Received: 21 September 2020 | Accepted: 27 January 2021

DOI: 10.1002/ppj2.20015

The Plant Phenome Journal OPEN ACCESS

ORIGINAL RESEARCH

PhenoImage: An open-source graphical user interface for plant image analysis

Feiyu Zhu[1] | **Manny Saluja**[2] | **Jaspinder Singh Dharni**[2] | **Puneet Paul**[2] | **Scott E. Sattler**[2,3] | **Paul Staswick**[2] | **Harkamal Walia**[2] | **Hongfeng Yu**[1]

[1] Dep. of Computer Science and Engineering, Univ. of Nebraska, Lincoln, NE 68588, USA

[2] Dep. of Agronomy and Horticulture, Univ. of Nebraska, Lincoln, NE 68583, USA

[3] Wheat, Sorghum and Forage Research Unit, USDA-ARS, Lincoln, NE 68583, USA

Correspondence
Hongfeng Yu, Dep. of Computer Science and Engineering, Univ. of Nebraska, Lincoln, NE 68588, USA.
Email: hfyu@unl.edu

Assigned to Technical Editor Carolyn Lawrence-Dill.

Funding information
National Science Foundation, USA, Grant/Award Number: 1736192

Abstract

High-throughput genotyping coupled with molecular breeding approaches have dramatically accelerated crop improvement programs. More recently, improved plant phenotyping methods have led to a shift from manual measurements to automated platforms with increased scalability and resolution. Considerable effort has also gone into developing large-scale downstream processing of the imaging datasets derived from high-throughput phenotyping (HTP) platforms. However, most available tools require some programming skills. We developed *PhenoImage*, an open-source graphical user interface (GUI) based cross-platform solution for HTP image processing intending to make image analysis accessible to users with either little or no programming skills. The open-source nature provides the possibility to extend its usability to meet user-specific requirements. The availability of multiple functions and filtering parameters provides flexibility to analyze images from a wide variety of plant species and platforms. *PhenoImage* can be run on a personal computer as well as on high-performance computing clusters. To test the efficacy of the application, we analyzed the LemnaTec Imaging system derived red, green, and blue (RGB) color intensity and plant pigmentation-based fluorescence shoot images from two plant species: sorghum [*Sorghum bicolor* (L.) Moench] and wheat (*Triticum aestivum* L.) differing in their physical attributes. In the study, we discuss the development, implementation, and working of the *PhenoImage*.

1 | INTRODUCTION

In the genomics and post-genomics era, technological advances in sequencing platforms have paved the way for high throughput genotyping (Furbank & Tester, 2011; Jackson et al., 2011). These developments coupled with molecular breeding approaches have enhanced the genetic understanding of plants, which has dramatically progressed the crop-improvement efforts (Moose & Mumm, 2008; Tester & Langridge, 2010; Varshney et al., 2009). However, precise and efficient phenotyping has been a challenge (Furbank & Tester, 2011). To tackle this problem, plant phenotyping technologies have achieved a huge leap in recent times—the shift from laborious and error-prone manual

Abbreviations: CPU, central processing unit; GI, greenness index; GUI, graphical user interface; HCA, hierarchical clustering analysis; HTP, high-throughput phenotyping; HPC, high-performance computing; HSV, hue, saturation, and value; IDs, identifications; PSA, projected shoot area; RGB, red, green, and blue; ROI, region of interest; WHC, water holding capacity; WL, water-limited treatment; WW, well-watered treatment.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2021 The Authors. *The Plant Phenome Journal* published by Wiley Periodicals LLC on behalf of American Society of Agronomy and Crop Science Society of America

measurements toward automation (Fiorani & Schurr, 2013; Gong & He, 2014). Automated imaging-based platforms have tremendously enhanced our ability to record a plant's physical and physiological attributes in a non-invasive manner. Despite these advances, phenotyping technologies still trail developments on the genomics front, especially the rate at which the phenotypic data is generated (Furbank & Tester, 2011; Gehan et al., 2017; Houle et al., 2010; Minervini et al., 2015; Sandhu et al., 2019). The major limit is not the ever-evolving sophisticated instrumentation for image capturing but with the downstream processing of large-scale phenotypic data, which is not easily accessible to many plant biologists.

High-throughput phenotyping (HTP) imaging platform refers to the accurate acquisition and analysis of multidimensional traits at the individual plant level in the context of this work (Yang et al., 2020). It is no surprise, these platforms generate a diversity of images corresponding to different light spectra such as RGB, near-infrared, plant pigmentation-based fluorescence, and hyperspectral. Thus, terabytes of digital information can be routinely generated through an imaging experiment. Currently, the website www.plant-image-analysis.org documents 179 image software tools (Lobet et al., 2013). Availability and usage of some of these tools are being restricted and adhere to proprietary rights, for instance, LemnaGrid-Scanalyzer3D by LemnaTec GmbH, Germany. On the other hand, several open-source tools designed for specific applications, ranging from cell to whole canopy analysis, are readily accessible in the public domain (Lobet et al., 2013). In addition to their broader functionalities, these have also opened new avenues to integrate third-party algorithms. Examples include HTPheno (developed as a plugin for ImageJ) (Hartmann et al., 2011), Plant Computer Vision or Plant CV (a community-based toolkit for plant phenotyping analysis) (Fahlgren et al., 2015; Gehan et al., 2017), Integrated Image Platform or IAP (Klukas et al., 2014), and Image Harvest (Knecht et al., 2016). Despite their power and flexibility, these tools may require some proficiency with programing language as a pre-requisite to process large-scale datasets. This is a challenge for many biologists with limited or no coding skills.

The availability of several affordable automated and semi-automated phenotyping platforms has increased their usage to score the traits of interest (Klukas et al., 2014; Li et al., 2014). Keeping this view in mind, we developed *PhenoImage*—an open-source, graphical user interface (GUI)-based cross-platform solution for large-scale data processing—which is not only convenient to use but highly precise and effective at the same time. The intuitive nature of the application will allow plant scientists with little or no knowledge of programming language to process phenotypic dataset on their personal computers. In addition, the application can facilitate parallel processing of large-scale image data on high-performance computing clusters. To test the efficacy of the application, we analyzed the LemnaTec Imaging system-derived RGB and plant pigmentation-based fluorescence images from sorghum [*Sorghum bicolor* (L.) Moench] and wheat (*Triticum aestivum* L.), which differ in their physical attributes. The availability of multiple functions and filtering parameters provides flexibility to analyze a wide variety of plant species. Images acquired from other phenotyping platforms or handheld devices can also be processed using *PhenoImage*.

Core Ideas

- *PhenoImage* is an application for analyzing images derived from high-throughput phenotyping.
- Using the tool, users can access image analysis with little or no programming skills.
- The open-source nature provides the possibility to further extend its usability.

2 | MATERIALS AND METHODS

2.1 | *PhenoImage* workflow

PhenoImage is a MATLAB-based application, i.e., compatible with multiple operating systems. The software is available in two versions, a regular version that requires MATLAB license and a standalone version that uses "MATLAB Compiler Runtime" and does not necessarily require a MATLAB license for its operation. Both versions can be downloaded from http://wrchr.org/phenolib/phenoimage. The same GUI application can support image data processing on a single central processing unit (CPU) and parallel processing on high-performance computing (HPC) clusters.

2.2 | Software development and implementation

The GUI for high throughput image analysis is based on MATLAB. The summary of image processing workflow of *PhenoImage* includes (a) file loading, (b and c) image cropping and filtering, (d) digital trait extraction using specific functions (based on the user's requirement), and (e) followed by image processing either on a local machine or HPC clusters (Figure 1). We have provided a step-by-step guide to use *PhenoImage* (see the *PhenoImage* Guide Document Supplemental file).

