Analysis of Machine Learning Techniques Applied to Coffee Bean Classification

Igor G. Lube[1], Gustavo M. Almeida[1]

[1]*Instituto Federal do Espírito Santo – Campus Serra*
Rodovia ES-010 - Km 6,5 – Manguinhos, 29173-087, Espírito Santo, Brazil
igorgl_@hotmail.com
gmaia@ifes.edu.br

Abstract. The coffee market is characterized by a set of activities of enormous complexity, dynamism, and a growing level of demand from consumers regarding the quality of the drink. This imposes high quality control on producer, consumer and exporter countries. Currently, the definition of the quality and, therefore, the value of coffee is based on manual classification, that is, a person plays the role of a trained (certified) classifier to qualify coffee samples. Thus, the current classification process suffers from the subjectivity of the classifiers and a great difficulty in standardizing the process due to possible inconsistencies in the process. Given this scenario, the present work proposes a comparison between three algorithms that classify coffee samples, considering shape and imperfections. The algorithms are classifiers, one based on MLP (Multi-Layer Perceptron), another in clustering by K-Means and the latter consists of a classifier based on Deep Learning and regional convolutional networks (R-CNN). The objective of this work is to compare which of the algorithms is more effective in classifying the grains according to the intrinsic defects present in the sample.

Keywords: Classification of coffee beans. Deep Learning. Mask R-CNN. K-Means. Multi-Layer Perceptron

1 Introduction

Coffee is the second most consumed beverage in the world according to the International Coffee Organization [1] which also reveals that Brazil has stood out in the production and export of coffee and currently occupies the first position in the world both in production and in the export of the product , which makes the product one of the most important sources of income in the country and in several municipalities, generating more than 8 million jobs throughout Brazil [2].

Due to the high demand of the international market for a quality product, the Brazilian coffee exported needs to undergo a strict quality control, which must be precise and fast, given the large volume of the product that must be evaluated.

One of the main difficulties faced by producers is the evaluation of the quality of their crops, which must be fast and reliable to monitor the volatility of the price of these commodities. The quality of the drink depends, among other factors, on the quality of the beans used and the amount of imperfections and impurities mixed with the good beans. Defects found in coffee samples are counted in order to define the type of coffee.

Currently, the evaluation is carried out visually by specialists and has the following disadvantages: the subjectivity of the process is great as each worker can have a different classification of the same sample, so that different workers obtain different diagnoses to the same sample; it is a slow process, as each grain must be evaluated individually; and the quality of the classification also suffers negative effects during the work shift for reasons inherent to human physiology and environmental conditions such as fatigue, stress, poor lighting, etc. Speed and reliability are extremely dependent on the professional, which makes it difficult to standardize this stage of the production chain, as each human being has its own particularities.

In general, automatic methods appear in machines and systems for application in the production process of seed processors. In this sense, the present work proposes a comparative analysis between three different computer vision techniques applied to the classification of coffee beans: Multi-layer Perceptron, Pizzaia et al [3]; Clustering by K-Means, Vasconcellos et al [4] and finally, regional convolutional neural networks through the Mask R-CNN network, Lube and Almeida [5].

2 Related Works

Although the problem has already been explored by groups that developed related products and patents, that is, classification of soybeans, corn, rice, barley, among others, there are not many academic works dealing with classification of coffee beans with computer vision in the literature.

In the work of Carrillo and Penaloza [6], classifiers based on the Mahalanobis distance were used to identify 6 classes of coffee beans in a sample, however, the average recall obtained in this method was only 67.7%.

Ahmad et al. [7] presented a machine to perform automatic separation of coffee beans by computer vision, however, they only separated the beans without classifying them, in addition to having an accuracy of only 78.32%.

Pizzaia et al. [3] propose a Multilayer Perceptron Neural Network (MLP) to classify coffee beans by sample shape, size and color. The images used contained good and bad coffee beans previously classified by experts. The grains were segmented by Otsu binarization [9] and then the areas, rounding and the average of the color values of each grain in the RGB layers were calculated. These characteristics were fed into the MLP network which classifies the beans as good or bad. The network achieved an average of 94% accuracy.

