofluorescent molecules being excited at a specific wavelength and emitting radiation at specific wavelengths. Plant cell wall autofluorescence is mainly linked to the presence of phenolic compounds, such as lignin and hydroxycinnamic acids. Hydroxycinnamic acids emit blue fluorescence under UV excitation at approximately 350 nm (Fulcher et al., 1971; Harris and Hartley, 1976; Lang et al., 1991), while lignin excited using

UV and visible light emitted blue, green and red fluorescence (Djikanović et al., 2007; Donaldson et al., 2010; Donaldson, 2013, 2020; Donaldson and Williams, 2018). The nature of the phenolic compounds, their variable relative proportions and the environment (pH, presence of quenching molecules, etc.) result in variable tissue fluorescence responses that can be interpreted as a fluorescence tissue signature.

The analysis of multispectral images requires specific analysis tools that can account for both the spatial and the spectral dimensions of the image. Using the chemometric approach, the first step in the analysis of multispectral images is to process the spectral dimension of the data (Geladi and Grahn, 2006; Ghaffari et al., 2019). Spectral information can be extracted either manually or automatically from regions in the images (de Juan et al., 2009). In many cases, regions in the image are segmented based on the spectral information (Salzer and Siesler, 2014).

In this study, we developed a quantitative histology approach to estimate the proportion of different tissues in maize stem sections and associated a chemical profile with each of these tissues. Two macroscopic imaging techniques without prior labelling of the tissues were used. Darkfield macroscopy was chosen to visualise the different tissues independently of their chemical composition. In parallel, multispectral autofluorescence macroscopy was used to associate a multispectral autofluorescence profile to the tissues with the aim of evaluating the relative distribution of lignin and hydroxycinnamic acids. In the darkfield images, tissues and cells are visualised based on the diffraction properties of the light by the cell walls. An image analysis workflow was implemented to identify the tissues and then extract 2D morphological descriptors. We propose that a simple stem model can be used to estimate the volume descriptors of the amount of rind, vascular bundles and parenchyma cell walls. We sought to measure multispectral autofluorescence pseudospectra in each tissue. Tissues were also segmented from the multispectral images using a "sum of intensities" image and a workflow similar to that of the darkfield images. Because the parenchyma near the rind has been revealed to have specific enzymatic degradation properties (Jung and Casler, 2006b; Devaux et al., 2018), two regions of parenchyma were considered, and we evaluated and compared their fluorescence properties, i.e., relative amounts of lignin and hydroxycinnamic acid. Two stem internodes of 14 inbred lines were analysed with the aim of demonstrating the feasibility of the method suggested here. Correlations between the histological descriptors and the amounts of phenolic compounds and digestibility measured at the stem level for the 14 inbred lines were examined.

MATERIALS AND METHODS

Plant Material and Stem Cell Wall Characterisation

Plant Material

Fourteen maize inbred lines selected for their contrasting digestibility were grown in Arras (France) in 2018. Twelve plants

per inbred line were harvested at the silage stage. The stems were separated from leaves, panicles and ears. The internode located under the main ear was collected for two plants per line and stored in 70% ethanol/water (v/v) for quantitative histology. The stems of the remaining 10 plants were pooled, chopped and oven dried (70°C). The dried stems were ground with a hammer mill to pass through a 1 mm screen for the analysis of phenolic compounds and cell wall enzymatic digestibility.

Chemical Analysis

Cell wall material was prepared from the 10 dried and ground pooled stems. The ground material was placed in 80% ethanol at 100°C in an automated solvent extractor (ASE 350, Dionex Sunnyvale, CA, United States; 6 min flow time, 2 mL/min flow rate, 150% flush, and 30 s purge). The ethanol insoluble material was taken as the cell wall estimate (Chazal et al., 2014) and therefore called cell wall content and expressed in percent of the dry matter. Using the automated solvent extractor, the standard deviation is less than 1%.

The Klason lignin content was measured according to Dence (1992). Ester-linked para-coumaric and ferulic acids were measured after mild alkaline hydrolysis as described by Ho-Yue-Kuang et al. (2016). Analyses were performed in duplicate, and the results are expressed as the percentage of dry matter.

Digestibility Measurement

The enzymatic digestibility was measured in duplicate on the extractive-free material using the Aufrère and Michalet-Doreau method (Aufrère and Michalet-Doreau, 1983). The technique involves three stages: (1) pretreatment with pepsin (pepsin Merck 2000 FIP U/g Art7190) in hydrochloric acid (0.2% pepsin in 0.1 N HCl in a water bath at 40°C for 24 h; (2) starch hydrolysis in a water bath in the same mixture for *exactly* 30 min at 80°C; and (3) attack by cellulase (cellulase Onozuka R 10 extracted from *Trichoderma viride*, Yakult Honsha Co. Ltd, Japan, 1 g/L in 0.05 M sodium acetate buffer, pH 4.6) after filtration and rinsing for 24 h in a water bath at 40°C. The final residue was weighed. Due to the low starch content (<2% of the dry matter content of the alcohol-insoluble material), cell wall digestibility was equated with dry matter digestibility and calculated as follows:

$$\text{IVCWD} = \frac{M1 - M2}{M1} * 100 \quad (1)$$

where M1 is the dry mass of the extractive-free sample and M2 is the dry mass of the residue after enzymatic degradation.

Image Acquisition

Sample Sectioning for Histological Analysis

For the two internodes retained for histology, a one cm long segment was sampled in the middle of the internode. For each segment, 150 µm thick cross-sections (called sections in the following) were cut in air with a gsl1 microtome (Design and production: Lucchinetti, Schenkung Dapples, Zurich, Switzerland) (Gärtner et al., 2014) and stored in 70% ethanol at 4°C until image acquisition. Prior to image acquisition, the sections were rehydrated in water overnight at 4°C to remove the air.

Darkfield Imaging

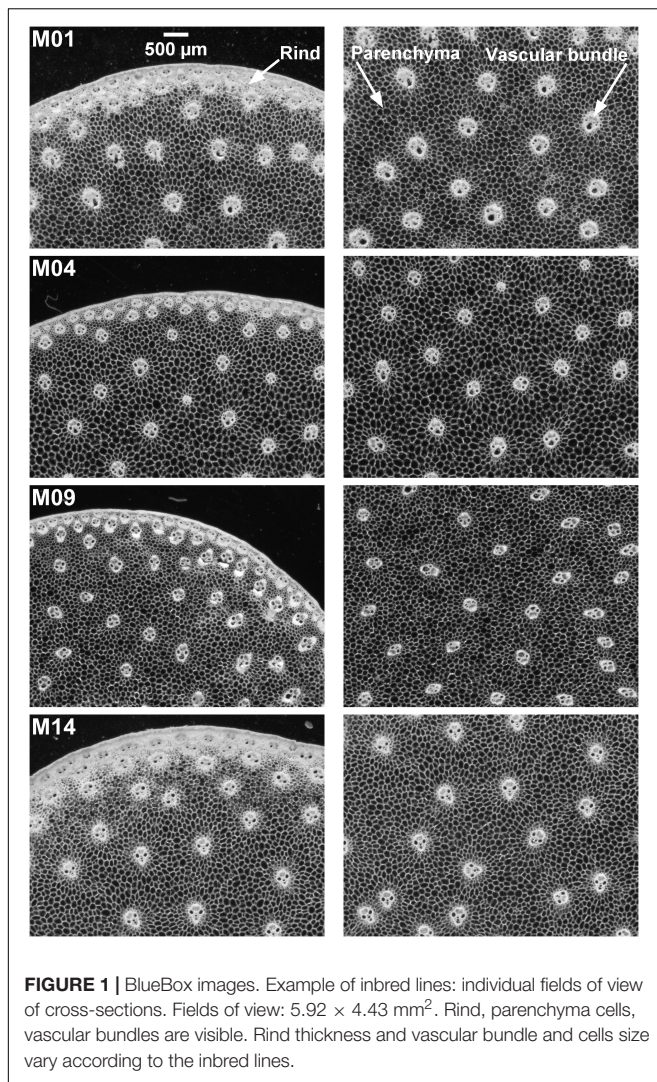
Images were acquired using the “BlueBox” macrovision acquisition prototype specially designed to observe plant tissue sections at the macroscopic scale without any prior labelling steps (Devaux et al., 2008, 2009). A monochrome CCD camera (Prosilica Digital Camera DCAM 1.31 – distributed by Alliance Vision, Montélimar, France) was equipped with a 1.2X magnification lens (Navitar Precise Eye, Rochester, NY, United States). With these settings, the images were $1,620 \times 1,220$ pixels and corresponded to a field of view of $5.92 \times 4.43 \text{ mm}^2$, with a pixel size of 3.63 µm. Grey levels were coded between 0 (black) and 255 (white). An optical fibre ring was connected to an intensity-controlled light source (SCHOTT DCRIV Light Source, Mainz, Germany) and placed under the samples to provide darkfield illumination. Motorised stages for positioning the camera and the samples allowed for the acquisition of large images. All elements were placed in a box to prevent outside light from entering. Homemade software developed under LabView was used for image acquisition.

Sections were placed between two round lamellae for observation. Mosaic images, called large images, were acquired to observe the entire sections. The largest images corresponded to a field of view of $20 \times 20 \text{ mm}^2$. Two sections per internode were imaged. Images of one internode were removed for M06 because air was still present after overnight rehydration. Finally, 54 large images were obtained. Examples of individual fields of view and large images can be seen in **Figures 1, 2**. Several tissues were observed within the stem sections: the rind, the vascular bundles and the pith parenchyma (Esau, 1977).

Multispectral Autofluorescence Imaging

Autofluorescence images were acquired using a Multizoom AZ100M fluorescence microscope (Nikon, Japan) equipped with a Q Imaging EXI Aqua monochrome camera plus an RGB-HM-S-IR filter wheel for colour image acquisition. The system provides $1,392 \times 1,040$ pixel RGB images with grey-level intensities coded using 16,386 values. The total magnification was set to X4 by combining the AZ-Plan Fluor 2X lens (NA: 0.2/WD: 45 mm) and a X2 optical zoom. With these settings, the pixel size was 2.78 µm and the field of view was $3.9 \times 2.9 \text{ mm}^2$. The microscope was equipped with a Prior Proscan II (Nikon, Japan) motorised stage, which allowed large image acquisition. The INTENSILIGHT (C-HGFI/C-HGFIE Precentred Fibre Illuminator Nikon, Japan) device with a mercury lamp ensured lighting for fluorescence imaging. Four fluorescence filter cubes corresponding to two UV excitations, namely, U1 and U2, and two visible excitations, namely, blue (BL) and green (GR), were placed inside the motorised filter wheel (**Supplementary Table 1**).

The acquisition software NIS-Elements (AR 5.02.02) allows automatised multispectral acquisition of large images. The multispectral sequence was designed to successively acquire the four RGB images corresponding to the four fluorescence filters for a given field of view before moving to the next field of view. The order of acquisition was GR, BL, U2, and U1, with exposure times set after viewing a few samples (**Supplementary Table 1**). After all acquisitions, the fluorescence intensity was found to be



much lower for the two visible filters than for the UV filters, and a multiplicative factor of 2 was applied to the RGB images of the blue and green filters.

The resulting multispectral images contained 12 channels by merging the RGB images recorded using the four filter cubes (Corcel et al., 2016). The channels were put in an order from high to low wavelengths: blue, red and green channels of each RGB image acquired with filters U1, U2, blue and green. The channels names were U1b, U1g, U1r, U2b, U2g, U2r, BLb, BLg, BLr, GRb, GRg, and GRr. Channel U1r was removed from the sequence because it contained unwanted reflection from the excitation Rayleigh band. The final multispectral image therefore contained 11 channels.

For morphological image acquisition, rehydrated sections were placed in water between two round lamellae for observation. One multispectral image per internode was acquired (except for the anomalous M06 internode), and for seven inbred lines, a second section of one internode was imaged for repetition. The final set contained 34 large multispectral images.

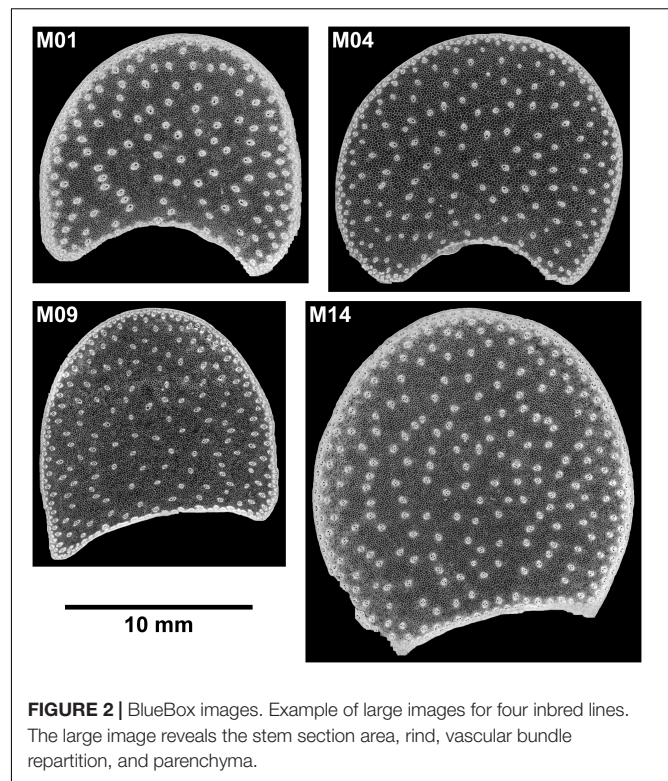


Image Analysis

Image analysis was performed in the MATLAB 2019b environment (Mathworks, Natick, MA, United States) using the image processing toolbox, dedicated homemade functions and scripts developed for BlueBox and macrofluorescence collections of images.

Image Representations

Displaying a set of large images is difficult, and the content of multispectral images cannot be viewed in a direct way. With the objective of comparing different inbred lines or sections and enabling details to be seen, a multiscale image representation was adopted. Zoom images corresponding to one field of view of the mosaic in the case of the BlueBox images were selected, and they showed details into the middle of the section and on the border of the section. Up to four large images with a resolution of $14 \mu\text{m}$ per pixel were compared. Low-resolution images ($24 \mu\text{m}$ per pixel) were finally retained to draw A4 300 dpi figures, with one image per inbred line.

In parallel, an RGB representation of the multispectral fluorescence image was implemented. The red channel of the RGB image was computed as the average of the red channels U2r, BLr, and GRr. The green channel of the RGB image was computed as the average of the green channels U1g, U2g, and BLg. The blue channel of the RGB image was computed as the average of the blue channels U1b and U2b. The RGB images were called *composite macrofluorescence images* in the following. Two grey-level images were also computed for segmentation purposes: the image “sum of fluorescence intensity of the 11 channels” and

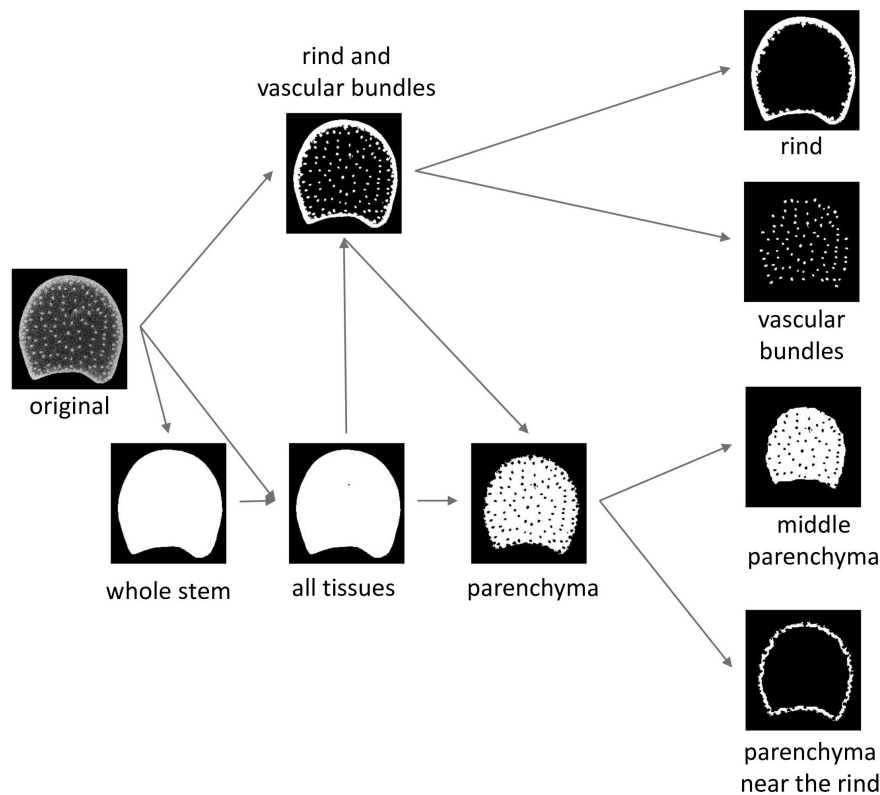


FIGURE 3 | Image segmentation workflow and resulting tissue regions of interests (ROIs). The “whole stem” ROI corresponds to the mask of the section. In the “all tissues” ROI, the holes encountered in some of the sections are segmented. Parenchyma ROI and rind and vascular bundle ROI are temporary ROIs necessary to compute tissue ROIs. Middle parenchyma ROI and rind parenchyma ROI are subregions of the parenchyma ROI.

the image “sum of visible fluorescence intensity” corresponding to the sum of the three channels BLg, BLr, and GRr.

Definition of Morphological and Autofluorescence Descriptors of Maize Stem Tissue

For each tissue, morphological and autofluorescence descriptors were defined. The stem area was retained as an absolute size descriptor. Relative areas were chosen to compare the rind, parenchyma and vascular bundle amounts. For the vascular bundles, the descriptors that were selected were the number density, individual surface area and elongation, which was defined as the width/length ratio. Parenchyma was also characterised by the cell size. Due to their different behaviour toward enzymatic degradation, two regions of parenchyma were considered (Jung and Casler, 2006b; Devaux et al., 2018): parenchyma near the rind and middle parenchyma. For each tissue, the average fluorescence properties were used to characterise the composition of cell walls. Measurements of the descriptors were performed after segmentation of the different tissues in the two types of images.

Segmentation of Tissues

Regions of interest (ROIs) corresponding to each tissue were identified for the dark field and macrofluorescence images. A semiautomated workflow was adapted from

Legland et al. (2014). The main steps summarised in **Figure 3** were similar for the two kinds of images. The specific implementations for the two kinds of images are given in **Supplementary Table 2**.

For some stems, the parenchyma was torn during cutting, resulting in the presence of holes. The *whole stem* ROI was obtained by thresholding followed by hole filling to observe the whole stem area. A second region, called *all tissue* ROI, was considered, in which the possible holes were segmented by a second thresholding operation. The objective was to obtain an ROI that avoided possible holes to measure the parenchyma cell size and fluorescence properties.

From the *all tissue* ROIs and the original images, intermediate ROIs were created to correspond to the *rind and vascular bundle* ROIs. Alternating filtering based on morphological openings and closings (Soille, 2003) was applied to contrast rinds and vascular bundles from the parenchyma. It was followed by automatic thresholding. In the resulting ROIs, some vascular bundles could be connected and the rind could be split into several fragments. Rind and vascular bundles were differentiated by size analysis. The size of the vascular bundles was determined from the mode of the size distribution of the segmented objects in the BlueBox image located at a distance greater than 1 mm from the epidermis. It was used to define a size threshold to extract the *individualised vascular bundle* ROIs. *Rind* ROIs were built by merging external

fragments larger than five times the value of the mode. In the BlueBox images, some vascular bundles were connected, and an additional region was computed to correspond to *all vascular bundle ROIs*, both connected and not connected.

The *parenchyma ROIs* were obtained as the logical difference between the *whole stem ROIs* and the *rind and vascular bundle ROIs*. The *parenchyma near the rind ROIs* and the *middle parenchyma ROIs* were obtained using *a priori* distances from the rind: below 500 μm for the parenchyma near the rind and over 1,000 μm for the middle parenchyma. Distances were chosen to contrast the two kinds of parenchymas.

Morphological Descriptors

Measuring the Raw Morphological Descriptors

Raw morphological descriptors could be directly measured from the segmented regions of interest as specified in **Table 1**. Areas were obtained by pixel counting; perimeters, vascular bundle length and width were obtained using the *regionprops* MATLAB function. Rind thickness was evaluated by granulometry using mathematical morphology transformations (Soille, 2003; Devaux et al., 2008; Legland et al., 2014, 2020); see “Parenchyma Cell Size” for an introduction to the method. The thickness distributions were obtained by applying opening transformations using squared structuring elements and a maximum size of 1,456 μm . The mean size of the distribution was taken as a measure of the average thickness of the rind.

However, some defects related to section cutting and image segmentation were observed, such as missing pieces of rind or bundles not separated. Therefore, we developed an estimate of the morphological features as described below.

Estimating Rind Area

In sections where some rind pieces were missing, the rind area $Ri(A)$ was estimated using the thickness $Ri(T)$ and the perimeter of the whole stem $St(P)$:

$$\widetilde{Ri(A)} = Ri(T) \times St(P) \quad (2a)$$

However, when examining undamaged sections, the estimated values were always higher than the measured values. The reason was an overestimation of the rind thickness due to vascular bundles that remained connected to the rind after segmentation. The average difference *diffEstMeas* between the estimated and

measured rind areas was computed from the undamaged images and used as a correction factor to estimate the rind areas:

$$\widetilde{Ri(A)} = Ri(T) \times St(P) - \text{diffEstMeas} \quad (2b)$$

Estimating Stem Area

The stem area was equal to the measure of the area of the whole stem ROI when the rind was preserved, and it was estimated when the rind was fragmented:

$$\widetilde{St(A)} = \begin{cases} St(A) : \text{rind preserved} \\ St(A) - Ri(A) + \widetilde{Ri(A)} : \text{rind fragments} \end{cases} \quad (3)$$

Computing the Relative Areas

Relative areas of measured parenchyma, estimated rind and measured vascular bundles were computed as percentages of the estimated whole stem area.

Average Morphology of Individual Vascular Bundles

Individual area and elongation measured for each vascular bundle were averaged to obtain one value per section: $\overline{Vi(A)}$ and $\overline{Vi(E)}$.

Density of the Number of Vascular Bundles

The number of vascular bundles ($Vb(N)$) was estimated as the total area of vascular bundles divided by the average area of vascular bundles:

$$Vb(N) = \frac{Vb(A)}{\overline{Vi(A)}} \quad (4)$$

The density of the number of vascular bundles ($Vb(D)$) was computed as the number $Vb(N)$ divided by the whole stem area.

Parenchyma Cell Size

Cell size was measured from the BlueBox greyscale image on the two parenchyma regions. Grey-level granulometry developed using mathematical morphology (Soille, 2003) was applied without segmenting the cells as described in Devaux et al. (2008) and Legland et al. (2020). Grey-level granulometry consists of successively applying size transformations of the image through a mask of known geometry, called a *structuring element* (Soille, 2003). The size and shape of the structuring element are chosen according to the characteristics of the image. In the BlueBox images (**Figure 1**), cells appeared as isotropic dark objects, and closing transformations using squared structuring elements were retained. Closing can be compared to sieving dark objects in the image: dark objects smaller than the structuring element are removed while preserving the size of larger objects. When closings of increasing size are applied, the sum of grey levels, measured after each operation, increases. The increase depends on the quantity of objects removed. The result is a granulometric curve expressed as a percentage of grey-level variations according to the closing step.

In the present work, closing transformations between 18 and 207 μm were applied by steps of 7.26 μm . Compared to the procedure described in Devaux et al. (2009), grey-level granulometry curves were postprocessed by subtracting the residual size variations caused by the general background of the

TABLE 1 | Measurement of the raw morphological descriptors: regions of interests (ROIs) used for measurement, morphological descriptors and acronyms.

Region of interest (ROI)	Morphological descriptors	Acronym
Whole stem ROI	Area	St(A)
	Perimeter	St(P)
Rind ROI	Area	Ri(A)
	Thickness	Ri(T)
All vascular bundles ROI	Area	Vb(A)
Individual vascular bundles ROI	Individual area	Vi(A)
	Individual Elongation = width/length	Vi(E)

Individual area or elongation means that measurements were performed for each segmented bundle.

image in the region of interest and renormalisation of the curves. Grey-level mean sizes and standard deviations were computed from the granulometric curves as described in Devaux and Legland (2014).

Measure of Autofluorescence Pseudospectra

For each pixel, 11 fluorescence intensity values were measured. The set of fluorescence intensities measured for individual pixels or averaged over a set of pixels was called pseudospectra (Corcel et al., 2016). Because no photon can be emitted at wavelengths higher than the excitation wavelength, for the two visible filters blue and green, the channels BLb, GRb, and GRr showed no signals. They were nevertheless maintained in the pseudospectra and were considered a baseline.

Average autofluorescence pseudospectra were measured for the four tissue ROIs: *rind*, *all vascular bundles*, *parenchyma near the rind* and *middle parenchyma*. A preliminary analysis revealed a channel-dependent background intensity. Three regions without any signal were manually selected in four images of the series. The background pseudospectrum was computed as their average pseudospectra. It was subtracted from all other measured pseudospectra. A section-dependent overall intensity effect was observed, which was probably due to variations in section thickness. A normalisation procedure was set, which is detailed in the “Results” section.

Data Analysis

The morphological descriptors and the autofluorescence pseudospectra were analysed based on a principal component analysis and variance analysis, followed by multiple comparisons of the estimated marginal means. Analyses were performed within the MATLAB 2019b environment (Mathworks, Natick, MA, United States) using the statistics and machine learning toolbox.

Principal component analyses were applied separately to the morphological descriptors and the autofluorescence pseudospectra. The morphological descriptors were normalised to describe the variations independently of the units, and the loadings were represented as correlation circles. In the case of pseudospectra, the variables were not normalised to avoid assigning importance to the baselines of the pseudospectra, and the loadings were represented in the form of pseudospectra.

Variance analyses were applied to morphological descriptors and principal components to determine their significance with regard to the 14 inbred lines studied. Multiple comparisons of the estimated means were applied to reveal the most contrasted lines. In the case of autofluorescence pseudospectra, analyses of variance were applied to the principal component scores to determine the effects of inbred lines, tissues and their interactions.

RESULTS

Variation in Cell Wall Phenolics and Digestibility Within the 14 Inbred Lines

The stems of the 14 lines were analysed for the cell wall content, lignin and hydroxycinnamic acid content of the cell walls, and

digestibility. The cell wall content represented on average 54.9% of the stem dry matter, with a coefficient of variation (CV) of 8.5% (Table 2). The content of esterified para-coumaric acid showed the highest variability, with an average value of 1.60% of the cell wall dry matter and a coefficient of variation of 15.56%. Lower variability was observed for Klason lignin and esterified ferulic acid contents. On average, the lignin content was 18.3%, with a coefficient of variation of 9.2%, and the esterified ferulic acid content was 0.61%, with a coefficient of variation of 9.22%. Cell wall digestibility ranged from 25.4 to 43.9%, with an average value of 33.6% and a coefficient of variation of 14.9%.

The values of the biochemical traits measured in this study are in the range of those reported in the literature for inbred lines (Jung and Buxton, 1994; Méchin et al., 2000; Barrière et al., 2009a; El Hage et al., 2018). Despite the observed variability within the 14 inbred lines for the measured traits, no correlation between these traits was found. Inbred Lines M06 and M05 showed the lowest cell wall lignin content, while the highest values were found for M02, M01, and M11. M04 had a low content of para-coumaric acid but a high amount of lignin. In contrast, M11 had both high amounts of lignin and para-coumaric acid, and M14 had a high amount of para-coumaric acid and an intermediate amount of lignin. M03 and M05 had intermediate values for lignin and para-coumaric acid, while M03 and M05 the highest and lowest values of ferulic acid content, respectively. The highest cell wall digestibility was found for Lines M05 and M06, which had the lowest amount of lignin. Although M01 and M02 had the highest lignin content, they showed intermediate cell wall digestibility.

In summary, our panel of inbred lines showed variability in the stem cell wall contents and phenolic composition and a lack of correlation between these biochemical traits.

Examples of Images From Details to the Collection

Four samples from inbred lines with contrasting morphologies were selected for a preliminary investigation of the dataset.

Zoom Images

Zoom images were selected to compare the border and the middle of the sections from four contrasting samples (Figures 1, 4). At this scale, details in the rind, vascular bundles, and parenchyma cells are visible. Cell walls appeared in white in the morphological images and had colours ranging from pink to blue in the autofluorescence images. Based on the colour representation of the autofluorescence images, blue fluorescence represents cell walls with mainly UV-induced fluorescence while pink or yellow fluorescence represents cell walls with visible-induced fluorescence. Differences between the four inbred lines were observed for all the tissues.

The rind is composed of vascular bundles and small cells of cortical parenchyma (Esau, 1977). In the morphological images of Lines M04 and M09, the vascular bundles were clearly visible, while for M01 and M14, the rind formed a larger white ribbon and the cortical cells could hardly be distinguished. The autofluorescence images showed that for M01 and M14, the cortical cells contained fragments that fluoresced red, while for M04 and M09, the cells seemed empty. The red fluorophore

TABLE 2 | Mean values for stem cell wall contents, Klason lignin (KL), hydroxycinnamic acid – esterified p-coumaric acid (Ester pCA), esterified ferulic acid (Ester FA) – and cell wall digestibility (IVCW digestibility) for the 14 inbred lines.

Inbred line	Cell wall % DM	KL % CW	Ester pCA %CW	Ester FA % CW	IVCW digestibility
M01	52.7	20.51	1.59	0.62	32.7
		<i>0.11</i>	<i>0.03</i>	<i>0.01</i>	<i>0.85</i>
		(90.27)	(7.00)	(2.73)	
M02	56.8	20.65	1.61	0.55	34.7
		<i>0.56</i>	<i>0.01</i>	<i>0.00</i>	<i>0.05</i>
		(90.53)	(7.06)	(2.41)	
M03	55.5	18.24	1.68	0.72	37.8
		<i>0.09</i>	<i>0.00</i>	<i>0.01</i>	<i>0.10</i>
		(88.37)	(8.14)	(3.49)	
M04	57	20.13	1.06	0.57	29.9
		<i>0.14</i>	<i>0.01</i>	<i>0.00</i>	<i>0.94</i>
		(92.51)	(4.87)	(2.62)	
M05	52	15.88	1.29	0.64	41
		<i>0.07</i>	<i>0.01</i>	<i>0.00</i>	<i>0.29</i>
		(89.16)	(7.24)	(3.59)	
M06	50	15.51	1.33	0.67	43.9
		<i>0.34</i>	<i>0.00</i>	<i>0.01</i>	<i>0.01</i>
		(88.58)	(7.60)	(3.83)	
M07	62	16.58	1.81	0.58	39
		<i>0.05</i>	<i>0.02</i>	<i>0.00</i>	<i>0.94</i>
		(87.4)	(9.54)	(3.06)	
M08	53.7	17.92	1.56	0.59	32.8
		<i>0.09</i>	<i>0.08</i>	<i>0.03</i>	<i>1.91</i>
		(89.29)	(7.77)	(2.94)	
M09	63.6	17.1	1.38	0.64	32.7
		<i>0.38</i>	<i>0.03</i>	<i>0.01</i>	<i>0.17</i>
		(89.44)	(7.22)	(3.35)	
M10	53.6	16.93	1.65	0.61	30.7
		<i>0.32</i>	<i>0.02</i>	<i>0.03</i>	<i>0.76</i>
		(88.22)	(8.60)	(3.18)	
M11	47.6	20.4	1.94	0.57	27.6
		<i>0.27</i>	<i>0.09</i>	<i>0.07</i>	<i>1.38</i>
		(89.04)	(8.47)	(2.49)	
M12	49.9	19.26	1.68	0.5	30.5
		<i>0.32</i>	<i>0.03</i>	<i>0.01</i>	<i>0.57</i>
		(89.93)	(7.84)	(2.33)	
M13	61.8	19.15	1.89	0.66	25.4
		<i>0.16</i>	<i>0.11</i>	<i>0.04</i>	<i>0.12</i>
		(88.25)	(8.71)	(3.04)	
M14	52.3	17.94	1.91	0.68	31.1
		<i>0.28</i>	<i>0.06</i>	<i>0.03</i>	<i>0.91</i>
		(87.38)	(9.30)	(3.31)	

% DM and % CW means that results are expressed in percent of dry matter and cell wall amount, respectively.

Numbers in *italic* correspond to the standard deviations. The cell wall % DM was measured in single (see “Materials and Methods” section). Number in brackets corresponds to the relative proportion of each cell wall phenolic compound expressed as percent of the sum of the cell wall phenolic compounds.

probably corresponded to residual chlorophyll (Donaldson and Williams, 2018; Donaldson, 2020), which resulted in a seemingly wider rind observed using the BlueBox system. In the rind, the

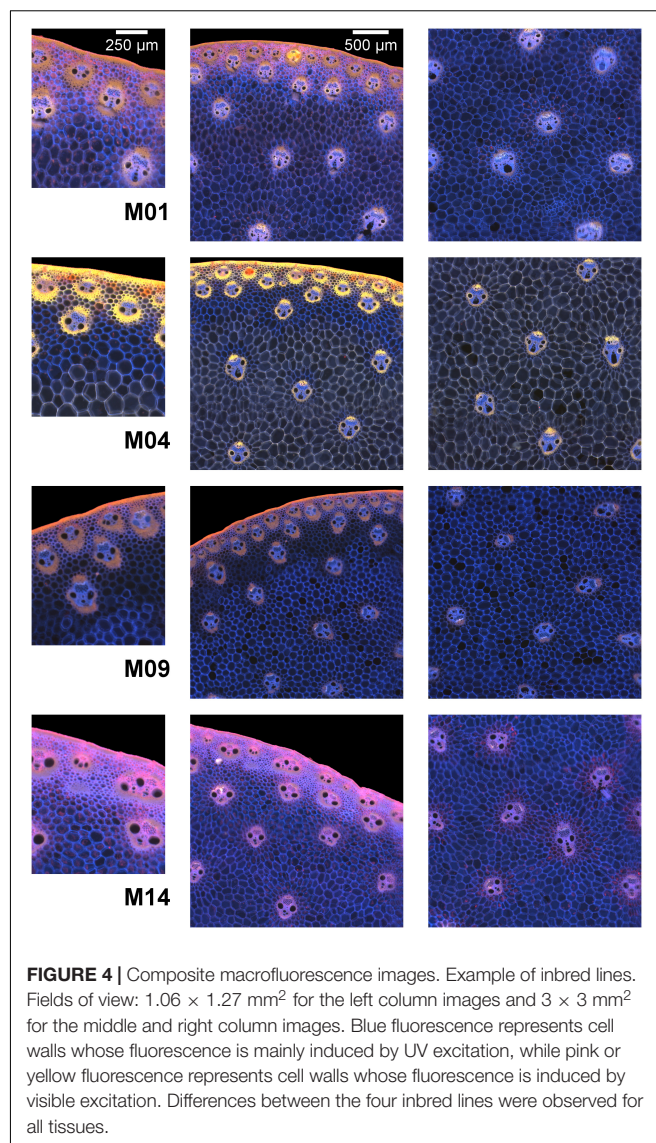
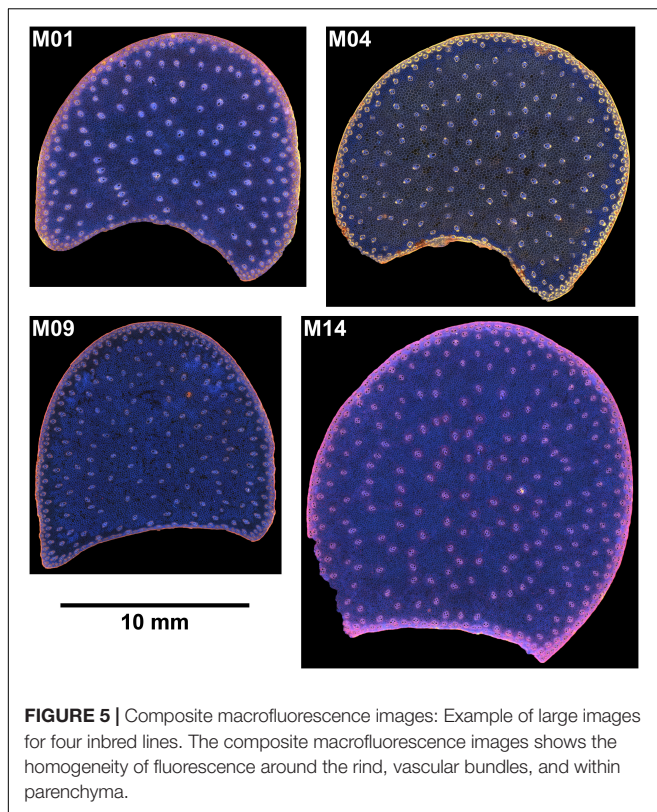


FIGURE 4 | Composite macrofluorescence images. Example of inbred lines. Fields of view: $1.06 \times 1.27 \text{ mm}^2$ for the left column images and $3 \times 3 \text{ mm}^2$ for the middle and right column images. Blue fluorescence represents cell walls whose fluorescence is mainly induced by UV excitation, while pink or yellow fluorescence represents cell walls whose fluorescence is induced by visible excitation. Differences between the four inbred lines were observed for all tissues.

lignified sclerenchyma sheaths of vascular bundles (Lopez and Barclay, 2017; El Hage et al., 2018) were thick and fluoresced considerably, with the colour varying from yellow for M04 and orange for M01 and M09 to pink for M14.

In the pith, the vascular bundle sizes and shapes differed according to the line, with the vascular bundles from M04 and M09 smaller than those from M01 and M14. M01 vascular bundles were round, whereas, vascular bundles of the other three lines were more elongated. The fluorescence colour of vascular bundles was less strong but consistent with that observed in the rind. Inside vascular bundles, blue fluorescence was observed for the phloem and vascular parenchyma.

The parenchyma cells were clearly visible at this scale, and their size was dependent on the line and the region in the section. The smallest cells were observed for M09, and the largest were observed for M04. Cells near the rind seemed smaller than those in the middle parenchyma. Parenchyma cell walls



fluoresced mainly in blue except for Line M04, which mainly showed yellow fluorescence. Specific fluorescence was observed in the parenchyma near the rind for M04 and M09. In the case of M09, the intensity was much lower, and in the case of M04, the fluorescence colour was blue compared to the yellow fluorescence of the middle parenchyma.

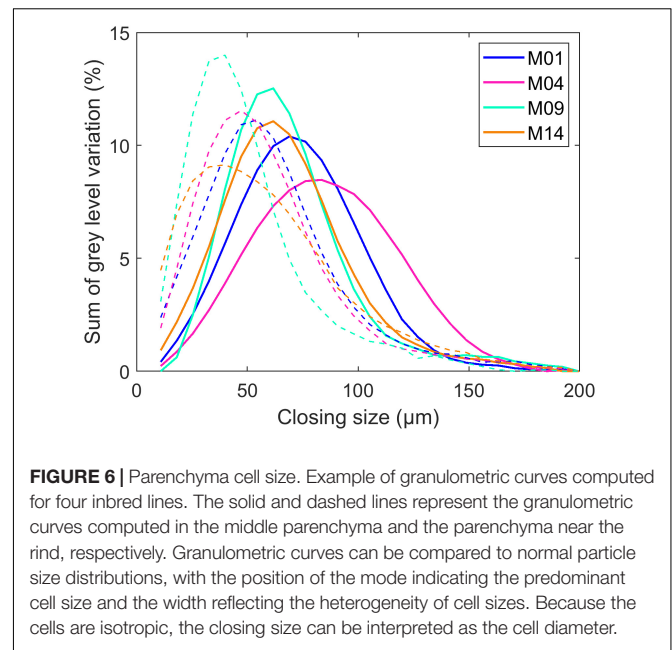
Large Image Scale

Large images were created from individual fields of view of the images (Figures 2, 5). The concave regions in the sections correspond to the location of the main ear. The large image reveals the stem section area, rind, vascular bundle repartition and parenchyma. The section area was the largest for Line M14 and the smallest for M09. The images show that the rind thickness was homogeneous all around the section as well as the vascular bundle size. The composite macrofluorescence images (Figure 5) showed largely homogeneous fluorescence around the rind, vascular bundles and parenchyma. In particular, the specific fluorescence found for the parenchyma near the rind for Lines M04 and M09 could be observed all around the sections. For M01, yellow fluorescence occurred in small places of the rind.

Extraction of Morphological Descriptors for the Four Examples of Inbred Lines

Proportions of Tissue Areas, Vascular Bundle Morphology, and Density

The proportion of tissues extracted for the four example lines are reported in Table 3: area of the stem section in cm²,



parenchyma, rind and vascular bundle areas, which are expressed as a percentage of the area of the stem section, vascular bundle density, mean area and elongation of individual bundles.

The stem area was two times larger for M14 than for M09 and similar for M01 and M04. The rind area was larger for M01 and M14, as expected from the images. Both lines also showed a larger vascular bundle total area together with a large area for individual vascular bundles. The smallest vascular bundles were observed for Line M09 along with the highest density, nevertheless resulting in a small relative total area. The proportion of parenchyma was consequently smaller for M01 and M14 and larger for M04 and M09. The vascular bundle shape did not vary much, as visually observed in the images.

Parenchyma Cell Size

Cell size was evaluated by grey-level granulometry without segmenting individual cells. The method was shown to be relevant to compare tissue sections from the BlueBox darkfield images (Devaux et al., 2008, 2009). Figure 6 shows the average granulometric curve computed for the parenchyma near the rind (dashed lines) and the middle parenchyma (solid lines). Granulometric curves can be compared to normal particle size distributions, with the position of the mode indicating the predominant cell size and the width reflecting the heterogeneity of cell sizes. Because the cells were isotropic, the closing size can be interpreted as the cell diameter.

In the middle parenchyma, the smallest cells were observed for Lines M09 and M14, with a cell diameter of approximately 60 μm, and the largest cells were observed for M04, with a cell diameter of approximately 85 μm, with M01 being intermediate. The distribution was more heterogeneous for M04, for which small cells were clearly distinguished around vascular bundles (Figure 1).

TABLE 3 | Morphological descriptors of the four examples of inbred lines.

Descriptor	M01	M04	M09	M14	Code
Proportion of tissue and vascular bundle morphology and density					
Stem area (cm ²)	1.45 <i>0.06</i>	1.48 <i>0.07</i>	1.21 <i>0.02</i>	2.36 <i>0.03</i>	St(A)
Parenchyma area (% of the stem area)	71.9 <i>1.4</i>	81.7 <i>0.5</i>	80.3 <i>0.1</i>	71.5 <i>1.7</i>	Pa(A)
Rind area (% of the stem area)	21.2 <i>2.0</i>	13.3 <i>0.4</i>	13.0 <i>0.5</i>	20.2 <i>1.0</i>	Ri(A)
Vascular bundle area (% of the stem area)	8.2 <i>0.2</i>	6.2 <i>0.2</i>	6.9 <i>0.2</i>	8.7 <i>0.4</i>	Vb(A)
Vascular bundle Density (number per cm ²)	63.4 <i>2.3</i>	81.7 <i>0.7</i>	116.6 <i>6.9</i>	62.1 <i>1.2</i>	Vb(D)
Vascular bundle Mean Area (mm ²)	0.130 <i>0.007</i>	0.076 <i>0.002</i>	0.059 <i>0.002</i>	0.141 <i>0.009</i>	Vi(A)
Vascular bundle elongation	0.777 <i>0.003</i>	0.802 <i>0.009</i>	0.738 <i>0.026</i>	0.790 <i>0.015</i>	Vi(E)
Parenchyma cell size					
Middle parenchyma Grey level mean size (μm)	70.0 <i>0.2</i>	81.8 <i>0.7</i>	65.5 <i>0.8</i>	64.2 <i>0.6</i>	Pm(Cd)
Middle parenchyma Standard deviation (μm)	27.9 <i>0.6</i>	31.6 <i>0.2</i>	29.1 <i>0.9</i>	29.4 <i>0.6</i>	Pm(Cs)
Parenchyma near the rind Grey level mean size (μm)	57.9 <i>0.9</i>	54.9 <i>0.2</i>	46.0 <i>0.6</i>	56.2 <i>1.7</i>	Pr(Cd)
Parenchyma near the rind Standard deviation (μm)	31.6 <i>1.0</i>	30.6 <i>0.4</i>	27.4 <i>0.5</i>	34.9 <i>1.3</i>	Pr(Cs)
Parenchyma cell wall density					
Middle parenchyma Cell wall density (%)	4.29 <i>0.01</i>	3.67 <i>0.03</i>	4.58 <i>0.06</i>	4.67 <i>0.05</i>	Pm(CD)
Parenchyma near the rind Cell wall density (%)	5.18 <i>0.08</i>	5.46 <i>0.02</i>	6.53 <i>0.09</i>	5.36 <i>0.16</i>	Pr(CD)
Tissue cell wall proportion					
Total cell wall Area [% St(A)]	32.7 <i>2.0</i>	23.1 <i>0.5</i>	24.1 <i>0.3</i>	32.4 <i>1.3</i>	CW(T)
Middle parenchyma cell wall area [% CW(T)]	6.3 <i>0.6</i>	8.8 <i>0.3</i>	10.1 <i>0.1</i>	7.7 <i>0.5</i>	Pm(Cw)
Parenchyma near the rind cell wall area [% CW(T)]	3.9 <i>0.2</i>	6.3 <i>0.1</i>	7.4 <i>0.1</i>	3.1 <i>0.1</i>	Pr(Cw)
Rind cell wall area [% CW(T)]	65.3 <i>2.0</i>	58.6 <i>0.8</i>	56.4 <i>1.2</i>	63.2 <i>0.7</i>	Ri(Cw)
Vascular bundle cell wall area [% CW(T)]	25.3 <i>1.2</i>	27.0 <i>0.7</i>	28.6 <i>1.1</i>	26.8 <i>0.6</i>	Vb(Cw)

Proportion of tissues as a percentage of the stem area, parenchyma cell size and cell wall density, and proportion of tissues as a percentage of the total cell wall. Mean values and standard errors measured for the two sections of the inbred lines. Numbers in italic correspond to the standard deviations.

Cells were found to be smaller in the parenchyma near the rind, which was also measured in Legland et al. (2020), who computed local granulometric curves in a section of maize stem.

M09 and M14 showed the smallest diameters of approximately 40 μm, and M01 showed the largest diameter of approximately 55 μm. M14 differed from the other lines by its greater heterogeneity in cell size. **Figure 1** shows that the cell walls were not always clearly contrasted due to the presence of cell content that may result in measuring small size reflecting the distance between cell wall and cell content together with cell size.

To summarise the granulometric curves, grey-level mean sizes and standard deviations were computed (Devaux and Legland, 2014; Legland et al., 2014; **Table 3**). The grey-level mean sizes were approximately 65–80 μm for the middle parenchyma and 45–60 μm for the parenchyma near the rind. The standard deviations of the granulometric curves were approximately 30 μm and depended on the inbred line, i.e., larger values were observed for M04 in the middle parenchyma and for M14 in the parenchyma near the rind.

Estimating the Proportions of Tissue Cell Walls

From 2D Images to Volumes of Cell Walls: Principles and Hypotheses

Morphological features were extracted with the objective of examining their relationships with data such as chemical composition data or wall digestibility, which are measured on stems. In the present work, we proposed estimating the volume of tissues from 2D images considering several approximations and hypotheses. First, we considered that the internode under the ear was representative of the stem (Méchin et al., 1998). The internode was considered as a cylinder, and the density value of the cell walls was constant regardless of the cell type. This means that the volume and mass of cell walls are proportional. With these assumptions, the relative proportions of tissues in the internode can be estimated from their relative areas in the internode sections.

In the case of rinds and vascular bundles, the area proportions of tissue largely reflect the quantity of cell walls because these are thick and the lumen of the cells is only slightly visible. Therefore, we approximated that the area of these tissues that corresponded mainly to their cell wall proportion. In contrast, the amount of parenchyma cell wall depends on the cell size. This amount was estimated using the parenchyma tissue area and the parenchyma cell wall density that was evaluated from the cell size as described below.

Parenchyma Cell Wall Density

A parenchyma cell wall density estimate was derived from the grey-level mean sizes. Parenchyma cells were modelled as spheres with a radius (R) that corresponded to the grey-level mean size divided by 2. The cell wall density is equal to the ratio between the wall volume and cell volume, and the wall volume is equal to the cell surface multiplied by the wall thickness. In this case, the cell wall density Cw(D) expressed in percentage of volume is equal to the following:

$$Cw(D) = \frac{3 \times CwThickness}{R} * 100 \quad (5)$$

In the present work, the cell wall thickness CwThickness was set to 0.5 μm (Jung and Engels, 2001). Consistent with the cell

size, the cell wall density was greater in the parenchyma near the rind (5–6%) than in the middle parenchyma (approximately 4%) (Table 3).

The total amount of cell wall in the parenchyma was assessed as the density of the cell wall multiplied by the parenchyma area, i.e., parenchyma cell wall areas, which are considered representative of their volume in the case of a cylindrical internode (Table 3). The middle parenchyma area was taken to estimate the cell wall amount. The area for the parenchyma near the rind was computed as the total parenchyma area minus the middle parenchyma area. It therefore also included the region between 500 and 1,000 μm .

Estimating Tissue Cell Wall Proportions

The total cell wall amounts were computed as the sum of the rind and vascular bundle areas plus the parenchyma cell wall amounts. Finally, the proportion of tissues was computed as the relative cell wall amounts. The values are reported in Table 3. Because of the different approximations, the rind and vascular bundle cell wall amounts could be somewhat overestimated. Nevertheless, these values were considered relevant to compare the lines. Table 3 shows that the rind was the major tissue, followed by the bundles. Depending on the cell size, stem diameter and parenchyma proportion, the contribution of parenchyma near the rind and middle parenchyma varied for the four inbred lines: the smallest contribution of the parenchyma near the rind was observed for M14 and the largest was observed for M09.

Tissue Pseudospectra of the Four Examples of Inbred Lines

Normalisation of Pseudo Spectra

Tissue-specific fluorescence pseudospectra were studied for the rinds, vascular bundles, and parenchyma near the rind and in the middle of the section. In the parenchyma, the pseudospectra depended on the fluorescence properties of the cell walls but also on the density of the cell wall. Parenchyma pseudospectra were therefore divided by the cell wall density. In this way, we expected to estimate the fluorescence that would have been measured on the walls alone, thereby avoiding the cell lumens. In addition, the overall fluorescence intensity was found to be section-dependent regardless of the tissue, which was attributed to uncontrolled thickness variations. A section normalisation factor was assessed as follows. For each section and for each tissue pseudospectra (rind, vascular bundles, parenchyma near the rind and middle parenchyma after correction for the cell wall density), the mean fluorescence intensity measured for the 11 channels was computed: $\overline{\text{Ri}}(\text{F})$, $\overline{\text{Vb}}(\text{F})$, $\overline{\text{Pr}}(\text{F})$, and $\overline{\text{Rm}}(\text{F})$. The normalisation factor of the section was computed as the mean fluorescence intensity:

$$Fn(\text{section}) = \frac{(\overline{\text{Ri}}(\text{F}) + \overline{\text{Vb}}(\text{F}) + \overline{\text{Pr}}(\text{F}) + \overline{\text{Rm}}(\text{F}))}{4} \quad (6)$$

Each tissue pseudospectrum was divided by this normalisation factor.

Spectral Information in the Pseudospectra

The resulting pseudospectra are shown in Figure 7 for the four example lines. In the pseudospectra, each value corresponds to the average fluorescence intensity of one of the 11 channels of the multispectral images. The colour images in Figures 4, 5 represent a summary of the 11 spectral fluorescence channels, while the pseudospectra represent the average spectral fluorescence behaviour computed over all pixels of the considered region of interest. The first five pseudospectra values report the intensity of UV-induced fluorescence, and the six others report the intensity of visible-induced fluorescence. As mentioned in the “Materials and Methods” section, no signal was observed in the channels BLb, GRb, and GRg, which were kept at baseline. Because the pseudospectra were normalised, only relative intensity variations can be discussed.

In the plant cell walls, not all the constituents are fluorescent for this range of excitation wavelengths. Polysaccharides are not fluorescent while lignin and hydroxycinnamic acids are the major natural fluorophores. To compare the normalised fluorescence intensities with the number of phenolic compounds (lignin + hydroxycinnamic acids), their relative amounts were calculated (Table 2). With this normalisation, Line M04 contained less para-coumaric acid than the three other lines, with a value of 4.87% compared to more than 7.00%, but more lignin, with a value of 92.51% compared to less than 90.3%. M14 was characterised by a high relative amount of para-coumaric acid (9.30%).

Hydroxycinnamic acids emit blue fluorescence with UV excitation at neutral pH, and lignin has a wide excitation range. Excitation with UV and blue light results in blue and green emission of lignins (Donaldson and Williams, 2018; Donaldson, 2020). Thus, a greater amount of visible fluorescence was assumed to correspond to samples that contained more lignin. Similarly, a greater amount of UV fluorescence was assumed to correspond to more hydroxycinnamic acids.

In addition, localisation or the lack of localisation of hydroxycinnamic acids and lignin could be responsible for the specific colour observed within a given line. Thus, the yellow colour of Line M04 observed in Figure 4 could be ascribed to a high relative amount of lignin together with a low relative amount of hydroxycinnamic acid. For M14, pink fluorescence could be ascribed to a low relative amount of lignin together with a high hydroxycinnamic acid content. In the following, tissue pseudospectra were examined to identify tissues that presented differences in phenolic compounds and differences between lines.

Tissue Pseudospectra of the Four Example Lines

For the parenchyma regions (Figures 7A,B), the UV-induced fluorescence was always stronger than the visible-induced fluorescence, showing that these tissues contained relatively less lignin and more hydroxycinnamic acids, which resulted in the generally blue-coloured parenchyma (Figures 4, 5). In contrast, rind and vascular bundles showed visible-induced fluorescence similar to UV-induced fluorescence (Figures 7C,D), which was consistent with the lignification of these tissues (Akin, 1989; Wilson et al., 1993; Hatfield et al., 1999; Zhang et al., 2013).

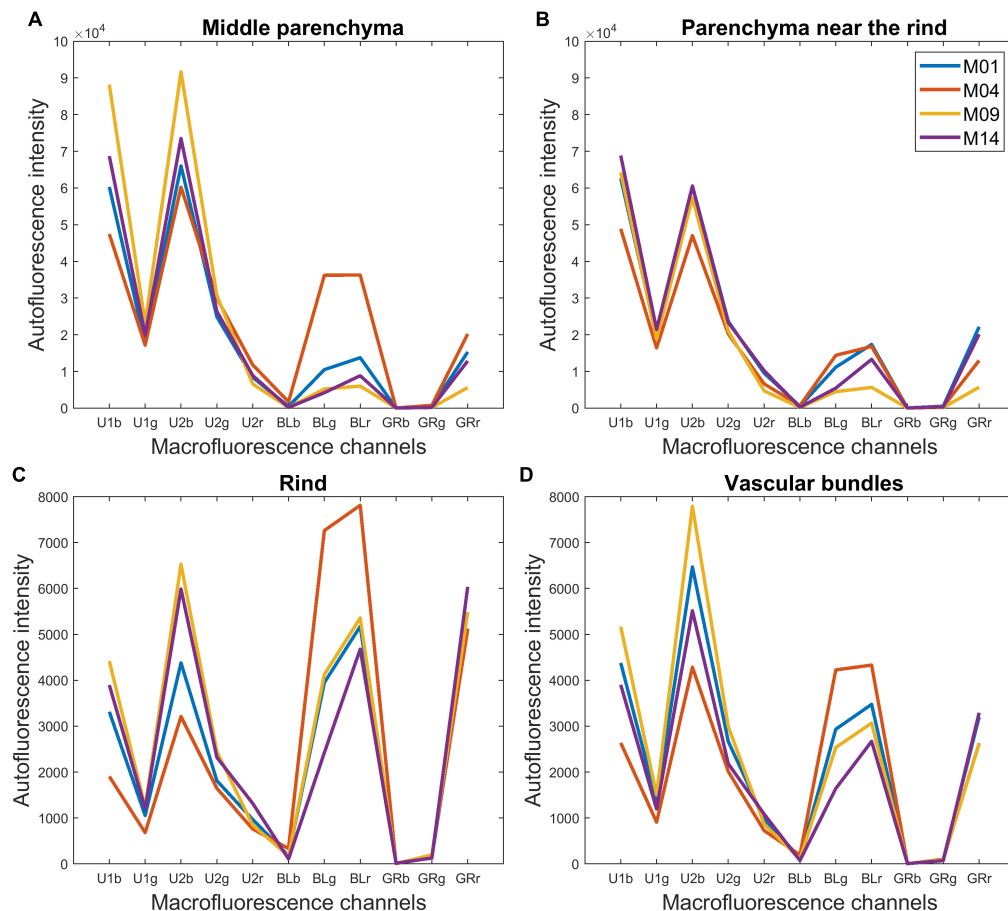


FIGURE 7 | Macrofluorescence analysis. Tissue-normalised pseudospectra of the four examples of inbred lines: **(A)** middle parenchyma, **(B)** parenchyma near the rind, **(C)** rind and **(D)** vascular bundles. The pseudospectra represent the average spectral fluorescence behaviour computed over all pixels of the considered region of interest. The first five values report the intensity of UV-induced fluorescence, and the six others report the intensity of visible-induced fluorescence. The pseudospectra were normalised and only relative intensity variations can be discussed.

After normalisation, the fluorescence of the parenchyma was approximately 10 times more intense than that of the rind and vascular bundles. This ratio is somewhat overestimated because cell size was considered for the parenchyma cells and not for rind and vascular bundles. Nevertheless, the result is consistent with the fact that the parenchyma cell walls were clearly visible despite the wall thickness between two cells ($1\ \mu\text{m}$) being much smaller than the pixel size ($2.78\ \mu\text{m}$). Willemse and Emons (1991) also reported lower UV autofluorescence for sclerenchyma walls than for parenchyma walls, which was even more pronounced when related to the cell wall area. Another explanation for the relatively lower fluorescence intensity of lignified tissues is that lignin fluorescence is a complex process involving different fluorophores with different fluorescence profiles and energy transfer processes. Lignin fluorescence can be quenched by interactions with other polymers inside the cell walls, especially UV-induced fluorescence (Donaldson, 2020).

Comparing the rind and vascular bundles, visible-induced fluorescence was higher in the rind regions for the four lines. The composite macrofluorescence images in **Figure 4** show

that visible-induced fluorescence was mainly observed in the sclerenchyma sheath of vascular bundles and that the sheath was much thicker in the rind than in the pith. In addition, the relative proportion of blue parenchyma was higher in the bundle than in the rind.

In the case of the M09 parenchyma, pseudospectra allow the quantification of the lower intensity of the parenchyma near the rind compared to the middle parenchyma. In the case of M04, visible-induced fluorescence was found to be much lower in the parenchyma near the rind than in the middle parenchyma. This finding corresponds to the blue and yellow-white fluorescences observed in **Figure 4** for the parenchyma near the rind and the middle parenchyma, respectively. For the other two lines, the pseudospectra of the two parenchyma were largely similar.

Looking more specifically at the lines, the intensity of the visible-induced fluorescence was much higher for M04 than for the other three lines in the rind, bundles and middle parenchyma. This finding is consistent with the high relative amount of lignin. It also suggests that a significant amount of lignin was found in the parenchyma cell walls for this line. The occurrence of an

equal intensity after blue excitation in the green BLg and red BLr channels led to the strong yellow fluorescence of the rind and vascular bundles and to the yellow–white fluorescence of the cell walls in the middle parenchyma (**Figure 4**).

The pink fluorescence of the rind and vascular bundles observed for M14 was due to a lower relative green fluorescence after blue excitation (Blg channel), which corresponded to the lower relative lignin content. This was also measured for parenchyma cell walls. More generally, for a given line, the relative proportions of fluorescence measured after blue excitation in channels green BLg and red BLr were similar for all tissues, which suggests that the signature of blue-induced lignin fluorescence would not be tissue-dependent but line-dependent.

The highest UV-induced fluorescence intensity was observed for M09 and M14 in the rind and the two parenchymas. The two lines contained the most hydroxycinnamic acids. In the case of M09, almost no visible-induced fluorescence was observed in the

parenchyma, suggesting that lignin was only found in the rind and vascular bundles.

In conclusion, the normalised pseudospectra were considered relevant to quantify the differences in the tissue fluorescence observed in the multispectral images.

Histological Variability Within the 14 Inbred Line Collections

Morphological Analysis

Descriptors Extracted for the 14 Inbred Lines

Examples of images acquired for each of the 14 inbred lines can be found in the **Supplementary Figure 1**. All descriptors were computed for the two stems of the 14 inbred lines. Average values are reported in the **Supplementary Table 3**. The main points are reported here. The area of the stem section ranged from 1.21 to 2.36 cm², with an average of 1.78 cm². The parenchyma

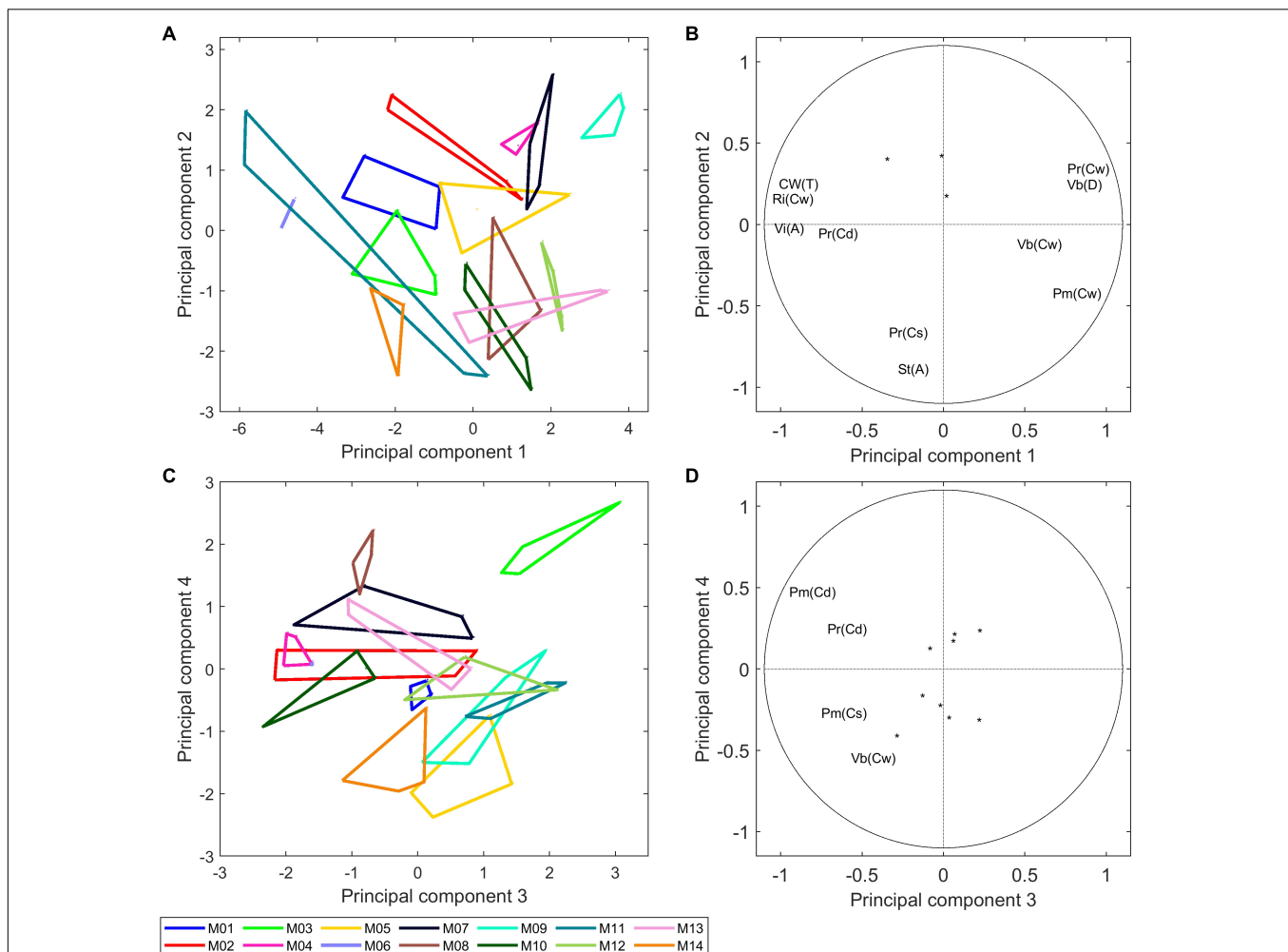


FIGURE 8 | Morphological analysis of 14 inbred lines: principal component analysis. **(A,B)** Similarity map and loadings of components 1 and 2 (44 and 15% of the total variance). **(C,D)** Similarity map and loadings of components 3 and 4 (14 and 10% of the total variance). Convex hulls were drawn for each inbred line. Loadings are shown as correlation circles. Considering that the variables in the middle of the correlation circles are not representative of principal components, they were represented as points. Only variables with correlation over 0.5 with the principal component are shown. The similarity maps reveal a considerable variability between inbred lines based on their morphological descriptors.

covered on average 76.1% of the total area of stem sections, and the coefficient of variation was 8% for the set of 14 inbred lines. Larger variations between inbred lines were observed for the rind and vascular bundle relative areas, with average values of 17.0 and 7.0%, respectively, and coefficients of variation of 27 and 20%, respectively. The vascular bundle density varied from 117 to 50 per cm² with an average value of 77. Legland et al. (2017) also studied maize internodes under the ears of four inbred lines grown under two irrigation conditions and found values ranging from 1.0 and 3.0 cm² for the area of the sections. The rind accounted for 10.3 to 16.8% of the section area, the vascular bundles accounted for 3.1 to 7.3% and the parenchyma accounted for 80–86%. The bundle density ranged between 111 and 66 per cm². Vo et al. (2020) compared internodes under the ear of six maize inbred lines and reported values ranging from 1.6 and 4.0 cm² for the area of the sections and 11–19 and 77–86% of the section area of the rind and pith parenchyma, respectively. The bundle density ranged between 76 and 42 per cm². In the internode sections of sorghum, a species that is very close to maize, Wilson et al. (1993) found that the rind accounted for 16.2% of the total section area, the parenchyma accounted for 79.2% and the vascular bundles in the pith parenchyma accounted for 4.7%. The values found in our work are on the same order of magnitude of those reported in these manuscripts.

In the present work, we estimated the contribution of the different tissues to the total wall content of the internodes on the basis of the tissue surface proportion in the sections and from a simple internode model. The rind and vascular bundles represented 61 and 25% of the total cell wall on average, respectively, with coefficients of variation of 8 and 12%, respectively. The middle parenchyma and parenchyma near the rind represented 9 and 5%, respectively, of the total cell wall, with high coefficients of variation of 25 and 30%, respectively. The values reported here for the relative contribution of tissues to the total cell walls were in the range of those reported by Wilson et al. (1993) for sorghum internodes. In this study, the tissues of one cultivar were manually separated and analysed individually. The rind and vascular bundles accounted for 68.7 and 11.4% of the total cell wall, respectively. The pith parenchyma accounted for 22% of the total cell walls.

Considering all descriptors, the coefficient of variation ranged between 6% (standard deviation of cell diameters) and 36% (average area of individual vascular bundles). An ANOVA test was run individually on the descriptors to test their ability to discriminate lines. All descriptors were found to be significant for the line effect, with *p* values lower than 0.01.

Principal Component Analysis

A principal component analysis was performed on the subset of 13 morphological descriptors of the 14 inbred lines, including the stem area, proportion of cell walls in the stem, relative proportions of cell wall ascribed to tissues, parenchyma mean cell diameters and standard deviations, and vascular bundle density and morphology. A variance analysis was applied to the principal components. The four first principal components accounted for 44, 15, 14, and 10% of the total variance, and the line effect of these components was highly significant.

Figure 8A shows the similarity map of components 1 and 2 and **Figure 8C** shows the similarity map of components 3 and 4 according to inbred lines. The corresponding loadings (**Figures 8B,D**) show the importance of the individual variables for the specified components.

Component 1 differentiates Lines M06, M11, M14, M01, and M03 based on the relatively high rind proportions [Ri(Cw)], high total cell wall amounts [Cw(T)] and large vascular bundle individual areas [Vi(A)], and it differentiates Lines M09, M07, M13, M12, and M04 based on the high vascular bundle densities [Vb(D)] and parenchyma cell wall amounts [Pm(Cw) and Pr(Cw)].

Component 2 differentiates lines based on their stem section area [St(A)], and it differentiated M09 and M07, which had a small section area, from M14 and M10, which had a larger section area. The two stems of M11 had very different stem diameters, with actual values of 2.6 and 1.5 cm². For all other lines, the two stems were largely similar, as revealed by the convex hulls. Component 2 mainly described the stem area variations, with Lines M14, M10, M13, and M08 showing larger stem diameters than Lines M09, M07, M02, and M04.

Figure 8B highlights the correlation between the morphological descriptors. Namely, the expected strong contribution of the rind to the total cell wall amount in the stem as well as the negative correlation with the parenchyma cell wall amounts. A negative correlation $r = -0.77$ was observed between the vascular bundle density and the average individual area of vascular bundles. Indeed, a general trend was observed among the 14 inbred lines, with M06, M11, M14, and M01 having large bundles over 0.1 mm² and less than 60 bundles per cm² and M09, M07, M12, and M05 having small bundles smaller than 0.1 mm² and more than 80 bundles per cm². Zhang et al. (2020) measured the area of individual vascular bundles and their density in the stem for 480 inbred lines and reported a negative correlation between these two descriptors.

Beyond examining the components individually, it is interesting to note the distribution of the 14 inbred lines that

TABLE 4 | Macrofluorescence analysis of 14 inbred lines.

Rind and vascular bundles				
	Component 1 72%	Component 2 16%	Component 3 11%	Component 4 1%
Line	12***	5***	19***	8***
Tissue	270***	–	17***	–
Interaction	–	–	–	–
Parenchyma tissues				
	Component 1 71%	Component 2 18%	Component 3 7%	Component 4 3%
Line	11***	9***	9***	15***
Tissue	70***	94***	22***	6*
Interaction	3**	9***	5***	4***

Variance analysis. Effects of inbred lines and tissues on the principal components. *F* value and significance. *, **, *** means that the probability was below 5%, 1% and 0.1%, respectively.

reveals their specificity and the great variability of the collection. The same comment can be applied to the similarity maps of components 3 and 4. In this case, the components described variations in the parenchyma cell diameters Pm(Cd), Pm(Cs), Pr(Cd), and the proportion of vascular bundles Vb(CW). Line M03 was clearly highlighted mainly because of the much lower relative number of vascular bundles due to its low density in number. On this similarity map, the other lines contrasted were M08, M04, M02, M05, M14, M11, and M12.

We investigated the correlation between the morphological descriptors and the relative amounts of chemical compounds to further explore their tissue origin, but no correlation was found.

Autofluorescence Variations According to Tissue and Lines

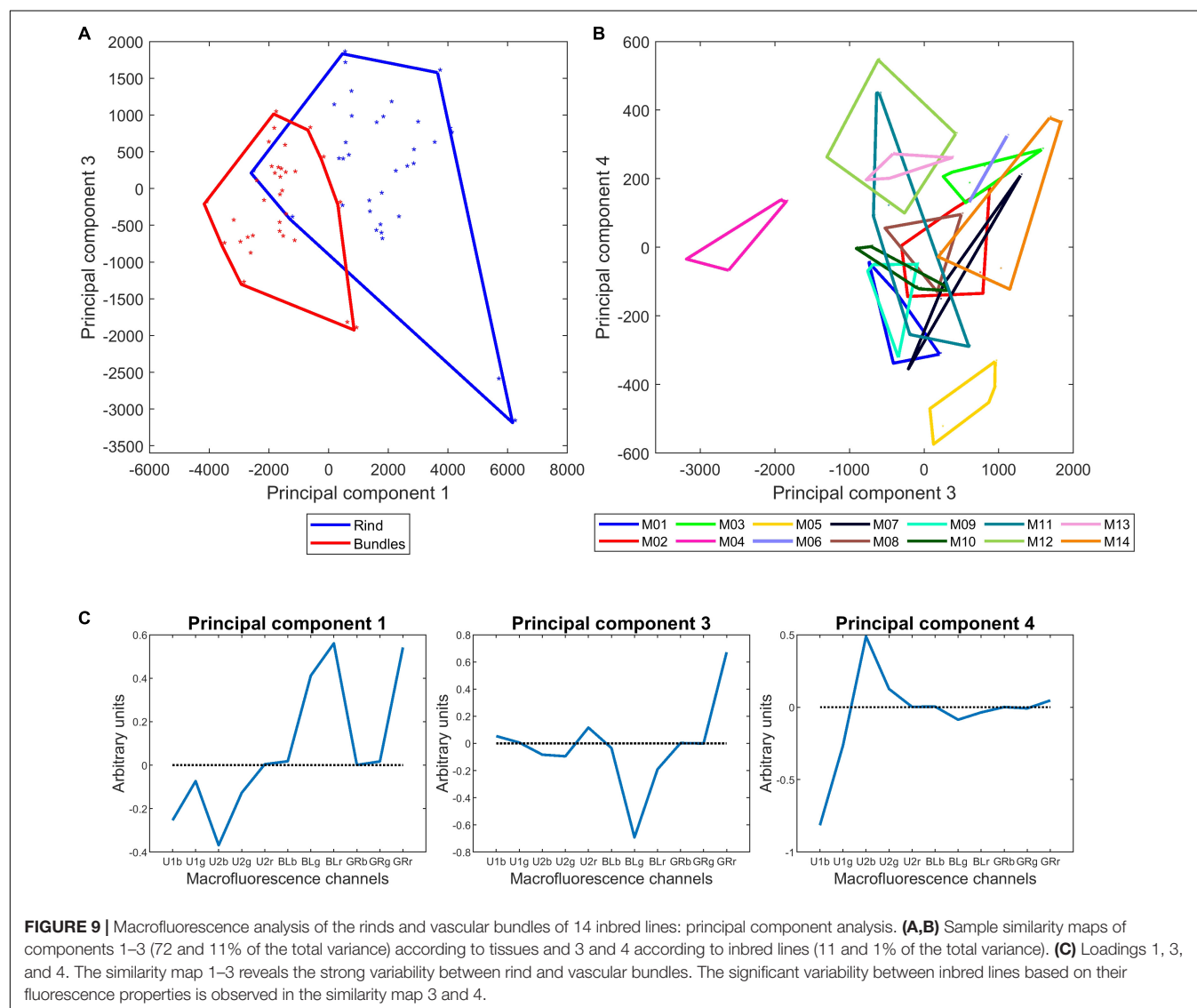
The fluorescence colour quantified in the pseudospectra should reveal more lignin or hydroxycinnamic acids and their localisation in some specific tissues. To compare the 14 inbred

lines of the study, multivariate analyses were performed on the tissue pseudospectra. Principal component analyses were performed to assess the relative importance of the tissue or line in determining the fluorescence properties. In a second step, the correlation between the relative amounts of phenolic compounds and the tissue pseudospectra was examined.