25782703, 2021, 1, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/ppj2.20015 by Test, Wiley Online Library on [11/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

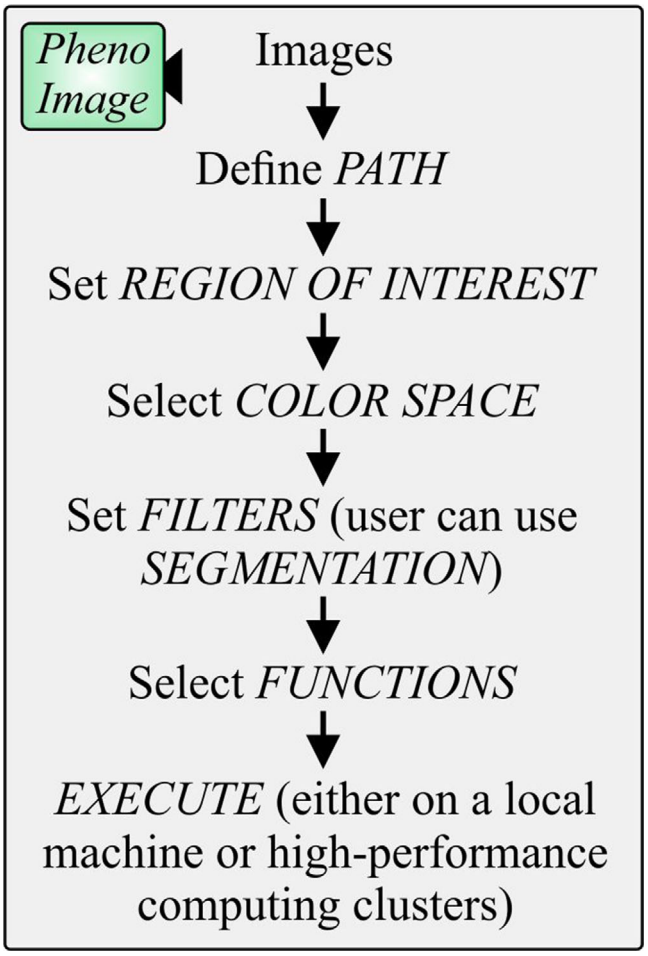

FIGURE 1 *PhenoImage* workflow. First, the path of the folder containing red, green blue (RGB) images is described. Then, a region of interest containing plant image is defined, followed by the selection of color space of preference. Afterward, filter parameters are determined, and functions corresponding to the digital traits of users' choice are selected. Then the processing of the images is executed using either a local central processing unit (CPU) or high-performance computing clusters

2.3 | File loading

The image files can be loaded by specifying regular expressions in *Path* using the following format: "FOLDER NAME*.png." This should allow the loading of all images under the respective folder. The application is compatible with widely used image formats such as jpg, png, and tiff. The spinner can be used to change the *Original Image* that is currently displayed (Figure 2).

The visible images (RGB) of plants can be obtained using any system such as LemmaTec or using standard digital cameras. If analyzing images using standard digital cameras, the user must ensure constant focal distance to have a similar scale for all the images corresponding to the same batch to facilitate precise comparison.

2.4 | Selection of ROI and image filtering

For selecting region of interest (ROI), the user can either crop the image interactively by dragging a marquee tool over the image or by typing the position of the ROI using this format, [X_min, Y_min, Width, Height], where X_min, Y_min is the coordinate of the upper-left corner, and Width and Height correspond to the size of the ROI. The ROI selected is fixed for the image analysis of the respective folder. Thus, it is recommended that the user selects a relatively larger ROI. This is important, especially during the analysis of plant growth dynamics in a temporal manner where plants tend to increase in size.

Next, image segmentation separates plant pixels from the background. For segmentation, a logical expression can be specified in *Filter* (Figure 2). The application supports (a) red, green, and blue (RGB), (b) hue, saturation, and value (HSV), and (c) Lab color spaces, which provide flexibility to the user to optimally segment the RGB images. Any combination of arithmetic and logical operation can be used to segment the plant. In terms of setting the filter, "&" means logical *AND*, and "|" means logical *OR*. For example, "r < 200 & g < 150" means finding pixels with red values less than 200 and green values less than 150. The *Processed Image* will be displayed after clicking the "Test" button (Figure 2).

If the user is unsure about predefining the filter, then the Segmentation feature can be utilized (Zhu et al., 2021). For this, click on "Foreground" and a new pop-up window having the original image appears. The user can select the zoom-in option from the task bar menu to enlarge the area of interest (i.e., plant tissue in this case). Once the area is zoomed in, the user can deselect the zoom-in option and scribble on the enlarged area of interest with a red mark (Figure 3). Next, background (i.e., pot, pot stand, plant background, etc.) is selected by clicking the "Background" button and scribble on the background using a green mark. Afterward, the user can click the "Segment" button to initiate the segmentation or subtraction of plant pixels from the background (Figure 3). We empirically segment the plant by finding plant pixels where the difference to the mean of selected foreground is less than 60. Implementation of the Segment function may take a few additional seconds. After segmentation, the *Processed Image* will show pixels corresponding only to the plant and the histograms corresponding only to the plant region will be displayed in *Channel 1*, *2*, *and 3*. The range of the histogram for each channel can be used to define Filter parameters. The Segmentation feature is helpful to define filter parameters in a similar manner for both RGB and plant pigmentation-based fluorescence images; however, histogram values for only the red channel need to be considered for setting the filter in the case of plant pigmentation-based fluorescence images.

25782703, 2021, 1, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/ppj2.20015 by Test, Wiley Online Library on [11/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

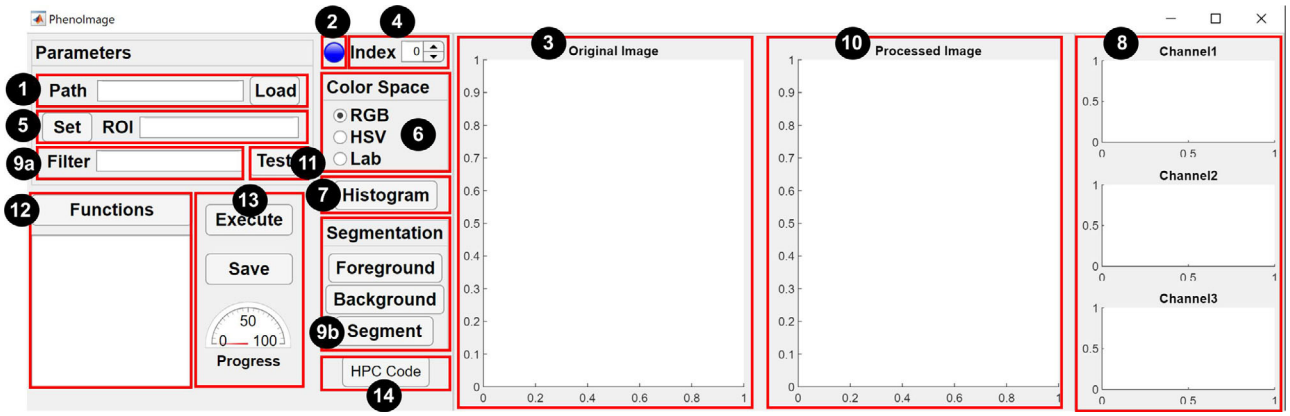

FIGURE 2 Graphical user interface (GUI) of *PhenoImage*. The numbers denote a step-by-step guide to use the application: (1) define *Path* of the folder containing plant images and click the "Load" button; (2) The light bulb shows status of the loading procedure, as the blue light bulb turns red while the loading is in progress and green when completed; (3) one of the images from the folder is displayed in the *Original Image* section; (4) the spinner can be used to change the current image shown in the *Original Image* space; (5) then the user must define *Region of interest* or *ROI* by dragging the cursor on the *Original Image*; (6) select the *Color space* of preference; (7 and 8) click on the "Histogram" button to visualize intensity of channels corresponding to the respective *Color space*; (9a and 9b) the user can either directly use the histogram values to define the Filter or the user can use the Segmentation function, where *Foreground* and *Background* need to be specified to precisely segment plant pixels from the background; (10 and 11) the user can click "Test" to view the *Processed Image* and if the user has decided on the Filter; (12) then selection of Functions is performed. The functions or the digital parameters that need to be extracted are based on user preference. (13) If running on a single machine, the user can *Execute* the function to process all images in the respective folder, and progress of batch processing can be viewed in the *Progress bar* and saved; or (14) high-performance computing clusters can be used to process the images

2.5 | Defining functions for plant trait analysis

For digital trait extraction, the user can select functions from a dialog window by clicking "Functions," where any user-defined functions can be selected. The selected functions will be listed in the text region and will take the segmented image as input to extract digital traits (Table 1). Some of the commonly used functions are defined below.

