Journal of Food Composition and Analysis 97 (2021) 103771

Contents lists available at ScienceDirect

Journal of Food Composition and Analysis

journal homepage: www.elsevier.com/locate/jfca

Original Research Article

Classification of fermented cocoa beans (cut test) using computer vision

Marciano M. Oliveira [a], Breno V. Cerqueira [a], Sylvio Barbon Jr. [b], Douglas F. Barbin [a,*]

[a] Department of Food Engineering, School of Food Engineering, University of Campinas, Campinas, SP, Brazil
[b] Department of Computer Science, State University of Londrina (UEL), Brazil

ARTICLE INFO

Keywords:
Chocolate
Cut-test
Food quality
Analytical method
Image analysis
Random decision forest

ABSTRACT

Fermentation of cocoa beans is a critical step for chocolate manufacturing, since fermentation influences the development of flavour, affecting components such as free amino acids, peptides and sugars. The degree of fermentation is determined by visual inspection of changes in the internal colour and texture of beans, through the cut-test. Although considered standard for evaluation of fermentation in cocoa beans, this method is time consuming and relies on specialized personnel. Therefore, this study aims to classify fermented cocoa beans using computer vision as a fast and accurate method. Imaging and image analysis provides hand-crafted features computed from the beans, that were used as predictors in random decision forests to classify the samples. A total of 1800 beans were classified into four grades of fermentation. Concerning all image features, 0.93 of accuracy was obtained for validation of unbalanced dataset, with precision of 0.85, recall of 0.81. Although the unbalanced dataset represents actual variation of fermentation, the method was tested for a balanced dataset, to investigate the influence of a smaller number of samples per class, obtaining 0.92, 0.92 and 0.90 for accuracy, precision and recall, respectively. The technique can evolve into an industrial application with a proper integration framework, substituting the traditional method to classify fermented cocoa beans.

1. Introduction

Cocoa beans are fermented and dried seeds of the *Theobroma cacao* tree and are widely consumed all over the world. The annual cocoa production worldwide is 4.2 million tons valued at $11.8 billion and has grown at a rate of 3% per year in the past decade (FAOSTAT, 2020; Beg et al., 2017). Cocoa beans are processed to obtain chocolate liquor, cocoa powder and cocoa butter, which are the main ingredients of chocolate and a vast range of products like cocoa beverages, ice cream and bakery products (Beg et al., 2017; Afoakwa et al., 2008).

However, before cocoa beans can be traded and processed into final industrial products, they must undergo post-harvest processing on farms and plantations. Post-harvest steps include opening the pod and removing cocoa beans, which will be subsequently fermented and dried (Kumari et al., 2018; Moreira et al., 2018; Castro-Alayo et al., 2019). Fermentation is a critical step for the development of flavour of commercial cocoa beans, affecting components such as free amino acids, peptides and sugars (Castro-Alayo et al., 2019; Moreira et al., 2018; Kadow et al., 2013; Amoa-Awua, 2015).

During fermentation and subsequent drying process, the polyphenol content (including anthocyanidins) decrease as some permeate out of the beans while others are oxidised and polymerise to insoluble high-molecular-weight compounds (tannins) (Castro-Alayo et al., 2019; Afoakwa et al., 2013). Also, the accumulation of acids during fermentation reduces the pH and is decisive for inactivation of the embryo, leading to enzymatic reactions that originate flavour precursors. This leads to changes in the internal colour of the dried cocoa beans, from dark grey (slaty) of the unfermented bean through the deep-purple colour of under fermented beans to the brown colour of the fully fermented bean (Castro-Alayo et al., 2019). Also, as fermentation progresses, ridges appear in the embryo surface and become larger and deeper. Therefore, it is important to identify these changes to determine if cocoa beans are properly fermented, originating samples of high quality, and lacking defects (Afoakwa et al., 2013; Kongor et al., 2016).

The change in cocoa beans colour during the fermentation is exploited in a simple method for determining the degree of fermentation, called cut-test. The cut-test is the simplest and still the most widely used method to assess the quality of a random sample of beans from a batch by visual inspection (Kongor et al., 2013). A regular cut-test consists in the evaluation of 300 beans, disposed in an inspection board (usually comprising spots for 100 beans), that are analysed individually (CAOBISCO/ECA/FCC, 2005; ISO 2451:2017, 2020). Despite

* Corresponding author.
E-mail address: dfbarbin@unicamp.br (D.F. Barbin).

https://doi.org/10.1016/j.jfca.2020.103771
Received 1 July 2020; Received in revised form 20 November 2020; Accepted 12 December 2020
Available online 15 December 2020

being simple and performed by trained personnel, this method is arguably accurate, since it is laborious, time-consuming and considered subjective and difficult to standardise (Kongor et al., 2013). Therefore, the industry urgently demands a rapid and accurate technique that can grade fermented cocoa with low cost and widely standardized.

Computer vision is a valuable technique that has been extensively applied for food quality assessment, including meat and meat products (Barbin et al., 2016; Barbon et al., 2016; Nolasco-Perez et al., 2019), fruits and vegetables (Pereira et al., 2018), nuts (Mathanker et al., 2011), as well as grains (Szczypiński and Zapotoczny, 2012). The combination of image analyses with multivariate statistics has become a powerful tool to deal with several problems in the food sector.

Recently, several authors have applied machine learning algorithms to determine the quality of cocoa beans using different analytical methods (León-Roque et al., 2016; Lawi, and Adhitya, 2018; Parra et al., 2018a, b; Barbin et al., 2018; Jimenez et al., 2018; Mite-Baidal et al., 2019). Among these methods, Random Forest (RF) is an ensemble learning method that provides similar or better prediction efficiency when compared to other classification algorithms, such as Support Vector Machines (SVM), C4.5, AdaBoost (AB), k-Nearest Neighbours (KNN), Logistic Regression (LR), Stochastic Gradient Boosting Trees (GBDT), Extreme Learning Machines (ELM), Sparse Representation-based Classification (SRC), and Deep Learning (DL) (Breiman, 2001; Zhang et al., 2017). Barbon et al. (2016) showed that RF was able to determine the storage time of pork, with prediction results (94.4 % accuracy) superior to those obtained by artificial neural networks (ANN), fuzzy-based classifiers, SVM, k-NN, and decision trees. In another study, Pereira et al. (2018) used RGB images and RF to predict the ripening of papaya, obtaining 94.7 % accuracy in the prediction.

