Package: GiNA (via r-universe)

September 7, 2024

Type Package

Title High Throughput Phenotyping

Version 1.0.1

Date 2016-03-29

Author Luis Diaz-Garcia and Giovanny Covarrubias-Pazaran, with collaborations of Brandon Schlautman, Walter Salazar and Juan Zalapa.

Maintainer Giovanny Covarrubias-Pazaran <covarrubiasp@wisc.edu>

Description Performs image segmentation in fruit or seeds pictures in order to measure physical features in a high-throughput manner for genome-wide association (GWAS) and genomic selection programs.

Depends R (>= 3.2.1), methods, stats, EBImage, png, parallel, doParallel, foreach

License GPL-3

URL http://www.wisc.edu

NeedsCompilation no

Repository CRAN

Date/Publication 2016-04-10 00:27:50

Contents

GiNA-package | *Software for high throughput phenotyping of fruit crops*

Description

GiNA is a package designed for automatic highthoroughput phenotyping analysis of images. The program was conceived to read files with .JPG and .TIFF extensions (which stands for Joint Picture Experts Group) and uses image segmentation based on the `EBImage` package in order to recognize fruits and compute its features such as area, perimeter, color (in 3 channels), shape and volume to be used for GWAS, QTL or descriptive studies. The core of the package relays in 3 functions; 1) `cutoff` is the function in charge of determining the best cut off value for segmentation, the color scale in R goes from 0 to 1 and in other siftwares from 0 to 265, the function detects the color of the background (black or white) and returns several pictures of segmentation at different cutoff values in order to pick the most suitable for your crop, 2) `pixArea` reads the pictures and returns a plot with the pixel area for each object in the picture in order to provide an estimate of the minimum value to use in the real analysis, the value obtained from this function should be passed to the `scanCRAN` function as the minArea to discard all noise from the picture, 3) `scanCRAN` is the core of the package and performs image analysis for all pictures stored in the folder path provided at the cutoff value selected. The output from the `scanCRAN` function is a list structure with each element of the list being a dataframe with all measurements for the fruits or objects found in the picture.

Additional plotting functions to visualize the data have been provided such as `color.plot` which displays the digital color detected by the computer for each picture, the input of this function is the output of the scanCRAN function. The function `jet.scan` returns a heat.map of the variability found in the pictures and takes the same input than 'color.plot' function. For additional information such as tutorials and most recent releases please visit our website http://cggl.horticulture.wisc.edu/software/.

We have spent valuable time developing this package, please cite it in your publication:

Diaz-Garcia L, Covarrubias-Pazaran G, Schlautman B, Zalapa J. GiNA: A flexible high throughput phenotyping tool. http://horticulture.wisc.edu/cggl/ZalapaLab/People.html. 2015.

Author(s)

Luis Diaz-Garcia, Giovanny Covarrubias-Pazaran, Brandon Schlautman, Juan Zalapa.

References

Diaz-Garcia L, Covarrubias-Pazaran G, Schlautman B, Salazar W, Zalapa J. (2015) GiNA: A flexible high throughput phenotyping tool. http://horticulture.wisc.edu/cggl/ZalapaLab/People.html. 2015.

Ben Hui Liu. 1998. Statistical Genomics. CRC Press LLC. ISBN 0-8493-3166-8.

See Also

http://cggl.horticulture.wisc.edu/software/

Examples

```
####=========================================####
#### For CRAN time limitations most lines in the
#### examples are silenced with one '#' mark,
#### remove them and run the examples
####=========================================####

###########################################
###              GiNA package         ####
###########################################

# XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX #
###########################################
### call library
library(GiNA)

### XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX #
###########################################
### EXPLORE YOUR DATA FIRST TO DETERMINE THE 'minArea' AND 'cutoff' VALUES
### indicate folder where sample pictures are, i.e.
### folder <- "C:/Users/zalapalab/Desktop/examp"

### for demonstration purposes we will call a sample picture
### included in the package, run the following 4 lines

data(GINA.sample) # RUN
display(GINA.sample)
writeImage(GINA.sample, "gina_cran.JPG") # RUN
folder <- getwd() # RUN

### XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX #
##############################################
### check the best cutoff values for your pictures

?cutoff
cutoff(folder, cores=1) # RUN

### go to your working directory and analyze the segmentation
### results, you will agree with us than 0.5 is the best value
### for segmentation in the 'cutoff_vals' folder FOR THIS SAMPLE PICTURE

### XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX #
############################################
### check the best minimum area (in pixels)
### to get rid of little trashes in the picture
### can be provided to the scanCRAN function if
### you don't like the default

?pixArea
#area <- pixArea(folder,cutoffvalue=0.5,cores=1) # RUN!!!

### XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX #
```

```
###########################################
### RUN the scanCRAN function providing your
### cutoff and minimum area values found previously

?scanCRAN
#folder <- getwd()
#mydata <- scanCRAN(folder,cutoffvalue=0.5,cores=1, fact = 1) # RUN!!!
#str(mydata) # RUN!!!

### XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX #
###########################################
### see colors of your fruits in each picture

#color.plot(mydata) # RUN!!!
#jet.scan(mydata, var=3) # RUN!!!

### XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX #
###########################################
```

---

brewer.pal   *Generate a sequence of colors for groups.*

Description

wrapper of brewer.pal function from RColorBrewer.

Usage

```
brewer.pal(n, name)
```

Arguments

n	Number of different colors in the palette, minimum 3, maximum depending on palette.
name	A palette name from the lists below.

Value

A vector of colors.

Examples

```
{
# Plot a colorbar with brewer.pal
mypalette<-brewer.pal(7,"Greens")
}
```

---

color.plot   *Analyzing the digital color of objects*

Description

This function was design to assess the color output of the scanCRAN function by drawing the digital color detected by the program in a way that the user can assess the colors that the computer is detecting for each object in the picture.

Usage

```
color.plot(x, cex.axis = 0.7, cex = 5)
```

Arguments

x   the data to be used for plotting, should be the output from the 'scanCRAN function.

cex.axis   the size of the dots in the plot

cex   the size of the font of the x labels

Value

If parameters are indicated correctly the function returns:

$plot a plot showing the digital color detected by the program for each of the objects found in the picture.

References

We have spent valuable time developing this package, please cite it in your publication:

Diaz-Garcia L, Covarrubias-Pazaran G, Schlautman B, Zalapa J. GiNA: A flexible high throughput phenotyping tool. http://horticulture.wisc.edu/cggl/ZalapaLab/People.html. 2015.

Examples

```
## x <- ScanCran(folder=0.8,minArea=6000)
## color.plot(x, cex.axis = 0.7, cex = 5)
```

---

cutoff   *Finding the optimal cutoff value for segmentation*

Description

This function reads the picture provided in a path and performs image segmentation in a gray scale version of the picture at different values of color between 0 and 1 depending of the background color. If the background is black the function will segment the picture at values of 0.01 to 0.12 and if a white background is used values 0.5 to 0.8 will be used in order to determine the optimal value of segmentation, such values can be modified by the user in the argument 'prefs'. The resulting images at different cutoff values will be stored in a folder created with the name 'cutoff_vals' in the same folder provided by the user.

Usage

```
cutoff(folder, cores = 1, prefs = NULL)
```

Arguments

folder	A path where the sample pictures are located.
cores	A scalar value indicating the number of cores that you want to use for paralleliz-ing the function, default is 1
prefs	a vector containing the values of the cutoffvalues to be used in the analysis.

Details

In image analysis segmentation can be done based on different parameters, the most common way to do it is by maximizing the contrast of a background with the object of study (i.e. fruits). Most software uses white or black backgrounds to maximize the contrast with the objects, but a cuttoff value needs to be set in order to get accurate estimates of the objects in the pictures. GiNA has been optimized to be used with any background but an initial cutoff value needs to be provided to the core function scanCRAN. The purpose of this function is to try different cutoff values and decide the optimal.

Value

If parameters are indicated correctly the function returns:

$pics pictures in a folder named "cutoff_vals" with different values for segmentation to help the user define the optimum cutoff value for the crop.

References

We have spent valuable time developing this package, please cite it in your publication:

Diaz-Garcia L, Covarrubias-Pazaran G, Schlautman B, Zalapa J. GiNA: A flexible high throughput phenotyping tool. http://horticulture.wisc.edu/cggl/ZalapaLab/People.html. 2015.

Examples

```
library(GiNA)

data(GINA.sample) # RUN
display(GINA.sample)
writeImage(GINA.sample, "gina_cran.JPG") # RUN
folder <- getwd() # RUN
?cutoff
cutoff(folder, cores=1) # RUN
```

GINA.sample *cranberry image*

Description

Sample image from the cranberry genomics program. The image displays 25 cranberry fruits from a single genotyped captured by a camera in a white background with black reference circles for standarization purposes.

Usage

```
data(GINA.sample)
```

Format

An image with 3 layers for green, red and blue.

Source

Cranberry Genomics Program, Madison, WI, USA.

References

Diaz-Garcia L, Covarrubias-Pazaran G, Schlautman B, Salazar W, Zalapa J. (2015) GINA: A flexible high throughput phenotyping tool. http://horticulture.wisc.edu/cggl/ZalapaLab/People.html. 2015.

Ben Hui Liu. 1998. Statistical Genomics. CRC Press LLC. ISBN 0-8493-3166-8.

Examples

```
####=======================================####
#### For CRAN time limitations most lines in the
#### examples are silenced with one '#' mark,
#### remove them and run the examples
####=======================================####
data(GINA.sample)
#display(GINA.sample)
#str(GINA.sample)
```

---

jet.scan   *Assessing variability for measurements using jet plot*

Description

Same than `color.plot` function, this function was designed to assess the results from the `scanCRAN` function by providing the output of such function and draw plots showing the variation found in the objects measured in each object.

Usage

```
jet.scan(mydata, var = 1)
```

Arguments

mydata   the data to be used for drawing the jet plot, is the output from the `scanCRAN` function

var   the variable to be plotted from several variables measured by the scanCRAN function.

Value

If parameters are indicated correctly the function returns:

$plot a plot showing the digital representation for the variable selected using a plot with jet colors to emphasize the variability within pictures.

References

We have spent valuable time developing this package, please cite it in your publication:

Diaz-Garcia L, Covarrubias-Pazaran G, Schlautman B, Zalapa J. GiNA: A flexible high throughput phenotyping tool. http://horticulture.wisc.edu/cggl/ZalapaLab/People.html. 2015.

Examples

```
## x <- scanCran(folder=0.8,minArea=6000)
## jet.scan(x, var = 1)
```

---

pixArea   *Finding the average area in pixels for objects*

Description

Even not required, the average value area in pixels for the objects analyzed is a parameter of interest when the objects to be analyzed are particularly small such as seeds or extremely small fruits. Therefore, the use of this function provides a plot of the are in pixels for all the objects in the picture for a given cutoff value provided. Once the value has determined this can be provided to the scanCRAN function in the argument 'minArea' to make the function perform faster.

Usage

```
pixArea(folder, cutoffvalue = 0.5, cores = 1, square = 10, fact = 0.25)
```

Arguments

folder	A path where the pictures are located
cutoffvalue	A scalar value where the fruits are distinguishable from the background in the color scale handled by R which is 0-1
cores	A scalar value indicating the number of cores that you want to use for paralleliz-ing the function
square	size of the square to brush smaller objects from the picture
fact	a scalar value between 0-1 indicating at what extent the picture should be re-duced in size in order to make computations faster, by default the pictures are sized at the 25 percent of their original size. If no reduction is required the user should set the argument to 1.

Value

If parameters are indicated correctly the function returns:

$plot a plot showing the size of different objects in the picture in number of pixels.

References

We have spent valuable time developing this package, please cite it in your publication:

Diaz-Garcia L, Covarrubias-Pazaran G, Schlautman B, Zalapa J. GiNA: A flexible high throughput phenotyping tool. http://horticulture.wisc.edu/cggl/ZalapaLab/People.html. 2015.

Examples

```
library(GiNA)
data(GINA.sample) # RUN
display(GINA.sample)
writeImage(GINA.sample, "gina_cran.JPG") # RUN
folder <- getwd() # RUN
area <- pixArea(folder,cutoffvalue=0.5,cores=1) # RUN!!!
```

---

scanCRAN   *High throughput phenotyping using scanCRAN*

Description

This function is the core of the GiNA package which is in charge of reading the pictures find in the folder provided and perform the image segmentation in order to find the objects in the picture (i.e. fruits) and measure parameters such as area, perimeter, shape, volume, gray color, red, green and blue color.

Usage

```
scanCRAN(folder, cutoffvalue=NULL, minArea = NULL,
        cores = 1, gray = TRUE, stand = c(0, 0),
        fact = 0.25)
```

Arguments

folder	A path where the pictures are located
cutoffvalue	A scalar value (0-1) where the fruits are distinguishable from the background in the color scale handled by R which is 0-1. if not passed the function will determine the color of the background and assign the value of 0.08 if black and 0.6 if white. We highly recommend to use the [cutoff] function to find the optimum segmentation value which will determine the success of your experiment.
minArea	A scalar number indicating the minimum size in pixels for a real object in the picture to avoid the program to confuse some little trash with an object. The function [cutoff] can provide an idea of the sizes of real objects and noise. If not passed the argument uses minArea=(length/40)*(width/40), where length and width refers to the dimensions of the picture
cores	A scalar value indicating the number of cores that you want to use for parallelizing the function
gray	a TRUE/FALSE value indicating if the user want the color to bereturned in grayscale besides the RGB color
stand	a vector with 2 scalar values for width and length of your references
fact	a scalar value between 0-1 indicating at what extent the picture should be reduced in size in order to make computations faster, by default the pictures are sized at the 25 percent of their original size. If no reduction is required the user should set the argument to 1.

Details

In order to exploit the advantages of this function we highly recommend the users to determine the best cutoff value with the use of the [cutoff] function to achieve a good segmentation of the objects. The package is sensitive to the prescense of shades in the picture, for that reason we highly recommend the use of black background and reference circles in a clear color (i.e. white).

This is a rough idea of how you should take the pictures, where the circles are your references and the dots are your fruits. Your fruits shouldn't block the references and YOUR REFERENCES SHOULD BE ALWAYS ON THE SIDES OF YOUR PICTURE.

O . . . . O

O . . . . . O

O . . . . . O

For additional information such as tutorials and most recent releases please visit our website http://cggl.horticulture.wisc.edu/s

Value

If parameters are indicated correctly the function returns:

$x3 a list of dataframes, one for each picture provided. The dataframe contains the measurements for each object (rows) found in the puncture for the variables (columns) area, perimeter, length, width, shape (ratio of length/width), volume and red, green, blue, grey color.

References

We have spent valuable time developing this package, please cite it in your publication:

Diaz-Garcia L, Covarrubias-Pazaran G, Schlautman B, Zalapa J. GiNA: A flexible high throughput phenotyping tool. http://horticulture.wisc.edu/cggl/ZalapaLab/People.html. 2015.

Examples

```
library(GiNA)
data(GINA.sample) # RUN
writeImage(GINA.sample, "gina_cran.JPG") # RUN
folder <- getwd()
mydata <- scanCRAN(folder,cutoffvalue=0.5,cores=1, fact = 1) # RUN!!!
```

Index