2.5.1 | Pixel count

After segmentation, only the pixels corresponding to the plant are kept, whereas pixels corresponding to other objects in the image are set to black. The tool counts the number of pixels that belong to the plant in the ROI.

2.5.2 | Pixel intensity

Pixel Intensity refers to the sum of the intensities of pixels in an image. As there are three channels (red, green, and blue), we calculate the pixel sum of each of the R, G, and B channels separately.

2.5.3 | Dimension

To define the dimensions (width and height), first a bounding box, which is based on the segmented pixels and encloses all pixels of the plant, is found (Figure 4). Then, the width and height of the bounding box are used to define the dimensions of the plant.

2.5.4 | Convex area

Convex Area is a feature that is related to the shape of the plant. The convex area is the area of the convex hull. The convex hull is the smallest convex polygon enclosing all the pixels of the plant (Figure 4).

2.5.5 | Image skeleton

We find the skeleton of the plant pixels using a skeletonization algorithm (Abeysinghe et al., 2008). The skeleton of the plant approximates the center lines of the stem and the leaves. Then, the number of pixels in the skeleton is obtained (Figure 4).

25782703, 2021, 1, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/ppj2.20015 by Test, Wiley Online Library on [11/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

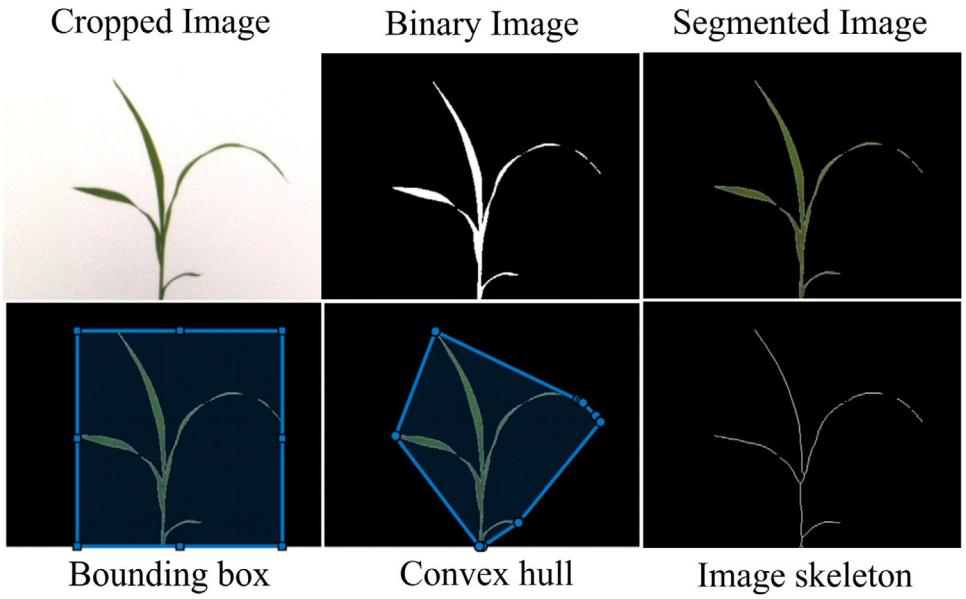

FIGURE 3 Representation of different features used by *PhenoImage*. The cropped image is derived from the original image after the selection of the region of interest. The binary image is a mask of the plant pixels where the plant pixels are set to 1 and the background is set to 0. The segmented image represents the segmented plant pixels from the background. The bounding box shown in the light blue color is calculated based on the segmented pixels and encloses all pixels of the plant. Convex hull signifies the smallest convex polygon enclosing all the pixels of the plant. The image skeleton approximates the center lines of the stem and the leaves and is calculated using a skeletonization algorithm

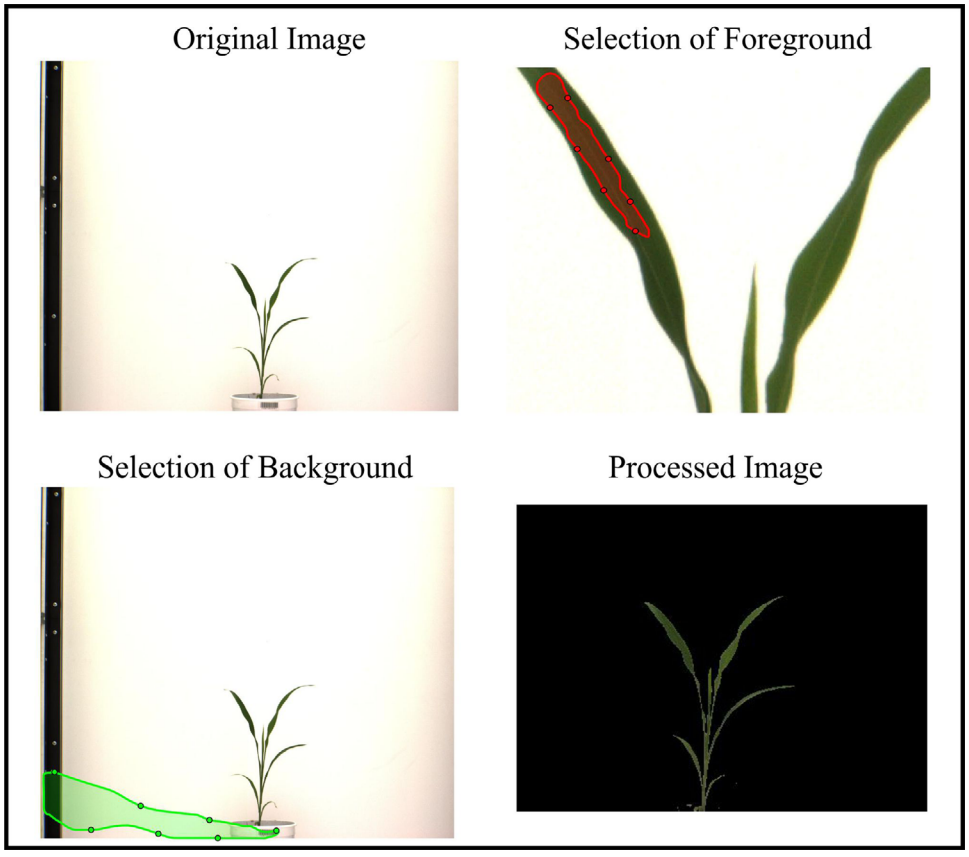

FIGURE 4 Segmentation. For subtracting plant pixels from the background, click on "Foreground" and a new pop-up window displaying the original image opens. The user can scribble on the area of interest with a red mark. Likewise, the user can define background (i.e., pot, pot stand, plant background, etc.) by scribbling on the background using a green mark. Afterward, the user can click the "Segment" button to initiate the segmentation of plant pixels from the background. As a result, the *Processed Image* will show pixels corresponding only to the plant

25782703, 2021, 1, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/ppj2.20015 by Test, Wiley Online Library on [11/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 List of the digital parameters that *PhenoImage* can extract

Function	Definition
Binary Central Moment	Central moment of the binary image of the plant
Central Moment	Central moment of the plant image
Convex Area	Area of the convex of plant pixels
Pixel Count	Number of plant pixels
Plant Height	Height of the bounding box of the plant pixels
Gravity Height	Height of the gravitational center
Histogram-Blue	Histogram of the blue channel
Histogram-Green	Histogram of the green channel
Histogram-Red	Histogram of the red channel
Pixel intensity mean-Blue	Mean of the intensities of the blue channel
Pixel intensity mean-Green	Mean of the intensities of the green channel
Pixel intensity mean-Red	Mean of the intensities of the red channel
Image Skeleton	Number of pixels of the skeleton of the plant
Skeleton Central Moment	Central moment of the skeleton of the plant
Pixel intensity sum-Blue	Sum of the intensities of the blue channel
Pixel intensity sum-Green	Sum of the intensities of the green channel
Pixel intensity sum-Red	Sum of the intensities of the red channel
Pixel intensity variance-Blue	Variance of the blue channel
Pixel intensity variance-Green	Variance of the green channel
Pixel intensity variance-Red	Variance of the red channel
Plant Width	Width of the bounding box of the plant pixels

2.5.6 | Image moment

Image moments can be used to evaluate the shape of the plant (Hu, 1962). We evaluate the image moment of the binary image or the segmented plant image. For the binary image, the plant pixel is considered as 1, and pixels of other objects in the image (e.g., pot, pot-holder, and background) are considered as 0. The moment of the binary image is only dependent on the positions of the pixels. The fourth-order central image moment (μ_{22}) is evaluated for the binary image in default. The user can easily modify the function to obtain image moments of other orders.