Therefore, this work aims to use computer vision as a fast and accurate method to classify cocoa beans in four grades of fermentation, using features extracted from cocoa bean images as predictors. The proposed approach could substitute the cut-test, using digital RGB imaging and Random Forest to evaluate the quality of fermented cocoa beans.

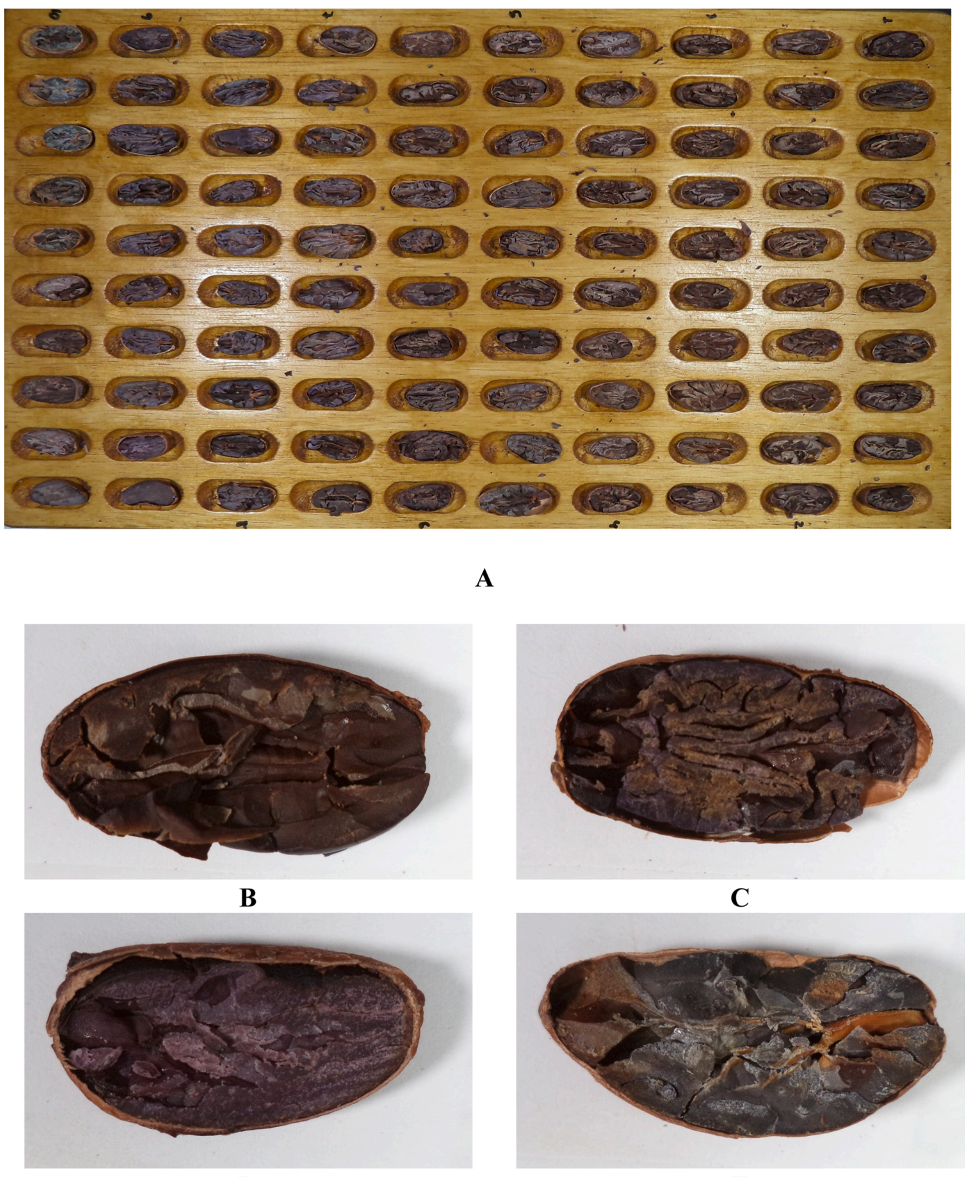

Fig. 1. Representative images of samples: A) Board with samples for cut-test; B) fully fermented sample; C) partially fermented sample; D) under-fermented sample; E) unfermented / slaty sample.

2. Materials and methods

2.1. Classification of cocoa bean samples

Fermented cocoa bean samples were collected from three different regions of Brazil (Bahia, Espírito Santo and Pará). Samples (n = 1800) of cocoa beans were cut lengthwise in two parts to expose the cotyledon. Each individual sample was analysed and classified regarding its colour and texture of the exposed surface, as in the cut-test (Fig. 1A), according to international standards adapted from Guehi et al., (2007), Yro et al. (2018), and Santos et al. (2019). Cocoa beans were classified as: fully fermented (ridged, brown - class 1; 1200 samples) (Fig. 1B), partially fermented (partially brown – class 2; 300 samples) (Fig. 1C), under fermented (purple – class 3; 200 samples) (Fig. 1D), unfermented (smoothed/slaty – class 4, 100 samples) (Fig. 1E). It is important to emphasize that partially fermented samples are not considered a commercial defect (CAOBISCO/ECA/FCC, 2005).

2.2. Determination of colour, pH, titratable acidity (TA)

A representative number of samples from each class were ground in a multipurpose mill (TE-631/4; Brazil) for 30 s at 10,000 rpm. Ground samples were placed in Petri dishes immediately after grinding and the colour was measured directly on the exposed surface using a Minolta colorimeter (Spectrophotometer CM-600d, Japan) after calibration with a standard ceramic tile. Colour was expressed as values for lightness (L*), redness (a*), and yellowness (b*) using the CIEL*a*b* colour system (Commission Internationale de l'Eclairage - CIE, 1978). In addition, chroma and hue angle was calculated to support the interpretation of variation among samples. Chroma (C*) was calculated as Eq. 1, while the hue angle was calculated according to Eq. 2:

$$C^* = (a^{*2} + b^{*2})^{1/2} \tag{1}$$

$$\text{Hue angle} = \arctan\,(b^*/a^*) \tag{2}$$

Subsequently, 5 g of ground cotyledons were homogenized in a centrifuge (Sorvall ST 8R Centrifuge, Germany) in 45 mL of boiling water for 300 s at 5000 rpm. The solution was filtered, and the pH measured using a pH meter (MB-10, Brazil). After, 25 mL of the solution was used to determine the titratable acidity with 0.1 N NaOH solution (Nazaruddin et al., 2006). All measurements were performed in triplicate and the data were analysed using the software Minitab (Minitab Inc., USA).

