Engineering Applications of Artificial Intelligence 125 (2023) 106736

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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

Cocoa beans classification using enhanced image feature extraction techniques and a regularized Artificial Neural Network model

Opoku Eric [a,*], Rose-Mary Owusuaa Mensah Gyening [a], Obed Appiah [b], Kate Takyi [a], Peter Appiahene [b]

[a] *Kwame Nkrumah University of Science and Technology, Ghana*
[b] *University of Energy and Natural Resources, Ghana*

ARTICLE INFO

Index terms:
Gray level cooccurrence matrix (GLCM)
Artificial neural networks
Computer-aided diagnostics
Cut-test
Machine learning
Cocoa

ABSTRACT

Cut-Test technique employs visual inspection of interior coloration, compartmentalization, and defects of beans for effective classification of cocoa beans. However, due to its subjective nature and natural variations in visual perception, it is intrinsically limited, resulting in disparities in verdicts, imprecision, discordance, and time-consuming and labor-intensive classification procedures. Machine Learning (ML) techniques have been proposed to address these challenges with significant results, but there is still a need for improvement. In this paper, we propose a color and texture extraction technique for image representation, as well as a generalized, less complex Neural Network model, to help improve the performance of machine classification of Cut-Test cocoa beans. A total of 1400 beans were classified into 14 grades. Experimental results on the equal cocoa cut-test dataset, which is the standard publicly available cut-test dataset, show that the novel extraction method combined with the developed Artificial Neural Networks provides a more homogeneous classification rate for all grades, obtaining 85.36%, 85%, 83%, and 83% for accuracy, precision, recall, and F1 measure, respectively. The proposed model outperforms other ML models, such as Support Vector Machines, Decision Trees, Random Forests, and Naïve Bayes, on the same dataset. Additionally, the proposed ANN model demonstrates relatively better generalization when compared with earlier work by Santos on the same dataset. The proposed techniques in this work are robust on the cut-test dataset and can serve as an accurate Computer-Aided Diagnostic tool for cocoa bean classification.

1. Introduction

Cocoa beans are fermented and dried seeds of the Theobroma cacao tree that are consumed worldwide. Cocoa production currently stands at 4.48 million tons per year, valued at $98.38 billion, and has been increasing at a 3% annual rate for the last decade (Voora et al., 2019; Beg et al., 2017). Cocoa beans are used to make chocolate, liquor, cocoa butter, and a variety of other products such as cocoa beverages, ice cream, and bread (Voora et al., 2019; Afoakwa et al., 2014; Dias, 2014). To be commercialized internationally, the beans must pass a series of quality tests that are mandated by global standards. One of the most popular methods for grading cocoa beans is the Cut-Test technique. This method has been widely used to test the quality of cocoa beans around the world and has been proven to be one of the best (Board, 2018; Santos et al., 2019). Cut-Test uses color changes registered during fermentation for grading (León-Roque et al., 2016). The test involves cutting beans lengthwise in half and having experts physically examine and evaluate the interior of the beans to determine coloring, compartmentalization, flaws, and other characteristics (León-Roque et al., 2016; Antunes et al., 2019; Catsberg and Dommelen, 1990). However, inspections by different experts can sometimes lead to disparities in grading due to human emotional variables such as weary eyes and biases, as well as the time-consuming and labor-intensive nature of the process (Aubain et al., 2020a; Owusu Ansah et al., 2018; Astika et al., 2010; Savakar, 2012; Majumdar et al., 1996; Essah et al., 2022). A possible solution to these challenges is Computer-Aided Diagnostics (CAD) (Kamilaris and Prenafeta-Boldú, 2018; Ismail et al., 2022).

The introduction and prevalence of digital images have paved the way for the application of Computer-Aided Classification or Diagnostics (Ismail et al., 2022), such as in the cocoa industry for cocoa bean inspection. CAD for cocoa bean inspection could harmonize classification and reduce interobserver variability, leading to more consistent and reliable classification and better decisions on grading cocoa beans which can provide a more effective and efficient solution to the grading of cocoa beans. In recent years, there has been an increasing number

* Corresponding author.
E-mail addresses: erik.opoku613@gmail.com (O. Eric), rmo.mensah@knust.edu.gh (R.-M.O.M. Gyening), obed.appiah@uenr.edu.gh (O. Appiah), takyikate@knust.edu.gh (K. Takyi), peter.appiahene@uenr.edu.gh (P. Appiahene).