2.6 | Execution on a local machine

2.6.1 | Image processing and results collection

The batch image processing can be initiated by clicking "Execute" (Figure 2). The light bulb located on the right side indicates the status of processing. For instance, the light bulb turns red as images are being processed and will turn green upon its completion. The Progress gauge will show the progress of the image processing on a percentage basis. A text file containing results from all the functions can be generated by clicking "Save." The user can specify the path and file name of the text file in a pop-up dialog window.

2.7 | Execution on HPC clusters

2.7.1 | Execution and code generation on HPC clusters

To accelerate the processing of a large number of images, one feasible way is to use a HPC cluster by distributing the workload among the cluster computing nodes and processing images in a parallel fashion. We design *PhenoImage* to be also executed on a HPC cluster and facilitate users to leverage the power of HPC without detailed HPC knowledge.

For executing jobs on HPC clusters, *slurm* (Yoo et al., 2003) is used to submit a batch job, which distributes the jobs using job identifications (IDs). If a job requires a small number of resources, the priority of execution of the job using *slurm* is comparably high. Thus, owing to less queuing time, we chose *slurm* for *PhenoImage*.

Further, to process images using HPC clusters, click the "Code" button (Figure 2), which generates a MATLAB script for processing images. A *slurm* file (an example is included in *PhenoImage*) is used to submit a job array to the cluster. Then, the job IDs and job size executed by *slurm* are used to partition the images so that each node processes only a part of the images. The job IDs and job size are passed to the MATLAB script generated by *PhenoImage* as input parameters. The user only needs to input the names of the files that need to be processed in a text file. For each node in the HPC cluster, the script reads all the filenames and processes a part of the images as specified by job ID and job size. Specifically, the script will process images with indices (JobID, JobID+JobSize, JobID+2* JobSize, …). The script

25782703, 2021, 1, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/ppj2.20015 by Test, Wiley Online Library on [11/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

contains the position of the ROI, the expression of the color filter, and the names of the functions that have been selected for digital trait extraction.

After submitting of the job using the *slurm* file, the result computed by each function on each node will be printed out and finally aggregated in one output file specified in the *slurm* file. The output file contains all the results for all the input images.

2.8 | Sorghum and wheat: A test case for *PhenoImage* validation

2.8.1 | Sorghum

Four seeds of sorghum genotype RTX430 were sown in each of the 10 5.6-L pots (22 cm diameter by 19.5 cm height) filled with 2.5 kg of a soil mix (consisting of two-thirds peat moss and one-third vermiculite and 1.4 kg lime). Plants were thinned to one seedling per pot 6 d after germination. For the first 21 d, all pots were watered to 85% water holding capacity (WHC). Afterward, water was withheld from half of the pots (water-limited treatment, WL) until 30% WHC is attained, and half of the pots were maintained at 80% WHC (well-watered treatment, WW; Supplemental Figure S1. During the entire experiment, the greenhouse was maintained at 28 ˚C for 13 h day, and 25 ˚C for 11 h night, , and 40–50% relative humidity (Saluja et al., 2020).

2.8.2 | Wheat

Seeds of wheat genotype Pavon were germinated in Petri dishes for 4 d in the dark at 25 ˚C. Uniformly germinated seeds were transplanted into 3-L pots (12 cm diameter by 19.5 cm height) and filled with 1.2 kg of Fafard germination soil (Sungro) supplemented with Osmocote fertilizer and Micromax micronutrients. Seedlings were grown for 7 d at 80% WHC. After 7 d, six seedlings each were maintained at 80% WHC for well-watered treatment and 30% WHC for water-limited treatment (Supplemental Figure S1). Growth conditions were maintained at 22 ˚C for 16 h day, and 16 ˚C for 8 h night. Afterward, plants were imaged every day for 15 d.

2.8.3 | Water holding capacity

For calculating WHC of the soil mix, 2.5 and 1.2 kg of soil mix for sorghum and wheat experiment, respectively, was oven-dried (60 ˚C for 7 d) and dry soil weight was measured. Then the soil mix was transferred to pots perforated at the bottom for drainage. To achieve the saturation point (weight of the soil at 100% WHC), the soil mix was saturated with water while covered at the top to prevent evaporation. Pots were weighed daily until no change in pot weight was observed. These computed values were then used to calculate the weight of soil at a particular WHC by using the following equation:

$$\text{Soil weight at a particular WHC} = \left[(\text{Soil weight at 100\% WHC} - \text{Dry soil weight}) \times \text{Required WHC}\right] + \text{Dry soil weight}$$

2.8.4 | Plant imaging

A high-throughput phenotyping facility (LemnaTec Imaging System) at Nebraska Innovation Campus, the University of Nebraska-Lincoln was used to evaluate sorghum (RTx430) and wheat (Pavon) plants by RGB and plant pigmentation-based fluorescence images. For sorghum plants, starting from the day water was withheld, both WW and WL pots were imaged every day until WL pots reached 30% WHC. Plants were imaged for 18 d (Supplemental Figure S1). Due to technical error during the experiment, the imaging system failed to acquire images on the Day 13 of imaging, so data corresponding to this day is missing from the downstream analysis. For wheat plants, imaging was performed for 15 d for WW and WL conditions (Supplemental Figure S1).

To reduce image occlusions, imaging was done from five different angles (side views at 0˚, 72˚, 144˚, 216˚, and 288˚; Supplemental Figure S2; Golzarian et al., 2011). Next, RGB and plant pigmentation-based fluorescence images from both the species were used as a test cases to validate *PhenoImage*. For validation and optimal segmentation of RGB images, the following filter parameters were used: $g < 150$ (for sorghum) and $g < 150$ & $h > 0.2$ & $h < 0.5$ & $s > 0.1$ & $v < 0.6$ & $a < -5$ (for wheat).

For plant pigmentation-based fluorescence images, plants were imaged at constant blue light (400–500 nm) and steady-state chlorophyll fluorescence was detected at 500–700 nm in a separate chamber. An ad-hoc image segmentation strategy was used to categorize image color ranges into 32 color classes (Campbell et al., 2015). Further, Hierarchical Cluster Analysis (HCA) was performed using Ward's method (JMP Pro13) to examine the temporal profile of the color classes with pixel intensities. For plant pigmentation-based fluorescence images, filters were defined using only the red pixels; sorghum $r > 150$, and wheat $r > 50$ & $r < 140$.