2.3. Image acquisition

Images were acquired immediately after visual analysis to avoid oxidation of the cocoa beans and changes in colour. The computer vision system consists of three main components: a digital camera (Sony DSC-H55, Brazil), an illumination source consisting of LED lamps (Natural daylight, 100 W), and a box with matte black internal walls to avoid specular reflections (Pereira et al., 2018). The camera lens and lighting source axis are positioned at an angle of approximately 45°. Images were acquired (.jpeg) using the following camera settings: 25 mm wide-angle lens, a setup of 3648 × 2376 pixels and spatial resolution of 72 ppi, with the exposure time of 0.1 s and f/4.5 opening (zoom and flash functions off).

2.4. Computer vision classification

In our work we built a Computer Vision Classification system based on three different steps (Fig. 2): a) image segmentation, b) image feature extraction and selection, and c) image classification.

2.4.1. Image segmentation

Initially, the images were processed to identify the region of interest (ROI), i.e., remove background and objects that are not part of the cocoa beans. Segmentation was performed using Otsu method (Pereira et al., 2018). A series of eroding and filling operations were performed to obtain a representative image of the cocoa bean after background removal. Image processing operations were carried out using Matlab software (Mathworks, USA).

2.4.2. Image feature extraction and selection

Images of cocoa beans were acquired in the RGB colour space. However, RGB colour space has some disadvantages, as it is uneven and the proximity of the colours does not indicate similarity between them. As a result, colour space transformations are effective to provide useful information. In this study, the transformation to the HSV colour space (hue, saturation and value) was used (Ohta et al., 1980).

Since HS-based spaces are developed based on the concept of visual perception of human eyes, their colour measurements are easy to use and have a good relationship with the visual significance of food surfaces, as confirmed by previous studies (Du and Sun, 2005).

Identification of the most important features reduces data to be analysed and computing costs, sometimes with better classification performances by avoiding useless information (Zheng et al., 2006). Selecting the most important features reduces the computational load for the model, in addition to avoiding unnecessary data with negligible relative importance. In our work, this selection was based on the relative importance of each feature given by the random forest itself (Fig. 3). After converting the RGB image to HSV, we obtained the 99 most relevant histogram bins from the H channel. In addition, 29 traditional image features (Table 1) were calculated for each channel (HSV), a total of 87 features, and these were also used to compose the predictors as in previous studies (Campos et al., 2019; Lopes et al., 2019). Also, the relative energy was extracted from each quarter of the H channel histogram, delivering 4 additional features. Our complete feature vector summary is shown in Fig. 3.

2.4.3. Sample classification

In the context of Computer Vision Classification, a machine learning algorithm is built to create a model to classify unknown images based on a given feature vector. In our current project, we choose Random Forest (RF) due to several advantages, such as improving regression and classification capabilities, reducing computational cost to train and predict, needing few hyperparameters for tuning, delivering an internal error estimate of generalization, dealing fairly well with large-scale problems

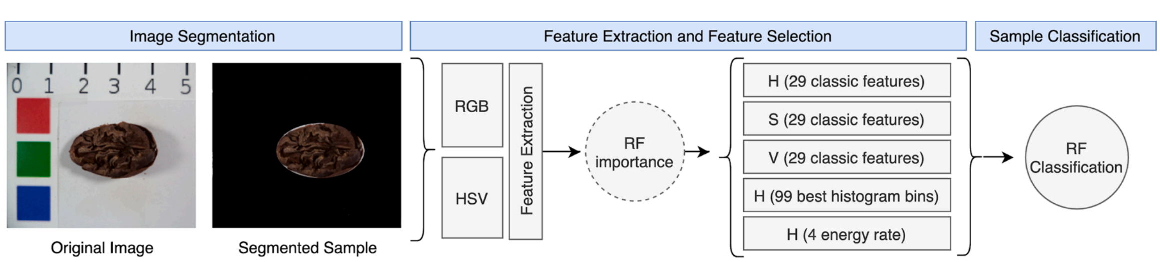

Fig. 2. Overview of proposed Computer Vision Classification, from image segmentation to sample classification.

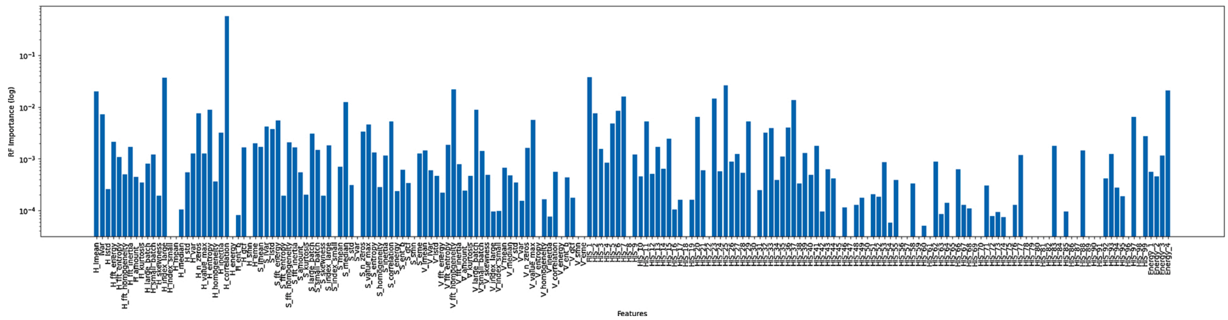

Fig. 3. Relative importance of each feature given by the random forest.

Table 1
Classic image features used in our experiments. The final feature vector was composed by adding some H histogram information, including the energy of each histogram quartile.