Principal Component Analysis of Tissue Pseudospectra

Because of the general intensity differences, principal component analyses were carried out separately on the rind and vascular bundle pseudospectra on the one hand and on the parenchyma pseudospectra on the other hand. A variance analysis was applied to the principal components to evaluate the effects of lines and tissues and their interaction.

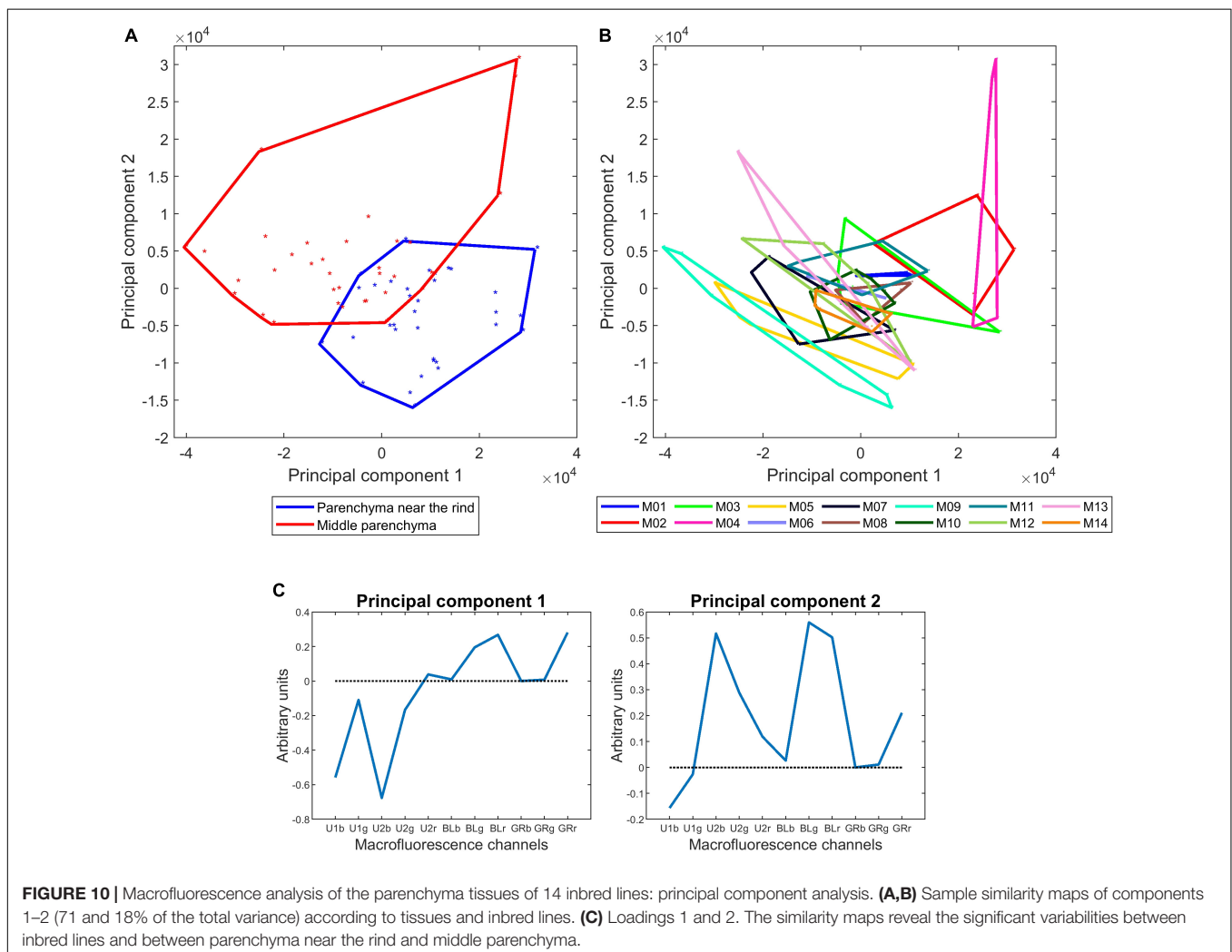
Table 4 reports the results of the variance analysis applied on the four first principal components computed for the rind and vascular bundles accounting for 72, 16, 11, and 1% of the total variance. For the four components, the line effect was significant.



Rind and vascular bundles differed on components 1 and 3, and no interaction was revealed. **Figure 9A** shows the similarity map of components 1 and 3 according to the tissues. **Figure 9B** shows the similarity map of components 3 and 4 according to the lines. **Figure 9C** shows and the loadings of components 1, 3, and 4. Loading 1 revealed the relative variations between UV- and visible-induced fluorescence. Loading 3 was based on the relative variations observed in visible-induced fluorescence, e.g., green emission after blue excitation (BLg channel) versus red emission after green excitation (GRr channel). Loading 4 showed a difference in the relative blue emission using UV excitation of U1 and U2 (U1b and U2b channels). The similarity map of components 1 and 3 shows that for all lines, the rind and vascular bundles differed mainly by their relative visible and UV-induced fluorescence, and to a lesser extent by a relatively higher red fluorescence emission of the rind after green excitation. This difference could be ascribed either to cortical parenchyma cell walls or to their content. On this map, the line effect was mainly caused by M04, which corresponded to the extreme points for the two tissue scatterplots. The similarity map of components 3

and 4 reveals the line effect among the 14 lines. Despite some overlap, contrasting fluorescence fingerprints were observed for some lines, such as M04, M05, M14, and M12. M14, M03, and M06 showed relatively higher red fluorescence after green excitation (GRr channel) (see also **Figures 7C,D** for M14). Line M05 was characterised by its relatively high blue emission after U1 excitation compared to M12 or M13.

Table 4 and **Figure 10** show the results obtained for the two parenchyma tissues. The first four components accounting for 71, 18, 7, and 3% of the total variance were found to be significant for both the lines, tissue and their interaction. Similarity maps and loadings are shown for components 1 and 2 (**Figures 10A,B**). The same map was drawn twice by considering the tissues or the lines. Figures for components 3 and 4 are given in **Supplementary Figure 2**. Loading 1 (**Figure 10C**) was partly similar to the one obtained for the rind and vascular bundles, thus showing the relative response after UV and visible excitation. Component 1 mainly described the differences between lines, with M04, M02, and M03 showing stronger visible-induced fluorescence and M09 and M05 showing stronger UV-induced fluorescence,



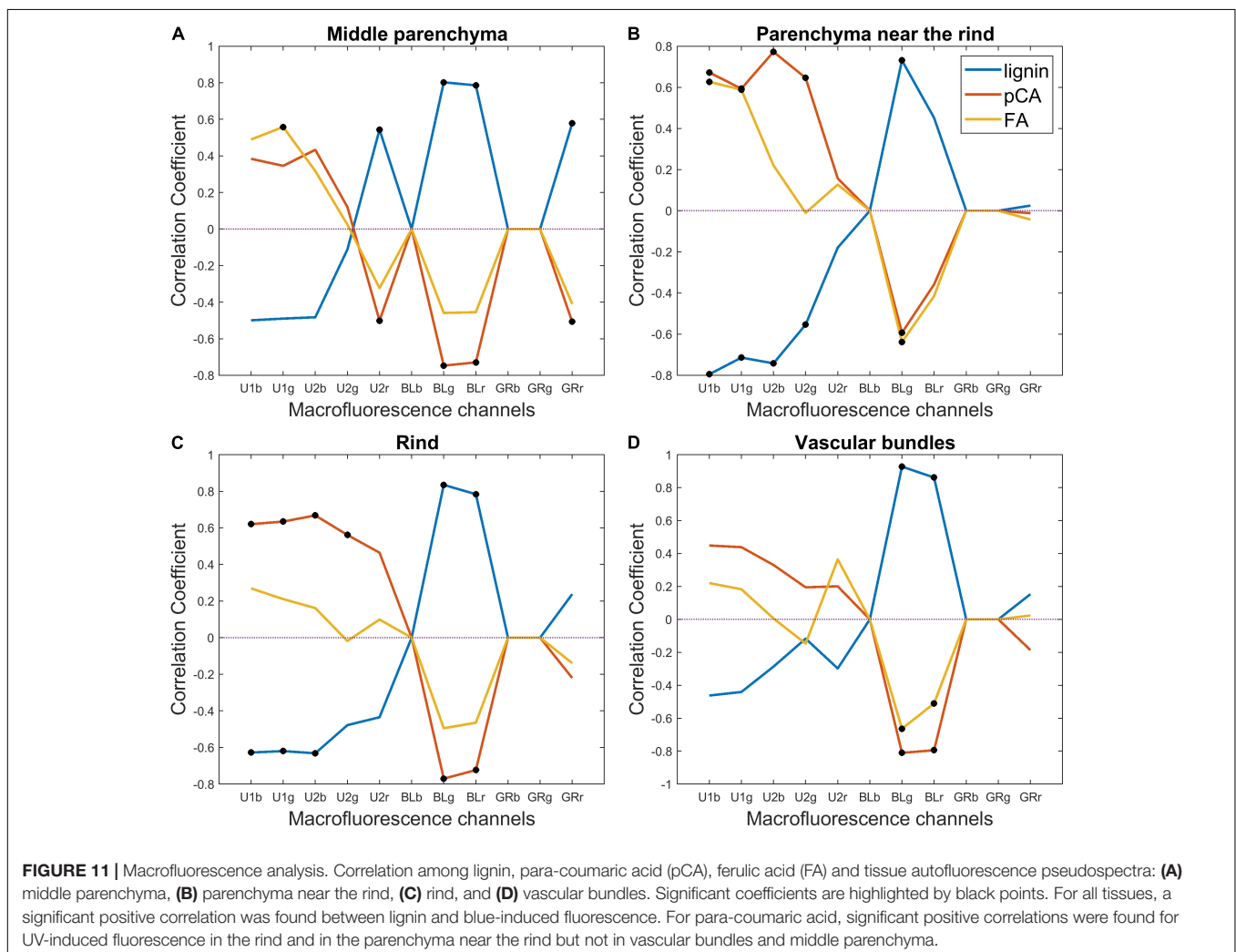
especially in the middle parenchyma. Loading 2 (**Figure 10C**) described the relative intensity of the two parenchyma tissues, thus attesting to the generally higher fluorescence in the middle parenchyma except for the U1-induced fluorescence, which was slightly higher in the parenchyma near the rind. The interaction was also significant, highlighting that the differences near the rind and middle parenchyma were enhanced for some lines, such as M09 and M05. As expected, M04 was found to be different from the other lines. Other lines, such as M05, were found to be characteristic. For this line, the visible fluorescence in the middle parenchyma was very low, which could be related to the low lignin content of this line. In addition, the intensity difference between the middle and near the rind parenchyma was important, the cell content was found near the rind, and relatively high U1-induced fluorescence compared to U2-induced fluorescence could be observed, especially near the rind.

In summary, this analysis revealed that fluorescence properties were primarily tissue-dependent but also clearly line-dependent. In particular, the relative fluorescence emission after blue excitation seemed to be similar within an inbred line regardless of the tissue. These differences should be related to variations

in the lignin content, composition and structure or in the phenolic acid contents.

Exploring the Correlation Between Biochemical Data and Tissue Fluorescence

Because the fluorescence properties of compounds depend on several factors, the interpretation of fluorescence variations is not straightforward. We investigated the correlation between tissue fluorescence pseudospectra and the relative amounts of fluorescent chemical compounds to further explore their tissue origin. The tissue fluorescence pseudospectra were normalised according to the individual sections, and this normalisation still allowed for the preservation of the relative intensity variations among the rind, vascular bundles and parenchyma. In parallel, the relative lignin, ferulic and para-coumaric acid amounts were considered after normalisation to their total amount (**Table 2**). A drawback is that this normalisation generates correlations between the variables. Thus, the correlation coefficient values were -0.94 between lignin and para-coumaric acid amounts, -0.54 between lignin and ferulic acid amounts, and 0.23 between para-coumaric and ferulic acid amounts.



Correlation coefficients were drawn according to the channels of the tissue pseudospectra (**Figure 11**). Significant coefficients are highlighted by black points. The correlation was always reversed for lignin and hydroxycinnamic acids, which is expected from the correlation induced by normalisation. For all tissues, a significant positive correlation was found between lignin and blue-induced fluorescence. This finding suggests that lignin is observed regardless of the tissue and that a high relative amount measured for the whole stem occurs in all tissues. In the case of the middle parenchyma, the lignin correlation is also related to red emission after U2 and green excitation. For para-coumaric acid, significant positive correlations were found for UV-induced fluorescence in the rind and in the parenchyma near the rind but not in vascular bundles and middle parenchyma. A significant positive correlation was found for ferulic acid in the parenchyma near the rind and to a lesser extent in the middle parenchyma. These results indicate that para-coumaric acids are mainly localised in the rind and parenchyma near the rind, and localisations of ferulic acids are mainly located in the parenchyma near the rind and in the middle parenchyma. In addition, the ferulic acid fluorescence signal was revealed mainly after U1 excitation, i.e., for shorter excitation wavelengths, while para-coumaric acid was revealed using both U1 and U2 filters.

DISCUSSION AND CONCLUSION

The objective of the present work was the histological quantification of the morphology and fluorescence signature of maize forage stem sections for a set of 14 inbred lines used as parents for maize hybrid production. In addition to the relationship with end-use properties such as digestibility, this work aimed to explore the methodological potential of two techniques, namely, darkfield and fluorescence imaging, to study maize stem collection. Macrovision was retained to acquire images of whole stem sections with a fair pixel resolution, thus allowing for the quantification of cell size together with tissue proportions. The INRAE BlueBox prototype is dedicated to morphological plant tissue imaging with good contrast without any labelling because of darkfield illumination. This kind of illumination, which is less sensitive to variations in the density of walls than the usual transmitted light, is more suited to cell size analysis directly from grey-level images without any cell segmentation. Automated fluorescence macroscopes for multispectral image acquisition are the tools of choice for studying the autofluorescence properties of phenolic compounds in plant tissues. Both techniques should be considered medium-throughput methods that allow for the acquisition of 1 cm² large images in approximately 5 and 20 mins per image for the BlueBox and the fluorescence microscope, respectively. In the present work, sections were observed after being stored in ethanol, i.e., ethanol soluble material was removed, which led to the emptying of the cells except in the rind or in the parenchyma near the rind, where residual cell contents remained for some inbred lines. The fluorescence properties of this material suggest that it could be residual chlorophyll.

Image analysis was implemented to extract morphological features and autofluorescence pseudospectra. Four tissues were studied: rind, vascular bundles, parenchyma near the rind and middle parenchyma. In the present work, serial sections were used for the two devices, and images were segmented separately. One improvement would be to acquire images exactly for the same section and develop a segmentation workflow that takes into account both fluorescence and darkfield properties. For both images, the segmentation workflow was based on the same image processing steps, i.e., grey-level and size thresholds and alternating filtering, and it was fully automated. The only user intervention was to validate and adapt if necessary the automatic or preset thresholds. The rind, vascular bundles and parenchyma regions of interest were extracted. The rind segmented from the darkfield images corresponded to a material-dense region, with small cells having thick walls and bundles with thick sclerenchyma sheaths. The segmentation was also partially dependent on the occurrence of cell contents in and near the rind. The parenchyma near the rind was not segmented in the same manner. The 500 μ m below the rind was examined to demonstrate the presence or absence of differences with the middle parenchyma. For this parenchyma, the cell size was smaller, and the fluorescence differences with the middle parenchyma were dependent on the line.

In this manuscript, 3D descriptors were estimated from 2D images given some assumptions. Indeed, the morphological descriptors were computed based on an estimation that considered the cell wall density of parenchyma tissues and the stem as a cylinder and the same histology of cross-sections all along the cylinder. The total derived cell wall amount can be interpreted as the relative areas of rind, bundles and parenchyma cell walls in the stem section. The fluorescence pseudospectra were also normalised by considering the parenchyma cell wall density.

We found a strong inbred line effect on the morphological descriptors. The most discriminant features were (1) the relative amount of rind and parenchyma tissues together with the density and size of individual bundles, (2) the stem area, and (3) the middle parenchyma cell diameter and distribution of the total vascular bundle amount. No correlation was observed between cell size and stem section, indicating that the diameter of the stem would rather depend on the number of cells in the parenchyma. A significant inverse correlation was observed between the vascular bundle size and density. Heckwolf et al. (2015) and Zhang et al. (2021) also found variations in stem diameter as well as in the area of the rind and pith of the inbred lines they analysed. Zhang et al. (2021) further analysed the variation in vascular bundle traits and reported wide phenotypic variations in vascular bundle size, number, and distribution density. Thirty of the phenotypic traits related to bundles showed high heritability, suggesting that the observed variations were at least partly of genetic origin. At the scale of whole cross-sections, they observed a negative correlation between the vascular bundle area and density, which was also observed in this study. However, neither of these studies considered the size of the cells in the parenchyma.

We did not find any correlation between the morphological descriptors and the phenolic composition of the 14 inbred

lines, which could be explained by several hypotheses. First, biochemical measures were obtained from the whole stem, including the node, and we analysed sections taken in the middle of the internode. Second, none of the three phenolic constituents, i.e., lignin, para-coumaric or ferulic acids, can be considered tissue-specific biochemical markers. This conclusion is consistent with the results of the fluorescence pseudospectral analysis, at least for lignin. Indeed, the correlation profiles of the relative amount of lignin with the tissue pseudospectra clearly showed that a higher level of lignin resulted in a higher visible-induced fluorescence in all tissues.

Specific fluorescence signatures with a predominant tissue effect have been identified, and the inbred line effect was also always found to be significant. The rind, as the most lignified tissue, showed strong visible-induced fluorescence. Our results suggest that the colour of the visible-induced fluorescence, which was line-dependent, may depend on the amount of colocalised lignin and para-coumaric acid.

The relative amount of para-coumaric acid was found to be significantly correlated with the UV-induced fluorescence intensity in the rind and in the parenchyma near the rind, while ferulic acid was significantly correlated mainly with the parenchyma near the rind. In grasses, para-coumaric acid is ester linked to lignin and, to a lesser extent, to hemicelluloses (Hatfield et al., 2017). Since the rind is highly lignified, the presence of para-coumaric acid was expected. More surprisingly, para-coumaric acid was present in the parenchyma near the rind.

The parenchyma near the rind was less fluorescent on average than the middle parenchyma, although the extent of the difference was dependent on the inbred lines. Farga staining is performed to reveal tissue lignification and often reveals this parenchyma region (El Hage et al., 2018). In this manuscript, the parenchyma Farga that was coloured in red was correlated to the lignin amount, cell wall digestibility, and para-coumaric acid content (to a lower extent). They did not find any correlation between the lignin content and the red intensity in the rind or the number or density of bundles in the stem.

To further interpret the differences in autofluorescence between the tissues, it would be useful to have additional information about the biochemical composition of the different tissues. For example, the amount of etherified ferulic acid was not determined. Farga or other lignin selective staining, such as Wiesner or Maïle staining, could be advantageously used to confirm the localisation of lignified tissues and reveal chemical differences in the lignin type (Méchin et al., 2005). Immunolabelling using antibodies would allow further identification and localisation of hydroxycinnamic acids (Philippe et al., 2007; Tranquet et al., 2009). Microspectroscopic

techniques, such as Raman or infrared imaging (Gierlinger, 2018), would allow further localisation of phenolic compounds together with cell wall polysaccharides.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because research partly funded by private company. Requests to access the datasets should be directed to M-FD, marie-francoise.devaux@inrae.fr.

AUTHOR CONTRIBUTIONS

MB, M-FD, BD, and FG planned and designed the research. MB performed all sample preparation and image acquisition. M-FD and MB performed the image analyses and chemometric analyses. M-FD wrote the computer code. DL helped for image segmentation. M-FD, MB, and FG interpreted the results and wrote the manuscript. CB, DL, and BD did a thorough review of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by INRAE (National Research Institute for Agriculture, Food and Environment). Fundings of MB was provided by Limagrain Europe and by the Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation of the French Government (CIFE 2018/1480).

ACKNOWLEDGMENTS

Maxime Martin, Simon Genin, and all persons of the Limagrain Arras Station cultivated and harvested the plants. Laurent Hélay and Camille Alvarado helped to harvest and prepare the samples. Anne-Laure Reguerre helped with the realization of the image plates.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.792981/full#supplementary-material>

REFERENCES

- Akin, D. E. (1989). Histological and physical factors affecting digestibility of forages. *Agron. J.* 81, 17–25. doi: 10.2134/agronj1989.00021962008100010004x
- Aufrère, J., and Michalet-Doreau, B. (1983). "In vivo digestibility and prediction of digestibility of some by-products," in *Proceedings of an EEC Seminar*, Melle Gontrode, 26–29.
- Barriere, Y. (2017). Brown-midrib genes in maize and their efficiency in dairy cow feeding. Perspectives for breeding improved silage maize targeting gene modifications in the monolignol and p-hydroxycinnamate pathways. *Maydica* 62, 1–19.
- Barrière, Y., Mechin, V., Riboulet, C., Guillaumie, S., Thomas, J., Bosio, M., et al. (2009b). Genetic and genomic approaches for improving biofuel production from maize. *Euphytica* 170, 183–202. doi: 10.1007/s10681-009-9923-6

- Barrière, Y., Mechin, V., Lafarguette, F., Manicacci, D., Guillon, F., Wang, H., et al. (2009a). Toward the discovery of maize cell wall genes involved in silage quality and capacity to biofuel production. *Maydica* 54, 161–198.
- Barros-Rios, J., Santiago, R., Malvar, R. A., and Jung, H.-J. G. (2012). Chemical composition and cell wall polysaccharide degradability of pith and rind tissues from mature maize internodes. *Anim. Feed Sci. Technol.* 172, 226–236. doi: 10.1016/j.anifeeds.2012.01.005
- Beć, K. B., Grabska, J., Bonn, G. K., Popp, M., and Huck, C. W. (2020). Principles and applications of vibrational spectroscopic imaging in plant science: a review. *Front. Plant Sci.* 11:1226. doi: 10.3389/fpls.2020.01226
- Boon, E. J. M. C., Struik, P. C., Tamminga, S., Engels, F. M., and Cone, J. W. (2008). Stem characteristics of two forage maize (*Zea mays* L.) cultivars varying in whole plant digestibility. III. Intra-stem variability in anatomy, chemical composition and in vitro rumen fermentation. *NJAS Wageningen J. Life Sci.* 56, 101–122. doi: 10.1016/S1573-5214(08)80019-X
- Carpita, N. C., and McCann, M. C. (2008). Maize and sorghum: genetic resources for bioenergy grasses. *Trends Plant Sci.* 13, 415–420. doi: 10.1016/j.tplants.2008.06.002
- Casler, M. D., and Jung, H. J. (2017). Lignin and etherified ferulates affect digestibility and structural composition of three temperate perennial grasses. *Crop Sci.* 57, 1010–1019. doi: 10.2135/cropsci2016.05.0311
- Chazal, R., Robert, P., Durand, S., Devaux, M.-F., Saulnier, L., Lapiere, C., et al. (2014). Investigating lignin key features in maize lignocelluloses using infrared spectroscopy. *Appl. Spectrosc.* 68, 1342–1347. doi: 10.1366/14-07472
- Cone, J. W., and Engels, F. M. (1993). The influence of ageing on cell wall composition and degradability of three maize genotypes. *Anim. Feed Sci. Technol.* 40, 331–342. doi: 10.1016/0377-8401(93)90062-O
- Corcel, M., Devaux, M.-F., Guillon, F., and Barron, C. (2016). Comparison of UV and visible autofluorescence of wheat grain tissues in macroscopic images of cross-sections and particles. *Comput. Electron. Agric.* 127, 281–288. doi: 10.1016/j.compag.2016.06.016
- de Juan, A., Maeder, M., Hancewicz, T., Duponchel, L., and Tauler, R. (2009). “Chemometric tools for image analysis,” in *Infrared and Raman Spectroscopic Imaging*, eds R. Sazler and H. W. Seiler (Hoboken, NJ: Wiley-VCH Verlag GmbH & Co. KGaA), 65–109.
- Dence, C. W. (1992). “The determination of lignin,” in *Methods in Lignin Chemistry*, eds S. Y. Lin and C. W. Dence (Berlin: Springer Berlin Heidelberg), 33–61.
- Devaux, M.-F., Bouchet, B., Legland, D., Guillon, F., and Lahaye, M. (2008). Macro-vision and grey level granulometry for quantification of tomato pericarp structure. *Postharvest Biol. Technol.* 47, 199–209. doi: 10.1016/j.postharvbio.2007.06.017
- Devaux, M.-F., Jamme, F., André, W., Bouchet, B., Alvarado, C., Durand, S., et al. (2018). Synchrotron time-lapse imaging of lignocellulosic biomass hydrolysis: tracking enzyme localization by protein autofluorescence and biochemical modification of cell walls by microfluidic infrared microspectroscopy. *Front. Plant Sci.* 9:200. doi: 10.3389/fpls.2018.00200
- Devaux, M. F., and Legland, D. (2014). “Grey level granulometry for histological image analysis of plant tissues,” in *Microscopy: Advances in Scientific Research and Education*, ed. A. Méndez-Vilas (Badajoz: Formatex Research Center), 624.
- Devaux, M.-F., Sire, A., and Papineau, P. (2009). *Macrovision et Analyse Granulométrique en Niveaux de Gris pour l'Analyse Histologique de Tissus Végétaux*. Cahier Techniques de l'INRA. INRAE – France. 93–100. Available online at: https://www6.inrae.fr/novae/content/download/3266/31715/version/1/file/93_Devaux_imag.pdf
- Ding, S. Y., Liu, Y. S., Zeng, Y., Himmel, M. E., Baker, J. O., and Bayer, E. A. (2012). How does plant cell wall nanoscale architecture correlate with enzymatic digestibility? *Science* 338, 1055–1060. doi: 10.1126/science.1227491
- Djikanović, D., Kalauzi, A., Radotić, K., Lapiere, C., and Jeremić, M. (2007). Deconvolution of lignin fluorescence spectra: a contribution to the comparative structural studies of lignins. *Russ. J. Phys. Chem. A* 81, 1425–1428. doi: 10.1134/S0036024407090142
- Donaldson, L. (2013). Softwood and hardwood lignin fluorescence spectra of wood cell walls in different mounting media. *IAWA J.* 34, 3–19. doi: 10.1163/22941932-00000002
- Donaldson, L. (2020). Autofluorescence in plants. *Molecules* 25:2393. doi: 10.3390/molecules25102393
- Donaldson, L., Radotić, K., Kalauzi, A., Djikanović, D., and Jeremić, M. (2010). Quantification of compression wood severity in tracheids of *Pinus radiata* D. Don using confocal fluorescence imaging and spectral deconvolution. *J. Struct. Biol.* 169, 106–115. doi: 10.1016/j.jsb.2009.09.006
- Donaldson, L., and Williams, N. (2018). Imaging and spectroscopy of natural fluorophores in pine needles. *Plants (Basel, Switzerland)* 7:10. doi: 10.3390/plants7010010
- El Hage, F., Legland, D., Borrega, N., Jacquemot, M. P., Griveau, Y., Coursol, S., et al. (2018). Tissue lignification, cell wall p-coumaroylation and degradability of maize stems depend on water status. *J. Agric. Food Chem.* 66, 4800–4808. doi: 10.1021/acs.jafc.7b05755
- Esau, K. (1977). *Anatomy of Seed Plants*. New York, NY: John Wiley & Sons Ltd.
- Fontaine, A. S., Bout, S., Barrière, Y., and Vermerris, W. (2003). Variation in cell wall composition among forage maize (*Zea mays* L.) inbred lines and its impact on digestibility: analysis of neutral detergent fiber composition by pyrolysis-gas chromatography-mass spectrometry. *J. Agric. Food Chem.* 51, 8080–8087. doi: 10.1021/jf0034321g
- Fulcher, R., O'Brien, T. P., and Lee, J. W. (1971). Studies on the aleurone layer I. Conventional and fluorescence microscopy of the cell wall with emphasis on phenol-carbohydrate complexes in wheat. *Aust. J. Biol. Sci.* 25, 23–34. doi: 10.1071/BI9720023
- Gärtner, H., Lucchinetti, S., and Schweingruber, F. H. (2014). New perspectives for wood anatomical analysis in dendrosciences: the GSL1-microtome. *Dendrochronologia* 32, 47–51. doi: 10.1016/j.dendro.2013.07.002
- Geladi, P., and Grahn, H. F. (2006). “Multivariate image analysis,” in *Encyclopedia of Analytical Chemistry*, ed. R. A. Meyers (Chichester: John Wiley & Sons).
- Ghaffari, M., Chateigner-Boutin, A.-L., Guillon, F., Devaux, M.-F., Abdollahi, H., and Duponchel, L. (2019). Multi-excitation hyperspectral autofluorescence imaging for the exploration of biological samples. *Anal. Chim. Acta* 1062, 47–59. doi: 10.1016/j.aca.2019.03.003
- Gierlinger, N. (2018). New insights into plant cell walls by vibrational microspectroscopy. *Appl. Spectrosc. Rev.* 53, 517–551. doi: 10.1080/05704928.2017.1363052
- Harris, P. J., and Hartley, R. D. (1976). Detection of bound ferulic acid in cell walls of the Gramineae by ultraviolet fluorescence microscopy. *Nature* 259, 508–510. doi: 10.1038/259508a0
- Hatfield, R. D., Rancour, D. M., and Marita, J. M. (2017). Grass cell walls: a story of cross-linking. *Front. Plant Sci.* 7:15. doi: 10.3389/fpls.2016.02056
- Hatfield, R. D., Wilson, J. R., and Mertens, D. R. (1999). Composition of cell walls isolated from cell types of grain sorghum stems. *J. Sci. Food Agric.* 79, 891–899.
- Heckwolf, S., Heckwolf, M., Kaeppler, S. M., de Leon, N., and Spalding, E. P. (2015). Image analysis of anatomical traits in stalk transections of maize and other grasses. *Plant Methods* 11:10. doi: 10.1186/s13007-015-0070-x
- Ho-Yue-Kuang, S., Alvarado, C., Antelme, S., Bouchet, B., Cézard, L., Le Bris, P., et al. (2016). Mutation in *Brachypodium cafeeic acid O-methyltransferase 6* alters stem and grain lignins and improves straw saccharification without deteriorating grain quality. *J. Exp. Bot.* 67, 227–237. doi: 10.1093/jxb/erv446
- Jung, H. G., and Casler, M. D. (2006a). Maize stem tissues: cell wall concentration and composition during development. *Crop Sci.* 46, 1793–1800. doi: 10.2135/cropsci2006.02.0086r
- Jung, H. G., and Casler, M. D. (2006b). Maize stem tissues: impact of development on cell wall degradability. *Crop Sci.* 46, 1801–1809. doi: 10.2135/cropsci2006.02-0086
- Jung, H. G., and Engels, F. M. (2001). Alfalfa stem tissues: rate and extent of cell-wall thinning during ruminal degradation. *Neth. J. Agric. Sci.* 49, 3–13. doi: 10.1016/S1573-5214(01)80012-9
- Jung, H. G., and Phillips, R. L. (2010). Putative seedling ferulate ester (sfe) maize mutant: morphology, biomass yield, and stover cell wall composition and rumen degradability. *Crop Sci.* 50, 403–418. doi: 10.2135/cropsci2009.04.0191
- Jung, H. J. G., and Buxton, D. R. (1994). Forage quality variation among maize inbred relationships of cell-wall composition and in-vitro degradability for stem internodes. *J. Sci. Food Agric.* 66:313. doi: 10.1002/JSEA.2740660308
- Lang, M., Stober, F., and Lichtenthaler, H. K. (1991). Fluorescence emission spectra of plant leaves and plant constituents. *Radiat. Environ. Biophys.* 30, 333–347. doi: 10.1007/BF01210517

- Legland, D., Devaux, M. F., and Guillon, F. (2014). Statistical mapping of maize bundle intensity at the stem scale using spatial normalisation of replicated images. *PLoS One* 9:10. doi: 10.1371/journal.pone.0090673
- Legland, D., El-Hage, F., Méchin, V., and Reymond, M. (2017). Histological quantification of maize stem sections from FASGA-stained images. *Plant Methods* 13:84. doi: 10.1186/s13007-017-0225-z
- Legland, D., Guillon, F., and Devaux, M.-F. (2020). Parametric mapping of cellular morphology in plant tissue sections by gray level granulometry. *Plant Methods* 16:63. doi: 10.1186/s13007-020-00603-7
- Lopez, F. B., and Barclay, G. F. (2017). "Chapter 4—Plant anatomy and physiology," in *Pharmacognosy*, eds S. Badal and R. Delgoda (Boston, MA: Academic Press), 45–60.
- McCann, M. C., and Carpita, N. C. (2015). Biomass recalcitrance: a multi-scale, multi-factor, and conversion-specific property. *J. Exp. Bot.* 66, 4109–4118. doi: 10.1093/jxb/erv267
- Méchin, V., Argillier, O., Barrière, Y., and Menanteau, V. (1998). Genetic variation in stems of normal and brown-midrib 3 maize inbred lines. Towards similarity for in vitro digestibility and cell wall composition. *Maydica* 43:205.
- Méchin, V., Argillier, O., Menanteau, V., Barrière, Y., Mila, I., Pollet, B., et al. (2000). Relationship of cell wall composition to in vitro cell wall digestibility of maize inbred line stems. *J. Sci. Food Agric.* 80, 574–580. doi: 10.1002/(SICI)1097-0010(200004)80:5<574::AID-JSFA575<3.0.CO;2-R
- Méchin, V., Argillier, O., Rocher, F., Hébert, Y., Mila, I., Pollet, B., et al. (2005). In search of a maize ideotype for cell wall enzymatic degradability using histological and biochemical lignin characterization. *J. Agric. Food Chem.* 53, 5872–5881. doi: 10.1021/jf050722f
- Melati, R. B., Shimizu, F. L., Oliveira, G., Pagnocca, F. C., de Souza, W., Sant'Anna, C., et al. (2019). Key factors affecting the recalcitrance and conversion process of biomass. *Bioenerg. Res.* 12, 1–20. doi: 10.1007/s12155-018-9941-0
- Meng, X., and Ragauskas, A. J. (2014). Recent advances in understanding the role of cellulose accessibility in enzymatic hydrolysis of lignocellulosic substrates. *Curr. Opin. Biotechnol.* 27, 150–158. doi: 10.1016/j.copbio.2014.01.014
- Morrison, T. A., Jung, H. G., Buxton, D. R., and Hatfield, R. D. (1998). Cell-wall composition of maize internodes of varying maturity. *Crop Sci.* 38:455. doi: 10.2135/cropsci1998.0011183X003800020031x
- Perrier, L., Rouan, L., Jaffuel, S., Clément-Vidal, A., Roques, S., Soutiras, A., et al. (2017). Plasticity of sorghum stem biomass accumulation in response to water deficit: a multiscale analysis from internode tissue to plant level. *Front Plant Sci* 8:1516. doi: 10.3389/fpls.2017.01516
- Philippe, S., Tranquet, O., Utille, J.-P., Saulnier, L., and Guillon, F. (2007). Investigation of ferulate deposition in endosperm cell walls of mature and developing wheat grains by using a polyclonal antibody. *Planta* 225, 1287–1299. doi: 10.1007/s00425-006-0422-x
- Ralph, J., Hatfield, R. D., Grabber, J. H., Jung, H.-J. G., Quideau, S., and Helm, R. F. (1998). "Cell wall cross-linking in grasses by ferulates and diferulates," in *Lignin and Lignan Biosynthesis*, eds N. G. Lewis and S. Sarkanen (Washington, DC: American Chemical Society).
- Salzer, R., and Siesler, H. W. (2014). *Infrared and Raman Spectroscopic Imaging*. Weinheim: John Wiley & Sons.
- Scobbie, L., Russell, W., Provan, G. J., and Chesson, A. (1993). The newly extended maize internode: a model for the study of secondary cell wall formation and consequences for digestibility. *J. Sci. Food Agric.* 61, 217–225. doi: 10.1002/jsfa.2740610213
- Soille, P. (2003). *Morphological Image Analysis: Principles and Applications*. Berlin: Springer-Verlag.
- Terrett, O. M., and Dupree, P. (2019). Covalent interactions between lignin and hemicelluloses in plant secondary cell walls. *Curr. Opin. Biotechnol.* 56, 97–104. doi: 10.1016/j.copbio.2018.10.010
- Tolivia, D., and Tolivia, J. (1987). Fasga: a new polychromatic method for simultaneous and differential staining of plant tissues. *J. Microsc.* 148, 113–117. doi: 10.1111/j.1365-2818.1987.tb02859.x
- Tranquet, O., Saulnier, L., Utille, J.-P., Ralph, J., and Guillon, F. (2009). Monoclonal antibodies to p-coumarate. *Phytochemistry* 70, 1366–1373. doi: 10.1016/j.phytochem.2009.06.019
- Vanholme, R., De Meester, B., Ralph, J., and Boerjan, W. (2019). Lignin biosynthesis and its integration into metabolism. *Curr. Opin. Biotechnol.* 56, 230–239. doi: 10.1016/j.copbio.2019.02.018
- Vo, L. T., Girones, J., Jacquemot, M.-P., Legée, F., Cezard, L., Lapiere, C., et al. (2020). Correlations between genotype biochemical characteristics and mechanical properties of maize stem-polyethylene composites. *Ind. Crops Prod.* 143:111925. doi: 10.1016/j.indcrop.2019.111925
- Willemse, M. T. M., and Emons, A. M. C. (1991). Autofluorescence and HPLC analyses of phenolics in *Zea mays* L. stem cell walls. *Acta Bot. Neerl.* 40, 115–124. doi: 10.1111/j.1438-8677.1991.tb01519.x
- Wilson, J. R., and Mertens, D. R. (1995). Cell wall accessibility and cell structure limitations to microbial digestion of forage. *Crop Sci.* 35, 251–259. doi: 10.2135/cropsci1995.0011183X003500010046x
- Wilson, J. R., Mertens, D. R., and Hatfield, R. D. (1993). Isolates of cell types from sorghum stems: digestion, cell wall and anatomical characteristics. *J. Sci. Food Agric.* 63, 407–417. doi: 10.1002/jsfa.2740630406
- Zhang, Y., Du, J., Wang, J., Ma, L., Lu, X., Pan, X., et al. (2018). High-throughput micro-phenotyping measurements applied to assess stalk lodging in maize (*Zea mays* L.). *Biol. Res.* 51:40. doi: 10.1186/s40659-018-0190-7
- Zhang, Y., Legay, S., Barrière, Y., Meichin, V. R., and Legland, D. (2013). Color quantification of stained maize stem section describes lignin spatial distribution within the whole stem. *J. Agric. Food Chem.* 61, 3186–3192. doi: 10.1021/jf400912s
- Zhang, Y., Legland, D., El Hage, F., Devaux, M.-F., Guillon, F., Reymond, M., et al. (2019). Changes in cell walls lignification, feruloylation and p-coumaroylation throughout maize internode development. *PLoS One* 14:e0219923. doi: 10.1371/journal.pone.0219923
- Zhang, Y., Ma, L., Wang, J., Wang, X., Guo, X., and Du, J. (2020). Phenotyping analysis of maize stem using micro-computed tomography at the elongation and tasseling stages. *Plant Methods* 16:2. doi: 10.1186/s13007-019-0549-y
- Zhang, Y., Wang, J., Du, J., Zhao, Y., Lu, X., Wen, W., et al. (2021). Dissecting the phenotypic components and genetic architecture of maize stem vascular bundles using high-throughput phenotypic analysis. *Plant Biotechnol. J.* 19, 35–50. doi: 10.1111/pbi.13437
- Zoghalmi, A., and Paës, G. (2019). Lignocellulosic biomass: understanding recalcitrance and predicting hydrolysis. *Front. Chem.* 7:874. doi: 10.3389/fchem.2019.00874

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Berger, Devaux, Legland, Barron, Delord and Guillon. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Using Unmanned Aerial Vehicle-Based Multispectral Image Data to Monitor the Growth of Intercropping Crops in Tea Plantation

Yujie Shi¹, Yuan Gao², Yu Wang¹, Danni Luo¹, Sizhou Chen¹, Zhaotang Ding^{1*} and Kai Fan^{1*}

¹ Tea Research Institute, Qingdao Agricultural University, Qingdao, China, ² Jinan Agricultural Technology Promotion Service Center, Jinan, China

OPEN ACCESS

Edited by:

Gregorio Egea,
University of Seville, Spain

Reviewed by:

Abbas Atefi,
California Polytechnic State University,
United States

Liujiun Li,
Missouri University of Science
and Technology, United States

*Correspondence:

Zhaotang Ding
dztttea@163.com
Kai Fan
fankaitea@163.com

Specialty section:

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

Received: 23 November 2021

Accepted: 20 January 2022

Published: 25 February 2022

Citation:

Shi Y, Gao Y, Wang Y, Luo D,
Chen S, Ding Z and Fan K (2022)
Using Unmanned Aerial
Vehicle-Based Multispectral Image
Data to Monitor the Growth
of Intercropping Crops in Tea
Plantation.
Front. Plant Sci. 13:820585.
doi: 10.3389/fpls.2022.820585

Aboveground biomass (AGB) and leaf area index (LAI) are important indicators to measure crop growth and development. Rapid estimation of AGB and LAI is of great significance for monitoring crop growth and agricultural site-specific management decision-making. As a fast and non-destructive detection method, unmanned aerial vehicle (UAV)-based imaging technologies provide a new way for crop growth monitoring. This study is aimed at exploring the feasibility of estimating AGB and LAI of mung bean and red bean in tea plantations by using UAV multispectral image data. The spectral parameters with high correlation with growth parameters were selected using correlation analysis. It was found that the red and near-infrared bands were sensitive bands for LAI and AGB. In addition, this study compared the performance of five machine learning methods in estimating AGB and LAI. The results showed that the support vector machine (SVM) and backpropagation neural network (BPNN) models, which can simulate non-linear relationships, had higher accuracy in estimating AGB and LAI compared with simple linear regression (LR), stepwise multiple linear regression (SMLR), and partial least-squares regression (PLSR) models. Moreover, the SVM models were better than other models in terms of fitting, consistency, and estimation accuracy, which provides higher performance for AGB (red bean: $R^2 = 0.811$, root-mean-square error (RMSE) = 0.137 kg/m², normalized RMSE (NRMSE) = 0.134; mung bean: $R^2 = 0.751$, RMSE = 0.078 kg/m², NRMSE = 0.100) and LAI (red bean: $R^2 = 0.649$, RMSE = 0.36, NRMSE = 0.123; mung bean: $R^2 = 0.706$, RMSE = 0.225, NRMSE = 0.081) estimation. Therefore, the crop growth parameters can be estimated quickly and accurately using the models established by combining the crop spectral information obtained by the UAV multispectral system using the SVM method. The results of this study provide valuable practical guidelines for site-specific tea plantations and the improvement of their ecological and environmental benefits.

Keywords: UAV, multispectral, machine learning, leaf area index, above-ground biomass

INTRODUCTION

Intercropping, as the essence of traditional agriculture, has the advantages of increasing yield and quality (Mao et al., 2014; Egesa et al., 2016), promoting the utilization of nutrient resources (Rivest et al., 2010; Crème et al., 2016; Davies et al., 2016), increasing biodiversity (Bainard et al., 2011; Sanaa et al., 2016), and reducing pests and weeds (Brooker et al., 2015; Lopes et al., 2016). Tea plants [*Camellia sinensis* (L.) O. Kuntze] are cultivated worldwide as an economical woody plant, which grow in warm, humid, and light scattering regions. The different intercropping patterns of tea plantations, such as tea-fruit and tea-soybean intercropping, will be more in line with the biological characteristics of tea plant growth by improving microenvironment and resource utilization. Previous studies have shown that diverse agroforestry-tea intercropping systems, such as tree/tea and soybean/tea cannot only regulate the ecological environment of tea plantation, improve the soil nutrition, but also reduce the occurrence of diseases and insect pests and grass, and achieve high yield and quality (Sedagathoor and Janatpoor, 2012; Li et al., 2019). However, the intercropping density and the growth status of intercropping crops have a great influence on the growth of tea plants (Natarajan and Willey, 1980; Huang et al., 2019). A better understanding of the growth and development of intercropping crops is of great significance for guiding young tea plantation intercropping techniques and improving planting benefits.

Aboveground biomass (AGB) and leaf area index (LAI) are two main parameters of crop growth, which can reflect the growth status of legumes intercropped in young tea plantations, thus contributing to production management in tea plantations (Li et al., 2015; Liu B. et al., 2017). Rapid and accurate estimation of these two parameters can provide a strong basis for the timely formulation of management measures for young tea plantations (Li B. et al., 2020). However, the traditional crop growth assessment method is based on destructive sampling, which is to manually collect data samples in the field, or use field measuring instruments to evaluate crops (Freeman et al., 2007; Yue et al., 2018; Afrasiabian et al., 2020). Although this method is accurate, it is destructive, labor-intensive, time-consuming, and not operationally feasible for large-scale spatial and temporal measurements (Wang et al., 2017). Another relatively new method is to use instruments for measurement, which is less destructive to crops, but external factors have a certain impact on experimental equipment, and it is also difficult to apply to rapid monitoring of field crops.

In recent years, high-throughput non-destructive plant phenotyping techniques based on UAV are becoming a powerful tool for crop monitoring, due to the advantages of convenient operation, high spatial and temporal resolution, and reasonable spatial coverage, such as crop plot detection (Liu H. et al., 2017), crop growth status monitoring (Pölönen et al., 2013; Harkel et al., 2019; Maimaitijiang et al., 2019), crop yield prediction (Zhou et al., 2017; Gilliot et al., 2020; Li B. et al., 2020), and plant water status assessment (Romero et al., 2018). Machine learning, as an important data analysis method, has

been used to establish crop remote sensing estimation models combined with spectral parameters of remote sensing images. For example, Jin et al. (2015) used a vegetation index (VI) and radar parameter to accurately estimate the LAI ($R^2 = 0.83$) and biomass ($R^2 = 0.90$) of winter wheat using partial least-squares regression (PLSR). Devia et al. (2019) used an unmanned aerial vehicle (UAV)-based multispectral system for aerial crop monitoring to combine seven VIs of rice growth in a multivariate regression model to estimate rice biomass. Furthermore, it was confirmed that this method could estimate crop biomass in a large area with an average correlation coefficient of 0.76. Han et al. (2019) pointed out that the random forest (RF) model derived from the crop surface model using VIs and crop height correlation indicators can predict corn biomass ($R^2 = 0.699$, root-mean-square error (RMSE) = 1.2), and its accuracy is slightly higher than that of the backpropagation artificial neural network (ANN) and stepwise multiple linear regression (SMLR) models. Qi et al. (2020) developed a model for the estimation of peanut LAI by using a backpropagation neural network (BPNN) with UAV-based multispectral image data ($R^2 = 0.968$, RMSE = 0.165). Tatsumi et al. (2021) constructed a tomato biomass estimation model using red-green-blue (RGB) and multispectral image data acquired from UAV with feature variable selection and machine learning and improved the estimation accuracy (rRMSE = 8.8–28.1%). Similarly, Jiang et al. (2019) established a model for the estimation of rice biomass by using RGB and multispectral image data obtained from UAV and further improved the estimation accuracy of the model by combining meteorological data with RF ($R^2 = 0.92$, RMSE = 126.28 g/m²).

However, there were few reports on the use of UAV-based multispectral image data combined with machine learning to monitor crop growth of tea plantations, and it is difficult to provide valuable data support and practical guidance for site-specific management decisions and the construction of smart tea plantations. Therefore, this study attempts to use UAV-based multispectral imagery combined with ground-measured sample data to explore the feasibility of estimating AGB and LAI using the spectral parameters in intercropping tea plantations. The spectral parameters sensitive to crop growth response were selected according to the correlation analysis. Then, the remote sensing monitoring models of intercropping crop growth parameters suitable for young tea plantation were constructed using machine learning, and the estimation performance of five machine learning models was evaluated: (1) Simple linear regression (LR), (2) SMLR, (3) PLSR models, (4) support vector machine (SVM), and (5) BPNN. We hypothesized that the SVM method can simulate both linear and non-linear relationships between multiple independent variables and one factor. Compared with other modeling methods, the SVM model should have a higher degree of explanation for AGB and LAI. It is hoped that the results of this study can provide basic data and theoretical support for the growth monitoring of crops in young tea plantations in order to provide valuable practical guidelines for site-specific tea plantations and the improvement of their ecological and environmental benefits.

MATERIALS AND METHODS

Study Area and Experimental Design

The field experiment was conducted at the tea research demonstration base of Qingdao Agricultural University (36°26' N, 120°34' E, average altitude 54.47 m a.s.l.). The area has a warm temperate continental monsoon climate, with precipitation mostly occurring during summer and autumn and a large temperature difference between day and night. The average annual temperature is 12.1° (the annual maximum/minimum temperature is 38.6/−18.6°), and the annual average precipitation is 708.9 mm. The experimental tea plantation covers an area of 100 m × 30 m, with a soil pH of 6.5. The location diagram of the experiment area is shown in **Figure 1A**.

The tea plants planted in the tea plantation are half-year seedlings and the variety was Zhongcha 108, with a total of 11 rows. In early June 2020, mung bean and red bean were planted in the tea plantation, and the varieties, namely, Zhonglv 4 and Qidonghong were used. Red bean (*Vigna angularis* L.) and mung bean (*V. radiata* (L.) Wilczek) were planted in rotation with 6 rows each. Each row was divided into 10 plots, with a total of 120 plots. The specific test design is shown in **Figure 1B**.

Unmanned Aerial Vehicle Imagery Data Acquisition and Preprocessing

Multispectral cameras and accessories were mounted on a UAV platform (DJI M200 V2, DJI, Shenzhen, China) during data collection. The UAV has four propellers, is equipped with two 7,660 mAh (22.8 V) batteries with a battery life of 38 min, and can maintain stability at low speed and low altitude; for the data acquisition, the takeoff mass was 5.5 kg. Images were taken at 25 m above ground level (AGL) at a speed of 1.5 m/s. The collection dates were July 24 and August 11, 2020.

Multispectral images were acquired using a multispectral camera MS600 (Yusense, Qindao, Shandong, China), which

has a resolution (effective pixels) of 1,280 × 960 pixels. The multispectral camera used in this experiment was equipped with six spectral wavebands, namely blue, green, red, red edge, and two near-infrared wavebands (**Table 1**). A downward light sensor system was installed horizontally on the top of the UAV to measure the environmental irradiance and the readings of post-calibrate reflectance. As another source of radiometric calibration data, the standard panel attached to the multispectral camera was used for image calibration on the ground before each flight. Images in this study were captured in sub-centimeter pixel resolution, and the flight survey was configured with an 80% side and 80% forward overlap. The original multispectral images obtained from each aerial photography operation were processed using Yusense map V1.0 software (Yusense, Qindao, Shandong, China) to generate a complete multispectral image. Then, the average digital number (DN) values of the six bands of each experimental cell are extracted using ENVI 5.2 software (Research Systems Inc., Boulder Co., United States) for subsequent processing.

Ground Data Acquisition

Field measurements were conducted on the same days as the UAV surveys to provide ground-truth data. To measure the LAI of red bean and mung bean accurately, a place with uniform crop growth (1 m × 1 m) in each plot was selected to measure the LAI using CI-110 plant canopy digital image analyzer (CID Bio-Science Inc., WA, United States). When measuring LAI, direct sunlight was avoided. First, a blank value was measured above the crop canopy, and then four values were randomly measured below the crop canopy. The average LAI of mung bean and red bean in the community was obtained by maintaining the lens level throughout the measurement, and the results are shown in **Figure 2**.

After the measurement of LAI, mung bean or red bean were randomly selected from experimental plots, which were intercepted from the height of 1 cm above the ground, and the

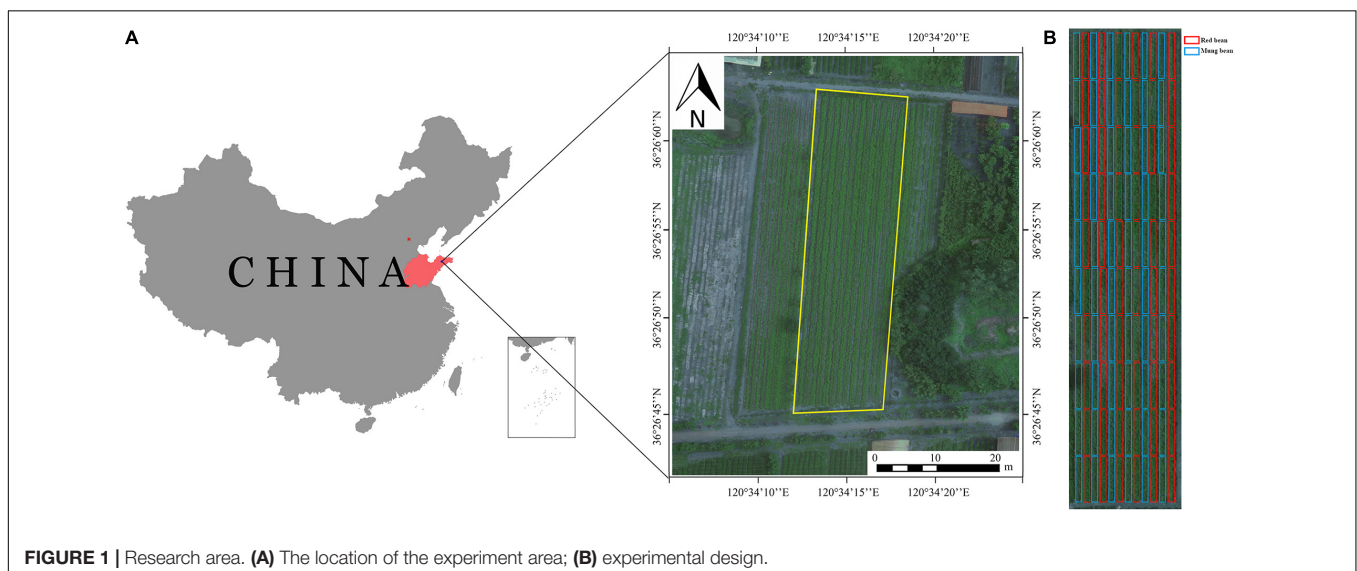








TABLE 1 | Center wavelength and full width at half maximum (FWHM) bandwidth of each spectral band of the multispectral camera.

Spectral band	Color	Sample	Center wavelength (nm)	Bandwidth FWHM (nm)
Blue	Blue		450	25
Green	Green		555	25
Red	Red		660	25
Red edge	Pink		710	25
Near infrared	Light purple		840	25
Near infrared	Purple		940	25

total number of plants in the sampling area was measured. The sample was placed in a paper bag, and the fresh biomass of the sample was measured immediately. The paper bag was placed in an oven at 80° for 24 h and maintained in a constant mass state. Then the sample was weighed to determine the dry mass to estimate the total biomass of the whole plot, and the results are shown in **Figure 2**.

Selection of Spectral Parameters

A spectral parameter should combine the reflectance of different bands with a VI in a certain way, which can reduce the influence of background environmental information on the crop canopy spectrum. According to previous studies, we selected 22 VIs and combined them with the 6 spectral bands of the MS600 multispectral camera to estimate the AGB and LAI of red bean and mung bean. Specific spectral parameters are shown in **Table 2**.

Data Analysis

In this study, 120 datasets of red bean and mung bean were collected. Each dataset was composed of ground measurement data and UAV remote sensing data. In data analysis, three-fourth

(90 datasets) and one-fourth (30 datasets) of the total data were divided into training sets and test sets, respectively. In the training sets, the LR method was used to establish growth parameter estimation models based on a single spectral parameter, and the SMLR method was used to establish growth parameter estimation models based on multiple spectral parameters. These two different established models were evaluated using the test datasets. The feasibility of the models was evaluated by the coefficient of determination (R^2), root-mean-square error (RMSE), and normalized RMSE (NRMSE). A larger R^2 value indicates a better model fit, while smaller RMSE and NRMSE values indicate a higher model accuracy. Finally, the estimation models of AGB and LAI were established by using three machine learning methods: PLSR, SVM, and BPNN. In the process of model building, the random 10-fold cross-validation method was used to divide 120 sample data into 10 parts. Each time, 90% of all samples was used to fit the model, and the remaining 10% was used as a test set to estimate performance metrics. This process was repeated ten times, and each model was run 100 times in total. The mean values of R^2 , RMSE, and NRMSE were calculated to evaluate the accuracy of AGB and LAI estimation models. The values of R^2 , RMSE, and NRMSE were calculated using the following formulas (1)–(3), respectively:

$$R^2 = 1 - \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - x_i)^2}{n}} \quad (2)$$

$$NRMSE = \frac{RMSE}{\bar{X}} \quad (3)$$

where x_i is the measured AGB or LAI for red bean and mung bean, \bar{x} is the average measured AGB or LAI, y_i is the AGB or LAI predicted by the model, and n is the number of data points.

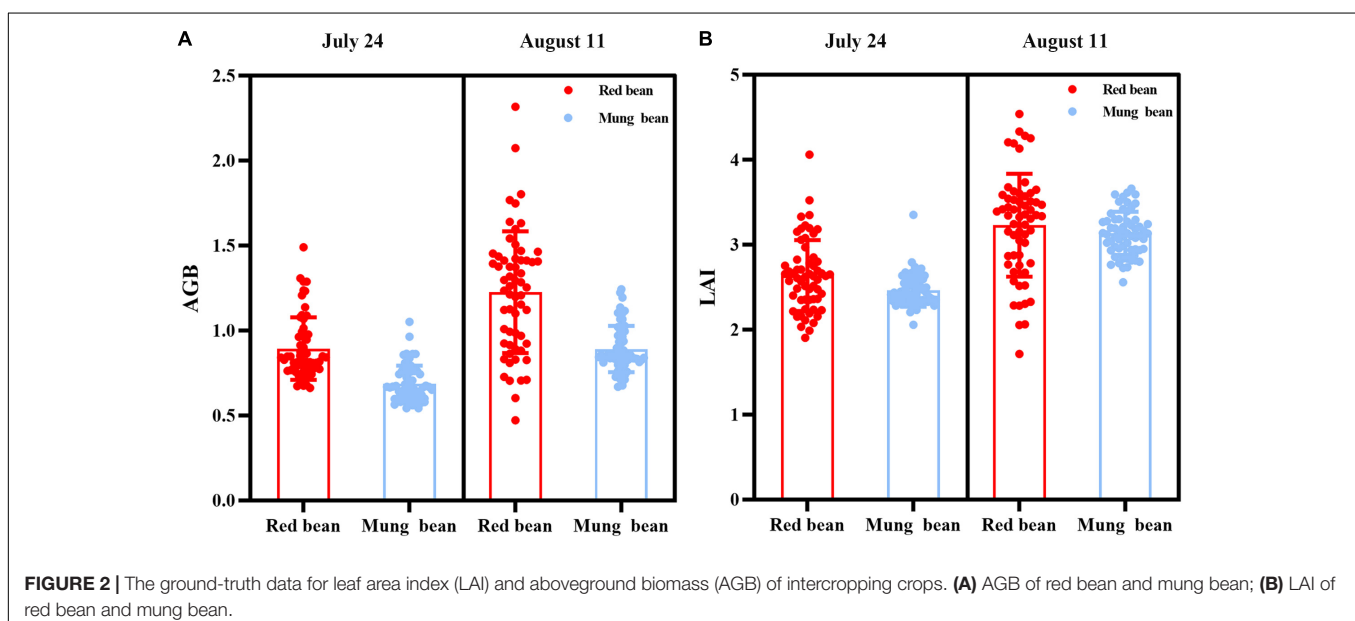


TABLE 2 | The spectral parameters used in this study.

Spectral parameters	Calculation formula	References
B.450	/	/
G.555	/	/
R.660	/	/
RE.710	/	/
NIR.840	/	/
NIR.940	/	/
DVI	NIR.840-G.555	Naito et al., 2017
NDVI	$(\text{NIR.840}-\text{R.660})/(\text{NIR.840}+\text{R.660})$	Rouse et al., 1974
EVI	$2.5 \cdot (\text{NIR.940}-\text{G.555})/(\text{NIR.940}+6 \cdot \text{R.660}-7.5 \cdot \text{B.450}+1)$	Prabhakara et al., 2015
GNDVI	$(\text{NIR.940}-\text{G.555})/(\text{NIR.940}+\text{G.555})$	Wang et al., 2007
PPR	$(\text{G.555}-\text{B.450})/(\text{G.555}+\text{B.450})$	Metternicht, 2003
SIPI	$(\text{NIR.940}-\text{B.450})/(\text{NIR.940}-\text{R.660})$	Penuelas et al., 1995
RECI	$\text{NIR.840}/\text{RE.710}-1$	Kanke et al., 2016
Red edge NDVI	$(\text{NIR.940}-\text{RE.710})/(\text{NIR.940}+\text{RE.710})$	Kanke et al., 2016
MERIS Terrestrial Chlorophyll Index (MTCI)	$(\text{NIR.840}-\text{RE.710})/(\text{RE.710}-\text{R.660})$	Panigada et al., 2010
Modified chlorophyll absorption ratio index (MCARI)	$[\text{RE.710}-\text{R.660}-0.2(\text{RE.710}-\text{R.660})]/(\text{RE.710}/\text{R.660})$	Wu et al., 2008
Triangular vegetation index (TVI)	$0.5 \cdot [120 \cdot (\text{NIR.840}-\text{G.555})-200 \cdot (\text{R.660}-\text{G.555})]$	Haboudane et al., 2004
Modified triangular vegetation index (MTVI2)	$1.5 \cdot [1.2 \cdot (\text{NIR.840}-\text{G.555})-2.5 \cdot (\text{R.660}-\text{G.555})]/[(12 \cdot \text{NIR.880}+1)^2-6 \cdot \text{NIR.880}-5 \cdot (\text{R.660})^2]-0.5]^{1/2}$	Haboudane et al., 2004
Transformed chlorophyll absorption reflectance index (TCARI)	$3 \cdot [(\text{RE.710}-\text{R.660})-0.2 \cdot (\text{RE.710}-\text{G.555})]/(\text{RE.710}/\text{G.555})$	Haboudane et al., 2004
Optimization of soil-adjusted vegetation index (OSAVI)	$1.16 \cdot (\text{NIR.840}-\text{R.660})/(\text{NIR.840}+\text{R.660}+0.16)$	Rondeaux et al., 1996
Ratio vegetation index (RVI1)	$\text{NIR.840}/\text{R.660}$	Kanke et al., 2016
PPR/NDVI	PPR/NDVI	Jin et al., 2017
SIPI/RVI1	$\text{SIPI}/\text{RVI1}$	Jin et al., 2017
Modified non-linear vegetation index (MNLI)	$1.5 \cdot [(\text{NIR.840})^2-\text{R.660}]/[(\text{NIR.840})^2+\text{R.660}+0.5]$	Yang Z. et al., 2008
Soil-adjusted vegetation index (SAVI)	$(\text{NIR.840}-\text{R.660})/(\text{NIR.840}+\text{R.660}+0.5)$	Pinty and Verstraete, 1992
Modified simple ratio (MSR)	$(\text{NIR.840}/\text{R.660}-1)/[(\text{NIR.840}/\text{R.660})^{1/2}+1]$	Wu et al., 2008
Non-linear vegetation index (NLI)	$[(\text{NIR.840})^2-\text{R.660}]/[(\text{NIR.840})^2+\text{R.660}]$	Goel and Qin, 1994
Renormalized difference vegetation index (RDVI)	$(\text{NIR.840}-\text{R.660})/(\text{NIR.840}+\text{R.660})^{1/2}$	Tucker, 1979

RESULTS

Correlation Analysis Between Spectral Parameters With Growth Parameters

To select the spectral parameters that are highly correlated with the growth parameters (AGB and LAI) of red bean and mung

bean, the correlation analysis between 28 spectral parameters and the growth parameters of red bean and mung bean (Figure 3) was carried out. For the AGB and LAI of red bean, the spectral parameters with the strongest correlation were RVI1 and red-edge chlorophyll index, and their correlation coefficients were 0.847 and 0.783, respectively. For the AGB and LAI of mung bean, the spectral parameters with the strongest correlation were RVI1 and B.450, and their correlation coefficients were 0.801 and 0.774, respectively. In general, most of the spectral parameters selected in this study had a strong correlation with the growth parameters, which can be used for the modeling and inversion of AGB and LAI of red bean and mung bean.

Estimation of Aboveground Biomass and Leaf Area Index Using Optimal Spectral Parameters Combined With Simple Linear Regression

To evaluate the direct relationship between spectral parameters and crop growth parameters, the LR method was used to establish AGB and LAI estimation models of red bean and mung bean in the training set using the optimal spectral parameters screened by correlation analysis (Table 3). Then, we verified the models with a test set (Figure 4). The training results showed that RVI1 could explain 76.1% (RMSE = 0.168 kg/m², NRMSE = 0.157) and 62.6% (RMSE = 0.088 kg/m², NRMSE = 0.113) of AGB variation in red bean and mung bean, respectively. As for LAI, the optimal spectral parameter red-edge chlorophyll index (RECI) could explain 63.4% (RMSE = 0.376, NRMSE = 0.129) of the LAI variation in red bean and B.450 could explain 59.1% (RMSE = 0.25, NRMSE = 0.09) of the LAI variation in mung bean. In addition, for growth parameters of red bean, these models deteriorated with the test dataset and the explanatory degree for AGB and LAI variation decreased to 52.4% (RMSE = 0.194 kg/m², NRMSE = 0.187) and 56.3% (RMSE = 0.357, NRMSE = 0.119), respectively (Figures 4A,C). In contrast, for growth parameters of mung bean, the models performed better with the test dataset and the explanatory degree for AGB and LAI variation increased to 66.3 and 62.1%, respectively. At the same time, the values of RMSE increased to 0.113 and 0.271, and the values of NRMSE increased to 0.138 and 0.096, respectively (Figures 4B,D).

Estimation Aboveground Biomass and Leaf Area Index Using Spectral Parameters Combined With Stepwise Multiple Linear Regression

To compare the growth parameter estimation models based on the optimal spectral parameters, we screened out 2–4 spectral parameters with a high correlation with the growth parameters of red bean and mung bean. Then, the SMLR method was used to establish AGB and LAI estimation models in the training set (Table 4). SMLR analysis showed that the models explained 85.7% (RMSE = 0.133 kg/m², NRMSE = 0.125) and 75.7% (RMSE = 0.073 kg/m², NRMSE = 0.093) of AGB variation in red bean and mung bean. Similar results were obtained for LAI.

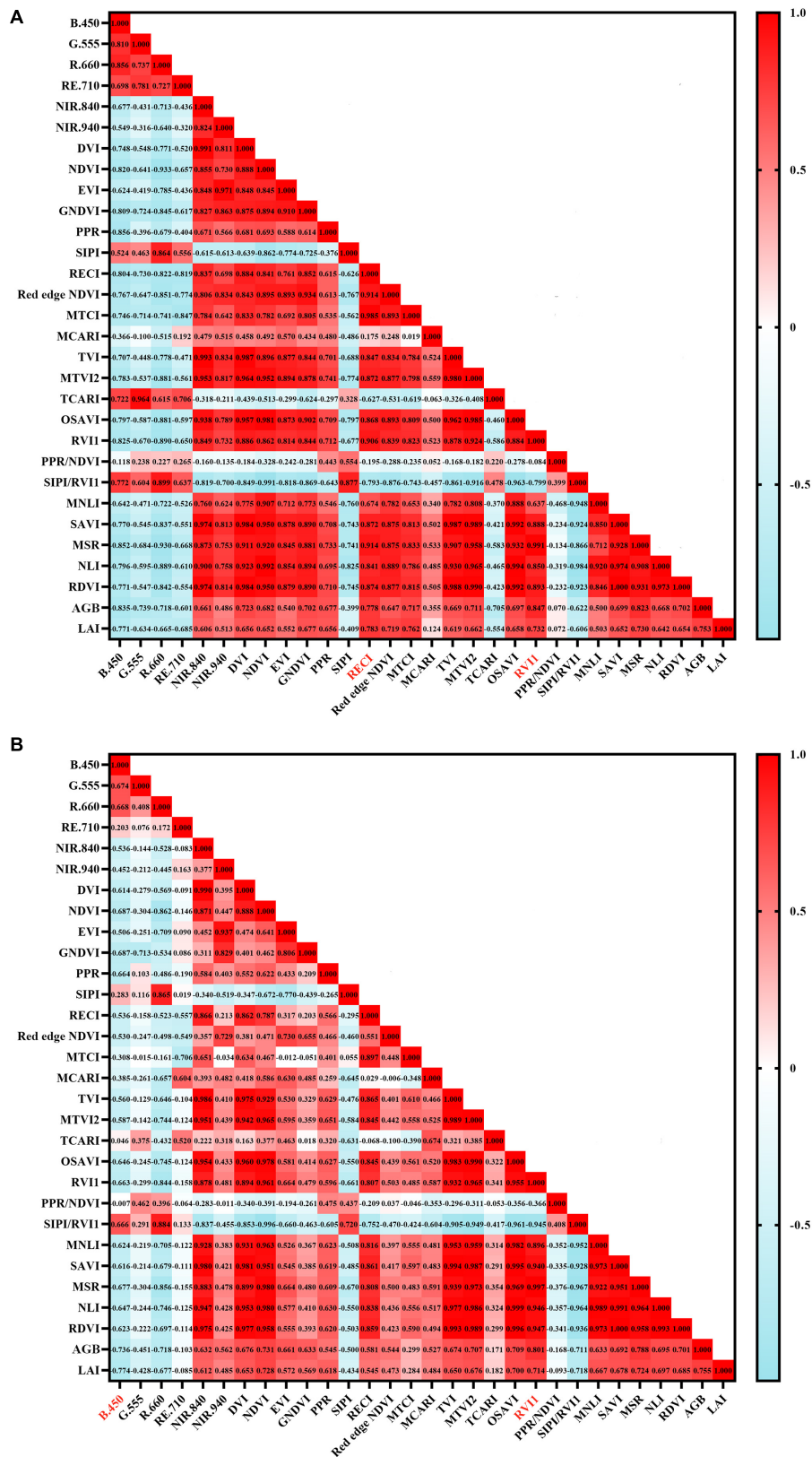
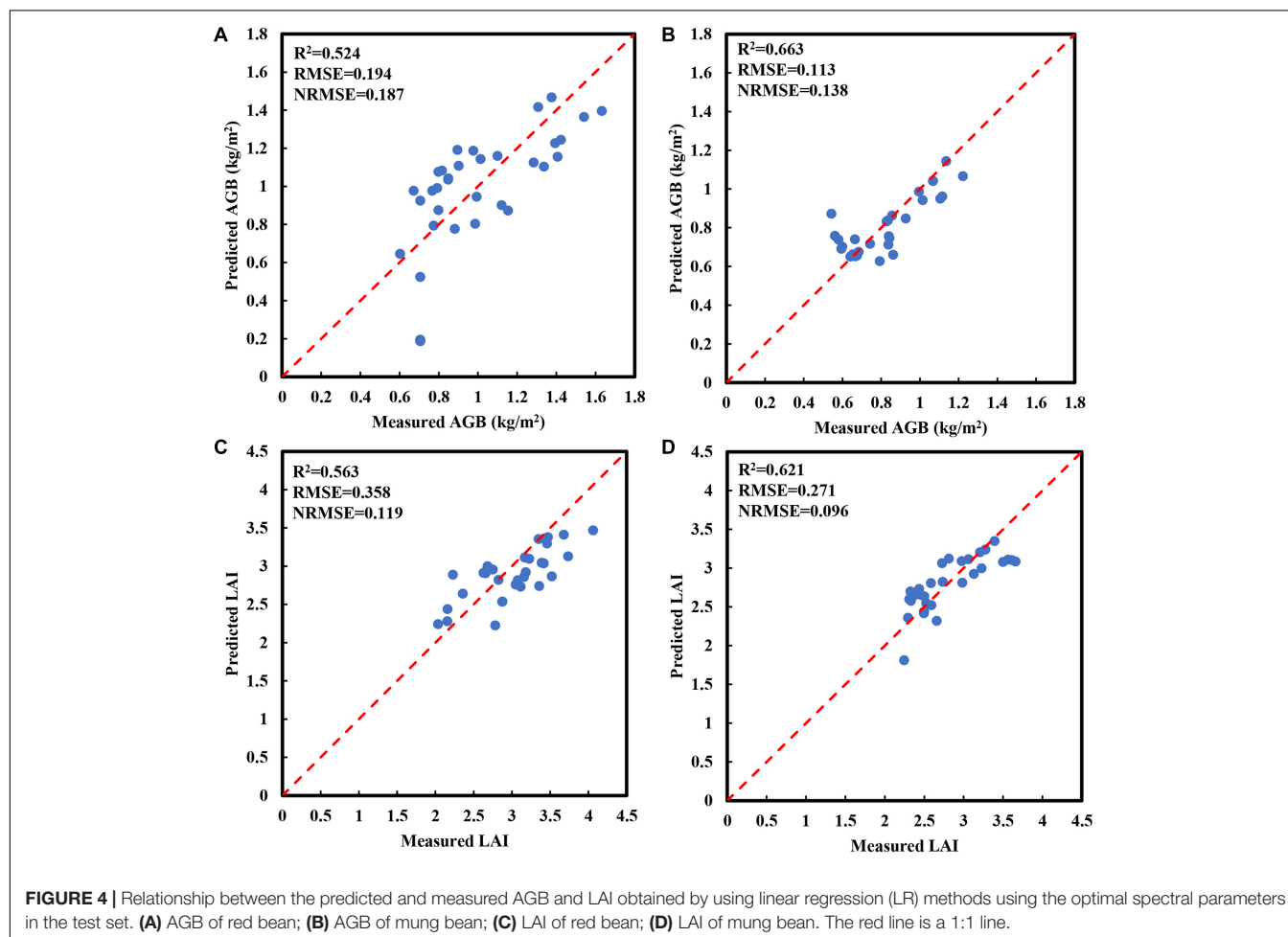


FIGURE 3 | Correlation coefficients between spectral parameters and growth parameters (AGB and LAI) of intercropped crops. **(A)** AGB and LAI of red bean; **(B)** AGB and LAI of mung bean.