https://doi.org/10.1016/j.engappai.2023.106736
Received 12 August 2022; Received in revised form 17 May 2023; Accepted 26 June 2023
Available online 14 July 2023

of studies focusing on the development of CAD systems for cocoa bean classification (León-Roque et al., 2016; Antunes et al., 2019; Aubain et al., 2020a; Barbon et al., 2016), with some demonstrating impressive results in terms of accuracy and precision (Asogwa and Orimoloye, 2021). CAD mostly uses Machine Learning (ML) algorithms, such as Support Vector Machines (SVMs), K-Nearest Neighbors (KNNs), K-means, and Artificial Neural Networks (ANNs), to effectively classify cocoa beans into various grades, such as Fermented beans, Unfermented beans, Moldy beans, Agglutinated beans, Slaty beans, Smoky beans, etc. (Dand, 1999). However, despite the successes of these machine learning tools in aiding the classification of cocoa bean quality, the vast majority of cocoa bean classification studies utilize fewer than six classes of beans, making it difficult to generalize the models developed in the cocoa industry to a wider context (Santos et al., 2019). Additionally, a considerable number of these studies fail to report precision, recall, and F1-measure values, which are critical indicators for assessing the effectiveness of a proposed model, particularly when the dataset used for training and testing the model's performance is imbalanced (Chawla et al., 2002; Kurniawati, 2019). Furthermore, the current accuracies obtained on the Cut-Test cocoa dataset have yet to achieve optimal results and are therefore subject to improvement (Jintawatsakoon and Inthachot, 2021; Anggraini et al., 2021). Improved learning-based self-adaptation techniques for classifying beans would be of great benefit to the cocoa industry and could facilitate the widespread adoption of CAD for classifying cocoa beans in the global market.

This paper proposes a new approach to improving machine classification of Cut-Test cocoa beans using effective color and texture feature extraction techniques, a generalized Artificial Neural Network model, and a comprehensive dataset to accurately and quickly grade cocoa beans. The method can be used in place of manual inspection of beans during the cut-test exercise to achieve consistent and reliable cocoa bean classification.

The paper is organized as follows: related work is provided in Section 2, methods are discussed in Section 3, and results and conclusions are discussed in Sections 4 and 5, respectively.

2. Related works

Machine Learning has had a huge impact on agriculture, as it is now used to deal with variations and uncertainties in crop production, forecast rainfall and recommends crop varieties for plantation, predict crop yields, and soil organic matter, and classify plant diseases, among other agricultural-related activities (Astika et al., 2010; Van Klompenburg et al., 2020; Barbin et al., 2016; Yudianto et al., 2021; Medar et al., 2019; Bhagawati et al., 2016). In recent years, the assessment of food quality, such as fish and meat products, vegetables, grains, fruits, beans, and other foods, has seen both classical and traditional machine learning algorithms implemented to achieve optimum results (Peres et al., 2021; Mathanker et al., 2011).

In this context, this section reports the use of Machine Learning techniques in conjunction with other methodologies that have been used to assess the quality of fermented cocoa beans. The classification of fermented cocoa beans using digital images was demonstrated recently by Aubain et al. (2020b). In their work, digital images of various classified fermented cocoa beans were processed to extract relevant predictors. RGB (Red, Green, and Blue) and HSV (Hue, Saturation, and Value) color spaces were used for their experiment. Three grades of beans – fermented beans, un-fermented beans, and moldy beans – were classified. In their experiment, a Multi-Class Integrated Least Square Support Vector (MCIL-SSV) was used, achieving an accuracy of 99.281% for the three fermented grades of cocoa beans. However, the paper did not present other metrics such as Precision, Recall, and F1-measure, which are crucial in machine learning to assess the strength of the model, especially on imbalanced datasets. A common error that may be related to the high accuracy is the issue of the accuracy paradox, which is largely related to imbalanced datasets used to train this model (Sturm, 2013). Additionally, the use of only three fermented grades as the dataset limits the model for industrial purposes in terms of generalization.

Lawi and Adhitya (2018) presented an interesting feature extraction approach for bean images based on image morphology. The beans' area, perimeter, major axis length, minor axis length, aspect ratio, circularity, roundness, and ferret diameter are the major features extracted from an image. Normal Beans, Broken Beans, Fractured Beans, and Skin Damage Beans were the four categories of cocoa beans graded. The multi-Ensemble Least-Support Vector Machine model was utilized, and the system achieved an accuracy of 99.705%. However, the work was primarily interested in the physical characteristics of the beans, not issues related to fermentation. Indeed, all the extracted features used by Lawi and Adhitya were related to shapes, and therefore such a model would not be ideal for the classification of fermented beans.

