Rice Seed Classification

Niranga Udumulla

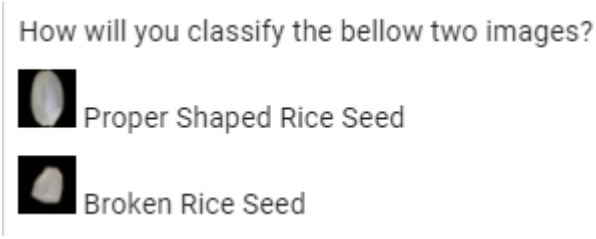
December 6, 2020

Abstract

In this report consist the classification of the rice seed into two categories using their images.This image set includes Proper Shaped Rice Seed and Broken Rice Seed images and use their area to find out the exact category.

1 Introduction

Following figure shows the 2 types of the rice seeds and 395 total images have used for this analysis.



2 Theory

In this lab demonstrate how to use machine learning based technique to identify the broken rice seeds from a set of rice seed images.

3 Procedure

We calculate the seed contain area of the image and which used to make a binary classification

3.1 Load Data

We have 397 seed images and we take 377 for training and rest of 20 images were used for testing the model.There are few libraries were installed for the analysis.

```
import numpy as np
import math, os, sys
```

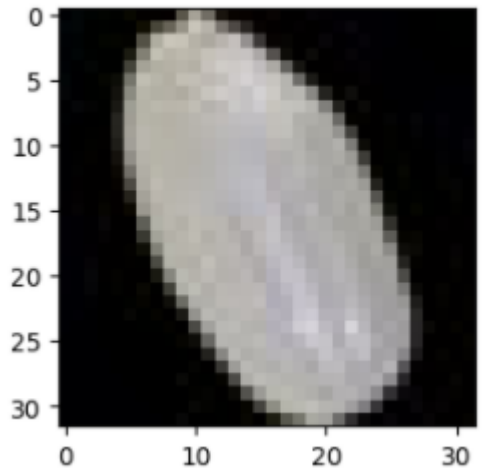
```
import itertools

import matplotlib.pyplot as plt
plt.style.use('default')
from scipy import ndimage

from skimage import measure, morphology
from skimage.io import imsave, imread
from skimage.color import rgb2gray
from skimage.filters import threshold_otsu
from skimage.transform import resize

from sklearn import svm, datasets
from sklearn.metrics import confusion_matrix
import pandas as pd
```

4 Analysis

Fllowing figure shows the single image of a proper rice seed.

```
#let's visualize a single file
image = imread("image/train/proper/100.jpg")
plt.figure(figsize=(3,3))
plt.imshow(image)
```

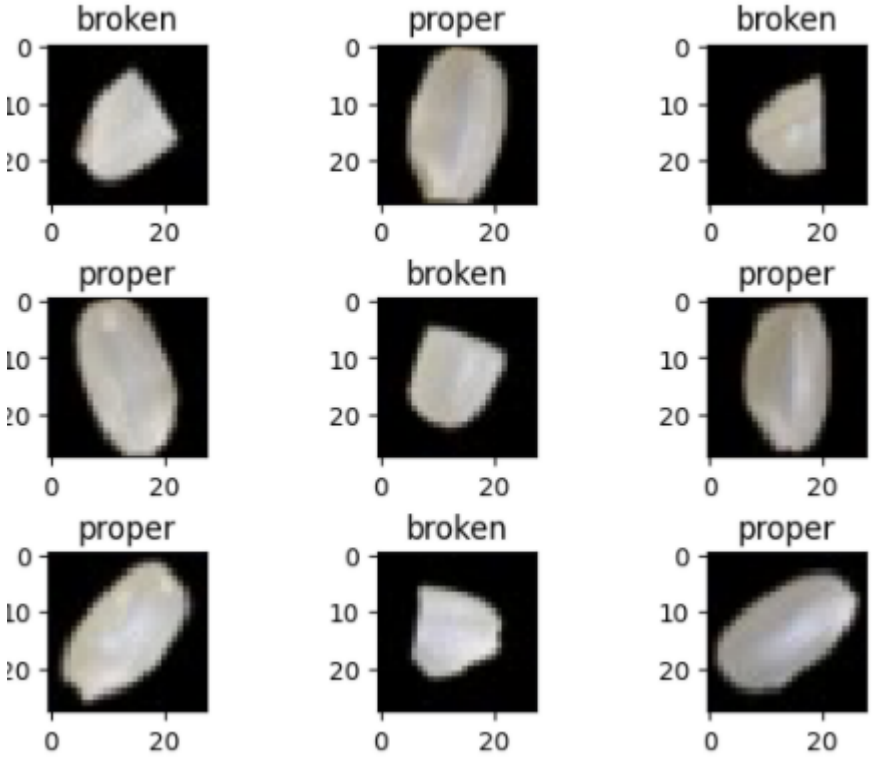
```
<matplotlib.image.AxesImage at 0x7f30be
```



. In the next step,we resize our images into 28*28 fixed resolution and labeled the proper images as 0 and broken seed images as 1.

