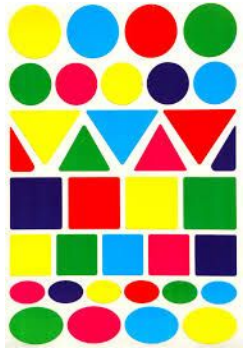
Image Processing for Plantix

Mehul Jain 2019A3PS1315H

What is an Image?



1) A 2D representation of vectors:

    1.1) Value of vector - RGBA values

    1.2) Direction of vector - column and row index of vector

All physical quantities are measured in pixels. Quantities - width, height, size, etc. For example,

Height = 271 pixels

Width = 186 pixels

Image is represented by matrix of dimension, 186 x 271

= no. of columns x no. of rows

= Size = (186, 271)

How matrix helps?

Matrix weightage

Matrix helps us understand matrix weightages which are used to preprocess images.

1) Each matrix represents a collection of pixels. Each pixel has its own Red, Green, Blue and Opacity or RGBA values. This is the weightage of one pixel.
2) Pixels combining together give rise to an image.

The image given here is generated using 2D arrays in Numpy.

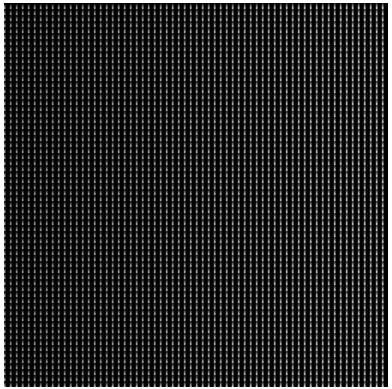
A representation a combination of each of the 256 pixels possible.

This is a 256 by 256 dimensional image.

The modification of weightages of matrix is done by filters.

Filter is a matrix, F when multiplied with image matrix, I which is

I.dot(F) = Pre processed Images

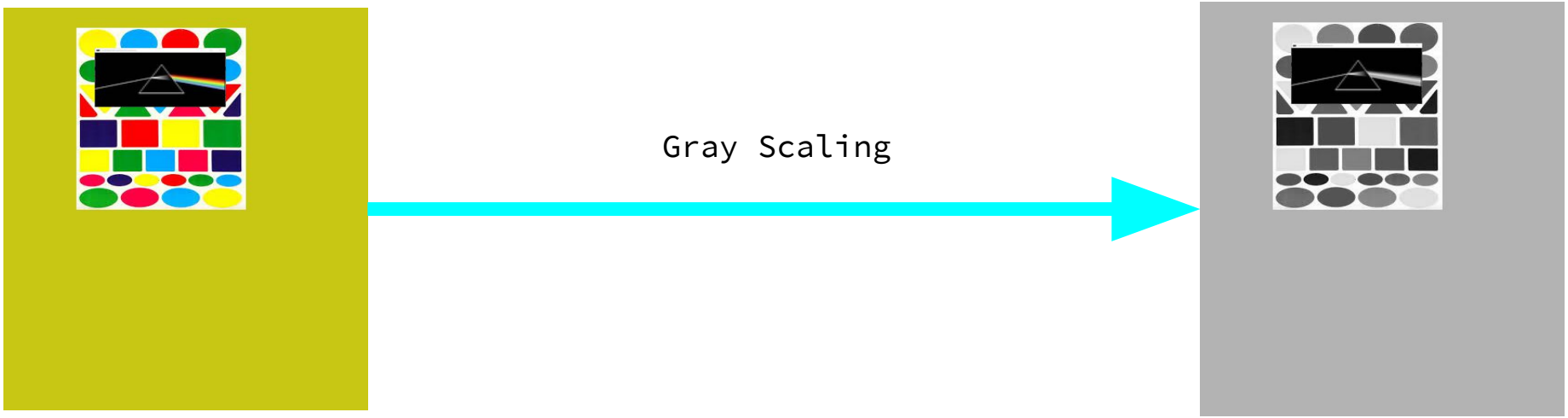


Preprocessing images

Gray Scaling Images

The most basic preprocessing function for images.

$$P(r,g,b)[i, j] = Average(P(r) + P(g) + P(b))[i,j]$$



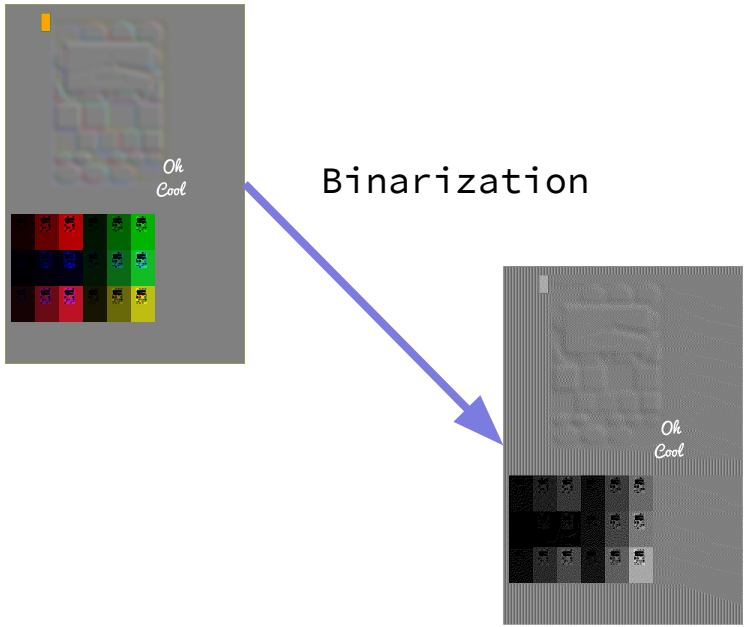
Binarizing images

Binarizing images is to compare each gray scale value of a pixel with a certain threshold. If the threshold, h is compared with pixel, P[i,j] then:

```
P[i,j] = 0 if GrayScale(P[i,j]) < h

P[i,j] = 255 if GrayScale(P[i,j]) >= h
```

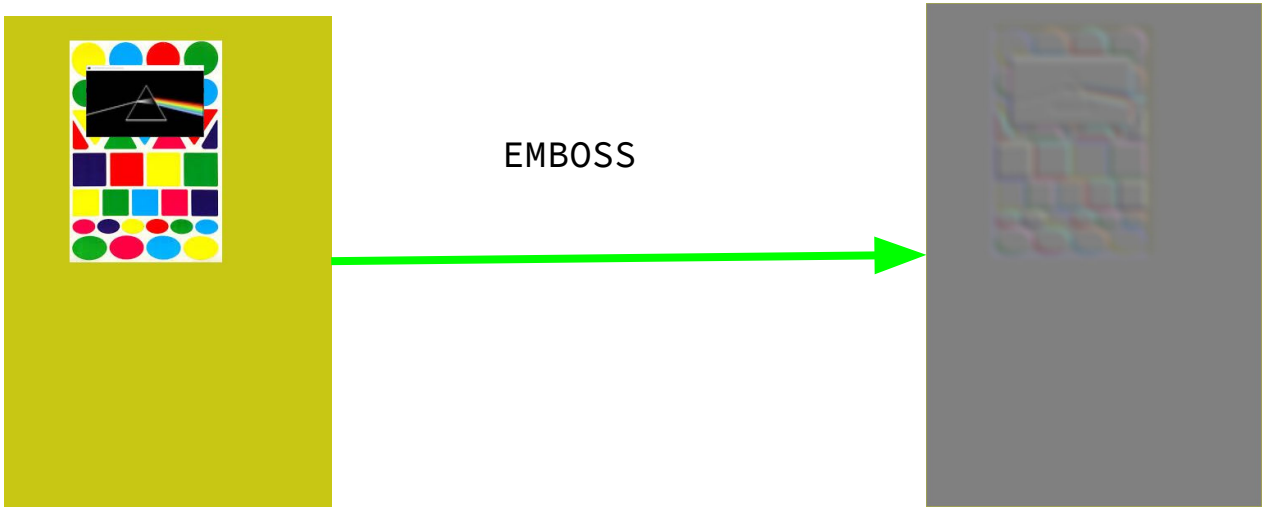
It is used in classification algorithms.



EMBOSSING IMAGES

Use of convolution calculations is used in embossing images. It involves convolution of small squares of images with the emboss filter.

Edge detection



Enhancing images

Enhancing is done to get channels of the images by dot product with filters.

The red matix filter is

(1,0,0,0,

0,0,0,0,

0,0,0,0)

The green matix filter is

(0,0,0,0,

0,1,0,0,

0,0,0,0)

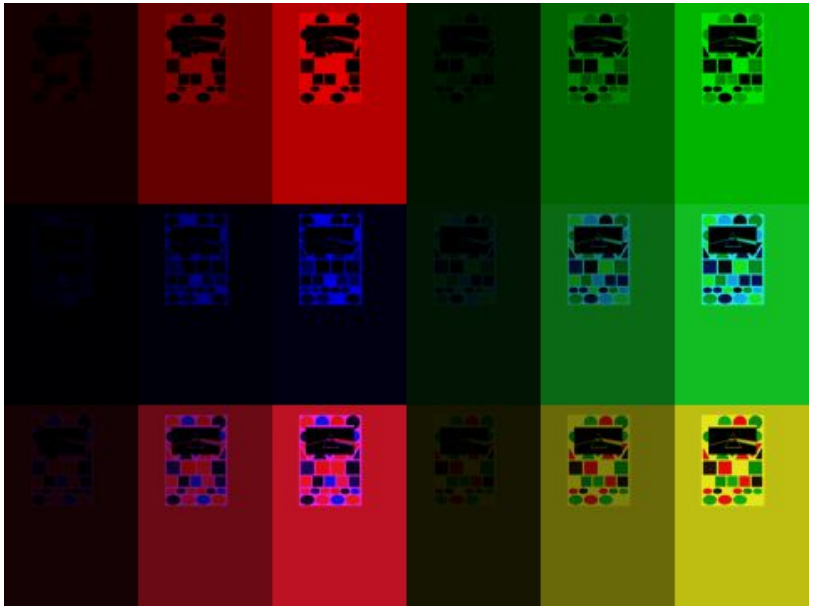
The blue matix filter is

(0,0,0,0,

0,0,0,0,

0,0,1,0)

Color Detection
In YOLO Algorithm



What to do after preprocessing?

Processing libraries

1) Processing libraries like OpenCV, Tesseract, Kraken, etc. are used on preprocessed images.
2) Involves Linear Algebra libraries as well as Calculus.

Note : Preprocessing methods used should match the goal of the problem statement so that processing libraries can find it easy to operate on.

For example:

1) RGBA images are needed for Kraken.
2) Binarized or Bicubic images are needed for Tesseract.

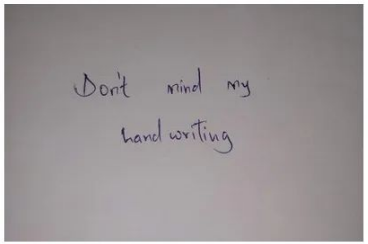
TESSERACT

This library is used to identify text from images of RGB format.

ANTIALIAS function is used to avoid distortion for better text recognition.

It uses Deep Neural Networks for identifying parts of a character.

Result:



The extracted text from the image above is: **ad oviling**

Before applying ANTI ALIAS function

The extracted text from the image above ts: ad oviling

After applying ANTI ALIAS function

Result:

The extracted text from the image above ts: ad oviling

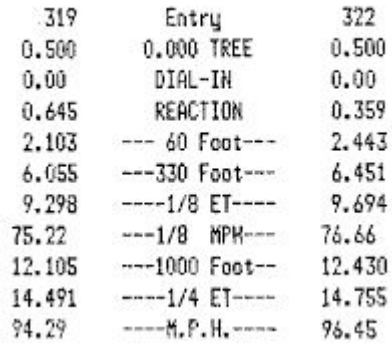
Kraken

Kraken is used in PDF page division.

Divides a Bicubic/Binary image into columns.

Images like bills, books and online books are divided into columns.

Regex formulas can be applied to the text in columns for filtering the text extracted.



319	Entry	322
0.500	0.000 TREE	0.500
0.00	DIAL-IN	0.00
0.645	REACTION	0.359
2.103	--- 60 Foot---	2.443
6.055	---330 Foot---	6.451
9.298	----1/8 ET----	9.694
75.22	---1/8 MPH---	76.66
12.105	---1000 Foot--	12.430
14.491	----1/4 ET----	14.755
94.29	----M.P.H.----	96.45

OpenCV

1) Identify a non structured data from an image data structure.
2) OpenCV is an image processing library generally used in Computer Vision projects.
3) It is a trained library making image processing faster.

It uses many algorithms ranging from the basic bilinear processing to complex edge detection libraries. It is written in C++.

The library of OpenCV in python is called cv2.

How to use OpenCV?

OpenCV uses Haar Cascade XML files for Object Detection. The steps involve:

1) Collecting negative and positive datasets.
2) Train the datasets separately.
3) Generate Haar Cascade file and use the algorithm generated on the files.

OpenCV uses padding followed by edge detection and it later classifies the generated algorithm for a negative and positive image. OpenCV is generally used for binary classification but can be used for multi class classification projects.

OUTPUT OF OPENCV

OUTPUT 1

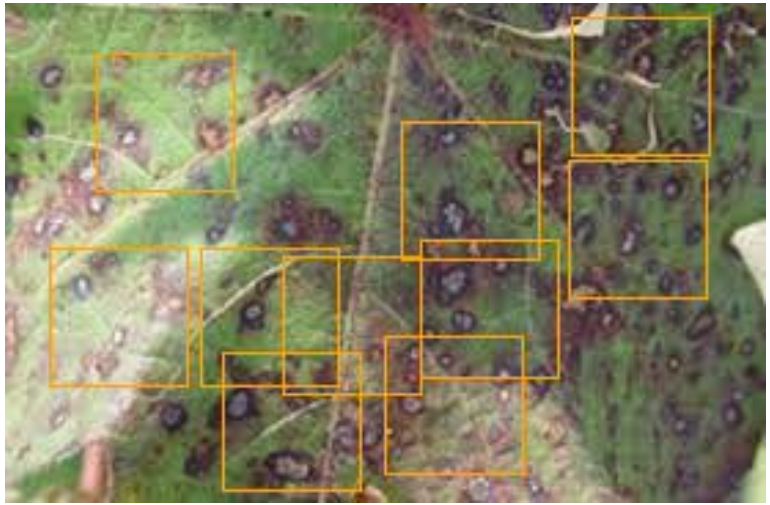
Face

Detection

Using default

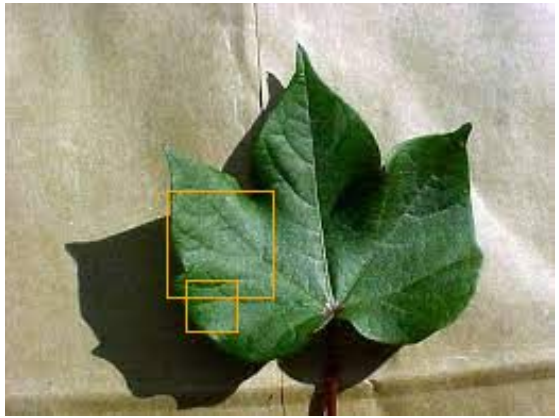
Face XML.

Output 2



Generated an XML file using

Cascade-Trainer-GUI



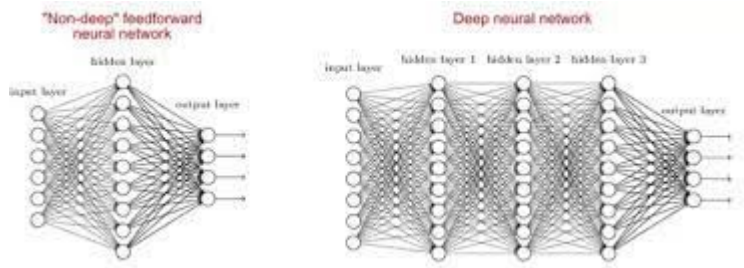
What are the disadvantages of OpenCV?

1) OpenCV generates algorithms based on negative and positive images. This suggests that it can be used for binary classification images. Training a Haar Cascade file is a tedious job as it requires a lot of images to generate an accurate algorithm. (Around 5 minutes for 50 images for leaf disease XML file). It is also not that accurate.
2) A multi class classification XML file can also be generated but it would require the problem to be broken down into many binary classification problems. This means that it would do pairwise operations resulting in quadratic time complexity. This results in slower operations and results.

What to use next?

1) Use ResNets, 2DConv, YOLO object detection algorithms.

Deeper neural networks give us better outputs. The term deeper neural networks means that the numbers parameters as well as the accuracy of weightage values is higher than processing libraries already available as open source.



Resources

1. IEEE papers provided by Bhavin Sir.
2. My own github account having all the source code for the images used in the presentation: https://github.com/mehul14062001/Practice-School-1
3. Cascade Trainer GUI app at: https://amin-ahmadi.com/cascade-trainer-gui/

   used for training own images to generate XML files.