25782703, 2021, 1, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/ppj2.20015 by Test, Wiley Online Library on [11/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.9 | Manual phenotyping and comparisons with other methods

Manual measurements were performed for both sorghum and wheat plants on the last day of imaging in a destructive manner. For this, fresh and dry shoots were weighed. Shoots were dried for 1 wk in an oven at 60 ˚C and weighed to determine the dry weight. The manually derived traits were correlated with digital traits derived from last day of imaging for sorghum and wheat (Day 18 and 15, respectively; Supplemental Figure S1). The RGB images from the last day of imaging were processed using *PhenoImage* (Supplemental Table S2), HTPheno (Hartmann et al., 2011), and OpenCV (Bradski, 2000) (Supplemental Table S3). For correlation, pixel count derived from *PhenoImage* and OpenCV, and object area from HTPheno were considered. In addition, sorghum plant images captured by standard smartphone camera (12-megapixel, f/1.8 aperture) and wheat spike images captured by a flatbed scanner (Epson Expression 12,000 XL at 600 dpi resolution) (Paul et al., 2020a) were analyzed using *PhenoImage*.

3 | RESULTS AND DISCUSSION

3.1 | Performance testing

We evaluated the performance of *PhenoImage* (Figure 1 and 2) with respect to the time required to process images. For this, we computed time required to generate data for two functions: convex area and pixel count, derived from RGB images, which had different levels of resolution ranging from 100 × 100 to 10,000 × 10,000 pixels (Supplemental Table S1). We observed that the application's performance at different resolutions depended on the function that is being evaluated (Figure 5). For example, the time taken to analyze convex area at the highest resolution (10,000 × 10,000 pixels; 1.646 s) increased by 53.15% compared with the lowest resolution (100 × 100 pixels; 0.030 s). On the other hand, the time taken to analyze pixel count increased by 16.20% with the increase in the resolution (i.e., 0.167 and 0.009 s for the highest and lowest resolution, respectively) (Figure 5; Supplemental Table S1).

3.2 | Sorghum and wheat: A test case for *PhenoImage* validation

The RGB images from sorghum (RTx430) and wheat (Pavon) were used for validating *PhenoImage* (Figure 1). These species were selected because of their visibly different physical attributes. Sorghum has one main shoot axis, which results in a relatively compact-looking phenotype compared with the wheat plant, which produces multiple tillers (Figure 5). The two species also differ in other parameters such as stalk diameter, leaf width, leaf length, and so forth. Plants from both species were used for HTP with the LemnaTec Imaging System, and images were processed using *PhenoImage*. After loading the images onto the application, the best filter parameters were determined empirically based on a histogram generated after segmenting the foreground (i.e., plant pixels) from the background.

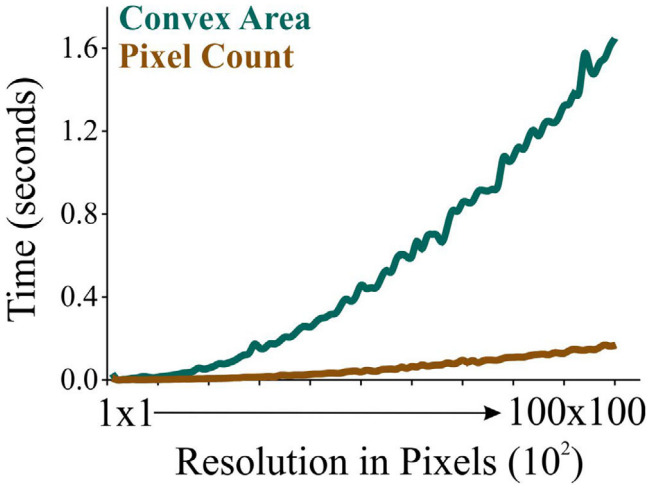

FIGURE 5 Performance testing of *PhenoImage*. The plot shows the time taken to process images and extract the respective digital trait (convex area and pixel count) from the images at different resolution

We assessed two digital traits derived from the RGB images—pixel count and convex area—which are representative of a plant's overall architecture. Pixel count was used as a proxy for projected shoot area (PSA) and represents the total number of pixels of a plant, whereas the convex area was the area of the convex hull and illustrates the smallest convex polygon enclosing all the pixels of a plant (Figure 3). For validation, the visible differences between the two species were assessed for plants of the same age (26-day-old) via imaging. As a result, we detected significant differences ($P < .001$) between sorghum and wheat plants with respect to PSA and the convex area (Figure 6). Interestingly, although wheat plants have a higher number of leaves than sorghum of the same age, sorghum plants had higher PSA and convex area, apparently due to the broader leaves of sorghum.

In addition, we evaluated *PhenoImage's* ability to process images captured using different imaging platforms such as a standard smartphone camera and a flatbed scanner (Supplemental Figure S3). Consequently, *PhenoImage* successfully extracted regions of interest (leaf pixels from sorghum images and spike pixels from wheat spike images), which shows that our application can process the images derived using

25782703, 2021, 1, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/ppj2.20015 by Test, Wiley Online Library on [11/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

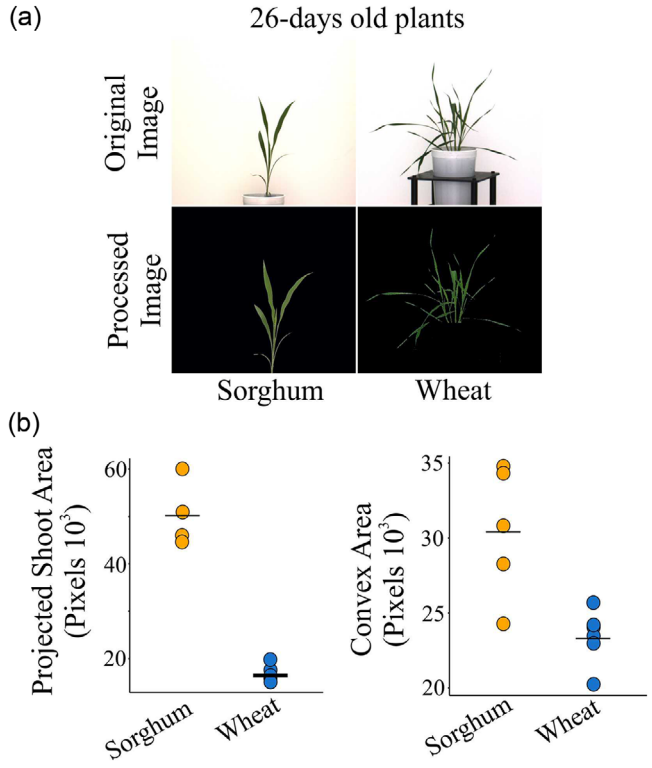

FIGURE 6 *PhenoImage* validation. (a) Original and processed RGB images from 26-day-old sorghum and wheat plants grown under well-watered conditions. (b) Two digital traits: convex area and projected shoot area (PSA), which represent a plant's architecture were derived using *PhenoImage*. Each dot represents individual plant replicate and horizontal line represents the group mean. Significant differences ($P < .001$) were detected between sorghum ($n = 5$) and wheat ($n = 6$) for both the digital traits. For statistics, Welch's t test (equal variance not assumed) was used

multiple platforms (Supplemental Figure S3). Also, these results highlighted *PhenoImage's* ability to extract different plant organs such as leaves and spikes.

3.3 | Comparison with manual measurements and other image processing methodologies

Next, we performed destructive phenotyping of sorghum and wheat plants at the last day of imaging (Day 18 and 15, respectively; Supplemental Figure S1). The harvested plants were used to manually record fresh and dry shoot weight. As expected, we observed significantly higher fresh and dry weight for sorghum compared to wheat (Supplemental Table S2).

Furthermore, we compared the manually recorded phenotypes with digital traits (pixel count) derived from RGB images. For this, the RGB images were processed using the in-house generated application, *PhenoImage*, and two publicly available tools, HTPheno and OpenCV. HTPheno is used as a plugin for ImageJ and does not involve programming language (Hartmann et al., 2011). The application does not allow calibration of color settings for image processing. On the other hand, OpenCV requires skills in Python programming language and does not offer GUI (Bradski, 2000). For both the plant species, we detected a high correlation for fresh and dry weight with PSA derived from RGB images processed using *PhenoImage*, which is at par with the well-established image processing platforms, HTPheno and OpenCV (Supplemental Figure S4, Supplemental Table S2 and S3). This illustrates the sensitivity of the *PhenoImage* application to estimate digital traits is equivalent to other image processing platforms with an added advantage that no coding skills are required to process the images.