No.	Name	Description
1	imean	Mean value of channel colour value
2	Ivar	Variance value of channel colour value
3	Istd	Standard deviation of channel colour value
4	fft_energy	Energy extracted from Fourier Domain
5	fft_entropy	Entropy extracted from Fourier Domain
6	fft_homogeneity	Homogeneity extracted from Fourier Domain
7	fft_inertia	Inertia extracted from Fourier Domain
8	amount_ group	Intensity feature based on the histogram: Number of batches of levels with at least 1 pixel on the histogram.
9	kurtosis	Kurtosis extracted from channel histogram
10	large_batch	Intensity feature based on the histogram: The largest batch of levels on the histogram.
11	small_batch	Intensity feature based on the histogram: The smallest batch of levels on the histogram.
12	skewness	Skewness from channel histogram
13	index_large	Intensity feature based on the histogram: Index of the largest batch of levels on the histogram.
14	index_small	Intensity feature based on the histogram: Index of the smallest batch of levels on the histogram.
15	mean	Mean of current channel intensity
16	median	Median value of current channel intensity
17	std	Intensity standard deviation of the current channel
18	var	Variance of current channel intensity
19	n_zeros	Intensity feature based on the histogram: Number of non-zero levels in the histogram.
20	value_max	Intensity of max value of the current channel intensity
21	entropy	Entropy of the current channel
22	homogeneity	Homogeneity of the current channel
23	inertia	Inertia of the current channel
24	correlation	Correlation of current channel
25	energy	Energy of current channel
26	ent_b	Entropy of the grayscale image.
27	gcf	Global Contrast Factor
28	smn	Statistical Naturalness measure
29	eme	Measure of Enhancement by Entropy

and facilitating implementation (Breiman, 2001). In addition, statistically, the random forest offers additional resources, such as measures of variable importance, called RF Importance, differential class weighting, missing value imputation, outlier visualization and detection, and unsupervised learning (Zhang and Ma, 2012; Belgiu and Drăguţ, 2016). We adjusted our RF using 5000 trees and a standard value of $m = \sqrt{p}$ were used for the number of features randomly sampled as candidates in each partition.

Due to the unbalanced nature of the data set, only overall accuracy may not provide sufficient information about the classifier's discrimination capacity regarding a specific class of samples, as it could be favouring a class with a higher probability of occurrence (majority) over another with a low probability (minority) of occurrence. Hence, the model was later applied to a balanced dataset, with the same number of samples for each class.

The samples were separated in calibration (or training set, 60 % of the images), and validation (or independent test set, 40 % of images) set, using Kennard Stone algorithm (Kennard and Stone, 1969). The calibration set was used to induce the models to adjust the hyperparameters (30 iterations), while the set of prediction tests was used to assess the performance of the classification.

2.4.4. Statistical analysis

The performance of the RF approach was evaluated using performance metrics-classification accuracy, precision, recall and F-measure (Powers, 2011; Farid et al., 2014; Cuadros-Rodríguez et al., 2016), using Python software. Accuracy is the ratio of correctly predicted observation of the total observations. It is an important figure of merit for symmetric data sets where values of false positive and false negatives are almost the same. Precision is the ratio of correctly predicted positive observations of the total predicted positive observations. Recall, also called sensitivity, is the ratio of correctly predicted positive observations to all observations in the respective actual class. F-Score is the weighted average of Precision and Recall. Therefore, this score takes both false positives and false negatives into account. F-score is usually more useful than accuracy for models that have an unbalanced class distribution. Accuracy is representative if a given model has false positives and false negatives with a similar cost. If the cost of false positives and false negatives are very different, it is preferable to consider both Precision and Recall (Fawcett, 2006; Ballabio, and Todeschini, 2009; Labatut, and Cherif, 2011).

3. Results and discussion

This section supports our discussion on the application of a computer vision classification system to assess cocoa samples by image. At first, we discuss physicochemical results for all classes of fermented samples. After, we compare samples regarding their colour and provide further insights about how to discriminate each class by colour variation between each class. Finally, we discuss the results of sample classification using features extracted from images, performance measurements and some important issues for industrial applications.

3.1. pH and titratable acidity (TA)

Fully fermented beans were significantly different ($p < 0.05$) in pH and TA from partially fermented and unfermented beans (Table 2). Samples from classes 2 and 3 (partially fermented and under fermented) were not significantly different. Fully fermented beans (class 1) presented the lowest values for pH and highest for TA. In contrast, unfermented cocoa beans (class 4) had the highest pH and lowest TA values. It is expected that the concentration of acids will increase as the fermentation proceeds, directly impacting the pH and TA values. The pH reduction and consequently increase in TA during fermentation are due to the diffusion of acids (predominantly acetic acid) through the cotyledon (Afoakwa et al., 2008, 2013). It is important that beans achieve

Table 2
CIEL*a*b* values (average and standard deviation) for the four classes of fermented cocoa beans.

Class / Sample size	L* (D65)	a* (D65)	b* (D65)	C* (D65)	Hue angle (°) (D65)	pH	TA
1 / 1200	32.63 ± 0.62[a]	14.96 ± 0.18[a]	17.38 ± 0.34[a]	22.94 ± 0.37[a]	49.28 ± 0.32[a]	5.17 ± 0.08[a]	0.054 ± 0.002[a]
2 / 300	32.62 ± 1.09[a]	12.85 ± 0.27[b]	9.27 ± 0.22[c]	15.85 ± 0.33[b]	35.80 ± 0.42[b]	5.77 ± 0.06[b]	0.039 ± 0.003[b]
3 / 200	36.47 ± 0.95[b]	11.91 ± 0.22[c]	3.81 ± 0.31[d]	12.51 ± 0.26[c]	17.74 ± 1.27[c]	5.92 ± 0.14[b]	0.034 ± 0.002[b]
4 / 100	44.94 ± 0.95[c]	7.29 ± 0.16[d]	14.69 ± 0.39[b]	16.40 ± 0.41[d]	63.58 ± 0.34[d]	6.48 ± 0.03[c]	0.022 ± 0.002[c]

Values are expressed as average ± standard deviation. Values in the same column with different letters are significantly different ($P < 0.05$). Class 1: fully fermented; Class 2: partially fermented; Class 3: under fermented; Class 4: unfermented.

the pH range between 5.0–5.5 during fermentation, to obtain fermented and dried cocoa beans with a high level of aromatic compounds. On the other hand, a pH range of 4.0–4.5 may indicate inferior cocoa beans and possibly results in low values of aromatic compounds (Afoakwa et al., 2011, 2013; Castro-Alayo et al., 2019).