Astika et al. (2010) attempted to address the issue of only classifying beans based on shape. In their work, the cocoa bean's shape and color properties were extracted using image processing techniques. The relationship between the input parameters, the bean's quality components, and the outputs were then developed using two ANN models: one for classifying bean breakage and the other for classifying fermentation status. The initial Artificial Neural Network (ANN) contained 35 input nodes and could sort beans into four different grades: whole beans, broken beans, bean fractions, and skin-damaged beans. The second ANN structure used six color parameters to divide the beans into three categories: fermented, non-fermented, and moldy beans. The first ANN could predict accuracies of 84%, 52%, 20%, and 20% for whole beans, broken beans, bean factions, and skin-damaged beans, respectively. The second ANN predicted 99%, 98%, 79%, and 92% accuracies for normal fermented beans, non-fermented beans, and moldy beans, respectively. The model registered 92% of accuracy on average. However, the researchers did not report the results of the various evaluation metrics that are accepted in the 22 domains of machine learning. Additionally, the dataset used for the experiment was not based on a cut-test dataset, which is difficult to perform classification on Board (2018).

Aubain et al. (2020a) classified cocoa beans using a Multi-class Support Vector Machine (SVM). In their work, cocoa beans were classified into three categories: Unfermented (UF), Partly Fermented (PF), and Whole Fermented (WF). The work reported accuracies of 99.17%, 97.50%, and 100% for Unfermented (UF), Partly Fermented (PF), and Whole Fermented (WF) beans, respectively. However, problems related to the fermentation of cocoa beans go beyond just three classes that the authors worked on. To be able to deal with the proper classification of cocoa beans, the model must be able to handle at least 10 different cocoa grades related to fermented cocoa beans, which are common in the industry. In an effort to address the challenges associated with most models presented by researchers to effectively classify cocoa beans, Adhitya et al. (2020) expanded the number of classes for their work. Whole beans, Fractured beans, Skin-damaged beans, Fermented beans, Unfermented beans, and Moldy beans were categorized. Gray Level Co-Occurrence Matrix (GLCM) and Convolutional image processing techniques were used for the extraction of textural information from cocoa beans images. The GLCM texture feature extraction with Support Vector and XGBoost classifiers achieved 61.04% and 65.08% accuracy, respectively, whereas using the Convolution feature extraction achieved 59.14% and 56.99% accuracy for Support Vector and XGBoost classifiers, respectively. It was discovered that GLCM texture feature extraction produced more consistent results than Convolution feature extraction. The results presented for the two extraction methods are low and indicate that dealing with a higher number of classes for CAD does not easily result in excellent accuracy values.

Antunes et al. (2019) studied an upgraded version of a computer vision system that classified cocoa beans by evaluating applications of classification algorithms and attribute selection. There were three phases to the system presented. In their experiment, 14 different classes

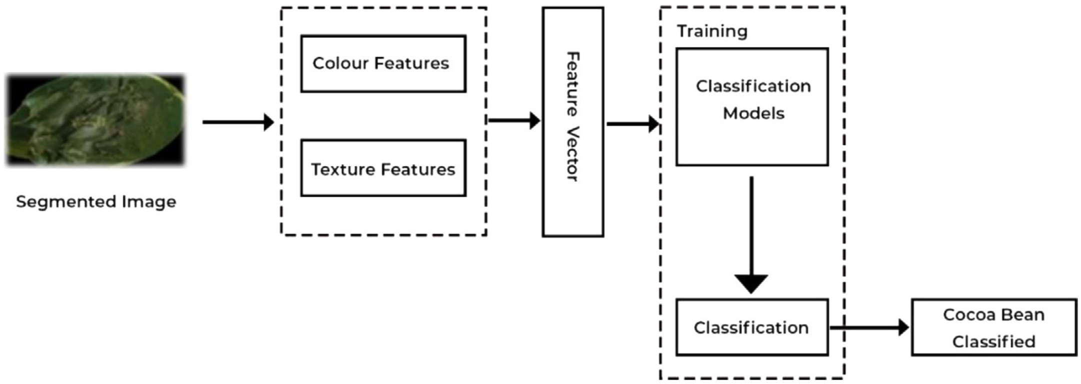


Fig. 1. Overview of the proposed feature extraction.

