

FIELDimageR: An R package to analyze orthomosaic images from agricultural field trials

Filipe Inácio Matias  | Maria V. Caraza-Harter | Jeffrey B. Endelman 

Department of Horticulture, University of Wisconsin, Madison, WI 53706, USA

Correspondence

Filipe Inácio Matias, Department of Horticulture, University of Wisconsin, Madison, WI 53706, USA.

Email: filipematiass23@gmail.com

Abstract

Remote sensing is revolutionizing the phenotyping of agricultural field trials, but for many researchers, the extraction of plot-level results is a bottleneck. We have developed the R package FIELDimageR as a user-friendly tool to analyze orthomosaic images containing many plots. The basic workflow involves cropping and rotating the image, followed by the creation of a shapefile based on the experimental design. The package includes functions to calculate the number of plants per plot, canopy cover percentage, vegetation indices, and plant height. FIELDimageR is publicly available as a GitHub repository (<https://github.com/filipematiass23/FIELDimageR>).

1 | INTRODUCTION

High-throughput phenotyping platforms (HTPPs) are a fast and non-invasive tool for phenotyping plant populations under field conditions (Shi, Thomasson, Murray, Pugh, & Rooney, 2016). The most common HTPPs in agriculture are ground-wheeled or aerial vehicles equipped with multiple sensors for imaging based on geographic information systems (GIS) (Araus & Cairns, 2014). The reflected light captured in these images can be used to draw inference about many traits, including plant structure (e.g., plant height, leaf area index), nutrient status, and the presence of abiotic and biotic stresses (Yang, Liu, Zhao, Li, & Huang, 2017).

To realize the potential benefits of remote sensing, a number of software tools are needed. After collecting overlapping images, tools such as Agisoft PhotoScan (2019) and Pix4Dmapper (Pix4D, 2019) can be used to stitch them together into a single orthomosaic image. Whereas the con-

struction of the orthomosaic is a largely automated process, the delineation of experimental units (i.e., plots) and extraction of plot-level features is more idiosyncratic. At present, many agricultural scientists are relying on (typically unpublished) scripts for use with specialized GIS software programs, such as QGIS (QGIS Development Team, 2019) or ArcGIS (ESRI, 2019). Our objective was to develop a more user-friendly, R-based pipeline to accomplish this task, building on existing image analysis packages, such as *raster* (Hijmans, 2019), *sp* (Bivand, Pebesma, & Gómez-Rubio, 2013), and *EBImage* (Pau, Fuchs, Sklyar, Boutros, & Huber, 2010). The package, called FIELDimageR, is being distributed under the GNU General Public License 2 at <https://github.com/filipematiass23/FIELDimageR>. In this Science Note, we illustrate key features of the software using images from the University of Wisconsin-Madison potato (*Solanum tuberosum* L.) breeding program. A more in-depth tutorial is available for download with the package.

Abbreviations: GIS, geographic information system; HTPP, high-throughput phenotyping platforms; NDRE, Normalized Difference Red Edge Index; NDVI, Normalized Difference Vegetation Index; SC, stand count; UAS, unmanned aerial systems; UAVs, unmanned aerial vehicles; UW, University of Wisconsin; VI, vegetation index.

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.

© 2020 The Authors. *The Plant Phenome Journal* published by Wiley Periodicals, Inc. on behalf of American Society of Agronomy and Crop Science Society of America.

five bands (Blue = R₄₄₅, Green = R₅₄₅, Red = R₆₅₀, Red-Edge = R₇₂₀, and NIR = R₈₄₀). Pix4Dmapper version 4.5.3 was used to make the orthomosaics. Unmanned aerial systems (UAS) surveys collected images (Figure 1) at solar noon on 8 July 2019 at the University of Wisconsin (UW) Hancock Agricultural Research Station (45°39'45" N, 89°17'50" W), with a flight altitude of 60 m above ground, 24 km h⁻¹ flight speed, and 75% image overlap (forward and side). The UAS surveys collected images (Figure 2) on 25 June 2019 at the UW Rhinelander Agricultural Research Station (44°07'11" N, 89°32'10" W) using the same flight parameters. The broad-sense heritability of different phenotypes on a plot basis (Table 1) was calculated by

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_e^2}$$

where the genetic (σ_g^2) and residual (σ_e^2) variances were estimated with *ASreml-R* (Butler, Cullis, Gilmour, & Gogel, 2009). For instance, three potato populations from different market categories were evaluated under an incomplete block design with 110 (Reds), 138 (Chips), and 91 (Russets) individuals.