Factors such as aeration, cocoa mass, fermentation period, type and duration of the heat source applied during drying, strongly influence the final pH, the titratable acidity and the acid concentration of the beans and, as a result, the quality of cocoa beans (Guehi et al., 2010). During the fermentation process, volatile (acetic, propionic, butyric, isobutyric and isovaleric) and non-volatile (citric, lactic, malic, succinic, oxalic and tartaric) acids are formed through the degradation of sugar by the metabolism of microorganisms in the pulp and, later, these compounds permeate the cocoa beans (cotyledon) (Afoakwa et al., 2008, 2013; Castro-Alayo et al., 2019).

3.2. Colour analysis

Cocoa beans became darker as the fermentation increased, with a progressive decrease in luminosity from unfermented samples to fully fermented samples (Table 2). According to Afoakwa et al. (2011), these changes may be due to the degradation of anthocyanins by enzymatic hydrolysis, which is accompanied by bleaching and subsequent browning of the cocoa beans.

Fully fermented samples (class 1) presented the lowest values for brightness ($L*$), and the highest values of $a*$ and $b*$. This indicates that these samples were slightly darker than the other classes, while higher values in the component $a*$ indicates higher values for "redness". This may have been caused by the decrease in the content of anthocyanins during fermentation, which normally gives the fermented cocoa beans a purple colour and may also be due to the diffusion of polyphenols together with the liquids of the cells (Wollgast and Anklam, 2000; Afoakwa et al., 2013).

Purple pigments were more pronounced in samples of classes 2 and 3, with lower values of $a*$ and $b*$, indicating these samples have more pigments of blue colour. On the other hand, samples of class 4 showed high values of $L*$ (lighter), lower values of $a*$ (less red) and higher values of $b*$ (more yellow), indicating a low content of purple pigments. These values may indicate that the cocoa beans at this stage have not been fermented or that fermentation was insufficient. In addition, studies report that the fermentation process cause more yellow pigments ($b*$) in cocoa beans, most likely due to the presence of oxidized polyphenols, as a result of enzymatic oxidation and polyphenol oxidase in the beans (Hansen et al., 1998; Mayer, 2006). The higher values of $C*$ for samples of classes 1 and 4 indicates the intensity in the colours of these samples, while the lower values for classes 2 and 3 indicates its darkness, while the higher values of hue angle for samples in class 4 indicates it is more yellow. Therefore, cocoa beans become darker with increased fermentation.

3.3. Classification accuracy and performance evaluation metrics

3.3.1. Unbalanced dataset

The high values for accuracy and precision obtained for the classification models using random forest indicate that the model performed well, and could thus be used for prediction of the grade of fermentation of cocoa beans (Table 3).

The results of the model's performance metrics for each class are shown in Table 4, for both calibration and validation models. Classes 1 and 3 showed greater accuracy compared to classes 2 and 4, achieving 0.976 and 0.915, respectively. Accuracy (also known as 'total accuracy') reflects the correct prediction to the total number of objects, that is, the "non-error rate".

The ratio of correctly predicted observations to the total predicted observations for the respective class, i.e., precision, was 1.00 for classes 1 and 3. Most importantly, for an unbalanced data set, the values of Recall and F-score were also high for these classes. The worst performance was obtained for class 2, with a recall of 0.643, for calibration, and 0.475 for validation.

It is important to observe that the dataset is unbalanced, with a smaller number of samples from classes 2 and 4, when compared to classes 1 and 3. This indicates classes 1 (fully fermented) and 3 (under fermented) would probably provide higher values for accuracy and precision, while other classes (partially fermented and unfermented) would be more frequently misclassified by classifiers sensitive to unbalanced data. Since those classes are classified incorrectly, they have an influence on the values of the performance metrics for their respective class due to their small number.

However, the results of the performance metrics of each class (Table 4) reflects the ability of RF algorithm's to obtain a realistic and consistent model to predict the quality of cocoa beans, even with an unbalanced data set. These results can be explained by the way RF discriminates classes, as it generates many classifiers and aggregates their results. RF builds many classification trees with diversity between them that ensures the predictive accuracy of individual trees. It is usually capable to obtain good performance when the number of trees is large enough (Zhang and Suganthan, 2014). Furthermore, these results also indicate that RF has the capacity to deal with this recognition problem since the smallest standard deviation in all metrics shows stability overall repetitions.

It is recommended to avoid the global metrics for assessing the classification performances of unbalanced datasets, unless there is the intention to give more importance to the correct classification of most represented classes. The results showed that the class of unfermented beans had the worst performance, due to the lower number of samples present in this class. In contrast, the fully fermented, partially fermented and under fermented beans classes performed well in the classification, with the fully fermented class having the best result for the metrics evaluated.

3.3.2. Balanced dataset

Although the real dataset could be well predicted with the proposed approach, we tested the classification model in a balanced dataset to observe the performance when all classes had the same number of samples. This method provides more detailed information about the random forest performance, as it avoids the influence of the number of samples from a given class to affect the accuracy.

The results have shown that the classification model performed similarly to the real dataset, with an accuracy of 0.918 (Table 3),

Table 3
Random forest classification performance of global calibration and validation sets – unbalanced and balanced datasets.

Unbalanced Dataset								
	Accuracy		Precision		Recall		F-measure	
	mean	SD	mean	SD	mean	SD	mean	SD
Cal	0.977	0.0003	0.989	0.0002	0.904	0.0015	0.944	0.0009
Val	0.927	0.0011	0.846	0.0039	0.810	0.0019	0.828	0.0028
Balanced Dataset								
Set	Accuracy		Precision		Recall		F-measure	
	mean	SD	mean	SD	mean	SD	mean	SD
Cal	0.994	0.002	0.995	0.001	0.995	0.001	0.995	0.001
Val	0.918	0.001	0.923	0.002	0.900	0.002	0.911	0.002

SD = Standard deviation; Cal = calibration set; Val = validation set.