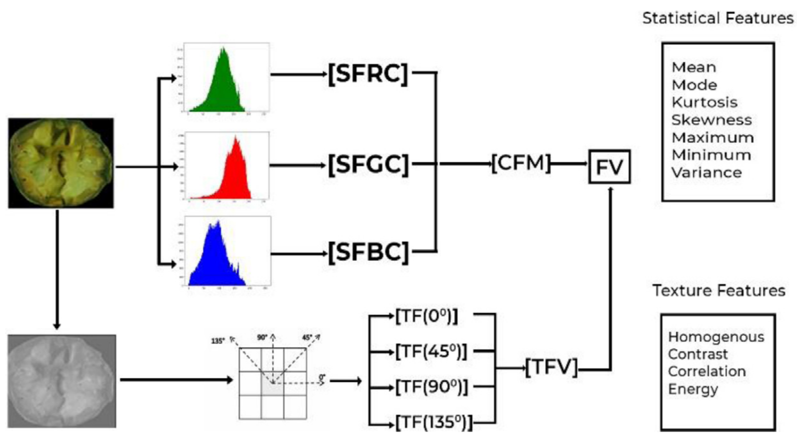


Fig. 2. Method for the feature extraction.
TF = Texture Feature, TFV= Texture Feature Vector, FV = Feature Vector, SFRC = Statistical Feature for Red Channel, SFGC = Statistical Feature for Green Channel, SFBC = Statistical Feature for Red Channel, CFM = Color Feature Map.

of cocoa beans were used. The 14 Haralick's Textural Characteristics, the average pixel values of each RGB layer, and grayscale-converted picture color features, as well as the area and circumference of the bean, were extracted (structural features). k-NN, Naive Bayes, Decision Tree, and Multi-Layer Perceptron Neural Network were the four supervised classification models tested. According to their findings, the most accurate model attained 92.9714% accuracy with a 0.2178 standard deviation, which is lower than the 99.64% accuracy reported by their previous studies. They believe, however, that the model that scored 92.9714% is superior in terms of generalization and has a lower level of complexity, thus it is a preferable version to utilize in a computer vision model. This supports the fact that accuracies are not enough to justify a model, especially when other metrics for evaluating a model have not been presented in works as seen mostly in the cocoa beans' classification papers.

Based on the reviewed literature, the majority of cocoa bean classification works use less than six (6) classes of beans. This makes it difficult to generalize the models that have been developed in the cocoa industry (Santos et al., 2019). More so, most of the works generally report the accuracy of their model without the precision, recall, and f1-measure values which are critical for evaluating the strength of a proposed model, especially where the dataset used for training and testing the performance of the model is imbalanced (Chawla et al., 2002; Kurniawati, 2019). A cocoa beans dataset with more than 10 classes can be more comprehensive since it will present almost all the possible states a fermented cocoa bean can be in. Models trained with such comprehensive datasets can be generalized as compared to earlier research with fewer grades of fermentation. Further, these works did not present effective analyses of the Precision, Recall, and F-Measure even on unbalanced datasets to be able to properly evaluate their model on the dataset. This, therefore, leaves room for more exploration to be done on the cocoa beans dataset to develop a robust Machine Learning model for the Cut-Test dataset. The main contribution of this paper is the use of an effective feature extraction technique for the cut-test dataset as well as a robust ML algorithm to effectively classify cocoa beans.

3. Methods

This section presents the methods and experimental setup, including the dataset and its management, the image classification pipeline, the baseline algorithms for the experiments, the principles employed to model these algorithms, and the evaluation techniques used as a unique classifier to differentiate and choose the optimal classifier. Two main effective techniques are presented in this work: feature selection techniques and a novel machine learning algorithm. These techniques are explained in the subsequent sections. Fig. 1 presents an overview of the framework for this experiment.

3.1. Feature extraction

Color and texture features are extracted from each image and stored as a feature vector to identify the most important features and reduce the data to be analyzed and computing costs, which can sometimes improve classification performance by avoiding unnecessary information. Reducing the computational load for the model and avoiding insignificant data is important for the Cut-Test, which relies heavily on the color and texture of cocoa beans. In this section, we describe the details of the feature extraction techniques as shown in Fig. 2.

Table 1
Color and texture information extracted for the image feature vector.

