UNIVERSIDADE FEDERAL
DE ALAGOAS

Universidade Federal de Alagoas - UFAL
Instituto de Computação - IC
Curso de Ciência da Computação

Prof. Tiago Vieira

YCbCr Leaf Segmentation with Deep Learning Classification

Eduardo Antônio de Lucena Lisboa

Maceió, 19 de setembro de 2021

Contents

1 Introduction

A common problem in Brazilian crops, as well in crops all around the world, is diseases and pests. All kinds of crops can suffer from such issues and preventing them is a highly sought subject.

Usually it can be hard to diagnose these diseases and pests in advance due to similarities between diseases or the size of the insect. Thus, identifying these diseases and treating them as fast as possible is fundamental to avoid crop losses.

One such plant is the Brazilian Arabica Coffee, which is known worldwide and one of Brazil's most important agricultural commodity, can be affected by diseases and pests and this can have enormous economical implications.

Being able to determine which part of the plant has a disease or a pest can help to identify where to focus the treatment.

2 Objective

2.1 Related Works

In the literature, the YCbCr color space, initially used only in television because it uses less bandwidth to be transmitted and can be perceived just as well as a standard RGB image but with much less color information, has been shown to have excellent potential of increasing the accuracy of image segmentation, being it with faces [1], satellite images [2] and plants, being the latter used for maturity evaluation [3] or disease detection [4].

For this segmentation achieve its full potential, it must be paired with other methods and algorithms, and all these methods rely on a basic strategy that is the use of textural features [5].

Another important strategy commonly used is using the *k-means* clustering algorithm, which has been shown several times to be an efficient algorithm in segmentation problems [6] including plant diseases [7].

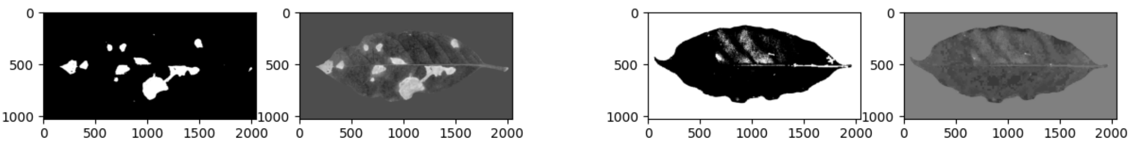
When it comes to the use of Neural Networks, it has had many applications such as identifying weeds, classifying land cover, fruit counting, plant leaf stress detection, etc [8] [9] [10] [11] [12]. And all of them have had extremely positive results [13], sometimes even achieving 100% accuracy rate in some diseases [10].

Combining both techniques seems a logical step in further improving detection, classification and segmentation of diseased and healthy plants, having the potential do greatly benefit agricultural producers all around the world.

2.2 Original Paper

The original paper proposed a method of segmentation of leaf images using the *YCbCr* color space combined with the *K-Means* clustering algorithm to identify the diseased part and highlight it for easier identification.

The problem with this method is that if an image of a healthy leaf goes through this processing, it yields a wrong segmentation which can be confusing and an unnecessary expense of processing power, specially if the amount of images processed is big, as it my be the case in large coffee crops. Example images can be seen on Fig 1.



(a) Correct Segmentation

(b) Wrong Segmentation

Figure 1: Segmentation Examples of Original Code

2.3 Proposed Improvement

To improve this paper, I created a Convolutional Neural Network model to, before an image is processed, predict if the image is a healthy or diseased leaf and only process it if the prediction is of a diseased leaf. Preventing the unnecessary processing and yielding less confuse and more accurate results.

3 Materials and Methods

3.1 The BRACOL Dataset

The models were trained using the BRACOL Dataset [14], which is in the public domain, consisting of 1747 images of size 2048 x 1024 pixels. But due to the corruption of the compressed archive, only 1216 images could be used.

The images in BRACOL are labeled with healthy or diseased and the severity of each disease present, being these diseases *Rust*, *Miner*, *Cercospora* and *Phoma*, and the severity is divided in 5 different levels, being 0 a healthy leaf and 4 a very compromised leaf. All images were captured in similar conditions of lighting and a single leaf per image. The leaves were collected at different times of the year.

3.2 The Models

The CNN model consisted of a *data augmentation* layer followed by a *normalization* layer, then two *2D convolutional* layers and a *2D max pooling* layer. After this come three more *2D convolutional* layers interspersed with three *2D max pooling* layers. Then we have a *dropout* layer, a *flattening* layer and three *dense* layers, the first two with *ReLU* activation and the last with *Softmax* activation.

The models were trained using 80% of the total amount of images in the dataset, being this percentage further divided in 80% for training and 20% for validation.

The hyperparameters adjusted between model trainings were *batch size* and *epochs*, with a fixed *learning rate* of $0{,}001$.

4 Results

4.1 Model

All the trained models were compared using Accuracy, Precision, Recall and F1 scores, then the mean accuracy and standard deviation were calculated using these results. The metrics scores can be seen in Table 1.