Table 4
Random Forest classification performance for each class – Unbalanced and balanced datasets.

Unbalanced dataset								
Calibration								
Classes	Accuracy		Precision		Recall		F-Score	
	Mean	SD	Mean	SD	Mean	SD	mean	SD
Class 1	0.998	0.001	0.976	0.001	0.998	0.001	0.987	0.001
Class 2	0.643	0.009	1.000	0.000	0.643	0.009	0.783	0.007
Class 3	0.994	0.000	0.974	0.001	0.994	0.000	0.983	0.001
Class 4	0.983	0.005	1.000	0.000	0.983	0.005	0.991	0.003
Validation								
Classes	Accuracy		Precision		Recall		F-Score	
	Mean	SD	Mean	SD	Mean	SD	mean	SD
Class 1	0.976	0.000	0.934	0.002	0.976	0.000	0.955	0.001
Class 2	0.475	0.020	0.496	0.016	0.475	0.020	0.485	0.016
Class 3	0.915	0.004	0.939	0.003	0.915	0.004	0.927	0.002
Class 4	0.894	0.005	0.976	0.000	0.894	0.005	0.933	0.003
Balanced dataset								
Calibration								
Classes	Accuracy		Precision		Recall		F-Score	
	Mean	SD	Mean	SD	Mean	SD	mean	SD
Class 1	1.000	0.000	1.000	0.000	1.000	0.000	1.000	0.000
Class 2	1.000	0.000	0.985	0.004	1.000	0.000	0.992	0.002
Class 3	0.985	0.004	1.000	0.000	0.985	0.004	0.992	0.002
Class 4	1.000	0.000	1.000	0.000	1.000	0.000	1.000	0.000
Validation								
Classes	Accuracy		Precision		Recall		F-Score	
	Mean	SD	Mean	SD	Mean	SD	mean	SD
Class 1	1.000	0.000	0.860	0.004	1.000	0.000	0.925	0.002
Class 2	0.770	0.001	0.809	0.002	0.771	0.001	0.790	0.005
Class 3	0.895	0.000	1.000	0.000	0.895	0.000	0.944	0.000
Class 4	0.949	0.000	1.000	0.000	0.949	0.000	0.974	0.000

compared to 0.927 for the unbalanced model.

Also, it was observed that the accuracy, precision and recall presented higher values for classes 2 and 4 (Table 4) when compared to the model for the unbalanced dataset. From these results, it was concluded that the texture and colour features extracted from H channel can be good enough for practical uses since both the accuracy and precision were high for both balanced and unbalanced dataset.

Usually, sampling techniques are used to select few cocoa beans, which are analysed to represent a much larger batch of beans (CAOBISCO/ECA/FCC, 2005; ISO 2451:2017, 2017). Although there are also changes in aromatic compounds during cocoa fermentation that are not identified by visual inspection, these changes also cause visual changes in the bean. Thus, imaging systems could expedite the inspection task after developing artificial intelligence framework for image processing; further images can be compared from different regions and countries to enhance the proposed method. Although hyperspectral imaging can provide useful chemical information of cocoa beans (Cruz-Tirado et al., 2020), RGB imaging systems are currently cheaper and can thus be applied by farmers and small processors as a low-cost alternative to the traditional cut-test.

This study showed the importance of establishing an approach for feature extraction from colour images, and determination of most important features, that could accurately predict the fermentation grade of cocoa beans. By acquiring the RGB images of cocoa samples, it was demonstrated that they are useful predictors for classification models, according to the grade of fermentation. This approach could thus be embedded in a computer vision system, providing a cutting-edge technology for the cocoa industry, with a low-cost device.

The system can be adapted with addition of further images (Santos et al., 2019) or datasets of beans containing other defects such as mould, over fermentation and internal insect infestation. Since these defects are more easily distinguishable by visual inspection (hence, by RGB digital imaging), it is expected that the model accuracy will be enhanced when these types of samples are included. The approach in this study could quickly process all the features from the images and identify the grade of fermentation, thus providing a fast overall classification of the sampling batch. It could thus use a larger and more representative sampling, or larger amount of samples from the same batch.

4. Conclusion

The cocoa industry and its food research field claims for automated and reliable methods for the evaluation of product quality. The current results showed that image analyses combined with random forest algorithm is a promising system to classify cocoa beans according to the grade of fermentation. Random forest can be used for reducing the dimension of the feature vector and provide good results for balanced and unbalanced data sets, adjusting to the real occurrence of data.

Our established model has the potential to match different industry and laboratory sizes. Using traditional and handcrafted features, we proposed an accurate solution that demands few computational cost and human intervention during the classification. The prototype created uses an algorithm developed and reported in the current study that can be implemented using vast image processing library and machine learning package, without overburden hardware and software used for classification tasks. Thus, our classification algorithm could then be embedded in a computer vision system and implemented in the cocoa processing and production line, providing a cutting edge, automatic solution. The industry could benefit from the proposed approach to accurately determine the grade of fermentation of cocoa beans.

CRediT authorship contribution statement

Marciano M. Oliveira: Conceptualization, Validation, Data curation, Writing - original draft, Methodology, Investigation. **Breno V. Cerqueira:** Formal analysis, Writing - review & editing, Software. **Sylvio Barbon:** Formal analysis, Methodology, Validation, Writing - review & editing, Software, Visualization. **Douglas F. Barbin:** Conceptualization, Methodology, Supervision, Validation, Resources, Writing - review & editing, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was financed in part by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001; and São Paulo Research Foundation (FAPESP) (project number 2015/24351-2). Marciano Marques Oliveira acknowledges scholarship funding from CAPES. Breno Vicente Cerqueira acknowledges scholarship funding from SAE/Unicamp. Authors are grateful to Mr. Felisberto Coutinho (Docel, Belém-PA, Brazil), Mrs. Arali Aguiar (Gencau, Tambau-SP, Brazil) and Dr. Uilson Lopes (Cocoa Research Center, CEPLAC/Ilhéus-BA, Brazil) for providing samples and technical assistance.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.jfca.2020.103771.