In the work by Vasconcellos et al. [4] a system using K-Means clustering is proposed to classify four types of coffee beans, namely the good, broca, parchment and black beans. The samples were treated with the White-Patch luminosity standardization algorithm. The area, rounding and average of the RGB channels of each grain were used in the classifier, obtaining a correctness rate of 90.74%.

The approach proposed by Lube and Almeida [5] uses Deep Learning in the form of a Convolutional Neural Network (R-CNN) applied in the Mask R-CNN technique proposed by He et al [8]. In this work, the images contained samples of coffee beans that were separated into 5 classes: good, broca, parchment, black and bad, the latter being a grouping of several defects not covered in the aforementioned classes. The technique achieved an average accuracy of 93.94% and an F-score of 82.24% for the five classes covered.

3 Theoretical foundation

3.1 This section describes in more detail the three techniques analyzed in this article, covering their differences and similarities.

3.2 Classification by a Multi-Layer Perceptron Network

The classification method proposed by Pizzaia et al. [3] uses an MLP to classify coffee beans. The grid uses as input parameters the color, area and roundness of each grain. The images used for training were composed of samples previously separated by experts that contained good grains and grains with various impurities and non-discriminated defects. For this reason, the classes covered in the classification were “good grains” and “bad grains”. The images went through a segmentation process by Otsu binarization [9], separating the grains from the background, the grain contours were obtained and the Regions of Interest were highlighted, characterizing the process as a supervised learning, thus the parameters color, area and roundness were calculated for each sample contained in the images. Figure 1 shows the flowchart of the developed algorithm.

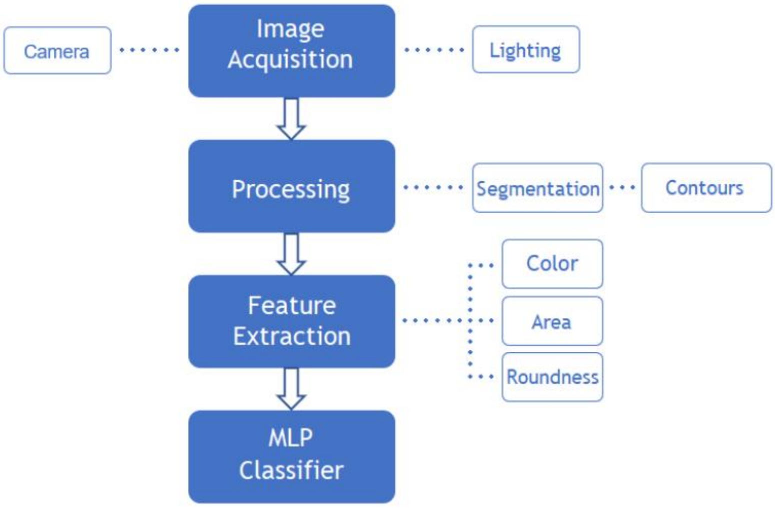

Figure 1. Flowchart of the MLP Classifier

The Area parameter is calculated by the number of pixels in each sample. Rounding defines how close each sample is to a perfect circle, in this way, the network can differentiate an intact grain from a broken grain, among other defects. Rounding is obtained through the equation:

$$\text{Rounding} = \frac{4\,.\,\pi\,.\,\text{area}}{perimeter^2}$$

The color parameter calculates the average of the RGB channels. These five parameters are the inputs used in the MLP classifier to obtain the "good grain" and "bad grain" classes.

The classifier was exposed to 848 samples, 748 for training and 100 samples for validation. The network has five neurons in the input layer (one for each parameter), a hidden layer with 100 sigmoid neurons using the Levenberg-Marquardt algorithm and an output layer containing 1 neuron with binary output being "0" for a good grain and "1" for a bad grain. In Figure 2 we have examples of a good grain (Fig. 2a) and a bad grain (Fig. 2b).