TABLE 3 | Performance indicators of the AGB and LAI estimation models established by the LR method using the optimal spectral parameters in the training set.

Growth parameters	Intercropping crops	Optimal spectral parameters	Regression equation	Modeling accuracy		
				R^2	RMSE	NRMSE
AGB (kg/m ²)	Red bean	RV11	AGB = 0.059*RV11+0.313	0.761	0.168	0.157
	Mung bean	RV11	AGB = 0.054*RV11+1.55	0.626	0.088	0.113
LAI	Red bean	RECI	LAI = 0.616*RECI+0.355	0.634	0.376	0.129
	Mung bean	B.450	LAI = -74.297*B.450+5.292	0.591	0.25	0.09

**TABLE 4** | Performance indicators of AGB and LAI estimation models established by the SMLR methods in the training set.

Growth parameters	Intercropping crops	Regression equation	Modeling accuracy		
			R^2	RMSE	NRMSE
AGB (kg/m ²)	Red bean	AGB = 0.155*RV11-27.913*B.450-0.964*MSR-5.09*G.555 + 2.748	0.857	0.133	0.125
	Mung bean	AGB = 0.231703*RV11-1.1639*MSR-15.0778*B.450-3.64563*R.660 + 1.7216	0.757	0.073	0.093
LAI	Red bean	LAI = 0.478338*RECI-53.7192*B.450 + 0.123683*RV11-1.12239*MSR+ 4.65337	0.698	0.351	0.121
	Mung bean	LAI = -49.2931*B.450-3.39808*SIPI/RV11 + 4.98799	0.672	0.227	0.081

These models explained 69.8% (RMSE = 0.351, NRMSE = 0.121) and 67.2% of LAI (RMSE = 0.227, NRMSE = 0.081) variation in red bean and mung bean, respectively.

To evaluate the performance of the AGB and LAI estimation models constructed using SMLR, we plotted the relationship

between the measured values and predicted values of AGB and LAI in the test dataset (Figure 5). Compared with the training set, the SMLR model showed a greater decrease in the explanatory degree of AGB variation, indicating that the estimation accuracy of the model decreased significantly. The

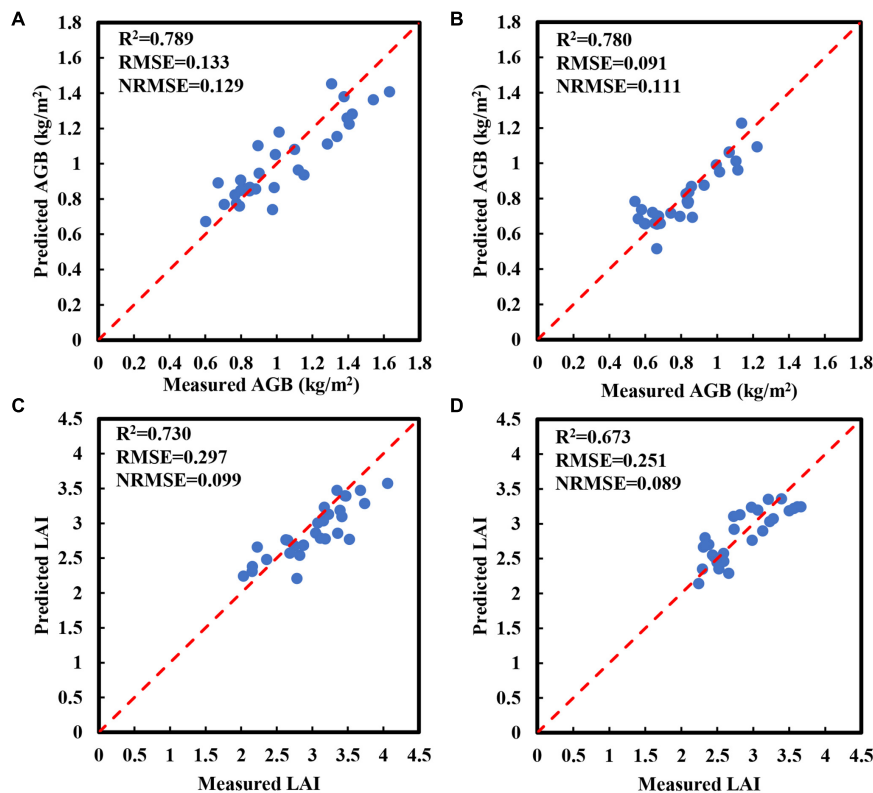


FIGURE 5 | Relationship between the predicted and measured AGB and LAI obtained by using the SMLR models within the test dataset. **(A)** AGB of red bean; **(B)** AGB of mung bean; **(C)** LAI of red bean; **(D)** LAI of mung bean. The red line is a 1:1 line.

NRMSE value increased to 0.129, indicating that the AGB estimation model of red bean was not stable. Compared with the training set, the SMLR model had lower explanatory power for AGB variation and higher NRMSE value, indicating that the accuracy of the estimation models of AGB of red bean decreased significantly and its stability was not good (Figure 5A). The accuracy of other models was basically consistent with the results of the training set, indicating that the stability of models was better. Compared with evaluation indexes of the LR models based on optimal spectral parameters, the R^2 values of SMLR models based on multispectral parameters increased, while the RMSE and NRMSE values decreased. These results indicated that the performance of SMLR models was better than LR models in estimating the growth parameters of red bean and mung bean.

Estimation of Aboveground Biomass and Leaf Area Index Using Spectral Parameters Combined With SVMs, Partial Least-Squares Regression, and Backpropagation Neural Network

In addition, to evaluate the performance of SVMs, PLSR, and BPNN in the estimation of crop growth parameters, we established AGB and LAI estimation models of red bean and mung bean by combining SVM, PLSR, and BPNN with

spectral parameters. To prevent overfitting caused by using too many independent variables when establishing models, we selected five spectral parameters with high correlation for each growth parameter for modeling and analyzing according to the results of correlation analysis (Supplementary Figure 1). The training results given in Figure 6 indicated that the SVM method showed better performance than other methods in the estimation of the AGB and LAI of red bean and mung bean. Compared with PLSR and BPNN models, SVM models had the highest R^2 values and relatively low RMSE and NRMSE values, indicating that SVM models had the highest accuracy in the estimation of the growth parameters of red bean and mung bean. Although BPNN also provided higher R^2 values in the estimation of the growth parameters of red bean and mung bean, the obtained RMSE and NRMS values were higher with high variability. In addition, the accuracy of estimating the AGB of the red bean by three methods was better than that of mung bean, but the performance was the opposite in LAI estimation. The SVM models obtained the highest values of R^2 and the lowest values of RMSE and NRMSE when estimating the growth parameters of red bean and mung bean in the test set (Figure 7). These results prove the excellent performance of the SVM models in estimating the growth parameters of red bean and mung bean. Similarly, the PLSR models were still the least applicable model for estimating the LAI and AGB of red bean and mung bean.

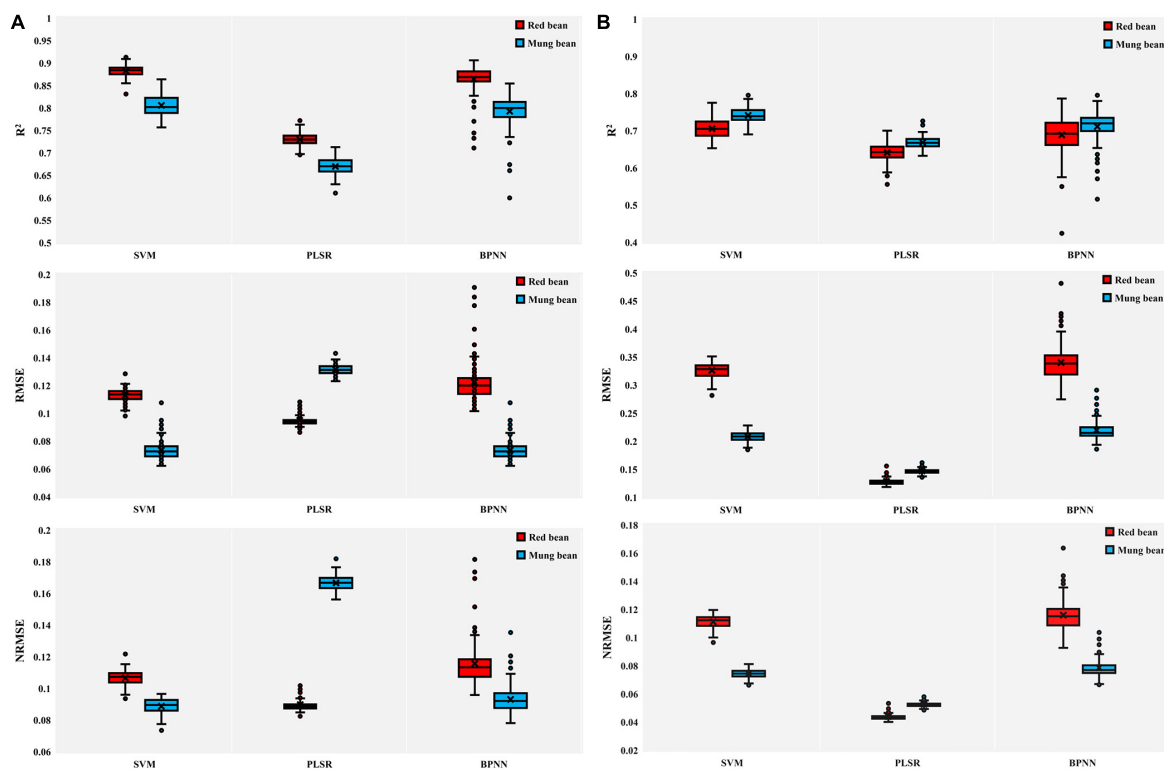


FIGURE 6 | Boxplots for the coefficient of determination (R^2), root-mean-square error (RMSE), and normalized RMSE (NRMSE) of the training results of SVM, PLSR, and BPNN models. **(A)** AGB of red bean and mung bean; **(B)** LAI of red bean and mung bean. The point plots indicate outliers encountered during the phase of the 100 different verifications repetitions and the black multiplication sign indicates the mean value.

To further compare the differences between the performance indicators calculated by using the training dataset and the test dataset in the three methods, we had drawn comparison charts of line segment connection (**Figures 8, 9**). For AGB of red bean, the SVM model showed high performance (**Figure 8A**). In the training set, the SVM model could explain 88.2% of the AGB variation in red bean, and the RMSE and NRMSE values were 0.113 and 0.116, respectively. In the test set, the explanatory degree of the SVM model for AGB variation decreased to 81.1%, RMSE and NRMSE increased to 0.137 and 0.134, respectively. Although the explanatory degree of AGB variation and RMSE value of the SVM models changed greatly, the prediction accuracy and stability of the models were better than that of PLSR and BPNN models. For AGB of mung bean, the SVM model showed better stability (**Figure 8B**). In both the training set and the test set, the SVM model had the highest explanatory degree (80.5 and 75.1%) of AGB variation and the lowest RMSE (0.070 and 0.078) and NRMSE (0.116 and 0.134). The difference between training results and test results was small, which is more stable than other models.

Similarly, the SVM model also showed high performance for LAI estimation of red bean and mung bean (**Figure 9**). In the training set and test set, the explanatory degrees of the SVM model for LAI variation were 70.5 and 64.9%, for RMSE were 0.326 and 0.360, and NRMSE were 0.116 and 0.134, respectively. The explanatory degrees of LAI for mung bean were 74.1 and

70.6%, RMSE were 0.208 and 0.225, NRMSE were 0.11 and 0.123, respectively. In terms of overall performance indicators, the SVM models had better accuracy than the PLSR model and BPNN model and had lower RMSE and NRMSE as well as small test differences.

DISCUSSION

The Spectral Data Obtained From Unmanned Aerial Vehicle Multispectral Image Can Reliably Reflect the Growth Status of Crops Intercropped in Tea Plantation

Monitoring the growth of intercropping crops in tea plantations and guiding the formulation of tea plantation management measures using UAV-based multispectral imagery are very attractive. The results indicated that a single spectral parameter can be used to estimate the AGB and LAI of crops. However, the optimal spectral parameters for estimating growth parameters of red bean and mung bean were different, among which RVII could accurately estimate AGB, while RECI and B.450 were more suitable for estimating LAI. The difference between optimal spectral parameters for estimating AGB and LAI indicated that different VIs showed different sensitivities to AGB and

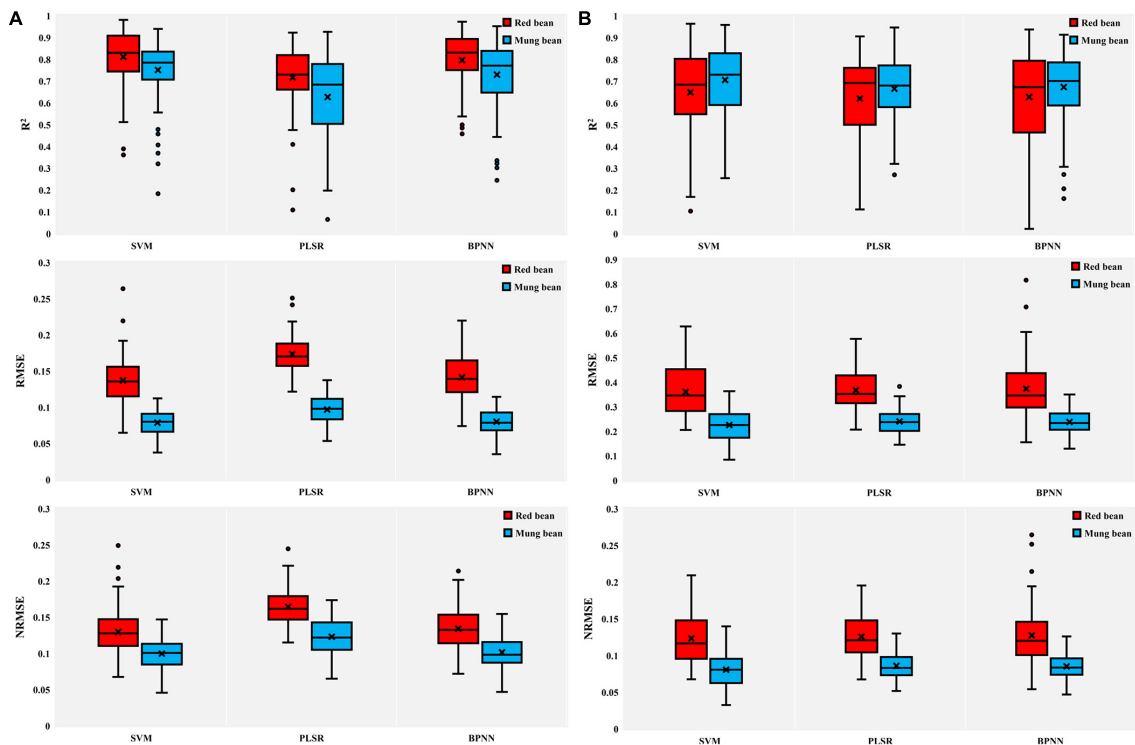


FIGURE 7 | Box plots of coefficient of determination (R^2), RMSE, and NRMSE of test results of SVM, PLSR, and BPNN. **(A)** AGB of red bean and mung bean; **(B)** LAI of red bean and mung bean. The point plots indicate outliers encountered during the phase of the 100 different test repetitions and the black multiplication sign indicates the mean value.

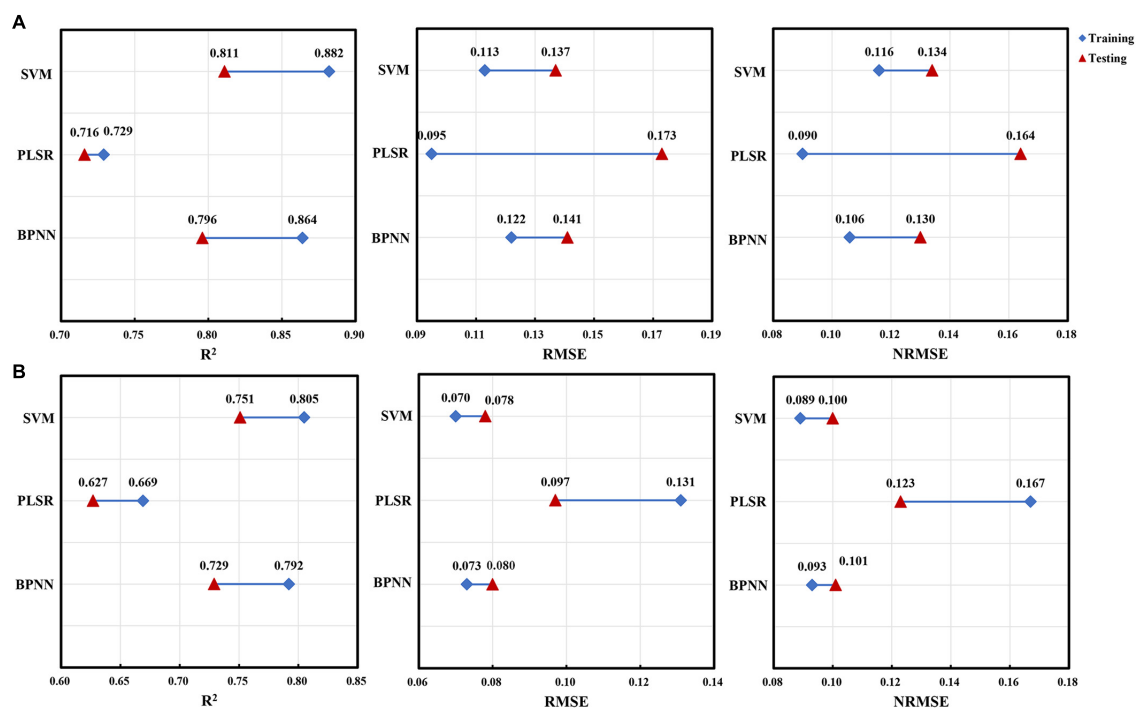


FIGURE 8 | The difference between the performance indicators for AGB estimation of red bean and mung bean using three machine learning methods within training and test datasets. **(A)** Red bean; **(B)** mung bean.

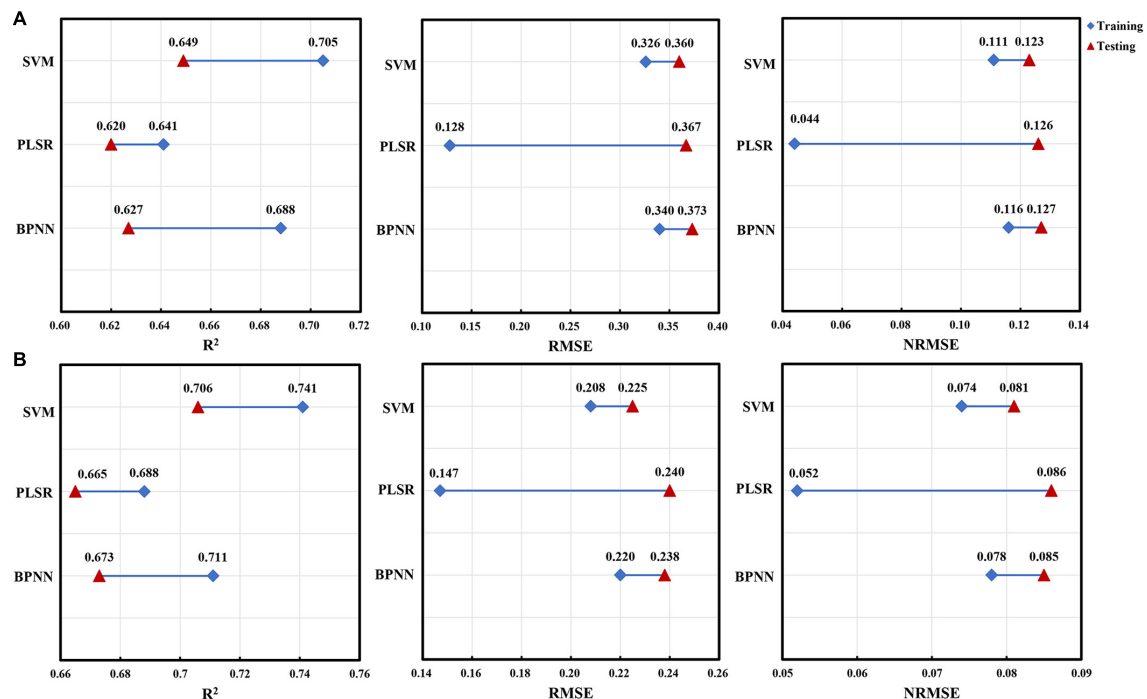


FIGURE 9 | The difference between the performance indicators for LAI estimation of red bean and mung bean using three machine learning methods within training and test datasets. **(A)** Red bean; **(B)** mung bean.

LAI changes in different crops. Similarly, Li W. et al. (2020) found that RVI had a strong correlation with wheat biomass and LAI in the process of using meteorological factors and spectral information to study the disease measurement model of winter wheat, and Liu et al. (2019) also proved that RVI is an important VI for estimating biomass of winter oilseed rape. These conclusions were consistent with our results.

In addition, it was reported that LAI and AGB could exert a certain influence on the spectral reflectance of crop canopy in near infrared (NIR) and visible spectrum (Anthony et al., 2012; Liu et al., 2012; Jin et al., 2015). Qi et al. (2020) found that red and near-infrared bands were sensitive bands for LAI in the process of estimating the LAI of peanuts by using UVA multispectral images. According to the calculation formula of spectral parameters in Table 4, RVI1 is composed of red band and near-infrared band, and the red-edge chlorophyll index is composed of the red-edge band and near-infrared band. Jin et al. (2015) found that enhanced VI (EVI) with the blue band could estimate LAI and biomass more accurately than other spectral parameters when estimating LAI and biomass of wheat using multitemporal optical and radar parameters. In this study, the optimal spectral parameter B.450 used to estimate the LAI of mung bean represents the blue band, which is consistent with this result. In contrast, in the remote sensing monitoring of sorghum growth and development based on UAV system, Li et al. (2018) found that NDVI and RDVI showed a good exponential correlation with biomass; Shafian (2018) also proved that there was a high correlation between NDVI and LAI. Although the calculation of these two spectral parameters has a

red band and near-infrared band, in our study, the correlation between these two spectral parameters and AGB and LAI of mung bean and red bean is not the highest, which may be due to some interference of shadow soil pixels in the process of extracting spectral parameters. Some studies also pointed out that the saturation problem of NDVI would reduce its function of predicting LAI under very high LAI values (Feng et al., 2020). However, the growth period of red bean and mung bean was relatively short and the growth rate was very fast, resulting in higher LAI data values collected later, which further leads to the low correlation between NDVI and the LAI of mung bean and red bean in this study.

Different Machine Learning Algorithms Combined With Spectral Data Can Effectively Estimate the Growth Parameters of Intercropping Crops in Tea Plantation

In addition to single spectral parameters, SMLR, PLSR, SVM, and BPNN algorithms were used to monitor the growth parameters of intercropping crops in tea plantations. The results showed that the SMLR and PLSR models performed significantly better than the LR models, which is consistent with LAI estimation of peanut (Qi et al., 2020) and LAI and AGB estimation of winter wheat (Tao et al., 2020). The reason is that SMLR models and PLSR models use more spectral information related to the variables of interest than single spectral parameter models (Qin et al., 2017; Wei et al., 2018).

In addition, compared with the LR models based on a single spectral parameter or SMLR and PLSR models based on multiple parameters, the SVM and BPNN models can realize non-linear mapping between input and output variables. Therefore, the performance of the SVM and BPNN models in the estimation of growth parameters was better than other models. When the two models were compared, the SVM models still maintained excellent performance. Both the training results and test results of models maintained a high explanatory degree for AGB and LAI variations of red bean and mung bean. Because the SVM method is suitable for small samples, the BPNN method is usually used for a large number of sampled data (Zhu et al., 2019). However, the sample size used to construct models in this study is small, which highlights the superiority of the SVM method. In conclusion, the SVM model can effectively estimate the growth parameters of intercropping crops in tea plantations, and the fitting, stability, and accuracy of this model are better than other models. The superior performance of the SVM method observed in this study is consistent with previous results. For example, Yang X. et al. (2008) found that the SVM method had good learning ability and robustness in estimating the LAI of rice, while Yue et al. (2017) also proved that SVM had strong adaptability in estimating AGB of grassland. However, other studies have shown that PLSR provides better results than SVM in estimating crop growth parameters (Marabel and Alvarez-Taboada, 2013). This difference might depend on the degree of non-linearity in the relationships, the degree of multilinearity and noise in the independent variables, and how accurately the SVM parameters can be tuned (Christoffer et al., 2013). However, our crop growth data precisely fit the advantages of SVM in simulating non-linear relationships, thus highlighting the superiority of the SVM model in estimating growth parameters.

CONCLUSION

Reasonable and reliable estimation of AGB and LAI is of great significance for monitoring crop growth and agricultural site-specific management decision-making. In this study, five machine learning algorithms (LR, SMLR, PLSR, SVM, and BPNN) were used to estimate AGB and LAI of red bean and mung bean in tea plantations based on the extracted multispectral image features collected by UAV remote sensing system. The results showed that the SVM and BPNN models, which can simulate non-linear relationships, were more accurate in estimating AGB and LAI of red bean and mung bean compared with simple LR, SMLR, and PLSR models. In particular, the SVM model provides higher performance in the estimation of AGB and

LAI of red bean and mung bean. Both RMSE and NRMSE of the training set and test set were smaller, and the explanatory degree for AGB and LAI variation was higher. It is proved that the use of UAV multispectral image data combined with machine learning methods can effectively monitor the growth status of crops in tea plantations and provide valuable practical guidelines for site-specific tea plantations and the improvement of their ecological and environmental benefits.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

YS carried out the experiment, collected and processed the data, and wrote the manuscript. ZD and KF raised the hypothesis underlying this study, designed the experiment, helped organize the manuscript structure, and directed the study. YG and YW participated in designing the experiment and reviewed the manuscript. SC and DL participated in designing the experiment and directed the study. All authors contributed to the study and approved the submitted version.

FUNDING

This study was subsidized by the Significant Application Projects of Agriculture Technology Innovation in Shandong Province (SD2019ZZ010), the Technology System of Modern Agricultural Industry in Shandong Province (SDAIT-19-01), and the Special Foundation for Distinguished Taishan Scholar of Shandong Province (No.ts201712057), the Livelihood Project of Qingdao City (19-6-1-64-nsh), the Project of Agricultural Science and Technology Fund in Shandong Province (2019LY002, 2019YQ010, and 2019TSLH0802), and the Project of Improved Agricultural Varieties in Shandong Province (2020LZGC010).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.820585/full#supplementary-material>

REFERENCES

- Afrasiabian, Y., Noory, H., Mokhtari, A., Nikoo, M. R., Pourshakouri, F., and Haghighatmehr, P. (2020). Effects of spatial, temporal, and spectral resolutions on the estimation of wheat and barley leaf area index using multi- and hyper-spectral data (case study: Karaj, Iran). *Precis. Agric.* 22, 660–688. doi: 10.1007/s11119-020-09749-9
- Anthony, N. R., Anatoly, G., Yi, P., Andrés, V., Timothy, A., and Donald, R. (2012). Green leaf area index estimation in maize and soybean: combining vegetation indices to achieve maximal sensitivity. *Agron. J.* 104, 1336–1347. doi: 10.2134/agronj2012.0065
- Bainard, L. D., Klironomos, J. N., and Gordon, A. M. (2011). Arbuscular mycorrhizal fungi in tree-based intercropping systems: a review of their abundance and diversity. *Pedobiologia* 54, 57–61.
- Brooker, R. W., Karley, A. J., Newton, A. C., Pakeman, R. J., and Schb, C. (2015). Facilitation and sustainable agriculture: a mechanistic approach to reconciling crop production and conservation. *Funct. Ecol.* 30, 98–107. doi: 10.1111/1365-2435.12496

- Christoffer, A., Andrew, K. S., Martin, S., Anas, F., and Wouter, V. (2013). Hyperspectral analysis of mangrove foliar chemistry using PLSR and support vector regression. *Int. J. Remote Sens.* 34, 1724–1743. doi: 10.1080/01431161.2012.725958
- Crème, A., Rumpel, C., Gastal, F., Maria, D., and Chabbi, A. (2016). Effects of grasses and a legume grown in monoculture or mixture on soil organic matter and phosphorus forms. *Plant Soil* 402, 117–128. doi: 10.1007/s11104-015-2740-x
- Davies, W. J., Tang, X., Haigang, L., Zhang, D., Shen, J., Rengel, Z., et al. (2016). Increased soil phosphorus availability induced by faba bean root exudation stimulates root growth and phosphorus uptake in neighbouring maize. *New Phytol.* 209, 823–831. doi: 10.1111/nph.13613
- Devia, C. A., Rojas, J. P., Petro, E., Martinez, C., Mondragon, I. F., Patino, D., et al. (2019). High-throughput biomass estimation in rice crops using UAV multispectral imagery. *J. Intell. Robot. Syst.* 96, 573–589. doi: 10.1007/s10846-019-01001-5
- Egesa, A. O., Njagi, S. N., and Muui, C. W. (2016). Effect of facilitative interaction of sorghum-cowpea intercrop on sorghum growth rate and yields. *J. Environ. Agric. Sci.* 9, 50–58.
- Feng, L., Zhang, Z., Ma, Y., Du, Q., Luck, B., Williams, P., et al. (2020). Alfalfa yield prediction using UAV-based hyperspectral imagery and ensemble learning. *Remote Sens.* 12:2028. doi: 10.3390/rs12122028
- Freeman, K. W., Girma, K., Arnall, D. B., Mullen, R. W., Martin, K. L., Teal, R. K., et al. (2007). By-plant prediction of corn forage biomass and nitrogen uptake at various growth stages using remote sensing and plant height. *Agron. J.* 99, 530–536. doi: 10.2134/agronj2006.0135
- Gilliot, J. M., Michelin, J., Hadjard, D., and Houot, S. (2020). An accurate method for predicting spatial variability of maize yield from UAV-based plant height estimation: a tool for monitoring agronomic field experiments. *Precis. Agric.* 22, 1–25. doi: 10.1007/s11119-020-09764-w
- Goel, N. S., and Qin, W. (1994). Influences of canopy architecture on relationships between various vegetation indices and LAI and FPAR: a computer simulation. *Remote Sens. Rev.* 10, 309–347. doi: 10.1080/02757259409532252
- Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., and Strachan, I. B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: modeling and validation in the context of precision agriculture. *Remote Sens. Environ.* 90, 337–352. doi: 10.1016/j.rse.2003.12.013
- Han, L., Yang, G., Dai, H., Xu, B., Yang, H., Feng, H., et al. (2019). Modeling maize above-ground biomass based on machine learning approaches using UAV remote-sensing data. *Plant Methods* 15:10. doi: 10.1186/s13007-019-0394-z
- Harkel, T., Bartholomeus, H., and Kooistra, L. (2019). Biomass and crop height estimation of different crops using UAV-based Lidar. *Remote Sens.* 12:17. doi: 10.3390/rs12010017
- Huang, J., Li, Y., Shi, Y., Wang, L., Zhou, Q., and Huang, X. (2019). Effects of nutrient level and planting density on population relationship in soybean and wheat intercropping populations. *PLoS One* 14:e02258. doi: 10.1371/journal.pone.0225810
- Jiang, Q., Fang, S., Peng, Y., Gong, Y., Zhu, R., Wu, X., et al. (2019). UAV-based biomass estimation for rice-combining spectral, TIN-based structural and meteorological features. *Remote Sens.* 11:890. doi: 10.3390/rs11070890
- Jin, X., Li, Z., Feng, H., Xu, X., and Yang, G. (2017). Newly combined spectral indices to improve estimation of total leaf chlorophyll content in cotton. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 7, 4589–4600. doi: 10.1109/JSTARS.2014.2360069
- Jin, X., Yang, G., Xu, X., Yang, H., Feng, H., Li, Z., et al. (2015). Combined multi-temporal optical and radar parameters for estimating LAI and biomass in winter wheat using HJ and RADARSAR-2 Data. *Remote Sens.* 7, 13251–13272. doi: 10.3390/rs71013251
- Kanke, Y., Tubana, B., Dalen, M., and Harrell, D. (2016). Evaluation of red and red-edge reflectance-based vegetation indices for rice biomass and grain yield prediction models in paddy fields. *Precis. Agric.* 17, 507–530. doi: 10.1007/s11119-016-9433-1
- Li, B., Xu, X., Zhang, L., Han, J., and Jin, L. (2020). Above-ground biomass estimation and yield prediction in potato by using UAV-based RGB and hyperspectral imaging. *ISPRS J. Photogramm. Remote Sens.* 162, 161–172. doi: 10.1016/j.isprsjprs.2020.02.013
- Li, W., Liu, Y., Chen, H., and Zhang, C. (2020). Estimation model of winter wheat disease based on meteorological factors and spectral information. *Food Product. Process. Nutr.* 2, 1–7. doi: 10.1186/s43014-020-0019-y
- Li, J., Shi, Y., Veeranampalayam-Sivakumar, A. N., and Schachtman, D. P. (2018). Elucidating sorghum biomass, nitrogen and chlorophyll contents with spectral and morphological traits derived from unmanned aircraft system. *Front. Plant Sci.* 9:1406. doi: 10.3389/fpls.2018.01406
- Li, J., Zhou, Y., Zhou, B., Tang, H., Chen, Y., Qiao, X., et al. (2019). Habitat management as a safe and effective approach for improving yield and quality of tea (*Camellia sinensis*) leaves. *Sci. Rep.* 9:433. doi: 10.1038/s41598-018-36591-x
- Li, X., Zhang, Y., Luo, J., Jin, X., and Yang, W. (2015). Quantification winter wheat LAI with HJ-1 CCD image 1 features over multiple growing seasons. *Int. J. Appl. Earth Obs. Geoinf.* 44, 104–112. doi: 10.1016/j.jag.2015.08.004
- Liu, B., Asseng, S., Wang, A., Wang, S., Tang, L., Cao, W., et al. (2017). Modelling the effects of post-heading heat stress on biomass growth of winter wheat. *Agric. For. Meteorol.* 247, 476–490. doi: 10.1016/j.agrformet.2017.08.018
- Liu, H., Yuan, Z., Zhang, J., and Shuai, G. (2017). “Highly efficient paddy classification using UAV-based orthorectified image,” in *Proceedings of the IGARSS 2017 - 2017 IEEE International Geoscience and Remote Sensing Symposium IEEE*, Fort Worth, TX.
- Liu, J., Pattey, E., and Jégo, G. (2012). Assessment of vegetation indices for regional crop green LAI estimation from Landsat images over multiple growing seasons. *Remote Sens. Environ.* 123, 347–358. doi: 10.1016/j.rse.2012.04.002
- Liu, Y., Liu, S., Li, J., Guo, X., and Lu, J. (2019). Estimating biomass of winter oilseed rape using vegetation indices and texture metrics derived from UAV multispectral images. *Comput. Electron. Agric.* 166:105026. doi: 10.1016/j.compag.2019.105026
- Lopes, T., Hatt, S., Xu, Q., Chen, J., Yong, L., and Francis, F. (2016). Wheat (*Triticum aestivum* L.)-based intercropping systems for biological pest control. *Pest Manage. Sci.* 72, 2193–2202. doi: 10.1002/ps.4332
- Maimaitijiang, M., Sagan, V., Sidike, P., Maimaitiyming, M., Hartling, S., Peterson, K. T., et al. (2019). Vegetation index weighted canopy volume model (CVM VI) for soybean biomass estimation from unmanned aerial system-based RGB imagery. *ISPRS J. Photogramm. Remote Sens.* 151, 27–41. doi: 10.1016/j.isprsjprs.2019.03.003
- Mao, L., Zhang, L., Zhao, X., Liu, S., van der Wer, W., Zhang, S., et al. (2014). Crop growth, light utilization and yield of relay intercropped cotton as affected by plant density and a plant growth regulator. *Field Crops Res.* 155, 67–76. doi: 10.1016/j.fcr.2013.09.021
- Marabel, M., and Alvarez-Taboada, F. (2013). Spectroscopic determination of aboveground biomass in grasslands using spectral transformations, support vector machine and partial least squares regression. *Sensors* 13, 10027–10051. doi: 10.3390/s130810027
- Metternicht, G. (2003). Vegetation indices derived from high-resolution airborne videography for precision crop management. *Int. J. Remote Sens.* 24, 2855–2877. doi: 10.1080/01431160210163074
- Naito, H., Ogawa, S., Valencia, M. O., Mohri, H., Urano, Y., Hosoi, F., et al. (2017). Estimating rice yield related traits and quantitative trait loci analysis under different nitrogen treatments using a simple tower-based field phenotyping system with modified single-lens reflex cameras. *ISPRS J. Photogramm. Remote Sens.* 125, 50–62. doi: 10.1016/j.isprsjprs.2017.01.010
- Natarajan, M., and Willey, R. W. (1980). Sorghum-pigeonpea intercropping and the effects of plant population density. *J. Agric. Sci.* 95, 59–65. doi: 10.1017/S0021859600029270
- Panigada, C., Rossini, R., Busetto, R., Meroni, R., Fava, R., and Colombo, R. (2010). Chlorophyll concentration mapping with MIVIS data to assess crown discoloration in the Ticino Park oak forest. *Int. J. Remote Sens.* 31, 3307–3332. doi: 10.1080/01431160903193497
- Penuelas, J., Baret, F., and Filella, I. (1995). Semi-empirical indices to assess carotenoids/chlorophyll A ratio from leaf spectral reflectances. *Photosynthetica* 31, 221–230.
- Pinty, B., and Verstraete, M. M. (1992). GEMI : a non-linear index to monitor global vegetation from satellites | Groenekenis. *Vegetatio* 110, 15–20. doi: 10.1007/bf00031911
- Pölonen, I., Saari, H., Kaivosoja, J., Honkavaara, E., and Pesonen, L. (2013). “Hyperspectral imaging based biomass and nitrogen content estimations from light-weight UAV,” in *Proceedings of SPIE - The International Society for Optical Engineering* 8887, Dresden. doi: 10.1117/12.2028624

- Prabhakara, K., McCarty, G. W., and Hively, W. D. (2015). Evaluating the relationship between biomass, percent groundcover and remote sensing indices across six winter cover crop fields in Maryland, United States. *Int. J. Appl. Earth Obs. Geoinf.* 39, 88–102. doi: 10.1016/j.jag.2015.03.002
- Qi, H., Zhu, B., Wu, Z., Liang, Y., and Zhang, L. (2020). Estimation of peanut leaf area index from unmanned aerial vehicle multispectral images. *Sensors* 20:6732. doi: 10.3390/s20236732
- Qin, H., Wang, C., Xi, X., Tian, J., and Zhou, G. (2017). Estimation of coniferous forest aboveground biomass with aggregated airborne small-footprint LiDAR full-waveforms. *Opt. Express* 25, A851–A869. doi: 10.1364/OE.25.00A851
- Rivest, D., Cogliastro, A., Bradley, R. L., and Olivier, A. (2010). Intercropping hybrid poplar with soybean increases soil microbial biomass, mineral N supply and tree growth. *Agrofor. Syst.* 80, 33–40. doi: 10.1007/s10457-010-9342-7
- Romero, M., Luo, Y., Su, B., and Fuentes, S. (2018). Vineyard water status estimation using multispectral imagery from an UAV platform and machine learning algorithms for irrigation scheduling management. *Comput. Electron. Agric.* 147, 109–117. doi: 10.1016/j.compag.2018.02.013
- Rondeaux, G., Steven, M., and Baret, S. F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* 55, 95–107. doi: 10.1016/0034-4257(95)00186-7
- Rouse, J. W., Haas, R. W., Schell, J. A., Deering, D. W., and Harlan, J. C. (1974). *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*. NASA/GSFCT Type III Final Report No. NASA-CR-139243. College Station, TX: Texas A&M University.
- Sanaa, W., Yves, P., Jean, T., Hervé, S., Ezékiel, B., Tasnime, M., et al. (2016). Impact of wheat/faba bean mixed cropping or rotation systems on soil microbial functionalities. *Front. Plant Sci.* 7:1364. doi: 10.3389/fpls.2016.01364
- Sedaghatpour, S., and Janatpour, G. (2012). Study on effect of soybean and tea intercropping on yield and yield components of soybean and tea. *J. Agric. Biol. Sci.* 7, 664–671.
- Shafian, S. (2018). Unmanned aerial systems-based remote sensing for monitoring sorghum growth and development. *PLoS One* 13:e0196605. doi: 10.1371/journal.pone.0196605
- Tao, H., Feng, H., Xu, L., Miao, M., and Fan, L. (2020). Estimation of crop growth parameters using UAV-based hyperspectral remote sensing data. *Sensors* 20:1296. doi: 10.3390/s20051296
- Tatsumi, K., Igarashi, N., and Xiao, M. (2021). Prediction of plant-level tomato biomass and yield using machine learning with unmanned aerial vehicle imagery. *Plant Methods* 17, 1–17. doi: 10.21203/rs.3.rs-344860/v1
- Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* 8, 127–150. doi: 10.1016/0034-4257(79)90013-0
- Wang, C., Sheng, N., Xi, X., Luo, S., and Sun, X. (2017). Estimating the biomass of maize with hyperspectral and LiDAR data. *Remote Sens.* 11, 1–12. doi: 10.3390/rs9010011
- Wang, F. M., Huang, J. F., Tang, Y. L., and Wang, X. Z. (2007). New vegetation index and its application in estimating leaf area index of rice. *Rice Sci.* 14, 195–203. doi: 10.1016/S1672-6308(07)60027-4
- Wei, C., Jian, Z., Chunxiang, C., and Haijing, T. (2018). Shrub biomass estimation in semi-arid sandland ecosystem based on remote sensing technology - ScienceDirect. *Glob. Ecol. Conserv.* 16:e00479. doi: 10.1016/j.gecco.2018.e00479
- Wu, C., Zheng, N., Tang, Q., and Huang, W. (2008). Estimating chlorophyll content from hyperspectral vegetation indices: modeling and validation. *Agric. For. Meteorol.* 148, 1230–1241. doi: 10.1016/j.agrformet.2008.03.005
- Yang, X., Huang, J., Wang, X., and Wang, F. (2008). The estimation model of rice leaf area index using hyperspectral data based on support vector machine. *Spectrosc. Spectr. Anal.* 28, 1837–1841. doi: 10.3964/j.issn.1000-0593.2008.08.034
- Yang, Z., Willis, P., and Mueller, R. (2008). Impact of band-ratio enhanced AWIFS image on crop classification accuracy. *J. Am. Chem. Soc.* 136, 16132–16135. doi: 10.1021/ja5095099
- Yue, J., Feng, H., Jin, X., Yuan, H., Li, Z., Zhou, C., et al. (2018). A comparison of crop parameters estimation using images from UAV-mounted snapshot hyperspectral sensor and high-definition digital camera. *Remote Sens.* 10:1138. doi: 10.3390/rs10071138
- Yue, J., Yang, G., Li, C., Li, Z., Wang, Y., Feng, H., et al. (2017). Estimation of winter wheat above-ground biomass using unmanned aerial vehicle-based snapshot hyperspectral sensor and crop height improved models. *Remote Sens.* 9:708. doi: 10.3390/rs9070708
- Zhou, X., Zheng, H. B., Xu, X. Q., He, J. Y., Ge, X. K., Yao, X., et al. (2017). Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery. *ISPRS J. Photogramm. Remote Sens.* 130, 246–255. doi: 10.1016/j.isprsjprs.2017.05.003
- Zhu, W., Sun, Z., Peng, J., Huang, Y., and Liao, X. (2019). Estimating maize above-ground biomass using 3D point clouds of multi-source unmanned aerial vehicle data at multi-spatial scales. *Remote Sens.* 11:2678. doi: 10.3390/rs1122678

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Shi, Gao, Wang, Luo, Chen, Ding and Fan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



OPEN ACCESS

EDITED BY
Gregorio Egea,
University of Seville, Spain

REVIEWED BY
Thomas Burks,
University of Florida, United States
Jun Lu,
Huazhong Agricultural University,
China

*CORRESPONDENCE
Yu Tang
yutang@gpnu.edu.cn

†These authors have contributed
equally to this work and share first
authorship

SPECIALTY SECTION
This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

RECEIVED 18 June 2022
ACCEPTED 12 July 2022
PUBLISHED 29 July 2022

CITATION
Hou C, Zhang X, Tang Y, Zhuang J,
Tan Z, Huang H, Chen W, Wei S, He Y
and Luo S (2022) Detection
and localization of citrus fruit based on
improved You Only Look Once v5s
and binocular vision in the orchard.
Front. Plant Sci. 13:972445.
doi: 10.3389/fpls.2022.972445

COPYRIGHT
© 2022 Hou, Zhang, Tang, Zhuang,
Tan, Huang, Chen, Wei, He and Luo.
This is an open-access article
distributed under the terms of the
Creative Commons Attribution License
(CC BY). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Detection and localization of citrus fruit based on improved You Only Look Once v5s and binocular vision in the orchard

Chaojun Hou^{1†}, Xiaodi Zhang^{1†}, Yu Tang^{2*}, Jiajun Zhuang¹,
Zhiping Tan², Huasheng Huang², Weilin Chen³, Sheng Wei⁴,
Yong He⁵ and Shaoming Luo³

¹Academy of Contemporary Agriculture Engineering Innovations, Zhongkai University of Agriculture and Engineering, Guangzhou, China, ²Academy of Interdisciplinary Studies, Guangdong Polytechnic Normal University, Guangzhou, China, ³School of Mechatronics Engineering and Automation, Foshan University, Foshan, China, ⁴Engineering Research Center for Intelligent Robotics, Jihua Laboratory, Foshan, China, ⁵College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou, China

Intelligent detection and localization of mature citrus fruits is a critical challenge in developing an automatic harvesting robot. Variable illumination conditions and different occlusion states are some of the essential issues that must be addressed for the accurate detection and localization of citrus in the orchard environment. In this paper, a novel method for the detection and localization of mature citrus using improved You Only Look Once (YOLO) v5s with binocular vision is proposed. First, a new loss function (polarity binary cross-entropy with logit loss) for YOLO v5s is designed to calculate the loss value of class probability and objectness score, so that a large penalty for false and missing detection is applied during the training process. Second, to recover the missing depth information caused by randomly overlapping background participants, Cr-Cb chromatic mapping, the Otsu thresholding algorithm, and morphological processing are successively used to extract the complete shape of the citrus, and the kriging method is applied to obtain the best linear unbiased estimator for the missing depth value. Finally, the citrus spatial position and posture information are obtained according to the camera imaging model and the geometric features of the citrus. The experimental results show that the recall rates of citrus detection under non-uniform illumination conditions, weak illumination, and well illumination are 99.55%, 98.47%, and 98.48%, respectively, approximately 2–9% higher than those of the original YOLO v5s network. The average error of the distance between the citrus fruit and the camera is 3.98 mm, and the average errors of the citrus diameters in the 3D direction are less than 2.75 mm. The average detection time per frame is 78.96 ms. The results indicate that our method can detect and localize citrus fruits in the complex environment of orchards with high accuracy and speed. Our dataset and codes are available at <https://github.com/AshesBen/citrus-detection-localization>.

KEYWORDS

citrus detection, citrus localization, binocular vision, YOLO v5s, loss function

Introduction

Citrus plays an essential role in the fruit industry around the world, with an annual production of approximately 140 million tons (Zheng et al., 2021; Noorizadeh et al., 2022). As the cost of fruit harvesting increases and the availability of skilled labor decreases in China, the traditional manual harvesting method is no longer practical (Gongal et al., 2015; Tang et al., 2021). Presently, fruit harvesting has become increasingly automated for labor-saving and large-scale agriculture (Onishi et al., 2019). The development of an automated citrus picking robot is an inevitable trend for fruit harvesting (Zhuang et al., 2018). In recent work, the development of automatic fruit picking with a robot involves two main tasks: (1) fruit detection and (2) fruit localization via computer vision. The accuracy of fruit detection and fruit localization directly determines the picking efficiency of the robot.

Fruit detection using computer vision has been investigated in numerous recent studies, and most have applied deep learning methods to achieve good performance and robustness (Yang et al., 2020; Chen et al., 2021; Yan et al., 2021). Wan and Goudos (2020) integrated multiclass classification into Faster R-CNN to detect oranges, apples, and mangoes. The improved model achieved a 90.72% mAP. Kang and Chen (2020) proposed a LedNet network with a feature pyramid network and an atrial space pyramid pool for mature apple detection; the recall rate and precision were 0.821 and 0.853, respectively. Chu et al. (2021) improved mask R-CNN by adopting a suppression branch to suppress the generation of nonapple fruit features. However, their method has poor detection performance under backlight conditions. He et al. (2020) developed a deep bounding box regression forest to describe the characteristics of immature citrus on three levels, which is beneficial for differentiating an object from the background. However, the detection speed is slow (0.759 s per frame), making it challenging to apply in real-time applications. For the real-time application of fruit harvesting, the detection speed should be at least 10–15 frames per second (Tu et al., 2020). YOLO series models have been used in various applications for fast detection speed with high accuracy (Jiang et al., 2020; Wang et al., 2021). Xiong et al. (2020) used a YOLO v2 model to detect green mango and reported a recall of 89.0%, a precision of 96.1%, and an average detection time of 0.08 s per frame. Liang et al. (2020) combined YOLO v3 and U-Net to detect litchi fruits and litchi stems at night for picking robots under different illuminations; 96.1% precision and 89.0% recall were achieved. However, the method has not yet been assessed in the daytime. Wang and He (2021) developed an improved YOLO v5 model to detect apple fruitlets using the channel pruning method. However, the network architecture must be manually

adjusted during detection. Notably, the target-background class imbalance is typically the main obstacle encountered in training convolutional neural networks (Buda et al., 2018). To address such class imbalance, Lin et al. (2020) designed a focal loss function to make the network pay more attention to hard samples in training, but the approach cannot push the object further from the background. Rahman et al. (2020) proposed polarity loss to improve focal loss. In the above studies, various deep learning methods have been proposed to detect fruit targets and have achieved good results. However, the detection performance deteriorates in unstructured growing environments with variable illumination conditions. For better accuracy, the disparity between citrus and background under variable illumination conditions and different occlusion states should be incorporated into the network structure.

The purpose of fruit localization is to determine the spatial coordinates of the detected fruit and its location information, such as posture and shape (Huang et al., 2019). Many fruit localization methods require a binocular stereo vision system. The depth map or point cloud image is captured to obtain three-dimensional (3D) localization of fruit. Yang et al. (2020) employed a mask R-CNN to detect citrus objects and branches and matched the color and depth maps to locate fruits and branches. The average error in the diameter of the fruit and the branch was less than 4 mm. However, the distance from the fruit to the camera was not provided in their work. Nguyen et al. (2016) used a Euclidean clustering algorithm to segment a single apple using a point cloud image. The results showed that the errors in the spatial coordinates and the diameter of the fruit were slightly less than 10 mm, but the 3D location information about apples was not the aim of their work. Xu et al. (2018) proposed the PointFusion structure to estimate the 3D object bounding box and its confidence from RGB image and point cloud information. The approach produces good results in the KITTI and SUN-RGBD datasets, with 78% AP. Since the information of the depth map or point cloud is incomplete, fruit localization often requires the use of empirical knowledge (Liu et al., 2017). Wang et al. (2017) adopted Otsu's method and a one-dimensional filter to remove occluded objects (leaves, branches, fruit particles, etc.) and employed ellipse fitting to extract a well-separated mango region. Finally, mango dimensions were calculated using depth information. Ge et al. (2020) developed a shape completion method to reconstruct the point clouds of strawberries; the average error of the center point of strawberries was 5.7 mm. However, the reconstructed error is larger in the case of the neighboring overlapping fruits. Note that an incomplete depth map makes it difficult to recover the missing depth value lost by variable illumination or the fruit region being occluded by randomly overlapping participants, such as neighboring fruits and other background

objects. Therefore, this paper aims to restore the depth map with high accuracy for locating fruits in unstructured orchard environments.

The objective of this work is to develop a novel method for the detection and localization of mature citrus fruits in natural orchards using a binocular camera. The pipelines of the study are to (1) design a new loss function to enhance the detection performance of the YOLO v5s network architecture under variable illumination conditions, (2) extract the fruit region in the RGB image and recover the missing value in the depth map under different occlusion states of citrus fruit, and (3) estimate the 3D localization of citrus fruits using the camera imaging model and the geometric features of citrus fruits. Our method can provide 3D localization information of citrus fruits, such as the diameters of citrus fruits in the 3D direction, the spatial coordinates of citrus fruits, the distance between citrus fruits and the camera, and the 3D bounding box of citrus fruits.

Materials and methods

Datasets

A variety of citrus named "Shantanju" was investigated in the hillside orchard of the Guangzhou Conghua Hualong Fruit and Vegetable Freshness Co. Ltd., located in Guangzhou, China (113°39'2.38"E, 23°33'12.48"N). A total of 4855 groups of images were captured in December 2020 and December 2021 before harvest. Image acquisition was performed using a binocular camera (Model ZED 2, Stereolab's Co. Ltd, USA) with a 1920 × 1080 pixel resolution under sunny and cloudy conditions. The distance between the camera and citrus was set to approximately 30~150 cm. Each group of images contains a left view (RGB image) and a depth map (grayscale image). Note that the right view images were also captured and used only to generate the depth map with the left view images. The depth map is provided with a Z value for every pixel (X, Y) in the left view image. According to the illumination of the citrus surface, images are divided into three groups: non-uniform illumination (non), weak illumination (weak), and well illumination (well). In total, 2913 images were randomly selected as the training dataset (train), 971 images were selected as the validation dataset (validation), and 971 images were selected as the test dataset (test), the number of citrus samples in each group is shown in [Table 1](#).

The hand of the harvesting robot is designed to pick citrus fruits that are in the correct position in front of the camera. In each left view image, the citrus fruits located near the center of the image were manually labeled with bounding boxes using Labelme software. [Figures 1A–C](#) provides examples of citrus images from each illumination group. The bounding boxes of labeled citrus are annotated with red rectangles.

The corresponding depth maps with labeled citrus are shown in [Figures 1D–F](#), where the grayscale of color is based on distance from the camera, i.e., closer objects are darker; further objects are lighter.

Detection and localization of citrus

An overview of our proposed method for citrus detection and localization is presented in [Figure 2](#). The main procedure involves the following steps: Firstly, an improved YOLO v5s is developed to detect citrus in the 2D bounding box. Secondly, Cr-Cb chromatic mapping, Otsu threshold algorithm, and morphology processing are used to extract citrus shape. The missing depth values are recovered by the kriging method. Finally, the 3D localization of citrus fruit is realized by geometric imaging model. Each procedure is described in detail in the following subsections.

Detection of the 2D bounding box of citrus fruit

YOLO (You Only Look Once) is a one-stage detection network that converts object detection into a regression problem using convolutional neural networks ([Wang et al., 2021, 2022](#)). YOLO v5, the latest version of the YOLO model ([Jocher and Stoken, 2021](#)), has a faster detection speed and higher accuracy than the previous version. The release of YOLO v5 consists of four different model sizes: YOLO v5s (smallest), YOLO v5m, YOLOv5l, and YOLO v5x (largest). The network structures of these four models are basically the same, but the numbers of modules and convolution kernels are different. Considering that the application scenario of this paper requires fast detection efficiency, the YOLO v5s model is selected as the basic network, and its structure is shown in [Figure 3A](#). The YOLO v5s network is divided into three parts. The first part is the backbone network, which is responsible for the feature extraction of the target. The second part is PANet, which generates feature pyramids for object scaling. The third part is the head network, which conducts the final detection.

In YOLO v5s, binary cross-entropy with a logit loss function ($Loss_B$) is used to calculate the class probability and objectness score for each sample, as follows:

$$Loss_B(x_i, y_i) = -y_i \log(\sigma(x_i)) - (1 - y_i) \log(1 - \sigma(x_i)), \quad (1)$$

where i is the sample index, x_i is the predicted likelihood, y_i stands for the ground truth, and $\sigma(\cdot)$ is the sigmoid function that maps the prediction x_i to the probability for the ground truth. In object detection tasks, the problem of unbalanced training sets is considerable ([Lin et al., 2020](#)), i.e., the background information in the dataset used for training is overrepresented compared to that of the target class. The sum of $Loss_B$ from the easy samples over the entire images can overwhelm the overall $Loss_B$ from the hard samples. Moreover, the training is inefficient, as most

TABLE 1 Dataset distribution.

	Non		Weak		Well		Total	
	Images	Samples	Images	Samples	Images	Samples	Images	Samples
Train	923	4569	814	2503	1176	4016	2913	11088
Validation	333	1636	255	809	383	1435	971	3880
Test	307	892	269	655	395	1052	971	2599

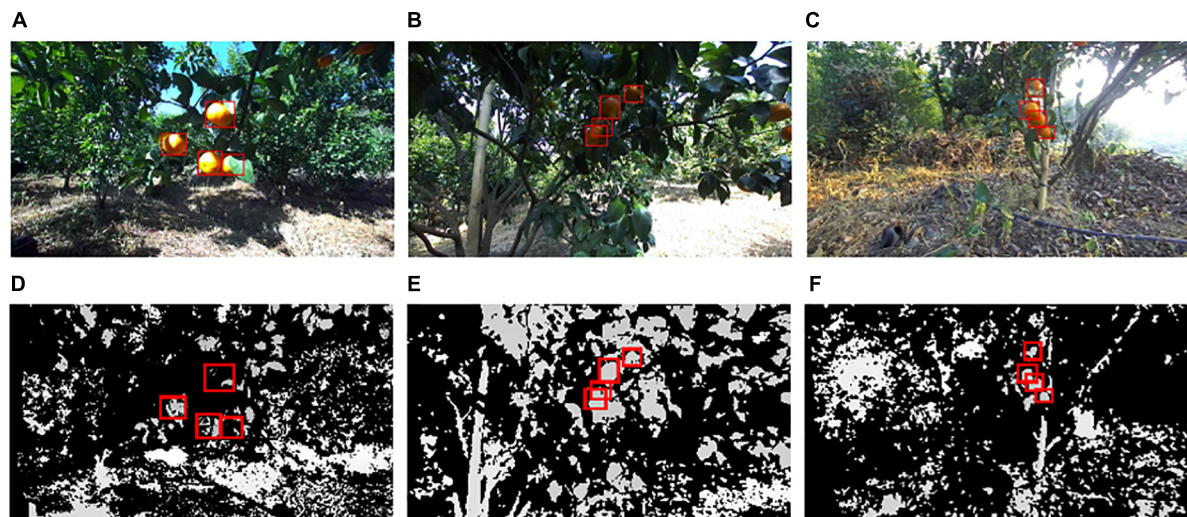


FIGURE 1

Examples of citrus images captured in three illumination conditions: (A) non, (B) weak, (C) well, (D) depth map of (A), (E) depth map of (B), and (F) depth map of (C).

locations are easy samples that do not contribute to learning. Furthermore, in our trial-and-error experiments, the hard negative samples, i.e., the citrus misclassified as background, are difficult to distinguish from the background under weak illumination or obvious occlusion. On the other hand, the hard positive samples, i.e., the background misclassified as a citrus target, exhibit similar characteristics to mature citrus due to the uncontrolled factors in the orchard environment.

To better differentiate citrus from the background under variable illumination conditions and different occlusion states, we design a new loss function, the polarity binary cross-entropy with logit loss ($Loss_{PB}$), to calculate the class probability and objectness score to penalize the hard samples. In particular, a penalty function f_p (Rahman et al., 2020) is developed to represent the disparity between the prediction for citrus and background. $Loss_{PB}$ is defined as follows:

$$\begin{cases} Loss_{PB}(x_i, y_i) = f_p(\sigma(x_i)) Loss_B(x_i, y_i) \\ f_p(z_i) = \frac{2}{1 + \exp(-\gamma(\bar{z}_i - z_i))} \end{cases} \quad (2)$$

where z_i is the probability of sample i being predicted as the true class, such as citrus target or background, $\bar{z}_i = 1 - z_i$ is the probability of sample i being misclassified as the incorrect

class, and γ is a slope parameter of the sigmoid function f_p (Figure 3B). f_p is used to calculate the disparity between the prediction for the true class and false class based on the value of $\bar{z}_i - z_i$. If the citrus target is misclassified as background, the prediction probability \bar{z}_i is greater than, such that a large value of $\bar{z}_i - z_i$ is obtained, and a large penalty will be assigned by f_p . In this case, the penalty value of $Loss_{PB}$ is larger than that of $Loss_B$, which helps to suppress the missed detection of citrus. Similarly, if the background is misclassified as citrus, a large penalty will be assigned by f_p due to the large value of $\bar{z}_i - z_i$, which will improve the false detection of citrus. On the other hand, if a citrus target or the background is predicted with a more reliable probability of z_i , the penalty value applied by f_p will be closer to 0 due to the small value of $\bar{z}_i - z_i$. In such a case, the penalty value of $Loss_{PB}$ is smaller than that of $Loss_B$ and is pushed toward zero. In general, a large penalty is applied to missed detection and false detection of citrus targets. Thus, f_p enforces a large margin to push predictions z_i and \bar{z}_i further apart.

Recall rate (R), precision (P), and F_β -score (F_β) are selected to evaluate the performance of the improved YOLO v5s in the test dataset:

$$R = \frac{TP}{TP + FN}, \quad (3)$$

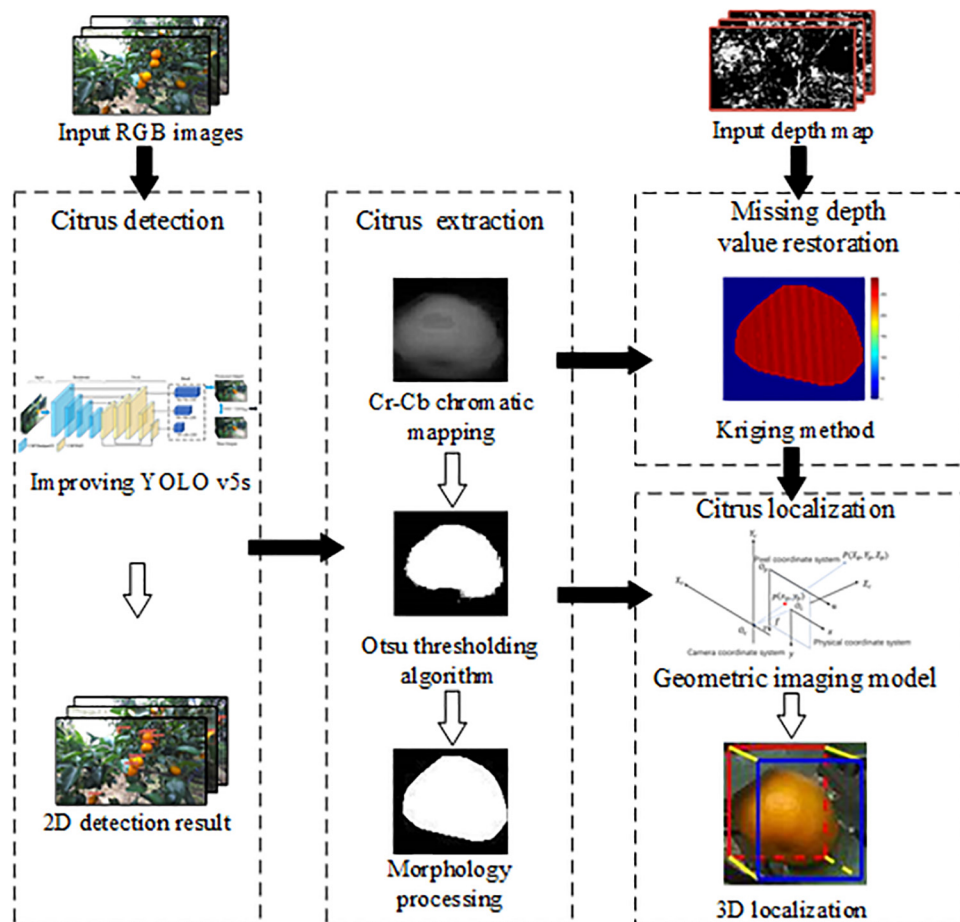


FIGURE 2
Flow diagram of our proposed method.

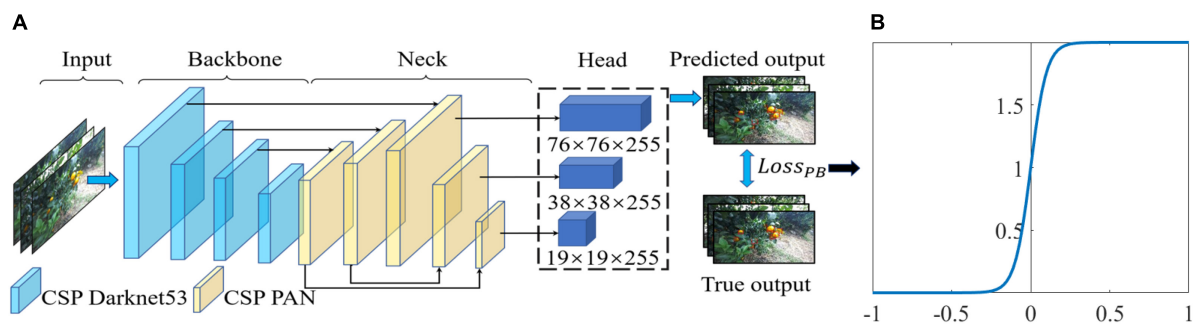


FIGURE 3
Citrus detection model based on You Only Look Once (YOLO) v5s: (A) network structure of improved YOLO v5s and (B) function graph of penalty function f_p .

$$P = \frac{TP}{TP + FP}, \quad (4)$$

$$F_\beta = (1 + \beta^2) \frac{P \times R}{\beta^2 \times P + R}, \quad (5)$$

where FN is the number of false negatives for the false detection of citrus samples, FP is the number of false positives for the missed detection of citrus samples, and TP is the number of true positives for the detected citrus samples. F_β uses a positive real number β to weigh the importance between R and P . In

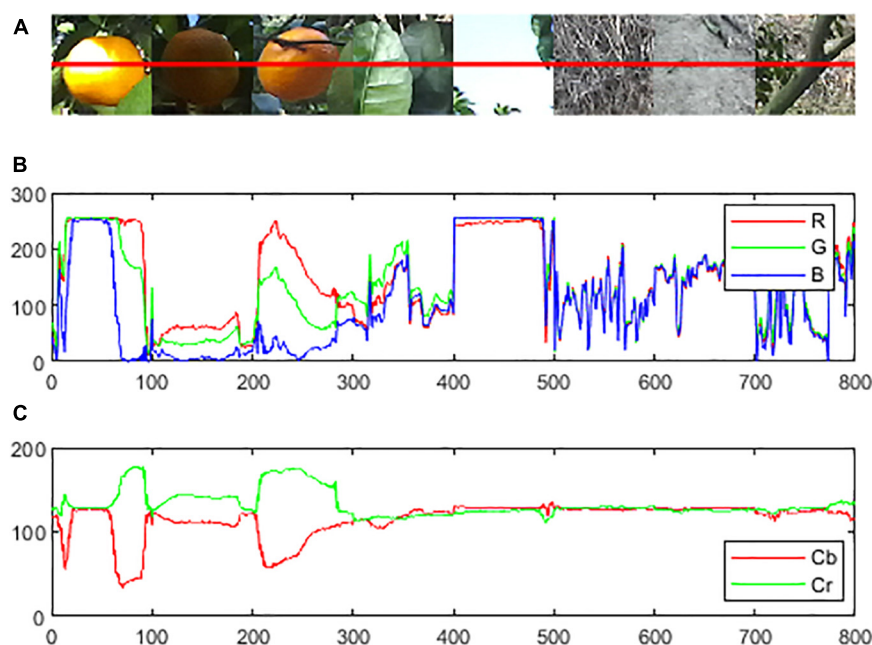


FIGURE 4

Examples of color curves of the citrus and background in different color spaces: (A) original RGB image, (B) color intensity on the line on R, G, and B elements in RGB color space, and (C) color intensity on the line on Cb and Cr elements in YCrCb color space.

this paper, β is set to 1 as F_1 by regarding R and P are equally necessary for our experiment.

Extraction of the citrus fruit region from the 2D bounding box

Image data captured in a natural orchard always contain multiple participants, e.g., grass, soil, lawn, leaves, branches, trunks, and sky. The citrus fruit region is difficult to extract exactly from the 2D bounding box predicted from the improved YOLO v5s. Fortunately, these participants have different color characteristics, so the different targets can be extracted based on their color information. Here, the proper color space is beneficial to robustly extract the citrus fruit region from the background. Zhuang et al. (2018) and Zhuang et al. (2019) adopted improved R-G chromatic mapping to extract fruit regions. In this paper, the input images are converted into the YCbCr color space for better contrast enhancement between the citrus fruit region and background.

As shown in Figure 4A, a horizontal red line was drawn across citrus fruits and the background. The color intensities of the pixels of the line are represented with the R curve (the red element of RGB), the G curve (the green element), and the B curve (the blue element) in Figure 4B. The Cr curve (the Cr element of YCbCr) and the Cb curve (the Cb element) are represented in Figure 4C. The intensity difference between the R curve and G curve is small in both the citrus region and background, and there are no obvious rules exhibited in

the B curve among the citrus fruit regions and backgrounds. However, the intensity difference between the Cr curve and Cb curve values within the citrus region is obviously greater than that of the background. Thus, Cr-Cb chromatic mapping is suitable to enhance the disparity between the citrus region and the background participants.

The Otsu thresholding algorithm is an appropriate method to segment the potential citrus regions from the background, where the best threshold value is selected by maximizing the variance between foreground and background. As shown in Figure 5, the Cr-Cb chromatic mapping has prominent bimodal characteristics in the intensity histogram under variable illumination, where the citrus fruit region contributes to the high value and background contributes to the lower value. Therefore, the best threshold value from Otsu is suitable to segment the citrus fruit region from the background.

The fruit region segmented by Otsu thresholding will not be complete in terms of shape due to the irregular growth situations of citrus fruit that are occluded by adjacent fruits, branches and leaves. To address this problem, the mathematical morphology operations of erosion, dilation, and hole filling are subsequently adopted to fill the gaps between detected regions, remove noise, fill small holes, and smooth the region's boundary. Then, the mathematical morphology operation of convex hull is used to estimate the occluded regions of the fruit from the partially compact region. In this way, the citrus fruit can be almost completely segmented from its corresponding 2D bounding box.

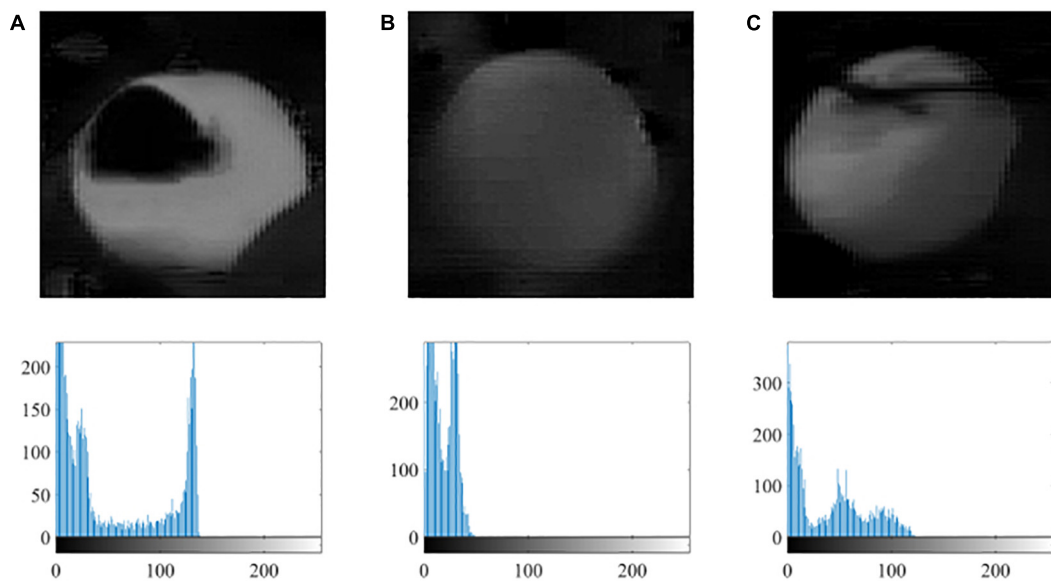


FIGURE 5

Examples of the citrus image after Cr-Cb chromatic mapping and its gray histogram under variable illumination: (A) non, (B) weak, and (C) well.

Recovery of missing depth values

To achieve the 3D localization result of citrus, it is essential to obtain a complete depth map of the whole citrus fruit region; however, an incomplete depth map is always obtained for two main reasons. First, the depth map is sparse in the case of binocular stereo conditions. The depth value is missing and set to zero for pixels where no depth information is sensed by the ZED camera, which may be caused by variable illumination, camera performance limitation, and shooting angle (Liu et al., 2017). Second, the depth values can be missing due to the occluded region estimated from the morphological processing. To restore the complete depth map of the citrus region, the kriging method is adopted to predict the missing depth value by adding the weight of the observed depth value.