3.4 | Temporal analysis of growth dynamics using *PhenoImage*

The image-based phenotyping platforms have enabled quantification of physiological and morphological features in a time-dependent manner. In this context, we performed temporal evaluation of sorghum and wheat growth dynamics under WW and WL conditions using HTP. The RGB and fluorescent derived images were processed using *PhenoImage* for testing sensitivity of the tool to detect subtle physiological changes over time.

The biomass of the plant increases with growth and development, which can be quantified by imaging, and environmental stresses in general slow growth and development (Chen et al., 2014; Röth et al., 2016). To evaluate the changes in plant size in a temporal manner, we traced PSA derived from RGB images under WW and WL conditions. For both sorghum and wheat, PSA showed a gradual increase over time under WW and WL; however, WL conditions exhibited lower PSA relative to WW conditions for the identical time-point (Figure 7).

Furthermore, we processed temporal dataset for wheat plants using HTPheno and PlantCV, which revealed significant differences between WW and WL conditions starting from Day 10–15 (HTPheno) and Day 8–10 and Day 13–15 (PlantCV; Supplemental Table S4). Also, we considered the ratio of WW and WL for the values derived from the three image processing platforms (PlantCV, HTPheno, and *PhenoImage*). We did not find statistical differences between the ratio derived from HTPheno and *PhenoImage* for the three randomly selected days (Day 3, 8, and 14). Contrastingly, significant statistical difference ($P < .05$) was observed for the WW and WL ratio derived from PlantCV and *PhenoImage* for all the tested days (Day 3, 8, and 14; Supplemental Table S4). This indicates that the results obtained from *PhenoImage* are more in line with PlantCV than HTPheno.

25782703, 2021, 1, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/ppj2.20015 by Test, Wiley Online Library on [11/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

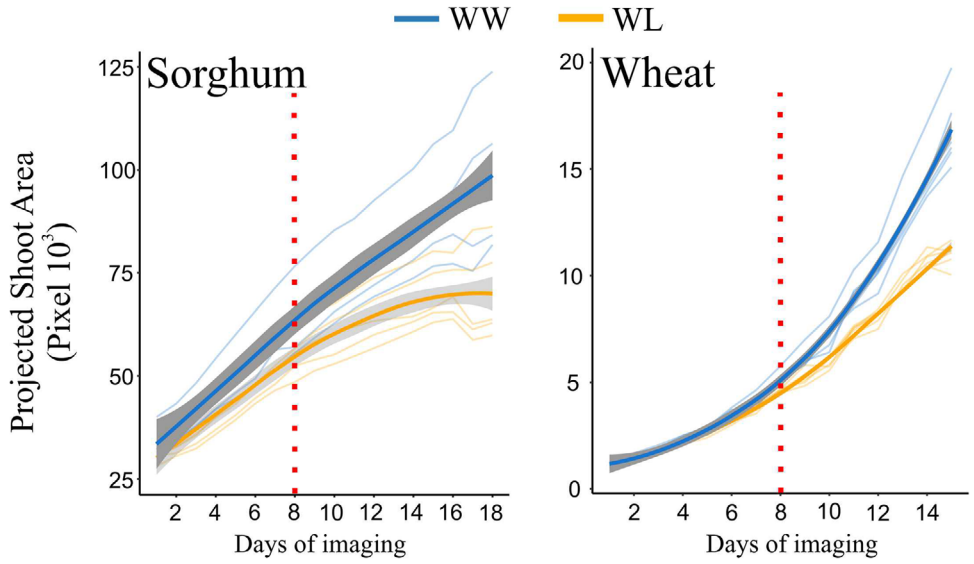

FIGURE 7 Temporal analysis of shoot growth dynamics. Sorghum and wheat plants subjected to well-watered (WW) and water-limited (WL) conditions were imaged for 18 and 15 d, respectively, using the LemnaTec Imaging System. Sorghum and wheat plants subjected to WL conditions reached 30% water holding capacity on Day 18 and 14, respectively. *PhenoImage* derived projected shoot area (PSA) showed significant differences between WW and WL conditions on Day 8 for both sorghum ($n = 5$) and wheat ($n = 6$). Each line represents the temporal trajectory of individual plants under well-watered (blue) and water-limited (orange) conditions. The central bold line represents the mean of WW and WL group and the grey band represents the standard error. For statistics, the *t* test was used. The vertical dashed red line represents the significance difference between WW and WL treatments from that day onward

We evaluated changes in pixel intensities corresponding to the "G" channel and chlorophyll florescence as an indicator of plant health. In principle, the "G" pixel intensity derived from the RGB images reflect the greenness of the plant, which is an indirect indicator of chlorophyll content (Wood et al., 2020). The greenness index (GI) was calculated using the following formula:

$$GI = \frac{N_G}{N_R + N_G + N_B}$$

where N_R, N_G, and N_B are pixel intensity for R, G, and B channel normalized to total pixel count for the respective time-point and treatment. For both sorghum and wheat, we observed higher GI under WW relative to WL conditions (Figure 8).

Furthermore, abiotic stresses such as heat stress or water limitation decreases photosynthetic efficiency and increases non-photochemical quenching resulting in enhanced chlorophyll fluorescence and heat dissipation (Zhao et al., 2017; Paul et al., 2020b). Therefore, we evaluated the dynamics of chlorophyll fluorescence for sorghum and wheat under WW and WL conditions. For this, total pixels corresponding to the red channel were classified into 32 color classes based on their fluorescence intensity. As the stress progressed, fluorescent intensity of pixels changed. To monitor the rearrangement of pixels over time and treatments, we performed HCA. As a result, we detected four clusters (I–IV) each for sorghum and wheat (Figure 9; left panel). For sorghum, the identified clusters distinguished changes related to both development and water treatments (WW and WL). Cluster I comprised fluorescence changes at early time points—1 to 5 days (d) of imaging—wherein cluster II and III were associated with later time points (d6–d17) under both WW and WL conditions (Figure 9). Furthermore, HCA clearly distinguished fluorescence changes linked with water treatment, as cluster II and III were predominant ones under WL and WW conditions, respectively (Figure 9). In the case of wheat, HCA distinguished development-driven fluorescence changes, as early time points (d1–d5) were represented by cluster I and II and late time points (d6–d15) were represented by cluster III and IV (Figure 9; right panel). However, a clear distinction between WW and WL conditions was not observed. These results are in line with previous findings documenting decreased chlorophyll content or photosynthetic activity as a possible penalty on a plant subjected to WL conditions (Mathobo et al., 2017).

Collectively, the results establish that *PhenoImage* can be used to analyze HTP-derived longitudinal phenotypic datasets (RGB and plant pigmentation-based fluorescence images) to detect the occurrence of subtle phenotypic changes in a plant's growth and development.