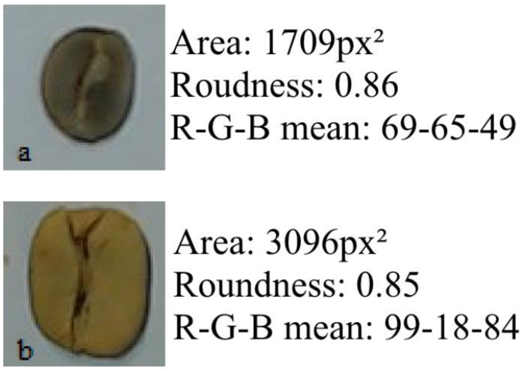

Figure 2. Examples of a good grain (a) and bad grain (b)

3.3 Classification by K-Means Clustering

In the classification proposed by Vasconcellos et al. [4] the method used was clustering using K-Means, Lloyd [10] which is an unsupervised learning method that occurs by creating k groups of previously known elements. Each Cluster encompasses data with similar characteristics that are mathematically verified using the Euclidean distance. This work is another classifier approach that uses the same database, but breaking down the bad grains into three new classes which are: black grain, broca grain and parchment grain. The "good grain" class was maintained.

The images with the samples were processed using the White-Patch algorithm, which aims to improve the brightness and standardize the treatment of the database.

The characteristics used in the training were: Area, Rounding and average of the RGB channels and their values obtained in the same way as the work by Pizzaia et al. [3]. Figure 3 exemplifies an image after treatment and ready to extract the characteristics that will be used in the network.

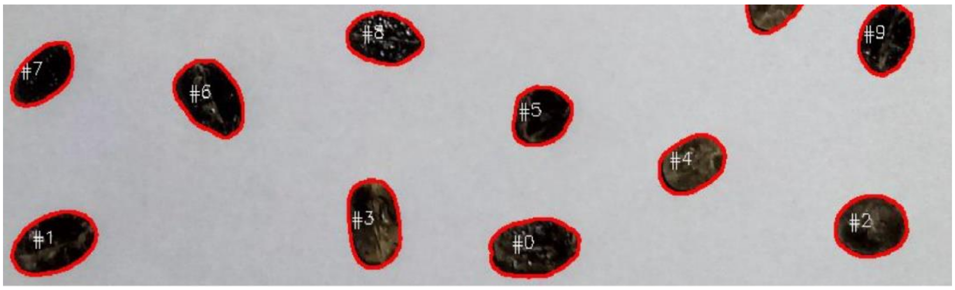

Figure 3. Contoured and numbered grains

3.4 Classification by Deep Learning using Mask R-CNN framework

The network proposed by Lube and Almeida [5] uses as a basis the Mask R-CNN network that was defined by He et al. [8] as a simple and flexible framework for instance segmentation, the Mask R-CNN is based on a framework proposed by Ren et al. [11] called Faster R-CNN, a system for object detection and semantic segmentation, which is fast, robust and intuitive.

Mask R-CNN adds a new mask prediction branch that works parallel to Faster R-CNN's existing bounding box classification and regression branch. This new branch is a small FCN (Full Connection Network) applied pixel by pixel in each region of interest (RoI) and allied to this FCN there is a layer called RoIAlign with the function of improving the precision of the created masks and decouple the branch from the masks from the class prediction branch, ensuring a 10% to 50% improvement in mask accuracy. In Figure 4, the framework of the Mask R-CNN architecture is shown.