Let I_O be the citrus region segmented by Otsu thresholding and I_C be the citrus fruit region extracted via the convex hull operation. We denote by I_{in} the set of pixels whose depth value is missing in I_C , such that the depth value is zero or the pixel is located outside of I_O . Let I_V be the set of pixels whose observed depth value is available in I_O . Therefore, the missing depth value in I_{in} can be obtained as follows:

$$\hat{Z}(s) = \sum_{p \in I_V} \lambda_p(s) Z(p), \quad \forall s \in I_{in}, \quad (6)$$

where $Z(p)$ is the observed depth value at pixel p and $\lambda_p(s)$ is the weight of $Z(p)$, which depends not only on the distance between the depth values but also on the position and overall spatial arrangement of the observed depth value around pixel s . Note that the kriging method is the best linear unbiased estimator to restore the missing depth value using observed depth values

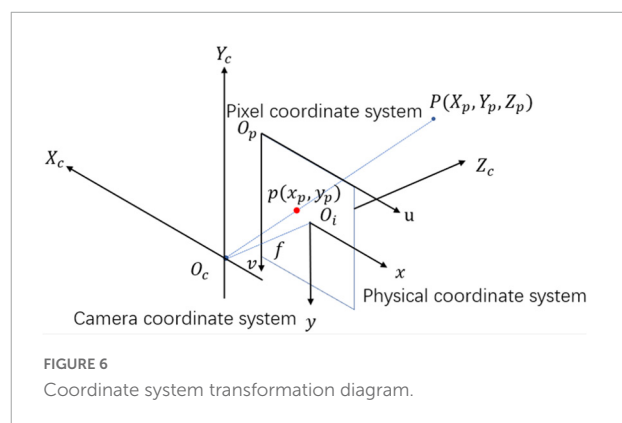


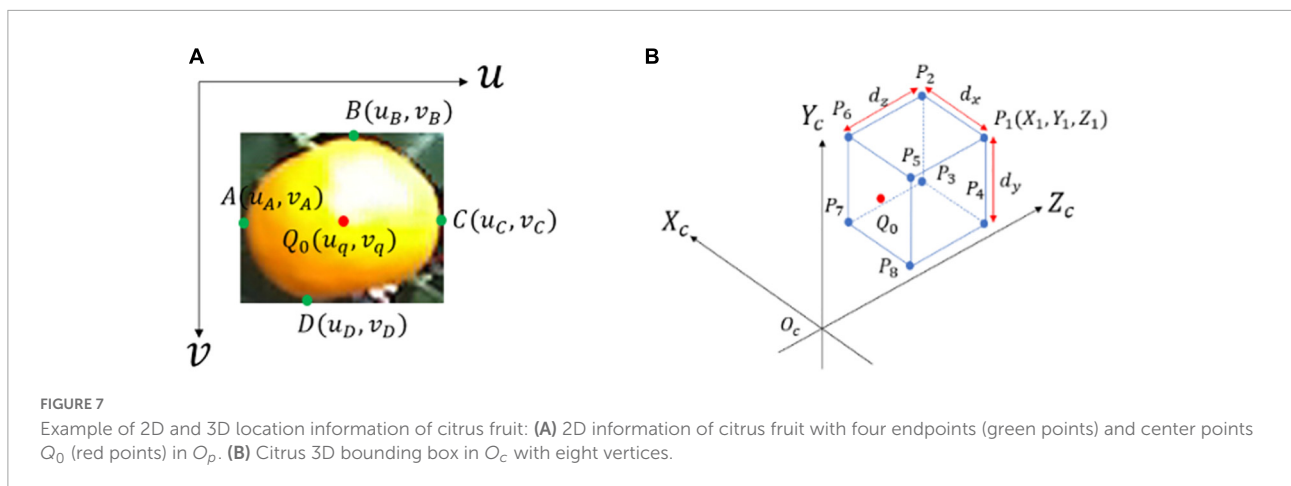
FIGURE 6

Coordinate system transformation diagram.

from the incomplete depth map. Therefore, all the missing depth values in I_C will be restored completely.

3D localization of citrus fruit

The 3D localization of citrus determines the spatial position and posture information, such as citrus diameter in the 3D direction d_x , d_y , and d_z , the spatial coordinates of citrus $Q_0(X_q, Y_q, Z_q)$, the distance between the citrus and camera d , the spatial coordinates of the citrus 3D bounding box P_1, P_2, \dots, P_8 , and the corresponding 2D coordinates of the 3D bounding box in the image plane p_1, p_2, \dots, p_8 . The 3D coordinates of a point in the real world must be precisely mapped to the 2D coordinates of a pixel in the imaging plane. Here, the transformation relation among the camera coordinate system O_c , the physical coordinate system O_i , and the pixel coordinate system O_p should be analyzed.



As illustrated in **Figure 6**, a physical coordinate system O_i is depicted with the origin in the imaging plane (unit: millimeter). The camera coordinate system O_c is created with the optical center as the coordinate origin. Note that the coordinates of the object in the real world are represented relative to O_c , and O_c reaches O_i through perspective projection transformation. Suppose the coordinates of point P in O_c are (X_p, Y_p, Z_p) , and the corresponding coordinates projected onto O_i are (x_p, y_p) . The relationship of point P between O_c and O_i is given by

$$\begin{cases} x_p = f \frac{X_p}{Z_p} \\ y_p = f \frac{Y_p}{Z_p} \end{cases} \quad (7)$$

where f is the camera focal length. As demonstrated in **Figure 6**, a pixel coordinate system O_p is depicted with the origin on the top-left vertex of the image (unit: pixel). The u - and v -axes are parallel to the x - and y -axes of O_i . Let the point (u_p, v_p) in O_p corresponding to the point (x_p, y_p) in O_i . The two coordinate values can be obtained as follows:

$$\begin{cases} u_p = \frac{x_p}{d_u} + u_0 \\ v_p = \frac{y_p}{d_v} + v_0 \end{cases} \quad (8)$$

where (u_0, v_0) represents the translation of the origin of O_i relative to the origin of O_p and d_u and d_v represent the actual size of the pixels in the u -axis and v -axis directions, respectively. According to Eqn. (7) and (8), the transformed relationship between O_p and O_c is given as

$$Z_p \begin{bmatrix} u_p \\ v_p \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_p \\ Y_p \\ Z_p \end{bmatrix}, \quad (9)$$

where $f_x = f/d_u$, $f_y = f/d_v$. Note that f , d_u , d_v , u_0 , and v_0 are the intrinsic camera parameters that can be provided from the factory parameters of the ZED camera, and Z_p is the observed depth value of the depth map.

As shown in **Figure 7A**, let A, B, C, and D be the leftmost, topmost, rightmost, and bottom-most endpoints of the citrus fruit region projected in O_i , respectively, which have coordinates (u_A, v_A) , (u_B, v_B) , (u_C, v_C) , and (u_D, v_D) . Denote (X_A, Y_A, Z_A) , (X_B, Y_B, Z_B) , (X_C, Y_C, Z_C) , and (X_D, Y_D, Z_D) as the corresponding spatial coordinates of points A, B, C, and D in O_c . According to Eqn. (9), the spatial coordinates of A, B, C, and D are given by

$$\begin{bmatrix} X_i \\ Y_i \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \end{bmatrix} \begin{bmatrix} u_i \\ v_i \\ Z_i \end{bmatrix}, \quad (10)$$

where i is A, B, C and D. Let d_x , d_y and d_z be the fruit diameter in the X_c -, Y_c -, and Z_c -axes, respectively. d_x and d_y are obtained according to the spatial coordinates of A, B, C, and D,

$$\begin{cases} d_x = X_C - X_A \\ d_y = Y_D - Y_B \end{cases} \quad (11)$$

In 3D perspective projection, the citrus fruit diameter d_z cannot be obtained directly from the image. Fortunately, the shape of a citrus fruit is similar to an ellipsoid; thus, the magnitudes of d_x , d_y , and d_z will be highly correlated. In this paper, d_z can be estimated by fitting a quadratic polynomial function of d_x and d_y :

$$\hat{d}_z = \beta_0 + \beta_1 d_x^2 + \beta_2 d_y^2 + \beta_3 d_x + \beta_4 d_y, \quad (12)$$

where $\beta_0, \beta_1, \dots, \beta_4$ are the regression coefficients of a polynomial that can be determined using the least-squares method.

Let $Q_0(u_q, v_q)$ be the center point of the citrus 2D bounding box (**Figure 7**), which indeed corresponds to the center of the citrus surface. The spatial coordinates (X_q, Y_q, Z_q) of Q_0 in O_c are obtained using Eqn. (10). Denote by d the Euclidean distance from Q_0 to the origin point, i.e., the distance between the citrus and camera,

$$d = \sqrt{X_q^2 + Y_q^2 + Z_q^2}. \quad (13)$$

The position and posture information for detected targets can usually be determined by the 3D bounding box (Xu et al., 2018). Let P_1, P_2, \dots, P_8 be the eight vertices of the citrus 3D bounding box (Figure 7B), which have coordinates of (X_i, Y_i, Z_i) for $i = 1, 2, \dots, 8$. In particular, (X_i, Y_i, Z_i) can be obtained from the relative geometrical position of P_i to Q_0 , e.g., (X_1, Y_1, Z_1) is inferred as follows:

$$\begin{cases} X_1 = X_q - d_x/2 \\ Y_1 = Y_q + d_y/2 \\ Z_1 = Z_q + d_z \end{cases} \quad (14)$$

To visualize the 3D bounding box of citrus in the image, the corresponding projected 2D coordinates are calculated. Let the eight vertex points p_1, p_2, \dots, p_8 be the corresponding P_1, P_2, \dots, P_8 projected on O_p , which have coordinates (u_i, v_i) , $i = 1, 2, \dots, 8$. They can be deduced by Eqn. (9). Therefore, the 3D localization for each citrus is summarized in Algorithm 1.

Algorithm 1 - Calculation of 3D localization for a citrus fruit.

Input: Citrus fruit region I_C and depth map I_d .
Output: $d_x, d_y, d_z, Q_0(X_q, Y_q, Z_q), d, (u_i, v_i)$ and (X_i, Y_i, Z_i) for $i = 1, 2, \dots, 8$.
S1: According to I_C , 2D coordinates of citrus region extreme points $A(u_A, v_A), B(u_B, v_B), C(u_C, v_C)$, and $D(u_D, v_D)$ are obtained.
S2: The spatial coordinates of $(X_A, Y_A, Z_A), (X_B, Y_B, Z_B), (X_C, Y_C, Z_C)$, and (X_D, Y_D, Z_D) are calculated by Eqn. (10).
S3: Citrus fruit diameter d_x and d_y are calculated by Eqn. (11), and d_z is estimated by Eqn. (12).
S4: According to I_d , the spatial coordinates of citrus $Q_0(X_q, Y_q, Z_q)$ are determined by Eqn. (10).
S5: The distance d between Q_0 and the origin point in O_c is obtained by Eqn. (13).
S6: The spatial coordinates (X_i, Y_i, Z_i) of citrus 3D bounding box are calculated by Eqn. (14).
S7: The 2D coordinates (u_i, v_i) of citrus 3D bounding box are calculated from (X_i, Y_i, Z_i) using Eqn. (10).

Results and discussion

The performance of the proposed method was evaluated on a workstation with an Intel Core i9-9920X processor with 3.50 GHz, 32 GB RAM, and an NVIDIA GeForce RTX 2080 GPU with 8 GB RAM. The operating system is Windows 10, and the software framework is PyTorch 1.8. All the algorithms were developed in Visual Studio Code 1.63 and MATLAB R2020a software.

Performance evaluation of citrus 2D detection

To evaluate the performance of citrus 2D detection using our proposed loss function, ($Loss_{PB}$), on YOLO v5s, three

loss functions, $Loss_B$, focal loss ($Loss_F$) (Lin et al., 2020), and polarity loss ($Loss_P$) (Rahman et al., 2020), were used for comparison. The YOLO v5s models were trained using the training dataset, and the hyperparameters of the model were fine-tuned using the validation dataset. The performance of the final model was evaluated using the test dataset. After several trial-and-error training runs, the learning rate was set to 0.0032, the batch size was set to 32, the IoU threshold was set to 0.5, the training epoch was 200 and γ was set to 20. All the input images were resized to 640×640 pixels. The network weights of YOLO v5s were initialized with the weights of the model pretrained on the COCO image dataset.

The detection results under three illumination conditions on the test dataset are provided in Table 2. With our proposed loss function, $Loss_{PB}$, we achieve the best improvement on the non-uniform illumination than weak illumination and well illumination, compared to $Loss_B$, $Loss_F$, and $Loss_P$. Specifically, under non-uniform illumination, the recall of our loss is 99.55%, which is an average improvement of 9.08% over $Loss_B$, 7.17% over $Loss_F$, and 5.38% over $Loss_P$. The precision of our loss is 95.79%, which is almost the same result as that of the other three loss functions, while the highest precision of 95.93% is obtained by $Loss_F$. The F_1 -score of our loss is 0.98, which is the highest.

Under weak illumination, the precision of our loss is 96.13%, which is 1.33% higher than that of $Loss_B$ and 1.04% higher than that of $Loss_F$ and $Loss_P$. The recall of our loss is 98.47%, and the F_1 -score is 0.97, both of which are better than those of the other loss functions. Under well illumination, the F_1 -score of our loss is 0.98, an average of 3%, 4%, and 2% higher than that of $Loss_B$, $Loss_F$ and $Loss_P$, respectively. The precision and recall of our loss are 96.64% and 98.48%, respectively, which are both the best highest.

Overall, for our loss, the recall is 98.85%, the precision is 96.22%, and the F_1 -score is 0.98, on average, under the three illumination conditions, values that are approximately 2–9% higher than those of $Loss_B$, about 1–6% higher than those of $Loss_F$, and approximately 1–4% higher than those of $Loss_P$. In terms of other metrics, the detection time per image (T) is similar for all loss functions and is consistent with the requirements of the picking robot (Tu et al., 2020).

Figure 8 shows the citrus samples detected by our loss function $Loss_{PB}$ but not $Loss_B$ under different illumination conditions. As listed in Table 2, the recall rate of $Loss_B$ under non-uniform illumination is the lowest at 90.47% than other illumination conditions. On the other hand, the recall rate of $Loss_{PB}$ performed the best at 99.55% over other illumination conditions. The reason may be twofold: (1) As shown in Figure 8, the illumination component is uniform on the surface of a citrus fruit under weak or well illumination conditions. Therefore, the total number of samples is larger under weak and well illumination than under non-uniform

TABLE 2 Detection results of You Only Look Once (YOLO) v5s using different loss functions in the test dataset.

Loss function	Illumination	P (%)	R (%)	F ₁	T (ms)	TP	FP	FN
Our Loss <i>Loss_{PB}</i>	Non	95.79	99.55	0.98	79.31	888	39	4
	Weak	96.13	98.47	0.97	75.49	645	26	10
	Well	96.64	98.48	0.98	81.04	1036	36	16
	Total	96.22	98.85	0.98	78.96	2569	101	30
<i>Loss_B</i>	Non	95.50	90.47	0.93	81.34	807	38	85
	Weak	94.80	91.91	0.93	78.63	602	33	53
	Well	96.07	93.06	0.95	83.16	979	40	73
	Total	95.56	91.88	0.94	81.33	2388	111	211
<i>Loss_F</i>	Non	95.93	92.38	0.94	79.91	824	35	68
	Weak	95.09	91.76	0.93	75.38	601	31	54
	Well	96.25	92.78	0.94	82.59	976	38	76
	Total	95.85	92.38	0.94	79.75	2401	104	198
<i>Loss_P</i>	Non	95.67	94.17	0.95	79.33	840	38	52
	Weak	95.09	94.66	0.95	75.53	620	32	35
	Well	96.06	95.06	0.96	81.73	1000	41	52
	Total	95.68	94.65	0.95	79.25	2460	111	139

The bold values means the best result on each metrics.



FIGURE 8 The missed detection of citrus samples of You Only Look Once (YOLO) v5s but detected by our proposed loss in different illumination on test data: (A) non, (B) weak, and (C) well.

illumination, making the YOLO v5s with *Loss_B* more likely to learn citrus with uniform color features. (2) It is likely that, compared with weak and well illumination, the color features of a citrus fruit under non-uniform will be hard to extract by the Yolo v5s with *Loss_B*, such that the most citrus sample cannot be detected. Using our loss function, the

citrus target under non-uniform illumination will be further pushed from the background. A large penalty is applied to missed detection from the penalty function *f_P* in the training process.

Figure 9 shows the detection results for different loss functions. Specifically, the red bounding box represents



FIGURE 9
Comparison of detection results using different loss functions: (A) Our loss, (B) $Loss_B$, (C) $Loss_F$, and (D) $Loss_P$.

the predicted output by models, the yellow bounding box represents the missed detection, and the blue bounding box represents the false detection. **Figure 9A** indicates that the YOLO v5s model with our loss function achieves the best citrus detection performance under all illumination

conditions, reducing the occurrence of both missed detection and false detection.

There are several examples of missed detection or false detection by other loss functions, as presented in **Figures 9B–D**. With such loss functions, some background objects, such as

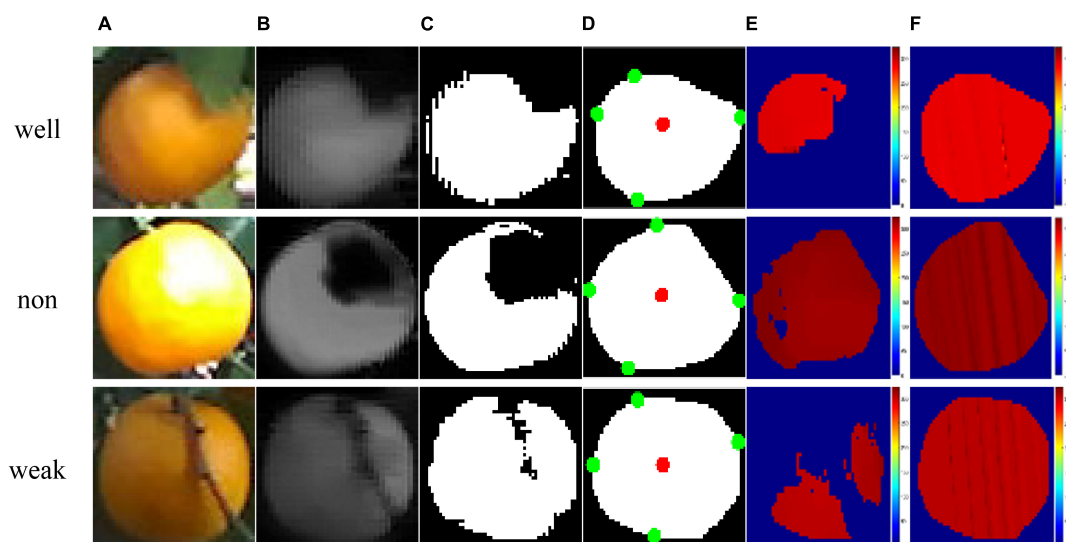


FIGURE 10

Results of samples under different illumination conditions: (A) RGB image, (B) Cr-Cb chromatic mapping, (C) Otsu segmentation, (D) morphological operations, where the red point is the center point and the green point is the maximum and minimum point of the citrus fruit, (E) color map of the original depth map, and (F) color map on depth map restored by the kriging method.

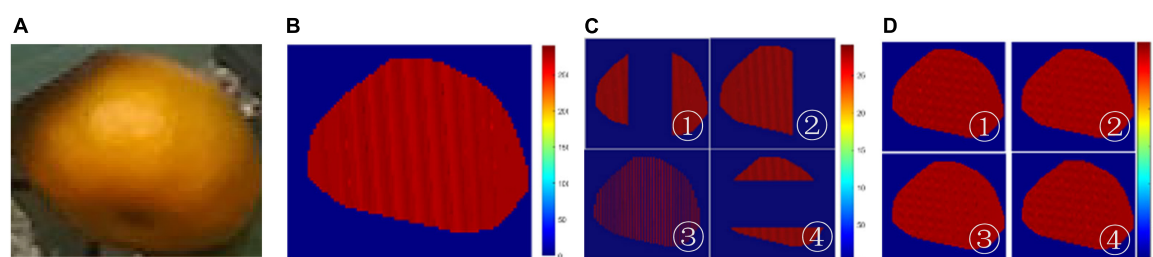


FIGURE 11

Experiment results using the Kriging method: (A) color map of depth values, (B) RGB image of a citrus fruit, (C) color map by setting depth values zero at random pixels, and (D) color map of restoration by kriging.

immature citrus and yellow insect-attracting boards, can lead to false detection of the citrus target. It is likely that immature green citrus has similar texture and shape properties as mature citrus, and the yellow insect-attracting board has similar color characteristics as citrus. On the other hand, citrus that is occluded by leaves, branches, or other background objects may be misclassified as background, i.e., missed detection. For such citrus fruits, it is likely that only a few features can be extracted from the image, resulting in a hard negative sample that is difficult to distinguish from the background.

Figure 9A shows that our proposed loss function achieves the best detection performance. Specifically, the penalty for false detection is enhanced by the penalty function f_p during the training process, and citrus targets are displaced from the background. As a result, the probability of missed detection is reduced substantially, and the detection performance of citrus is thus improved. Note that $Loss_p$ uses a penalty

function similar to f_p and also achieves better performance than that of $Loss_F$ and $Loss_B$. Indeed, it was developed based on $Loss_F$. However, $Loss_F$ cannot push the object further from the background, which may not be an effective improvement on our dataset.

Performance evaluation of citrus region extraction and depth value restoration

Figure 10 illustrates the results of citrus region extraction and depth map restoration under variable illumination conditions. Under the well illumination conditions, the citrus occluded by leaves is shown in the first row of Figure 10A. The results of Cr-Cb chromatic mapping and Otsu thresholding are presented in Figures 10B–C. Image noise, holes, and weakly



FIGURE 12
Examples of 3D bounding boxes for citrus fruits.

connected regions can exist in the binary image obtained via Otsu thresholding. The citrus region is likely blurred, mainly due to the far distance from the camera. The result of morphological processing is shown in **Figure 10D**. The image noise was completely removed, and contour smoothing was achieved, such that the majority of the citrus region occluded by the leaves was filled perfectly. As shown in **Figure 10E**, the depth map of the extracted citrus fruit region after the convex hull operation is incomplete, i.e., the area of missing values covers approximately large than half of the area of the citrus fruit region, which may be caused by camera performance limitations. As shown in **Figure 10F**, the missing depth values are restored using the kriging method, thereby estimating the complete depth values of the fruit region.

The results of citrus fruit extraction and depth map restoration under the non-uniform illumination conditions are presented in the second row of **Figure 10**. The shape of the extracted citrus region is obviously incomplete, which may result from overexposure to the citrus surface. As shown in **Figure 10D**, the incomplete part was restored by morphological

operations. Subsequently, the missing depth values in the citrus region (**Figure 10E**) were recovered, as shown in **Figure 10F**. Similarly, the results under weak illumination conditions are illustrated in the third row of **Figure 10**. The citrus fruit region occluded by branches is extracted almost completely, as shown in **Figure 10D**. Due to the lack of light and other factors, the depth map of the extracted citrus region is sparse, as shown in **Figure 10E**. After using the kriging method, the missing depth values are effectively restored, as shown in **Figure 10F**.

To evaluate the accuracy of the kriging method to recover depth values on the occluded citrus region, an experiment was conducted by simulating the restoration using the incomplete depth map. **Figure 11** shows the results of using the kriging method on an extracted citrus region. **Figure 11A** is the complete depth map of **Figure 11B**. **Figure 11C** shows that the incomplete depth map was generated by setting the corresponding depth values to zero with four schemes. About 50% of the pixels are set as missing values. Specifically, the incomplete depth maps ① and ④ were created by setting the pixels of the central part to zero in the vertical and horizontal directions. The incomplete depth map ② was created by setting the pixels of the right part to zero, and ③ was created by setting the interleaving pixels to zero. As shown in **Figure 11D**, the missing values are recovered using the kriging method, such that the depth map of the fruit region is completely restored.

Compared with the original depth map of **Figure 11B**, the average restoration error of depth map ①, ②, ③, and ④ is 2.29, 2.15, 2.08, and 2.31 mm, respectively, such that the average of the all the restoration errors is 2.21 mm. The minimum error was performed in the depth map ③, indicating that the estimate of missing depth value is recovered with high accuracy when the depth values are only missing randomly in the depth map. On the other hand, the maximum error was performed in the depth map ① and ④, indicating that the restoration error is large when the missing depth values are in the most discontinuous part of the depth map. In total, the mean relative error is 1.36%, indicating that the kriging method effectively restored the depth map with high accuracy.

Performance evaluation of citrus 3D localization

Citrus diameter d_x , d_y , and d_z , coordinates of citrus $Q_0(X_q, Y_q, Z_q)$, the distance between the citrus and camera d , the 3D coordinates of the citrus 3D bounding box (X_i, Y_i, Z_i) , and its 2D coordinates (u_i, v_i) are calculated using **Algorithm 1**. Specifically, to obtain the regression model for d_z , as mentioned in Eqn. (12), a total of 137 citrus samples were collected in the orchard. The diameter d_x , d_y , and d_z of each fruit were measured

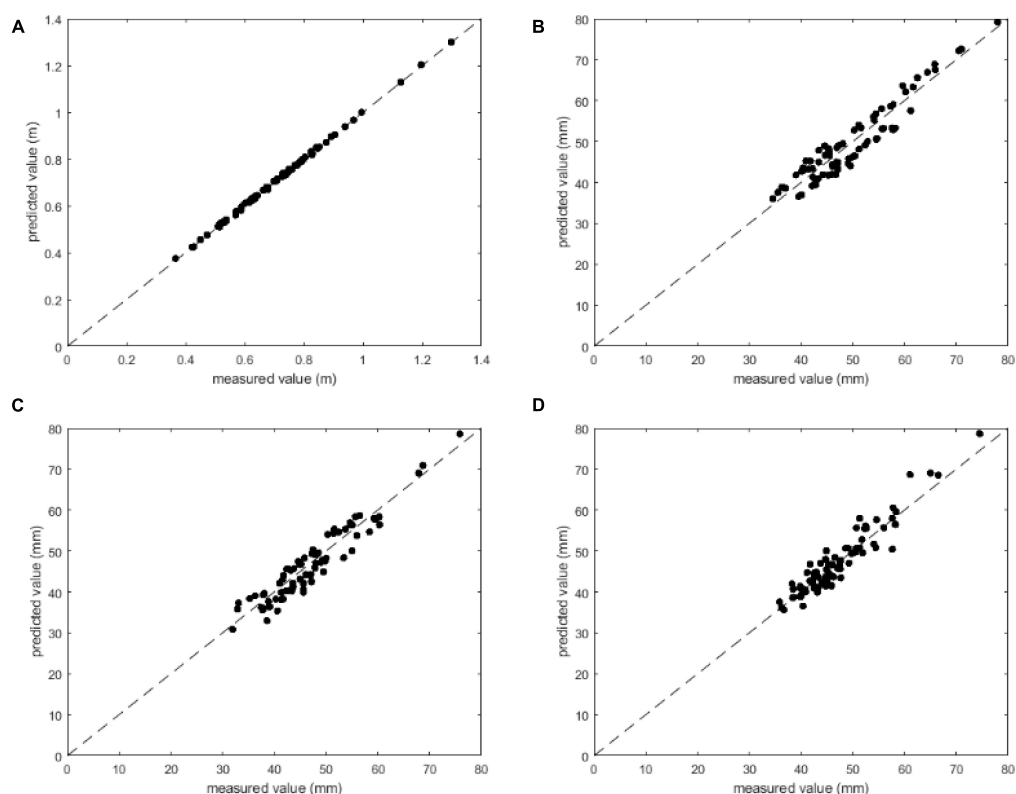


FIGURE 13

Comparison between the measured values and predicted values: (A) d , (B) d_x , (C) d_y , and (D) d_z .

by a Vernier caliper (Pro skit, PD-151). The quadratic polynomial function fitted for d_z is determined as follows:

$$\hat{d}_z = 16.0728 + 0.0028d_x^2 + 0.0018d_y^2 + 0.0264d_x + 0.4133d_y, \quad (15)$$

where the root mean square error (RMSE) is 4.51 mm and the coefficient of determination $R^2 = 0.940$, indicating a good model for estimating d_z .

Figure 12 shows the result of 3D bounding boxes predicted for each citrus fruit. The boxes are drawn by connecting the adjacent vertices (u_i , v_i), for $i = 1, 2, \dots, 8$, with a straight line. The front face of the 3D bounding box was drawn by the blue rectangle, the back face of the 3D bounding box was drawn by the red rectangle, and the side face of the 3D bounding box was drawn by the yellow line. The citrus fruits near the center of the image are correctly detected with the 3D bounding boxes. Moreover, the four edge lines (yellow lines) of the 3D bounding box disappear in the center of the image, which is consistent with the principle of parallel perspective (Cai et al., 2021). Thus, our proposed method achieves accurate localization results.

To evaluate the localization accuracy of citrus 3D localization, 22 images of citrus fruits were considered. The distance between the citrus and the camera d was measured by a laser rangefinder (UNI-T, UT392B), and citrus diameters

d_x , d_y , and d_z were measured with a Vernier caliper (Pro skit, PD-151). A scatter plot of the measured values and the values predicted by our method is presented in Figure 13. Our method obtains good accuracy for predicting d , d_x , d_y , and d_z : the closer the measured values and the predicted value are to the 45-degree line, the higher the accuracy. Figure 13A shows the best prediction and fewer errors between the measured value and predicted values for d , where all the plotted points lie almost on the 45-degree line. Furthermore, Figures 13B–D shows that the predicted values of d_x , d_y , and d_z are generally close to the 45-degree line, indicating that our proposed method is able to achieve accurate localization results.

Overall, the average error of distance d between the citrus and camera is 3.98 mm, which is better than the 15 mm achieved in Wang et al. (2016). The average errors of citrus diameters d_x , d_y , and d_z were 2.75, 2.52, and 2.11 mm, respectively, which is almost the same precision as (Yang et al., 2020) and better than the 10 mm achieved in (Nguyen et al., 2016) and the 4.9 mm achieved in (Wang et al., 2017).

Our method can accurately locate citrus under variable illumination and different occlusion conditions in natural orchards. The distance d can be used to determine the extension length of the robot hand, and the coordinates of citrus $Q_0(X_q, Y_q, Z_q)$ can be used to manipulate the robot hand's series of joints or articulations. The diameter d_x , d_y , d_z and

the 3D bounding box (X_i , Y_i , Z_i) can be used to finetune the posture of grasping structures.

Conclusion

This paper aims to address the problem of the lower detection rate for mature citrus under variable illumination and occlusion conditions. We proposed a novel method to detect and localize citrus fruits in natural orchards using binocular cameras and deep learning. The main conclusions are as follows:

1. A new loss function $Loss_{PB}$ for YOLO v5s is proposed to calculate the loss value for class probability and objectness score, with a penalty function f_p developed to account for the disparity between citrus and background. As a result, the citrus detection performance of our loss function is improved by pushing the citrus further from the background in the training process, even under variable illumination and different occlusion conditions. The recall values of the three groups of illumination conditions were 99.55%, 98.47%, and 98.48%, the precision values were 95.79%, 96.13%, and 96.64%, respectively, and the F_1 -scores were close to 0.98. The average detection time was 78.97 ms per image. Compared with the original YOLO v5s, the performance improvement was 2–9% on average.
2. Based on the detected 2D bounding box for a citrus, the potential fruit region of mature citrus was segmented completely using Cr-Cb chromatic mapping, Otsu thresholding and morphology processing. In particular, the difference in color intensity between citrus targets and background objects is enhanced using Cr-Cb chromatic mapping, which helps to extract the complete shape of citrus fruit using Otsu thresholding and morphology processing.
3. To recover the missing depth values in the citrus region under different occlusion states, the kriging method was applied based on the spatial proximity among neighboring points. The experimental results show that the average error of the restored depth values was 2.02 mm and the relative error was 1.26%, indicating that the method can accurately restore the depth map of citrus fruit.
4. Based on the ellipsoid characteristic of citrus fruit, the 3D localization information of citrus is accurately determined using the camera imaging model and a restored depth map. The experimental results show that the average error of the distance d between the citrus fruit and the camera was 3.98 mm, and the average errors of the citrus diameter d_x , d_y and d_z were 2.75, 2.52, and 2.11 mm, respectively, which is better than the results achieved in other research.

Our method can provide 3D citrus position data under variable illumination and different occlusion conditions in

natural orchards. Future work will focus on few-shot learning and reduce the number of citrus fruits in the training dataset to improve citrus detection and localization.

Data availability statement

The original contributions presented in this study are publicly available. This data can be found here: <https://github.com/AshesBen/citrus-detection-localization>.

Author contributions

All authors contributed to the method and result of the study, dataset generation, model training and testing, analysis of results, and the drafting, revising, and approving of the contents of the manuscript.

Funding

We acknowledged support from the Natural Science Foundation of Guangdong Province, China (Grant Nos. 2020B1515120070 and 2021A1515010824), the Planned Science and Technology Project of Guangdong Province, China (Grant Nos. 2019B020216001, 2019A050510045, and 2021A0505030075), the National Natural Science Foundation of China (Grant No. 32071895), the Key Project of Universities in Guangdong Province, China (Grant No. 2020ZDZX1061), the Innovation Team Project of Universities in Guangdong Province, China (Grant No. 2021KCXTD010), the Planned Science and Technology Project of Guangzhou, China (Grant Nos. 202002020063, 202007040007, and 202103000028), and the Rural Revitalization Strategy Project of Guangdong Province, China (Grant No. 2019KJ138).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Buda, M., Maki, A., and Mazurowski, M. A. (2018). A systematic study of the class imbalance problem in convolutional neural networks. *Neural Netw.* 106, 249–259. doi: 10.1016/j.neunet.2018.07.011
- Cai, W., Liu, D., Ning, X., Wang, C., and Xie, G. (2021). Voxel-based three-view hybrid parallel network for 3D object classification. *Displays* 69:102076. doi: 10.1016/j.displa.2021.102076
- Chen, M., Tang, Y., Zou, X., Huang, Z., Zhou, H., and Chen, S. (2021). 3D global mapping of large-scale unstructured orchard integrating eye-in-hand stereo vision and SLAM. *Comput. Electron. Agric.* 187:106237. doi: 10.1016/j.compag.2021.106237
- Chu, P., Li, Z., Lammers, K., Lu, R., and Liu, X. (2021). Deep learning-based apple detection using a suppression mask R-CNN. *Pattern Recognit. Lett.* 147, 206–211. doi: 10.1016/j.patrec.2021.04.022
- Ge, Y., Xiong, Y., and From, P. J. (2020). Symmetry-based 3D shape completion for fruit localisation for harvesting robots. *Biosyst. Eng.* 197, 188–202. doi: 10.1016/j.biosystemseng.2020.07.003
- Gongal, A., Amatya, S., Karkee, M., Zhang, Q., and Lewis, K. (2015). Sensors and systems for fruit detection and localization: a review. *Comput. Electron. Agric.* 116, 8–19. doi: 10.1016/j.compag.2015.05.021
- He, Z., Xiong, J., Chen, S., Li, Z., Chen, S., Zhong, Z., et al. (2020). A method of green citrus detection based on a deep bounding box regression forest. *Biosyst. Eng.* 193, 206–215. doi: 10.1016/j.biosystemseng.2020.03.001
- Huang, M., Lu, Q., Chen, W., Qiao, J., and Chen, X. (2019). Design, analysis, and testing of a novel compliant underactuated gripper. *Rev. Sci. Instrum.* 90:045122. doi: 10.1063/1.5088439
- Jiang, Z., Zhao, L., Li, S., and Jia, Y. (2020). Real-time object detection method based on improved YOLOv4-tiny. *ArXiv [preprint]*
- Jocher, G., and Stoken, A. (2021). *ultralytics/yolov5: v5.0*.
- Kang, H., and Chen, C. (2020). Fast implementation of real-time fruit detection in apple orchards using deep learning. *Comput. Electron. Agric.* 168:105108. doi: 10.1016/j.compag.2019.105108
- Liang, C., Xiong, J., Zheng, Z., Zhong, Z., Li, Z., Chen, S., et al. (2020). A visual detection method for nighttime litchi fruits and fruiting stems. *Comput. Electron. Agric.* 169:105192. doi: 10.1016/j.compag.2019.105192
- Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P. (2020). Focal loss for dense object detection. *IEEE Trans. Pattern Anal. Mach. Intell.* 42, 318–327. doi: 10.1109/TPAMI.2018.2858826
- Liu, W., Chen, X., Yang, J., and Wu, Q. (2017). Robust color guided depth map restoration. *IEEE Trans. Image Process.* 26, 315–327. doi: 10.1109/TIP.2016.2612826
- Nguyen, T. T., Vandevoorde, K., Wouters, N., Kayacan, E., De Baerdemaeker, J. G., and Saeys, W. (2016). Detection of red and bicoloured apples on tree with an RGB-D camera. *Biosyst. Eng.* 146, 33–44. doi: 10.1016/j.biosystemseng.2016.01.007
- Noorizadeh, S., Golmohammadi, M., Bagheri, A., and Bertaccini, A. (2022). Citrus industry: phytoplasma-associated diseases and related challenges for Asia, America and Africa. *Crop Prot.* 151:105822. doi: 10.1016/j.cropro.2021.105822
- Onishi, Y., Yoshida, T., Kurita, H., Fukao, T., Arihara, H., and Iwai, A. (2019). An automated fruit harvesting robot by using deep learning. *Robomech J.* 6:13. doi: 10.1186/s40648-019-0141-2
- Rahman, S., Khan, S. H., and Barnes, N. (2020). Polarity loss for zero-shot object detection. *ArXiv [preprint]*
- Tang, Y., Dananjayan, S., Hou, C., Guo, Q., Luo, S., and He, Y. (2021). A survey on the 5G network and its impact on agriculture: challenges and opportunities. *Comput. Electron. Agric.* 180:105895. doi: 10.1016/j.compag.2020.105895
- Tu, S., Pang, J., Liu, H., Zhuang, N., Chen, Y., Zheng, C., et al. (2020). Passion fruit detection and counting based on multiple scale faster R-CNN using RGB-D images. *Precis. Agric.* 21, 1072–1091. doi: 10.1007/s11119-020-09709-3
- Wan, S., and Goudos, S. (2020). Faster R-CNN for multi-class fruit detection using a robotic vision system. *Comput. Netw.* 168:107036. doi: 10.1016/j.comnet.2019.107036
- Wang, C., Wang, Y., Liu, S., Lin, G., He, P., Zhang, Z., et al. (2022). Study on pear flowers detection performance of YOLO-PEFL model trained with synthetic target images. *Front. Plant Sci.* 13:911473. doi: 10.3389/fpls.2022.911473
- Wang, C., Zou, X., Tang, Y., Luo, L., and Feng, W. (2016). Localisation of litchi in an unstructured environment using binocular stereo vision. *Biosyst. Eng.* 145, 39–51. doi: 10.1016/j.biosystemseng.2016.02.004
- Wang, D., and He, D. (2021). Channel pruned YOLO V5s-based deep learning approach for rapid and accurate apple fruitlet detection before fruit thinning. *Biosyst. Eng.* 210, 271–281. doi: 10.1016/j.biosystemseng.2021.08.015
- Wang, J., Chen, Y., Gao, M., and Dong, Z. (2021). Improved YOLOv5 network for real-time multi-scale traffic sign detection. *ArXiv [preprint]*
- Wang, Z., Walsh, K. B., and Verma, B. (2017). On-tree mango fruit size estimation using RGB-D Images. *Sensors* 17:2738. doi: 10.3390/s17122738
- Xiong, J., Liu, Z., Chen, S., Liu, B., Zheng, Z., Zhong, Z., et al. (2020). Visual detection of green mangoes by an unmanned aerial vehicle in orchards based on a deep learning method. *Biosyst. Eng.* 194, 261–272. doi: 10.1016/j.biosystemseng.2020.04.006
- Xu, D., Anguelov, D., and Jain, A. (2018). “PointFusion deep sensor fusion for 3D bounding box estimation,” in *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (Piscataway, NJ: IEEE). doi: 10.1109/CVPR.2018.00033
- Yan, B., Fan, P., Lei, X., Liu, Z., and Yang, F. (2021). A real-time apple targets detection method for picking robot based on improved YOLOv5. *Remote Sens.* 13:1619. doi: 10.3390/rs13091619
- Yang, C. H., Xiong, L. Y., Wang, Z., Wang, Y., Shi, G., Kuremot, T., et al. (2020). Integrated detection of citrus fruits and branches using a convolutional neural network. *Comput. Electron. Agric.* 174:105469. doi: 10.1016/j.compag.2020.105469
- Zheng, Z., Xiong, J., Lin, H., Han, Y., Sun, B., Xie, Z., et al. (2021). A method of green citrus detection in natural environments using a deep convolutional neural network. *Front. Plant Sci.* 12:705737. doi: 10.3389/fpls.2021.705737
- Zhuang, J., Hou, C., Tang, Y., He, Y., Guo, Q., Zhong, Z., et al. (2019). Computer vision-based localisation of picking points for automatic litchi harvesting applications towards natural scenarios. *Biosyst. Eng.* 187, 1–20. doi: 10.1016/j.biosystemseng.2019.08.016
- Zhuang, J. J., Luo, S. M., Hou, C. J., Tang, Y., He, Y., and Xue, X. Y. (2018). Detection of orchard citrus fruits using a monocular machine vision-based method for automatic fruit picking applications. *Comput. Electron. Agric.* 152, 64–73. doi: 10.1016/j.compag.2018.07.004



OPEN ACCESS

EDITED BY

Gregorio Egea,
University of Seville, Spain

REVIEWED BY

Jay Kant Pratap Singh Yadav,
Ajay Kumar Garg Engineering College,
India
Xiangnan Li,
Northeast Institute of Geography and
Agroecology (CAS), China

*CORRESPONDENCE

Kun Li
likun0818@wust.edu.cn

SPECIALTY SECTION

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

RECEIVED 13 May 2022

ACCEPTED 11 July 2022

PUBLISHED 05 August 2022

CITATION

Liu G, Hou Z, Liu H, Liu J, Zhao W and
Li K (2022) TomatoDet: Anchor-free
detector for tomato detection.
Front. Plant Sci. 13:942875.
doi: 10.3389/fpls.2022.942875

COPYRIGHT

© 2022 Liu, Hou, Liu, Liu, Zhao and Li.
This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License](#)
(CC BY). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

TomatoDet: Anchor-free detector for tomato detection

Guoxu Liu¹, Zengtian Hou², Hongtao Liu¹, Jun Liu^{3,4},
Wenjie Zhao¹ and Kun Li^{3,4*}

¹Goertek College of Science and Technology Industry, Weifang University, Weifang, China, ²School of Intelligent Manufacturing, Weifang University of Science and Technology, Weifang, China, ³Weifang Key Laboratory of Blockchain on Agricultural Vegetables, Weifang University of Science and Technology, Weifang, China, ⁴School of Computer, Weifang University of Science and Technology, Weifang, China

The accurate and robust detection of fruits in the greenhouse is a critical step of automatic robot harvesting. However, the complicated environmental conditions such as uneven illumination, leaves or branches occlusion, and overlap between fruits make it difficult to develop a robust fruit detection system and hinders the step of commercial application of harvesting robots. In this study, we propose an improved anchor-free detector called TomatoDet to deal with the above challenges. First, an attention mechanism is incorporated into the CenterNet backbone to improve the feature expression ability. Then, a circle representation is introduced to optimize the detector to make it more suitable for our specific detection task. This new representation can not only reduce the degree of freedom for shape fitting, but also simplifies the regression process from detected keypoints. The experimental results showed that the proposed TomatoDet outperformed other state-of-the-art detectors in respect of tomato detection. The F_1 score and average precision of TomatoDet reaches 95.03 and 98.16%. In addition, the proposed detector performs robustly under the condition of illumination variation and occlusion, which shows great promise in tomato detection in the greenhouse.

KEYWORDS

tomato detection, anchor-free, CenterNet, deep learning, harvesting robots

1. Introduction

Tomato harvesting is a labor-intensive work, which needs a lot of human resources. It is also very time consuming and includes much tedious work. However, with the development of urbanization and aging of society, the people in the countryside have decreased a lot, and the labor cost continues to increase, resulting in a big labor shortage in farming work (Yue et al., 2015). On the other side, intelligent agriculture is developing fast in the past decades, which is an ideal substitute of human resources for farming work. Among the various technologies applied in the agriculture, the fruit harvesting robot is one of the prominent artificial intelligent techniques. It has huge potential efficiency in fruit harvesting, which can bring high profit as well as liberating the labor force. Thus, it is of great value and significance to develop harvesting robots.

A harvesting robot usually consists of two components—a vision system and an eye-hand coordination system (Zhao et al., 2016a). The vision system plays a key role in

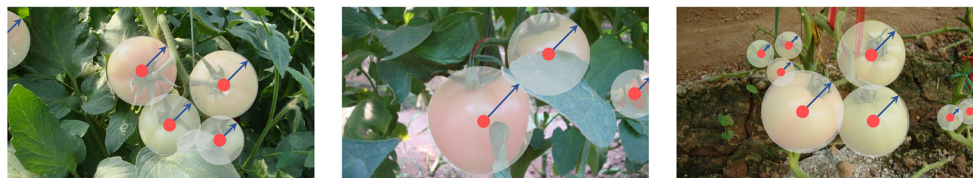


FIGURE 1

A tomato is modeled as a center point of its bounding circle. The radius of the bounding circle can be inferred from the keypoint at the center.

the whole system, since the first critical step for the harvesting robot is to detect fruits autonomously. This step determines the detection and subsequent picking accuracy of harvesting robots. Thus, it is very crucial to develop a robust fruit detection algorithm of the vision system. However, at present, no harvesting robot has been commercialized successfully due to either low detection accuracy or low detection speed. Many factors have hindered the pace of harvesting robot development such as uneven illumination, occlusion, overlap, and some other unpredictable factors (Gongal et al., 2015).

To deal with the above challenges, many researchers have studied fruit detection over the past years. In the early years, some researchers used threshold discriminant methods for fruit detection based on color, shape, texture, or fusion of them (Linker et al., 2012; Kelman and Linker, 2014; Wei et al., 2014), and achieved reasonable detection results. Bulanon et al. (2002) used an optimal threshold extracted from the intensity histogram of a red-color-difference enhanced image for apple recognition. The results showed that the success rate exceeds 88%. This method is restricted to ripe apples which present different color to the background. Okamoto and Lee (2009) employed hyperspectral imaging for detection of green citrus. The method is separated into pixel-wise segmentation process using pixel discrimination functions and fruit recognition process with thresholds selected by trial and error. This method greatly relies on the selection of several optimal thresholds, and thus is lack of robustness when the fruit environment changes. Inspired by the eigenface concept, Kurtulmus et al. (2011) proposed a novel eigenfruit feature for green citrus detection, combined with color and circular gabor texture. Although intrinsic texture features are used other than only color features, the method still confuses some fruits with background and does nothing with severe occluded fruits. Zhao C. et al. (2016) developed a cascaded pixel segmentation method for immature citrus detection in natural environment. Three color feature maps and a block matching method are adopted to identify potential fruit pixels. Finally, an SVM classifier is used to remove false detections. Nevertheless, with only color feature for segmentation in the early stage, many fruits are missed by the method due to similarity between green fruits and background. Zhao et al. (2016b) proposed a multi color feature fusion method

based on wavelet transformation for mature tomato recognition. The detection accuracy reaches 93%. However, since only color features are employed, the method is inferred to be sensitive to illumination variation. These methods greatly rely on the selection of suitable thresholds, making them sensitive to the changes in the form of fruit presentation, such as illumination variation and occlusion.

With the development of machine learning, many researchers tried to apply them to fruit detection, such as adaboost, support vector machine (SVM) or other statistical classifiers (Kurtulmus et al., 2014; Lv et al., 2014; Yamamoto et al., 2014), and get better results than the threshold discriminant methods. Zhao et al. (2016c) used an adaboost classifier associated with haar features for tomato detection. An average pixel value feature is adopted for the removal of false detections. More than 96% of tomatoes are detected in their study. Li et al. (2017) proposed to use an SVM trained on histogram-based features for green and ripe tomato recognition. Prior to detection, the fast normalized cross correlation method is used to extract the potential tomato regions. Finally, the circular hough transform and color analysis are combined to obtain tomato positions. Behroozi-Khazaei and Maleki (2017) proposed to use an artificial neural network optimized by genetic algorithm for grape cluster detection. Also, the genetic algorithm is adopted for color feature selection, which subsequently serves as input to the network. A Bi-Layer schema was proposed for automatic detection of ripening tomatoes by Wu et al. (2019). In their method, a weighted relevance vector machine is used for tomato recognition based on six color-related features and five textural features. A detection rate of 94.90% is reported in the results. Liu et al. (2019) developed a coarse-to-fine method for ripe tomato detection in the greenhouse. First, a naïve bayes classifier is used to identify potential tomato area, on which an SVM classifier combined with histogram of oriented gradients is applied to recognize tomatoes. At last, a color analysis method is proposed to remove false detection. The machine learning methods usually achieve better performance than threshold discriminant methods. However, the low-level abstraction capabilities of hand-crafted features make it difficult to adapt these methods to complicated environmental change.

The emergence of deep learning methods especially convolutional neural networks provides a new paradigm for computer vision tasks, including fruit detection tasks (Sa et al., 2016; Tian et al., 2019; Zheng et al., 2021). These methods can learn feature representations directly from the data and can be trained end-to-end. Nevertheless, the detection accuracy and robustness still need to be improved to enable real commercial applications under complicated conditions as discussed above.

To address the above problems, this study proposes an effective anchor-free detector called TomatoDet for tomato detection. The proposed model represents a tomato by the center point of its bounding circle, as shown in Figure 1. First, to improve the expression ability of the backbone network, an attention mechanism is introduced to guide the network to pay more attention to the region of interest (ROI), especially small tomatoes. Second, a bounding circle is adopted for tomato localization instead of the traditional bounding box, which is commonly used for general object localization.

Our main contribution is three-fold as follows:

1. The Convolutional Block Attention Module is introduced into the backbone network of CenterNet (Zhou et al., 2019) called Attentive-DLA34 to boost the representation power.
2. A circle representation for tomato detection is adopted to adapt the traditional detection methods to our specific detection task. The new circle representation not only reduces the degree of freedom for shape fitting, but also simplifies the regression process from detected keypoints.
3. Extensive experiments are conducted on tomato datasets. We show that the proposed TomatoDet achieves better performance in terms of both accuracy and robustness, compared to the original CenterNet and other state-of-the-art object detectors.

2. Related work

In recent years, deep learning methods have shown continuous performance improvements on fruit detection. A “MangoYOLO” detector was proposed for fruit detection and fruit load estimation by Koirala et al. (2019). This model combines the advantages of YOLOv2 (Redmon and Farhadi, 2017) and YOLOv3 (Redmon and Farhadi, 2018), which has both high detection speed and detection accuracy. It outperforms other methods such as Faster R-CNN (Ren et al., 2015), YOLOv2 (Redmon and Farhadi, 2017), YOLOv3 (Redmon and Farhadi, 2018), and SSD (Liu et al., 2016), on their Mango dataset. Bresilla et al. (2019) improved YOLO (Redmon et al., 2016) model for apples and pears detection. First, the grid-scale is scaled up twice to fit the size of the fruits. Second, the model is pruned to improve the detection speed while not degrading the accuracy. Afonso et al. (2020) applied Mask R-CNN to the tomato dataset for detection. Several neural networks are used as backbone for feature extraction.

The best F₁ score reaches over 94% in their report. Liu G. et al. (2020) proposed a YOLO-Tomato for tomato detection based on YOLOv3 (Redmon and Farhadi, 2018). A dense architecture is incorporated to the backbone to facilitate feature reuse, and a circular bounding box is adopted to optimize the non-maximum suppression process. The model achieves a competing performance compared to state-of-the-art detection methods. Zheng et al. (2021) improved YOLOv4 (Bochkovskiy et al., 2020) for green citrus detection. First, the backbone network is trimmed to reduce detection time. Then, a novel Bi-PANet is proposed to fuse features from different layers. With the modifications, the detection accuracy is reported to be 86% on their dataset. Zhang et al. (2021) developed an edge-device oriented lightweight model for fruit detection. The structure of the original CSPNet is lightened to boost detection speed, and a deep-shallow feature fusion model is proposed to enhance the expression ability of the network. Tested on three types of edge devices, the average detection precision reaches 93, 84.7, and 85% for oranges, tomatoes, and apples, respectively. Wei et al. (2022) proposed a green fruit detection model based on D2Det. By incorporating MobileNetV2, feature pyramid networks and region proposal network structure into the original model, the detection accuracy of green fruits in orchard environments was greatly improved. Chen et al. (2022) improved YOLOv4 for the detection of citrus by incorporating an attention mechanism and a depthwise separable convolution module. In addition, a pruning algorithm was applied to remove the influence of irrelevant latent factors of the data.

Although exciting results are achieved by the above methods, there is still much room for optimization of the networks to improve detection performance. Moreover, the above methods are all anchor-based methods, which commonly perform nearly exhaustive anchor classification over the image and have many hyperparameters for anchor design, reducing the detection efficiency.

3. Materials and methods

3.1. Image acquisition

The images used in this study are captured using a digital camera (Sony DSC-W170, Tokyo, Japan) with a resolution of $3,648 \times 2,056$ pixels in a Tomato Production Base, which is located in Shouguang City, Shandong Province, China. The datasets are collected under various environment conditions including sunlight, shading, occlusion, and overlap, etc. Some examples captured under different conditions are shown in Figure 2.

To verify the proposed method, the datasets are split into two subsets—a training set and a test set. The training set contains 725 images, and 241 images are included in the test set. Totally, 966 images are used in this study. For data labeling, a tool

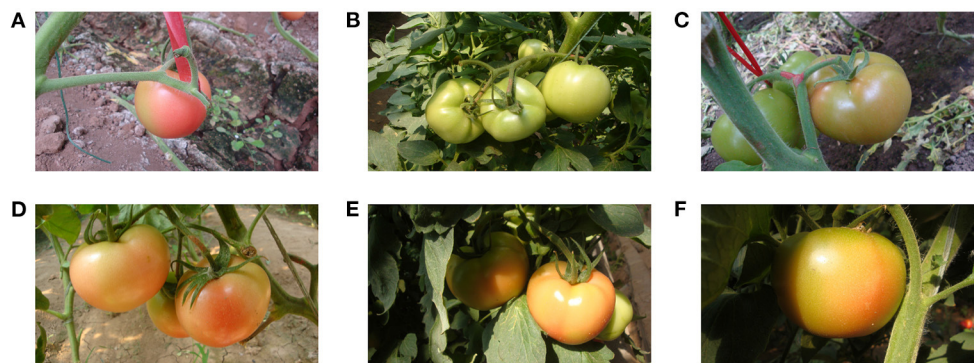


FIGURE 2
Some tomato samples with different growing circumstances: (A) a single tomato, (B) a cluster of tomatoes, (C) occlusion case, (D) overlap case, (E) shading case, and (F) sunlight case.

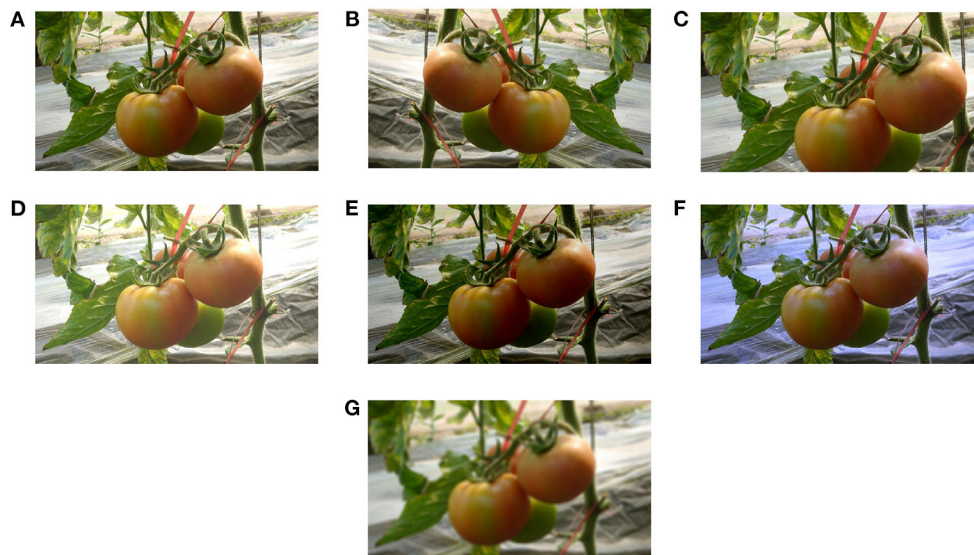


FIGURE 3
Data augmentation of tomato images: (A) original image, (B) horizontal flip, (C) scaling and cropping, (D) high brightness, (E) low brightness, (F) color balancing, and (G) blur processing.

called Label-Tomato has been developed to annotate images with proposed bounding circles based on Python. The output format of Label-Tomato is txt files, which include the numbers and locations of tomatoes for each image.

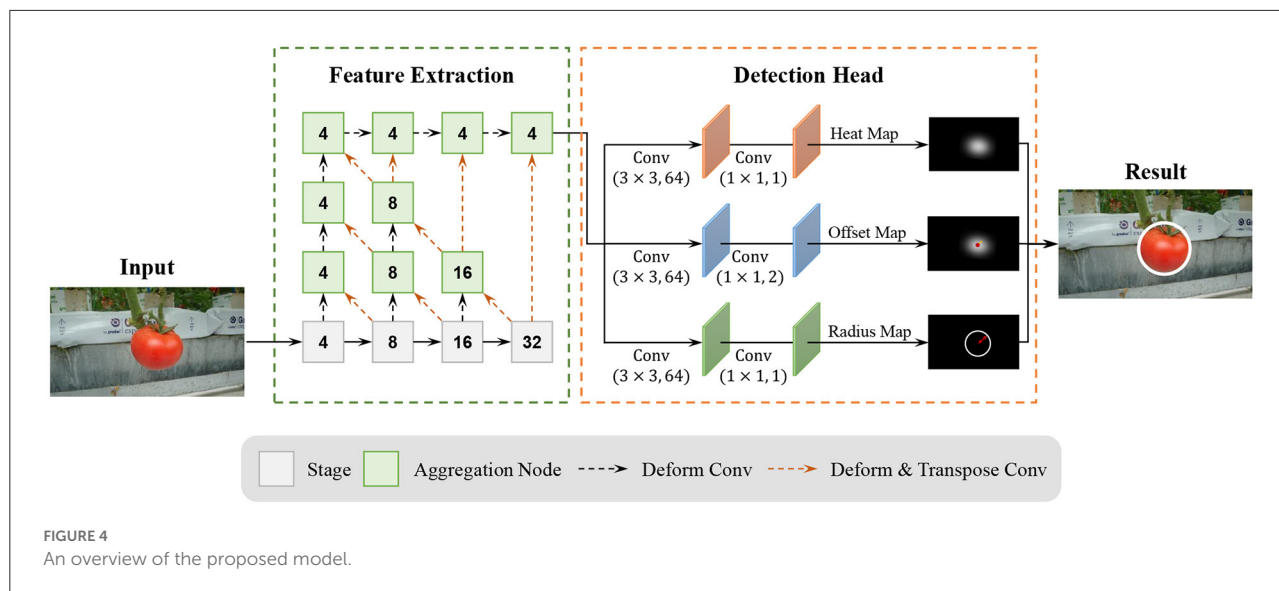
3.2. Data augmentation

To avoid over-fitting of the model in the training process, the data augmentation is used in this study to simulate real-life interference and enhance the richness of the collected datasets. Several image processing technologies are adopted for augmentation - horizontal flip, scaling and cropping, brightness

transformation, color balancing and image blurring, as shown in Figure 3. For the brightness transformation, we use a factor falling in the range [0.6, 1.4] to change the intensity of the pixels in the image randomly. This process can simulate different weather factors on the image intensity. For the scaling and cropping operation, we follow the same process as in Liu G. et al. (2020). To eliminate the effect of lighting on color rendering, we adopt the gray world algorithm (Lam, 2005) for color balancing. Finally, we randomly blur the augmented images by flip, scaling and cropping, brightness transformation, and color balancing to simulate indistinct images caused by camera movement. After data augmentation, the whole number of resultant images is shown in Table 1.

TABLE 1 The number of training images after data augmentation.

	Original	Flip	Scaling and cropping	Brightness	Color	Blur	Total
No. of tomato images	725	725	725	1,450	725	725	5,075



3.3. Overview of tomatoDet

Our tomato detection model, called TomatoDet, pools several concepts from the past work with our novel idea to improve the detection performance. An overview of the proposed model is shown in Figure 4. The proposed TomatoDet is based on CenterNet and consists of two modules. The first module is used for feature extraction. It adopts Deep Layer Aggregation-34 (DLA34) (Yu et al., 2018) as the backbone and incorporates Convolutional Block Attention Module (CBAM) (Woo et al., 2018) to improve the feature expression ability and guide the network to focus on small-scale tomato targets. The second module is the detection head. The architecture of the detection head is like that of CenterNet, except that we use a radius head instead of the height and width head for bounding circle prediction. More details are presented in Sections 3.4 and 3.5.

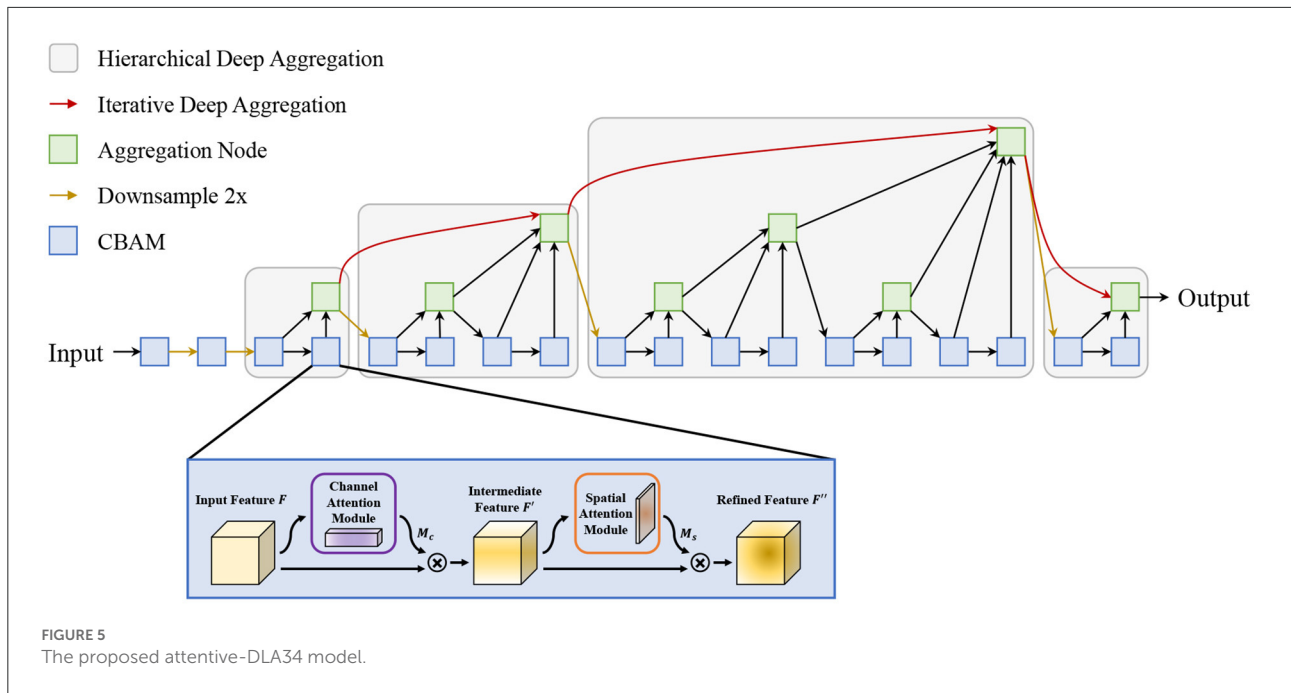
3.4. The proposed attentive-DLA34 backbone

In this study, an attentive Deep Layer Aggregation network (Attentive-DLA34) is proposed as the base backbone for feature extraction. The DLA is inspired by dense connection and feature pyramid and has two main structures: the iterative

deep aggregation (IDA) and the hierarchical deep aggregation (HDA). The IDA is mainly used for feature fusion across resolutions and scales while the HDA focuses on semantic fusion, i.e., aggregating features from different channels and depths in a tree-based structure. Based on these two structures, the DLA could make better use of spatial and semantic information for recognition and localization. However, the complicated conditions make it challenging to detect tomatoes in a natural environment, not to mention the existence of a large number of small tomatoes. To mitigate this problem, we introduce an attention mechanism—Convolutional Block Attention Module (CBAM)—into the backbone network to guide it to pay more attention to the region of interest (ROI). The architecture of the proposed Attentive-DLA34 model is shown in Figure 5.

As shown in Figure 5, we replace the original layers in each stage with CBAM to focus its attention on tomato areas. For CBAM, it is divided into a channel attention module and a spatial attention module in a sequential manner. First, the channel attention module takes the input and infers a 1D channel attention map. Then, the multiplication output of the input and the attention map is inputted to the spatial attention module to get the final output feature map in the same way. The detailed operation can be depicted in Equations (1) and (2):

$$F' = M_c(F) \otimes F \quad (1)$$



$$F'' = M_s(F') \otimes F' \quad (2)$$

where \otimes indicates element-wise multiplication, $F \in R^{C \times H \times W}$ is the input feature map, $M_C \in R^{C \times 1 \times 1}$ denotes the generated channel attention map, and $M_s \in R^{1 \times H \times W}$ denotes the generated spatial attention map. F'' is the final output by CBAM.

3.5. Circle representation

For general object detection, a bounding box is usually adopted for object localization. However, this type of detection representation is not optimal for specific objects which have a particular shape. In this study, since our detection target is tomato, which is roughly circular, it is better to use bounding circles instead of bounding boxes for localization. It has three folds of advantages. Firstly, compared with bounding boxes, bounding circles could better match the shape of tomatoes. Secondly, the representation of a circle is simpler than that of a box, which makes it easier for the network to learn. Lastly, the circle is invariant to rotation.

3.5.1. From point to bounding circle

For an input image $I \in R^{W \times H \times 3}$ with width W and height H , the target is to produce a keypoint heatmap $\hat{Y} \in [0, 1]^{\frac{W}{K} \times \frac{H}{K} \times C}$, where K is the downsampling ratio of output and C is the number of classes. A prediction from the heatmap $\hat{Y}_{x,y,c} = 1$ denotes a detected keypoint, and $\hat{Y}_{x,y,c} = 0$ denotes background. Following Law and Deng (2018), the ground truth

of the keypoints is mapped onto a heatmap Y using a 2D Gaussian kernel as in Equation (3):

$$Y_{x,y,c} = \exp\left(-\frac{(x - \tilde{p}_x)^2 + (y - \tilde{p}_y)^2}{2\sigma_p^2}\right) \quad (3)$$

where \tilde{p}_x and \tilde{p}_y are the equivalent groundtruth keypoints of prediction, and they are downsampled by the factor K from the original keypoint p and are then discretized. σ_p is a kernel standard deviation.

After getting the peaks of the heatmap for tomatoes, the top N peaks are selected among all the detected responses whose value is greater or equal to its eight-connected neighbors. We define $\hat{\mathcal{P}} = \{(\hat{x}_i, \hat{y}_i)\}_{i=1}^N$ as the set of N detected center points. The confidence of the detected bounding circle is represented by the keypoint values $\hat{Y}_{x_i, y_i, c}$, and the center point \hat{p} and radius \hat{r} of the bounding circle is denoted as follows:

$$\hat{p} = (\hat{x}_i + \Delta\hat{x}_i, \hat{y}_i + \Delta\hat{y}_i) \quad (4)$$

$$\hat{r} = \hat{R}_{\hat{x}_i, \hat{y}_i} \quad (5)$$

where $(\Delta\hat{x}_i, \Delta\hat{y}_i) = \hat{O}_{\hat{x}_i, \hat{y}_i} \in R^{\frac{W}{K} \times \frac{H}{K} \times 2}$ is the offset prediction and $\hat{R}_{\hat{x}_i, \hat{y}_i} \in R^{\frac{W}{K} \times \frac{H}{K} \times C}$ is the radius prediction.

3.5.2. Bounding circle IOU

The intersection-over-union (IOU) is commonly used to evaluate the similarity of two bounding boxes. In this study,

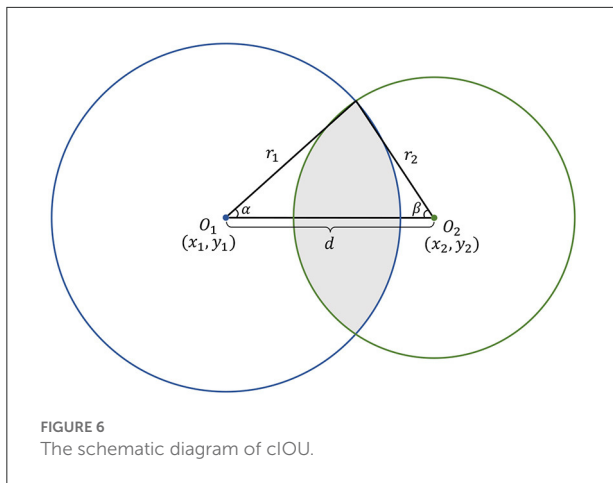


FIGURE 6
The schematic diagram of cIOU.

we introduce a circle IOU (cIOU) for evaluation of two bounding circles.

As shown in Figure 6, denoting the center coordinates of two intersected circles O_1 and O_2 be (x_1, y_1) and (x_2, y_2) , respectively, the distance between two centers d can be represented in Equation (6) and satisfies the condition $|R - r| \leq d \leq |R + r|$.

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (6)$$

The angles α and β can be calculated as:

$$\alpha = \cos^{-1} \frac{r_1^2 + d^2 - r_2^2}{2r_1 d} \quad (7)$$

$$\beta = \cos^{-1} \frac{r_2^2 + d^2 - r_1^2}{2r_2 d} \quad (8)$$

Then, the intersection area $A_{O_1 \cap O_2}$ and union area $A_{O_1 \cup O_2}$ of circles O_1 and O_2 can be derived as in Equations (9) and (10).

$$A_{O_1 \cap O_2} = \alpha r_1^2 + \beta r_2^2 - \frac{1}{2} r_1^2 \sin 2\alpha - \frac{1}{2} r_2^2 \sin 2\beta \quad (9)$$

$$A_{O_1 \cup O_2} = \pi r_1^2 + \pi r_2^2 - A_{O_1 \cap O_2} \quad (10)$$

Consequently, the cIOU can be represented as follows:

$$\text{cIOU} = \frac{(2\alpha - \sin 2\alpha)r_1^2 + (2\beta - \sin 2\beta)r_2^2}{(2\pi - 2\alpha + \sin 2\alpha)r_1^2 + (2\pi - 2\beta + \sin 2\beta)r_2^2} \quad (11)$$

3.6. Loss function

The loss function of TomatoDet in the training stage consists of three parts, i.e., the keypoint heatmap loss, bounding circle

radius loss and center offset loss. The keypoint heatmap loss L_{hm} is based on focal loss (Lin et al., 2017) as in Equation (12).

$$L_{hm} = -\frac{1}{N} \sum_{x,y,c} \begin{cases} (1 - \hat{Y}_{x,y,c})^\alpha \log \hat{Y}_{x,y,c} & \text{if } Y_{x,y,c} = 1 \\ (1 - Y_{x,y,c})^\beta (\hat{Y}_{x,y,c})^\alpha \log (1 - \hat{Y}_{x,y,c}) & \text{otherwise} \end{cases} \quad (12)$$

where N is the number of keypoints in an image, and α and β are hyper-parameters for the focal loss. In this study, α and β are set to be 2 and 4 following Zhou et al. (2019).

To rectify the keypoint location error resulting from the discretization of downsampling, an offset loss L_{off} is designed to measure the difference between the predicted offset \hat{O} and the groundtruth O based on L1 loss.

$$L_{off} = \frac{1}{N} \sum_p |\hat{O}_p - O_p| \quad (13)$$

The tomato radius is regressed from the center points optimized by the radius loss L_r in Equation (14).

$$L_r = \frac{1}{N} \sum_{k=1}^N |\hat{R}_{p_k} - r_k| \quad (14)$$

where \hat{R}_{p_k} and r_k denotes the predicted and groundtruth radius of the k th tomato, and N represents the number of results.

Above of all, the total loss of TomatoDet is denoted as in Equation (15).

$$L_{det} = L_{hm} + \lambda_{off} L_{off} + \lambda_r L_r \quad (15)$$

where $\lambda_{off} = 1$ and $\lambda_r = 0.1$ are used in our experiment to balance different losses, referring to Zhou et al. (2019).

3.7. Experimental setup

The experiments are performed on a Ubuntu 16.04 with an Intel(R) Core(TM) i7-9700 K CPU@3.60 GHz. It is accelerated by an NVIDIA GeForce GTX 1080Ti GPU. The proposed TomatoDet model is implemented on Pytorch.

The model is trained on an input resolution of 512×512 pixels. It is trained with a batch size of 8 and an initial learning rate of $1.25e-4$ for 140 epochs. The learning rate is then dropped 10 at 90 and 120 epochs, respectively.

To evaluate the performance of the proposed method, recall (R), precision (P), and F_1 score are used as the criterion indexes. They are defined in Equations (16)–(18):

$$P = \frac{TP}{TP + FP} \quad (16)$$

$$R = \frac{TP}{TP + FN} \quad (17)$$

$$F_1 = \frac{2 \times P \times R}{P + R} \quad (18)$$

where TP, FP, and FN represent true positives (correct detections), false positives (false detections), and false negatives (missing detections), respectively.

Besides, the average precision (AP) is adopted in this study to evaluate the overall detection performance. AP is defined as follows:

$$AP = \sum_n (r_{n+1} - r_n) p_{interp}(r_{n+1}) \quad (19)$$

$$p_{interp}(r_{n+1}) = \max_{\tilde{r}: \tilde{r} \geq r_{n+1}} p(\tilde{r}) \quad (20)$$

where $p(\tilde{r})$ is the measured precision at recall \tilde{r} .

4. Results and discussion

4.1. Ablation study

In this study, an attention mechanism and a circle representation are incorporated to the proposed detector. In order to evaluate the effectiveness of each component, an ablation study is performed on the tomato dataset. The results of the ablation experiments are shown in Table 2 and Figure 7.