Seq.	Feature category	Operation	Description
1		Min (I_R)	Minimum pixel value in R channel
2		Max (I_R)	Maximum pixel value in R channel
3		Mode (I_R)	Highest frequency pixel value in the R channel
4		Var (I_R)	The variance pixel value of the R color channel
5		Mean (I_R)	Mean value pixel of R color channel
6		Skewness (I_R)	Skewness extracted from the R color channel
7		Kurtosis (I_R)	Kurtosis extracted from the R color channel
8		min (I_G)	Minimum pixel value in G color channel
9		Max (I_G)	Maximum pixel value in G color channel
10		Mode (I_G)	Highest frequency pixel value in G color channel
11	Color features	Var (I_G)	The variance pixel value of the G color channel
12		Mean (I_G)	Mean value pixel of G color channel
13		Skewness (I_G)	Skewness extracted from G color channel
14		Kurtosis (I_G)	Kurtosis extracted from G color channel
15		Min (I_B)	Minimum pixel value in B color channel
16		Max (I_B)	Maximum pixel value in B color channel
17		Mode (I_B)	Highest frequency pixel value in the B color channel
18		Var (I_B)	The variance pixel value of B color channel
19		Mean (I_B)	Mean value pixel of G color channel
20		Skewness (I_B)	Skewness extracted from B color channel
21		Kurtosis (I_B)	Kurtosis extracted from B color channel
22		Contr (I_0)	Contrast the property of GLCM with an angle direction of 0^0
23		Contr (I_{45})	Contrast property of GLCM with an angle direction of 45^0
24		Contr (I_{90})	Contrast property of GLCM with an angle direction of 90^0
25		Contr (I_{135})	Contrast property of GLCM with an angle direction of 135^0
26		Ener (I_0)	Energy property of GLCM with an angle direction of 0^0
27		Ener (I_{45})	Energy property of GLCM with an angle direction of 45^0
28		Ener (I_{90})	Energy property of GLCM with an angle direction of 90^0
29	Texture Features	Ener (I_{135})	Energy property of GLCM with an angle direction of 135^0
30		Corr (I_0)	Correlation property of GLCM with an angle direction of 0^0
31		Corr (I_{45})	Correlation property of GLCM with an angle direction of 45^0
32		Corr (I_{45})	Correlation property of GLCM with an angle direction of 90^0
33		Corr (I_{135})	Correlation property of GLCM with an angle direction of 135^0
34		Homo (I_0)	Homogeneity property of GLCM with an angle direction of 0^0
35		Homo (I_{45})	Homogeneity property of GLCM with an angle direction of 45^0
36		Homo (I_{90})	Homogeneity property of GLCM with an angle direction of 90^0
37		Homo (I_{135})	Homogeneity property of GLCM with an angle direction of 135^0

3.1.1. color features

Image quantization is initially employed to reduce the number of colors in the images to 64 distinct colors for the effective representation of pixels (Zheng et al., 2006; Ohta et al., 1980). The color channels of each bean image are separated into three matrices, each representing a color channel (R, G, B). The histogram for each channel is extracted and stored in three 1D vectors. Seven statistical values: variance, mode, mean, maximum, minimum, skewness, and kurtosis are then extracted from each of the 1D vectors. In total, a color feature vector of size 21 is used to represent the color features of each bean. Eq. (1) to Eq. (3) represent the equations used for the color features for variance, skewness, and kurtosis, respectively.

$$\sigma = \sqrt{\frac{1}{N} + \sum_{i=1}^{N}(x_i - \mu)^2} \tag{1}$$

$$\frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{(\frac{1}{n}\sum_{i=1}^{n}(x_{i-\bar{x}})^2)^{3/2}} \tag{2}$$

$$\frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{(\frac{1}{n}\sum_{i=1}^{n}(x_{i-\bar{x}})^2)^2} - 3 \tag{3}$$

3.1.2. Texture feature

Gray Level Co-occurrence Matrices (GLCMs), proposed by Halalick in 1973, have been used to extract texture information from beans. GLCMs are mathematical tools used to extract texture information from images by analyzing the spatial relationship between pixels. To create a GLCM, an image is divided into cells, and the frequency of different gray level combinations is counted based on their spatial relationship, defined by the distance and angle between the pixels. Statistical measures such as contrast, correlation, energy, and homogeneity can be calculated from the GLCM and used to describe the texture of the image and differentiate it from others. GLCMs are commonly used in image processing and computer vision. Upon visually assessing the images for this task, we found GLCM to be particularly suitable for capturing fine details, patterns, and variations in textures, which are essential in applications where texture analysis plays a vital role, such as the present study. Moreover, GLCM has demonstrated successful applications in various fields, including medical imaging (Indra and Jusman, 2021), remote sensing (Oliveira et al., 2021), material inspection, and pattern recognition. Similar predictors utilizing GLCM as a feature extractor have been used on different datasets in similar work (Campos et al., 2019; Lopes et al., 2019), further supporting its effectiveness in the domain of study.