References

Afoakwa, E.O., Paterson, A., Fowler, M., Ryan, A., 2008. Flavor formation and character in cocoa and chocolate: a critical review. Crit. Rev. Food Sci. Nutr. 48 (9), 840–857.

Afoakwa, E.O., Quao, J., Takrama, J., Budu, A.S., Saalia, F.K., 2011. Chemical composition and physical quality characteristics of Ghanaian cocoa beans as affected by pulp pre-conditioning and fermentation. J. Food Sci. Technol. 47 (1), 3–11.

Afoakwa, E.O., Kongor, J.E., Takrama, J., Budu, A.S., 2013. Changes in nib acidification and biochemical composition during fermentation of pulp pre-conditioned cocoa (*Theobroma cacao*) beans. Int. Food Res. J. 20 (4), 1843–1853.

Amoa-Awua, W.K., 2015. Methods of Cocoa Fermentation and Drying, in R.F. Schwandand G.H. Fleet, Cocoa and Coffee Fermentations, first ed. CRC Press - Taylor &Francis Group, NY, USA, pp. 71–116.

Ballabio, D., Todeschini, R., 2009. Multivariate classification for qualitative analysis. In: Sun, D.-W. (Ed.), Infrared Spectroscopy for Food Quality Analysis and Control, pp. 83–104.

Barbin, D.F., Mastelini, S.M., Barbon Jr., S., Campos, G.F., Barbon, A.P.A., Shimokomaki, M., 2016. Digital image analyses as an alternative tool for chicken quality assessment. Biosyst. Eng. 144, 85–93.

Barbin, D.F., Maciel, L.F., Bazoni, C.H.V., Ribeiro, M., da, S., Carvalho, R.D.S., Bispo, E., da, S., Spíndola, M.P., Hirooka, E.Y., 2018. Classification and compositional characterization of different varieties of cocoa beans by near infrared spectroscopy and multivariate statistical analyses. J. Food Sci. Technol. 55 (7), 2457–2466. https://doi.org/10.1007/s13197-018-3163-5.

Barbon, A.P.A., Barbon Jr., S., Mantovani, R.G., Fuzyi, E.M., Peres, L.M., Bridi, A.M., 2016. Storage time prediction of pork by computational intelligence. Comput. Electron. Agric. 127, 368–375.

Beg, M.S., Ahmad, S., Jan, K., Bashir, K., 2017. Status, supply chain and processing of cocoa-A review. Trends Food Sci. Technol. 66, 108–116.

Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: a review of applications and future directions. ISPRS J. Photogramm. Remote. Sens. 114, 24–31.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.

Campos, G.F.C., Mastelini, S.M., Aguiar, G.J., Mantovani, R.G., de Melo, L.F., Barbon, S., 2019. Machine learning hyperparameter selection for contrast limited adaptive histogram Equalization. EURASIP J. Image Video Process. 2019 (1), 59.

CAOBISCO/ECA/FCC, 2005. In: End, M.J., Dand, R. (Eds.), Cocoa Beans: Chocolate and Cocoa Industry Quality Requirements.

Castro-Alayo, E.M., Idrogo-Vásquez, G., Siche, R., Cardenas-Toro, F.P., 2019. Formation of aromatic compounds precursors during fermentation of Criollo and Forastero cocoa. Heliyon 5 (1), e01157. https://doi.org/10.1016/j.heliyon.2019.e01157.

Commission Internationale de l'Eclairage - CIE, 1978. Recommendations on Uniform Colour Spaces, Color-difference Equations, Psychometric Colour Terms. CIE Publication. Commission Internationale de l'Eclairage, CIE.

Cruz-Tirado, J.P., Fernández Pierna, J.A., Rogez, H., Barbin, D.F., Baeten, V., 2020. Authentication of cocoa (*Theobroma cacao*) bean hybrids by NIR-hyperspectral imaging and chemometrics. Food Control 118. https://doi.org/10.1016/j.foodcont.2020.107445).

Cuadros-Rodríguez, L., Pérez-Castaño, E., Ruiz-Samblás, C., 2016. Quality performance metrics in multivariate classification methods for qualitative analysis. Trac Trends Anal. Chem. 80, 612–624.

Du, C.J., Sun, D.W., 2005. Comparison of three methods for classification of pizza topping using different colour space transformations. J. Food Eng. 68, 277–287.

FAOSTAT, 2020. Production/Yield Quantities of Cocoa, Beans in World, 2020. Retrieved from p. Retrieved June 14. http://www.fao.org/faostat/en/#data/QC/visualize.

Farid, D.M., Zhang, L., Rahman, C.M., Hossain, M.A., Strachan, R., 2014. Hybrid decision tree and naïve Bayes classifiers for multi-class classification tasks. Expert Syst. Appl. 41 (4), 1937–1946.

Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recognit. Lett. 27 (8), 861–874.

Guehi, S.T., Dabonne, S., Ban-Koffi, L., Kedjebo, D.K., Zahouli, G.I.B., 2010. Effect of turning beans and fermentation method on the acidity and physical quality of raw cocoa beans. Adv. J. Food Sci. Technol. 2 (3), 163–171.

Hansen, C.E., del Olmo, M., Burri, C., 1998. Enzyme activities in cocoa beans during fermentation. J. Sci. Food Agric. 77 (2), 273–281.

ISO 2451:2017 - Cocoa beans — Specification and quality requirements. 449 Retrieved 23 September 2020, from https://www.iso.org/standard/68202.html.

Jimenez, J.C., Amores, F.M., Solórzano, E.G., Rodríguez, G.A., La Mantia, A., Blasi, P., Loor, R.G., 2018. Differentiation of Ecuadorian National and CCN-51 cocoa beans and their mixtures by computer vision. J. Sci. Food Agric. 98 (7), 2824–2829. https://doi.org/10.1002/jsfa.8790.

Kadow, D., Bohlmann, J., Phillips, W., Lieberei, R., 2013. Identification of main fine flavour components in two genotypes of the cocoa tree (Theobroma cacao L.). J. Appl. Bot. Food Qual. 86 (1).

Kennard, R.W., Stone, L.A., 1969. Computer aided design of experiments. Technometrics 11 (1), 137–148.