From Table 2, we can see that the incorporation of the attention mechanism brought a significant improvement of all the indexes including the recall, precision, F₁ score and average precision (AP). The F₁ score and AP increases by 1.33 and 1.36%, respectively. This verifies the advantages of the proposed attentive-dla34 backbone, which optimizes the focus of the network and boosts the representation power. We also performed a contrast experiment to verify the effectiveness of the circle representation. With circle representation, the F₁ score and AP increases by 1.44 and 1.23%, respectively, as shown in Table 2. This benefits from the intrinsic shape fitting of the new circle representation to tomatoes, which can reduce the degree of freedom of the rectangle representation and simplify the regression process from detected keypoints. We also show the precision-recall (PR) curves of different components in Figure 7. The markers indicate the points where recall and precision are obtained when the confidence threshold equals 0.6. It can be

TABLE 2 Ablation study on the major components of TomatoDet.

Attention module	Circle representation	Recall (%)	Precision (%)	F ₁ (%)	AP (%)
		91.56	92.98	92.26	95.75
✓		92.87	94.32	93.59	97.11
	✓	92.98	94.43	93.70	96.98
✓	✓	94.30	95.77	95.03	98.16

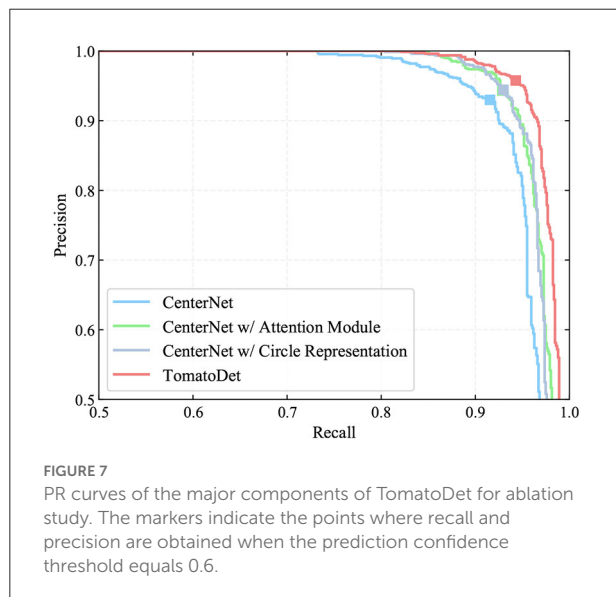
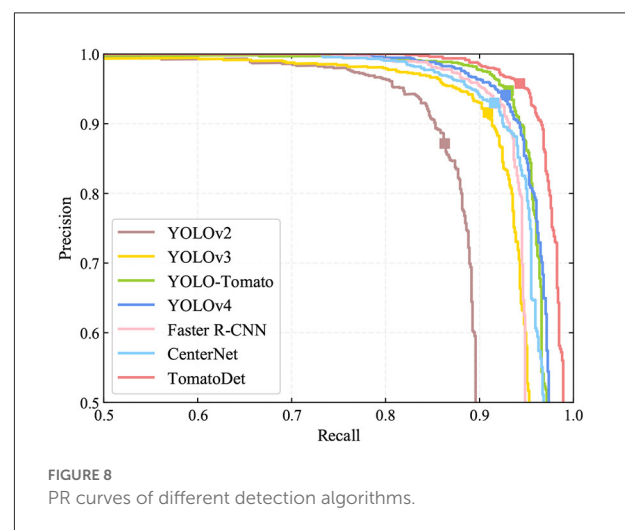


TABLE 3 Tomato detection results of different algorithms.

Methods	Recall (%)	Precision (%)	F ₁ (%)	AP (%)	(ms)
YOLOv2	86.18	87.24	86.71	88.46	30
YOLOv3	90.89	91.60	91.24	94.06	45
YOLO-Tomato	93.09	94.75	93.91	96.40	54
YOLOv4	92.76	94.11	93.43	96.59	25
Faster R-CNN	91.78	92.89	92.33	94.37	231
CenterNet	91.56	92.98	92.26	95.75	32
TomatoDet	94.30	95.77	95.03	98.16	35



seen that the detection performance improves significantly with the incorporation of different components.

4.2. Comparison of different methods

To verify the performance of the proposed TomatoDet model, we designed a comparative experiment of the state-of-the-art detection algorithms, including YOLOv2 (Redmon and Farhadi, 2017), YOLOv3 (Redmon and Farhadi, 2018), YOLO-Tomato (Liu G. et al., 2020), YOLOv4 (Bochkovskiy et al., 2020), Faster R-CNN (Ren et al., 2015), CenterNet (Zhou et al.,

2019), and the proposed model. Among all of these algorithms, the Faster R-CNN is a two-stage detector, and the others are one-stage detectors. Moreover, CenterNet and the proposed TomatoDet are anchor-free detectors, while the remaining are all anchor-based methods.

The recall, precision, F_1 score, average precision (AP), and average detection time are the evaluation indicators, as shown in Table 3. The precision-recall (PR) curves of different detection models are shown in Figure 8. In terms of detection performance, one can see that the proposed TomatoDet is superior to the other five methods. The F_1 score of TomatoDet is 95.03%. It is 1.12% higher than that of YOLO-Tomato,

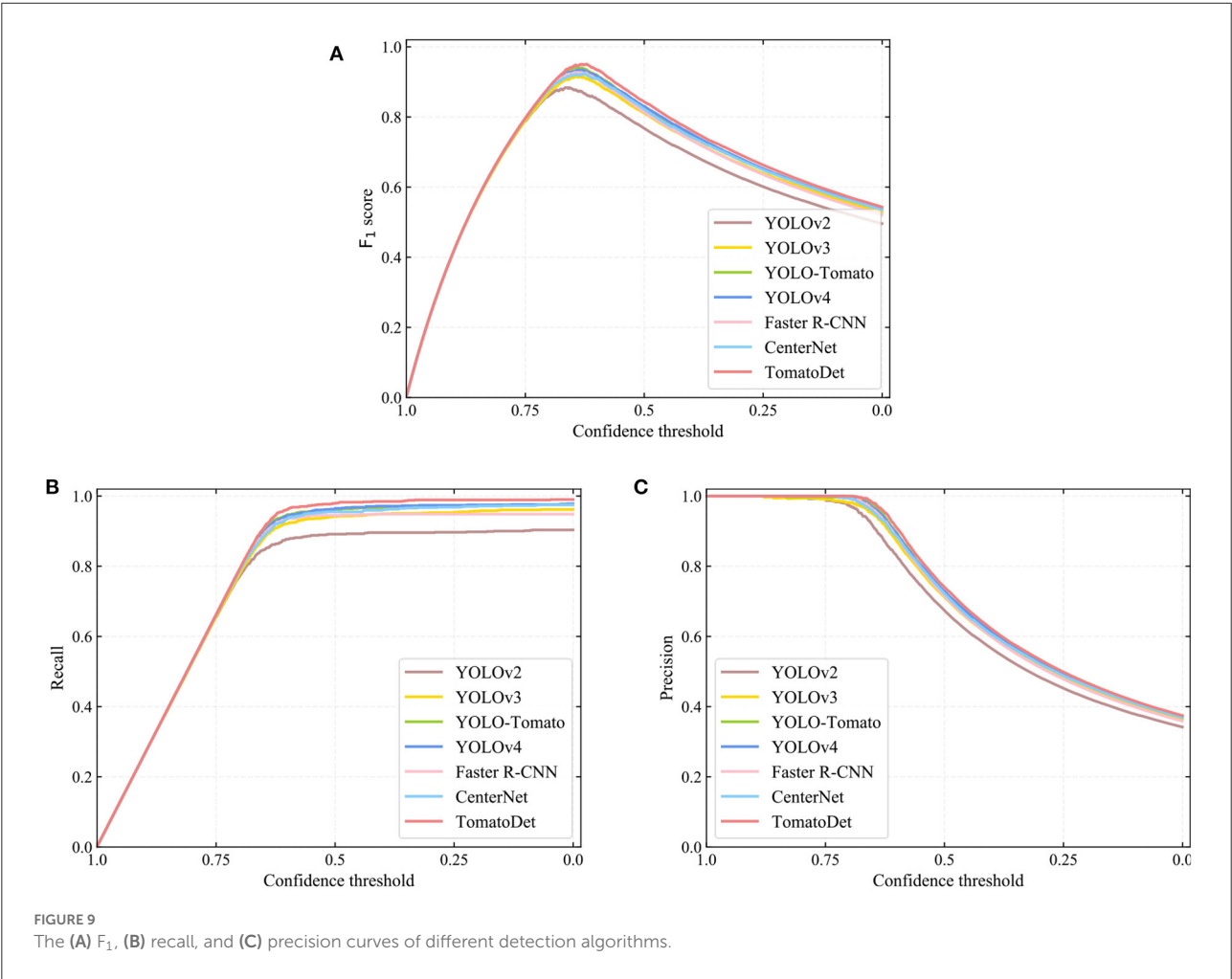


TABLE 4 Performance of the proposed TomatoDet under different lighting conditions.

Illumination	Tomato count	Correctly identified		Falsely identified		Missed	
		Amount	Rate(%)	Amount	Rate (%)	Amount	Rate (%)
Sunlight	487	460	94.46	22	4.56	27	5.54
Shading	425	400	94.12	16	3.85	25	5.88

which obtains the second-best performance. In terms of AP, TomatoDet performs 1.76 and 1.57% better than YOLO-Tomato and YOLOv4, respectively. Compared to CenterNet, the proposed TomatoDet is about 2.8 and 2.4% higher in terms of F₁ score and AP, respectively. We also show the F₁, recall and precision curves in Figure 9, separately. In accordance with the PR curves, they demonstrate the superiority of the proposed TomatoDet over other methods. This verifies the effectiveness of the proposed modifications. The introduction of CBAM guides the model to pay more attention to the ROI and thus improves the feature expression ability of the network. Besides, the adoption of bounding circles makes it easier to regress from center points to the size as the bounding circle only has one parameter, i.e., radius. Furthermore, bounding circles could match the shape of tomatoes better in nature and improve the IOU. The average detection time of the proposed model reaches 0.036 s per image. It is about 0.2 s less than Faster R-CNN and almost the same as the YOLOv2 model. The experimental results show that the proposed TomatoDet could detect tomatoes in complex environments in real-time with strong robustness.

4.3. Qualitative analysis

To better understand the prediction ability of our proposed TomatoDet, the output feature is visualized. Figure 10 shows

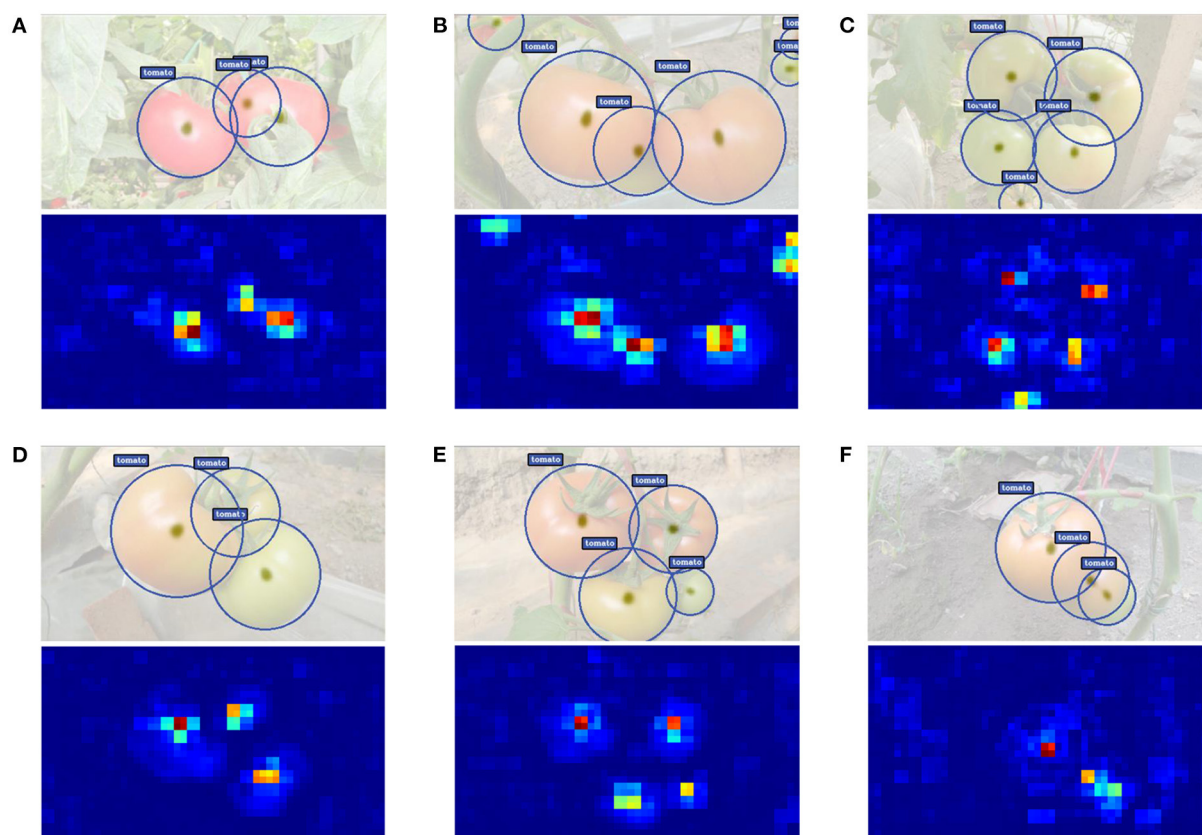
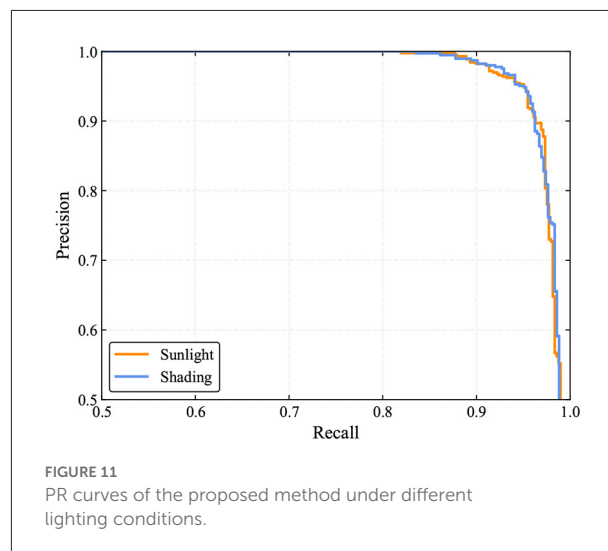


FIGURE 10
(A–F) Some examples of detection results along with the output heatmap.

some examples of detection results along with the output heatmap. From the second row of the subfigures, one can see that through the proposed attentive-DLA34 backbone, the heatmap almost only fires at the area of tomatoes, including small and severe occluded ones. This benefits from the combination of CBAM and DLA34, which emphasizes the meaningful features throughout the network and thus boosts the representation power. Further, the keypoints for tomatoes are extracted from the peaks of the heatmap and are then regressed to the radius of the proposed bounding circle, which reduces the degree of freedom of fitting compared to the traditional bounding boxes, as is shown in the first row of the subfigures.

4.4. Performance of the proposed model under different lighting conditions

In the natural environment, tomatoes may be exposed to different lighting conditions due to uneven illuminations. The performance of the proposed TomatoDet under different lighting conditions is evaluated in this study. Among all the tomatoes in the test set, 425 tomatoes are in shading conditions, while 487 tomatoes are in sunlight conditions. The correct identification rate (or recall), false identification rate and missing rate are used as evaluation indicators.

As shown in Table 4, 460 out of 487 tomatoes are correctly identified by the TomatoDet under sunlight conditions. The counterpart is 400 out of 425 for the shading conditions. The correct identification rates are comparable. The false identification rates are 4.56 and 3.85% for sunlight and shading conditions, respectively. This means that some of the detections are falsely recognized as tomatoes, which in fact are leaves, branches, or other backgrounds. This occurs when the background presents similar color and shape to tomatoes. The above results show that the proposed method is robust under different lighting conditions in real scenes. From Figure 11, one can see that the PR curves under sunlight and shading conditions are comparable, showing the robustness of the proposed method to different lighting conditions. Some examples are shown in Figure 12.

4.5. Performance of the proposed model under different occlusion conditions

In the greenhouse, tomatoes are inevitably obscured by leaves or branches and overlap with each other. This will have a certain impact on tomato detection. In this study, we also evaluate the performance of the proposed method under different occlusion conditions. As in YOLO-Tomato (Liu G. et al., 2020), depending on the degree of occlusion or overlap,

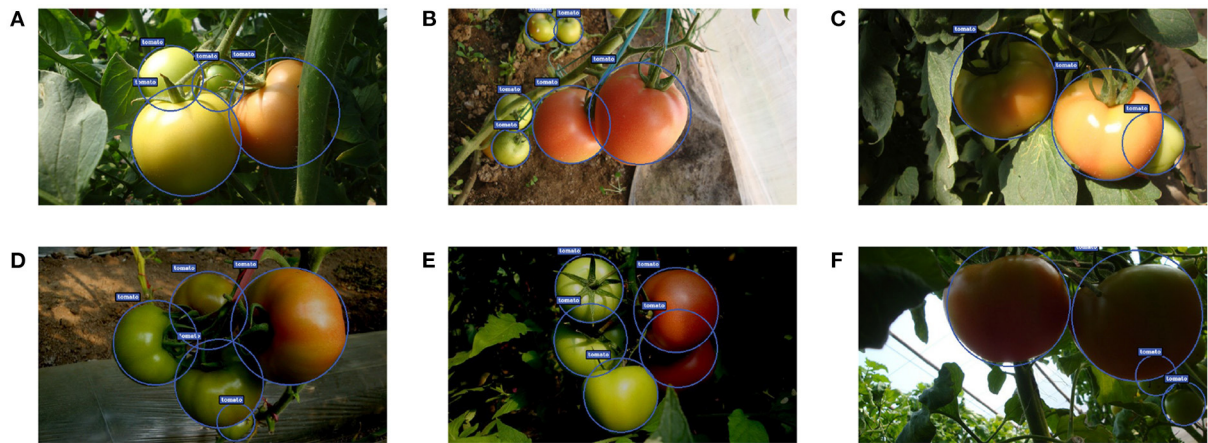


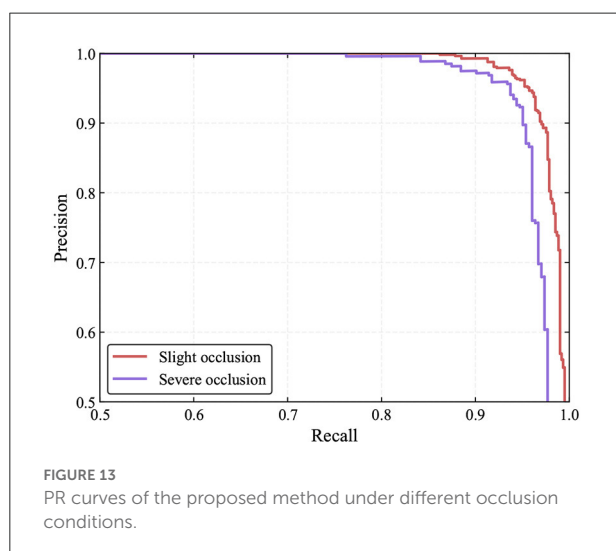
FIGURE 12 Some examples of the detection results under different lighting conditions: (A–C) sunlight conditions, and (D–F) shading conditions.

TABLE 5 Performance of the proposed TomatoDet under different occlusion conditions.

Occlusion condition	Tomato count	Correctly identified		Falsely identified		Missed	
		Amount	Rate (%)	Amount	Rate (%)	Amount	Rate (%)
Slight case	609	576	94.58	22	3.68	33	5.42
Severe case	303	284	93.73	16	5.33	19	6.27

we classify tomatoes as slight and severe occlusion cases. Severe cases refer to tomatoes being blocked by leaves, branches, or other tomatoes by more than 50% degrees. Conversely, tomatoes are regarded as slight cases. The detection results are shown in Table 5 and Figure 13.

Based on the above experiments, one can see that the detection performance for tomatoes under slight occlusion cases is marginally better than that of tomatoes under severe cases. This shows that occluded and overlapped tomatoes cause inaccurate detections. Nevertheless, most of the occluded and overlapped tomatoes can be detected by our model correctly. This is achieved by the accurate keypoints estimation resulting from the implicit contextual information utilization of the convolutional neural networks since the networks learn hierarchical features through multiple levels of abstraction.



However, it is believed that the detection performance of occluded tomatoes can be further improved by exploiting contextual information explicitly (Liu L. et al., 2020). Figure 14 shows some examples of detection results for both cases.

5. Conclusions and future work

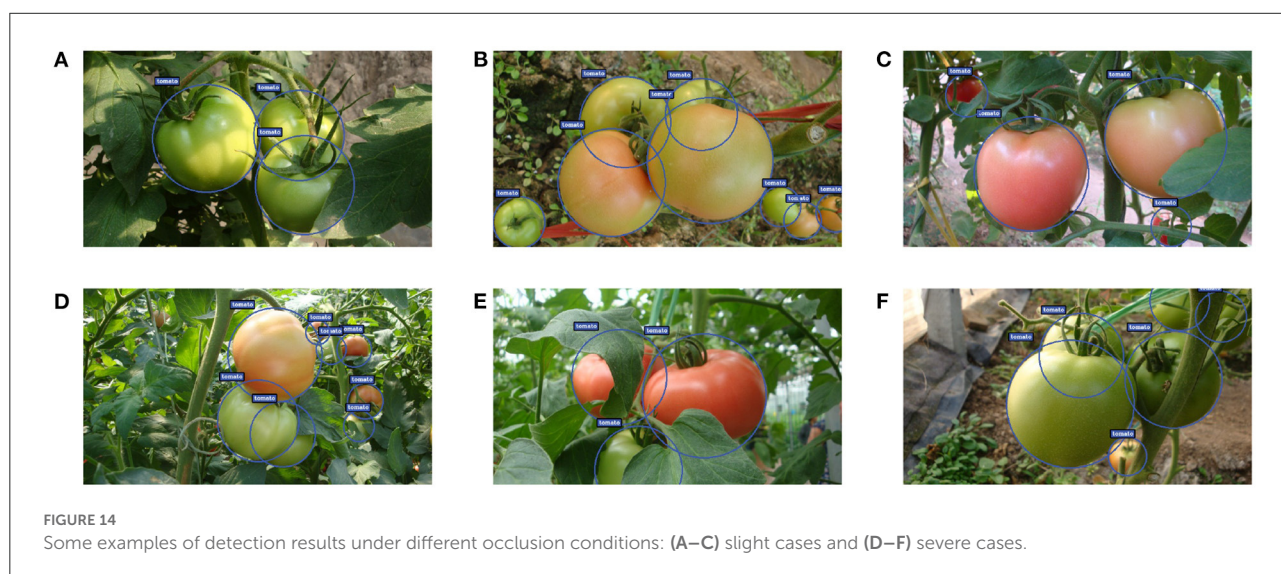
In this study, we propose TomatoDet, an improved anchor-free detector for tomato detection based on CenterNet. The proposed detector incorporates an attention mechanism to optimize the focus of the network and thus boost the representation power. In addition, a circle representation is introduced to adapt the detector to our specific detection task. With circle representation, the degree of freedom for tomato fitting is reduced and the regression process from keypoints to the size is simplified.

The experimental results show that the proposed TomatoDet is superior to other state-of-the-art detectors for tomato detection in the greenhouse. It can also detect tomatoes under different lighting and occlusion conditions with strong robustness.

Although the proposed model has achieved a good performance on the tomato datasets, there is still much space for further development. They can be summarized as follows:

When the overlap or occlusion area is high, the detection rate will drop. One possible solution is to incorporate contextual information such as branches or leaves to improve the detection accuracy.

The experimental dataset is relatively small and more data are needed for training and verification in the future study.



Moreover, the characteristics of tomatoes in different growing stages will be analyzed to realize multi-stage tomato detection.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

GL conceived the research idea. GL and ZH designed the methodology. JL and HL performed the experiments and analysis. GL and KL wrote the original draft. WZ and KL revised the manuscript. KL supervised the experiments. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the Weifang Science and Technology Development Plan (2021GX054),

Doctoral Research Foundation of Weifang University (2022BS70), and Natural Science Foundation of Shandong Province (ZR2021QC173).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Afonso, M., Fonteijn, H., Fiorentin, F. S., Lensink, D., Mooij, M., Faber, N., et al. (2020). Tomato fruit detection and counting in greenhouses using deep learning. *Front. Plant Sci.* 2020, 1759. doi: 10.3389/fpls.2020.571299
- Behroozi-Khazaei, N., and Maleki, M. R. (2017). A robust algorithm based on color features for grape cluster segmentation. *Comput. Electron. Agric.* 142, 41–49. doi: 10.1016/j.compag.2017.08.025
- Bochkovskiy, A., Wang, C.-Y., and Liao, H.-Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*. doi: 10.48550/arXiv.2004.10934
- Bresilla, K., Perulli, G. D., Boini, A., Morandi, B., Corelli Grappadelli, L., and Manfrini, L. (2019). Single-shot convolution neural networks for real-time fruit detection within the tree. *Front. Plant Sci.* 10, 611. doi: 10.3389/fpls.2019.00611
- Bulanon, D. M., Kataoka, T., Ota, Y., and Hiroma, T. (2002). AE-automation and emerging technologies: a segmentation algorithm for the automatic recognition of fuji apples at harvest. *Biosyst. Eng.* 83, 405–412. doi: 10.1006/bioe.2002.0132
- Chen, W., Lu, S., Liu, B., Chen, M., Li, G., and Qian, T. (2022). CitrusYOLO: a algorithm for citrus detection under orchard environment based on YOLOV4. *Multim. Tools Appl.* 1–27. doi: 10.1007/s11042-022-12687-5
- Gongal, A., Amatya, S., Karkee, M., Zhang, Q., and Lewis, K. (2015). Sensors and systems for fruit detection and localization: a review. *Comput. Electr. Agric.* 116, 8–19. doi: 10.1016/j.compag.2015.05.021
- Kelman, E. E., and Linker, R. (2014). Vision-based localisation of mature apples in tree images using convexity. *Biosyst. Eng.* 118, 174–185. doi: 10.1016/j.biosystemseng.2013.11.007
- Koirala, A., Walsh, K., Wang, Z., and McCarthy, C. (2019). Deep learning for real-time fruit detection and orchard fruit load estimation: benchmarking of “mangoyolo”. *Precis. Agric.* 20, 1107–1135. doi: 10.1007/s11119-019-09642-0
- Kurtulmus, F., Lee, W. S., and Vardar, A. (2011). Green citrus detection using “eigenfruit,” color and circular Gabor texture features under natural outdoor conditions. *Comput. Electr. Agric.* 78, 140–149. doi: 10.1016/j.compag.2011.07.001
- Kurtulmus, F., Lee, W. S., and Vardar, A. (2014). Immature peach detection in colour images acquired in natural illumination conditions using statistical classifiers and neural network. *Precis. Agric.* 15, 57–79. doi: 10.1007/s11119-013-9323-8
- Lam, E. Y. (2005). “Combining gray world and retinex theory for automatic white balance in digital photography,” in *Proceedings of the Ninth International Symposium on Consumer Electronics, 2005 (ISCE 2005)* (Macau), 134–139. doi: 10.1109/ISCE.2005.1502356
- Law, H., and Deng, J. (2018). “CornerNet: detecting objects as paired keypoints,” in *Proceedings of the European Conference on Computer Vision (ECCV)* (Munich), 734–750. doi: 10.1007/978-3-030-01264-9_45
- Li, H., Zhang, M., Gao, Y., Li, M., and Ji, Y. (2017). Green ripe tomato detection method based on machine vision in greenhouse. *Trans. Chinese Soc. Agric. Eng.* 33, 328–334. doi: 10.11975/j.issn.1002-6819.2017.z1.049
- Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P. (2017). “Focal loss for dense object detection,” in *Proceedings of the IEEE International Conference on Computer Vision (Venice)*, 2980–2988. doi: 10.1109/ICCV.2017.324
- Linker, R., Cohen, O., and Naor, A. (2012). Determination of the number of green apples in RGB images recorded in orchards. *Comput. Electr. Agric.* 81, 45–57. doi: 10.1016/j.compag.2011.11.007
- Liu, G., Mao, S., and Kim, J. H. (2019). A mature-tomato detection algorithm using machine learning and color analysis. *Sensors* 19:2023. doi: 10.3390/s19092023
- Liu, G., Nouaze, J. C., Touko Mbouembe, P. L., and Kim, J. H. (2020). YOLO-tomato: a robust algorithm for tomato detection based on YOLOV3. *Sensors* 20:2145. doi: 10.3390/s20072145
- Liu, L., Ouyang, W., Wang, X., Fieguth, P., Chen, J., Liu, X., et al. (2020). Deep learning for generic object detection: a survey. *Int. J. Comput. Vis.* 128, 261–318. doi: 10.1007/s11263-019-01247-4
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., et al. (2016). “SSD: single shot multibox detector,” in *European Conference on Computer Vision (Amsterdam: Springer)*, 21–37. doi: 10.1007/978-3-319-46448-0_2
- Lv, Q., Cai, J., Liu, B., Deng, L., and Zhang, Y. (2014). Identification of fruit and branch in natural scenes for citrus harvesting robot using machine vision and support vector machine. *Int. J. Agric. Biol. Eng.* 7, 115–121. doi: 10.3965/j.jabe.20140702.014

- Okamoto, H., and Lee, W. S. (2009). Green citrus detection using hyperspectral imaging. *Comput. Electr. Agric.* 66, 201–208. doi: 10.1016/j.compag.2009.02.004
- Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). “You only look once: unified, real-time object detection,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Las Vegas, NV), 779–788. doi: 10.1109/CVPR.2016.91
- Redmon, J., and Farhadi, A. (2017). “YOLO9000: better, faster, stronger,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Honolulu), 7263–7271. doi: 10.1109/CVPR.2017.690
- Redmon, J., and Farhadi, A. (2018). YOLOV3: an incremental improvement. *arXiv preprint arXiv:1804.02767*. doi: 10.48550/arXiv.1804.02767
- Ren, S., He, K., Girshick, R., and Sun, J. (2015). “Faster r-CNN: towards real-time object detection with region proposal networks,” in *Advances in Neural Information Processing Systems* (Montreal, QC), 28.
- Sa, I., Ge, Z., Dayoub, F., Upcroft, B., Perez, T., and McCool, C. (2016). Deepfruits: a fruit detection system using deep neural networks. *Sensors* 16:1222. doi: 10.3390/s16081222
- Tian, Y., Yang, G., Wang, Z., Wang, H., Li, E., and Liang, Z. (2019). Apple detection during different growth stages in orchards using the improved YOLO-V3 model. *Comput. Electr. Agric.* 157, 417–426. doi: 10.1016/j.compag.2019.01.012
- Wei, J., Ding, Y., Liu, J., Ullah, M. Z., Yin, X., and Jia, W. (2022). Novel green-fruit detection algorithm based on D2D framework. *Int. J. Agric. Biol. Eng.* 15, 251–259. doi: 10.25165/j.ijabe.20221501.6943
- Wei, X., Jia, K., Lan, J., Li, Y., Zeng, Y., and Wang, C. (2014). Automatic method of fruit object extraction under complex agricultural background for vision system of fruit picking robot. *Optik* 125, 5684–5689. doi: 10.1016/j.ijleo.2014.07.001
- Woo, S., Park, J., Lee, J.-Y., and Kweon, I. S. (2018). “CBAM: Convolutional block attention module,” in *Proceedings of the European Conference on Computer Vision (ECCV)* (Munich), 3–19. doi: 10.1007/978-3-030-01234-2_1
- Wu, J., Zhang, B., Zhou, J., Xiong, Y., Gu, B., and Yang, X. (2019). Automatic recognition of ripening tomatoes by combining multi-feature fusion with a bi-layer classification strategy for harvesting robots. *Sensors* 19:612. doi: 10.3390/s19030612
- Yamamoto, K., Guo, W., Yoshioka, Y., and Ninomiya, S. (2014). On plant detection of intact tomato fruits using image analysis and machine learning methods. *Sensors* 14, 12191–12206. doi: 10.3390/s140712191
- Yu, F., Wang, D., Shelhamer, E., and Darrell, T. (2018). “Deep layer aggregation,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Salt Lake City, UT), 2403–2412. doi: 10.1109/CVPR.2018.00255
- Yue, Z., Li, S., and Feldman, M. W. (2015). *Social Integration of Rural-Urban Migrants in China: Current Status, Determinants and Consequences*, Vol. 13. World Scientific. doi: 10.1142/9428
- Zhang, W., Liu, Y., Chen, K., Li, H., Duan, Y., Wu, W., et al. (2021). Lightweight fruit-detection algorithm for edge computing applications. *Front. Plant Sci.* 12:740936. doi: 10.3389/fpls.2021.740936
- Zhao, C., Lee, W. S., and He, D. (2016). Immature green citrus detection based on colour feature and sum of absolute transformed difference (SATD) using colour images in the citrus grove. *Comput. Electr. Agric.* 124, 243–253. doi: 10.1016/j.compag.2016.04.009
- Zhao, Y., Gong, L., Huang, Y., and Liu, C. (2016a). A review of key techniques of vision-based control for harvesting robot. *Comput. Electr. Agric.* 127, 311–323. doi: 10.1016/j.compag.2016.06.022
- Zhao, Y., Gong, L., Huang, Y., and Liu, C. (2016b). Robust tomato recognition for robotic harvesting using feature images fusion. *Sensors* 16:173. doi: 10.3390/s16020173
- Zhao, Y., Gong, L., Zhou, B., Huang, Y., and Liu, C. (2016c). Detecting tomatoes in greenhouse scenes by combining adaboost classifier and colour analysis. *Biosyst. Eng.* 148, 127–137. doi: 10.1016/j.biosystemseng.2016.05.001
- Zheng, Z., Xiong, J., Lin, H., Han, Y., Sun, B., Xie, Z., et al. (2021). A method of green citrus detection in natural environment using a deep convolutional neural network. *Front. Plant Sci.* 12:705737. doi: 10.3389/fpls.2021.705737
- Zhou, X., Wang, D., and Krähenbühl, P. (2019). Objects as points. *arXiv preprint arXiv:1904.07850*. doi: 10.48550/arXiv.1904.07850



OPEN ACCESS

EDITED BY
Gregorio Egea,
University of Seville, Spain

REVIEWED BY
Soodabeh Einafshar,
Agricultural Research, Education and
Extension Organization (AREEO), Iran
Javid A. Parry,
Department of Environmental Science
GDC Eidgah Srinagar, India

*CORRESPONDENCE
Amjad M. Husaini
amjadhusaini@skuastkashmir.ac.in

SPECIALTY SECTION
This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

RECEIVED 16 May 2022
ACCEPTED 19 July 2022
PUBLISHED 12 August 2022

CITATION
Husaini AM, Haq SA, Shabir A, Wani AB
and Dedmari MA (2022) The menace
of saffron adulteration: Low-cost rapid
identification of fake look-alike saffron
using Foldscope and machine learning
technology.
Front. Plant Sci. 13:945291.
doi: 10.3389/fpls.2022.945291

COPYRIGHT
© 2022 Husaini, Haq, Shabir, Wani and
Dedmari. This is an open-access
article distributed under the terms of
the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution
or reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

The menace of saffron adulteration: Low-cost rapid identification of fake look-alike saffron using Foldscope and machine learning technology

Amjad M. Husaini^{1*}, Syed Anam Ul Haq¹, Asma Shabir¹,
Amir B. Wani¹ and Muneer A. Dedmari²

¹Genome Engineering and Societal Biotechnology Lab, Division of Plant Biotechnology,
Sher-e-Kashmir University of Agricultural Sciences and Technology of Kashmir, Srinagar, India,
²Technische Universität München, Munich, Germany

Saffron authenticity is important for the saffron industry, consumers, food industry, and regulatory agencies. Herein we describe a combo of two novel methods to distinguish genuine saffron from fake in a user-friendly manner and without sophisticated instruments. A smartphone coupled with Foldscope was used to visualize characteristic features and distinguish “genuine” saffron from “fake.” Furthermore, destaining and staining agents were used to study the staining patterns. Toluidine blue staining pattern was distinct and easier to use as it stained the papillae and the margins deep purple, while its stain is lighter yellowish green toward the central axis. Further to automate the process, we tested and compared different machine learning-based classification approaches for performing the automated saffron classification into genuine or fake. We demonstrated that the deep learning-based models are efficient in learning the morphological features and classifying samples as either fake or genuine, making it much easier for end-users. This approach performed much better than conventional machine learning approaches (random forest and SVM), and the model achieved an accuracy of 99.5% and a precision of 99.3% on the test dataset. The process has increased the robustness and reliability of authenticating saffron samples. This is the first study that describes a customer-centric frugal science-based approach to creating an automated app to detect adulteration. Furthermore, a survey was conducted to assess saffron adulteration and quality. It revealed that only 40% of samples belonged to ISO Category I, while the average adulteration percentage in the remaining samples was 36.25%. After discarding the adulterants from crude samples, their quality parameters improved significantly, elevating these from ISO category III to Category II. Conversely, it also means that Categories II and III saffron are more prone to and favored for adulteration by fraudsters.

KEYWORDS

Crocus sativus, Foldscope, microscopy, adulteration, fraud, machine learning, deep learning, image processing

Introduction

Saffron (*Crocus sativus* L.) is a highly remunerative cash crop and a source of luxury spice obtained from handpicked flowers as dried crimson stigmas (Kafi et al., 2018). According to Food and Agricultural Organization (FAO), it forms “a loosely matted mass of dark, reddish-brown flattened threads, among which a few narrower yellow ones can be distinguished. The upper, enlarged part of the flattened threads is the stigma of the flower, the lower narrower portion is the style” (Husaini et al., 2010a). Saffron bioactive compounds have immense therapeutic properties useful for coronary artery diseases, neurodegenerative disorders, bronchitis, asthma, diabetes, fever, and colds. It has the potential to help tackle problems associated with severe acute respiratory syndrome (COVID-19) patients and post-COVID-19 problems (Ahmed and Husaini, 2021). It can help manage stress and anxiety during isolation, quarantine, and lockdowns (Husaini et al., 2021). Owing to all these beneficial properties and as an immunity booster, saffron extracts may be added to some drug formulations in future (Husaini et al., 2022). These properties and their importance in religious rituals of many communities make it costly and hence prone to adulterations. Some have even advocated its cultivation in kitchen gardens to ensure the supply of pure saffron for household use (Husaini and Wani, 2020).

The best quality saffron is usually sold in filaments (Melnik et al., 2010; Nehvi and Yasmin, 2021); therefore, in the present study, we focused on filamentous saffron. Different kinds of fake products sold under the name of “saffron” are reported in the literature (Husaini et al., 2010b; Heidarbeigi et al., 2015). The most common fraudulent practice includes artificial dyeing of some selected plant materials, making these look similar to saffron. According to a study on saffron sold in India, only 52% are genuine, 30% are poor grade, and 17% are adulterated (Husaini et al., 2010a). This menace of saffron adulteration is mushrooming as a white-collar fraud at a tremendous pace (Husaini et al., 2010a, 2013; Er et al., 2017).

According to the ISO 3632 standards (ISO, 2010, 2011), up to 1% (w/w) of foreign material is permitted in third-class products. Several chromatographic and spectroscopic methods are used for detecting saffron adulterants (Alonso et al., 1998; Lozano et al., 1999; Haghighi et al., 2007; Sabatino et al., 2011; Er et al., 2017). Moreover, several biotechnological and molecular methods are also employed to detect plant adulterants in saffron (Ma et al., 2001; Javanmardi et al., 2011; Marieschi et al., 2012; Babaei et al., 2014; Torelli et al., 2014; Petrakis et al., 2015; Kumari et al., 2021). These methods are useful for detecting low amounts (up to 1%) of bulking materials and are ideal for checking the purity of the product. However, all these analytical methods are too complicated, expensive, and need sophisticated instrumentation and higher skill levels of experts. The lack

of sophisticated laboratories and ineffective law enforcement adds to the constraints in handling cheating by retailers. Some affordable methods like simple microscopy or spectrometry have their limitations too. While a conventional microscope is a fairly expensive instrument seldom used by common people, UV-vis spectrophotometric method used in labs does not detect saffron contamination up to 20% (w/w) (Sabatino et al., 2011). None of the methods developed so far is easily accessible to customers or retailers. Hence, there is a need to invent faster, low-cost screening methods for detecting saffron adulteration and fictitious look-alike versions of saffron, and make these easily accessible to the end-users.

There is a need to have a customer-centric rather than a lab-centric approach. We want to change the standpoint of looking at the problem by bringing the customer directly into the screening procedure. Customers should be able to check the authenticity of a particular sample on a retail scale because saffron being expensive, is generally sold in small packings of 1–5 grams. The present study focuses on “fictitious look-alike” versions of saffron sold in the markets under the names of fictitious brands on a commercial scale around religious places, spice markets, individual retailers, and the unorganized sector. We aim to share a customer-friendly technology that is the first of its kind and does not depend on sophisticated instrumentation.

Recent developments in “frugal science” have made a monocular origami-based low-cost optical microscope, called Foldscope, commercially available and easily accessible (Cybulski et al., 2014; Moreno-Roman and Bobick, 2022). Similarly, there are tremendous advances in artificial intelligence-based solutions and machine learning (Ben Ayed and Hanana, 2021; Janiesch et al., 2021; Vijayakumar and Balakrishnan, 2021; Greener et al., 2022). We explored these developments and developed two methods for the self-detection of fake saffron by customers and retailers. One method uses Foldscope in combination with chemical staining and destaining technique for developing a printed poster to detect fake look-alike saffron through visual comparison. The effect of different dyes on the staining pattern of the samples was studied so that even school children could use this technique. The second method uses deep learning for image classification to automatically identify genuine from fake look-alike saffron samples. The method uses Foldscope and a mobile application (app) to automate the process without using any invasive procedure. It is time-efficient and can be used by people who do not have much knowledge about the domain (Saffron). To the best of our knowledge, this is the first time that Foldscope and machine learning have been used to authenticate saffron as fake/genuine and provide user-friendly testing access to a broader audience.

Materials and methods

Survey and collection of samples for analysis

The collection of plant material and all experiments were performed following relevant institutional, national, and international guidelines and legislation. Direct interactions were done with saffron farmers, traders, and consumers/tourists to know their experiences and to find out the nature of adulterants being commonly used by fraudsters. Only a few cooperated in giving some basic information about adulteration methods.

Categorization of genuine and fake look-alike samples

A total of nine diverse classes of samples were used in the present study (Table 1). Seven classes belonged to the fictitious look-alike saffron filaments collected from open markets in India, while two classes belonged to genuine saffron grades (known as “Laccha” in native Kashmiri [“saffron in filaments” as per ISO3632 and IS5453] and “Mongra” [saffron processed using a technique indigenous to Kashmir or “saffron in cut filaments” as per ISO3632 and IS5453]) (Figure 1). “Laccha” is the vernacular Kashmiri term for saffron filament with style and “Mongra” for a locally processed grade of cut filaments lacking style.

Microscopic study

Samples were observed under stereo-microscope (Olympus SZX16 using software LCmicro-2016-17 version) and Foldscope. Foldscope is an origami-based optical microscope developed at the University of Stanford, USA and designed to cost <US\$1. It weighs about 8–10 g and provides a magnification of 140×. It does not require external power and can survive being dropped from a three-story building (Cybulski et al., 2014; Joshi and Bhosale, 2018). With good resolution cell phone cameras, direct imaging is possible. Alternatively, the image can be viewed on a frosted sheet (thin velum) which can be placed above the lens.

To develop the Foldscope-based method, we used 2,250 filaments belonging to nine classes, with at least 250 individual filaments from each class for microscopic study. Single strands of dry, intact filaments were placed directly on clean, dry glass slides and covered with transparent cello-tape (Supplementary Figure S1f). These were observed under Foldscope in natural sunlight. Images were obtained by coupling the Foldscope with a cell phone (iPhone SE) using a custom magnetic coupler. The Foldscope magnification is 140×, which was further enhanced digitally by the zooming function of the smartphone having a 12-megapixel resolution camera.

Observations were recorded for morphological features like (a) papillose protuberances; (b) margins; (c) serrations; (d) texture; (e) dyeing patterns; (f) pubescence; (g) pollen grains.

De-pigmentation

In order to distinguish between artificially dyed samples and genuine ones, the dry filaments were de-pigmented by suspending in methanol (100%) for 4 h, followed by washing with 1:1 (methanol: water) 3–4 times.

Staining

For staining, dry filaments were placed in a 1% staining solution of each staining agent (toluidine blue, safranin O, iodine, fast green, crystal violet) for 1–2 min. Staining was followed by washing with water to remove excess stain. Filaments were placed on clean glass slides. Filaments were covered with coverslip after putting a drop of water on them and observed under Foldscope. Photographs were taken with a smartphone.

Machine learning

Before using machine learning, we tried a simpler image processing technique. As genuine saffron sample images have papillae on stigma, while being absent in fake saffron samples, we converted all images to single channel image, i.e., gray image, and employed the Canny edge detector multi-stage image processing algorithm to detect the edges in the image. We counted the number of contours on the edges in the image based on the highest gradient difference. Subsequently we used neural networks to fine tune the process.

Neural networks (NNs) are a subset of ML and basic components of the deep learning algorithm. Convolutional neural network (CNN) is a special form of NN that performs better with high-dimensional data like images and videos, and it allows faster training and reduces model complexity. We used deep neural network architecture based on ResNet18 (He et al., 2016) and Densenet121 (Huang et al., 2017) networks, with different model parameters and states. We used Python version 3.7.13 as a programming language, OpenCV version 4.1.2 for image processing, fastai version 1.0.61, backed by PyTorch version 1.11.0, as a deep learning framework, and scikit-learn version 1.0.2 as a machine learning library. We used NVIDIA Tesla K80 GPU with 12 GBs of memory capacity and Intel Xeon 2.20 GHz CPU with 12 GBs of memory capacity. We modified the last layer of networks and used it as a binary classifier with two labels, separating genuine from fake saffron samples.

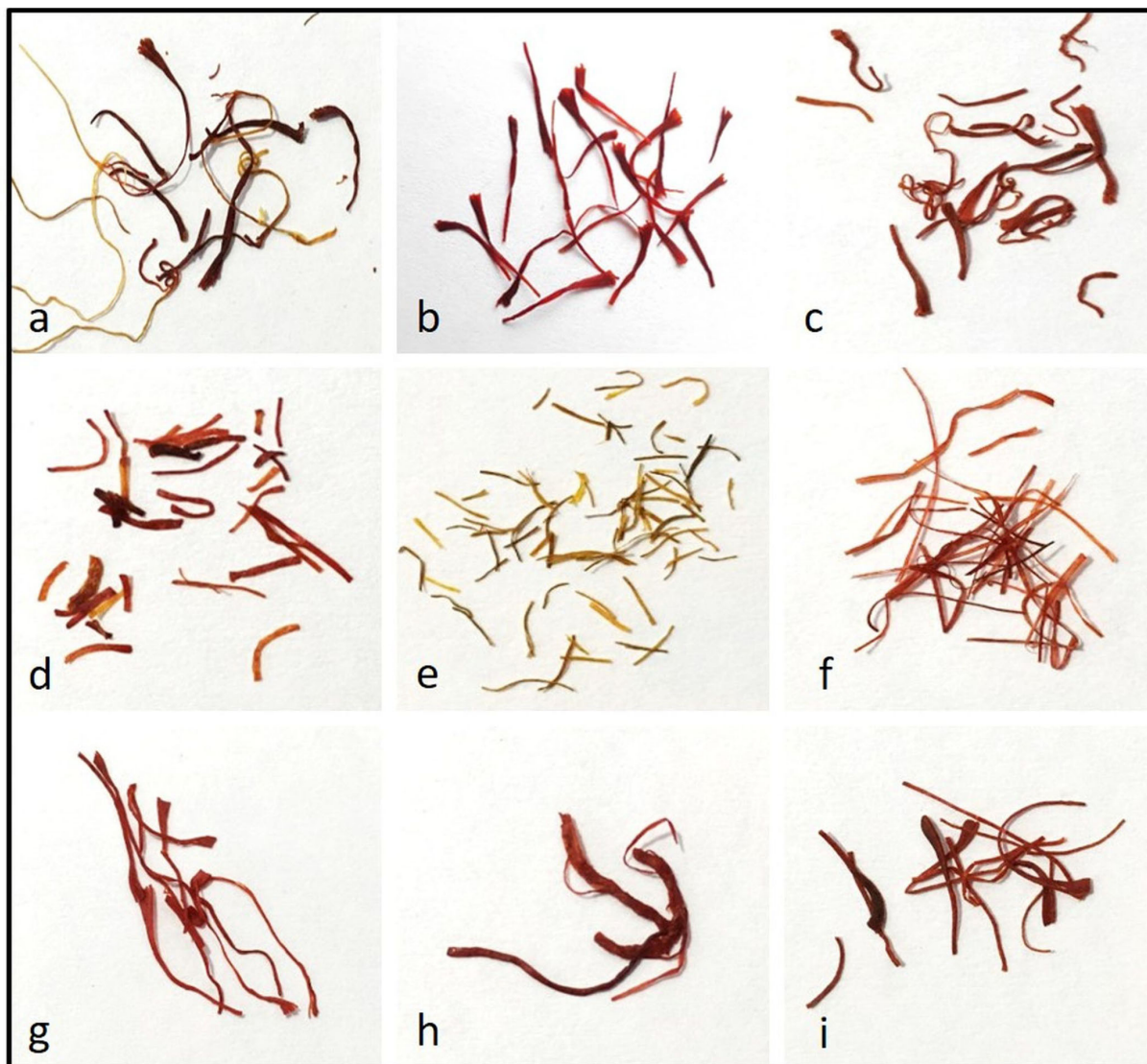


FIGURE 1

Photographs of closely resembling samples as visualized by a naked eye: Genuine saffron (a) saffron in filaments (Laccha), (b) saffron in cut filaments (Mongra); fake samples, (c) unknown look-alike, (d) unknown look-alike, (e) neon yellow dyed corn silk, (f) maroon red dyed corn silk, (g) unknown look-alike, (h) saffron stamen, and (i) paper strips.

Baseline and comparative methods

To ablatively test the introduction of ML algorithm for the classification of fake and genuine saffron samples, we compared NN with the conventional ML algorithms, i.e., random forest and SVM. Evaluation of the different NNs and their variants was also done. For training the random forest model, we used 100 trees in the forest and two as the minimum number of splits for the internal node. The maximum depth of the tree is expanded until all leaves contain less than the minimum number of split samples. Gini impurity was employed to measure the quality of

a split. SVM model was trained using radial basis function (RBF) kernel. During the training process, model leveraged squared hinge loss for the optimization, with 0.001 as the tolerance for the stopping criteria.

Three experiments per network were conducted for deep neural networks, namely ResNet18 and DenseNet121. First, the model weights were randomly initialized using the Kaiming initialization (He et al., 2015) approach. Second, convolutional layers of the deep neural network were initialized with pre-trained ImageNet (Deng et al., 2009) weights. The layers were kept in a frozen state during training, while the last two layers

were initialized with random weights and updated as training proceeded. And finally, all the layers of both the networks with pre-trained weights were unfrozen, and during training, all the weights belonging to the layers got updated based on the loss function optimization.

Random lightning and contrast changes were performed to further enhance the training and generalize the process. Real-time data augmentation of random flip with 50% probability was also applied to training samples. These augmentations help overcome the overfitting of the model on training data and improve the overall model accuracy. Models were trained using an Adam optimizer with a learning rate of 0.001 and momentum of 0.9. Binary cross entropy was used as a loss function due to the binary classification nature of the task, i.e., a sample can be either fake or genuine. All the variations of both neural networks were trained for 15 epochs.

Data acquisition and pre-processing

During the data acquisition phase, we randomly captured multiple images of the same filament for different samples to incorporate the variation that might arise due to clicking the image *via* smartphone from different angles, orientation, contrast, etc. It helps to make the dataset more generalized, to avoid neglecting the possible scenarios in the real-world, where users can click images with uncertainty toward any assumption, which in turn resulted in 3,794 images in total; out of these, 1,434 images belong to genuine samples and 2,360 images to fake samples. The dataset consisted of dried, with, and without stained saffron sample images to include the tolerance toward different processing performed on saffron strands.

In the pre-processing step, each image was resized to $224 \times 244 \times 3$ (width \times height \times color channels) dimension to decrease the computational load and create a uniformly sized dataset. After resizing, the pixel values of images were subtracted by the mean and divided by the standard deviation. This process brings all image pixel values between 0 and 1 range and allows faster convergence later in the model training.

For the experimentation purpose, the saffron dataset was split into a training dataset containing $\sim 72\%$ (2,732 samples), a validation dataset containing $\sim 18\%$ (683 samples), and a test dataset consisting of $\sim 10\%$ (379 samples). The training dataset was used to train the above-mentioned models in a supervised learning fashion, where input was a saffron sample image and the label was the class it belongs to, i.e., fake or genuine. The validation dataset was handy to validate and select the best-trained model in an unbiased manner while finetuning the model hyperparameters. Last but not least, since the test dataset was not used during the training phase, it depicts the real-world behavior and allows for the evaluation of the final model.

Machine learning algorithms for classification

Machine learning algorithms are mainly categorized into unsupervised, supervised, and reinforcement learning. The classification task falls under the supervised learning algorithm, where training takes place based on the pre-labeled data. During training, the algorithm learns the pattern from the labeled data (Veronese et al., 2013). Once trained, the algorithm assigns a new label to the new and unseen data and classifies the sample. There are multiple ML-based classification algorithms available, and, in this work, we explored random forest (Verma and Achutha, 2016), support vector machine (SVM) (Le et al., 2012), and multiple variations of CNN.

Random forest is an ensemble learning method and consists of a number of decision trees. Each decision tree predicts a class associated with the data sample, and the class with the maximum number of votes is taken as a final prediction. This combined approach adds robustness toward errors linked with the individual decision tree predictions.

SVM algorithm takes data samples during training as input and tries to find the optimal hyperplane in an N-dimensional space, where N is the number of input features. This hyperplane is a decision boundary and distinguishes the data samples into different classes.

Performance evaluation

For measuring the performance of all the trained ML-based classification models, accuracy and precision were employed. The classification accuracy was calculated as:

$$accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)} \quad (1)$$

True positive (TP) represents the saffron samples correctly classified as genuine samples in the above equation. False positive (FP) represents the saffron samples falsely classified as genuine samples. True negative (TN) represents the saffron samples correctly classified as fake samples, while false negative (FN) represents the genuine saffron samples falsely classified as fake.

For the evaluation, accuracy focuses on the fraction of the classification of samples, both fake and genuine, corrected as predicted by the model.

The classification precision is represented as:

$$precision = \frac{TP}{TP + FP} \quad (2)$$

Precision performance metric quantifies the number of correctly classified genuine saffron samples by the trained model.

Assessment of adulteration and quality estimation

Eighty crude saffron samples were collected from eight cities/towns of Jammu and Kashmir, India. Ten samples were collected from each city/town, with two packets of 1 g each bought from each vendor. All the 80 samples were then screened for adulteration using Foldscope. The quantum of adulteration in each sample was determined and expressed as adulteration percentage, and then averaged for each location. Spectrophotometer-based quality of the crude and the pure samples was determined to categorize these into Categories I, II, and III as per the ISO 3632.

Results and discussion

Limitations of saffron quality and adulteration detection methods

The quality and the commercial value of saffron are based on an estimation of coloring power, bitter taste, and aroma (Carmona et al., 2007; Kafi et al., 2018). It is certified in the international trade market following the International Organization for Standardization (ISO) 3632 Normative (Husaini et al., 2010b). Regardless of the fraudulent practice, it is challenging to identify commercial frauds in saffron because changes in physical, chemical, and organoleptic characteristics are not easily identifiable (Koocheki and Milani, 2020). Artificial intelligence technique-based artificial neural network and electronic nose have been used for quality control of saffron using its aroma fingerprint and distinguishing it from the samples mixed with safflower or corn stigma up to a proportion of 50% (w/w) (Heidarbeigi et al., 2015). The technique can detect adulterated saffron with a percentage classification accuracy of 86.87%.

An electronic nose is used to determine the geographic origins of saffron with 90% of confidence (Carmona et al., 2006). The principle of this detection is based on the differences in dehydration techniques followed in different countries and the consequent changes in the composition of volatile compounds of saffron. Several fake and original products like sunflower oil, corn oil, sesame oil, tea, and coffee have also been detected using the electronic nose (Hai and Wang, 2006; Mildner-Szkudlarz and Jeleń, 2008; Son et al., 2009). The most significant limitation with the users of such electronic sensors (e-nose and e-tongue) is the requirement of strictly controlling sample preparation, sampling, and data processing. At the same time, training a sensory panel is time-consuming and expensive. Moreover, these sensors are very sensitive to temperature, humidity, pressure, gas velocity, and vapor concentration (Tan and Xu, 2020).

Several studies have combined many techniques and used multiple types of sensors through the fusion technique to overcome the above-discussed limitations, but with limited

success (Kiani et al., 2018). Contrary to the previously discussed methods, our aim is not to develop a method for detecting adulterants or extraneous “powdery material” in the saffron “powder” or the mixing of “different grades” of genuine saffron. Our paper focuses on the “filamentous” adulterants and fictitious “look-alike” versions of saffron commercially sold in markets.

Stigma papillae are the characteristic morphological markers of genuine saffron

The stigma of *C. sativus* consists of three orange-red trumpet-shaped lobes, and it is papillate on the rim, and the average length is 3 cm. *C. sativus* pollen tube growth in intra- and interspecific pollinations has been studied in detail under a microscope (Chichiriccò, 1984; Chichiriccò and Caiola, 1986). The stigma surface of saffron is of the dry type, as in the case of many other *Crocus* species (Heslop-Harrison and Heslop-Harrison, 1975; Heslop-Harrison, 1977; Caiola and Chichiriccò, 1991). While the stigmas of *Crocus sativus* and its allies *C. cartwrightianus*, *C. thomasi*, and *C. hadriaticus* have been studied in detail for reproductive biology (Caiola et al., 2000), there is no emphasis on using it as a distinct morphological marker for identifying genuine saffron, once dried or processed.

In the present study, we first visualized the filaments of sample classes under a stereo-microscope for a wider field of vision. We observed some differences among the filaments, though these were more conspicuous toward their apices (Supplementary Figure S2). We used Foldscope to focus on the apex area and could easily identify papillae in both commercially available genuine saffron sample classes 1 & 2 (Supplementary Figure S3). Our results show that it is possible to distinguish genuine saffron from its commonly used adulterants or fake look-alikes by detecting the presence of distinct papilla on their trumpet-shaped upper rim (Figures 2b–d,f–h).

While the sample classes we studied showed textural differences between real saffron and the fake ones, serrated top margins and/or pollen grains were seen in authentic as well as spurious sample classes (3, 8) (Supplementary Figure S1). Therefore, unlike the common belief of saffron vendors, the presence of pollen is not a distinctive feature of genuine saffron. Sample class 3 (Supplementary Figure S3c) and sample class 7 (Supplementary Figure S3g) showed serrated top margins, a feature common with authentic saffron; however, both lacked the distinct finger-like projections “papillae.” Similarly, while a large number of pollen grains were seen in sample class 8 and pollen-like granules in sample 3, the papillae were absent in both. Sample classes 4, 5, and 6 (Supplementary Figures S3d–f) featured smooth margins with no serrations, papillae, or pollen grains. Sample class 9 showed smooth, wide top margins with narrower stalk, but lacked the characteristic papillae.

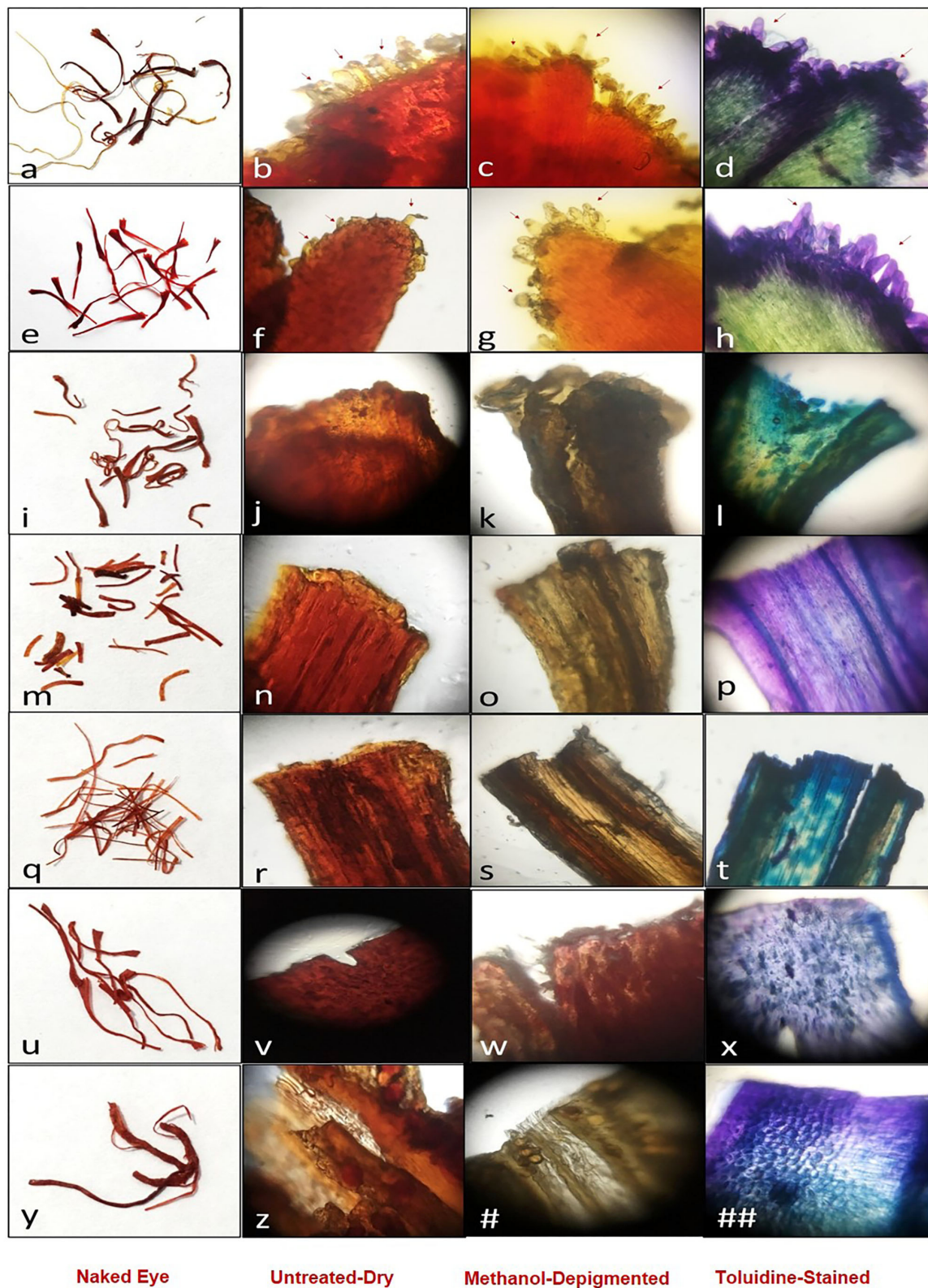


FIGURE 2

Poster for the identification of genuine/fake saffron: Genuine saffron (a–d) saffron in filament (Laccha), (e–h) saffron in cut filament (Mongra); fake samples (i–##).

Pollen grains were also absent. Overall, sample classes 1, 2, 3, 7, and 9 closely resembled trumpet/funnel-like structure, typical of most stigma, while the remaining four sample classes showed flattened top and margins. The results are summarized in Table 1.

Creating a poster by de-pigmentation and differential staining

Microscopy is generally used in combination with staining, particularly by school children. Most dyes stain tissues with differing intensities of the same color. However, certain basic dyes stain tissue components with colors other than that of the dye. Such a staining reaction is called metachromasy and is highly selective. Only certain tissue structures stain metachromatically and are said to exhibit metachromasia (Culling et al., 2014).

In the present study, we used staining to further expand the inventory of distinct visual color markers. Toluidine blue (also known as toloum chloride, methylaniline, or aminotoluene) is used to specifically stain certain components of mucosal lesions and tissue sections owing to its metachromatic property and was first applied for *in vivo* staining of uterine cervical carcinoma *in situ* by Reichart in 1963 (Siddiqui et al., 2006). Inspired by this, we used it to stain filaments, which incidentally showed a distinct differential staining pattern in the case of “genuine” saffron. It stained the papillae and the margins deep purple, while it is lighter yellowish green toward the central axis (Figure 2).

Moreover, when we try to remove the color of filaments, the de-pigmentation of genuine saffron requires more extensive washing than the samples which had been artificially dyed (Figures 2c,g,k,o,s,w,#). Genuine saffron retains most of the color despite washing 3–4-times. These significant findings were used to create a poster showing all the major visual features that a person can use for the manual validation of a given sample (Figure 2).

Toluidine blue imparts differential stain only to saffron stigma and not fake look-alikes

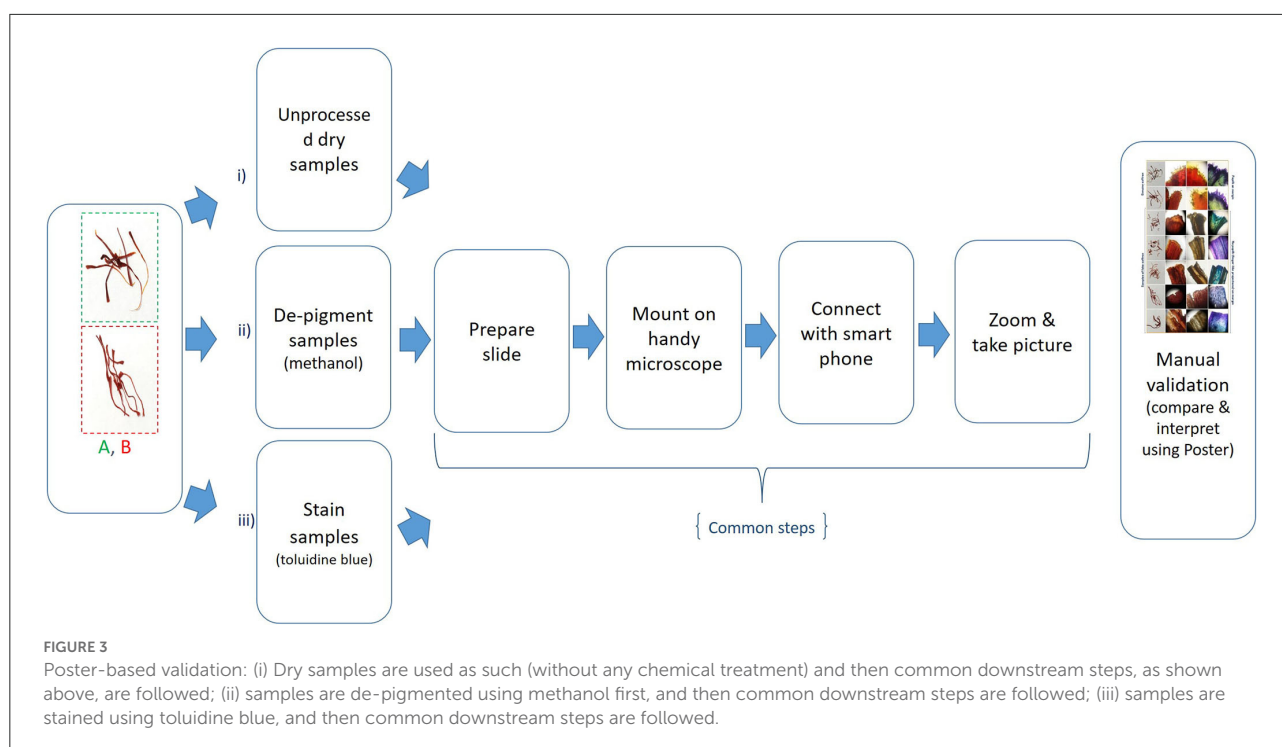
The staining of samples with toluidine blue in the present study clearly shows the distinct staining pattern based on the histology of papillae (Figure 2). Toluidine blue is partially soluble in both water and alcohol, and selectively stains acidic tissue components like sulfates, carboxylates, and phosphate radicals (Epstein et al., 1992; Gandolfo et al., 2006). Because of these properties, the differences in color intensity of the papillae,

TABLE 1 Distinctive morphological features in saffron and look-alikes as visualized using Foldscope.

Sample Class and description	Characteristic features
Class 1: “Saffron in filaments” (ISO3632)/Laccha (IS5453)	Serrated distal trumpet shaped top with distinct papillae; striated texture dotted with pit like structures; presence/absence of large pollen grains.
Class 2: “Saffron in cut filaments” (ISO3632)/Saffron processed using a technique indigenous to Kashmir “Mongra” (IS5453)	Serrated distal trumpet shaped top with distinct papillae; presence/absence of large pollen grains.
Class 3: Fake-1 “Unknown look-alike”	Serrated top exhibiting typical trumpet shaped structure; epidermal papillose protuberances absent; pollen like granules present. It may be stigma of another flower.
Class 4: Fake-2 “Unknown look-alike”	Serrations as well as other distinguishing features of authentic saffron absent. Long flattened structure with smooth margins and even width.
Class 5: Fake-3 “Corn silk dyed maroon”	Resembles stigma lobes; Smooth margins with no distal serrations or papillose protuberances. Long flattened structure with even width.
Class 6: Fake-4 “Corn silk dyed neon yellow”	Resembles style; Serrations as well as the papillose protuberances absent; Long flat structure with smooth margins and even width; pollen grains absent.
Class 7: Fake-5 “Unknown look-alike”	Wide and serrated top margins with narrow stalk resembling funnel. Papillose protuberances and texture distinctive of saffron stigma absent; Pollen grains absent. The sample structure disintegrates quickly in solvents and reveals a single thin, long thread like fiber.
Class 8: Fake-6 “Dyed saffron stamen”	Top margin wide and trumpet shaped; Large number of pollen grains present along the stalk and top regions; Papillose protuberances absent.
Class 9: Fake-7 “Dyed paper strips”	Smooth and flat, funnel shaped wide top margin with a narrow stalk; no serrations; It disintegrates quickly in solvents and reveals a single thin, long thread like fiber.

their base, and the tissue toward the central axis of saffron stigma are well depicted in toluidine blue staining.

It is known that saffron papillae possess a thick cell wall, covered with a continuous cuticle under which electron-dense material is visible. The papillae contain a large central vacuole, a scarce endoplasmic reticulum, numerous mitochondria and chromoplasts, and virus-like inclusions at the base (Caiola and



Chichiriccò, 1991; Caiola et al., 2000). These features are absent in the artificially created fake look-alikes and therefore, get stained uniformly across the whole tissue.

As a metachromatic dye, toluidine absorbs light at different wavelengths, varying with concentration and surroundings, and can change its color without changing its chemical structure. This color change is brought about by the specialized physical changes in the form of stacking of dye cations at regions of the high density of anionic groups in the tissue. Stacking causes a hypsochromic shift (shortens the wavelength of maximum absorption) so that the maximum wavelength of the transmitted light is longer, which makes the observed color look different (Kumar and Kiernan, 2010). The color shift in Figure 2 from a blue or violet dye to a greenish-yellow could represent the polymerization of the dye to varying degrees. The tissue at the margin may have an absorption maximum at 630 nm due to its orthochromatic nature, which, therefore, stains blue. In contrast, the inside tissue stains greenish-yellow, and its absorption spectrum may be closer to 540 nm (Culling et al., 2014). However, these differences are not noticed when the samples are stained with safranin O/iodine/fast green/crystal violet, and all the other staining agents show more or less uniform staining patterns. Similarly, the artificial dyes used by fraudsters to dye the look-alikes of saffron cannot generate the differential pattern shown by toluidine blue.

Based on the above findings, a method was developed for the manual validation of a given sample, and the workflow is shown in Figure 3.

Machine learning-based approach is quick and robust in detecting fake saffron

Several studies have been conducted on artificial intelligence for identification and classification tasks. However, only a few relevant contributions employ “image classification” in plants. Kurtulmuş et al. (2016) demonstrated using a neural network to classify pepper seed variety based on images. Likewise, Islam et al. (2020) performed flower classification by employing a convolutional neural network on eight different types of flowers and achieved 85% accuracy. The approaches mentioned above focused on identifying the different variety of spices and flowers, while in the present study, we explored end-to-end neural network learning to distinguish the genuine saffron from the fake using image of the sample.

Even when the presence/absence of papillae in genuine/fake saffron is distinct, the approach based on the Canny edge detector multi-stage image processing algorithm did not perform well (Rong et al., 2014). This is due to many factors, like using only gray images (which is needed for edge detection and finding contours) and throwing a lot of information about color schema, texture, etc., and uncertainty in deciding the threshold value of contours, contours created by other structures other than papillae. Moreover, in the machine learning-based approach, we did not process images, and the model learned from all the features available in the sample images, which allowed us to get robustness in performance and much more accurate results (Janiesch et al., 2021; Greener et al., 2022).

TABLE 2 Quantitative comparison of ML-based classification models for saffron classification based on accuracy and precision.

Model	Training precision	Training accuracy	Validation precision	Validation accuracy	Test precision	Test accuracy
Random Forest	1.0	1.0	0.94	0.94	0.95	0.95
SVC	0.95	0.95	0.90	0.90	0.95	0.95
Resnet18 with random weight initialization	0.96	0.93	0.91	0.91	0.93	0.92
Resnet18 Pretrained—unfreeze	0.97	0.96	0.93	0.92	0.95	0.93
Resnet18 pretrained	0.99	0.99	0.99	0.99	0.99	0.99
Densenet121 with random weight initialization	0.96	0.95	0.95	0.93	0.97	0.93
Densenet121 Pretrained—unfreeze	0.95	0.94	0.91	0.89	0.94	0.90
Densenet121 pretrained	0.99	0.99	0.99	0.99	0.99	0.99

We compared all the ML-based saffron classification models on validation and test datasets (Table 2). We observed that the deep neural networks (ResNet18 and DenseNet121) performed better than random forest (RF) and support vector machine (SVM)-based approaches (Kremic and Subasi, 2016; Speiser et al., 2019; Nandhini and Ashokkumar, 2022; Zhou et al., 2022). Also, the best models (two in total) out of the three variations per deep neural network are the models using the pre-trained model weights and freezing all the layers, apart from the last two layers. Both the models recorded 99.5% accuracy with a precision of 99.3% on the test dataset, which is pretty decent. The models show 99.5% accuracy and precision of 99.1% on the validation dataset. This shows that the models are generalized and behave almost the same on the validation and the test dataset. Further, this approach has accelerated the decision-making process regarding the genuineness of a sample image as it takes less than a second per photograph.