25782703, 2021, 1, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/ppj2.20015 by Test, Wiley Online Library on [11/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

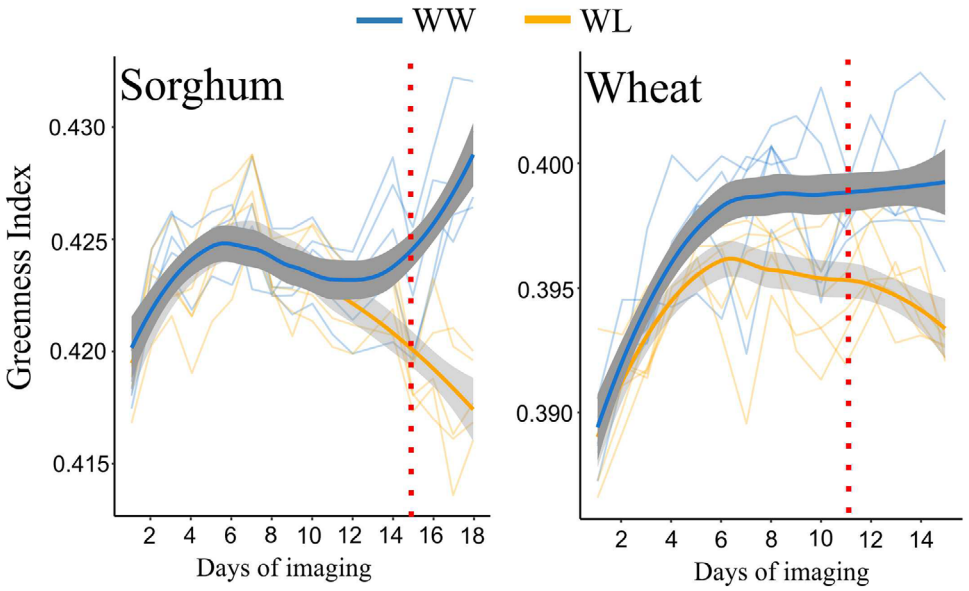

FIGURE 8 Temporal changes in intensity of green color represented by Greenness Index (GI) for sorghum and wheat under well-watered (WW) and water-limited (WL) conditions. Sorghum ($n = 5$) and wheat ($n = 6$) plants subjected to WW and WL conditions were imaged for 18 and 15 d, respectively, using the LemnaTec Imaging System. Each line represents the temporal trajectory of individual plants under well-watered (blue) and water-limited (orange) conditions. The central bold line represents the mean of WW and WL group and the grey band represents the standard error

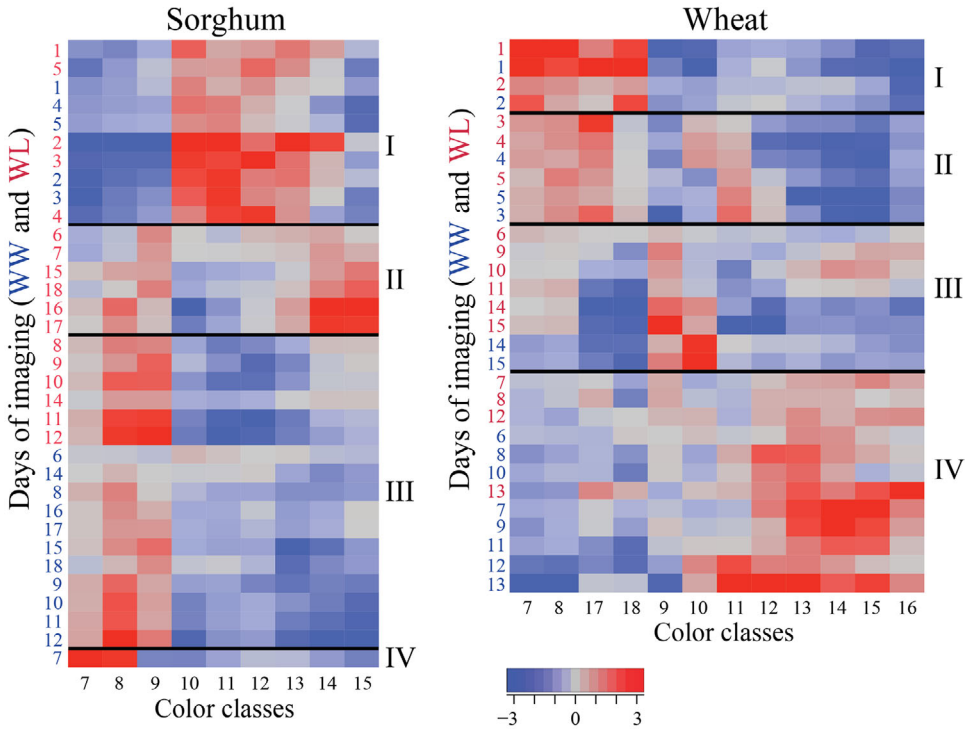

FIGURE 9 Hierarchical Cluster Analysis of fluorescence color classes for sorghum (left panel) and wheat (right panel). Normalized pixel counts corresponding to different color classes were clustered (I–IV) using Ward's method in JMP Pro13 under well-watered (WW) and water-limited (WL) conditions. Days of imaging under WW- and WL-treated plants were represented by blue and red-colored numerals, respectively

25782703, 2021, 1, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/ppj2.20015 by Test, Wiley Online Library on [11/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

4 | CONCLUSION

PhenoImage offers an exhaustive and robust analysis of large-scale plant phenotyping data. The intuitiveness of the application allows scientists with little programming experience to process large-scale datasets on their computers. The tool can also support parallel high performance computing clusters. The availability of multiple functions and filtering parameters provides flexibility to analyze a wide variety of plant species. Moreover, open-source nature provides the possibility to extend the usability of the tool to meet specific user requirements. The current version of the application is designed for analyzing aboveground plant images. However, we plan to extend its usability to examine other tissues such as root or panicles.

AUTHOR CONTRIBUTIONS

Equal contributions by these four first co-authors: **Feiyu Zhu**, **Manny Saluja**, **Jaspinder Singh Dharni**, and **Puneet Paul**. **Feiyu Zhu**: Conceptualization, Software. **Manny Saluja**: Investigation, Validation, Visualization, Writing-review & editing. **Jaspinder Singh Dharni**: Investigation, Validation, Writing-review & editing. **Puneet Paul**: Conceptualization, Investigation, Validation, Visualization, Writing-original draft, Writing-review & editing. **Scott E. Sattler**: Supervision, Writing-review & editing. **Paul Staswick**: Supervision, Writing-review & editing. **Harkamal Walia**: Supervision, Funding acquisition, Writing-review & editing. **Hongfeng Yu**: Supervision, Funding acquisition, Writing-review & editing.

ACKNOWLEDGMENTS

We thank the Nebraska Innovation Campus, University of Nebraska, Lincoln, for their support for the imaging experiments. This work was supported by National Science Foundation Award no. 1736192 to H. Walia and H. Yu, and by a gift to P. Staswick from Kamterter Products, LLC, Waverly, NE.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

SOFTWARE AVAILABILITY

PhenoImage is available in two different versions: (a) Standalone application: this version does not require MATLAB license for its operation; and (b) Regular application: this version does require MATLAB: https://www.mathworks.com/products/matlab.html. Both versions are available at http://wrchr.org/phenolib/phenoimage.

ORCID

Puneet Paul https://orcid.org/0000-0001-8220-8021
Paul Staswick https://orcid.org/0000-0003-2798-0275
Harkamal Walia https://orcid.org/0000-0002-9712-5824
Hongfeng Yu https://orcid.org/0000-0002-0596-8227

REFERENCES

Abeysinghe, S., Baker, M., Chiu, W., & Ju, T. (2008). Segmentation-free skeletonization of grayscale volumes for shape understanding. In Shape modeling and applications. *IEEE International Conference on Shape Modeling and Applications 2008, Proceedings, SMI*. (pp. 63–71).

Bradski, G. (2000). The OpenCV library. *Doctor Dobbs Journal*, *25*, 120–126.