In this experiment, four co-occurrence matrices were extracted from each grayscale image generated from each cocoa bean image in the directions 0°, 45°, 90°, and 135°, respectively. To minimize the effect of the background on the models, the frequency of the co-occurrence matrix location M(1,1) was set to zero. The contrast, energy, correlation, and homogeneity were extracted from the co-occurrence matrices and used as the texture information for each bean. The energy, contrast, correlation, and homogeneity from each of the four GLCMs were combined and stored in a vector of size 16. These values were used to represent the texture in the Texture Vector (TV). Eqs. (4) to (7) represent the equations for contrast, correlation, energy, and homogeneity, respectively (see Table 1).

$$\sum_{ij} |i - j|^2 p(i, j) \tag{4}$$

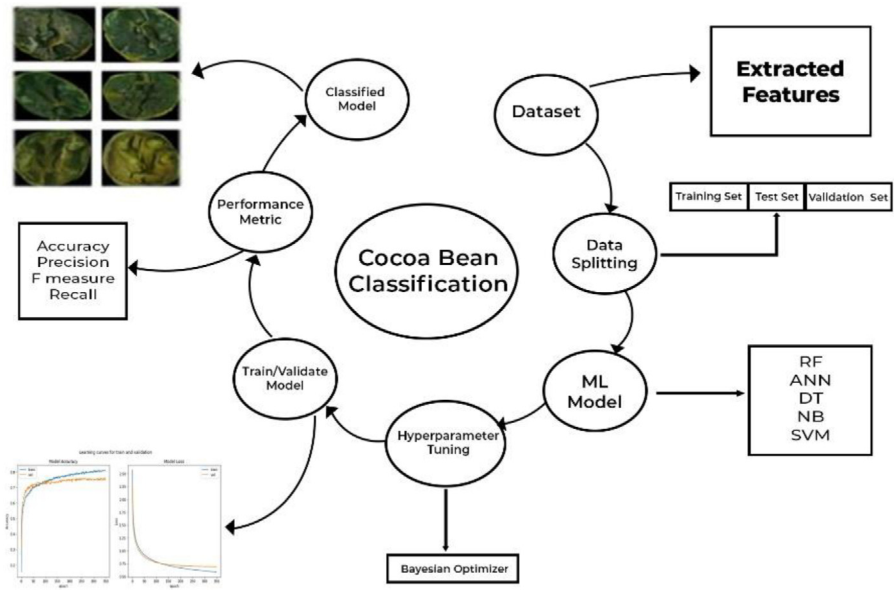


Fig. 3. Flow diagram of the ML models implementation.

$$\sum_{ij} \frac{(i-\mu i)(j-\mu j)p(i,j)}{\sigma_{i\sigma_j}} \tag{5}$$

$$\sum_{ij} p(i,j)^2 \tag{6}$$

$$\sum_{ij} \frac{p(i,j)}{1+|i-j|} \tag{7}$$

3.2. Classification of cocoa beans using machine learning models

In this section, we present the five algorithms used in this study: Support Vector Machine, Random Forest, Decision Tree, Naive Bayes, and ANN. Four of the algorithms (SVM, Random Forest, Decision Tree, and Naive Bayes) were implemented in SkLearn as baseline models. This was done intentionally to evaluate the effect of the extracted features on commonly used classical machine learning algorithms presented in the literature for similar tasks. We also compare the built ANN model in this work to see its effectiveness. Further, Fig. 3 is the flow diagram of the ML model implementation for the classification of cocoa beans.

3.3. Data splitting

The feature-extracted dataset was divided into training and test subsets, with the training subset further split into a training dataset and a validation dataset. The division of the total training dataset into these three subsets was crucial for this study. The validation dataset was used for successive parameter tuning of the various models, allowing for the exploration of a large number of parameter values to achieve the best possible model that exhibits good behavior and avoids overfitting. This is an important step in ensuring that the final model used for testing is able to generalize well to unseen data. In this study, 70% of the dataset was used for training, 10% for validation, and 20% for testing the model. This division of the dataset is a common practice in machine learning and has been used in numerous studies (e.g., Alaa et al., 2020; Liu and Zhang, 2020) (see Table 2).

3.4. Hyperparameter tuning

We performed hyperparameter tuning using the Bayesian optimization technique (Snoek et al., 2012). The validation accuracy was used as the objective to maximize the Bayesian Optimizer. We ran the optimizer for 100 iterative steps and recorded the best set of parameters that gave the highest validation accuracy. Table 3 presents the hyperparameters and their values for the various models.

Table 2
Data splitting for the training, validating, and testing of the models.

Training data	Validation data	Test data
70%	10%	20%

Table 3
The hyperparameters for each model and their value after tuning.

Model	Hyperparameter	Value
SVM	C	2
	degree	10
	Kernel	Linear
	random_state	10
Random Forest	Max_dept	209
	n_estimators	499
	n_jobs	1
	random_state	10
Decision Tree	Criterion	Gini
	Splitter	best
	Random state	10
Naïve Bayes	Var_smoothing	
	Priors	
	Number of layers	2
	Activation functions	Relu, Softmax
ANN	Epochs	500
	Dense	2
	Kernel_regularizer	L1L2
	Regularization factor	0.001
	Batch size	5
	Loss function	Categorical_Crossentropy
	Learning rate	0.001
	Dropout rate	0.2

3.5. Training and evaluation of the neural network

We developed five different Artificial Neural Network models, the architecture of which is illustrated in Fig. 4. The initial model consisted of two dense layers, each followed by a ReLU activation, a dropout layer with a value of 0.2, a learning rate of 0.001, and a SoftMax layer.