Core Ideas

- Developed user-friendly software to extract plot-level results from orthomosaic images.
- Plant stand counts can be estimated based on the watershed transformation.
- Potato maturity was estimated more reliably by remote sensing than by visual rating.

3 | RESULTS AND DISCUSSION

3.1 | Preparing the image for data extraction

The workflow in FIELDimageR (Figure 1 and 3) begins with an orthomosaic image and ends with a table of plot-level features. The *fieldCrop* function (Figure 3, Line 2) is used to crop the image to the area of interest, by clicking the four corners in the plotting window of RStudio (Supplemental Figure S1). The exact corners of the trial do not need to be selected at this time, as further refinement of the trial region is accomplished later in the pipeline. Although not necessary, at this point one may wish to create a mask over the soil by

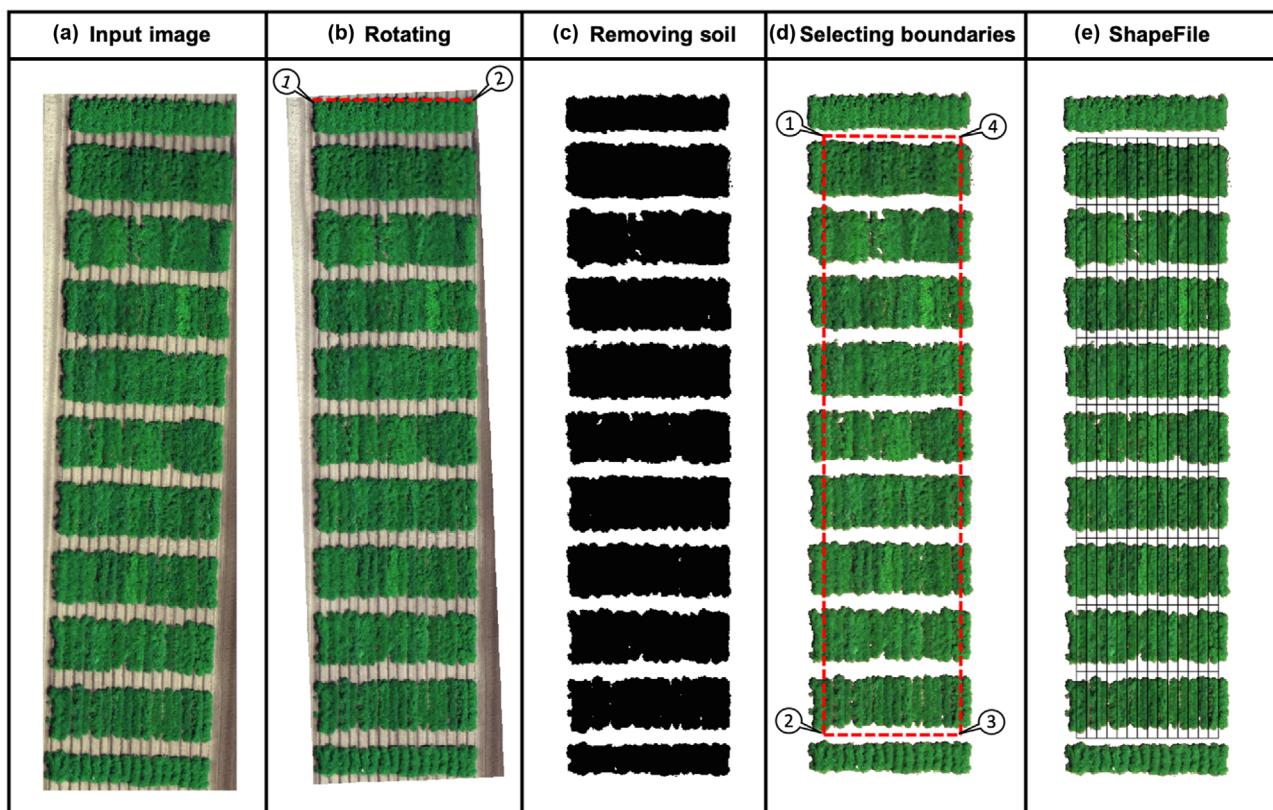


FIGURE 1 FIELDimageR basic steps: (a) Input image (*raster::stack*); (b) Rotating the image (*fieldRotate*); (c) Removing soil using vegetation index (*fieldMask*); (d) Building the plot shapefile by selecting points at the experiment boundaries to eliminate border plots (*fieldShape*); (e) Final shapefile to extract information for each plot (*fieldInfo*)