Campbell, M. T., Knecht, A. C., Berger, B., Brien, C. J., Wang, D., & Walia, H. (2015). Integrating image-based phenomics and association analysis to dissect the genetic architecture of temporal salinity responses in rice. *Plant Physiology*, *168*, 1476–1489. https://doi.org/10.1104/pp.15.00450

Chen, D., Neumann, K., Friedel, S., Kilian, B., Chen, M., Altmann, T., & Klukas, C. (2014). Dissecting the phenotypic components of crop plant growth and drought responses based on high-throughput image analysis w open. *Plant Cell*, *26*, 4636–4655. https://doi.org/10.1105/tpc.114.129601

Fahlgren, N., Gehan, M. A., & Baxter, I. (2015). Lights, camera, action: High-throughput plant phenotyping is ready for a close-up. *Current Opinion in Plant Biology*, *24*, 93–99. https://doi.org/10.1016/j.pbi.2015.02.006

Fiorani, F., & Schurr, U. (2013) Future scenarios for plant phenotyping. *Annual Review of Plant Biology*, *64*, 267–291. https://doi.org/10.1146/annurev-arplant-050312-120137

Furbank, R. T., & Tester, M. (2011) Phenomics: Technologies to relieve the phenotyping bottleneck. *Trends in Plant Science*, *16*, 635–644. https://doi.org/10.1016/j.tplants.2011.09.005

Gehan, M. A., Fahlgren, N., Abbasi, A., Berry, J. C., Callen, S. T., Chavez, L., Doust, A. N., Feldman, M. J., Gilbert, K. B., Hodge, J. G., Hoyer, J. S., Lin, A., Liu, S., Lizárraga, C., Lorence, A., Miller, M., Platon, E., Tessman, M., & Sax, T. (2017). PlantCV v2: Image analysis software for high-throughput plant phenotyping. *PeerJ*, *5*, e4088. https://doi.org/10.7717/peerj.4088

Golzarian, M. R., Frick, R. A., Rajendran, K., Berger, B., Roy, S., Tester, M., & Lun, D. S. (2011). Accurate inference of shoot biomass from high-throughput images of cereal plants. *Plant Methods*, *7*, 2. https://doi.org/10.1186/1746-4811-7-2

Gong, P., & He, C. (2014). Uncovering divergence of rice exon junction complex core heterodimer gene duplication reveals their essential role in growth, development, and reproduction. *Plant Physiology*, *165*, 1047–1061. https://doi.org/10.1104/pp.114.237958

Hartmann, A., Czauderna, T., Hoffmann, R., Stein, N., & Schreiber, F. (2011). HTPheno: An image analysis pipeline for high-throughput plant phenotyping. *BMC Bioinformatics*, *12*, 148. https://doi.org/10.1186/1471-2105-12-148

Houle, D., Govindaraju, D. R., & Omholt, S. (2010). Phenomics: The next challenge. *Nature Reviews Genetics*, *11*, 855–866. https://doi.org/10.1038/nrg2897

Hu, M. K. (1962). Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, *8*, 179–187.

Jackson, S. A., Iwata, A., Lee, S.-H., Schmutz, J., & Shoemaker, R. (2011). Sequencing crop genomes: Approaches and applications. *New Phytologist*, *191*, 915–925. https://doi.org/10.1111/j.1469-8137.2011.03804.x

25782703, 2021, 1, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/ppj2.20015 by Test, Wiley Online Library on [11/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Klukas, C., Chen, D., & Pape, J. M. (2014). Integrated analysis platform: An open-source information system for high-throughput plant phenotyping. *Plant Physiology*, *165*, 506–518. https://doi.org/10.1104/pp.113.233932

Knecht, A. C., Campbell, M. T., Caprez, A., Swanson, D. R., & Walia, H. (2016). Image Harvest: An open-source platform for high-throughput plant image processing and analysis. *Journal of Experimental Botany*, *67*, 3587–3599. https://doi.org/10.1093/jxb/erw176

Li, L., Zhang, Q., Huang, D., Li, L., Zhang, Q., & Huang, D. (2014). A review of imaging techniques for plant phenotyping. *Sensors*, *14*, 20078–20111. https://doi.org/10.3390/s141120078

Lobet, G., Draye, X., & Périlleux, C. (2013). An online database for plant image analysis software tools. *Plant Methods*, *9*. https://doi.org/10.1186/1746-4811-9-38

Mathobo, R., Marais, D., & Steyn, J. M. (2017). The effect of drought stress on yield, leaf gaseous exchange and chlorophyll fluorescence of dry beans (*Phaseolus vulgaris* L.). *Agricultural Water Management*, *180*, 118–125. https://doi.org/10.1016/j.agwat.2016.11.005

Minervini, M., Scharr, H., & Tsaftaris, S. A. (2015). Image analysis: The new bottleneck in plant phenotyping [applications corner]. *IEEE Signal Processing Magazine*, *32*, 126–131. https://doi.org/10.1109/MSP.2015.2405111

Moose, S. P., & Mumm, R. H. (2008). Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiology*, *147*, 969–977. https://doi.org/10.1104/pp.108.118232

Paul, P., Dhatt, B. K., Sandhu, J., Hussain, W., Irvin, L., Morota, G., Staswick, P., & Walia, H. (2020a) Divergent phenotypic response of rice accessions to transient heat stress during early seed development. *Plant Direct*, 1–13.

Paul, P., Mesihovic, A., Chaturvedi, P., Ghatak, A., Weckwerth, W., Böhmer, M., & Schleiff, E. (2020b). Structural and functional heat stress responses of chloroplasts of *Arabidopsis thaliana*. *Genes*, *11*, 1–20. https://doi.org/10.3390/genes11060650

Röth, S., Paul, P., & Fragkostefanakis, S. (2016). Plant heat stress response and thermotolerance. In P. Jaiwal, R. Singh, & O. Dhankher (Eds.), *Genetic manipulation in plants for mitigation of climate change* (pp. 15–41). Springer.

Saluja, M., Zhu, F., Yu, H., Walia, H., & Sattler, S. E. (2020). Loss of COMT activity reduces lateral root formation and alters the response to water limitation in sorghum *brown midrib (bmr) 12* mutant. *New Phytologist*, *229*, 2780–2794. https://doi.org/10.1111/nph.17051

Sandhu, J., Zhu, F., Paul, P., Gao, T., Dhatt, B. K., Ge, Y., Staswick, P., Yu, H., & Walia, H. (2019) PI-Plat: A high-resolution image-based 3D reconstruction method to estimate growth dynamics of rice inflorescence traits. *Plant Methods* 15, *162*. https://doi.org/10.1186/s13007-019-0545-2

Tester, M., & Langridge, P. (2010). Breeding technologies to increase crop production in a changing world. *Science*, *327*, 818–822. https://doi.org/10.1126/science.1183700

Varshney, R. K., Nayak, S. N., May, G. D., & Jackson, S. A. (2009). Next-generation sequencing technologies and their implications for crop genetics and breeding. *Trends in Biotechnology*, *27*, 522–530. https://doi.org/10.1016/j.tibtech.2009.05.006

Wood, N. J., Baker, A., Quinnell, R. J., & Camargo-Valero, M. A. (2020). A simple and non-destructive method for chlorophyll quantification of chlamydomonas cultures using digital image analysis. *Frontiers in Bioengineering and Biotechnology*, *8*. https://doi.org/10.3389/fbioe.2020.00746

Yang, W., Feng, H., Zhang, X., Zhang, J., & Doonan, J. (2020). Crop phenomics and high-throughput phenotyping: Past decades, current challenges, and future perspectives. *Molecular Plant*, *13*, 187–214. https://doi.org/10.1016/j.molp.2020.01.008

Yoo, A. B., Jette, M. A., & Grondona, M. (2003). SLURM: Simple Linux utility for resource management. In *Lecture notes in computer science (including subseries Lecture notes in artificial intelligence and lecture notes in bioinformatics)*, *2862*, 44–60. https://doi.org/10.1007/10968987_3

Zhao, X., Chen, T., Feng, B., Zhang, C., Peng, S., Zhang, X., Fu, G., & Tao, L. (2017). Non-photochemical quenching plays a key role in light acclimation of rice plants differing in leaf color. *Frontiers in Plant Science*, *7*, 1968. https://doi.org/10.3389/fpls.2016.01968

Zhu, F., Paul, P., Hussain, W., Wallman, K., Dhatt, B. K., Sandhu, J., Irvin, L., Morota, G., Yu, H., & Walia, H. (2021) SeedExtractor: An open-source GUI for seed image analysis. *Frontiers in Plant Science*, *11*, 581546. https://doi.org/10.3389/fpls.2020.581546

SUPPORTING INFORMATION

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

How to cite this article: Zhu F, Saluja M, Dharni JS, Paul P, et al. *PhenoImage*: An open-source graphical user interface for plant image analysis. *The Plant Phenome J*. 2021;*4:*e20015. https://doi.org/10.1002/ppj2.20015

25782703, 2021, 1, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/ppj2.20015 by Test, Wiley Online Library on [11/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License