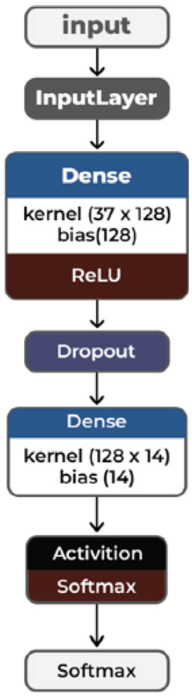


Fig. 4. The architecture of the Proposed Artificial Neural Network model.

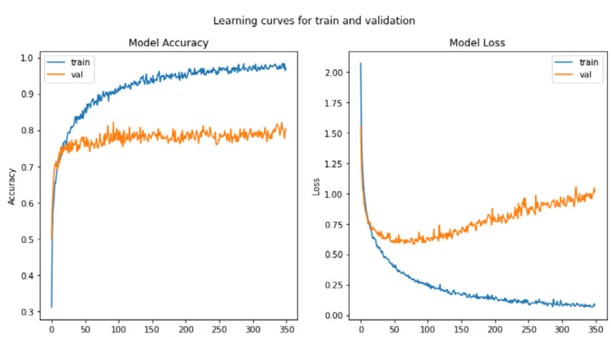


Fig. 5. Learning curve depicting that the model is overfitting.

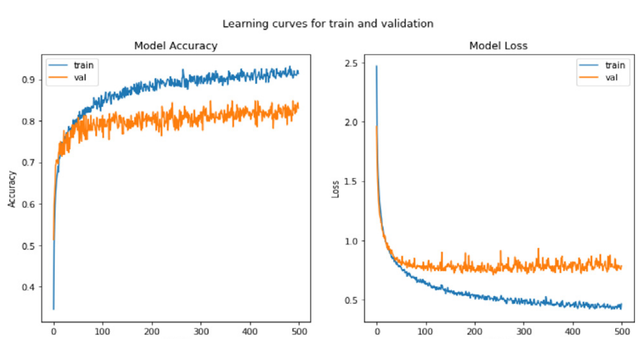


Fig. 6. Learning curve for regularized model.

After training for 500 epochs with a batch size of 16, we observed from Fig. 5 that the learning curve of the cross-entropy losses was decreasing, and accuracies were increasing for some epochs until the validation loss began to increase, even though the training loss was still decreasing. This indicates overfitting to the training data and suggests that the model may not generalize well to unseen data.

To address this concern, we thoroughly explored different techniques to strike a better balance between model complexity and generalization performance. Initially, we applied L1 and L2 regularization individually to combat overfitting in these specific scenarios. However, it is important to note that the improvement achieved was not substantial. This lack of significant improvement could potentially be attributed to certain factors. One factor is that L1 regularization tends to encounter challenges when handling correlated features, as it tends to discard features that may be important. On the other hand, L2 regularization mitigates this issue by distributing the impact of correlated features across all of them. Therefore, the modest improvement observed could be partially attributed to the presence of correlated features. Additionally, the dataset used for feature extraction contained outliers, which can negatively affect the performance of L2 regularization. Due to its penalty on the square of coefficients, L2 regularization is sensitive to outliers. In contrast, L1 regularization demonstrates more robustness in the presence of outliers, thanks to its absolute value penalty. Nonetheless, despite these considerations, we were unable to achieve more favorable results.

The L1L2 regularization, also known as Elastic Net regularization, has demonstrated significant results and is therefore preferred over using L1 or L2 regularization individually for preventing overfitting in this work. The dataset used in our study exhibits a high degree of correlation among its features, and L1L2 regularization is well-suited for handling such scenarios. It accomplishes this by encouraging sparse solutions, effectively performing feature selection, while also allowing for a degree of coefficient sharing among correlated features. Moreover, the combination of L1 and L2 regularization in L1L2 regularization offers improved generalization performance compared to using L1 or L2 alone. By leveraging the strengths of both regularization techniques, L1L2 regularization strikes a balance between controlling model complexity and ensuring a good fit for the data.