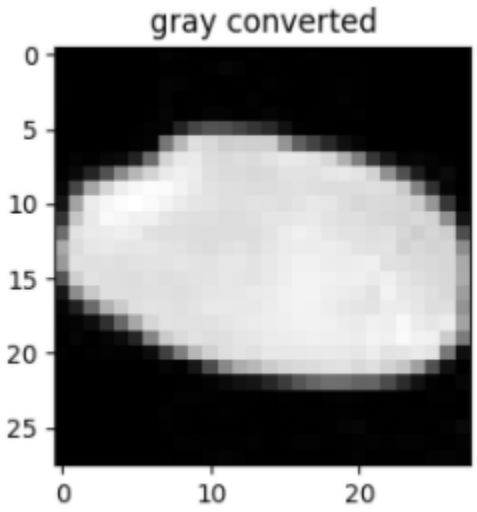
```
train dataset shape is: (377, 28, 28, 3) (377,)
test dataset shape is: (20, 28, 28, 3) (20,)
```

Lets plot random images from our image set.

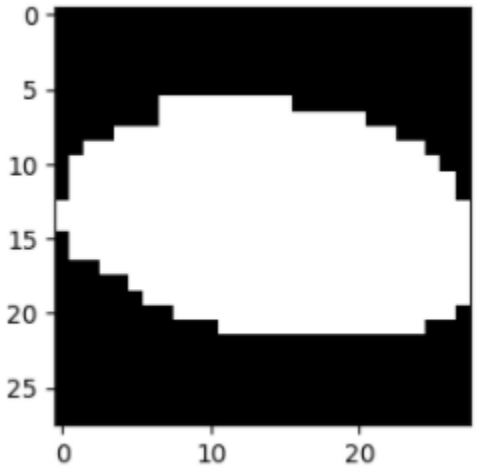


After that we convert single image into 2 another versions such gray conversion and binary conversion.

4.1 Gray conversion



4.2 Binary conversion

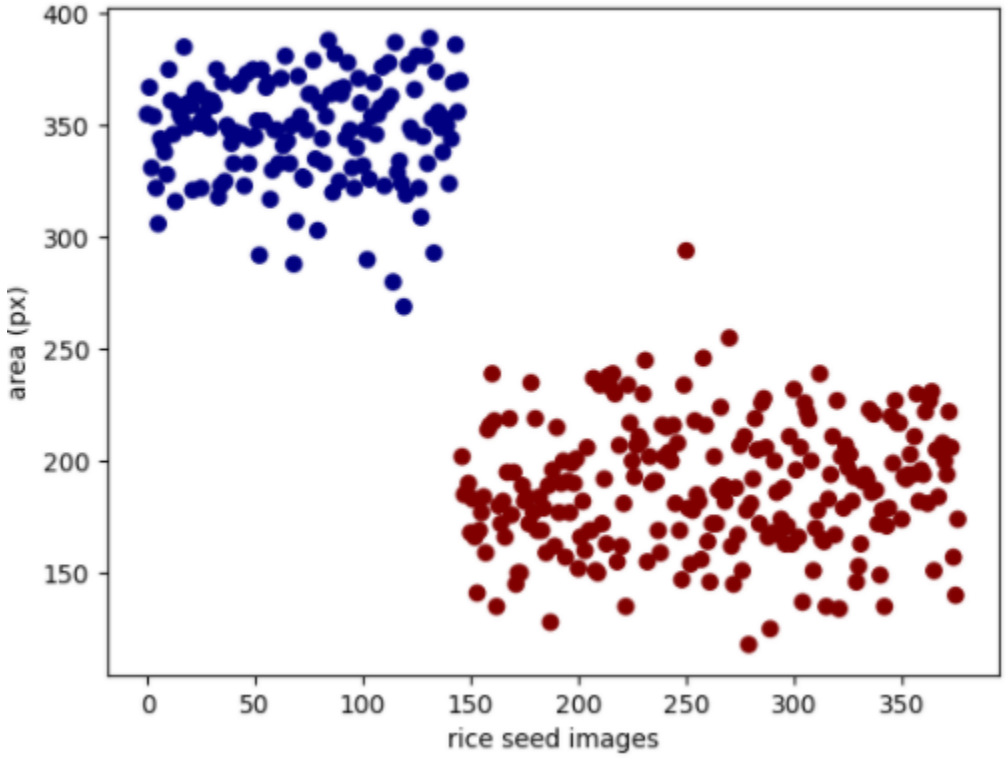


Now that we have a nice binary image, we can isolate the region of rice seed which corresponds to the white region of the image above.Then, this binary image used to calculate the area of white region and find the major and minor axis length of the images.