Figure 4 illustrates the contrastive results based on the confusion matrix. We observed that the deep neural network-based approaches (ResNet18 and DenseNet121 with pre-trained model weights and freezing all the layers, apart from the last two layers) demonstrate maximal performance with 235 fake samples correctly detected out of 236, and 142 rightly predicted as genuine out of 143 genuine saffron samples on the test dataset. Only one sample from fake and genuine was wrongly classified, as highlighted (Figure 4).

The machine learning-based classification approach automated and simplified the process to make detecting fake/genuine saffron quicker. The workflow for the process is shown in Figure 5.

Adulteration and quality estimation

Eighty market samples (1g each) were used to assess adulteration and quality. The geographical coordinates and the locations of the eight cities/towns from where these samples were procured are shown using the ESRI ArcGIS map (Figure 6A). The names of the sites and the localities are: (1) Budgam (Chadoora, Budgam, Magam, Beerwah, Humhama), (2) Jammu city (Trikuta Nagar, Gandhi Nagar, Raghunath Bazar, Janipur, Chani Himat), (3) Kangan (Kangan town, Dursuma, Wussan, Preng, Cherwan), (4) Katra (Katra town, Dhar Vaishno Devi, Akhli, Bhargal, Arli, Hansali), (5) Kishtwar (Poochal, Matta, Janwas, Dool, Ohli), (6) Pahalgam (Pahalgam town, Ashmukam, Salar, Dirhama, Batkoot), (7) Pulwama (Lethpora, Pampore, Awantipora, Namalabal, Konibal), and (8) Srinagar (Lalchowk, Dalgate, Sonwar, Dargah, Nowhatta).

Out of a total of 80 market samples, the number of adulterated samples was 48 (Figure 6B). Twenty samples from Pulwama and Srinagar showed a cumulative average adulteration of <1.5%, while the 30 samples from Budgam, Kishtwar, and Jammu city showed around 16%. The samples (30) from Katra, Pahalgam, and Kangan were highly adulterated, with an average of 40% adulteration. The saffron bought from Pulwama and Srinagar showed adulteration between 0.12 and 2.36% (among the 20 samples), and for the quality, they fall into Categories I and II of ISO 3632 (1 & 2) standards (Figure 6B).

The overall percentage of saffron samples which belonged to Categories II and III is 60% (Figure 6B). While the adulteration among these adulterated samples ranged from 2.09 to 71.23%, and their mean adulteration percentage was

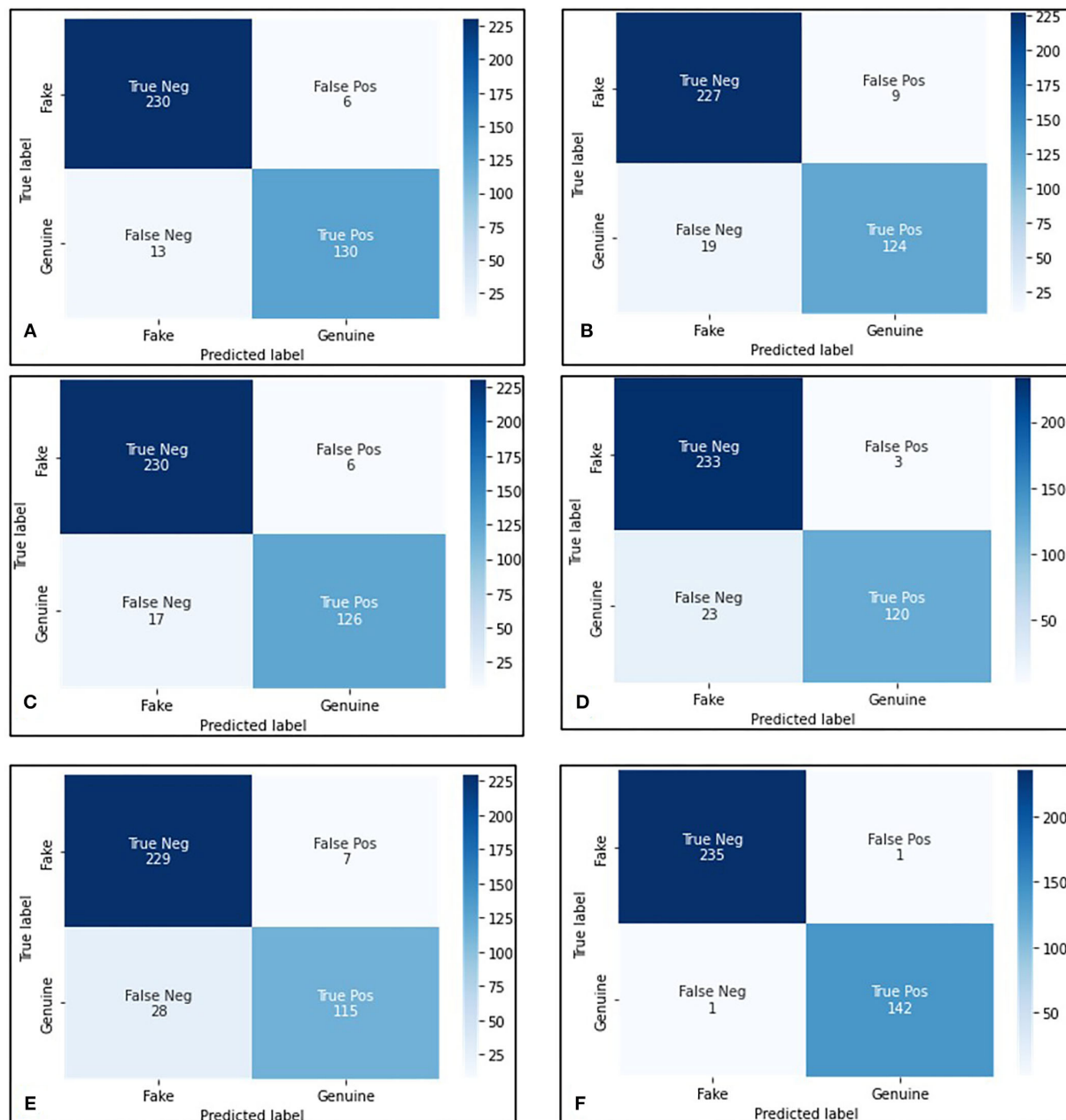
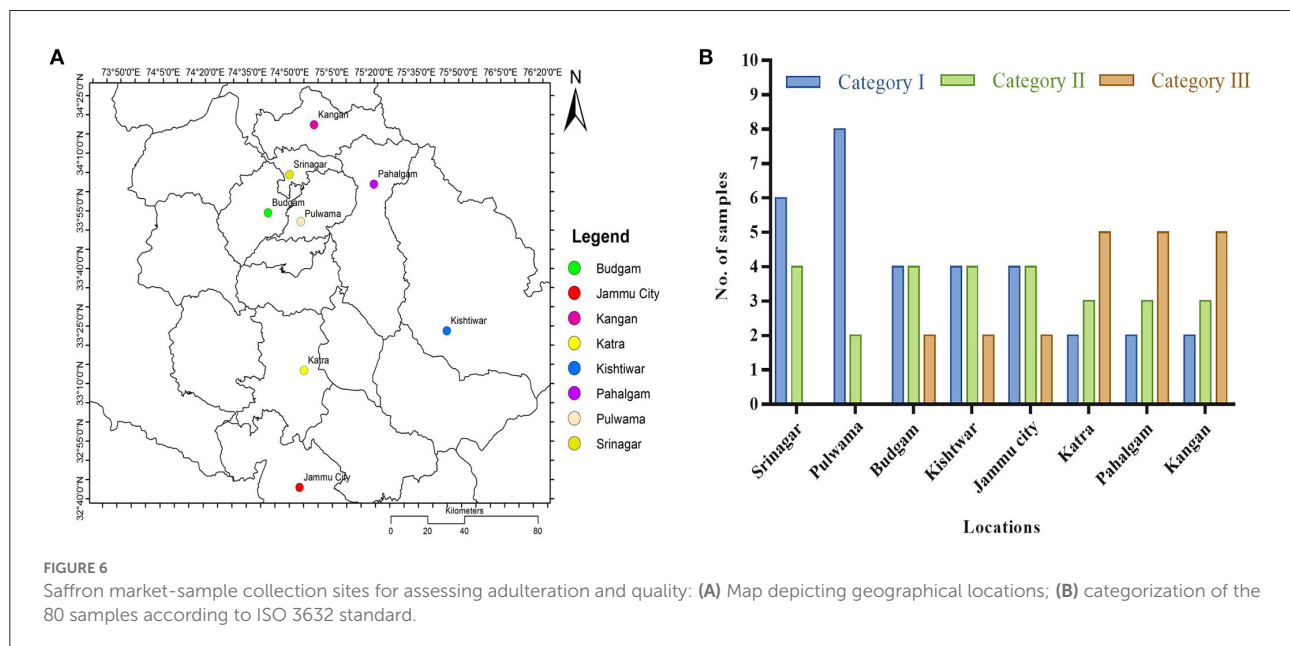
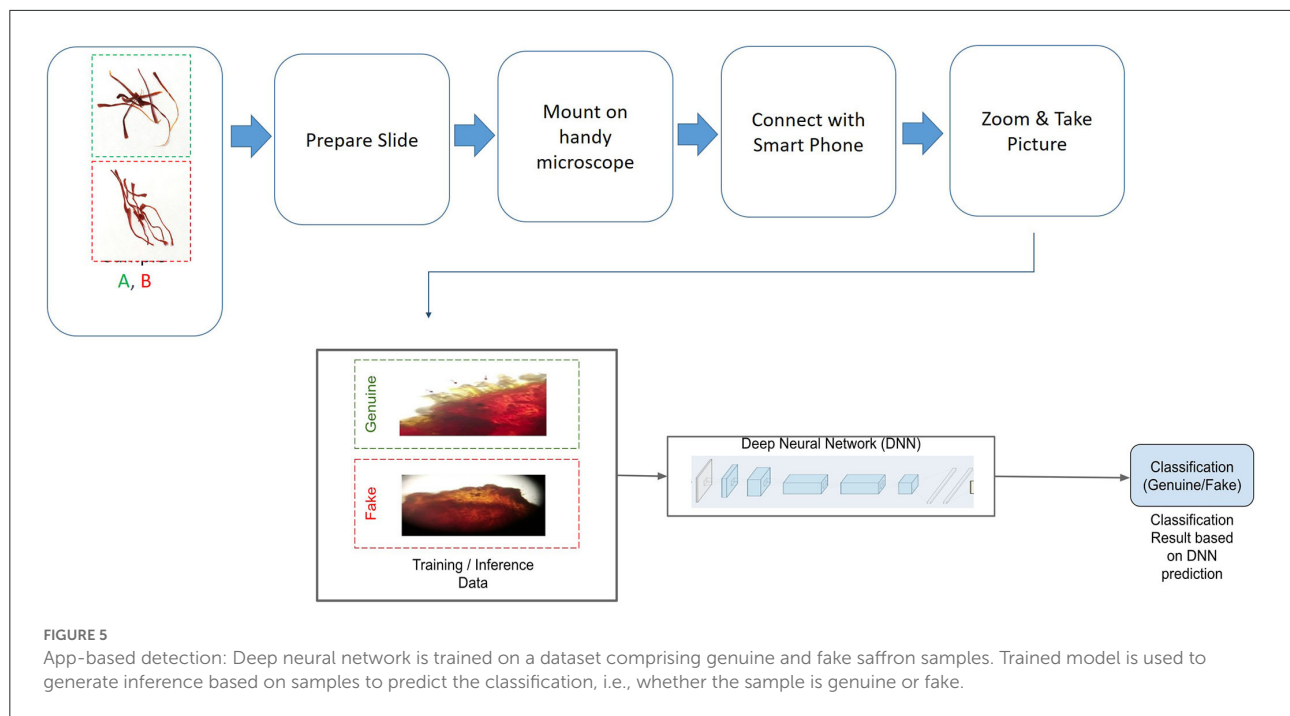


FIGURE 4
Quantitative analysis based on the test dataset. Confusion Matrix for: (A) random forestand SVM, (B) ResNet18 with random weight initialization, (C) ResNet18 pre-trained—network is not in freeze state, (D) DenseNet121 with random weight initialization, (E) DenseNet121 pre-trained—network is not in freeze state, and (F) ReNet18 and DenseNet121 pre-trained network, with all the layers in freeze state, but the last two layers.

36.25%. Surprisingly, adulteration strongly correlated with the location of sample procurement. Srinagar and Pulwama showed minimal adulteration problems, perhaps because of stricter law enforcement agencies or more awareness among the sellers and buyers. This location-dependent adulteration shows that fraudsters know that cheating would go undetected at places

where demand is more due to the tourist rush while its supply is limited.

While the saffron quality depends on many factors (Husaini, 2014), a significant reduction in the quality of color (crocin), bitterness (picrocrocin), and aroma (safranal) was recorded in the adulterated samples which are inversely proportional



to the adulteration percentage (Figure 7). After removing the adulterant from the crude market sample, the sample quality improved significantly, pushing some from Category III to Category II (Figure 7B). It infers that Categories II and III saffron are more likely to be adulterated by fraudsters (Figures 6B, 7A).

Practical applications

Saffron is a costly spice used as a routine in common people's religious rituals and local cuisines (Husaini and Wani, 2020). However, it is evident from the above data that saffron adulteration and fraud are a big menace. People prefer to buy

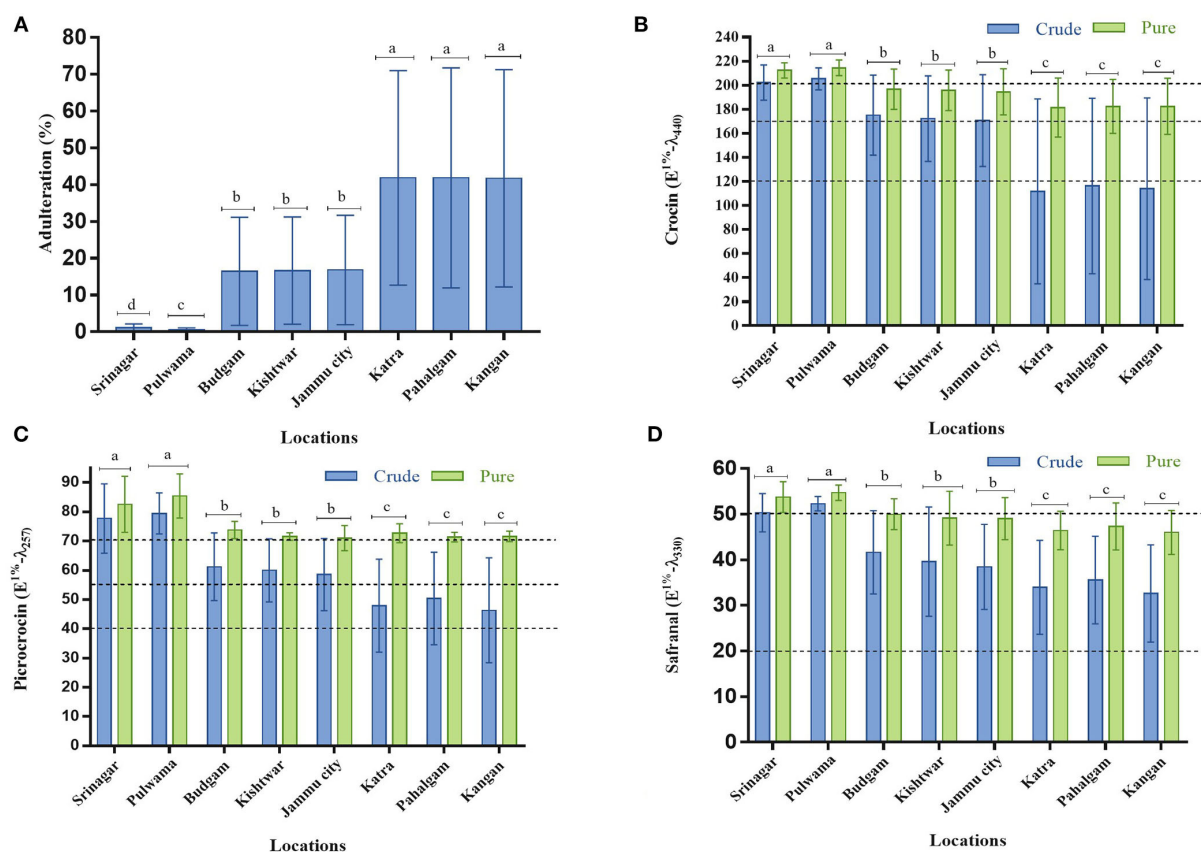


FIGURE 7

Assessment of saffron quality obtained from different locations of Jammu and Kashmir (India) in pure and adulterated samples: (A) Adulteration percent, (B) crocin, (C) picrocrocin, and (D) safranal content. The dotted lines represent the maximum and minimum values of crocin, picrocrocin, and safranal for the categorization of saffron quality as per ISO 3632 standards. Values are the mean of 10 replicates and expressed as mean \pm S.D (standard deviation). The letters a, b and c indicate a statistically significant difference at $p \leq 0.05$ probability level between different locations. Bars with no common letters are significantly different ($p \leq 0.05$).

it as “filaments” because it is easier to use from the dosage point of view, and there are lesser chances of adulteration than in the powdered form. However, unfortunately, some fraudsters have even found ways to “create” fake saffron-like filaments.

Saffron dealers can use the two methods developed in the present study (Figures 3, 6) to showcase the authenticity of saffron to their customers without much botheration. The customer can himself check the adulteration percentage in a random sample and identify fake saffron by using a simple application on a mobile phone. It would act as an additional check for the fraudsters who manage to get fake GI tags, holograms, and certifications and sell their products in the unorganized sector while going undetected by the law enforcement agencies (Husaini et al., 2010a). Further, genuine retailers can convince the customers about the authenticity of their products by showing these visual markers using Foldscope or by installing the mobile app on their smartphone, thereby promoting genuine retail business.

Furthermore, we created a kit for commercial use and quality control laboratories (the patent is under process). The kit constitutes a poster showing distinct and unique markers in saffron stigma, a Foldscope, slides, cello-tape, methanol, and toluidine solution. These methods have the potential to be put to use in the European Science Foundation-sponsored COST Action FA1101 (Saffron-omics: Omics technologies for crop improvement, traceability, determination of authenticity, adulteration, and origin). The results of the de-pigmentation and the staining procedures can be used to update the relevant sections of “test methods” in the identification test and microscopic examination of the International Standards Organization ISO 3632 (1 & 2).

Conclusion

Individual consumers prefer to buy “saffron in filaments” (Laccha), or “saffron in cut filaments” (Mongra) because it is

easier to use from the dosage point of view, and there are lesser chances of adulteration than in the powdered form. However, fraudsters have invented fake saffron-like filaments to cheat customers. We developed two Foldscope-based techniques for the identification of pure saffron. One technique uses a simple visual comparison of distinct markers (papillae of stigma) with a poster, and the other uses an automated approach through a mobile application. Machine learning simplifies the process and automates the detection of fake/genuine saffron samples. It enables end-users not to worry about identifying samples from magnified images themselves. This approach accelerates the whole identification process and takes less than a second per sample after acquiring its image. While large-scale testing of saffron quality using sophisticated methods in specialized laboratories shall always be required for the saffron industry, we have added a new dimension by bringing the customer to the forefront.

Adulteration remains a significant challenge to the saffron industry. The quality of a saffron sample decreases significantly with an increase in adulteration percentage. A critical observation of the present study is that the ISO Category I saffron is not subjected to adulteration, perhaps because elite customers are willing to pay higher prices and generally buy from trusted sources. Saffron belonging to the ISO Categories II and III is more prone to adulteration as it is available freely everywhere as innocent customers prefer to buy saffron at cheaper rates from untrusted sources!

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

AH conceived the idea, designed, directed, performed, and coordinated the research experiments with the active participation of SH, AS, and AW. MD planned and performed machine learning to analyze and automate the process. AH and MD wrote the paper together. All authors contributed to the article and approved the submitted version.

References

- Ahmed, S. A., and Husaini, A. M. (2021). Investigating binding potential of carotenoid pathway bioactive molecules for ACE2 receptor of SARS-CoV-2: possibility of a saffron based remedy for novel coronavirus! *J. Hort. Postharvest Res.* 4, 69–78. doi: 10.22077/jhpr.2021.4462.1224
- Alonso, G., Salinas, M., and Garijo, J. (1998). Method to determine the authenticity of aroma of saffron (*Crocus sativus* L.). *J. Food Prot.* 61, 1525–1528. doi: 10.4315/0362-028X-61.11.1525

Acknowledgments

AH is thankful to DBT Government of India for the financial support. Furthermore, AH is highly grateful SKUAST-K administration, and the staff of the GE&SB Lab (namely Muneeza, Asma, Huma, Amir), Dr. Anup Raj, Faculty of Forestry, SKUAST-K and Dr. Manu Prakash, University of Stanford, for their support and cooperation.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.945291/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

Pollen on (a) freshly cut saffron stigma; (b) sample class 1; (c) sample class 2; (d) sample class 3; (e) sample class 8; (f) saffron stigma mounted on a slide in Foldscope.

SUPPLEMENTARY FIGURE 2

Stereomicroscopic images of the apices of filament sample classes at a magnification of 80× (a–i).

SUPPLEMENTARY FIGURE 3

Foldscope images of top region/margin of (a) sample class 1 showing papillae; (b) sample class 2 showing papillae; (c) sample class 3 showing serrations; (d) sample class 4 showing smooth margin with no serrations; (e) sample class 5 showing even margin without serrations; (f) sample class 6 showing smooth edges without serrations; (g) sample class 7; (h) sample class 8; (i) sample class 9 showing smooth surface.

- Babaei, S., Talebi, M., and Bahar, M. (2014). Developing an SCAR and ITS reliable multiplex PCR-based assay for safflower adulterant detection in saffron samples. *Food Control* 35, 323–328. doi: 10.1016/j.foodcont.2013.07.019
- Ben Ayed, R., and Hanana, M. (2021). Artificial intelligence to improve the food and agriculture sector. *J. Food Qual.* 2021, 5584754. doi: 10.1155/2021/5584754
- Caiola, M. G., and Chichiricò, G. (1991). Structural organization of the pistil in saffron (*Crocus sativus* L.). *Isr. J. Plant Sci.* 40, 199–207.
- Caiola, M. G., Di Somma, D., and Lauretti, P. (2000). Comparative study of pollen and pistil in *Crocus sativus* L. (Iridaceae) and allied species. *Annali Di Botanica* 58. doi: 10.4462/annbotrm-9065
- Carmona, M., Martínez, J., Zalacain, A., Rodríguez-Mendez, M. L., de Saja, J. A., and Alonso, G. L. (2006). Analysis of saffron volatile fraction by TD–GC–MS and e-nose. *Eur. Food Res. Technol.* 223, 96–101. doi: 10.1007/s00217-005-0144-5
- Carmona, M., Zalacain, A., Salinas, M., and Alonso, G. (2007). A new approach to saffron aroma. *Crit. Rev. Food Sci. Nutr.* 47, 145–159. doi: 10.1080/10408390600626511
- Chichiricò, G. (1984). Karyotype and meiotic behaviour of the triploid *Crocus sativus* L. *Caryologia* 37, 233–239. doi: 10.1080/00087114.1984.10797702
- Chichiricò, G., and Caiola, M. G. (1986). *Crocus sativus* pollen germination and pollen tube growth in vitro and after intraspecific and interspecific pollination. *Can. J. Bot.* 64, 2774–2777. doi: 10.1139/b86-369
- Culling, C. F. A., Allison, R., and Barr, W. (2014). *Cellular Pathology Technique*. Butterworth-Heinemann: Elsevier.
- Cybulski, J. S., Clements, J., and Prakash, M. (2014). Foldscope: origami-based paper microscope. *PLoS ONE* 9, e98781. doi: 10.1371/journal.pone.0098781
- Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., Fei-Fei, L. (2009). “ImageNet: a large-scale, hierarchical, image, database,” in: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 248–255.
- Epstein, J. B., Scully, C., and Spinelli, J. (1992). Toluidine blue and Lugol's iodine application in the assessment of oral malignant disease and lesions at risk of malignancy. *J. Oral Pathol. Med.* 21, 160–163. doi: 10.1111/j.1600-0714.1992.tb00094.x
- Er, S. V., Eksi-Kocak, H., Yetim, H., and Boyaci, I. H. (2017). Novel spectroscopic method for determination and quantification of saffron adulteration. *Food Anal. Methods* 10, 1547–1555. doi: 10.1007/s12161-016-0710-4
- Gandolfo, S., Pentenero, M., Brocchetto, R., Pagano, M., Carrozzo, M., and Scully, C. (2006). Toluidine blue uptake in potentially malignant oral lesions in vivo: clinical and histological assessment. *Oral Oncol.* 42, 88–94. doi: 10.1016/j.oraloncology.2005.06.016
- Greener, J. G., Kandathil, S. M., Moffat, L., and Jones, D. T. (2022). A guide to machine learning for biologists. *Nat. Rev. Mol. Cell Biol.* 23, 40–55. doi: 10.1038/s41580-021-00407-0
- Haghighi, B., Feizy, J., and Kakhki, A. H. (2007). LC determination of adulterated saffron prepared by adding styles colored with some natural colorants. *Chromatographia* 66, 325–332. doi: 10.1365/s10337-007-0321-8
- Hai, Z., and Wang, J. (2006). Detection of adulteration in camellia seed oil and sesame oil using an electronic nose. *Eur. J. Lipid Sci. Technol.* 108, 116–124. doi: 10.1002/ejlt.200501224
- He, K., Zhang, X., Ren, S., and Sun, J. (2015). “Delving deep into rectifiers: Surpassing human-level performance on imagenet classification,” in *Proceedings of the IEEE International Conference on Computer Vision (IEEE)*, 1026–1034. Available online at: http://openaccess.thecvf.com/content_iccv_2015/html/He_Delving_Deep_into_ICCV_2015_paper.html
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). “Deep residual learning for image recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (IEEE)*, 770–778. Available online at: http://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html
- Heidarbeigi, K., Mohtasebi, S. S., Foroughirad, A., Ghasemi-Varnamkhasti, M., Rafiee, S., and Rezaei, K. (2015). Detection of adulteration in saffron samples using electronic nose. *Int. J. Food Propert.* 18, 1391–1401. doi: 10.1080/10942912.2014.915850
- Heslop-Harrison, J., and Heslop-Harrison, Y. (1975). Fine structure of the stigmatic papilla of *Crocus*. *Micron* (1969) 6, 45–52. doi: 10.1016/0047-7206(75)90006-0
- Heslop-Harrison, Y. (1977). The pollen-stigma interaction: pollen-tube penetration in *Crocus*. *Ann. Bot.* 41, 913–922. doi: 10.1093/oxfordjournals.aob.a085387
- Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q. (2017). “Densely connected convolutional networks,” in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4700–4708. doi: 10.1109/CVPR.2017.243
- Husaini, A., Bhat, M., Kamili, A., and Mir, M. (2013). Kashmir saffron in crisis. *Curr. Sci.* 104, 686–687.
- Husaini, A. M. (2014). Challenges of climate change: omics-based biology of saffron plants and organic agricultural biotechnology for sustainable saffron production. *GM Crops Food* 5, 97–105. doi: 10.4161/gmcr.29436
- Husaini, A. M., Hassan, B., Ghani, M. Y., Teixeira da Silva, J., and Kirmani, N. A. (2010a). Saffron (*Crocus sativus* Kashmirianus) cultivation in Kashmir: practices and problems. *Funct. Plant Sci. Biotechnol.* 4, 108–115. Available online at: [http://www.globalsciencebooks.info/Online/GSBOOnline/images/2010/FPSB_4\(SI2\)/FPSB_4\(SI2\)108-115o.pdf](http://www.globalsciencebooks.info/Online/GSBOOnline/images/2010/FPSB_4(SI2)/FPSB_4(SI2)108-115o.pdf)
- Husaini, A. M., Jan, K. N., and Wani, G. A. (2021). Saffron: a potential drug-supplement for severe acute respiratory syndrome coronavirus (COVID) management. *Heliyon* 7, e07068. doi: 10.1016/j.heliyon.2021.e07068
- Husaini, A. M., Kamili, A. N., Wani, M., Teixeira da Silva, J., and Bhat, G. (2010b). Sustainable saffron (*Crocus sativus* Kashmirianus) production: technological and policy interventions for Kashmir. *Funct. Plant Sci. Biotechnol.* 4, 116–127. Available online at: [http://www.globalsciencebooks.info/Online/GSBOOnline/images/2010/FPSB_4\(SI2\)/FPSB_4\(SI2\)116-127o.pdf](http://www.globalsciencebooks.info/Online/GSBOOnline/images/2010/FPSB_4(SI2)/FPSB_4(SI2)116-127o.pdf)
- Husaini, A. M., Ul Haq, S. A., and Jiménez, A. J. L. (2022). Understanding saffron biology using omics- and bioinformatics tools: stepping towards a better *Crocus* phenome. *Mol. Biol. Rep.* 49, 5325–5340. doi: 10.1007/s11033-021-07053-x
- Husaini, A. M., and Wani, A. B. (2020). Prospects of organic saffron kitchen gardens as a source of phytochemicals for boosting immunity in common households of semi-arid regions: a case study of trans-Himalayan Kashmir valley. *J. Pharmacogn. Phytochem.* 9, 237–243. doi: 10.22271/phyto.2020.v9.i6d.12889
- Islam, S., Foyzal, M. F. A., and Jahan, N. (2020). “A computer vision approach to classify local flower using convolutional neural network,” in: *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)* (IEEE), 1200–1204. doi: 10.1109/ICICCS48265.2020.9121143
- ISO. (2010). *ISO International Standard ISO 3632-2: Saffron (Crocus sativus L.) Test Methods*. Geneva: International Organization for Standardization. Available online at: <https://www.iso.org/standard/44526.html>
- ISO. (2011). *ISO International Standard ISO 3632-1: Saffron (Crocus sativus L.) Specification*. Geneva: International Organization for Standardization. Available online at: <https://www.iso.org/standard/44523.html>
- Janiesch, C., Zschech, P., and Heinrich, K. (2021). Machine learning and deep learning. *Electron. Mark.* 31, 685–695. doi: 10.1007/s12525-021-00475-2
- Javanmardi, N., Bagheri, A., Moshtaghi, N., Sharifi, A., and Hemati Kakhki, A. (2011). Identification of Safflower as a fraud in commercial Saffron using RAPD/SCAR marker. *J. Cell Mol. Res.* 3, 31–37. Available online at: <https://www.sid.ir/en/Journal/ViewPaper.aspx?ID=249430>
- Joshi, N., and Bhosale, S. (2018). Interpretation of physical properties like crystallinity of maize starch powder effectively by Foldscope. *Int. J. Pharm. Drug Anal.* 6, 592–598.
- Kafi, M., Kamili, A. N., Husaini, A. M., Ozturk, M., and Altay, V. (2018). “An expensive spice saffron (*Crocus sativus* L.): A case study from Kashmir, Iran, and Turkey,” in *Global Perspectives on Underutilized Crops*, eds M. Ozturk, K. Hakeem, M. Ashraf and M. Ahmad (Cham: Springer), 109–149. doi: 10.1007/978-3-319-77776-4_4
- Kiani, S., Minaei, S., and Ghasemi-Varnamkhasti, M. (2018). Instrumental approaches and innovative systems for saffron quality assessment. *J. Food Eng.* 216, 1–10. doi: 10.1016/j.jfoodeng.2017.06.022
- Koocheki, A., and Milani, E. (2020). “Saffron adulteration,” in *Saffron*, eds A. Koocheki and M. Khajeh-Hosseini (Sawston: Cambridge: Woodhead Publishing), 321–334.
- Kremic, E., and Subasi, A. (2016). Performance of random forest and SVM in face recognition. *Int. Arab J. Inf. Technol.* 13, 287–293. Available online at: <http://iajit.org/PDF/Vol.13,%20No.2/8468.pdf>
- Kumar, G., and Kiernan, J. A. (2010). Special stains and H and E second edition education guide| SPECIAL STAINS and H and E. *Dako North Am. Carpint. Calif.* 14, 1–158.
- Kumari, L., Jaiswal, P., and Tripathy, S. S. (2021). Various techniques useful for determination of adulterants in valuable saffron: a review. *Trends Food Sci. Technol.* 111, 301–321. doi: 10.1016/j.tifs.2021.02.061
- Kurtulmuş, F., Alibaş, I., and Kavdir, I. (2016). Classification of pepper seeds using machine vision based on neural network. *Int. J. Agric. Biol. Eng.* 9, 51–62. Available online at: <http://www.ijabe.org/index.php/ijabe/article/view/1790>
- Le, T. H., Tran, H. S., and Nguyen, T. T. (2012). “Applying multi support vector machine for flower image classification,” in *International Conference on Context-Aware Systems and Applications* (Springer), 268–281.

- Lozano, P., Castellar, M., Simancas, M., and Iborra, J. (1999). A quantitative high-performance liquid chromatographic method to analyse commercial saffron (*Crocus sativus* L.) products. *J. Chromatogr. A* 830, 477–483. doi: 10.1016/S0021-9673(98)00938-8
- Ma, X. Q., Zhu, D., Li, S. P., Dong, T., and Tsim, K. W. K. (2001). Authentic identification of stigma Croci (stigma of *Crocus sativus*) from its adulterants by molecular genetic analysis. *Planta Med.* 67, 183–186. doi: 10.1055/s-2001-11533
- Marieschi, M., Torelli, A., and Bruni, R. (2012). Quality control of saffron (*Crocus sativus* L.): development of SCAR markers for the detection of plant adulterants used as bulking agents. *J. Agric. Food Chem.* 60, 10998–11004. doi: 10.1021/jf303106r
- Melnik, J. P., Wang, S., and Marcone, M. F. (2010). Chemical and biological properties of the world's most expensive spice: Saffron. *Food Res. Int.* 43, 1981–1989. doi: 10.1016/j.foodres.2010.07.033
- Mildner-Szkudlarz, S., and Jeleń, H. H. (2008). The potential of different techniques for volatile compounds analysis coupled with PCA for the detection of the adulteration of olive oil with hazelnut oil. *Food Chem.* 110, 751–761. doi: 10.1016/j.foodchem.2008.02.053
- Moreno-Roman, P., and Bobick, K. (2022). Foldscope: increasing science accessibility worldwide. *Micros. Today* 30, 42–45. doi: 10.1017/S1551929522000633
- Nandhini, S., and Ashokkumar, K. (2022). An automatic plant leaf disease identification using DenseNet-121 architecture with a mutation-based henry gas solubility optimization algorithm. *Neural Comput. Appl.* 34, 5513–5534. doi: 10.1007/s00521-021-06714-z
- Nehvi, F., and Yasmin, S. (2021). Policy and research initiatives for promotion of saffron farming system and trade for doubling farmers' income. *J. Hort. Postharvest Res.* 4, 89–110. doi: 10.22077/JHPR.2021.4208.1199
- Petrakis, E. A., Cagliani, L. R., Polissiou, M. G., and Consonni, R. (2015). Evaluation of saffron (*Crocus sativus* L.) adulteration with plant adulterants by ¹H NMR metabolite fingerprinting. *Food Chem.* 173, 890–896. doi: 10.1016/j.foodchem.2014.10.107
- Rong, W., Li, Z., Zhang, W., and Sun, L. (2014). "An improved CANNY edge detection algorithm," in: *2014 IEEE International Conference on Mechatronics and Automation (IEEE)*, 577–582. doi: 10.1109/ICMA.2014.6885761
- Sabatino, L., Scordino, M., Gargano, M., Belligno, A., Traulo, P., and Gagliano, G. (2011). HPLC/PDA/ESI-MS evaluation of saffron (*Crocus sativus* L.) adulteration. *Nat. Prod. Commun.* 6, 1934578X1100601220. doi: 10.1177/1934578X1100601220
- Siddiqui, I. A., Farooq, M. U., Siddiqui, R. A., and Rafi, S. T. (2006). Role of toluidine blue in early detection of oral cancer. *Pak. J. Med. Sci.* 22, 184. Available online at: <https://pjms.com.pk/issues/aprjun06/article/article19.html>
- Son, H.-J., Kang, J.-H., Hong, E.-J., Lim, C.-L., Choi, J.-Y., and Noh, B.-S. (2009). Authentication of sesame oil with addition of perilla oil using electronic nose based on mass spectrometry. *Kor. J. Food Sci. Technol.* 41, 609–614. Available online at: <https://www.koreascience.or.kr/article/JAKO200916955022756.page>
- Speiser, J. L., Miller, M. E., Tooze, J., and Ip, E. (2019). A comparison of random forest variable selection methods for classification prediction modeling. *Expert Syst. Appl.* 134, 93–101. doi: 10.1016/j.eswa.2019.05.028
- Tan, J., and Xu, J. (2020). Applications of electronic nose (e-nose) and electronic tongue (e-tongue) in food quality-related properties determination: a review. *Artif. Intell. Agric.* 4, 104–115. doi: 10.1016/j.aiia.2020.06.003
- Torelli, A., Marieschi, M., and Bruni, R. (2014). Authentication of saffron (*Crocus sativus* L.) in different processed, retail products by means of SCAR markers. *Food Control* 36, 126–131. doi: 10.1016/j.foodcont.2013.08.001
- Verma, R., and Achutha, J. C. (2016). Flower species classification using random forest classifier. *Int. J. Comb. Res. Dev.* 5, 1056–1060. Available online at: http://www.ijcrrd.com/files/Vol_5_issue_8/y13.pdf
- Veronese, E., Castellani, U., Peruzzo, D., Bellani, M., and Brambilla, P. (2013). Machine learning approaches: from theory to application in schizophrenia. *Comput. Math. Methods Med.* 2013, 867924. doi: 10.1155/2013/867924
- Vijayakumar, V., and Balakrishnan, N. (2021). Artificial intelligence-based agriculture automated monitoring systems using WSN. *J. Ambient Intell. Humaniz. Comput.* 12, 8009–8016. doi: 10.1007/s12652-020-02530-w
- Zhou, Q., Zhu, W., Li, F., Yuan, M., Zheng, L., and Liu, X. (2022). Transfer learning of the ResNet-18 and DenseNet-121 model used to diagnose intracranial hemorrhage in CT scanning. *Curr. Pharm. Des.* 28, 287–295. doi: 10.2174/1381612827666211213143357



OPEN ACCESS

EDITED BY

Miha Humar,
University of Ljubljana, Slovenia

REVIEWED BY

Michael Altgen,
Aalto University, Finland
Boštjan Lesar,
University of Ljubljana, Slovenia
Ladislav Reinprecht,
Technical University of Zvolen, Slovakia
Erni Ma,
Beijing Forestry University, China

*CORRESPONDENCE

Andrea Ponzeccchi
apo@ign.ku.dk
Lisbeth Garbrecht Thygesen
lgt@ign.ku.dk

SPECIALTY SECTION

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

RECEIVED 05 July 2022

ACCEPTED 15 August 2022

PUBLISHED 06 September 2022

CITATION

Ponzeccchi A, Thybring EE, Digaitis R,
Fredriksson M, Solsona SP and
Thygesen LG (2022) Raman
micro-spectroscopy of two types
of acetylated Norway spruce wood
at controlled relative humidity.
Front. Plant Sci. 13:986578.
doi: 10.3389/fpls.2022.986578

COPYRIGHT

© 2022 Ponzeccchi, Thybring, Digaitis,
Fredriksson, Solsona and Thygesen.
This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License](#)
(CC BY). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Raman micro-spectroscopy of two types of acetylated Norway spruce wood at controlled relative humidity

Andrea Ponzeccchi^{1*}, Emil E. Thybring¹, Ramūnas Digaitis^{2,3,4},
Maria Fredriksson², Sara Piqueras Solsona¹ and
Lisbeth Garbrecht Thygesen^{1*}

¹Bioresource Chemistry and Technology, Department of Geoscience and Natural Resource Management, University of Copenhagen, Frederiksberg, Denmark, ²Division of Building Materials, Lund University, Lund, Sweden, ³Department of Wood Technology, Norwegian Institute of Bioeconomy Research, Ås, Norway, ⁴Department of Wood and Biomaterials, Danish Technological Institute, Taastrup, Denmark

Water is a key element for wood performance, as water molecules interact with the wood structure and affect important material characteristics such as mechanical properties and durability. Understanding wood-water interactions is consequently essential for all applications of wood, including the design of wood materials with improved durability by chemical modification. In this work, we used Raman micro-spectroscopy in combination with a specially designed moisture chamber to map molecular groups in wood cell walls under controlled moisture conditions in the hygroscopic range. We analyzed both untreated and chemically modified (acetylated to achieve two different spatial distributions of acetyl groups within the cell wall) Norway spruce wood. By moisture conditioning the specimens successively to 5, 50, and 95% relative humidity using deuterium oxide (D₂O), we localized the moisture in the cell walls as well as distinguished between hydroxyl groups accessible and inaccessible to water. The combination of Raman micro-spectroscopy with a moisturizing system with deuterium oxide allowed unprecedented mapping of wood-water interactions. The results confirm lower moisture uptake in acetylated samples, and furthermore showed that the location of moisture within the cell wall of acetylated wood is linked to the regions where acetylation is less pronounced. The study demonstrates the local effect that targeted acetylation has on moisture uptake in wood cell walls, and introduces a novel experimental set-up for simultaneously exploring sub-micron level wood chemistry and moisture in wood under hygroscopic conditions.

KEYWORDS

Raman micro-spectroscopy, wood, acetylation, biological imaging, chemical modification, moisture, relative humidity, water

Introduction

Durability is a factor that often limits the service life of wood and wood products, especially for in-soil and outdoor applications. The use of wood in outdoor environments is challenging as the wood cell wall will be degraded by decay fungi when exposed to prolonged humid conditions. Moisture plays a key role in this process, as it is essential for fungi to colonize and consume the lignocellulosic cell walls (Ringman et al., 2019; Brischke and Alfredsen, 2020). As a hygroscopic material, wood takes up water from its surroundings both in vapor and liquid state. Hydroxyl (OH) groups are the main water sorption sites in wood cell walls (Simpson, 1980) and these are present throughout the lignocellulosic matrix. Chemical modification of wood is a way to improve the durability, often by limiting the hygroscopicity of the material (Dong et al., 2020). Modifications of wood can limit the access of water molecules in the cell wall by bulking the available space and/or by reducing the number of accessible sorption sites in the structure (Thybring and Fredriksson, 2021). The most utilized wood modification processes are acetylation, thermal modification and furfurylation (Rowell, 2006; Mantanis, 2017; Hill et al., 2021; Zelinka et al., 2022). Acetylation of wood by reaction with acetic anhydride substitutes a fraction of the hydroxyl groups with the more voluminous acetyl groups (Çelen et al., 2007). Since the cell wall chemistry of wood is heterogeneous, chemical modification may not affect all domains evenly. The spatial distribution of a chemical modification can also be deliberately controlled by tuning the reaction conditions (Digaitis et al., 2021) or the reaction path (Keplinger et al., 2015). Chemical changes in wood cell walls as a result of modification are often studied by Raman micro-spectroscopy because it is non-invasive and offers high spatial resolution (Agarwal, 2009, 2019; Gierlinger et al., 2012, 2013; Gierlinger, 2018). The chemical characterization has so far been conducted predominantly on water-saturated wood specimens and information related to non-saturated states is limited (Guo et al., 2017). In this study we introduce a novel combination of Raman micro-spectroscopy and controlled moisture conditioning of wood in a unique, custom-built moisture chamber. With this experimental setup we are able to study water within wood cell walls under controlled, non-saturated environmental conditions. Moreover, by use of deuterium exchange, water-accessible and non-accessible hydroxyl groups can be distinguished from each other. This allows visualization of the moisture distribution within cell walls of native and modified wood. Here, we demonstrate this setup and semi-quantitatively assess the distribution of acetyl groups and moisture within native and two types of acetylated wood cell walls of Norway spruce to illustrate local effects of acetylation.

Materials and methods

Wood material

Wood specimens of untreated, pyridine treated (controls), uniformly acetylated and interface acetylated Norway spruce [*Picea abies* (L.) Karst.] mature sapwood with dimensions $10 \text{ (longitudinal)} \times 5 \times 5 \text{ mm}^3$ were employed for this study. The material originated from experimental forests in the southern parts of Sweden and is further described by Fredriksson et al. (2016). To minimize variation between specimens all specimens were cut from the same board. The modification procedures are described in detail by Digaitis et al. (2021). Briefly, the uniform acetylation was achieved by impregnating the samples in a 1:4 (v/v) mixture of acetic anhydride (VWR Chemicals, Radnor, United States) and pyridine (Merck, Darmstadt, Germany) and subsequent heating at 80°C for 60 min. The interface acetylation was achieved using a solution of pure acetic anhydride and carrying out the reaction at 75°C for 24 h. Control specimens were treated with pure pyridine at 80°C for 3 h. The mass gain caused by the modification was evaluated as the relative mass change:

$$R_{\text{mod}} = \frac{m_{\text{dry}} - m_{\text{dry},0}}{m_{\text{dry},0}} \quad (1)$$

where m_{dry} (g) is the dry mass after modification and $m_{\text{dry},0}$ (g) is the dry mass before modification.

The recorded R_{mod} (g/g) for the interface acetylated specimen used in this study was 0.113 g/g. The mean R_{mod} of 10 uniformly acetylated specimens was 0.142 g/g. The pyridine extraction gave on average a negative R_{mod} of 0.023 g/g, indicating a mass loss, possibly due to removal of extractives from the wood.

Raman measurements with controlled humidity

Using a microtome (RM2255, Leica Biosystems, Wetzlar, Germany), three 16 μm cross-sections were produced per specimen, in total $3 \times 4 = 12$ cross-sections. Four cross-sections at the time, one per each type of wood material, were placed on the moisture chamber (detailed description of the moisture chamber used is provided in **Supplementary Material**, Section 1) and wetted with a drop of deuterium oxide (99.98% D_2O , Sigma-Aldrich, Munich, Germany). The samples were then covered with a borosilicate glass slide (thickness #1), the edges of which were sealed with nail polish. The fully assembled and loaded with wood cross-sections moisture chamber, with open inlet and outlet channels, was then vacuum dried for 12 h at 60°C. Afterward, the moisture chamber was connected to a humidity microcontroller (ACE flow 2.0, SolGelWay, France) to adjust the flow of a wet flux of saturated deuterium

oxide (D₂O) vapor and a dry flux of air at 0% RH. The water saturated flux was achieved with a bubbling system that included two flasks, a warming plate and D₂O (**Supplementary Figure 1**). To ensure that all water-accessible hydroxyl groups were deuterated, the samples were preconditioned at 95% RH for 12 h. Then, the samples were equilibrated at 5, 50, and 95% RH for 12 h at each humidity level, and at each of these humidity levels, Raman images were captured. A total of 36 Raman images were captured, describing 12 different latewood tracheids at three hygroscopic states, belonging to four types of wood material.

The confocal Raman microscope (alpha 300R, WITec GmbH, Ulm, Germany) was equipped with a UHTS 300 spectrometer and a 100x oil immersion objective (Zeiss “N-Achroplan,” NA = 1.2, transmittance of 73%, Carl Zeiss GmbH, Jena, Germany). A linear polarized 532 nm NdYag was used at a 10 mW laser power and with 0.1 s of integration time per spectrum to avoid sample degradation (**Prats-Mateu et al., 2018**). Even though the same tracheid was imaged three times, no signs of degradation were observed in the spectra. The images were acquired from cross sections that were previously aligned with the tangential direction parallel to the laser polarization (**Gierlinger et al., 2013**). Raman scattered light was detected with a back-illuminated charge-coupled device camera, air cooled with Peltier cooling to -60°C and with a 600 g/mm grating, resulting in a spectral resolution of 3.8 cm^{-1} . Images were acquired with a diffraction limited lateral spatial resolution of approximately $0.3\text{ }\mu\text{m}$.

Treatment and data analysis of Raman scattering data

The treatment and reduction of all Raman scattering data were carried out in Matlab ver. 2020b (Mathworks, Natick, Massachusetts, United States). Prior to analysis, spectra were subjected to (1) image size reduction, specifically shaped for each image, to reduce the size of the dataset; (2) removal of the part of the spectrum not useful for the analysis, consisting in the wavenumbers below 300 cm^{-1} and above $3,720\text{ cm}^{-1}$ approximately; (3) cosmic ray removal by use of median filtering (Matlab built-in function `medfilt1` using default settings); (4) Alternating Least Squares (ALS) baseline correction according to **Eilers and Boelens (2005)**, which has been shown to cope well with fluorescence contribution (**De Juan et al., 2014**), with parameters $\lambda = 10^5$ and $p = 0.0005$. Due to the heterogeneous distribution of wood polymers in the wood cell walls, the data were clustered using k-means cluster analysis (as implemented in Matlab), which successfully separated lignin rich parts of the cell wall, i.e., the cell corner and middle lamella (CCML), the cellulose rich secondary cell wall (S2), and the empty lumina of tracheids and ray cells

(LUMEN). A normalization to equal length (2-norm of each spectrum) was used before clustering, as it made the k-means clustering perform better based on our visual inspection of the clustering results.

For the spectra assigned to the cell wall cluster (CELL WALL = S2 + CCML clusters), average spectra were computed and Raman peak heights or areas were estimated using a linear baseline, individually set for each Raman band. Estimation of peak areas was preferred over peak heights when possible, i.e., when the peak of interest was an isolated peak and not a shoulder. For the sake of visual comprehension, in addition to the pre-processing, the average spectra of the cell walls in **Figure 1** have been furthered baseline corrected (ALS, $\lambda = 10^4$ and $p = 0.0002$). Peak areas were estimated with trapezoidal numerical integration (Matlab `trapz` function), and peak heights by the height of the baselined corrected peak, using a linear baseline individually set for each Raman band of interest. Peak areas were estimated for: (1) O-D stretching at $2,490\text{ cm}^{-1}$ (**Hofstetter et al., 2006**), calculated in the range $2,300\text{--}2,685\text{ cm}^{-1}$ and assigned to the absorbed deuterium oxide (D₂O) and the deuterated hydroxyls (O-D); (2) C=O carbonyl stretch at $1,738\text{ cm}^{-1}$ (**Marchessault and Liang, 1962**; **Adebajo et al., 2006**), in the range between $1,710$ and $1,780\text{ cm}^{-1}$ and assigned to acetylation and (3) O-H stretching at $3,450\text{ cm}^{-1}$ (**Wiley and Atalla, 1987**), in the range between $3,150\text{--}3,650\text{ cm}^{-1}$ and assigned to inaccessible hydroxyl groups. Peak heights were estimated for: (1) the mean aromatic ring stretching at $1,601\text{--}1,604\text{ cm}^{-1}$ (**Gierlinger and Schwanninger, 2007**), using a baseline in the range between $1,545$ and $1,710\text{ cm}^{-1}$ and assigned to the symmetric CC stretch of the aromatic ring of lignin substructures (**Bock and Gierlinger, 2019**); (2) the mean of the C-H stretch at $2,898\text{--}2,902\text{ cm}^{-1}$ (**Gierlinger et al., 2013**) using a baseline in the range $2,785\text{--}3,040\text{ cm}^{-1}$; (3) the maximum height of the C=C and C=O stretch calculated between $1,660$ and $1,664\text{ cm}^{-1}$ (**Bock and Gierlinger, 2019**), using a baseline in the range between $1,648$ and $1,710\text{ cm}^{-1}$ and assigned to the lignin substructures such as coniferyl alcohol and coniferyl aldehyde. The peak heights and areas were normalized over the aromatic ring stretching peak height at $1,601\text{ cm}^{-1}$ to compensate for the differences in band intensity due to changes in focal plane. Before normalization, to discard outliers given by negative values and values close to zero, peak areas lower than 1 and peak heights lower than 0.01 were set to 1. Only a small fraction of the areas and heights were rejected as outliers by means of this sorting method ($<1\%$). For the spectra assigned to the CCML cluster, a threshold clustering was applied to further distinguish between the cell corners (CC) and the S1–S3 layers and the middle lamella (S1S3ML). The spectra with ratio $2,898\text{ cm}^{-1}/1,601\text{ cm}^{-1}$ higher than x where assigned to S1S3ML, the rest to the CC cluster. x was individually set for each image after visual inspection, and varied between 1.8 and 2.

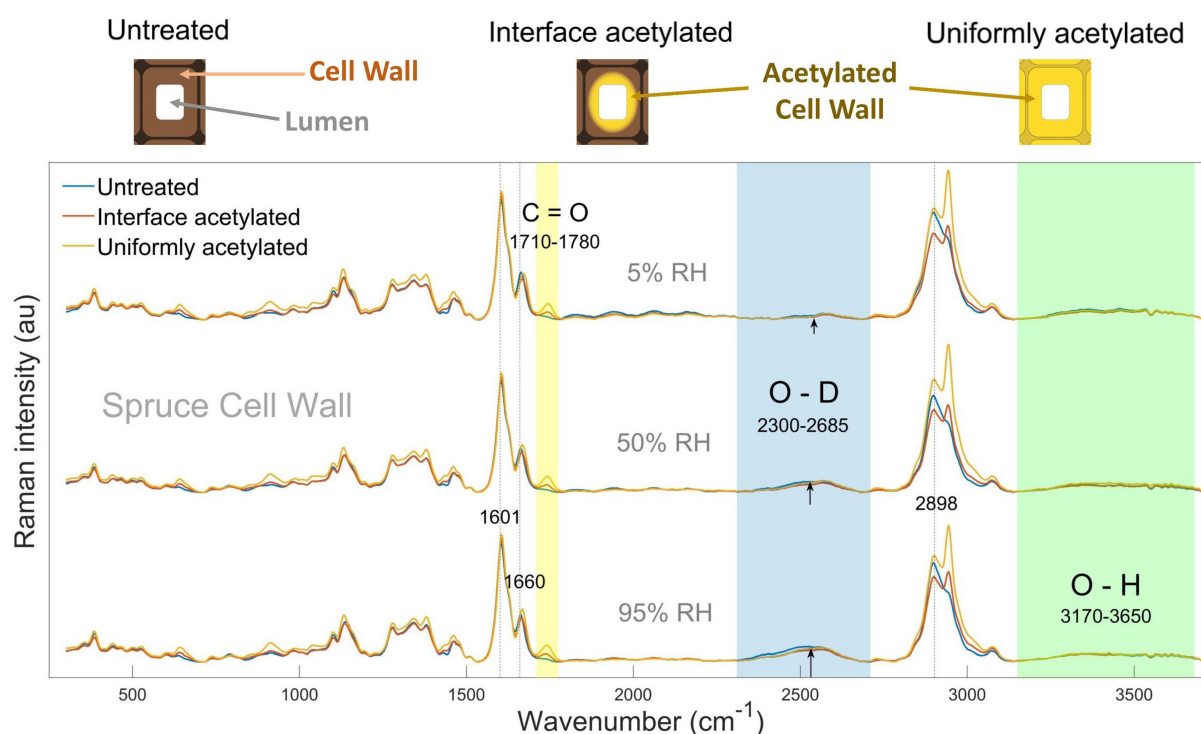


FIGURE 1

In the top, schematic representations of untreated, interface and uniformly acetylated cell walls are shown. Below, average Raman spectra calculated for untreated (blue), interface acetylated (red), and uniformly acetylated (yellow) spruce cell walls (i.e., CELL WALL cluster), conditioned in D₂O vapor at 5% (top), 50% (middle), and 95% RH (bottom). Wavenumber regions of interest assigned to acetylation, moisture, and water-inaccessible hydroxyls are shaded, respectively, in yellow, blue, and green. Peaks of interests are also highlighted by dotted lines, and marked with their exact wavenumber. The spectra for pyridine control samples are omitted for clarity as they were similar to the untreated wood, see [Supplementary Figure 2](#).

Due to the high content of noise, a statistical test was done to exclude unreliable information from the main peaks of interest, identified as the C=O and O-D stretching vibrations. The total raw sum of C=O (1,710–1,780 cm⁻¹) and O-D (2,300–2,685 cm⁻¹) counts were linearly baseline corrected. By evaluating the baselined raw sum of the peaks over the noise of the spectrum, each image pixel was labeled as significant or not, regarding the O-D and CO signals. The non-significant pixels contributed as null values in the averages. The pixels assigned to a mere fluctuation of noise were the ones in which the following expression was not fulfilled:

$$\frac{p_{\text{raw,events}}}{\sigma_{\text{noise}}} > 3.5 \quad (2)$$

with $p_{\text{raw,events}}$ the total sum of the counts (raw spectra, linearly baselined) and considering the background to be $0 \pm \sigma_{\text{noise}}$ events. The σ_{noise} was computed as the standard deviation of the difference between the raw and the reduced signal (PCA, first 3 components). Considering the Poisson statistic of the event of Raman Stokes scattering from a functional group, the $3.5\sigma_{\text{noise}}$ threshold is a cautious one (Barlow, 1993). A null value was assigned to the

spectra belonging to the pixels of lumina of tracheids and ray cells, as well as to the pixels in which the peak of interest was not significantly greater than the background noise.

Results and discussion

General observations about acetylation, hydroxyl groups and moisture

The average spectra for the acetylated samples showed higher intensity than the one for untreated wood at approximately 645, 910, 1,735, and 2,941 cm⁻¹, as previously reported by Digaitis et al. (2021) for acetylated spruce wood cell walls (Figure 1). These peaks were, respectively, assigned to O-C=O in plane deformation, H-C=C and H-C=O bending, C=O carbonyl stretching vibration and C-H stretching vibration, which are related to acetylation (Wiley and Atalla, 1987; Adebajo et al., 2006; Bock and Gierlinger, 2019). As

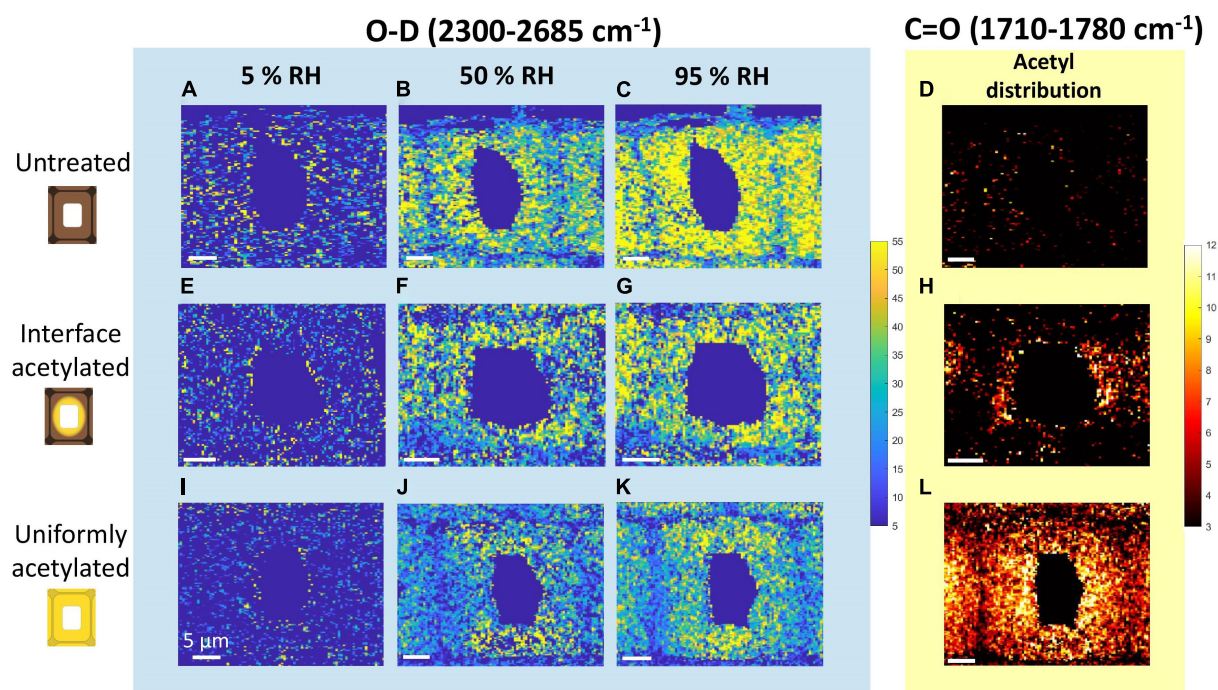


FIGURE 2

Intensity maps of the Raman peaks assigned to moisture and acetylation of untreated (A–D), interface acetylated (E–H) and uniformly acetylated wood (I–L). The three columns of maps on the left show O–D stretching ($2,300\text{--}2,685\text{ cm}^{-1}$) intensity scanned at 5% (first column), 50% (second column), and 95% RH (third column). The column on the right shows intensity maps of C=O stretching ($1,710\text{--}1,780\text{ cm}^{-1}$). For further clarity the backgrounds of the O–D and C=O intensity maps are color-coded as the band in Figure 1, i.e., respectively, blue and yellow. All the maps of O–D share the same intensity scale, as do all maps of C=O.

expected, these peaks did not vary significantly over the three hygroscopic states of the same types of wood. Among those, the peak area assigned to the C=O carbonyl stretching vibration, calculated between $1,710$ and $1,780\text{ cm}^{-1}$ (yellow band in Figure 1), was used as peak of interest to characterize acetylation (Adebajo et al., 2006). The C=O peak intensity is visibly the highest in uniformly acetylated wood (highest degree of wood cell wall acetylation, $R_{\text{mod}} = 0.142\text{ g/g}$), while the interface acetylated had the second highest peak intensity (lower degree of wood cell wall acetylation, $R_{\text{mod}} = 0.113\text{ g/g}$), in all the three hygroscopic states. This is because the interface acetylation only acts at the lumen-cell wall interface, i.e., the somewhat lower peak height compared to uniformly acetylated wood is a dilution effect of the spectral averaging.

The O–D signal (blue shading in Figure 1) is from the deuterated hydroxyl groups and moisture within the cell walls. Since each D_2O water molecule contains two O–D functionalities that contributes to the Raman signal, the measured O–D signal reflects the accessible hydroxyls plus two times the concentration of water molecules. Please refer to Supplementary Material, Section 3 for an extended discussion of this point. In the average spectra, the O–D signal was seen at 5% RH and it increased for higher RH levels for all

the types of wood studied. These observations indicate that successful deuteration and moisture uptake in the cell walls was achieved.

The O–H band (green shading in Figure 1) derives from the un-deuterated hydroxyl groups. Since the wood specimens were exposed to both liquid D_2O and high D_2O vapor pressure for prolonged time, it can be assumed that all water-accessible hydroxyl groups were deuterated. Consequently, the O–H signal relates to the hydroxyls inaccessible to water, and these groups are mainly found inside the cellulose microfibrils (Hofstetter et al., 2006; Salmén and Bergström, 2009). No outstanding variations can be pointed out from this band, neither between the different types of wood nor between the hygroscopic states, except for a slightly higher O–H peak at 5% RH for the untreated samples.

Distribution of acetyl groups and deuterium within cell walls

Figure 2 shows the intensity maps of the peak areas assigned to O–D (Figures 2A–C,E–G,I–K) and C=O stretching (Figures 2D,H,L) for one tracheid from each

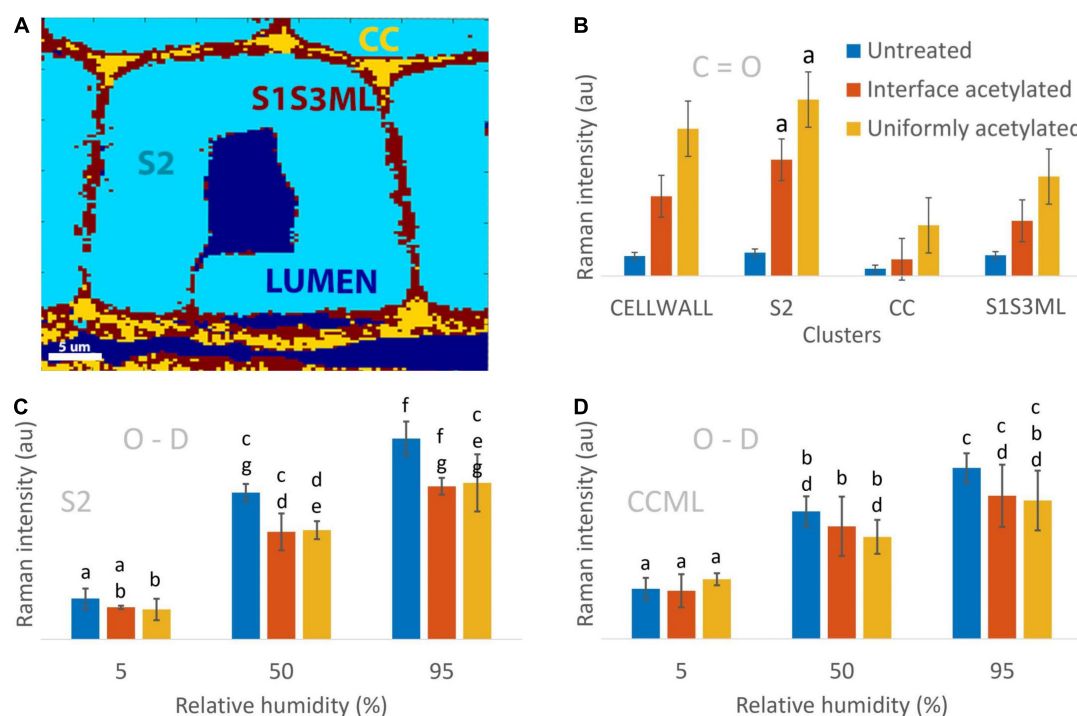


FIGURE 3

Quantification of the Raman O-D and C=O signals in the different clusters of the cell wall. (A) Example of the wood cell clustering depicting LUMEN (tracheid and ray cell lumina), S2 (secondary cell wall), CC (cell corners), S1S3ML (primary and tertiary cell wall, and middle lamella). The CELLWALL cluster (not shown) is the sum of the S2, CC and S1S3ML clusters, i.e., the whole cell wall. (B) Average Raman C=O intensity computed for untreated, interface acetylated and uniformly acetylated wood by averaging the images from the three replicates and the three hygroscopic states for each cluster (9 images contributed to the average and standard deviation of each bar). (C,D) Average Raman O-D intensity calculated for the three materials in the three hygroscopic states for S2 and CCML clusters. All Raman intensities are normalized over the lignin peak height at $1,601\text{ cm}^{-1}$. In (B–D) untreated samples are illustrated by blue bars, interface acetylated samples by red and uniformly acetylated samples by yellow. The bars include the standard deviation. The bars were compared using a two-tailed pair *t*-test for significance with the null hypothesis of no difference between the two set of data and $\alpha = 0.05$. The letter a on top of the bars in (B) indicates that the interface and uniformly acetylated belonging to the S2 cluster were the only results not rejecting the null hypothesis of equal signal for all three groups.

treatment. Due to the complex composition of wood and the lateral resolution of Raman micro-spectroscopy, the spectra from wood specimens often contain overlapping information. However, for this study the intensity maps of O-D and C=O vibrations could with high certainty be related to deuterated hydroxyls plus moisture, and acetyl esters, respectively.

In interface acetylated wood (Figure 2H), the C=O distribution was mainly concentrated in the cell wall area around the lumen, while it was basically absent in untreated wood (Figure 2D) and uniformly distributed in the uniformly acetylated wood (Figure 2L). These maps also indicate that the maximum C=O signal is of the same magnitude for the two different types of acetylation. As also seen from the average spectra in Figure 1, the intensity maps of C=O confirm that intensity and distribution between different hygroscopic states of the same tracheid does not vary significantly, as no acetyl esters are introduced or washed out during the moisture conditioning of the samples.

All types of wood materials showed an increase in the O-D signal over the whole cell wall, when going from nearly dry (5% RH) to more moist hygroscopic states. Since the amount of water-accessible hydroxyl groups does not vary over the three hygroscopic states (Altgen and Rautkari, 2021), the difference between the images of the same material were solely due to the moisture uptake. The major difference between modified and untreated wood was in the intensity and distribution of the O-D signal. The untreated wood (Figures 2A–C) had higher and more even distribution of the O-D signal over the secondary cell wall at every hygroscopic state than what was seen for the interface acetylated (Figures 2E–G) and uniformly acetylated wood (Figures 2I–K). Furthermore, the two types of acetylated wood showed a more uneven distribution of the O-D signal in the cell wall than the untreated wood, and lower values of O-D seemed to be associated with higher values of C=O signal intensity, i.e., the degree of acetylation (Figure 2H with Figure 2H and Supplementary Figure 4). Overall, the O-D

intensity maps showed that the moisture uptake was reduced in the acetylated samples, illustrating the usefulness of the experimental setup.

Quantification of acetyl esters and moisture present within individual cell wall layers

Finally, the clustering analysis that identified and localized different areas of the wood cell walls (**Figure 3A**) was used to quantify the contribution of moisture and acetylation separately for each of those regions. **Figure 3B** shows the mean value of the C=O stretching from cell walls of every kind of wood material and specific clusters. In each of the clusters analyzed, the average C=O stretching peak area (associated with acetylation) was highest for the uniformly acetylated samples and lowest for the untreated wood, with the interface acetylation in between, however, without being statistically significant (**Figure 3B**). To evaluate the reliability of the obtained results, the average Raman C=O signal was compared with the bound acetyl content found in literature. This was done by taking the ratio of the C=O signal of acetylated and untreated wood. For the cell wall cluster, this ratio was 7.4 ± 3.3 for uniformly acetylated and 4.0 ± 1.3 for interface acetylated wood. For comparison, the ratio of bound acetyl concentration in uniformly acetylated and untreated wood is theoretically expected to be 11.9 ± 0.8 and 9.4 ± 0.6 (**Supplementary Table 1**), which align with experimental data for acetylation of *Radiata pine* (Beck et al., 2017, 2018). Thus, the theoretical ratios were somewhat higher but of the same magnitude as the ratios of the Raman C=O signals.

In the S2 cluster (**Figure 3C**), the mean O-D values at each of the hygroscopic states of untreated and interface acetylated wood were statistically the same, as well as between interface and uniformly acetylated wood, while results for the untreated wood were different. This is in contrast to the CCML cluster (**Figure 3D**) where mean O-D values of interface acetylated, uniformly acetylated and untreated wood were all statistically comparable between the same hygroscopic states. Overall, the trends of **Figures 3C,D** make us speculate that the interface acetylated wood reduces the moisture uptake relatively more in the secondary cell wall than the uniformly acetylated wood, even though the statistics can only partially confirm this claim. The low significance of the data regarding the mean O-D signal suggests that, considering the evidences from the average spectra (**Figure 1**), the intensity maps (**Figure 2**) and the trends from the bar charts (**Figures 3C,D**), three replicas is not enough to overcome the great variability of the material.

To further evaluate the obtained results, the Raman O-D signal at 5% RH was compared with the expected O-D concentration from experimentally determined hydroxyl accessibility and predicted residual moisture. The ratio of

the O-D signal of acetylated samples and untreated wood was compared with predicted O-D concentration at 5% RH (**Supplementary Table 3**). Whereas the latter gave ratios of 0.7 ± 0.1 for interface acetylated and 0.4 ± 0.1 for uniformly acetylated wood, the ratios based on the Raman O-D signal were found to be 0.8 ± 0.3 and 0.9 ± 0.2 for interface and uniformly acetylated wood, respectively. Thus, while the Raman data suggests a decreasing O-D concentration for the acetylated materials, the uniformly acetylated wood exhibited a more intense O-D signal at 5% RH than the predicted values.

The Raman O-D signal from the cell wall cluster (not shown) for each type of material was also compared with the predicted O-D concentration in the different moist states based on experimental data from Digaitis et al. (2021). The values showed the O-D Raman signal and O-D concentration at 50 and 95% RH, normalized over the corresponding values at 5% RH to be of the same order of magnitude (**Supplementary Table 2**).

Overall, this study demonstrates the effect of acetylation on moisture uptake locally in wood cell walls, and illustrates the possibilities for simultaneously exploring sub-micron level wood chemistry and moisture in wood under hygroscopic conditions.