Table 1: Accuracy Results of Best Model

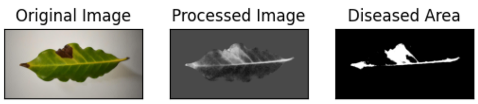
Metrics	
Accuracy	97,18 %
Precision	98,66 %
Recall	98,00 %
F1	98,33 %
Accuracy Mean	98,04 %
Standard Deviation	0,0064

4.2 Combined Results

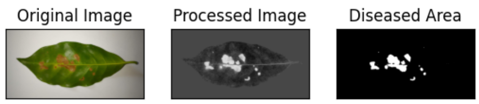
The model was then combined with the original segmentation code. Images considered healthy were ignored and a message is displayed in the terminal, as can be seen on Figure 2c and only the images of considered diseased leaves were processed and the resulting segmentation saved in a PNG file, as can be seen on Figures 2a and 2b.

5 Conclusion

The image segmentation algorithm, who already had good results, allied with the prediction model, which had a satisfactory accuracy, can certainly be used in the field by agricultural workers all over the world, as a computer software, internet website or even a smartphone application. Potentially being of great help in preemptively identifying and enabling the early treatment of pests and diseases, thus avoiding any major economical impacts that these diseases and pests might cause.



(a) Processed Image 2



(b) Processed Image 10

```
Image 1 is a healthy leaf!

Image 10 is a diseased leaf!
Image saved in "Plots" folder.

Image 2 is a diseased leaf!
Image saved in "Plots" folder.
```

(c) Terminal Message

Figure 2: Combined Results

References

[1] K. H. B. Ghazali, J. Ma, R. Xiao, and S. A. lubis, "An innovative face detection based on YCgCr color space," *Physics Procedia*, vol. 25, pp. 2116–2124, 2012. [Online]. Available: https://doi.org/10.1016/j.phpro.2012.03.358

[2] P. Ganesan, V. Rajini, B. S. Sathish, V. Kalist, and S. K. K. Basha, "Satellite image segmentation based on YCbCr color space," *Indian Journal of Science and Technology*, vol. 8, no. 1, p. 35, Jan. 2015. [Online]. Available: https://doi.org/10.17485/ijst/2015/v8i1/51281

[3] R. Dhanesha, C. L. S. Naika, and Y. Kantharaj, "Segmentation of arecanut bunches using YCgCr color model," in *2019 1st International Conference on Advances in Information Technology (ICAIT)*. IEEE, Jul. 2019. [Online]. Available: https://doi.org/10.1109/icait47043.2019.8987431

[4] V. Singh and A. Misra, "Detection of plant leaf diseases using image segmentation and soft computing techniques," *Information Processing in Agriculture*, vol. 4, no. 1, pp. 41–49, Mar. 2017. [Online]. Available: https://doi.org/10.1016/j.inpa.2016.10.005

[5] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973. [Online]. Available: https://doi.org/10.1109/tsmc.1973.4309314

[6] J. Baghel and P. K. Jain, "K-means segmentation method for automatic leaf disease detection," 2016.

[7] T. G. Devi, P. Neelamegam, and A. Srinivasan, "Plant leaf disease detection using k means segmentation," 2018.

[8] J. Liu and X. Wang, "Plant diseases and pests detection based on deep learning: a review," *Plant Methods*, vol. 17, p. 22, 2021.

[9] L. X. B. Sorte, C. T. Ferraz, F. Fambrini, R. dos Reis Goulart, and J. H. Saito, "Coffee leaf disease recognition based on deep learning and texture attributes," *Procedia Computer Science*, vol. 159, pp. 135–144, 2019. [Online]. Available: https://doi.org/10.1016/j.procs.2019.09.168

[10] J.-R. Xiao, P.-C. Chung, H.-Y. Wu, Q.-H. Phan, J.-L. A. Yeh, and M. T.-K. Hou, "Detection of strawberry diseases using a convolutional neural network," *Plants*, vol. 10, no. 1, 2021.

[11] J. G. Esgario, R. A. Krohling, and J. A. Ventura, "Deep learning for classification and severity estimation of coffee leaf biotic stress," *Computers and Electronics in Agriculture*, vol. 169, p. 105162, 2020.

[12] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers in Plant Science*, vol. 7, Sep. 2016. [Online]. Available: https://doi.org/10.3389/fpls.2016.01419

[13] S. K. Noon, M. Amjad, M. A. Qureshi, and A. Mannan, "Use of deep learning techniques for identification of plant leaf stresses: A review," *Sustainable Computing: Informatics and Systems*, vol. 28, p. 100443, 2020.

[14] R. Krohling, J. Esgario, and J. A. Ventura, "Bracol - a brazilian arabica coffee leaf images dataset to identification and quantification of coffee diseases and pests," https://www.doi.org/10.17632/yy2k5y8mxg.1, 2019.