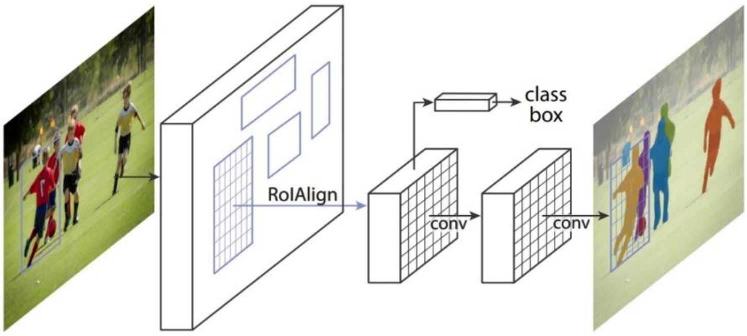

Figure 4. Mask R-CNN framework

The coffee beans database was assembled from the images used in the works cited in sections 3.2 and 3.3. Samples contain good grains and defective grains (broca, parchment, black and bad). The image treatment was by segmentation, in which an image with several grains was subdivided into smaller images with only one grain and with a dark background. Due to the low amount of samples needed to learn with Deep Learning, it was necessary to use the Data Augmentation process that applies rotations, mirroring, among other strategies to increase the database and, consequently, the network's accuracy.

Figure 5. Grains before (left) and after (right) segmentation

The algorithm of this work received the images with annotations for training, but the database is insufficient to train a network from scratch. To solve this impasse, the technique known as fine tuning was used, which is a transfer learning of characteristics from a network trained with a larger database to a smaller one.

The VGG Image Annotator software (VIA tool) was used to produce the Regions of Interest (RoI), which contain the annotations with the location of the grains and their respective classes, which are input parameters of the network together with the images with the samples.

The proposed network was trained by transfer of learning from a pre-trained network based on the Microsoft COCO database: Common Objects on Context, Lin et al [12]. With fine tuning it is possible to have a network with good results even with a reasonably small database. The net was trained with 606 samples of segmented grains, during 45 seasons, with 100 steps per season and a minimum confidence of 60%. The learning rate was set at 0.001.

3.5 Grain Classes

Grain classes have characteristics that differentiate them by attributes such as color, appearance and uniformity as shown in Tab 1.

Table 1. Characteristics of coffee beans

Class	Characteristics
Good	Color and appearance uniformity and free from significant imperfections
Bad	Broken, non-uniform, small grains and various imperfections
Broca	dark canals due to the presence of the coffee berry borer in the planting
Black	Blackish color due to high level of undesirable fermentation
Parchment	Yellowish parchment surrounding the grain that has not been completely removed

The MLP network was trained with the classes "Good" and "Bad", the latter being a composition of the classes "Broca", "Dark" and "Marinheiro", since for this work, there was no separate sample in terms type-specific grains, such as drill, dark, and sailor. The classification network by clustering used for training the “bad” class broken down into the other three mentioned above and, finally, the Mask R-CNN network was trained with all classes present in Tab. 1.

4 Results and discussions

In this section, the results of each classifier and the comparative analysis of their results will be presented. Table 2 shows the performance metrics for each classifier type.

Table 2. Comparison of classifier performance (%)

	ACCURACY			PRECISION			RECALL			F-SCORE		
	MLP	K-M	Mask	MLP	K-M	Mask	MLP	K-M	Mask	MLP	K-M	Mask
Good	94,10	94,44	93,13	95,12	94,32	83,33	92,85	95,40	58,82	93,97	94,86	68,97
Bad	94,10	-	91,60	93,18	-	68,18	95,35	-	78,95	94,25	-	73,17
Broca	-	90,12	96,18	-	73,17	100	-	85,71	86,49	-	78,95	92,75
Black	-	100,0	94,66	-	100,0	87,72	-	100,0	100	-	100,0	93,46
Parchment	-	95,68	98,47	-	100,0	87,50	-	66,67	87,50	-	80,00	87,50
Average	94,10	95,06	94,81	94,15	91,87	85,35	94,10	86,67	82,35	94,11	88,45	83,17

The MLP classifier has an average value greater than 94% in all metrics, but with classes limited to "good" and "bad" grains which makes the direct comparison imprecise when related to other classifiers, however for the simpler task of just selecting positive samples, that is, good grains, this proposal is effective and has a low computational cost, but requires steps to calculate the input parameters (color, rounding and area).

K-Means clustering stands out for being a technique with unsupervised learning and also needs five input parameters for grain classification. The authors of the article mention in their conclusion the low efficiency in classifying “broca” grains, which can be confirmed in the precision and recall metrics present in Tab. 2. The “black” grain metrics are the highlight of this technique and due to the subdivision of the "bad" class the performance values of this class are not shown in the table.