Conclusion

A novel experimental set-up was introduced in this study for simultaneous exploration of the sub-micron level cell wall chemistry and moisture in wood under hygroscopic conditions. Analysis of both cell wall-lumen interface acetylated and uniformly acetylated latewood cells of Norway spruce illustrated the local effect of acetyl esters on moisture uptake in different regions of the cell wall, at various levels of relative humidity and with sub-microscale resolution. The results collectively point to the conclusion that moisture is reduced more in highly acetylated areas of the cell wall.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

AP, ET, RD, LT, MF, and SS: conceptualization. AP, MF, SS, and RD: methodology and material production. AP, LT, and ET: writing—original draft preparation. AP: data acquisition and data analysis. LT, ET, and SS: supervision. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

LT thank Aase og Ejnar Danielsens Fond for funding part of the research work. RD would like to acknowledge the Royal Swedish Agricultural Academy for funding which led to the development of the moisture chamber used in this study grant no. GFS2019-0079. MF thank the Swedish Research Council FORMAS grant no. 2018-00387.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Adebajo, M. O., Frost, R. L., Klopogge, J. T., and Kokot, S. (2006). Raman spectroscopic investigation of acetylation of raw cotton. *Spectrochim. Acta Part A Mol. Biomol. Spectros.* 64, 448–453. doi: 10.1016/j.saa.2005.07.045
- Agarwal, U. P. (2009). “Raman spectroscopic characterization of wood and pulp fibers,” in *Characterization of lignocellulosic materials*, ed. T. Q. Hu (Oxford: Blackwell Publishing), 17–35. doi: 10.1002/9781444305425.ch2
- Agarwal, U. P. (2019). Analysis of cellulose and lignocellulose materials by raman spectroscopy: A review of the current status. *Molecules* 24:1659. doi: 10.3390/molecules24091659
- Altgen, M., and Rautkari, L. (2021). Humidity-dependence of the hydroxyl accessibility in Norway spruce wood. *Cellulose* 28, 45–58. doi: 10.1007/s10570-020-03535-6
- Barlow, R. J. (1993). *Statistics: A guide to the use of statistical methods in the physical sciences*. Hoboken, NJ: Wiley.
- Beck, G., Strobusch, S., Larnøy, E., Miltz, H., and Hill, C. (2017). Accessibility of hydroxyl groups in anhydride modified wood as measured by deuterium exchange and saponification. *Holzforschung* 72, 17–23. doi: 10.1515/hf-2017-0059
- Beck, G., Thybring, E. E., and Thygesen, L. G. (2018). Brown-rot fungal degradation and de-acetylation of acetylated wood. *Intl. Biodeterior. Biodegradation* 135, 62–70. doi: 10.1016/j.ibiod.2018.09.009
- Bock, P., and Gierlinger, N. (2019). Infrared and Raman spectra of lignin substructures: Coniferyl alcohol, abietin, and coniferyl aldehyde. *J. Raman Spectros.* 50, 778–792. doi: 10.1002/jrs.5588
- Brischke, C., and Alfredsen, G. (2020). Wood-water relationships and their role for wood susceptibility to fungal decay. *Appl. Microbiol. Biotechnol.* 104, 3781–3795. doi: 10.1007/s00253-020-10479-1
- Çelen, Y., Harper, D. P., and Labbé, N. (2007). “Chemical modification of wood by acetic anhydride,” in *Proceedings of the ACS National Meeting Book of Abstracts (Issue I)*. Chicago, IL.
- De Juan, A., Piqueras, S., Maeder, M., Hancewicz, T., Duponchel, L., and Tauler, R. (2014). Chemometric tools for image analysis. *Infrared and Raman Spectros. Imaging Second Edition* 9783527336, 57–110. doi: 10.1002/9783527678136.ch2
- Digaitis, R., Thybring, E. E., Thygesen, L. G., and Fredriksson, M. (2021). Targeted acetylation of wood: A tool for tuning wood-water interactions. *Cellulose* 28, 8009–8025. doi: 10.1007/s10570-021-04033-z
- Dong, Y., Wang, K., Li, J., Zhang, S., and Shi, S. Q. (2020). Environmentally benign wood modifications: A review. *ACS Sustain. Chem. Eng.* 8, 3532–3540. doi: 10.1021/acssuschemeng.0c00342
- Eilers, P. H. C., and Boelens, H. F. (2005) *Baseline correction with asymmetric least squares smoothing: Leiden university medical centre report*. Leiden: Leiden University, 1–5.
- Fredriksson, M., Wadsö, L., Johansson, P., and Ulvcróna, T. (2016). Microclimate and moisture content profile measurements in rain exposed Norway spruce (*Picea abies* (L.) Karst.) joints. *Wood Material Sci. Eng.* 11, 189–200. doi: 10.1080/17480272.2014.965742
- Gierlinger, N. (2018). New insights into plant cell walls by vibrational microspectroscopy. *Appl. Spectros. Rev.* 53, 517–551. doi: 10.1080/05704928.2017.1363052
- Gierlinger, N., and Schwanninger, M. (2007). The potential of raman microscopy and raman imaging in plant research. *Spectroscopy* 21, 69–89. doi: 10.1155/2007/498206
- Gierlinger, N., Keplinger, T., and Harrington, M. (2012). Imaging of plant cell walls by confocal Raman microscopy. *Nat. Protoc.* 7, 1694–1708. doi: 10.1038/nprot.2012.092
- Gierlinger, N., Keplinger, T., Harrington, M., and Schwanninger, M. (2013). “Raman imaging of lignocellulosic feedstock,” In *Cellulose - Biomass conversion*, eds T. A. de Ven, and J. Kadla (London: IntechOpen). doi: 10.5772/50878
- Guo, X., Wu, Y., and Yan, N. (2017). Characterizing spatial distribution of the adsorbed water in wood cell wall of Ginkgo biloba L. by μ -FTIR and confocal Raman spectroscopy. *Holzforschung* 71, 415–423. doi: 10.1515/hf-2016-0145
- Hill, C., Altgen, M., and Rautkari, L. (2021). Thermal modification of wood—a review: Chemical changes and hygroscopicity. *J. Materials Sci.* 56, 6581–6614. doi: 10.1007/s10853-020-05722-z
- Hofstetter, K., Hinterstoisser, B., and Salmén, L. (2006). Moisture uptake in native cellulose—The roles of different hydrogen bonds: A dynamic FT-IR study using Deuterium exchange. *Cellulose* 13, 131–145. doi: 10.1007/s10570-006-9055-2
- Keplinger, T., Cabane, E., Chanana, M., Hass, P., Merk, V., Gierlinger, N., et al. (2015). A versatile strategy for grafting polymers to wood cell walls. *Acta Biomaterialia* 11, 256–263. doi: 10.1016/j.actbio.2014.09.016
- Mantanis, G. I. (2017). Chemical modification of wood by acetylation or furfurylation: A review of the present scaled-up technologies. *BioResources* 12, 4478–4489. doi: 10.15376/BIORES.12.2.MANTANIS
- Marchessault, R. H., and Liang, C. Y. (1962). The infrared spectra of crystalline polysaccharides. VIII. Xylans. *J. Polym. Sci.* 59, 357–378. doi: 10.1002/pol.1962.1205916813
- Prats-Mateu, B., Bock, P., Schöffenegger, M., Toca-Herrera, J. L., and Gierlinger, N. (2018). Following laser induced changes of plant phenylpropanoids by Raman microscopy. *Sci. Rep.* 8, 1–12. doi: 10.1038/s41598-018-30096-3
- Ringman, R., Beck, G., and Pilgård, A. (2019). The importance of moisture for brown rot degradation of modified wood: A critical discussion. *Forests* 10, 1–22. doi: 10.3390/f10060522

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.986578/full#supplementary-material>

- Rowell, R. M. (2006). Chemical modification of wood: A short review. *Wood Material Sci. Eng.* 1, 29–33. doi: 10.1080/17480270600670923
- Salmén, L., and Bergström, E. (2009). Cellulose structural arrangement in relation to spectral changes in tensile loading FTIR. *Cellulose* 16, 975–982. doi: 10.1007/s10570-009-9331-z
- Simpson, W. (1980) Sorption theories applied to wood. *Wood Fiber Sci.* 12, 183–195.
- Thybring, E. E., and Fredriksson, M. (2021). Wood modification as a tool to understand moisture in wood. *Forests* 12, 1–18. doi: 10.3390/f12030372
- Wiley, J. H., and Atalla, R. H. (1987). Band assignments in the raman spectra of celluloses. *Carbohydr. Res.* 160, 113–129. doi: 10.1016/0008-6215(87)80306-3
- Zelinka, S. L., Altgen, M., Emmerich, L., Guigo, N., Keplinger, T., Kymäläinen, M., et al. (2022). Review of wood modification and wood functionalization technologies. *Forests* 13:1004.



OPEN ACCESS

EDITED BY

Gregorio Egea,
University of Seville, Spain

REVIEWED BY

Mohd. Anwar Khan,
Sher-e-Kashmir University of
Agricultural Sciences and Technology
of Kashmir, India
Klára Kosová,
Crop Research Institute (CRI), Czechia

*CORRESPONDENCE

Haiguang Wang
wanghaiguang@cau.edu.cn

SPECIALTY SECTION

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

RECEIVED 25 July 2022

ACCEPTED 20 September 2022

PUBLISHED 03 October 2022

CITATION

Jiang Q, Wang H and Wang H (2022)
Two new methods for severity
assessment of wheat stripe rust
caused by *Puccinia striiformis*
f. sp. *tritici*.
Front. Plant Sci. 13:1002627.
doi: 10.3389/fpls.2022.1002627

COPYRIGHT

© 2022 Jiang, Wang and Wang. This is
an open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use,
distribution or reproduction is
permitted which does not comply with
these terms.

Two new methods for severity assessment of wheat stripe rust caused by *Puccinia striiformis* f. sp. *tritici*

Qian Jiang, Hongli Wang and Haiguang Wang*

College of Plant Protection, China Agricultural University, Beijing, China

Accurate severity assessment of wheat stripe rust caused by *Puccinia striiformis* f. sp. *tritici* is of great significance for phenotypic determination, prediction, and control of the disease. To achieve accurate severity assessment of the disease based on the actual percentages of lesion areas in the areas of the corresponding whole diseased leaves, two new methods were proposed for severity assessment of the disease. In the Adobe Photoshop 2022 software, the acquired images of single diseased leaves of each severity class of the disease were manually segmented, and the numbers of the leaf region pixels and lesion pixels of each diseased leaf were obtained by pixel statistics. After calculation of the actual percentages of lesion areas in the areas of the corresponding whole diseased leaves based on the obtained pixel numbers, the training sets and testing sets were constructed for each severity class by using the system sampling method with two sampling ratios of 4:1 and 3:2. Then the mean and standard deviation of the actual percentages of lesion areas contained in each training set were calculated, respectively. For each sampling ratio, two methods, one based on the midpoint value of the means of the actual percentages of lesion areas corresponding to two adjacent severity classes and the other based on the distribution range of most of the actual percentages of lesion areas, were used to determine the midpoint-of-two-adjacent-means-based actual percentage reference range and the 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas for each severity class. According to the determined reference ranges, the severity of each diseased leaf in the training sets and testing sets was assessed. The results showed that high assessment accuracies (not lower than 85%) for the training sets and testing sets were achieved, demonstrating that the proposed methods could be used to conduct severity assessment of wheat stripe rust based on the actual percentages of lesion areas. This study provides a reference for accurate severity assessments of plant diseases.

KEYWORDS

wheat stripe rust, severity, disease assessment, reference range, lesion area, normal distribution method

Introduction

Stripe rust (yellow rust) caused by *Puccinia striiformis* f. sp. *tritici* (*Pst*) is an important disease of wheat worldwide (Li and Zeng, 2002; Line, 2002; Chen, 2005; Wellings, 2011; Chen et al., 2014; Wang X. J. et al., 2014; Ali et al., 2017). It was estimated that this disease could cause yield losses of 5.47 million tons of wheat (equivalent to US\$979 million) each year in the worldwide wheat-growing regions (Beddow et al., 2015). In the United States in 2000, 2001, 2002, and 2003, the total losses of wheat in the top 12 states with the most severe wheat yield losses resulting from stripe rust were approximately 1.20, 5.24, 1.06, and 11.75 million tons, respectively (Chen, 2005). As an air-borne disease, in China, wheat stripe rust has the characteristics of high epidemic frequency, wide occurrence range, and severe damage to wheat production, and it has been pandemic in wheat-growing regions for many times, especially in 1950, 1964, 1990, and 2002, reducing the yield of wheat by 6, 3.2, 1.8, and 1.3 billion kg, respectively (Li and Zeng, 2002; Wan et al., 2003; Wan et al., 2004; Wan et al., 2007). In China, wheat stripe rust is one of the most important and devastating wheat diseases and is always a serious threat to the safety of wheat production, critically affecting China's food security (Li and Zeng, 2002; Chen et al., 2014; Wang X. J. et al., 2014; Wang et al., 2022). To carry out the surveys and monitoring of wheat stripe rust is a key way to obtain the information on the occurrences of the disease, which can provide basic supports for disease prediction, resistant variety identification, disease management, and so on.

During the surveys of wheat stripe rust, generally, the main disease indicators surveyed include incidence, severity, disease index, etc. Among these indicators, the severity is applied to describe disease intensity or the degree of infection of a plant unit (e.g., a plant, leaf, fruit, branch, stem, or other plant part) and it is of great significance for disease quantification (Nutter et al., 1991; Bock et al., 2022). For wheat stripe rust, according to the Rules for Monitoring and Forecast of the Wheat Stripe Rust (*Puccinia striiformis* West.) (National Standard of the People's Republic China, GB/T 15795–2011), eight severity classes are classified based on the percentages of lesion areas in the areas of the corresponding whole wheat leaves. In this severity grading standard, the disease intensity between two adjacent severity classes is taken as its nearest percentage severity class, and the disease intensity of a diseased leaf with the severity lower than 1% is recorded as the severity class of 1%. Shang et al. (1990) designed a standard area diagram set for the severity assessment of wheat stripe rust, and this diagram set plays an important role in accurate severity assessment of the disease. The severity assessment of wheat stripe rust is an important part in disease surveys, concerning diseased plant phenotyping, disease prediction and forecast, and disease control decision-making. Therefore, the severity assessment should be conducted in strict

accordance with the severity grading standard to ensure the assessment accuracy and to facilitate the exchange and sharing of the corresponding assessment information.

At present, the severity assessment of wheat stripe rust is conducted mainly by using visual observation method (i.e., naked eye observation method) that is heavily dependent on experienced personnel. In addition to the method above, disease severity of wheat stripe rust can be assessed by using the methods based on image processing technology (Jiang et al., 2021), remote sensing technology (Huang et al., 2004; Wang et al., 2007; Zhao et al., 2014; Wang et al., 2016), and near infrared spectroscopy technology (Li et al., 2015). In some cases, the severity of wheat stripe rust can be assessed based on the disease incidence obtained via disease survey (Dong et al., 1990).

During *Pst* infects into wheat leaves, infection sites on wheat leaves are required. The actual area occupied by each infection site may be larger than the area of each lesion with the disease symptom (usually the uredinium produced at the infection site). In the severity grading standard of wheat stripe rust (i.e., the Rules for Monitoring and Forecast of the Wheat Stripe Rust (*Puccinia striiformis* West.)) as described above, the percentage of the lesion area in the area of a whole diseased wheat leaf corresponding to one of eight severity classes is not the actual percentage of the lesion area in the area of the whole leaf. The percentage of the lesion area in the area of a whole diseased wheat leaf corresponding to a severity class in the severity grading standard is greater than the actual percentage of the lesion area in the area of the whole leaf. This makes it very difficult to accurately assess the severity of wheat stripe rust in practice. Shang et al. (1990) measured the areas of wheat leaves using a leaf area meter, and obtained the actual coverage rate of all the uredinia on a wheat leaf of each severity class using a uredinium parameters based calculating method and a method via actual measurement of the amplified image of the wheat leaf with the most severe disease symptom selected in the field. The results obtained by Shang et al. (1990) showed that the actual uredinium coverage rates for the severity classes of 1%, 5%, 10%, 20%, 40%, 60%, 80%, and 100% were 0.35%, 1.75%, 3.5%, 7%, 14%, 21%, 28%, and 35%, respectively, indicating that the actual percentage of the lesion area in the area of a whole diseased wheat leaf corresponding to one of eight severity classes is quite different from the corresponding percentage of the lesion area in the area of the whole leaf determined by using the severity grading standard of wheat stripe rust. In addition, due to the relatively small size and great shape changes of the *Pst* uredinia, it is easy to induce errors in the actual measurements of the coverage areas of the uredinia, and thus there may be some induced errors in the actual uredinium coverage rates for the severity classes obtained by Shang et al. (1990). Therefore, it is necessary to develop a more convenient and accurate method for determining the actual percentage of uredinium coverage area in a diseased wheat leaf area. Moreover, the actual uredinium

coverage rate of each severity class obtained by Shang et al. (1990) is a fixed value, but most of the actual uredinium coverage rates in practice are between the fixed values of two adjacent severity classes, inducing great difficulties and inconvenience to the severity assessments. Therefore, under these circumstances, when disease severity is assessed by comparing the actual percentage of lesion area in the area of a whole diseased wheat leaf to the percentage of lesion area in the severity grading standard, great assessment errors may be induced and disease severity class may be incorrectly assessed.

The visual observation method is widely utilized to assess the severity of wheat stripe rust, it is time-consuming and laborious, and it has high requirements of experience of an assessor or a rater. When this method is utilized to carry out disease severity assessment in practice, it is not easy to conduct the assessment and to obtain accurate assessment results. Due to the influence of the human vision and experience, using this method, different assessors/raters may obtain different assessed severity class for the same diseased wheat leaf. Therefore, before carrying out the disease severity assessment in practice, an assessor or a rater is required to be trained to master the severity grading standard and the severity grading method, aiming to ensure the accuracy and reliability of the severity assessment results.

In comparison with disease severity of wheat stripe rust, it is easier to investigate disease incidence by determining whether a wheat leaf is diseased. The quantitative relationship between incidence and severity (*I-S* relationship) can be established after investigations of the incidence and severity of the disease, and then the severity can be speculated according to the incidence. Nevertheless, the *I-S* relationship is greatly affected by many factors such as the incidence, the growth stage of wheat, and the distribution of lesions on wheat leaf layers (Dong et al., 1990). Therefore, the application of the established *I-S* relationship equation/model has great limitations, limiting the application of the severity assessing method based on the disease incidence.

Studies on severity assessment of wheat stripe rust based on remote sensing technology, near infrared spectroscopy technology, image processing technology, and other information technologies, have been paid attention to. The severity assessment of wheat stripe rust based on remote sensing technology and near infrared spectroscopy technology is still in the experimental research phase. Due to the high price of the required instruments and the low practical applicability, the related methods based on remote sensing technology and near infrared spectroscopy technology are rarely applied in practical productions. With the rapid development of image acquisition technology and image processing technology, more and more recognition methods (Li et al., 2012; Wang M. L. et al., 2014; Guo et al., 2015; Hu et al., 2018) and severity assessment methods (Bao et al., 2021; Jiang et al., 2021) of wheat stripe rust based on image processing technology are utilized in research and practical applications.

At present, the methods based on image processing technology to assess the severity of wheat stripe rust can be divided into two categories; one is to directly identify the severity classes based on the extracted disease image features (Bao et al., 2021), and the other is to segment the lesion/lesions, calculate the lesion area and the area of a whole diseased wheat leaf (or the number of the lesion pixels and the number of pixels of the whole diseased leaf), calculate the actual percentage of the lesion area in the area of the whole diseased leaf, and assess the severity of the corresponding diseased leaf by comparing the actual percentage of the diseased area to the percentages for the eight severity classes in the disease severity grading standard (Jiang et al., 2021). In the current research and applications of plant disease severity assessment by using image processing technology, the situation that the percentage of the lesion area in the area of a whole diseased plant unit corresponding to each severity class in the severity grading standards of some plant diseases (such as wheat stripe rust and wheat leaf rust caused by *Puccinia triticina*) is not the actual percentage of the lesion area in the area of the whole diseased plant unit, is not taken into account. Thus the accuracies of the severity assessments of these plant diseases based on image processing technology are seriously affected, resulting in great errors or complete errors in the disease severity assessments. This is also the main reason for the low accuracy obtained in assessing the severity of these plant diseases based on the ratio of lesion area to the total area of a plant unit by using image processing technology, which limits the practical applications of the related technology.

To solve the difficulties in assessing the severity of wheat stripe rust and the problems in severity assessment of the disease based on the actual percentage of lesion area in the area of the corresponding whole diseased wheat leaf, and to improve the severity assessment accuracy, it is necessary to explore a simple, easy-to-operate, and rapid method with high accuracy for assessing the disease severity, which is of great significance for the survey, monitoring, prediction and forecast, and control of the disease. In this study, by using image processing software, the leaf region and lesion region in the acquired image of each single diseased wheat leaf were obtained via image segmentation operations, and the numbers of the lesion pixels and the pixels of the whole diseased leaf were achieved by pixel statistics. Then the actual percentage of the lesion area in the area of the whole diseased leaf was calculated, and the mean of the actual percentages of lesion areas corresponding to each severity class was calculated subsequently. Based on the midpoint value of the means of two adjacent severity classes, the reference range of the actual percentages of lesion areas corresponding to each severity class was determined for severity assessments. Furthermore, by referring to the method for determining the reference ranges in the field of medicine, the reference ranges at different levels (90%, 95%, and 99%) of the actual percentages of lesion areas corresponding to each severity class were estimated based on the distribution range of the actual percentages of lesion areas of most of the diseased wheat leaves

belonging to each severity class. According to the midpoint-of-two-adjacent-means-based actual percentage reference ranges and the reference ranges at different levels for the eight severity classes, the severity of each of the acquired diseased wheat leaves with the percentages of lesion areas was assessed, and the assessment performance of each reference range was evaluated by using the assessment accuracy. In this study, it is aimed to explore two new methods for severity assessment of wheat stripe rust based on the actual percentages of the lesion areas in the areas of the corresponding whole diseased leaves, to provide a reference for severity assessments of plant diseases based on the ratios of lesion areas to the total areas of plant units, and to provide supports for the automatic severity assessments of plant diseases based on image processing technology.

Materials and methods

In this study, two new methods for severity assessment of wheat stripe rust were developed according to the procedures

and steps as shown in Figure 1, and then the constructed data sets were used to evaluate the new methods.

Acquisition of single wheat leaf images of each severity class of wheat stripe rust

According to the Rules for Monitoring and Forecast of the Wheat Stripe Rust (*Puccinia striiformis* West.) as described above, wheat leaves with typical symptoms of wheat stripe rust with severity levels of 1%, 5%, 10%, 20%, 40%, 60%, 80%, and 100% were collected from the diseased wheat plants that were obtained by using artificial spray inoculation method in Shangzhuang Experimental Station of China Agricultural University, Beijing, China and an artificial climate chamber in the Laboratory of Macro-Phytopathology, China Agricultural University, Beijing, China. Each diseased leaf was expanded as flat as possible and fixed on a sheet of A4 white paper with the lesion side facing up by using double sided sticky tape. Images of the diseased leaves were taken with a Nikon D700 digital camera

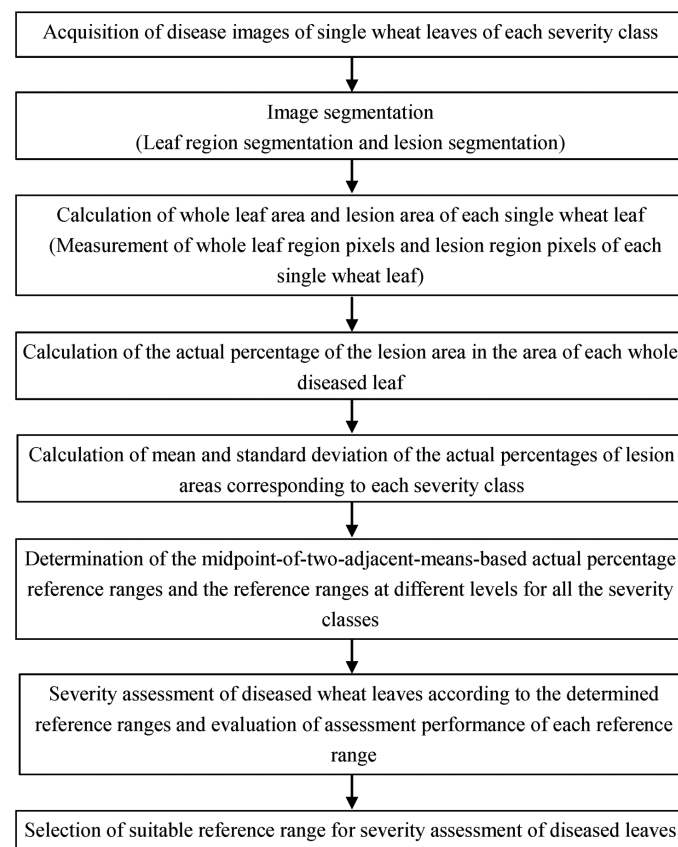


FIGURE 1

Work flow diagram for determining the reference ranges for disease severity assessment based on the actual percentages of lesion areas corresponding to each severity class and assessing the severity of wheat stripe rust.

(Nikon Corp., Tokyo, Japan), a HUAWEI P30 smartphone, and an iPhone 6S smartphone, and the sizes of the corresponding acquired images were 4256×2832, 3648×2736, and 4032×3024 pixels, respectively. One image was taken for each diseased leaf, 50 single diseased leaves of each severity class were used to be photographed, and a total of 400 single diseased leaf images were acquired. All the acquired images were in the JPEG format.

Manual image segmentation and pixel statistics of leaf regions and lesion regions of diseased wheat leaves

Manual image segmentation and pixel statistics of leaf regions and lesion regions of diseased wheat leaves were conducted by using the Adobe Photoshop 2022 software (Adobe Systems Incorporated, San Jose, CA, USA). In the software, a single leaf image (as shown in [Figures 2A–H](#)) of wheat stripe rust was opened, then the whole leaf region was selected with the quick selection tool, and subsequently the pixel number of the whole leaf region was viewed in the histogram panel and was recorded in a sheet in Microsoft Excel 2016. Inverse selection was carried out, then the selected region was filled with black color, and, finally, the image was saved in the JPEG format and the TIFF format, respectively (as shown in [Figures 2I–P](#)). When the quick selection tool was used, the ‘Enhance Edge’ was not selected, and for the brush options, in most cases, the size was set to 5 pixels, the hardness was set to 0%, the spacing was set to 25%, the angle was set to 35°, and the roundness was set to 100%. After completing the image segmentation of the diseased leaf, the diseased leaf layer was duplicated to form a new layer, and the original diseased leaf layer was named background and the new layer was named Layer 1 in the Adobe Photoshop 2022 software. Then the background layer was hidden, and Layer 1 was shown and selected. Repeatedly, the magic wand tool was used to select the non-lesion regions and the corresponding selected regions were filled with black color, so as to complete the initial segmentation of the lesion/lesions. After completing the initial segmentation, if there was still any non-lesion region that was not shown as black, the region was circled by using the lasso tool and was subsequently filled with black color, so as to complete the secondary segmentation of the lesion/lesions. The background layer was shown, Layer 1 was selected, and whether any lesion region was completely segmented or not was checked by repeatedly showing and hiding Layer 1. If there was still any lesion region shown as black, Layer 1 was selected and hidden, the background layer was shown, and then the region was circled by using the lasso tool and was subsequently removed. Until any lesion region was completely segmented, the non-lesion region was clicked by using the magic wand tool, then the inverse selection was carried out, and subsequently the pixel number of the lesion region/regions was viewed in the

histogram panel and was recorded in the sheet in Microsoft Excel 2016. Finally, Layer 1 was saved in the JPEG format and the TIFF format, respectively (as shown in [Figures 2Q–X](#)). When the magic wand tool was used, the sample size was set to point sample and the tolerance value was set to a number between 0 and 35. According to the actual selection of the lesion region, the tolerance value can be adjusted and the ‘Contiguous’ option can be selected. When the lasso tool was used, the feather value was set to 0 pixel. In all the above processes, the options ‘Anti-alias’ and ‘Sample All Layers’ were not selected. The numbers of the lesion pixels and the whole diseased leaf pixels for each diseased wheat leaf image were obtained by using the method as described above.

Calculation of the actual percentage of the lesion area in the area of each whole diseased wheat leaf

For each diseased wheat leaf image, based on the pixel number of the whole leaf region and the pixel number of the lesion region/regions recorded in the sheet in Microsoft Excel 2016, the percentage of the pixel number of the lesion region/regions in the pixel number of the whole leaf region, i.e., the actual percentage of the lesion area in the area of the whole diseased leaf, was calculated according to the following Formula (1).

$$r = \frac{A_d}{A_l} \times 100\% \quad (1)$$

where r is the actual percentage of the lesion area in the area of the whole diseased leaf for a single diseased wheat leaf, A_d is the pixel number of the lesion region/regions in the single diseased wheat leaf image, and A_l is the pixel number of the whole leaf region in the single diseased wheat leaf image.

Normal distribution tests on the data of the actual percentages of lesion areas corresponding to each severity class and the data of the reconstructed data sets after sampling

Normal distribution test on the data of the actual percentages of lesion areas in the corresponding whole leaf areas at the severity level of 1%, 5%, 10%, 20%, 40%, 60%, 80%, or 100% was conducted by using the UNIVARIATE procedure in the SAS 9.4 software (SAS Institute Inc. Cary, NC, USA). The results showed that 50 actual percentages of the lesion areas in the corresponding whole leaf areas for each severity class had a normal distribution. After 50 actual percentages of the lesion areas in the corresponding whole leaf areas for each severity class were sorted from large to small, the

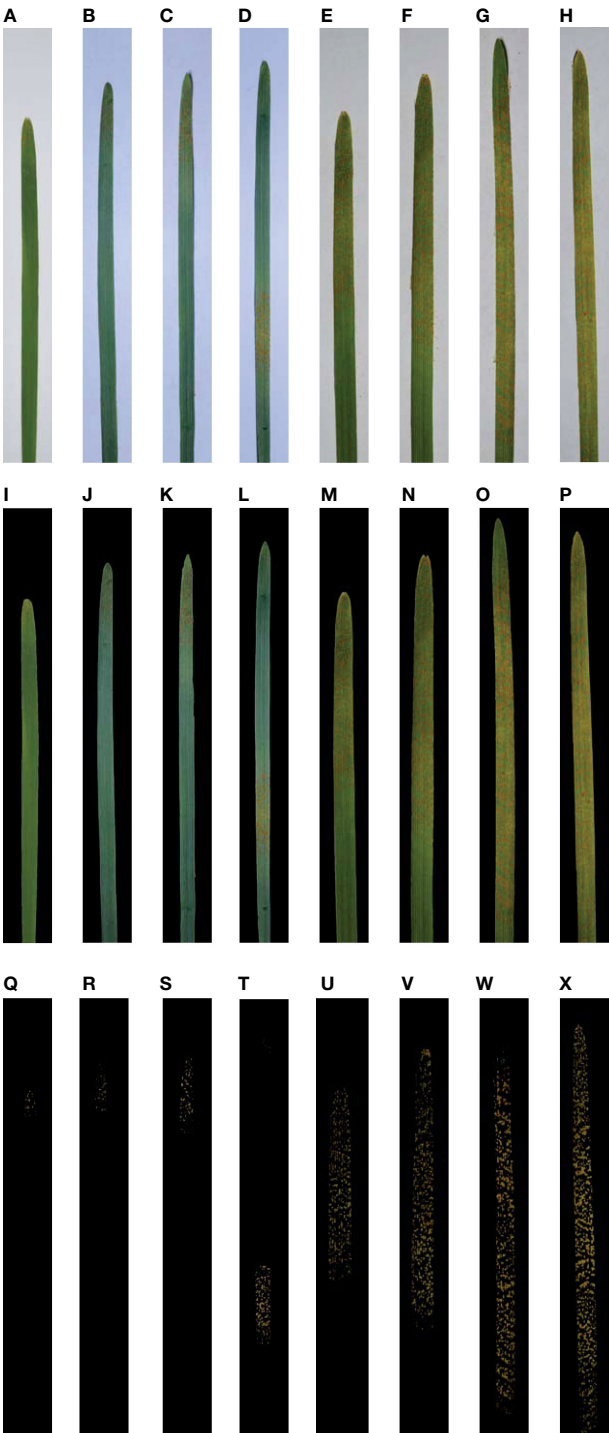


FIGURE 2
Single wheat leaf images of each severity class of wheat stripe rust and the corresponding leaf region images and lesion images after segmentation. All the images were shown after being cropped uniformly so that they could be demonstrated clearly. **(A–H)** Single diseased wheat leaf images of the severity classes of 1%, 5%, 10%, 20%, 40%, 60%, 80%, and 100%, respectively; **(I–P)** Segmented images of leaf regions for the single diseased wheat leaf images of the severity classes of 1%, 5%, 10%, 20%, 40%, 60%, 80%, and 100%, respectively; **(Q–X)** Segmented lesion images for the single diseased wheat leaf images of the severity classes of 1%, 5%, 10%, 20%, 40%, 60%, 80%, and 100%, respectively.

training and testing sets were constructed based on the data sampled from the 50 actual percentages by using the system sampling method with the ratio of the number of specimens in the training set to the number of specimens in the testing set equal to 4:1 or 3:2. For the sampling ratio of 4:1, the corresponding constructed training and testing sets were recorded as Train40_R and Test10_R, respectively. Train40_R was composed of the 40 specimens obtained by using the system sampling method with the sampling ratio equal to 4:1 when the severity class was *R*, and Test10_R was composed of the 10 remaining specimens. For the sampling ratio of 3:2, the corresponding constructed training and testing sets were recorded as Train30_R and Test20_R, respectively. Train30_R was composed of the 30 specimens obtained by using the system sampling method with the sampling ratio equal to 3:2 when the severity class was *R*, and Test20_R was composed of the 20 remaining specimens. *R* was the percentage of the lesion area in the area of the whole diseased leaf of the corresponding severity class in the severity grading standard of wheat stripe rust, so the value of *R* could be 1%, 5%, 10%, 20%, 40%, 60%, 80%, or 100%. Normal distribution tests on the data of the training sets (Train40_R and Train30_R) for the severity class of *R* were conducted by using the UNIVARIATE procedure in the SAS software. The results showed that the actual percentages of lesion areas contained in each training set had a normal distribution.

Calculation of the mean and standard deviation of the actual percentages of lesion areas corresponding to each severity class

The mean (\bar{r}) and standard deviation (*s*) of the actual percentages of lesion areas contained in each training set (Train40_R or Train30_R) for the severity class of *R* were calculated, respectively. The value of \bar{r} for each severity class was treated as the representative value of the actual percentage of lesion area in the area of a whole diseased wheat leaf of the corresponding severity class.

Determination of the reference ranges of the actual percentages of lesion areas in the corresponding whole leaf areas for all the severity classes

The reference ranges of the actual percentages of lesion areas in the corresponding whole leaf areas for all the severity classes of wheat stripe rust were determined by using the following two methods.

Method 1: The actual percentage reference range corresponding to each severity class was determined by taking

the midpoint value (r_{midpoint}) of the means of the actual percentages of lesion areas of two adjacent severity classes as the demarcation point, and this midpoint-of-two-adjacent-means-based actual percentage reference range was treated as one kind of the reference ranges of the actual percentages of lesion areas in the corresponding whole leaf areas for the severity class. Based on the actual percentages of lesion areas contained in each training set (Train40_R or Train30_R), the midpoint value r_{midpoint} of the means of two adjacent severity classes was treated as the demarcation point, and then the r_{midpoint} value was regarded as the upper limit of the actual percentages of lesion areas corresponding to the lower severity class in the two adjacent severity classes and was regarded as the lower limit of the actual percentages of lesion areas corresponding to the higher severity class. For the lowest severity class (1%), the lowest actual percentage of lesion area in the corresponding whole diseased leaf is greater than 0%. Thus the midpoint-of-two-adjacent-means-based actual percentage reference ranges corresponding to the severity class of *R* based on the training sets Train40_R and Train30_R were determined for the severity assessment of wheat stripe rust. The actual percentage of the lesion area corresponding to demarcation point was calculated according to the following Formula (2).

$$r_{\text{midpoint}} = \frac{\bar{r}_a + \bar{r}_b}{2} \quad (2)$$

where r_{midpoint} is the midpoint value of the means of the actual percentages of lesion areas of two adjacent severity classes, \bar{r}_a is the mean of the actual percentages of lesion areas corresponding to the lower severity class of *a* in the two adjacent severity classes, and \bar{r}_b is the mean of the actual percentages of lesion areas corresponding to the higher severity class of *b* in the two adjacent severity classes.

Method 2: The reference ranges at different levels for all the severity classes were determined by referring to the method for determining the medical reference ranges. In this study, since the actual percentages of the lesion areas in the corresponding whole leaf areas contained in each training set (Train40_R or Train30_R) for each severity class had a normal distribution, the normal distribution method (Sun and Xu, 2014) was used to determine the reference ranges of the actual percentages of the lesion areas for all the severity classes. According to the normal distribution method (Sun and Xu, 2014), for each severity class, the upper and lower limits of the bilateral 100(1- α)% reference range were calculated by using the formula $\bar{r} \pm u_{\alpha/2}s$, and the unilateral 100(1- α)% reference range was determined by using the formula $> \bar{r} - u_{\alpha}s$ or $< \bar{r} + u_{\alpha}s$. In the formulas, \bar{r} is the mean of the actual percentages of lesion areas in the corresponding whole leaf areas for a severity class, *s* is the standard deviation of the actual percentages of lesion areas in the corresponding whole leaf areas for the severity class, and u_{α} is the standard normal deviate corresponding to the α value. In this study, based on the values of \bar{r} and *s* of the actual percentages of lesion areas in the

corresponding whole leaf areas contained in each training set (Train40_R or Train30_R) for the severity class of R , the 90% ($\alpha=0.1$), 95% ($\alpha=0.05$), and 99% ($\alpha=0.01$) reference ranges of the actual percentages of lesion areas corresponding to the severity class of 5%, 10%, 20%, 40%, 60%, 80%, or 100% were determined according to the formulas $\bar{r} \pm 1.64s$, $\bar{r} \pm 1.96s$, and $\bar{r} \pm 2.58s$, respectively. In particular, for the 90% ($\alpha=0.1$), 95% ($\alpha=0.05$), and 99% ($\alpha=0.01$) reference ranges of the severity class of 1%, the lower limits were greater than 0%, and the corresponding upper limits were calculated by using the formulas $\bar{r} + 1.28s$, $\bar{r} + 1.64s$, and $\bar{r} + 2.33s$, respectively.

For a sampling ratio (4:1 or 3:2), if the 90%, 95%, or 99% reference ranges of the actual percentages of lesion areas of adjacent severity classes of wheat stripe rust obtained by using the normal distribution method overlapped, the normal distribution curves of the actual percentage data contained in the training sets of all the severity classes were drawn with the corresponding values of \bar{r} and s by using the normal distribution probability density function (*normpdf*) in the software MATLAB 2019b (MathWorks, Natick, MA, USA), and then the intersection point of the normal distribution curves of the actual percentage data contained in the training sets of two adjacent severity classes was obtained by using the function *solve* in the software. The abscissa value of the intersection point was denoted as $\bar{r}a-b$ where a was the lower one and b was the higher one in the two adjacent severity classes. The upper limit of the reference range of the actual percentages of lesion areas corresponding to the severity class of a and the lower limit of the reference range of the actual percentages of lesion areas corresponding to the severity class of b were determined based on the $\bar{r}a-b$ value. Subsequently, the probability of the interval composed of the lower and upper limits of the reference range for a severity class was calculated by using the function *normspec* in the software MATLAB 2019b, and was regarded as the corresponding actual probability of this reference range.

For a sampling ratio, if the normal distribution curve of the actual percentage data contained in the training set of a severity class of R had no intersection point with that of the actual percentage data contained in the training set of any adjacent severity class, or had an intersection point, but the $\bar{r}a-b$ value lay outside the interval corresponding to the 90%, 95%, or 99% probability of the normal distribution curve of the actual percentage data contained in the training set of the severity class of R , the 90%, 95%, or 99% reference range of the actual percentages of lesion areas for the severity class of R was estimated according to the formula as described above. If the normal distribution curve of the actual percentage data contained in the training set of a severity class of R only had an intersection point with that of the actual percentage data contained in the training set of the lower adjacent severity class, and the $\bar{r}a-b$ value lay inside the interval corresponding to the 90%, 95%, or 99% probability of the normal distribution curve of the actual percentage data contained in the training set of the

severity class of R , the $\bar{r}a-b$ value was regarded as the lower limit of the 90%, 95%, or 99% reference range of the actual percentages of lesion areas for the severity class of R , then the upper limit of the corresponding reference range was calculated according to the formula as described above, and, subsequently, the probability of the interval composed of the lower and upper limits of the reference range for this severity class calculated by using the function *normspec* in the software MATLAB 2019b was regarded as the corresponding actual probability of this reference range. If the normal distribution curve of the actual percentage data contained in the training set of a severity class of R only had an intersection point with that of the actual percentage data contained in the training set of the higher adjacent severity class, and the $\bar{r}a-b$ value lay inside the interval corresponding to the 90%, 95%, or 99% probability of the normal distribution curve of the actual percentage data contained in the training set of the severity class of R , the $\bar{r}a-b$ value was regarded as the upper limit of the 90%, 95%, or 99% reference range of the actual percentages of lesion areas for the severity class of R , then the lower limit of the corresponding reference range was calculated according to the formula as described above, and, subsequently, the probability of the interval composed of the lower and upper limits of the reference range for this severity class calculated by using the function *normspec* in the software MATLAB 2019b was regarded as the corresponding actual probability of this reference range. If the normal distribution curve of the actual percentage data contained in the training set of a severity class of R had an intersection point with that of the actual percentage data contained in the training set of each of the two adjacent severity classes, and the abscissa values of the two intersection points lay inside the interval corresponding to the 90%, 95%, or 99% probability of the normal distribution curve of the actual percentage data contained in the training set of the severity class of R , the two abscissa values were regarded as the lower and upper limits of the 90%, 95%, or 99% reference range of the actual percentages of lesion areas for the severity class of R , respectively, and then the probability of the interval composed of the lower and upper limits of the reference range for this severity class calculated by using the function *normspec* in the software MATLAB 2019b was regarded as the corresponding actual probability of this reference range.

In this study, according to the formulas as described above, the estimated 95% reference ranges of the actual percentages of lesion areas for some adjacent severity classes overlapped, and the estimated 99% reference ranges of the actual percentages of lesion areas for all the adjacent severity classes overlapped. Therefore, the 95% or 99% reference ranges of the actual percentages of lesion areas of the adjacent severity classes were obtained according to the method as described above. Although the corresponding probabilities of the reference ranges changed, in a convenient manner, the reference ranges were still called the 95% or 99% reference ranges. In this study, for the two sampling

ratios, the calculation methods of the 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas corresponding to each severity class of wheat stripe rust are shown in Table 1.

Severity assessment of each of the acquired diseased wheat leaves with the percentages of lesion areas

According to the midpoint-of-two-adjacent-means-based actual percentage reference ranges and the 90%, 95%, and 99% reference ranges obtained based on the actual percentage data contained in the training set Train40_R, the severity assessment of each diseased wheat leaf with the actual percentage of lesion area in the area of the corresponding whole diseased leaf contained in the training set Train40_R and the testing set Test10_R was conducted. In the same way, according to the midpoint-of-two-adjacent-means-based actual percentage reference ranges and the 90%, 95%, and 99% reference ranges obtained based on the actual percentage data contained in the training set Train30_R, the severity assessment of each diseased wheat leaf with the actual percentage of lesion area in the area of the corresponding whole diseased leaf contained in the training set Train30_R and the testing set Test20_R was carried out. Then the accuracy of severity assessments of the diseased wheat leaves with the actual percentages of lesion areas contained in each data set of a severity class was calculated by using the following Formula (3).

$$\text{Accuracy} = \frac{n_p}{n} \times 100\% \quad (3)$$

where accuracy is the severity assessment accuracy of the diseased wheat leaves with the actual percentages of lesion areas contained in each data set of a severity class, n_p is the number of the diseased leaves correctly assessed, and n is the total number of the diseased leaves assessed.

Results

The range, mean, and standard deviation of the actual percentage data contained in each of the training sets corresponding to each severity class

Based on the actual percentages of the lesion areas in the areas of the corresponding whole diseased wheat leaves, the range, mean (\bar{r}), and standard deviation (s) of the actual percentage data contained in each training set (Train40_R or Train30_R) for the severity class of R were achieved as shown in Table 2. The results showed that the range composed of the minimum and maximum actual percentages for the severity class of R (1%, 5%, 10%, 20%, 40%, 60%, 80%, or 100%) obtained

based on the actual percentage data contained in Train40_R was the same as that obtained based on the actual percentage data contained in Train30_R. The ranges of actual percentages for the severity classes of 1%, 5%, 10%, 20%, 40%, 60%, 80%, and 100% were [0.06%, 0.78%], [0.85%, 1.64%], [1.73%, 3.29%], [3.65%, 6.31%], [6.76%, 13.88%], [14.22%, 18.43%], [18.90%, 24.15%], and [24.54%, 36.49%], respectively. In Train40_R and Train30_R, the maximum actual percentage of lesion area in the corresponding whole leaf area for the severity class of 100% was 36.49%. Based on the actual percentage data contained in each training set that was called Train40_R, the means of the actual percentages of lesion areas for the severity classes of 1%, 5%, 10%, 20%, 40%, 60%, 80%, and 100% were 0.40%, 1.27%, 2.50%, 4.92%, 9.89%, 16.61%, 21.23%, and 30.52%, respectively. Based on the actual percentage data contained in each training set that was called Train30_R, the means of the actual percentages of lesion areas for the severity classes of 1%, 5%, 10%, 20%, 40%, 60%, 80%, and 100% were 0.40%, 1.28%, 2.50%, 4.92%, 9.87%, 16.61%, 21.23%, and 30.53%, respectively. The results showed that there was no obvious difference between the \bar{r} values or the s values of the actual percentages of lesion areas in the areas of the corresponding whole diseased leaves contained in Train40_R and Train30_R at the same severity level (severity class of R). The results demonstrated that for each severity class of wheat stripe rust, there was great difference between the actual percentage of lesion area in the area of a whole diseased leaf and the percentage of the lesion area in the area of the whole diseased leaf corresponding to the severity class in the severity grading standard of the disease as described above.

The determined reference ranges of the actual percentages of lesion areas in the corresponding whole leaf areas for all the severity classes

For each sampling ratio (4:1 or 3:2), a total of four sets of reference ranges of the actual percentages of lesion areas in the corresponding whole leaf areas for all the severity classes of wheat stripe rust, including the midpoint-of-two-adjacent-means-based actual percentage reference ranges, the 90% reference ranges, the 95% reference ranges, and the 99% reference ranges, were determined as shown in Table 3.

Based on the actual percentage data contained in the training sets Train40_R and Train30_R, the determined midpoint-of-two-adjacent-means-based actual percentage reference ranges corresponding to the severity class of R , as shown in Table 3, were obtained by taking the values of r_{midpoint} of the means of the actual percentages of lesion areas in the corresponding whole leaf areas of two adjacent severity classes as the demarcation points. Based on the actual percentage data contained in each training set that was called Train40_R, the midpoint-of-two-adjacent-means-based actual percentage reference ranges for

TABLE 1 Calculation methods of the 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas corresponding to each severity class of wheat stripe rust.

Severity class	90% reference range	95% reference range	99% reference range
1%	(0%, $\bar{r}+1.28s$]	(0%, $\bar{r}+1.64s$]	(0%, $\bar{r}_{1\%-5\%}$]
5%	[$\bar{r}-1.64s$, $\bar{r}+1.64s$]	[$\bar{r}-1.96s$, $\bar{r}+1.96s$]	($\bar{r}_{1\%-5\%}$, $\bar{r}_{5\%-10\%}$]
10%	[$\bar{r}-1.64s$, $\bar{r}+1.64s$]	($\bar{r}_{5\%-10\%}$, $\bar{r}+1.96s$]	($\bar{r}_{5\%-10\%}$, $\bar{r}_{10\%-20\%}$]
20%	[$\bar{r}-1.64s$, $\bar{r}+1.64s$]	[$\bar{r}-1.96s$, $\bar{r}+1.96s$]	($\bar{r}_{10\%-20\%}$, $\bar{r}_{20\%-40\%}$]
40%	[$\bar{r}-1.64s$, $\bar{r}+1.64s$]	($\bar{r}_{20\%-40\%}$, $\bar{r}+1.96s$]	($\bar{r}_{20\%-40\%}$, $\bar{r}_{40\%-60\%}$]
60%	[$\bar{r}-1.64s$, $\bar{r}+1.64s$]	[$\bar{r}-1.96s$, $\bar{r}_{60\%-80\%}$]	($\bar{r}_{40\%-60\%}$, $\bar{r}_{60\%-80\%}$]
80%	[$\bar{r}-1.64s$, $\bar{r}+1.64s$]	($\bar{r}_{60\%-80\%}$, $\bar{r}+1.96s$]	($\bar{r}_{60\%-80\%}$, $\bar{r}_{80\%-100\%}$]
100%	[$\bar{r}-1.64s$, $\bar{r}+1.64s$]	[$\bar{r}-1.96s$, $\bar{r}+1.96s$]	($\bar{r}_{80\%-100\%}$, $\bar{r}+2.58s$]

the severity classes of 1%, 5%, 10%, 20%, 40%, 60%, 80%, and 100% were (0, 0.84%], (0.84%, 1.89%], (1.89%, 3.71%], (3.71%, 7.41%], (7.41%, 13.25%], (13.25%, 18.92%], (18.92%, 25.88%], and (25.88%, 100%], respectively. Based on the actual percentage data contained in each training set that was called Train30_R, the midpoint-of-two-adjacent-means-based actual percentage reference ranges for the severity classes of 1%, 5%, 10%, 20%, 40%, 60%, 80%, and 100% were (0, 0.84%], (0.84%, 1.89%], (1.89%, 3.71%], (3.71%, 7.40%], (7.40%, 13.24%], (13.24%, 18.92%], (18.92%, 25.88%], and (25.88%, 100%], respectively. The results showed that for the severity class of *R*, the midpoint-of-two-adjacent-means-based actual percentage reference range obtained based on the actual percentage data in Train40_R was similar to that obtained based on the actual percentage data in Train30_R. The midpoint-of-two-adjacent-means-based actual percentage reference range for each severity class obtained

based on the actual percentage data contained in each corresponding training set by using the method as described above, had relatively large difference with the range (as shown in Table 2) composed of the minimum and maximum actual percentages for the corresponding severity class.

For the sampling ratio of 4:1, based on the actual percentage data contained in each training set that was called Train40_R, the obtained 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas corresponding to each severity class and the actual probabilities for the corresponding reference ranges are shown in Table 3. For the sampling ratio of 3:2, based on the actual percentage data contained in each training set that was called Train30_R, the obtained 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas corresponding to each severity class and the actual probabilities for the corresponding reference ranges are also

TABLE 2 Statistics of the actual percentage data of the lesion areas in the areas of the corresponding whole diseased leaves contained in each training set (Train40_R or Train30_R) of the severity class of *R* including the range of actual percentages of lesion areas, mean, and standard deviation.

Data set	Severity class	The range of actual percentages	Mean	Standard deviation
Train40 _{1%}	1%	[0.06%, 0.78%]	0.40%	0.19%
Train40 _{5%}	5%	[0.85%, 1.64%]	1.27%	0.23%
Train40 _{10%}	10%	[1.73%, 3.29%]	2.50%	0.42%
Train40 _{20%}	20%	[3.65%, 6.31%]	4.92%	0.78%
Train40 _{40%}	40%	[6.76%, 13.88%]	9.89%	1.97%
Train40 _{60%}	60%	[14.22%, 18.43%]	16.61%	1.21%
Train40 _{80%}	80%	[18.90%, 24.15%]	21.23%	1.41%
Train40 _{100%}	100%	[24.54%, 36.49%]	30.52%	3.19%
Train30 _{1%}	1%	[0.06%, 0.78%]	0.40%	0.19%
Train30 _{5%}	5%	[0.85%, 1.64%]	1.28%	0.23%
Train30 _{10%}	10%	[1.73%, 3.29%]	2.50%	0.43%
Train30 _{20%}	20%	[3.65%, 6.31%]	4.92%	0.78%
Train30 _{40%}	40%	[6.76%, 13.88%]	9.87%	1.95%
Train30 _{60%}	60%	[14.22%, 18.43%]	16.61%	1.20%
Train30 _{80%}	80%	[18.90%, 24.15%]	21.23%	1.43%
Train30 _{100%}	100%	[24.54%, 36.49%]	30.53%	3.21%

TABLE 3 The determined reference ranges of the actual percentages of lesion areas in the corresponding whole leaf areas for all the severity classes based on the actual percentage data in the training sets Train40_R and Train30_R and the actual probabilities for the corresponding 90%, 95%, and 99% reference ranges.

Data set	Severity class	Midpoint-of-two-adjacent-means-based actual percentage reference range	90% reference range	Actual probability corresponding to the 90% reference range	95% reference range	Actual probability corresponding to the 95% reference range	99% reference range	Actual probability corresponding to the 99% reference range
Train40 _{1%}	1%	(0, 0.84%]	(0%,0.64%]	90%	(0%, 0.71%]	95%	(0%, 0.80%]	96.47%
Train40 _{5%}	5%	(0.84%, 1.89%]	[0.89%, 1.65%]	90%	[0.82%, 1.72%]	95%	(0.80%, 1.75%]	96.11%
Train40 _{10%}	10%	(1.89%, 3.71%]	[1.81%, 3.19%]	90%	(1.75%, 3.32%]	93.75%	(1.75%, 3.43%]	94.95%
Train40 _{20%}	20%	(3.71%, 7.41%]	[3.64%, 6.20%]	90%	[3.39%, 6.45%]	95%	(3.43%, 6.60%]	95.63%
Train40 _{40%}	40%	(7.41%, 13.25%]	[6.66%,13.12%]	90%	(6.60%,13.75%]	92.75%	(6.60%,13.88%]	93.11%
Train40 _{60%}	60%	(13.25%, 18.92%]	[14.63%, 18.59%]	90%	[14.24%, 18.80%]	93.98%	(13.88%, 18.80%]	95.28%
Train40 _{80%}	80%	(18.92%, 25.88%]	[18.92%, 23.54%]	90%	(18.80%, 23.99%]	93.24%	(18.80%, 24.46%]	94.66%
Train40 _{100%}	100%	(25.88%, 100%]	[25.29%, 35.75%]	90%	[24.27%, 36.77%]	95%	(24.46%, 38.75%]	96.63%
Train30 _{1%}	1%	(0, 0.84%]	(0%,0.64%]	90%	(0%, 0.71%]	95%	(0%, 0.81%]	96.69%
Train30 _{5%}	5%	(0.84%, 1.89%]	[0.90%, 1.66%]	90%	[0.83%, 1.73%]	95%	(0.81%, 1.75%]	95.90%
Train30 _{10%}	10%	(1.89%, 3.71%]	[1.79%, 3.21%]	90%	(1.75%, 3.34%]	93.41%	(1.75%, 3.44%]	94.50%
Train30 _{20%}	20%	(3.71%, 7.40%]	[3.64%, 6.20%]	90%	[3.39%, 6.45%]	95%	(3.44%, 6.60%]	95.55%
Train30 _{40%}	40%	(7.40%, 13.24%]	[6.67%, 13.07%]	90%	(6.60%, 13.69%]	92.82%	(6.60%, 13.88%]	93.33%
Train30 _{60%}	60%	(13.24%, 18.92%]	[14.64%, 18.58%]	90%	[14.26%, 18.78%]	93.96%	(13.88%, 18.78%]	95.33%
Train30 _{80%}	80%	(18.92%, 25.88%]	[18.88%, 23.58%]	90%	(18.78%, 24.03%]	93.16%	(18.78%, 24.48%]	94.51%
Train30 _{100%}	100%	(25.88%, 100%]	[25.27%, 35.79%]	90%	[24.24%, 36.82%]	95%	(24.48%, 38.81%]	96.53%

shown in Table 3. On the whole, the obtained 90%, 95%, or 99% reference range for the severity class of *R* based on the actual percentage data in Train40_R had small difference with the corresponding 90%, 95%, or 99% reference range for the severity class of *R* based on the actual percentage data in Train30_R. For the obtained 90%, 95%, and 99% reference ranges based on the actual percentage data contained in each training set of a severity class of *R* (1%, 5%, 10%, 20%, 40%, 60%, 80%, or 100%), the 95% and 99% reference ranges of the severity class had small difference, but both of them had relatively large differences with the 90% reference range of the corresponding severity class.

The results indicated that based on the actual percentages of the lesion areas in the corresponding whole leaf areas contained in each training set (Train40_R or Train30_R), the 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas corresponding to each severity class obtained by using the normal distribution method had relatively obvious differences with the obtained midpoint-of-two-adjacent-means-based actual percentage reference range for the corresponding severity class. Moreover, the obtained 90%, 95%, and 99%

reference ranges of the actual percentages of lesion areas for the severity class had certain differences with the range (as shown in Table 2) composed of the minimum and maximum actual percentages for the corresponding severity class.

Severity assessment results for the acquired diseased wheat leaves with the actual percentages of lesion areas according to the determined reference ranges

For the sampling ratio of 4:1, according to the determined reference ranges based on the actual percentage data contained in the training sets for all severity classes of wheat stripe rust, including the midpoint-of-two-adjacent-means-based actual percentage reference ranges and the 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas, the results of severity assessment of each diseased wheat leaf contained in the corresponding training sets are shown in Table 4. For the sampling ratio of 3:2, according to the

midpoint-of-two-adjacent-means-based actual percentage reference range and the 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas determined based on the actual percentage data contained in each training set that was called Train30_R, the results of severity assessment of each diseased wheat leaf contained in the corresponding training set are also shown in Table 4. The results demonstrated that satisfactory assessment accuracies for the training sets could be achieved by using each set of the determined reference ranges for all the severity classes of wheat stripe rust, and that the assessment accuracy for each training set (Train40_R or Train30_R) was not lower than 85%. For the sampling ratio of 4:1, according to the midpoint-of-two-adjacent-means-based actual percentage reference ranges and the 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas based on the actual percentage data contained in the training sets for all the severity classes, among the assessment accuracies for all the corresponding training sets, the lowest accuracies were 85.00%, 87.50%, 95.00%, and 95.00%, respectively. For the sampling ratio of 3:2, according to the midpoint-of-two-adjacent-means-based actual percentage reference ranges and the 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas based on the actual percentage data contained in the training sets for all the severity classes, the lowest accuracies were 86.67%, 90.00%, 96.67%, and 96.67%, respectively, among the assessment accuracies for all the corresponding training sets. On the whole, for the sampling ratio of 4:1 or 3:2, the severity assessment results of all the diseased wheat leaves contained in the training set of a severity

class according to the 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas for the corresponding severity class indicated that the 99% reference range had the best assessment performance and that the assessment performance of the 95% reference range ranked second. Furthermore, the severity assessment results of all the diseased wheat leaves contained in the training set of a severity class according to the midpoint-of-two-adjacent-means-based actual percentage reference range and the 90% reference ranges of the actual percentages of lesion areas for the corresponding severity class indicated that the two reference ranges had the similar assessment performances.

According to the midpoint-of-two-adjacent-means-based actual percentage reference range and the 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas determined based on the actual percentage data in the training set Train40_R for a severity classes of *R*, the results of severity assessments of the diseased leaves contained in the corresponding testing set Test10_R are shown in Table 5. According to the midpoint-of-two-adjacent-means-based actual percentage reference range and the 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas determined based on the actual percentage data in the training set Train30_R for a severity class of *R*, the results of severity assessments of the diseased leaves contained in the corresponding testing set Test20_R are also shown in Table 5. The results demonstrated that satisfactory assessment accuracies for the testing sets could be achieved according to the midpoint-of-two-adjacent-means-based actual percentage reference ranges

TABLE 4 Severity assessment results of the diseased wheat leaves with the actual percentages of lesion areas contained in each training set of all the severity classes of wheat stripe rust according to the determined reference ranges.

Severity class	Training set	Assessment accuracy based on the midpoint-of-two-adjacent-means-based actual percentage reference range	Assessment accuracy based on the 90% reference range	Assessment accuracy based on the 95% reference range	Assessment accuracy based on the 99% reference range
1%	Train40 _{1%}	100.00%	95.00%	97.50%	100.00%
	Train30 _{1%}	100.00%	93.33%	96.67%	100.00%
5%	Train40 _{5%}	100.00%	95.00%	100.00%	100.00%
	Train30 _{5%}	100.00%	93.33%	100.00%	100.00%
10%	Train40 _{10%}	87.50%	90.00%	100.00%	100.00%
	Train30 _{10%}	86.67%	90.00%	96.67%	96.67%
20%	Train40 _{20%}	97.50%	95.00%	100.00%	100.00%
	Train30 _{20%}	96.67%	96.67%	100.00%	100.00%
40%	Train40 _{40%}	85.00%	95.00%	95.00%	100.00%
	Train30 _{40%}	86.67%	96.67%	96.67%	96.67%
60%	Train40 _{60%}	100.00%	92.50%	97.50%	95.00%
	Train30 _{60%}	100.00%	93.33%	96.67%	100.00%
80%	Train40 _{80%}	97.50%	90.00%	97.50%	100.00%
	Train30 _{80%}	96.67%	93.33%	96.67%	100.00%
100%	Train40 _{100%}	90.00%	87.50%	100.00%	100.00%
	Train30 _{100%}	90.00%	90.00%	100.00%	100.00%

TABLE 5 Severity assessment results of the diseased wheat leaves with the actual percentages of lesion areas contained in each testing set of all the severity classes of wheat stripe rust according to the determined reference ranges.

Severity class	Testing set	Assessment accuracy based on the midpoint-of-two-adjacent-means-based actual percentage reference range	Assessment accuracy based on the 90% reference range	Assessment accuracy based on the 95% reference range	Assessment accuracy based on the 99% reference range
1%	Test10 _{1%}	100.00%	90.00%	100.00%	100.00%
	Test20 _{1%}	100.00%	95.00%	100.00%	100.00%
5%	Test10 _{5%}	100.00%	90.00%	100.00%	100.00%
	Test20 _{5%}	100.00%	95.00%	100.00%	100.00%
10%	Test10 _{10%}	90.00%	90.00%	100.00%	100.00%
	Test20 _{10%}	90.00%	90.00%	95.00%	95.00%
20%	Test10 _{20%}	100.00%	100.00%	100.00%	100.00%
	Test20 _{20%}	100.00%	95.00%	100.00%	100.00%
40%	Test10 _{40%}	90.00%	100.00%	100.00%	100.00%
	Test20 _{40%}	85.00%	95.00%	95.00%	100.00%
60%	Test10 _{60%}	100.00%	90.00%	100.00%	100.00%
	Test20 _{60%}	100.00%	90.00%	100.00%	100.00%
80%	Test10 _{80%}	100.00%	90.00%	100.00%	100.00%
	Test20 _{80%}	100.00%	90.00%	100.00%	100.00%
100%	Test10 _{100%}	90.00%	90.00%	100.00%	100.00%
	Test20 _{100%}	90.00%	85.00%	100.00%	100.00%

and the 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas in the corresponding whole leaf areas for all the severity classes of wheat stripe rust, and that the assessment accuracy for each testing set (Test10_R or Test20_R) was not lower than 85%. For the sampling ratio of 4:1, according to the midpoint-of-two-adjacent-means-based actual percentage reference ranges and the 90% reference ranges of the actual percentages of lesion areas based on the actual percentage data contained in the training sets for all the severity classes, the lowest accuracies were both 90.00% among the assessment accuracies for all the corresponding testing sets; according to the 95% and 99% reference ranges of the actual percentages of lesion areas based on the actual percentage data contained in the training sets for all the severity classes, the assessment accuracies for all the corresponding testing sets were 100.00%. For the sampling ratio of 3:2, according to the midpoint-of-two-adjacent-means-based actual percentage reference ranges and the 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas based on the actual percentage data contained in the training sets for all the severity classes, the lowest accuracies were 85.00%, 85.00%, 95.00%, and 95.00%, respectively, among the assessment accuracies for all the corresponding testing sets. Overall, for the sampling ratio of 4:1 or 3:2, the severity assessment results of all the diseased wheat leaves contained in the testing set of a severity class according to the midpoint-of-two-adjacent-means-based actual percentage reference range and the 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas for the corresponding severity class demonstrated that the 99%

reference range had the best assessment performance, the assessment performance of the 95% reference range ranked second, that of the 90% reference range ranked third, and that of the midpoint-of-two-adjacent-means-based actual percentage reference range ranked last. The assessment performance of the midpoint-of-two-adjacent-means-based actual percentage reference range, the 90% reference range, the 95% reference range, or the 99% reference range determined based on the actual percentage data in the training set Train40_R for a severity class of *R* when the reference range was used to assess all the diseased wheat leaves contained in the testing set Test10_R, was better than that of the corresponding reference range determined based on the actual percentage data in the training set Train30_R for the severity class of *R* when it was used to assess all the diseased wheat leaves contained in the testing set Test20_R.

The results demonstrated that according to the two developed methods based on the reference ranges of the percentages of lesion areas for severity assessment of wheat stripe rust in this study, high accuracy can be obtained in the severity assessments of the diseased leaves, indicating that the two methods were suitable for the severity assessment of the disease. In the practical applications, the midpoint-of-two-adjacent-means-based actual percentage reference ranges can be used to carry out severity assessment of wheat stripe rust, or according to the accuracy requirements for the severity assessment results, a set of reference ranges can be selected for severity assessment of the disease from the 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas corresponding to all the severity classes.

Discussion

In this study, two new methods for severity assessment of wheat stripe rust were proposed based on the actual percentages of lesion areas in the areas of the corresponding whole wheat leaves. The main characteristics of the two proposed methods are shown in Table 6. By using the methods, the suitable reference range selected from the midpoint-of-two-adjacent-means-based actual percentage reference ranges and the 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas corresponding to all the severity classes of wheat stripe rust can be directly used to assess the severity of each diseased wheat leaf with the actual percentage of lesion area in the area of the corresponding whole leaf. The two methods are simple, easy-to-operate, rapid, and accurate. The methods are applicable to all plant diseases for which the severity is classified according to the ratio of lesion area to the area of the corresponding whole diseased plant unit. The method for determination of the midpoint-of-two-adjacent-means-based actual percentage reference ranges corresponding to all the disease severity classes and the method for determination of the 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas corresponding to all the disease severity classes, are provided for severity assessments of plant diseases. The basis of the two methods for disease severity assessment is very intuitive and in line with human psychological cognitive habits. The two methods are very convenient for practical operations and can improve the accuracy of plant disease severity assessment, resulting in more reliable plant disease information for diseased plant phenotyping, disease prediction and forecast, and disease management. The two methods are conducive to solve the

classification difficulties in assessing the severity of plant diseases. Especially, during severity assessments of plant diseases according to the ratio of lesion area to the area of the whole diseased plant unit, for some plant diseases such as wheat stripe rust and wheat leaf rust, the ratio of the lesion area to the area of the whole diseased plant unit corresponding to a severity class in the disease severity grading standard is not the actual ratio of the lesion area to the area of the whole diseased plant unit, which can induce great errors or complete errors in the severity assessment results. This problem was well solved in this study, which provided a basis and methodological reference for accurate severity assessments of plant diseases and was of great significance for survey, monitoring, prediction, and control of plant diseases.

In this study, 50 single diseased wheat leaf images for each severity class of wheat stripe rust were acquired and the actual percentages of the lesion areas in the corresponding whole leaf areas were obtained. For each severity class of the disease, the training sets and testing sets were constructed by using the system sampling method with two sampling ratios of 4:1 and 3:2. The representative values of the actual percentages of lesion areas corresponding to each severity class for the two sampling ratios had no obvious difference. For the sampling ratio of 4:1 or 3:2, high assessment accuracies for the training set and testing set were achieved according to the midpoint-of-two-adjacent-means-based actual percentage reference range and the 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas in the corresponding whole leaf areas for each severity class. In comparison, by and large, the assessment performance of the midpoint-of-two-adjacent-means-based actual percentage reference range, the 90% reference range, the 95% reference range, or the 99% reference range determined

TABLE 6 The main characteristics of the two proposed methods for determining the reference ranges of the actual percentages of lesion areas in the corresponding whole leaf areas for all the severity classes of wheat stripe rust.

Method	Devices used to acquire images	Image assessment method	Mathematical algorithms used for image evaluation	Statistical evaluation of the data obtained from the images	The determined reference range	Reliability of the obtained results
Method 1	Nikon D700 digital camera, HUAWEI P30 smartphone, and iPhone 6S	Manual image segmentation and pixel statistics in the Adobe Photoshop software.	Calculation of the actual percentage (r) of the lesion area in the area of the whole diseased leaf by using the formula: $r = \frac{A_d}{A_l} \times 100\%$ where A_d is the pixel number of the lesion region/regions in the diseased leaf image, and A_l is the pixel number of the whole leaf region in the diseased leaf image.	Mean, standard deviation, the midpoint value of the means of the actual percentages of lesion areas of two adjacent severity classes.	The midpoint-of-two-adjacent-means-based actual percentage reference ranges.	Assessment accuracy $\geq 85.00\%$.
Method 2	smartphone.			Mean, standard deviation, normal distribution test, normal distribution method for determining the bilateral $100(1-\alpha)\%$ reference ranges and the unilateral $100(1-\alpha)\%$ reference ranges by combining the normal distribution probability density function (<i>normpdf</i>) and the functions <i>solve</i> and <i>normspec</i> in the MATLAB software.	The 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas.	Assessment accuracy $\geq 85.00\%$ for the 90% reference ranges, and $\geq 95.00\%$ for the 95% and 99% reference ranges.

based on the actual percentage data in the training set Train40_R constructed by using the sampling ratio of 4:1 was better than that of the corresponding reference range determined based on the actual percentage data in the training set Train30_R constructed by using the sampling ratio of 3:2. If more images of the single diseased wheat leaves for each severity class of the disease can be acquired, the more ideal reference ranges for each severity class may be obtained by using the proposed methods in this study, and thus the better severity assessment results may be achieved.

The results obtained in this study showed that the actual percentage of lesion area in the area of a whole diseased leaf corresponding to each severity class of wheat stripe rust had great difference with the percentage of the lesion area in the area of the whole diseased leaf corresponding to the severity class in the disease severity grading standard, which is consistent with the results obtained by Shang et al. (1990). The maximum actual percentage of the lesion area in the area of the whole diseased wheat leaf with the most severe disease symptom among the collected diseased wheat leaves obtained by using image processing technology in this study was 36.49%, and it was higher than the maximum actual uredinium coverage rate of 35% obtained by Shang et al. (1990) via actual measurement of the selected wheat leaf with the most severe disease symptom. The maximum actual percentage of lesion area obtained in this study should be more close to the true value of the percentage of the lesion area in the area of the whole diseased wheat leaf with the most severe disease symptom of wheat stripe rust. Therefore, it is believed that in this study, whether the sampling ratio was 4:1 or 3:2, each set of the determined midpoint-of-two-adjacent-means-based actual percentage reference ranges could cover all possible actual percentages of lesion areas in the corresponding whole leaf areas for all the severity classes of wheat stripe rust, and each set of the determined 99% reference ranges of the actual percentages of lesion areas could basically cover all possible actual percentages of lesion areas in the corresponding whole leaf areas for all the severity classes of the disease. Each set of the determined 90% reference ranges of the actual percentages of lesion areas or each set of the determined 95% reference ranges of the actual percentages of lesion areas could basically meet the accuracy requirements of severity assessment of wheat stripe rust, although there were gaps between the reference ranges of some adjacent severity classes. If necessary, an actual percentage of lesion area falling into a gap can be assessed as the severity class corresponding to the nearest reference range according to the nearest percent estimate principle (by taking the value of the midpoint of the gap as the demarcation point). In practice, a set of the determined midpoint-of-two-adjacent-means-based actual percentage reference ranges or a set of the determined 99% reference ranges of the actual percentages of lesion areas can be selected, aiming to use a set of reference ranges that can cover all possible actual percentages of lesion areas in the corresponding whole

leaf areas for all the severity classes to carry out disease severity assessment.

In this study, when determining the reference ranges for disease severity assessment, the standard deviation was directly used, rather than the standard error. The difference between reference range and confidence interval should be paid attention to. The confidence interval is the estimation interval of a population parameter obtained by the sample statistics. When determining a confidence interval, the standard error is directly used, rather than the standard deviation. In this study, based on the constructed training sets, the 90%, 95%, and 99% confidence intervals were also estimated (as shown in [Supplementary Table 1](#)), and then by using these different confidence intervals as the reference ranges, the severity assessments of the diseased wheat leaves with the actual percentages of lesion areas in the areas of the whole diseased leaves contained in the corresponding training sets and testing sets were conducted, but the obtained assessment accuracies (as shown in [Supplementary Table 2](#)) were not high. The results indicated that in order to obtain satisfactory severity assessment results, the reference ranges for disease severity assessments should be determined by using the methods proposed in this study.

In the field of medicine, the reference ranges of the normal values of various medical indicators are the normal fluctuation ranges of the corresponding indicators of the vast majority of normal people, and they are used to evaluate whether the measured corresponding indicators are normal and can provide a basis for disease diagnosis, health assessment, and disease treatment (Horn and Pesce, 2003; Sun and Xu, 2014; Haeckel et al., 2021; Yang et al., 2022). There are many methods to determine the medical reference ranges (Horn and Pesce, 2003; Sun and Xu, 2014; Haeckel et al., 2021; Yang et al., 2022). A medical reference range is usually determined by using normal distribution method or percentile method, mainly depending on whether the related data conform to a normal distribution (Horn and Pesce, 2003; Sun and Xu, 2014; Haeckel et al., 2021). When the related data of the corresponding indicator conform to a normal distribution or can be transformed into a normal distribution via data transformation, normal distribution method can be used to estimate the reference range, otherwise, when the data do not conform to a normal distribution, percentile method can be used (Horn and Pesce, 2003; Sun and Xu, 2014; Haeckel et al., 2021). In this study, the reference ranges at different levels (90%, 95%, and 99%) of the actual percentages of lesion areas corresponding to each severity class of wheat stripe rust were estimated by referring to the method for determining medical reference ranges, and the determined 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas corresponding to each severity class can be considered as the fluctuation ranges of the actual percentages of lesion areas of 90%, 95%, and 99% of diseased leaves of the corresponding severity class, respectively. In this study, 50 actual percentages of the lesion areas in the areas of the corresponding

whole leaves of each severity class of wheat stripe rust had a normal distribution, and the actual percentages of lesion areas contained in each constructed training set conformed to a normal distribution, so the normal distribution method was used to determine the 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas corresponding to each severity class of the disease. When the method proposed in this study is used to determine the reference ranges at different levels of the actual percentages of lesion areas corresponding to each severity class of a plant disease, the determination method of the reference ranges should be modified or changed if the actual percentage data do not conform to a normal distribution. The actual percentage data can be transformed into a normal distribution through data transformation and then the normal distribution method can be used to determine the reference ranges, or the other methods including the percentile method can be used (Horn and Pesce, 2003; Sun and Xu, 2014; Haeckel et al., 2021).