Kongor, J.E., Takrama, J.F., Budu, A.S., Mensah-Brown, H., Afoakwa, E.O., 2013. Effects of fermentation and drying on the fermentation index and cut test of pulp pre-conditioned ghanaian cocoa (Theobroma cacao) beans. J. Food Sci. Eng. 3, 625–634.

Kongor, J.E., Hinneh, M., de Walle, D.Van, Afoakwa, E.O., Boeckx, P., Dewettinck, K., 2016. Factors influencing quality variation in cocoa (Theobroma cacao) bean flavour profile — a review. Food Res. Int. 82, 44–52. https://doi.org/10.1016/j.foodres.2016.01.012.

Kumari, N., Grimbs, A., D'Souza, R.N., Verma, S.K., Corno, M., Kuhnert, N., Ullrich, M.S., 2018. Origin and varietal based proteomic and peptidomic fingerprinting of Theobroma cacao in non-fermented and fermented cocoa beans. Food Res. Int. 111, 137–147. https://doi.org/10.1016/j.foodres.2018.05.010.
Labatut, V., Cherif, H., 2011. Evaluation of performance measures for classifiers comparison. Ubiquitous Comput. Commun. J. 6, 21–34.
Lawi, A., Adhitya, Y., 2018. Classifying Physical Morphology of Cocoa Beans Digital Images using Multiclass Ensemble Least-Squares Support Vector Machine. J. Phys. Conf. Series 979 (1), 012029. IOP Publishing.
León-Roque, N., Abderrahim, M., Nuñez-Alejos, L., Arribas, S.M., Condezo-Hoyos, L., 2016. Prediction of fermentation index of cocoa beans (Theobroma cacao L.) based on colour measurement and artificial neural networks. Talanta 161, 31–39.
Lopes, J.F., Ludwig, L., Barbin, D.F., Grossmann, M.V.E., Barbon, S., 2019. Computer vision classification of barley flour based on spatial pyramid partition ensemble. Sensors 19 (13), 2953.
Mathanker, S.K., Weckler, P.R., Bowser, T.J., Wang, N., Maness, N.O., 2011. AdaBoost classifiers for pecan defect classification. Comput. Electron. Agric. 77, 60–68.
Mayer, A.M., 2006. Polyphenol oxidases in plants and fungi: going places? A review. Phytochemistry 67 (21), 2318–2331.
Mite-Baidal, K., Solís-Avilés, E., Martínez-Carriel, T., Marcillo-Plaza, A., Cruz-Ibarra, E., Baque-Bustamante, W., 2019. Analysis of computer vision algorithms to determine the quality of fermented cocoa (Theobroma cacao): systematic literature review. In: In 2nd International Conference on ICTs in Agronomy and Environment. Springer, Cham, pp. 79–87.
Moreira, I.M., da, V., Vilela, L., de, F., Santos, C., Lima, N., Schwan, R.F., 2018. Volatile compounds and protein profiles analyses of fermented cocoa beans and chocolates from different hybrids cultivated in Brazil. Food Res. Int. 109, 196–203. https://doi.org/10.1016/j.foodres.2018.04.012.
Nazaruddin, R., Seng, L.K., Hassan, O., Said, M., 2006. Effect of pulp preconditioning on the content of polyphenols in cocoa beans (Theobroma cacao) during fermentation. Ind. Crops Prod. 24 (1), 87–94.
Nolasco-Perez, I.M., Rocco, L.A.C.M., Cruz-Tirado, J.P., Pollonio, M.A.R., Barbon, S., Barbon, A.P.A.C., Barbin, D.F., 2019. Comparison of rapid techniques for classification of ground meat. Biosyst. Eng. 183, 151–159.
Ohta, Y.I., Kanade, T., Sakai, T., 1980. Colour information for region segmentation. Comput. Graph. Image Process. 13 (3), 222–241.
Parra, P., Negrete, T., Llaguno, J., Vega, N., 2018a. Computer vision methods in the process of fermentation of the cocoa bean. In: In 2018 IEEE Third Ecuador Technical Chapters Meeting (ETCM). IEEE, pp. 1–6.
Parra, P., Negrete, T., Llaguno, J., Vega, N., 2018b. Computer vision techniques applied in the estimation of the cocoa beans fermentation grade. In: In 2018 IEEE ANDESCON. IEEE, pp. 1–10.
Pereira, L.F.S., Barbon Jr., S., Valous, N.A., Barbin, D.F., 2018. Predicting the ripening of papaya fruit with digital imaging and random forests. Comput. Electron. Agric. 145, 76–82.
Powers, D.M.W., 2011. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness & correlation. J. Mach. Learn. Technol. 2, 37–63.
Santos, F.A., Palmeira, E.S., Jesus, G.J., 2019. An image dataset of cut-test-classified cocoa beans. Data Brief 24, 103916.
Szczypiński, P.M., Zapotoczny, P., 2012. Computer vision algorithm for barley kernel identification, orientation estimation and surface structure assessment. Comput. Electron. Agric. 87, 32–38.
Wollgast, J., Anklam, E., 2000. Review on polyphenols in Theobroma cacao: changes in composition during the manufacture of chocolate and methodology for identification and quantification. Food Res. Int. 33 (6), 423–447.
Yro, A., N'zi, C.E., Kpalma, K., 2018. Cocoa beans fermentation degree assessment for quality control using machine vision and multiclass SVM classifier. Int. J. Innovation Appl. Studies, ISSR J. 24 (4), 1711–1717 ffhal-01986522f.
Zhang, L., Suganthan, P.N., 2014. Random forests with ensemble of feature spaces. Pattern Recognit. 47 (10), 3429–3437.
Zhang, C., Liu, C., Zhang, X., Almpanidis, G., 2017. An up-to-date comparison of state-of-the-art classification algorithms. Expert Syst. Appl. 82, 128–150.
Zhang, C., Ma, Y. (Eds.), 2012. Ensemble Machine Learning: Methods and Applications. Springer Science & Business Media.
Zheng, C., Sun, D.W., Zheng, L., 2006. Recent developments and applications of image features for food quality evaluation and inspection–a review. Trends Food Sci. Technol. 17 (12), 642–655.