As the R-CNN Mask is a Deep Learning algorithm, it needs a sample bank with large volume, the “bad” class was harmed by this aspect, as several samples were taken from it that gained their own classes (broca, black and parchment) and the rest of the samples as small, broken and misshapen grains did not have a sufficient volume for training the net. The main highlight were the brocaded grains that presented values above 95% in accuracy and precision, which demonstrates the ability of the network to classify these grains correctly even with a great similarity in relation to the so-called "good" grains.

5 Conclusions

This work aimed to demonstrate through performance metrics and analysis of computer vision techniques presented that there is no classifier that is superior to others in all situations. The MLP classifier is very effective in identifying good grains using a simple and low computational cost network. In some situations this type of classifier already produces sufficient performance to meet the desired goals. K-Means clustering is satisfactory when the task requires the classification of grains with very distinct characteristics between classes, as the technique groups similar ones into clusters, making it difficult to differentiate similar samples, such as good and broca grains.

The Mask R-CNN network requires a large-volume database, but is capable of distinguishing subtle differences between samples if it is trained with a sufficient number of images. This classifier is the most generalist and scalable among the three studied and also the most promising, as its deep learning allows to increase the number of classes without changing the input parameters, just adding more samples and their annotations. Impurities such as "sticks", "stones" and "clods" that are also part of the classification of coffee beans, could be easily added to the classifier and would obtain promising results, which cannot be said of other techniques that use color, rounding and area for classification task.

As future work, the authors are seeking partnership with cooperatives in the State of ES, in order to obtain samples of all types of defects in coffee beans, in addition to what we already have, so that a classifier with the capacity to identify all types of coffee beans and also increase the metrics presented by Mask-RCNN, as it is known that the greater the volume of images in the dataset, the more efficient the network becomes.

.**Authorship statement.** The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

[1] International Cofee Organization, year = 2019, url = http://www.ico.org/prices/new-consumption-table.pdf, urldate = 2020-08-10.
[2] Ministry of Agriculture, Livestock and Supply, Brazil, year = 2019, url = https://www.gov.br/agricultura/pt-br/assuntos/politica-agricola/cafe/cafeicultura-brasileira, urldate = 2020-08-05.
[3] J. P. L. Pizzaia, I. R. Salcides, G. M. Almeida, R. Contarato, R. Almeida. Arabica coffee samples classification using a Multilayer Perceptron neural network. In 13th IEEE International Conference on Industry Applications (INDUSCON), pp. 80-84, 2018.
[4] B. M. Vasconcellos, G. M. Almeida, R. B. Contarato, R. Almeida. Classificação de amostras de café Usando visão Computacional. In Simpósio Brasileiro de Automação Inteligente, pp. 2505-2508, 2019.
[5] I. G. Lube, G. M. Almeida. Classification of coffee beans using Deep Learning. In: XLI CILAMCE, 2020, Foz do iguaçu.
[6] E. Carrillo and A.A. Penaloza. "Artifical Vision to assure coffee-Excelso beans quality". Proc. Euro American Conference on Telematics and Information System, 2009.
[7] U. Ahmad et. al. "The development of automatic coffee sorting sytem based on image processing and artificial neural networks". 2010.
[8] K. He, G. Gkioxari, P. Dollar and R. Girshick. Mask r-cnn. In Proceedings of the IEEE international conference on computer vision, pp. 2961–2969, 2017.
[9] N. Otsu, "A threshld selection method from gray-level histograms". Automatica, 11(285-296), pp. 23-27, 1975.
[10] S.P., Lloyd, "Least squares quantization in PCM". Information Theory, IEEE Transactions on 28.2, 129-137,1982.
[11] S. Ren, K. He, R. Girshick and J. Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems, pp. 91–99, 2015.
[12] T. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, P. Dollar and C. L. Zitnick. Microsoft coco: Common objects in context. In European conference on computer vision, pp. 740–755, 2014.