In this study, the images of the single diseased wheat leaves of all the severity classes of wheat stripe rust were acquired by using digital camera and smartphones, the segmented leaf images and the segmented lesion images were obtained by using manual image segmentation method in the Adobe Photoshop 2022 software, then the numbers of the whole leaf region pixels and lesion region pixels of each single wheat leaf were achieved by viewing the histogram panel in the software, and subsequently the actual percentage of the lesion area in the area of each whole diseased leaf was calculated for further data processing. In terms of obtaining the actual ratios of lesion areas to the areas of the corresponding whole diseased plant units, in addition to the method of obtaining the actual percentages of lesion areas in the areas of the corresponding whole diseased wheat leaves used in this study, automatic image processing methods can be used to carry out disease image segmentation and obtain the actual ratios of lesion areas to the areas of the corresponding whole diseased plant units by programming or by using the developed software and packages such as APS Assess (Lamari, 2008), ImageJ (Schneider et al., 2012), Leaf Doctor (Pethybridge and Nelson, 2015), and the pliman package (Olivoto et al., 2022), and in some situations, the graph paper method (Li et al., 2011) and the paper-weighing method (Li et al., 2011) can be used to achieve the actual ratios of lesion areas. After obtaining the actual ratios of lesion areas to the areas of the corresponding whole diseased plant units for a plant disease, the reference ranges corresponding to all the plant disease severity classes can be determined according to the methods proposed in this study and then can be used to carry out the disease severity assessment, or the disease severity assessment can be directly carried out according to the severity grading standard established based on the actual ratios of lesion areas to the areas of the corresponding whole diseased plant units.

At present, in the studies and practical applications of plant disease severity assessment based on image processing technology, the severity classes are determined according to the ratios of segmented lesion areas to the areas of the corresponding whole diseased plant units (Chen et al., 2008; Guan et al., 2010; Li et al., 2011; Barbedo, 2014; Shrivastava et al., 2015; Jiang et al., 2021) or identified by using the established recognition models based on the extracted image features (Bai et al., 2011; Wang et al., 2017; Bao et al., 2018; Bao et al., 2021). However, in the reported studies on the severity assessment of some plant diseases such as wheat stripe rust based on the ratios of lesion areas to the areas of the corresponding whole disease plant units, it was not taken into account that the actual ratios of lesion areas for each disease severity class are obviously lower than the corresponding ratios of lesion areas of the estimated severity class according to the corresponding severity grading standard. The previous understanding of plant disease severity in the plant disease severity assessment can be corrected by using the two methods for disease severity assessment proposed in this study, which will greatly improve the accuracy of plant disease severity assessment and the reliability of plant disease monitoring and early warning information based on image processing technology. Some basis and research ideas for the realization of automatic assessment of plant disease severity based on image processing technology were provided in this study, which is conducive to the automation and intellectualization of plant disease severity assessment and is helpful to improve the levels of disease survey, disease monitoring and early warning, and disease management, thus providing more reliable supports for diseased plant phenotyping, disease monitoring, disease prediction and forecast, and disease control strategy making.

Conclusion

Two new methods were developed based on the reference ranges of the actual percentages of lesion areas for severity assessment of wheat stripe rust in this study. Based on the acquired single diseased wheat leaf images of all the severity classes of the disease, the actual percentage of the lesion area in the area of the corresponding whole diseased leaf for each disease image was obtained by using image processing technology, the training sets and testing sets were constructed by using the system sampling method with two sampling ratios, then the methods were developed for determination of the midpoint-of-two-adjacent-means-based actual percentage reference ranges and the reference ranges of the actual percentages of lesion areas at different levels for all the severity classes, and simultaneously the corresponding detailed reference ranges were provided. The satisfactory assessment accuracies for the training and testing sets were achieved according to the determined midpoint-of-two-adjacent-means-based actual percentage reference ranges

and the estimated 90%, 95%, and 99% reference ranges of the actual percentages of lesion areas for all the severity classes. In this study, two simple and practical methods were provided for the severity assessment of wheat stripe rust and a reference was provided for accurate severity assessments of plant diseases.

Data availability statement

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding author.

Author contributions

HGW contributed conception of the study and designed the experiments. QJ and HLW performed the experiments. QJ and HGW analyzed the data. QJ and HGW wrote the draft of the manuscript. All authors contributed to manuscript revision, read and approved the final version of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Key Basic Research Program of China (Grant No. 2021YFD1401001 and

2018YFD0200402) and the National Natural Science Foundation of China (Grant No. 32072357).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.1002627/full#supplementary-material>

References

- Ali, S., Rodriguez-Algaba, J., Thach, T., Sørensen, C. K., Hansen, J. G., Lassen, P., et al. (2017). Yellow rust epidemics worldwide were caused by pathogen races from divergent genetic lineages. *Front. Plant Sci.* 8 (1057). doi: 10.3389/fpls.2017.01057
- Bai, T. C., Zhang, R., Meng, H. B., and Wang, L. (2011). A method of estimating red jujube blade disease severity based on computer vision. *J. Tarim Univ.* 23, 72–78. doi: 10.3969/j.issn.1009-0568.2011.04.012
- Bao, W. X., Lin, Z., Hu, G. S., Liang, D., Huang, L. S., and Yang, X. J. (2021). Severity estimation of wheat leaf diseases based on RSTCNN. *Trans. Chin. Soc. Agric. Mach.* 52, 242–252, 263. doi: 10.6041/j.issn.1000-1298.2021.12.026
- Bao, W. X., Zhao, J., Zhang, D. Y., and Liang, D. (2018). Recognition of wheat leaf diseases based on elliptic metric learning. *Trans. Chin. Soc. Agric. Mach.* 49, 20–26. doi: 10.6041/j.issn.1000-1298.2018.12.003
- Barbedo, J. G. A. (2014). An automatic method to detect and measure leaf disease symptoms using digital image processing. *Plant Dis.* 98, 1709–1716. doi: 10.1094/PDIS-03.14-0290-RE
- Beddow, J. M., Pardey, P. G., Chai, Y., Hurley, T. M., Kriticos, D. J., Braun, H. J., et al. (2015). Research investment implications of shifts in the global geography of wheat stripe rust. *Nat. Plants* 1, 15132. doi: 10.1038/nplants.2015.132
- Bock, C. H., Pethybridge, S. J., Barbedo, J. G. A., Esker, P. D., Mahlein, A. K., and Del Ponte, E. M. (2022). A phytopathometry glossary for the twenty-first century: Towards consistency and precision in intra- and inter-disciplinary dialogues. *Trop. Plant Pathol.* 47, 14–24. doi: 10.1007/s40858-021-00454-0
- Chen, X. M. (2005). Epidemiology and control of stripe rust [*Puccinia striiformis* f. sp. *tritici*] on wheat. *Can. J. Plant Pathol.* 27, 314–337. doi: 10.1080/07060660509507230
- Chen, W. Q., Wellings, C., Chen, X. M., Kang, Z. S., and Liu, T. G. (2014). Wheat stripe (yellow) rust caused by *Puccinia striiformis* f. sp. *tritici*. *Mol. Plant Pathol.* 15, 433–446. doi: 10.1111/mpp.12116
- Chen, Z. L., Zhang, C. L., Shen, W. Z., and Chen, X. X. (2008). Grading method of leaf spot disease based on image processing. *J. Agric. Mech. Res.* 30(73–75), 80. doi: 10.3969/j.issn.1003-188X.2008.11.022
- Dong, J. Z., Zeng, S. M., and Ma, Q. X. (1990). The relationships between incidence and severity of leaf rust, powdery mildew and yellow rust of wheat. *Acta Phytopathol. Sin.* 20, 55–60. doi: 10.13926/j.cnki.apps.1990.01.021
- Guan, H., Zhang, C. L., and Zhang, C. Y. (2010). Grading method of cucumber leaf spot disease based on image processing. *J. Agric. Mech. Res.* 32, 94–97. doi: 10.3969/j.issn.1003-188X.2010.03.025
- Guo, Q., Wang, L. W., Dong, F. M., Nie, C. W., Sun, S. F., and Wang, J. H. (2015). Identification of wheat stripe rust and powdery mildew using orientation coherence feature. *Trans. Chin. Soc. Agric. Mach.* 46, 26–34. doi: 10.6041/j.issn.1000-1298.2015.01.005
- Haeckel, R., Wosniok, W., and Streichert, T. (2021). Review of potentials and limitations of indirect approaches for estimating reference limits/intervals of quantitative procedures in laboratory medicine. *J. Lab. Med.* 45, 35–53. doi: 10.1515/labmed-2020-0131
- Horn, P. S., and Pesce, A. J. (2003). Reference intervals: An update. *Clin. Chim. Acta* 334, 5–23. doi: 10.1016/S0009-8981(03)00133-5
- Huang, M. Y., Huang, W. J., Liu, L. Y., Huang, Y. D., Wang, J. H., Zhao, C. J., et al. (2004). Spectral reflectance feature of winter wheat single leaf infected with stripe rust and severity level inversion. *Trans. Chin. Soc. Agric. Eng.* 20, 176–180. doi: 10.3321/j.issn.1002-6819.2004.01.042
- Hu, W. W., Zhang, W., and Liu, L. Z. (2018). Identification of wheat leaf diseases based on variance-SFFS algorithm. *J. Hunan Agric. Univ. (Nat. Sci.)* 44, 225–228. doi: 10.13331/j.cnki.jhau.2018.02.021
- Jiang, X. M., Feng, H. K., Chang, H., Yang, G. J., and Yang, X. D. (2021). Classification method of wheat stripe rust disease degree based on digital image. *Jiangsu Agric. Sci.* 49, 109–115. doi: 10.15889/j.issn.1002-1302.2021.23.020

- Lamari, L. (2008). *Assess 2.0: Image analysis software for disease quantification* (Saint Paul, MN, USA: The American Phytopathological Society).
- Li, G. L., Ma, Z. H., and Wang, H. G. (2011). An automatic grading method of severity of single leaf infected with grape downy mildew based on image processing. *J. China Agric. Univ.* 16, 88–93. doi: 10.11841/j.issn.1007-4333.2011.06.014
- Li, G. L., Ma, Z. H., and Wang, H. G. (2012). Image recognition of wheat stripe rust and wheat leaf rust based on support vector machine. *J. China Agric. Univ.* 17, 72–79. doi: 10.11841/j.issn.1007-4333.2012.02.013
- Line, R. F. (2002). Stripe rust of wheat and barley in north America: A retrospective historical review. *Annu. Rev. Phytopathol.* 40, 75–118. doi: 10.1146/annurev.phyto.40.020102.111645
- Li, X. L., Qin, F., Zhao, L. L., Li, J. H., Ma, Z. H., and Wang, H. G. (2015). Identification and classification of disease severity of wheat stripe rust using near infrared spectroscopy technology. *Spectrosc. Spect. Anal.* 35, 367–371. doi: 10.3964/j.issn.1000-0593(2015)02-0367-05
- Li, Z. Q., and Zeng, S. M. (2002). *Wheat rusts in China* (Beijing: China: Agriculture Press).
- Nutter, F. W. Jr., Teng, P. S., and Shokes, F. M. (1991). Disease assessment terms and concepts. *Plant Dis.* 75, 1187–1188.
- Olivoto, T., Andrade, S. M. P., and Del Ponte, E. M. (2022). Measuring plant disease severity in R: Introducing and evaluating the pliman package. *Trop. Plant Pathol.* 47, 95–104. doi: 10.1007/s40858-021-00487-5
- Pethybridge, S. J., and Nelson, S. C. (2015). Leaf doctor: A new portable application for quantifying plant disease severity. *Plant Dis.* 99, 1310–1316. doi: 10.1094/PDIS-03-15-0319-RE
- Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. doi: 10.1038/nmeth.2089
- Shang, H. S., Ren, W. L., and Wang, S. M. (1990). A standard grading for wheat stripe rust disease. *Plant Prot.* 16, 31.
- Shrivastava, S., Singh, S. K., and Hooda, D. S. (2015). Color sensing and image processing-based automatic soybean plant foliar disease severity detection and estimation. *Multimed. Tools Appl.* 74, 11467–11484. doi: 10.1007/s11042-014-2239-0
- Sun, Z. Q., and Xu, Y. Y. (2014). *Medical statistics, 4th edn* (Beijing: People's Medical Publishing House).
- Wan, A. M., Chen, X. M., and He, Z. H. (2007). Wheat stripe rust in China. *Aust. J. Agric. Res.* 58, 605–619. doi: 10.1071/AR06142
- Wang, H. L., Jiang, Q., Cao, S. Q., Sun, Z. Y., and Wang, H. G. (2022). Effects of cultivar mixture on wheat stripe rust caused by *Puccinia striiformis* f. sp. *tritici* and wheat yield. *J. China Agric. Univ.* 27, 1–21. doi: 10.11841/j.issn.1007-4333.2022.04.01
- Wang, X. J., Ma, Z. H., Jiang, Y. Y., Shi, S. D., Liu, W. C., Zeng, J., et al. (2014). Modeling of the overwintering distribution of *Puccinia striiformis* f. sp. *tritici* based on meteorological data from 2001 to 2012 in China. *Front. Agr. Sci. Eng.* 1, 223–235. doi: 10.15302/J-FASE-2014025
- Wang, H. G., Ma, Z. H., Wang, T., Cai, C. J., An, H., and Zhang, L. D. (2007). Application of hyperspectral data to the classification and identification of severity of wheat stripe rust. *Spectrosc. Spect. Anal.* 27, 1811–1814. doi: 10.3964/j.issn.1000-0593.2007.09.037
- Wang, M. L., Niu, X. J., Zhang, H. M., Zhao, J. B., and He, D. J. (2014). Research on feature extraction and recognition of common diseases of wheat leaf. *Comput. Eng. Appl.* 50, 154–157. doi: 10.3778/j.issn.1002-8331.1308-0316
- Wang, H., Qin, F., Ruan, L., Wang, R., Liu, Q., Ma, Z. H., et al. (2016). Identification and severity determination of wheat stripe rust and wheat leaf rust based on hyperspectral data acquired using a black-paper-based measuring method. *PloS One* 11, e0154648. doi: 10.1371/journal.pone.0154648
- Wang, G., Sun, Y., and Wang, J. X. (2017). Automatic image-based plant disease severity estimation using deep learning. *Comput. Intell. Neurosci.* 2017, 2917536. doi: 10.1155/2017/2917536
- Wan, A. M., Zhao, Z. H., Chen, X. M., He, Z. H., Jin, S. L., Jia, Q. Z., et al. (2004). Wheat stripe rust epidemic and virulence of *Puccinia striiformis* f. sp. *tritici* in China in 2002. *Plant Dis.* 88, 896–904. doi: 10.1094/PDIS.2004.88.8.896
- Wan, A. M., Zhao, Z. H., and Wu, L. R. (2003). Reviews of occurrence of wheat stripe rust disease in 2002 in China. *Plant Prot.* 29, 5–8. doi: 10.3969/j.issn.0529-1542.2003.02.001
- Wellings, C. R. (2011). Global status of stripe rust: A review of historical and current threats. *Euphytica* 179, 129–141. doi: 10.1007/s10681-011-0360-y
- Yang, D., Su, Z. H., and Zhao, M. (2022). Big data and reference intervals. *Clin. Chim. Acta* 527, 23–32. doi: 10.1016/j.cca.2022.01.001
- Zhao, J. L., Huang, L. S., Huang, W. J., Zhang, D. Y., Yuan, L., Zhang, J. C., et al. (2014). Hyperspectral measurements of severity of stripe rust on individual wheat leaves. *Eur. J. Plant Pathol.* 139, 407–417. doi: 10.1007/s10658-014-0397-6



OPEN ACCESS

EDITED BY
Gregorio Egea,
University of Seville, Spain

REVIEWED BY
Giorgio Mariano Balestra,
University of Tuscia, Italy
Sara Francesconi,
University of Tuscia, Italy

*CORRESPONDENCE
Fernanda Leiva
Fernanda.leiva@slu.se

SPECIALTY SECTION
This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

RECEIVED 02 August 2022
ACCEPTED 22 September 2022
PUBLISHED 18 October 2022

CITATION
Leiva F, Zakieh M, Alamrani M,
Dhakal R, Henriksson T, Singh PK
and Chawade A (2022)
Phenotyping Fusarium head
blight through seed morphology
characteristics using RGB imaging.
Front. Plant Sci. 13:1010249.
doi: 10.3389/fpls.2022.1010249

COPYRIGHT
© 2022 Leiva, Zakieh, Alamrani, Dhakal,
Henriksson, Singh and Chawade. This is
an open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use,
distribution or reproduction is
permitted which does not comply
with these terms.

Phenotyping Fusarium head blight through seed morphology characteristics using RGB imaging

Fernanda Leiva^{1*}, Mustafa Zakieh¹, Marwan Alamrani¹,
Rishap Dhakal¹, Tina Henriksson², Pawan Kumar Singh³
and Aakash Chawade¹

¹Department of Plant Breeding, Swedish University of Agricultural Sciences, Lomma, Sweden,
²Lantmännen Lantbruk, Svalöv, Sweden, ³International Maize and Wheat Improvement Center
(CIMMYT), Texcoco, Mexico

Fusarium head blight (FHB) is an economically important disease affecting wheat and thus poses a major threat to wheat production. Several studies have evaluated the effectiveness of image analysis methods to predict FHB using disease-infected grains; however, few have looked at the final application, considering the relationship between cost and benefit, resolution, and accuracy. The conventional screening of FHB resistance of large-scale samples is still dependent on low-throughput visual inspections. This study aims to compare the performance of two cost-benefit seed image analysis methods, the free software “SmartGrain” and the fully automated commercially available instrument “Cgrain Value™” by assessing 16 seed morphological traits of winter wheat to predict FHB. The analysis was carried out on a seed set of FHB which was visually assessed as to the severity. The dataset is composed of 432 winter wheat genotypes that were greenhouse-inoculated. The predictions from each method, in addition to the predictions combined from the results of both methods, were compared with the disease visual scores. The results showed that Cgrain Value™ had a higher prediction accuracy of $R^2 = 0.52$ compared with SmartGrain for which $R^2 = 0.30$ for all morphological traits. However, the results combined from both methods showed the greatest prediction performance of $R^2 = 0.58$. Additionally, a subpart of the morphological traits, namely, width, length, thickness, and color features, showed a higher correlation with the visual scores compared with the other traits. Overall, both methods were related to the visual scores. This study shows that these affordable imaging methods could be effective to predict FHB in seeds and enable us to distinguish minor differences in seed morphology, which could lead to a precise performance selection of disease-free seeds/grains.

KEYWORDS

Fusarium head blight, seed phenotyping, seed morphological characters, wheat, visual scores, SmartGrain, Cgrain Value™

Introduction

In the countries of the Baltic Sea region, the most widely cultivated crop is winter wheat (*Triticum aestivum* L.), (Shiferaw et al., 2013; Chawade et al., 2018). While efforts are made to achieve sustainable intensification of high grain yields in wheat production, the emergence and increase in the virulence of plant pathogens conversely leave the nutritional integrity and production of wheat grains at risk (Castro Aviles et al., 2020). The decrease in grain quality and protein content negatively impacts the use of the grains and therefore affects food security and safety (Asseng et al., 2019). Fusarium head blight (FHB), mainly caused by the fungus *Fusarium graminearum* Schwabe [teleomorph: *Gibberella zeae* (Schwein) Petch], is one of the wheat diseases with a major impact on wheat grain yield and quality. FHB can dramatically reduce grain quality and yield through the formation of sterile and wizened florets. FHB-infected grains suffer from major marketing, consumption, and processing constraints, which is the buildup of mycotoxins—mainly deoxynivalenol (DON) (Del Ponte et al., 2022). DON inhibits protein synthesis, cutting off normal cell function, which is hazardous for the consumption of humans and animals (Polak-Śliwińska and Paszczyk, 2021). FHB disease management strategies rely on integrating several cultural practices such as fungicide treatment, crop rotation, mixed culture, and tillage (Gilbert and Haber, 2013). However, growing FHB-resistant cultivars is seen as a more sustainable and durable strategy for mitigating disease epidemics, thus avoiding large economic losses. Hence, identifying sources of novel resistance is a key component in pre-breeding activities that can be introgressed to develop commercial FHB-resistant cultivars.

The resistance components for FHB, commonly known as resistance types, have been defined into type I to type V (Mesterhazy, 2020): type I is resistance to initial infection, type II is resistance to disease spread (Schroeder and Christensen, 1963), type III is resistance to damage of Fusarium-damaged kernels (FDK), type IV is resistance to the buildup of DON toxins, and type V is tolerance. Traditionally, studies on FHB resistance have relied on measuring the symptoms in spikes and kernels (resistance types II and III). Type II is assessed by rating the visual symptoms on the spikes, which appear as bleached, yellowish or discolored, and stunted (Zakieh et al., 2021; Steed et al., 2022). FDK is quantified traditionally by estimating the amount of visibly damaged kernels, which appear smaller, shriveled, and in a range of colors from pale pink to brown (Delwiche et al., 2010), according to a predetermined scale for visual assessments or by employing manual tools (Ackerman et al., 2022). Comparisons between both types of resistance (resistance types II and III) have revealed that it would be more efficient and consistent to estimate FHB than the degree of colonization on the spike (Agostinelli, 2009; Balut et al., 2013; Khaeim et al., 2019; Ackerman et al., 2022). However, screening

by either manual or visual assessments is a labor- and time-consuming process for rating genotypes, is biased due to the subjectivity of visual assessments, and has low reproducibility among experiments (Barbedo et al., 2015; Khaeim et al., 2019). As a result of the previously cited limitations, the use of image analysis approaches has been investigated to evaluate FDK, particularly in estimating morphological characteristics. However, the existing different imaging approaches have their disadvantages and trade-off in terms of costs, time expenses, resolution, and precision when considering an application (Saccon et al., 2017).

Among the investigated methods, Iwata and Ukai (2002) and Iwata et al. (2010) investigated changes in grain shape using elliptic Fourier descriptors of two- and three-dimensional features from vertically and horizontally located seed images. Despite the accuracy reached, there are limitations in terms of image resolution and regarding the manual handling of samples during the procedure. Menesatti et al. (2009) presented a method to classify FHB in wheat-infected kernels—according to the shape criteria—into the following groups: chalky, shriveled, or healthy. The method proved to be functional to categorize kernels as chalky or healthy, but not for shriveled or gravely affected samples. Jirsa and Polišenská (2011) developed a model for the identification of Fusarium-damaged wheat kernels using image analysis. The characterization of healthy or damaged kernels based on color parameters revealed a high accuracy compared with the shape and DON content parameters. However, image processing was done with manual selections and comparing only 40 kernels—either heavily damaged or healthy—without considering any halfway stage. Similarly, the use of hyperspectral imaging for detecting Fusarium sp. in seeds has been previously investigated (Delwiche et al., 2010; Shahin and Symons, 2011; Bauriegel and Herppich, 2014; Barbedo et al., 2015; Femenias et al., 2022; Rangarajan et al., 2022; Yipeng et al., 2022). The methods have been shown to be accurate and have identified more factors involved in FDK. A more advanced technique based on X-ray computed tomography has been implemented for evaluating seed shape in finer detail (Gomes and Duijn, 2017; Liu et al., 2020). Nevertheless, inconsistencies because of specular reflection, correct wavelength selection, kernel orientation, selection of reference parameter, costs of acquisition devices, and the storage requirement for highly dimensional and massive data sets may be limiting the application of these methods (Dissing et al., 2013; Lu et al., 2020).

In the face of the constraints cited earlier, automated and light-weight free software for grain image analysis have been developed (Wang et al., 2009; Komyshev et al., 2017; Colmer et al., 2020; Zhu et al., 2021); some examples of them are GrainScan (Whan et al., 2014), which analyzes size and color features, and SmartGrain (Tanabata et al., 2012), which analyzes size and shape features. Both software are instantaneous in image recognition despite the position, overlapping, or the

number of seeds. Alternatively, commercially available imaging instruments for grain image analysis combine hardware and software, including WinSEEDLE (Regent Instruments Inc.), Seed Count (Next Instrument Pty Ltd.), Vibe QM3 Grain Analyzer (VIBE), and Cgrain ValueTM (Cgrain AB). The instruments use optical or flatbed scanners to extract features such as size, shape, and color in the color representation hue, saturation, and light (HSL). However, SeedCount and Vibe QM3 Grain Analyzer only scan the top surface of the samples, thus omitting morphological characteristics that are not in the viewing area. A more advanced instrument is Videometer Lab (Videometer A/S, Denmark), which provides rapid color, shape, and texture measurements. Videometer Lab is ideal to use in analyzing kernel surfaces, but it requires certain expertise and allows the analysis of only a few samples at once.

In this context, this paper has three objectives; first is to investigate the applicability of low-cost digital image analysis to predict FHB infection in harvested grains through morphological traits. This will offer more insight into the traits that are correlated to the degree of FDK. The second objective is to compare the applicability of the two methods used for grain image analysis—SmartGrain, and Cgrain ValueTM—in terms of consistency and throughput. The third one is to illustrate the processing chain and result interpretation with a descriptive data analysis.

Materials and methods

Plant material

Wheat kernel samples were collected from an experiment under accelerated indoor growth conditions (Zakieh et al., 2021) using winter wheat genotypes from two different sources. The first source consisted of 338 genotypes (breeding set) provided by the Swedish agricultural cooperative (Lantmännen Lantbruk, Svalöv, Sweden). The second source consisted of 181 germplasm genotypes (genebank set) provided by the Nordic Genetic Resource Center (Nordgen), with highly diverse plant materials including landraces and old cultivars.

Experimental design/growth and inoculation protocol

Plants were grown following an augmented block design in a climate-controlled chamber. After germination, the plants were subjected to a vernalization period of 57 days at 3°C with 8 h of daily light at medium–high light intensity (LI) of 250 $\mu\text{mol m}^{-2} \text{s}^{-1}$. At the end of the vernalization period, the climatic conditions were adjusted with a gradual increase in temperature and LI for the acclimatization of the plants to the next phase of accelerated growth conditions. Once the acclimatization period was concluded, the

plants were left to grow at a constant temperature of 22°C. The accelerated growth conditions were adapted by exposing the plants to a prolonged daily light duration of 22 h, with LI at 400 $\mu\text{mol m}^{-2} \text{s}^{-1}$ of uniform light intensity from LED light plates. Under these accelerated growth conditions, the plants were watered daily and fertilized weekly using first a combination of a high-phosphate and high-nitrogen soluble fertilizer SW-BOUYANT 7-1-5 + Mikro + KH_2PO_4 , then only with a high-nitrogen fertilizer, and finally with a high-potassium soluble fertilizer Yara Tera Kristalon NPK 12-5-30 with S and Mikro.

After completing the anthesis stage, at 33 days post-acclimatization, the plants were moved to a glasshouse chamber with relative humidity (rh) of 60% and a constant temperature of 24°C for 24 h to allow their adaptation to the new growth conditions prior to inoculation. Thereafter, the winter wheat spikes were spray-inoculated with an inoculum suspension prepared from the harvested spore of *F. graminearum* and *F. culmorum*, with a concentration of 5×10^5 spore/ml. Subsequently, the plants were left to incubate at 90% rh with 16/8 h dark/light cycle at a constant temperature of 24°C for 48 h before adjusting the climatic conditions back to 60% rh. The plants were eventually left to grow under the latter conditions for 24 days before harvesting the seeds. Eight isolates from *F. graminearum* and *F. culmorum* species were used in inoculating the plants provided by the Swedish agricultural cooperative Lantmännen Lantbruk. An inoculum preparation was carried out by incubating the fungal spores at 24°C for 4 days in dark conditions to allow for mycelial growth on SNA media plates. Later, the fungal plates were exposed to near ultra-violet UV radiation for 10 h to induce macroconidia formation. Afterward, the fungal plates were incubated for 4 days at 24°C in dark conditions. Finally, macroconidia spores were collected to make the inoculation suspension with the provided concentration after adding the surfactant Tween[®]20 0.002% (v/v) final volume of the inoculum. A more detailed protocol is described in Zakieh et al. (2021).

FHB visual assessment

In order to evaluate FHB resistance on a large number of genotypes, a modified visual scoring of the FHB disease severity method was adopted. The method took into account the incidence of all FHB symptoms across the main tiller spike of each genotype. Therefore, disease severity was assessed as the percentage score of infected spikelets relative to all spikes, regardless of symptom continuity on the same spike. FHB development was scored at 6, 8, 10, and 12-days post-inoculation (dpi) (Stack and McMullen, 1998). The FHB disease severity scores varied between 100 to 5% for the most susceptible phenotypes and the most resistant ones, respectively. Finally, the results of the visual scores were validated by association mapping, thus identifying the quantitative trait loci of FHB resistance (Appendix 1).

Seed shape parameters

Two different grain phenotyping methods were employed in this study: an automated imaging instrument with software and hardware named Cgrain Value™ which is commercially available (Cgrain AB) and the free software named SmartGrain developed by Tanabata et al. (2012) and can be downloaded from the Quantitative Plant website (Lobet, 2017). The implementation of both methods is described in the following sections.

SmartGrain

For image acquisition, the seeds were captured with a low-cost image protocol acquisition from a top-view angle of 55 cm above the seeds and placed manually on a flat surface using a digital single-lens reflex camera Canon EOS 1300D (Canon U.S.A. Inc., Huntington, NY, USA), which has a resolution of 18 megapixels, mounted on a Kaiser RS-1 repro stand. The camera was tethered to the software digiCamControl (Istvan, 2014) with optimal exposure settings based on the best seed view, F-Stop 1/160, exposure time 1/10, and ISO 800. The seeds were placed manually per genotype uniformly on a blue cardboard that was used as a background on a stand aside from a 15-cm ruler for further analysis. Digital images were stored with $3,456 \times 2,304$ -pixel resolution in JPEG format (Figure 1, top images).

The image analysis was thereafter carried out using SmartGrain software following its default protocol (Tanabata

et al., 2012). Briefly, the image scale was set up by taking a known sample from the ruler and registering it on the software. Then, the segmentation method by color was chosen, the precision sensibility was set at the minimum value of “1”, and the seed detection intensity was at a maximum value of “4” to obtain all possible shape details; the rest of the parameters were set to default. Finally, all the processed images were saved as TIFF files, and the results were saved in a CSV format. The software provides seven morphological characteristics: area seed (AS), perimeter length (PL), length (L), width (W), length-to-width ratio (LWR), circularity of the seed (CS), distance between the intersection of length and width, and the center of gravity (DS). AS corresponds to the total number of pixels of the segmented seed, this parameter estimates the seed size. PL refers to the length measurement of the seed outline. L corresponds to the major length measurement in the axis and W to the minor length axis measurement. CS estimates how round the region of interest is (seed), and it is calculated as $\frac{4 \times \pi \times AS}{PL^2}$. LWR is calculated by $\frac{L}{W}$, and it provides an idea of the seed shape between rectangular and circular depending on the value. The distance between the transverse axis from the outline of the seed (IS) and the center of gravity (CG) is used to estimate DS [described in detail by Tanabata et al. (2012)].

Cgrain Value™

For single kernel analysis, seeds were scanned with Cgrain Value™, which is an analytical imaging instrument. The device

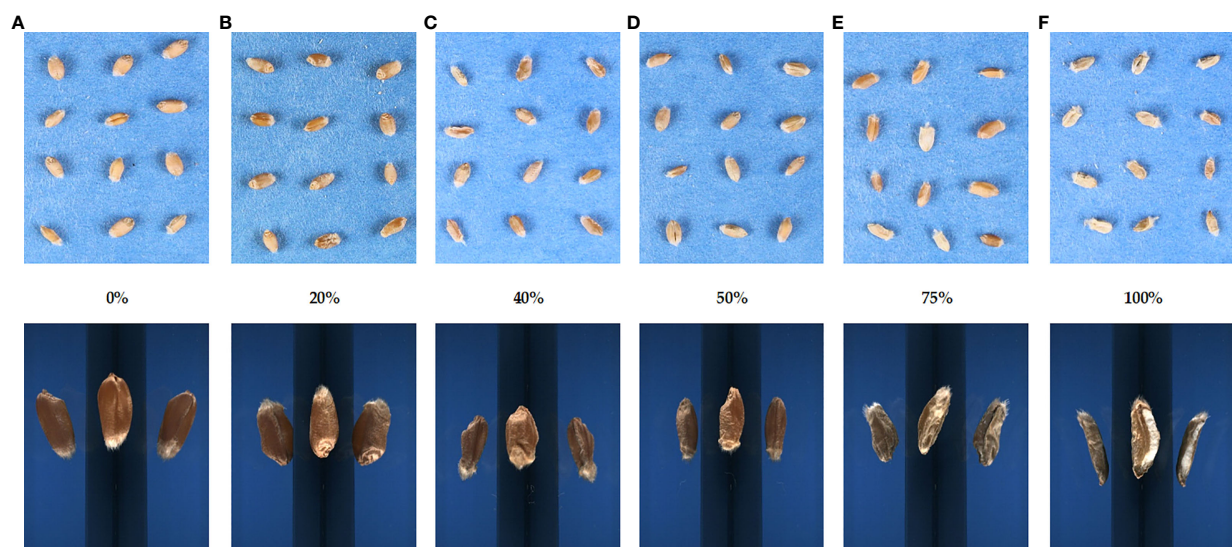


FIGURE 1
Images of the different levels of Fusarium head blight severity on winter wheat seeds. The rating of disease severity ranged from (A) 0 to (F) 100%. Scoring was based on the proportion of total infected spikes to the total amount of spikes. The top images were obtained for the SmartGrain analysis, and the bottom images were acquired using the Cgrain Value™ instrument.

inspects each kernel through a unique mirror design covering more than 90% of the grains' surfaces in every image. The analysis starts by pouring into the metal bowl of the Cgrain Value™ a batch of seeds per line and per genotype. The seeds rotate into the bowl and then, one by one, are photographed and analyzed simultaneously. After the analysis is completed, three different reports are created (result file, stat file, and image file). The result file consists of the morphological characteristics for each batch of seeds (seed count, thousand kernels, *etc.*), the stat file provides data per individual seed of a group (length, width, *etc.*), and the image file corresponds to the single seed images acquired (Figure 1, bottom images).

The instrument provides nine morphological attributes: length (L), width (W), thickness (T), average width (AVG.W), volume (V), weight (WT), light, hue, and saturation. Parameters such as L, W, and T are estimated by taking the longitudinal measurement of the axis major, higher minor, and minor, respectively. In the case of AVG.W, as the seed is received as a three-dimensional image, the measurement is referring to the mean of the average curvature. V corresponds to the seed volume obtained from the 3D image. For WT, the device has an internal balance, so while acquiring the image, it also weighs the grain. Color parameters, hue, saturation, and light are also determined by the instrument; it specifies the color base of a sample, how saturated it is, and how bright it is, respectively.

Statistical analysis

Statistical analyses were conducted using R (Team, R. C, 2013). The visual scorings of the last time-point on infected spikes, including cultivars with zero symptoms, were included in a file together with the mean values per genotype of the results given by Cgrain Value™ and SmartGrain. Each replicate of the data set was filtered by missing data (NA). Those with NA along the four replicates were removed and those with presence in more than one replicate were substituted using FactoMineR (Lê et al., 2008) and missMDA (Josse and Husson, 2016) packages. Then, using the Agricolae R package (De Mendiburu, 2014), the checks in each augmented block were used to adjust the means for each trait per replicate, the model of which is as follows:

$$y_{il} = u + G_{il} + \beta_1 + \epsilon_{il}$$

where y_{il} corresponds to the adjusted means of the i^{th} wheat cultivar in the l^{th} block, u is the general mean value, G_{il} is the effect of the i^{th} wheat genotype in the l^{th} block, β_1 is the l^{th} block effect, and ϵ_{il} is the residual. Subsequently, using the adjusted means, the best linear unbiased estimates (BLUEs) was calculated using the randomized complete block design option in META-R 6.04 (Alvarado et al., 2015) based on the following model:

$$y_{ijm} = u + S_j + G_{ijm} + R_m + \epsilon_{ijm}$$

where y_{ijm} corresponds to the BLUE of the i^{th} genotype from the j^{th} population in the m^{th} replicate, u is the general mean value, S_j is the effect of the j^{th} source of material, G_{ijm} is the effect of the i^{th} genotype in the m^{th} replicate, R_m is the m^{th} replicate of the effect, and ϵ_{ijm} is the residual effect. The source of wheat genotypes S_j was considered the grouping factor.

The BLUEs data previously centered were used to predict FHB using a multiple regression model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \epsilon$$

Where for $i=n$ observations: y_i corresponds to the dependent variable, x_i to the explanatory variables, β_0 corresponds to y -intercept (constant term), β_p corresponds to the slope coefficients for each explanatory variable, and ϵ corresponds to the error of the model (also known as the residuals). Three models were created using the morphological traits provided by both methods (Cgrain Value™ and SmartGrain) as independent variables and visual scorings as the dependent variable. One model combines all the traits, and two others use the traits provided by each method. To build each model, the data set was partitioned employing the function "createDataPartition" of the caret package (Kuhn et al., 2020) into 70% for model training (training set) and the remaining 30% for evaluating model performance (test set). Subsequently, the model was fitted to the training set, and it predicted the responses using the test set. To evaluate the quality of the predictions and mitigate the possibility of errors due to the random data partitioning, the cross-validation was executed 100 times, which means resampling the data set, and the mean of the criterion was taken as the final result.

Results

This study examined a total of 16 morphological traits, including size, color, and shape of winter wheat grains from the genebank and breeding sets with different levels of FHB infection. Nine traits were obtained with the instrument Cgrain Value™ and seven traits with the software SmartGrain. The distribution of all the morphological traits measured by the two methods showed a Gaussian distribution (Figure 2). In order to understand the association between these traits and FHB resistance, a comparison with the traits measured of 80 FHB susceptible and resistant genotypes was performed. For this purpose, five genotypes per replicate (four replicates) from both sets, breeding and genebank, were selected based on the FHB severity scores on the spikes, genotypes scored as 0% (visually non-infected or resistant), and ones scored as 100% (visually infected or susceptible). Among the infected and non-infected selected groups, there was a 22.61% reduction in V and 11.32% in AS. Other parameters also showed a reduction, such as T_RAW at 10.60%, W at 8.30% in both methods, and WT at

22.63%. Additionally, L was reduced according to the results by 1.96% in Cgrain Value™ and 2.26% in SmartGrain. Similarly, CS and PL showed a decrease, but in less proportions with 4.60 and 3.25%, respectively. The minimum seed L measured was 4.59 mm for non-infected and 4.50 mm for infected genotypes. On the other hand, color parameters expressed major changes compared with all the other morphological traits. Hue and the light increased with the infection by 19.91 and 8.28%, respectively, while saturation decreased at about 15.52% (Table 1). According to the analysis of variance (two-way

ANOVA), the morphological traits L, W, T_RAW, light, and hue were highly significant ($P < 0.001$), likewise with V, CS, and saturation ($P < 0.01$), indicating a clear association with FHB disease severity level. Meanwhile, the parameters WT, AS, LWR, PL, and DS did not indicate any significance but still showed slight differences between infected and non-infected grains.

Additionally, a principal component analysis (Figure 3) was performed to show the response of all the seed traits studied regarding the disease infection and how they correlate to each other. The proportion of total variance on the two first principal

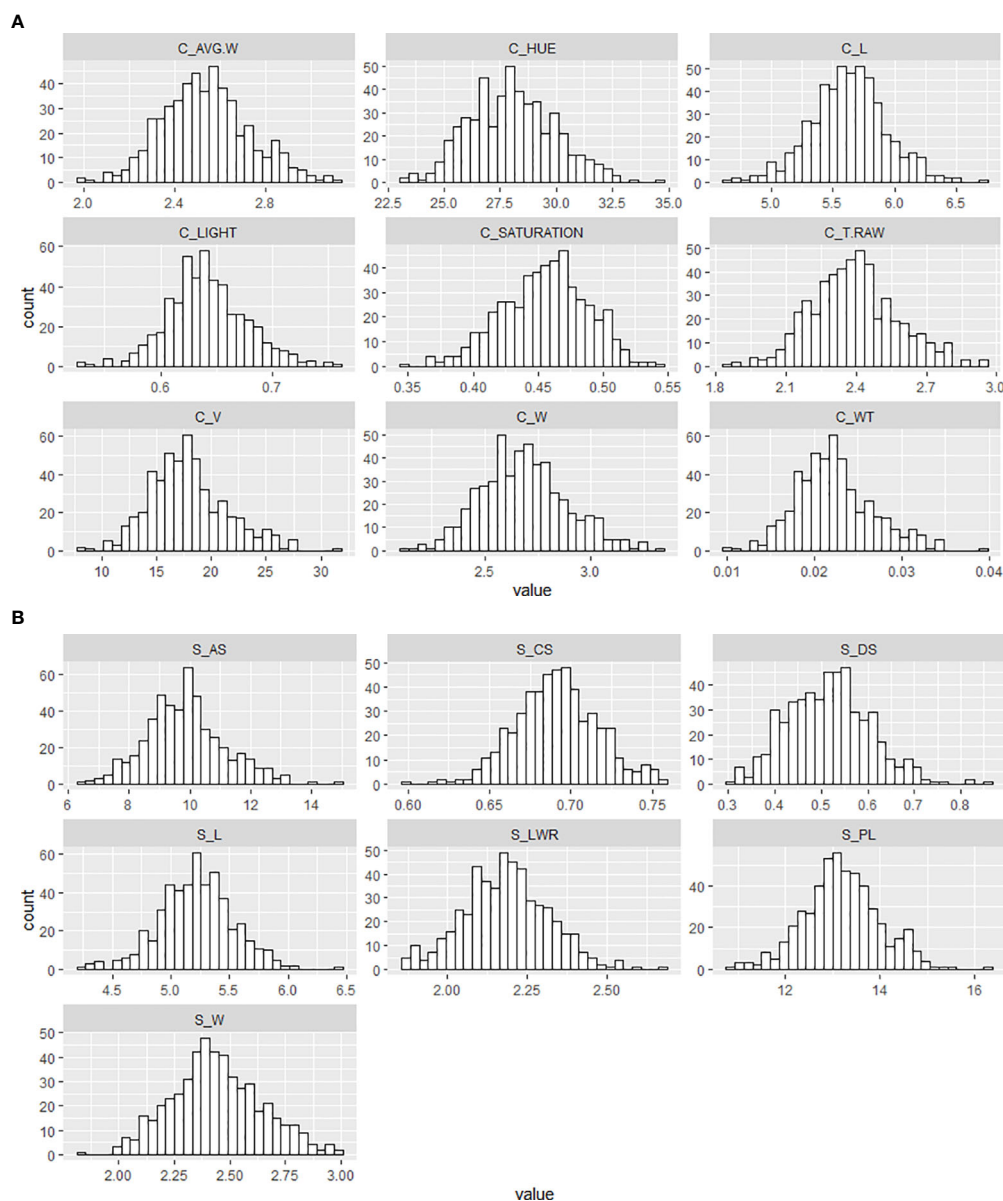


FIGURE 2

Frequency distribution of the different morphological traits of wheat genotypes seeds from the breeding and genebank sets collected with (A) the Cgrain Value™ instrument and (B) the SmartGrain software.

TABLE 1 Descriptive statistics showing differences between the seed shape characters of 80 genotypes from genebank and breeding set under non-infection (0%) and full infection (100%) FHB symptoms, with five genotypes of each one per replicate.

Description	Level	a) CGRAIN VALUE™								
		L	W	T.RAW	AVG.W	V	WT	HUE	SAT	LIGHT
Mean	Non_Infected	5.6	2.76	2.47	2.61	19.18	0.02	25.78	0.48	0.62
	Infected	5.49	2.53	2.2	2.36	14.84	0.01	30.81	0.4	0.68
% Reduction		1.96	8.29	10.6	9.41	22.61	25	-19.51	16.52	-9.67
Max	Non_Infected	6.88	3.7	3.245	3.41	38.9	0.04	30.46	0.55	0.715
	Infected	6.46	3.13	2.93	3.03	26.6	0.03	38.99	0.51	0.81
Min	Non_Infected	4.59	2.18	1.98	2.08	10.66	0.01	23.45	0.43	0.55
	Infected	4.5	2.05	1.88	1.96	7.1	0.008	24.88	0.3	0.58
SD	Non_Infected	0.52	0.36	0.3	0.32	6.74	0.008	1.32	0.02	0.04
	Infected	0.45	0.23	0.22	0.23	4.02	0.005	3.01	0.05	0.05
SE	Non_Infected	0.08	0.05	0.048	0.05	1.06	0.001	0.21	0.004	0.006
	Infected	0.07	0.04	0.036	0.04	0.63	0.0007	0.47	0.008	0.007
CV (%)		9.44	13.02	12.26	12.53	35.15	35.15	5.14	5.8	6.79
Description	Level	b) SMARTGRAIN								
		AS	PL	L	W	LWR	CS	DS		
Mean	Non_Infected	9.77	12.91	5.08	2.44	2.13	0.7	0.48		
	Infected	8.66	12.49	4.97	2.23	2.25	0.67	0.51		
% Reduction		11.32	3.25	2.26	8.27	-5.64	4.6	-6.9		
Max	Non_Infected	17.36	17.15	6.57	3.71	2.53	0.8	0.85		
	Infected	13.63	15.54	6.25	2.95	2.65	0.73	1.01		
Min	Non_Infected	3.41	7.91	3.2	1.39	1.53	0.63	0.24		
	Infected	3.01	7.31	2.88	1.36	1.88	0.61	0.23		
SD	Non_Infected	3.21	2.16	0.81	0.48	0.17	0.03	0.13		
	Infected	2.55	1.95	0.79	0.38	0.15	0.02	0.18		
SE	Non_Infected	0.5	0.34	0.12	0.07	0.02	0.005	0.02		
	Infected	0.4	0.3	0.12	0.06	0.02	0.004	0.02		
CV (%)		32.85	16.72	16.11	19.72	8.41	4.76	28.27		

a) Cgrain Value™ size, shape and color characteristics, (L) [mm], Width (W) [mm], Raw Thickness (T.RAW) [mm], Mean Width (AVG.W) [mm], Weight (WT) [g], Hue, Saturation, and Light; b) SmartGrain size and shape characteristics, Area size (AS) [mm²], Perimeter length (PL) [mm], Length (L) [mm], Width (W) [mm], Length to width ratio (LWR), Circularity (CS) Distance between IS and CG (DS) [mm].

components and correlations represents 60.50 and 19.90%, respectively, of the total variance. The LWR trait was shown to be the higher positive in the first principal component; similarly, hue was shown to be positive but in a lesser proportion. In the same component but with negative loading, we found CS as the variable with the highest contribution; the traits W from both methods, AVG.W, and T_RAW were also projected onto this component with a loading of a slightly lesser norm. Although saturation was also projected onto this component, it was shown to be the smallest loading. On the other hand, in the second principal component, the traits DS and L from both methods, PL, AS, V, and WT showed a high positive loading with similar proportions, whereas the trait light was the only one with a negative loading into the second principal component and the one with less projection among all the traits. In general, all the seed morphological traits assessed expressed variability and influence in the two principal components. In addition, as can be observed in the graph, the variation of LWR has an opposite

projection to the CS trait, expressing a good indicator to study the deformation of the grains caused by the disease infection.

Considering Table 1, the mean values for the same morphological traits measured by both methods (L and W) across the two sets, genebank and breeding, were similar. The difference between infected and non-infected seeds was 0.11 mm in L in both methods and between 0.21 and 0.25 mm in W and AVG_W. Both methods provide important parameters for seed morphology studies. Cgrain Value™ provides V and WT values and color information. Although these are important characteristics for different study purposes, mainly for identifying FHB-infected kernels, SmartGrain, in turn, provides information such as PL, AS, and CS that can show variabilities between infected and non-infected seeds. Here the BLUES for all the measured parameters were correlated with each other and in association with the visual scorings on the spikes (Figure 4). A moderate to high positive correlation was found with the color parameter hue, and a low positive

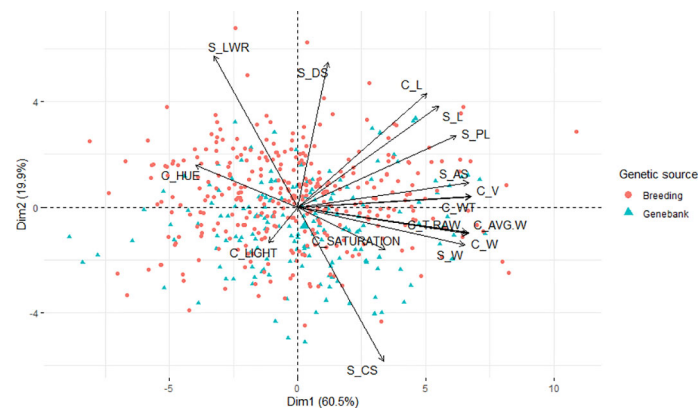


FIGURE 3

Principal component analysis biplot of the morphological traits collected with Cgrain ValueTM and SmartGrain of the breeding and genebank seeds infected with different levels of Fusarium head blight.

correlation with light was given by Cgrain ValueTM and LWR as well as given by SmartGrain ($r = 0.65$, $r = 0.36$, and $r = 0.27$, respectively). Negative correlations were also found between the visual evaluations of symptoms and the other characteristics in different levels of strength of association. There was no correlation between FHB visual scoring and DS ($r = 0.01$).

The multiple linear regression model developed to identify the contributions of the 16 different morphological traits provided by Cgrain ValueTM and SmartGrain expressed a high

moderate prediction ($R^2 = 0.58$), (Figure 5A). Aiming to identify which of both methods used in this study provides a higher prediction and also to identify the best morphological traits to predict FHB, two more models were constructed: one for the results given by Cgrain ValueTM and another one for the results of SmartGrain. The model of Cgrain ValueTM traits showed a moderate prediction ($R^2 = 0.52$), (Figure 5B). On the other hand, the model of SmartGrain traits showed medium-low prediction ($R^2 = 0.30$), (Figure 5C), clearly showing that the first model had

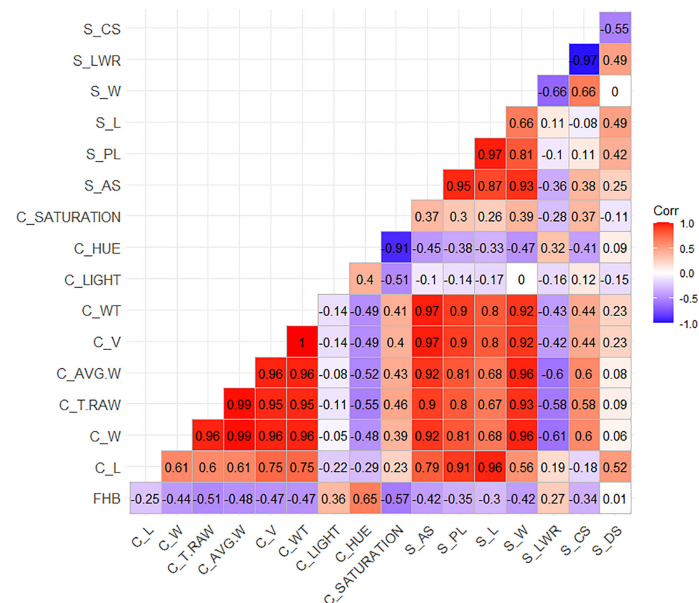


FIGURE 4

Sorted upper triangle correlation matrix among the morphological attributes of the wheat genotype seeds from the breeding and genebank sets collected with the Cgrain ValueTM and the SmartGrain software.

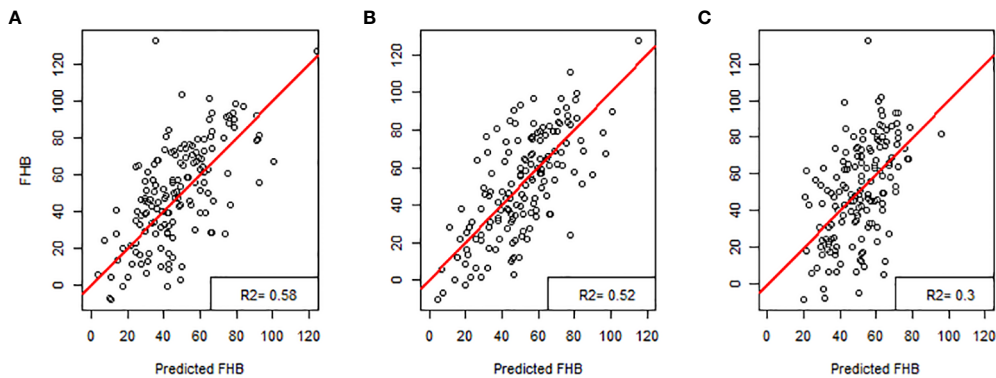


FIGURE 5
Regression models for predicting Fusarium head blight in wheat: (A) all the characteristics obtained with Cgrain Value™ and SmartGrain, (B) Cgrain Value™ morphological traits, and (C) SmartGrain morphological traits.

a higher prediction than separately. In addition, the morphological parameters that are the most suitable to assess FHB in grains above all the 16 evaluated were identified. According to the regression model and the ANOVA analysis, the parameters that provided more information about the disease are the length, width, thickness, average width, circularity, and the color parameters in the color representation HSL (Table 2). The sensitivity test showed that these variables provide the highest value of R -square, ($R^2 = 0.52$). These morphological traits are enumerated from most significant to least significant in Figure 6.

Discussion

This study compared the potential performances of two different image-based methods to predict FHB. The results of both indicated that morphological seed traits are functional for predicting FHB among two different sets of genotypes evaluated. Furthermore, a comparison of the applicability of the two methods was properly addressed by evaluating the cost, accuracy, and time efficiency—for instance, to extract dimension, shape, and color parameters, Cgrain Value™ utilizes a unique mirror design to inspect all possible angles of

TABLE 2 Summary of the multiple linear regression model combining all the 16 morphological characteristics provided by Cgrain Value™ and SmartGrain.

Model summary				
Morphological traits	Sum sq	Mean sq	F-value	Pr (>F)
C_L	23,829	23,829	64.587	6.99E-15 ***
C_W	51,079	51,079	138.446	< 2e-16 ***
C_T.RAW	40,500	40,500	109.772	< 2e-16 ***
C_AVG.W	2,013	2,013	5.456	0.0199 *
C_V	2,603	2,603	7.055	0.00816 **
C_WT	680	680	1.843	0.17526
C_LIGHT	31,656	31,656	85.802	< 2e-16 ***
C_HUE	39,386	39,386	106.752	< 2e-16 ***
C_SATURATION	2,649	2,649	7.18	0.00762 **
S_AS	178	178	0.483	0.48734
S_PL	624	624	1.691	0.1941
S_L	3,027	3,027	8.204	0.00436 **
S_W	45	45	0.121	0.72828
S_LWR	0	0	0.001	0.9802
S_CS	1,651	1,651	4.476	0.03489 *
S_DS	539	539	1.461	0.22731

The most significant characteristics concerning the Fusarium head blight disease infection according to the P-value has an *. (No significance $P>0.05$; * $P\leq0.05$; ** $P\leq0.01$; *** $P\leq0.001$).

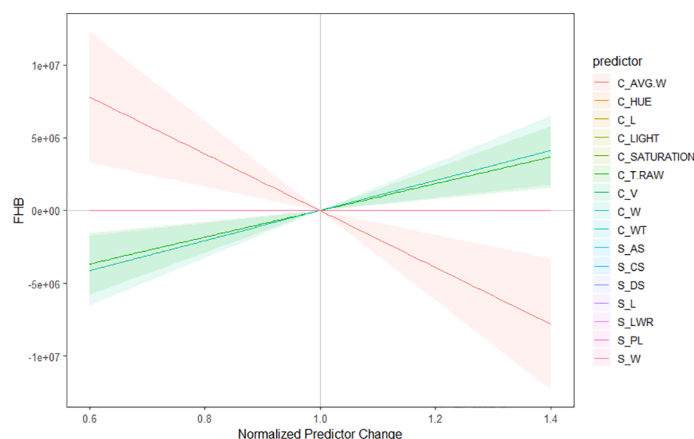


FIGURE 6

Sensitivity plot of the morphological characteristics to predict Fusarium head blight in wheat. The parameters are organized from the best predictors to the less significant to predict the disease. Color lines indicate the significance, considering red as the most important predictor and pink as the less important one. The highlighted regions reflect the correlation of the parameters among each other.

individual kernels in the sample. Additionally, image capture and processing are instantaneous, thanks to the hardware and software combination. Conversely, image acquisition using the SmartGrain system was carried out over a relatively long period, yet image processing was done relatively fast. However, compared with Cgrain ValueTM, the earlier approach is cheaper considering the cost of the tools used in image capture, requiring a simple RGB camera, a static frame, and the free software.

On the other hand, the morphological traits, based on the statistical analysis results, that showed significant correlations to the visual scores were color traits in the HSL color representation and thickness from Cgrain ValueTM, length and width, from both methods (Figures 5, 6). Although the other measured morphological traits were not significantly correlated to the visual scores, infected grains still expressed differences in these traits that may be ultimately informative about seed health and refine the prediction (Table 1). Nevertheless, DS was not correlated and did not express significant differences in infected seeds of FHB, but it could prove useful in other applications.

The evaluated visual scores of the symptoms associated with FHB—bleached, yellowish or discolored, and stunted spikes—were previously validated by the identification of several loci by genome-wide association studies (GWAS) (Appendix 1), in a previous study with the same plants and visual scorings (Zakieh et al., 2021). The proposed methods aim to replace costly and labor-intensive genetic analysis.

Therefore, the prediction of both methods studied here appears to be consistent for FHB with the assigned traits concerning the phenotype–genotype association. Previous investigations showed a high correlation between symptoms

that are present on wheat heads and the rate of kernel damage (Góral et al., 2018). Therefore, it is feasible to reference the estimated visual scores of disease severity to establish similar results of association/disassociation with the corresponding assessments of grain traits following the methodology in this study.

An important aspect to highlight is that the percentage of disease severity can be assessed, where, in contrast to disease spread from the point of inoculation, it offers less intensive labor by spray inoculation of a larger number of wheat genotypes. Additionally, unlike point-inoculated wheat spikelets, spray-inoculated spikes allow for evaluating the degree of damage caused by the disease to all kernels of the infected spike. Within this work frame, whole spike kernels are investigated for their characteristics rather than the damage to a limited number of kernels caused by Fusarium colonization from the point of inoculation. This, in turn, is expected to shorten the period for disease resistance assessment, lower its cost, and be less labor demanding.

Conclusion

The results indicated that the traits with a higher correlation to FHB were length, width, thickness, and especially color values in HSL color representation. Moreover, Cgrain ValueTM was advantageous to SmartGrain in terms of the time required for image capture and outperformed the latter when applied to a large number of samples, yet SmartGrain processes samples fast and is cheaper in comparison to Cgrain ValueTM. Although the disease prediction showed a low–moderate accuracy for SmartGrain and a high–moderate accuracy for Cgrain ValueTM and the results of both methods combined, this is attributed to the prediction reference,

which corresponds to FHB disease severity scorings done on the spikes. However, the novelty of this study resides in the accuracy reached even with a different reference source, but which is directly related. Additionally, as the plant material genotypes and visual scores were validated by GWAS analysis, then the results presented here are phenotype-genotype-associated.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

Author contributions

AC conceived the study. TH developed the breeding population set. MZ provided the material and the scores of the disease severity. MA performed the image and data acquisition with SmartGrain. RD performed the image and data acquisition with Cgrain ValueTM. FL analyzed the data and wrote the draft. All authors contributed to the article and approved the submitted version.

References

- Ackerman, A. J., Holmes, R., Gaskins, E., Jordan, K. E., Hicks, D. S., Fitzgerald, J., et al. (2022). Evaluation of methods for measuring fusarium-damaged kernels of wheat. *Agronomy* 12 (2), 532. doi: 10.3390/agronomy12020532
- Agostinelli, A. M. (2009). *Phenotypic and genotypic selection for head scab resistance in wheat*. University of Kentucky Master's Theses. 582. Available at: https://uknowledge.uky.edu/gradschool_theses/582
- Alvarado, G., López, M., Vargas, M., Pacheco, Á., Rodríguez, F., Burgueño, J., et al. (2015). META-R (Multi Environment Trial Analysis with R for Windows) Version 5.0. CIMMYT Research Data & Software Repository Network 23, 2015.
- Asseng, S., Martre, P., Maïorano, A., Rötter, R. P., O'Leary, G. J., Fitzgerald, G. J., et al. (2019). Climate change impact and adaptation for wheat protein. *Global Change Biol.* 25 (1), 155–173. doi: 10.1111/gcb.14481
- Balut, A. L., Clark, A. J., Brown-Guedira, G., Souza, E., and Van Sanford, D. A. (2013). Validation of Fhb1 and QFhs. nau-2DL in several soft red winter wheat populations. *Crop Sci.* 53 (3), 934–945. doi: 10.2135/cropsci2012.09.0550
- Barbedo, J. G., Tibola, C. S., and Fernandes, J. M. (2015). Detecting fusarium head blight in wheat kernels using hyperspectral imaging. *Biosyst. Eng.* 131, 65–76. doi: 10.1016/j.biosystemseng.2015.01.003
- Bauriegel, E., and Herppich, W. B. (2014). Hyperspectral and chlorophyll fluorescence imaging for early detection of plant diseases, with special reference to fusarium spec. infections on wheat. *Agriculture* 4 (1), 32–57. doi: 10.3390/agriculture4010032
- Castro Aviles, A., Alan Harrison, S., Joseph Arceneaux, K., Brown-Guidera, G., Esten Mason, R., and Baisakh, N. (2020). Identification of qtls for resistance to fusarium head blight using a doubled haploid population derived from southeastern united states soft red winter wheat varieties ags 2060 and ags 2035. *Genes* 11 (6), 699. doi: 10.3390/genes11060699
- Cgrain AB Cgrain value TM, the new standard for analysis grain quality. Available at: www.cgrain.se.
- Chawade, A., Armoniené, R., Berg, G., Brazauskas, G., Frostgård, G., Geleta, M., et al. (2018). A transnational and holistic breeding approach is needed for sustainable wheat production in the Baltic Sea region. *Physiol. Plantarum* 164 (4), 442–451. doi: 10.1111/ppl.12726
- Colmer, J., O'Neill, C. M., Wells, R., Bostrom, A., Reynolds, D., Websdale, D., et al. (2020). SeedGerm: a cost-effective phenotyping platform for automated seed imaging and machine-learning based phenotypic analysis of crop seed germination. *New Phytol.* 228 (2), 778–793. doi: 10.1111/nph.16736
- Del Ponte, E. M., Moreira, G. M., Ward, T. J., O'Donnell, K., Nicolli, C. P., Machado, F. J., et al. (2022). Fusarium graminearum species complex: A bibliographic analysis and web-accessible database for global mapping of species and trichothecene toxin chemotypes. *Phytopathology* 112 (4), 741–751. doi: 10.1094/PHYTO-06-21-0277-RVW
- Delwiche, S. R., Kim, M. S., and Dong, Y. (2010). "Damage and quality assessment in wheat by NIR hyperspectral imaging," in *Sensing for agriculture and food quality and safety II: SPIE*, Vol.7676, 45–52.
- De Mendiburu, F. (2014). *Agricolae: statistical procedures for agricultural research. R. Package Version 1* (1), 1–4. doi: 10.7287/peerj.preprints.1404v1
- Dissing, B. S., Papadopoulou, O. S., Tassou, C., Ersbøll, B. K., Carstensen, J. M., Panagou, E. Z., et al. (2013). Using multispectral imaging for spoilage detection of pork meat. *Food Bioprocess Technol.* 6 (9), 2268–2279. doi: 10.1007/s11947-012-0886-6
- Femenias, A., Llorens-Serentill, E., Ramos, A. J., Sanchis, V., and Marín, S. (2022). Near-infrared hyperspectral imaging evaluation of fusarium damage and DON in single wheat kernels. *Food Control* 142, 109239. doi: 10.1016/j.foodcont.2022.109239
- Gilbert, J., and Haber, S. (2013). Overview of some recent research developments in fusarium head blight of wheat. *Can. J. Plant Pathol.* 35 (2), 149–174. doi: 10.1080/07060661.2013.772921
- Gomes, F. G. Jr, and Duijn, B. V. (2017). Three-dimensional (3-d) X-ray imaging for seed analysis. *Seed Testing Int.* 154, 48–52. Available at: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85084478678&partnerID=40&md5=f95af32d6e967fe8630289cf7b49a4f0>
- Góral, T., Wiśniewska, H., Ochodźki, P., Nielsen, L. K., Walentyn-Góral, D., and Stepień, Ł. (2018). Relationship between fusarium head blight, kernel damage, concentration of fusarium biomass, and fusarium toxins in grain of winter wheat inoculated with fusarium culmorum. *Toxins* 11 (1), 2. doi: 10.3390/toxins11010002

Funding

This study was supported by funding from the SLU Grogrund (SLU.ltv.2019.1.1.1-623), Nordic Council of Ministers (PPP #6P3), and NordForsk (#84597). Formas (#2020-01828).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Istvan, D. (2014). "DigiCamControl software". 2.1.2. ed.
- Iwata, H., Ebana, K., Uga, Y., Hayashi, T., and Jannink, J.-L. (2010). Genome-wide association study of grain shape variation among *oryza sativa* l. germplasm based on elliptic Fourier analysis. *Mol. Breed.* 25 (2), 203–215. doi: 10.1007/s11032-009-9319-2
- Iwata, H., and Ukai, Y. (2002). SHAPE: a computer program package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors. *J. Heredity* 93 (5), 384–385. doi: 10.1093/jhered/93.5.384
- Jirsa, O., and Polišenská, I. (2011). Identification of fusarium damaged wheat kernels using image analysis. *Acta Universitatis Agriculturae Silviculturae Mendelianae Brunensis* 59 (5), 125–130. doi: 10.11118/actaun201159050125
- Josse, J., and Husson, F. (2016). missMDA: a package for handling missing values in multivariate data analysis. *J. Stat. Software* 70, 1–31. doi: 10.18637/jss.v070.i01
- Khacim, H. M., Clark, A., Pearson, T., and Van Sanford, D. (2019). Methods of assessing fusarium damage to wheat kernels. *Al-Qadisiyah J. For Agric. Sci. (QJAS) (P-ISSN: 2077-5822 E-ISSN: 2617-1479)* 9 (2), 297–308. doi: 10.33794/qjas.Vol9.Iss2.91
- Komyshchev, E., Genaev, M., and Afonnikov, D. (2017). Evaluation of the SeedCounter, a mobile application for grain phenotyping. *Front. Plant Sci.* 7, 1990. doi: 10.3389/fpls.2016.01990
- Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., et al. (2020). Package 'caret'. *R. J.* 223, 7. doi: 10.18637/jss.v028.i05
- Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *J. Stat. Software* 25, 1–18. doi: 10.18637/jss.v025.i01
- Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., et al. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28 (18), 2397–2399. doi: 10.1093/bioinformatics/bts444
- Liu, W., Liu, C., Jin, J., Li, D., Fu, Y., and Yuan, X. (2020). High-throughput phenotyping of morphological seed and fruit characteristics using X-ray computed tomography. *Front. Plant Sci.* 11, 601475. doi: 10.3389/fpls.2020.601475
- Lobet, G. (2017). Image analysis in plant sciences: publish then perish. *Trends Plant Sci.* 22 (7), 559–566. doi: 10.1016/j.tplants.2017.05.002
- Lu, B., Dao, P. D., Liu, J., He, Y., and Shang, J. (2020). Recent advances of hyperspectral imaging technology and applications in agriculture. *Remote Sens.* 12 (16), 26–59. doi: 10.3390/rs12162659
- Menesatti, P., Antonucci, F., Costa, C., Santori, A., Niciarelli, I., and Infantino, A. (2009). Application of morphometric image analysis system to evaluate the incidence of fusarium head blight wheat infected kernels. *Proc. Bornimer Agrartechnische Berichte* 69, 157–161. Available at: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84876683478&partnerID=40&md5=44c21b369e2d5fcd1b165442fd8dbb7d>
- Mesterhazy, A. (2020). Updating the breeding philosophy of wheat to fusarium head blight (FHB): Resistance components, QTL identification, and phenotyping—a review. *Plants* 9 (12), 1702. doi: 10.3390/plants9121702
- Polak-Śliwińska, M., and Paszczyk, B. (2021). Trichothecenes in food and feed, relevance to human and animal health and methods of detection: A systematic review. *Molecules* 26 (2), 454. doi: 10.3390/molecules26020454
- Rangarajan, A. K., Whetton, R. L., and Mouazen, A. M. (2022). Detection of fusarium head blight in wheat using hyperspectral data and deep learning. *Expert Syst. Appl.* 208, 118240. doi: 10.1016/j.eswa.2022.118240
- Saccon, F. A., Parcey, D., Paliwal, J., and Sherif, S. S. (2017). Assessment of fusarium and deoxynivalenol using optical methods. *Food Bioprocess Technol.* 10 (1), 34–50. doi: 10.1007/s11947-016-1788-9
- Schroeder, H., and Christensen, J. (1963). Factors affecting resistance of wheat to scab caused by *Gibberella zeae*. *Phytopathology* 53 (1), 831–838.
- Shahin, M. A., and Symons, S. J. (2011). Detection of fusarium damaged kernels in Canada Western red spring wheat using visible/near-infrared hyperspectral imaging and principal component analysis. *Comput. Electron. Agric.* 75 (1), 107–112. doi: 10.1016/j.compag.2010.10.004
- Shiferaw, B., Smale, M., Braun, H.-J., Duveiller, E., Reynolds, M., and Muricho, G. (2013). Crops that feed the world 10. past successes and future challenges to the role played by wheat in global food security. *Food Secur.* 5 (3), 291–317. doi: 10.1007/s12571-013-0263-y
- Stack, R. W., and McMullen, M. P. (1998). A visual scale to estimate severity of fusarium head blight in wheat. NDSu Extension Circular: n. pag
- Steed, A., King, J., Grewal, S., Yang, C.-Y., Clarke, M., Devi, U., et al. (2022). Identification of fusarium head blight resistance in triticum timopheevii accessions and characterization of wheat-t. timopheevii introgression lines for enhanced resistance. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.943211
- Tanabata, T., Shibuya, T., Hori, K., Ebana, K., and Yano, M. (2012). SmartGrain: High-throughput phenotyping software for measuring seed shape through image analysis. *Plant Physiol.* 160 (4), 1871–1880. doi: 10.1104/pp.112.205120
- Team, R. C. (2013). *R: A language and environment for statistical computing*.
- Wang, L., Uilecan, I. V., Assadi, A. H., Kozmik, C. A., and Spalding, E. P. (2009). HYPOTrace: image analysis software for measuring hypocotyl growth and shape demonstrated on arabidopsis seedlings undergoing photomorphogenesis. *Plant Physiol.* 149 (4), 1632–1637. doi: 10.1104/pp.108.134072
- Whan, A. P., Smith, A. B., Cavanagh, C. R., Ral, J.-P. F., Shaw, L. M., Howitt, C. A., et al. (2014). GrainScan: a low cost, fast method for grain size and colour measurements. *Plant Methods* 10 (1), 1–10. doi: 10.1186/1746-4811-10-23
- Yipeng, L., Wenbing, L., Kaixuan, H., Wentao, T., Ling, Z., Shizhuang, W., et al. (2022). Determination of wheat kernels damaged by fusarium head blight using monochromatic images of effective wavelengths from hyperspectral imaging coupled with an architecture self-search deep network. *Food Control* 135, 108819. doi: 10.1016/j.foodcont.2022.108819
- Zakieh, M., Gaikpa, D. S., Leiva Sandoval, F., Alamrani, M., Henriksson, T., Odilbekov, F., et al. (2021). Characterizing winter wheat germplasm for fusarium head blight resistance under accelerated growth conditions. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.705006
- Zhu, F., Paul, P., Hussain, W., Wallman, K., Dhatt, B. K., Sandhu, J., et al. (2021). SeedExtractor: an open-source GUI for seed image analysis. *Front. Plant Sci.* 11, 581546. doi: 10.3389/fpls.2020.581546

Appendix 1

Quantitative trait loci (QTL) detected in genome-wide association studies employing seven models at $p = 0.0001$ ($\text{LOD} \geq 4$) for Fusarium head blight severity in winter wheat from the breeding, genebank, and combined sets (Zakieh et al., 2021). Chr., chromosome; FAF, favorable allele frequencies. The asterisk means also detected by these models at $p = 0.0002$. A, detected above Bonferroni corrected threshold ($\alpha = 0.05$). B, the marker effects are estimated for only GLM, MLM, and CMLM and FarmCPU in GAPIT (Lipka et al., 2012).

QTL	Marker	Chr.	Position (cM)	FAF	Effect	Model (s)	Set
SLUfhhbchr1B.1	BS00021877_51	1B	154.58	0.06	NA	Blink	Combined
SLUfhhbchr2A.2	BobWhite_c16923_64	2A	125.33	0.06	NA	Blink; (SUPER)*	Combined
SLUfhhbchr3A.3	Kukri_rep_c89183_282	3A	15.05	0.64	27.84 to 28.10	GLM, CMLM	Combined
SLUfhhbchr3B.4	wsnp_Ex_c34975_43204180	3B	67.45	0.95 (CS), 0.94 (BS), 0.97 (GS)	65.78 to 82.47	GLM, MLM, CMLM, SUPER, MLMM, FarmCPU, Blink	All
	Kukri_c18009_398a	3B	67.67	0.95	78.20 to 80.15	GLM, MLM, CMLM, SUPER	Combined
	wsnp_Ex_c5378_9505533	3B	68.71	0.94	NA	SUPER	Combined
SLUfhhbchr3D.5a	RFL_Contig4591_1759	3D	0.00	0.94	51.94 to 54.69*	MLMM; (GLM, MLM, CLM, SUPER, Blink)*	Combined
	RAC875_rep_c115090_5	3D	0.00	0.02	NA	Blink	Breeding
SLUfhhbchr3D.5b	JD_c7714_954	3D	143.01	0.04	NA	Blink, SUPER	Genebank
SLUfhhbchr5A.6	RAC875_rep_c106118_339	5A	39.02	0.03	-31.55 to -29.40	GLM, MLM, SUPER, MLMM	Combined
SLUfhhbchr6A.7	Tdurum_contig46670_911	6A	128.26	0.96	NA	SUPER	Combined
SLUfhhbchr7A.8	Kukri_c11530_92	7A	232.11	0.84	44.1	CMLM, SUPER, MLMM	Combined
	RAC875_c12733_1509a	7A	228.37	0.83	40.41 to 45.14	GLM, MLM, CMLM, SUPER, MLMM, FarmCPU, Blink	Combined
SLUfhhbchr7B.9	wsnp_Ex_c351_689415	7B	143.23	0.02	NA	Blink, SUPER	Breeding
	RAC875_c8752_1079	7B	158.98	0.84	39.97*	SUPER; (CMLM)*	Combined

Frontiers in Plant Science

Cultivates the science of plant biology and its applications

The most cited plant science journal, which advances our understanding of plant biology for sustainable food security, functional ecosystems and human health.

Discover the latest Research Topics

[See more →](#)

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact