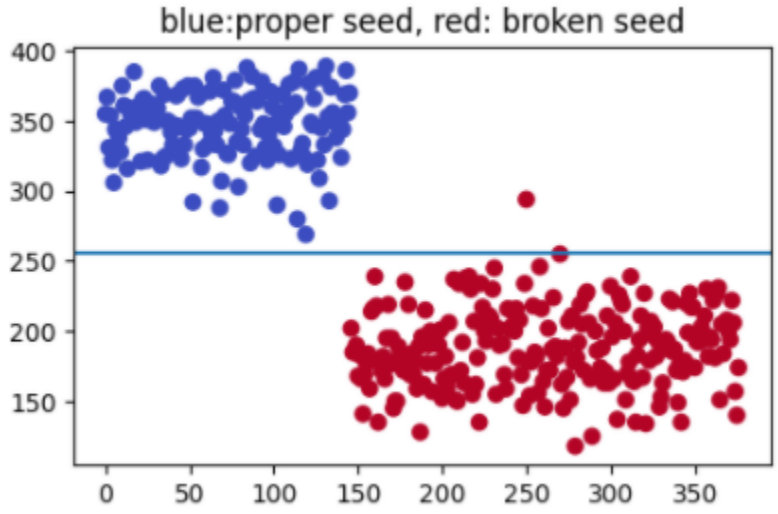
```
label_im, nb_labels = ndimage.label(binary)
regionprops = measure.regionprops(label_im, intensity_image=gray)
regionprop = regionprops[0]

print("area is",regionprop.area)
print("major axis length is", regionprop.major_axis_length)
print("minor axis length is", regionprop.minor_axis_length)
```

This procedure use for every training image and calculate the area of white region and then plot that area values in a same plot.Then we made a threshold area to separate 2 types of images

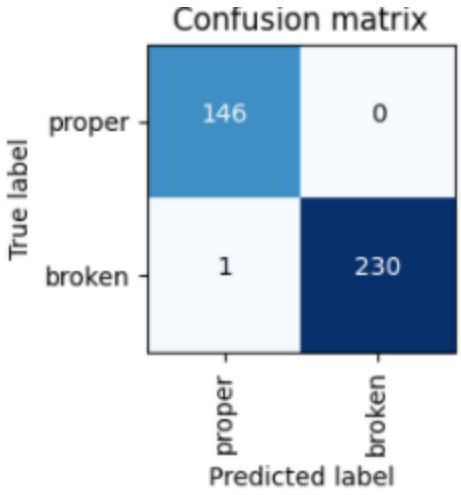
4.3 Threshold Function for Area



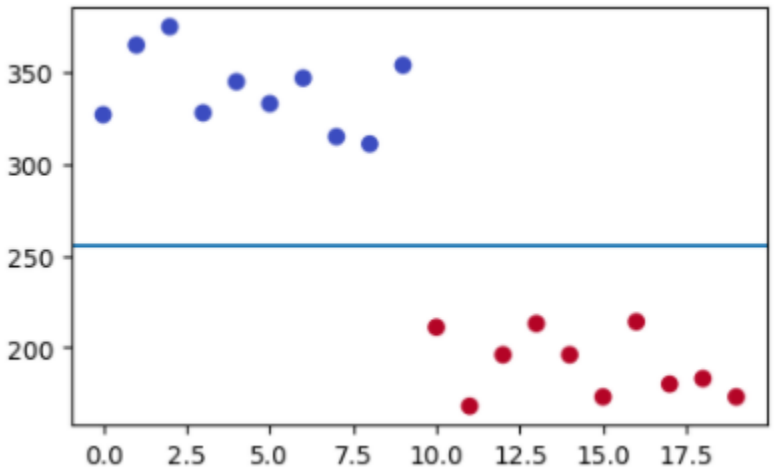
Above figure used to find the threshold value.



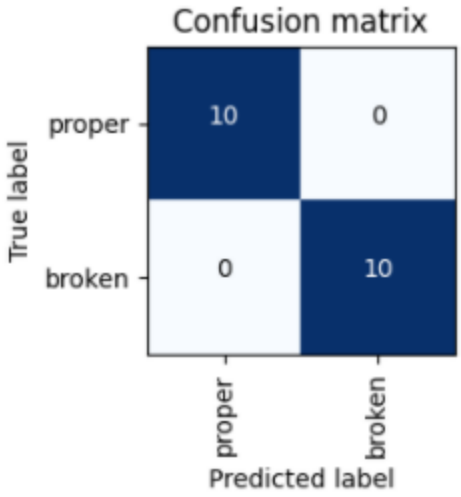
plot a confusion matrix to check correctly and incorrectly classified images.



This threshold value used to evaluate our test set.



Again plot the confusion matrix for testing part.



Here we can see,our threshold value is perfect for classifying the images.Then we can make a classifier to categories the images and it can be used to identify the seed type using their images.