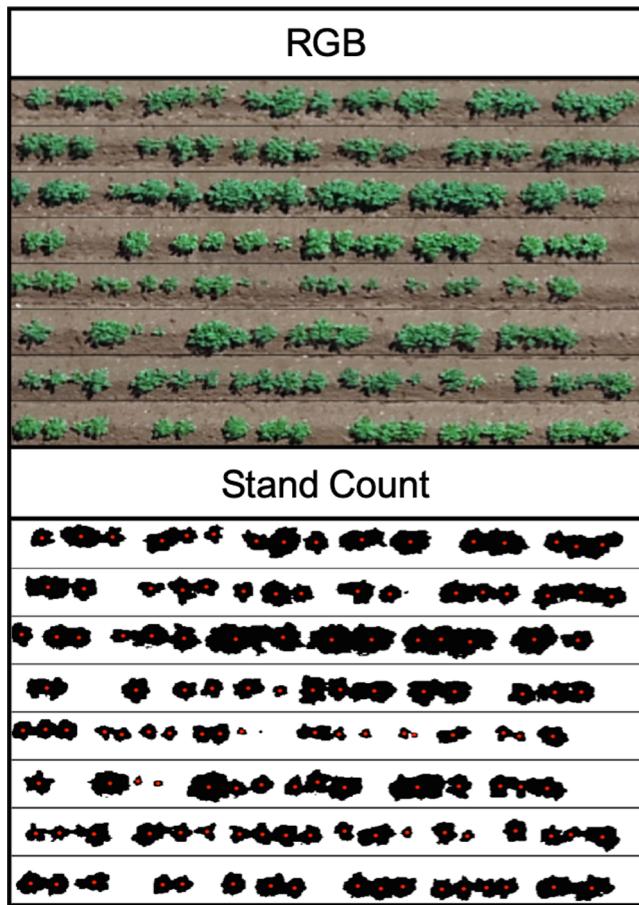


FIGURE 2 Counting the number of plants per plot. Top: RGB image of potato plants (each row is a different genotype). Bottom: Binary mask created with *fieldMask*, and each red dot represents a distinct plant identified with *fieldCount*

classifying pixels as soil vs. plant. The *fieldMask* function (Figure 3, Line 4) accomplishes this classification based on the vegetation index specified by the user. A number of common vegetation indexes have already been defined in the package (Supplemental Table S1), or the user can define a new index (*myIndex*).

When creating a shapefile to facilitate extraction of plot-level features, the goal is to overlay the field design on the orthomosaic. FIELDImageR assumes a row-column design, and for convenience the grid is constructed to be aligned with the plotting window. As a result, the user needs to first rotate the image using the *fieldRotate* function (Figure 3, Line 3). The rotation angle is determined by clicking on two points along the edge of the trial (Figure 1b), and the user indicates whether the line segment should be horizontal or vertical in the rotated image. Next, using the *fieldShape* function (Figure 3, Line 5), the user selects the four corners of the region that will be subdivided evenly based on the number of rows and columns in the design (Figure 1d). The grid is constructed solely from geometric calculations without regard to the

TABLE 1 Broad-sense heritability of plant maturity in three potato breeding populations (Reds, Chips, and Russets) measured by different methods: visual rating, normalized difference vegetation index (NDVI), and normalized difference red edge index (NDRE)

Trial	Visual rating	NDVI	NDRE
Reds	0.56	0.54	0.75
Chips	0.53	0.86	0.71
Russets	0.36	0.51	0.64

presence of alleys or other features in the image (Figure 1e). By default, plot numbers are assigned from left to right, starting in the upper left corner of the grid. Alternatively, the user can also provide a matrix of plot numbers or labels. When multiple orthomosaics (e.g., collected over time) are available with the same coordinate reference system (e.g., by using ground control points or RTK GPS), a shapefile created from one image can be used to extract information from all of them.

3.2 | Counting plants

The number of plants per plot, or stand count, is an important measure of seed quality and early vigor and can be used to adjust yield calculations. The *fieldCount* function (Figure 3, Line 6) uses image segmentation routines from the EBImage package (Pau et al., 2010) to identify distinct plants. First, the distance map transformation is applied to the binary mask created by *fieldMask*, generating a grayscale image of the minimum distance to each background pixel. Next, this grayscale image is interpreted as a topographic map, and the watershed transformation is used to identify the boundaries between the watershed basins (Soille, 2013), which are interpreted as different plants. As shown in Figure 2, *fieldCount* has some ability to identify distinct plants even when they are touching, with accuracy varying from 78 to 100% (Supplementary Table S2). When there are differences in emergence time and early vigor between genotypes (as in Figure 2), it is unlikely that images collected on a single date can be used to estimate stand count accurately for all plots; multiple collection dates will be needed.

3.3 | Vegetation indices

Vegetation indices (VIs) are designed to relate spectral differences in reflectance to plant characteristics. The *fieldIndex* function (Figure 3, Line 8) was developed to calculate the specified VI for every pixel in the image. The *fieldInfo* function (Figure 3, Line 9) calculates the average VI for each plot based on the shapefile. One of the most commonly used VIs is the normalized difference vegetation index (NDVI)

```

# Required package:
library(FIELDImageR)

# 1. Importing image using raster::stack()
mosaic <- raster::stack("mosaic.tif")

# 2. Cropping the field trial from the original image:
mosaic.Crop <- fieldCrop(mosaic=mosaic)

# 3. Rotating the field:
mosaic.Rotated <- fieldRotate(mosaic=mosaic.Crop)

# 4. Removing soil using vegetation index:
mosaic.RemSoil <- fieldMask(mosaic=mosaic.Rotated)

# 5. Building the plot shapefile (ex: 14 columns and 9 rows):
mosaic.Shape <- fieldShape(mosaic=mosaic.RemSoil$newMosaic, ncols=14, nrows=9)

# 6. Counting the number of plants:
mosaic.SC <- fieldCount(mosaic=mosaic.RemSoil$mask, fieldShape= mosaic.Shape$fieldShape)

# 7. Evaluating canopy percentage (greenness):
mosaic.Canopy<- fieldArea(mosaic=mosaic.RemSoil$mask, fieldShape= mosaic.SC$fieldShape)

# 8. Building vegetation indices:
mosaic.Indices <- fieldIndex(mosaic=mosaic.Shape$cropField,
                               Red=1, Green=2, Blue=3, RedEdge=4, NIR=5,
                               index=c("NDVI", "NDRE"))

# 9. Extracting data from field:
mosaic.Info <- fieldInfo(mosaic=mosaic.Indices, fieldShape= mosaic.Canopy$fieldShape)

```

FIGURE 3 FIELDImageR command line pipeline

(Rouse, Haas, Schell, & Deering, 1973), which is a good predictor of aboveground biomass but tends to saturate as canopy cover exceeds 90% (Barnes, Clarke, Richards, Colaizzi, & Haberland, 2000). The normalized difference red edge index (NDRE) has been proposed as more sensitive to plant N status under these conditions (Jain, Ray, Singh, & Panigrahy, 2007; Nigon, Mulla, Rosen, Cohen, & Alchanatis, 2015). In the UW-Madison potato breeding program, plant maturity has traditionally been assessed based on a visual rating (1–9) of senescence around 100 days after planting (DAP). Table 1 compares the broad-sense H^2 of the visual rating compared to NDVI and NDRE for three different trials in 2019, representing the major market classes in the United States. In two trials, both NDVI and NDRE had higher H^2 than the visual rating. In the third trial, only NDRE was more heritable. These results suggest remote sensing can improve the reliability of maturity assessment for the potato breeding program, as has been observed for other traits in maize (*Zea mays* L.) (Anderson, Murray, Malambo, Ratcliff, & Popescu, 2019; Niu, Zhang, Zhang, Han, & Peng, 2019) and sorghum [*Sorghum bicolor* (L.) Moench] (Guo, Zheng, Potgieter, Diot, & Watanabe, 2018; Hu, Chapman, Wang, Potgieter, & Duan, 2018).

3.4 | Plant height

Plant height is strongly correlated with life span, seed mass, time to maturity, and ability to compete for light (Moles, Warton, Warman, Swenson, & Laffan, 2009). Despite its relative simplicity of measurement, plant height can be labor- and time-consuming to collect for research programs with large fields. An alternative way to estimate plant height is using image analysis through the Canopy Height Model (CHM) (Anderson et al., 2019). This model is based on the difference between the Digital Surface Model (DSM) from the soil base (before sprouting) and any moment in the crop growth cycle. The DSM files are one output product from the orthomosaic step (Agisoft PhotoScan, 2019; Pix4D, 2019). This methodology was used with FIELDImageR to estimate plant height for the three UW-Madison potato trials mentioned above with H^2 estimates ranging from 0.70 to 0.85.

3.5 | Processing time

Image size and resolution can drastically impact the processing time. To accelerate the analysis, we recommend two

approaches: (a) reduce the image resolution using the *aggregate* function from package *raster*, and (b) use multiple cores by setting the parameter *n.core* (available for *fieldInfo*, *fieldCount*, and *fieldArea*). For example, Figure 1 has a resolution of 1.7 by 1.7 cm and required 115 s to extract and compute plot averages with *fieldInfo* (MacOS 10.14.5, 3.2 GHz Intel Core i5 with 32 GB RAM, four physical cores). Reducing the resolution to 3.5 by 3.5 cm decreased the time to 45 s without appreciable impact on the results (correlation was 0.99). The processing time can be reduced even further by using multiple cores: 25 and 16 s using two and four cores, respectively.

4 | CONCLUSION

The FIELDimageR package offers plant scientists a convenient set of utilities to extract remote-sensing phenotypes from orthomosaic images of field trials. The vignette distributed with the package illustrates the functionality of the software in more detail. Features not addressed in this Science Note include the ability to estimate plant canopy percentage (*fieldArea* function), non-grid shapefiles (*fieldPolygon* function), and spatial visualization of traits (*fieldPlot* function).

ACKNOWLEDGMENTS

We thank Lin Song for assisting with the literature review of the vegetation indices.

AUTHOR CONTRIBUTIONS

FIM designed the study. FIM and MVCH collected and analyzed the data. FIM, MVCH, and JBE drafted the manuscript. JBE supervised the whole study. All authors read and approved the final version of the manuscript for publication.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

ORCID

Filipe Inácio Matias 

<https://orcid.org/0000-0002-4414-2866>

Jeffrey B. Endelman 

<https://orcid.org/0000-0003-0957-4337>

REFERENCES

- Agisoft PhotoScan. (2019). Agisoft Metashape 1.6.2. Retrieved from <http://www.agisoft.com/downloads/installer/>
- Anderson, S. L., Murray, S. C., Malambo, L., Ratcliff, C., Popescu, S., Cope, D., ... Thomasson, J. A. (2019). Prediction of maize grain yield before maturity using improved temporal height estimates of unmanned aerial systems. *The Plant Phenome Journal*, 2, 190004. <https://doi.org/10.2135/tppj2019.02.0004>
- Araus, J. L., & Cairns, J. E. (2014). Field high-throughput phenotyping: The new crop breeding frontier. *Trends in Plant Science*, 19, 52–61. <https://doi.org/10.1016/j.tplants.2013.09.008>
- Barnes, E. M., Clarke, T. R., Richards, S. E., Colaizzi, P. D., Haberland, J., Li, H., ... Moran, M. S. (2000). Coincident detection of crop water stress, nitrogen status and canopy density using ground based multispectral data. In P. C. Robert, R. H. Rust, and W. E. Larson (Eds.), *Proceedings of the Fifth International Conference on Precision Agriculture, Bloomington, MN, 16–19 July* [CD]. Madison, WI: ASA.
- Bivand, R. S., Pebesma, E., & Gómez-Rubio, V. (2013). *Applied spatial data analysis with R: Second Edition*. New York: Springer.
- Butler, D. G., Cullis, B. R., Gilmour, A. R., & Gogel, B. J. (2009). *ASReml-R reference manual mixed models for S language environments*. Brisbane, QLD: Department of Primary Industries and Fisheries.
- ESRI. (2019). *ArcGIS Desktop: Release 10.7.1*. Redlands, CA: Environmental Systems Research Institute.
- Guo, W., Zheng, B., Potgieter, A. B., Diot, J., Watanabe, K., Noshita, K., ... Chapman, S. C. (2018). Aerial imagery analysis—quantifying appearance and number of sorghum heads for applications in breeding and agronomy. *Frontiers in Plant Science*, 9, 1544. <https://doi.org/10.3389/fpls.2018.01544>
- Hijmans, R. J. (2019). Geographic data analysis and modeling. *R Package version*, 2, 9–23. Retrieved from <https://cran.r-project.org/package=raster/raster.pdf>
- Hu, P., Chapman, S. C., Wang, X., Potgieter, A., Duan, T., Jordan, D., ... Zheng, B. (2018). Estimation of plant height using a high throughput phenotyping platform based on unmanned aerial vehicle and self-calibration: Example for sorghum breeding. *European Journal of Agronomy*, 95, 24–32. <https://doi.org/10.1016/j.eja.2018.02.004>
- Jain, N., Ray, S. S., Singh, J. P., & Panigrahy, S. (2007). Use of hyperspectral data to assess the effects of different nitrogen applications on a potato crop. *Precision Agriculture*, 8, 225–239. <https://doi.org/10.1007/s11119-007-9042-0>
- Moles, A. T., Warton, D. I., Warman, L., Swenson, N. G., Laffan, S. W., Zanne, A. E., ... Leishman, M. R. (2009). Global patterns in plant height. *Journal of Ecology*, 97, 923–932. <https://doi.org/10.1111/j.1365-2745.2009.01526.x>
- Nigon, T. J., Mulla, D. J., Rosen, C. J., Cohen, Y., Alchanatis, V., Knight, J., & Rud, R. (2015). Hyperspectral aerial imagery for detecting nitrogen stress in two potato cultivars. *Computers and Electronics in Agriculture*, 112, 36–46. <https://doi.org/10.1016/j.compag.2014.12.018>
- Niu, Y., Zhang, L., Zhang, H., Han, W., & Peng, X. (2019). Estimating above-ground biomass of maize using features derived from UAV-based RGB imagery. *Remote Sensing*, 11, 1261. <https://doi.org/10.3390/rs1111261>
- Pau, G., Fuchs, F., Sklyar, O., Boutros, M., & Huber, W. (2010). EBImage—an R package for image processing with applications to cellular phenotypes. *Bioinformatics*, 26, 979–981. <https://doi.org/10.1093/bioinformatics/btq046>
- Pix4D. (2019). Measure from images. Retrieved from <https://www.pix4d.com>
- QGIS Development Team. (2019). QGIS—a free and open source geographic information system. Retrieved from <https://qgis.org/>
- Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. (1973). Monitoring vegetation systems in the Great Plains with ERTS (Earth Resources Technology Satellite). *Journal of the American Statistical Association*, 68, 397–415. <https://doi.org/10.1080/0162145708834683>

- Resources Technology Satellite). *Third Earth Resources Technology Satellite-1 Symposium*. Retrieved from <https://ntrs.nasa.gov/search.jsp?R=19740022614>
- Shi, Y., Thomasson, J. A., Murray, S. C., Pugh, N. A., Rooney, W. L., Shafian, S., ... Yang, C. (2016). Unmanned aerial vehicles for high-throughput phenotyping and agronomic research. *PLOS ONE*, 11, e0159781. <https://doi.org/10.1371/journal.pone.0159781>
- Soille, P. (2013). *Morphological image analysis: Principles and applications*. Berlin: Springer-Verlag.
- Yang, G., Liu, J., Zhao, C., Li, Z., Huang, Y., Yu, H., ... Yang, H. (2017). Unmanned aerial vehicle remote sensing for field-based crop phenotyping: Current status and perspectives. *Frontiers in Plant Science*, 8, 1111. <https://doi.org/10.3389/fpls.2017.01111>

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: Matias FI, Caraza-Harter MV, Endelman JB. FIELDimageR: An R package to analyze orthomosaic images from agricultural field trials. *The Plant Phenome J.* 2020;3:e20005.
<https://doi.org/10.1002/ppj2.20005>