We used L1L2 elastic net regularization with a shrinkage factor of 0.001 on the weights of the first fully connected layer of the initial model. This aims to reduce the parameter search space by restricting the range of values that the regularized weights can assume. After training, the validation loss no longer diverged, as indicated in Fig. 6. However, the test accuracy decreased, so we explored other methods for achieving better generalization. Secondly, we implemented early stopping during training when we did not observe a significant change in model performance, particularly the validation loss, after 10 epochs. This helps to avoid overfitting, as the validation and training curves flatten at epoch 140, as shown in Fig. 7. Thirdly, we reduced the model complexity to a single fully connected layer with a ReLU activation and a SoftMax layer. This resulted in a non-overfitting model that achieved good accuracy and better generalization to unseen data, as depicted in Fig. 8. Finally, we increased the epochs of the simple model to 800 and maintained the single fully connected layer with a ReLU activation and a SoftMax layer. We observed that the model was non-overfitting, as indicated in Fig. 9, and the accuracy of the model with a larger number of epochs increased compared to the epochs used in the previous simple model.

3.6. Training and evaluation of the neural network

In this step, we used Artificial Neural Networks as a function approximator to learn the classification rule for mapping the input features of the data to their corresponding labels. The learning process for the neural network is iterative. Training examples are fed to the model, which predicts the labels for each data point in each epoch. We used the cross-entropy loss as the criterion to measure the rate of misclassification for each data point. For a complete pass through the batch of data, we compute the cost, which is the total sum of losses incurred on each training data point in the batch. To improve the model in the iterative step, the gradient of the loss on the batch of data with respect to the model's parameters is computed and used in the gradient

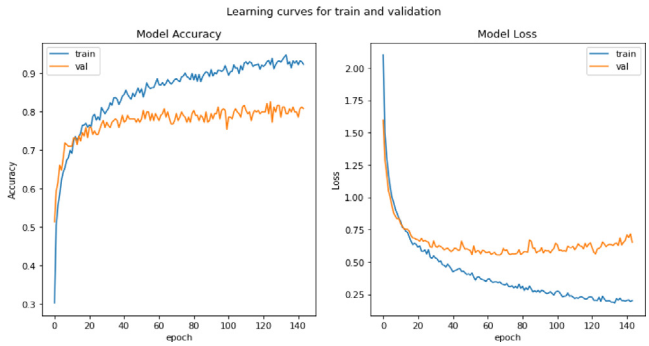


Fig. 7. Learning curve for early stopped model.

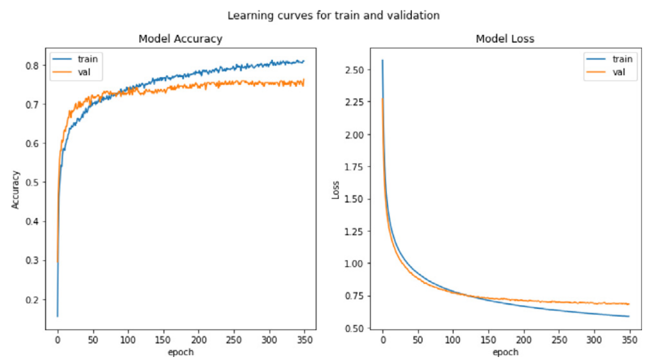


Fig. 8. Learning curve for simple model.

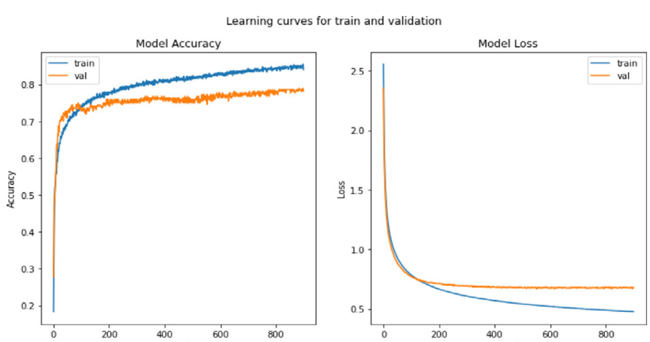


Fig. 9. The learning curve for a simpler model trained for longer (800 epochs).

descent step to update the model's parameters. Essentially, the gradient tells us the direction to step in order to reduce the loss and extent of the learning rate (Manoj et al., 2022).

After a complete pass of the data, the model is evaluated on both the training and validation sets, and the total prediction accuracy and misclassification losses are recorded. We plotted the accuracies and losses for some epochs to visualize how well the model was learning. Ideally, we want the losses for both training and validation to decrease while accuracies increase as training proceeds until the curve flattens out at the point of convergence. At the end of the training, we evaluated our model and reported our performance on the held-out test set.

3.7. Evaluation metrics

In this study, the performance of the baseline and the Artificial Neural Networks models on the cocoa cut-test dataset has been evaluated using Accuracy, F-measure, Recall, and Precision. These evaluation metrics were selected because they have a high level of use in related classification tasks (Antunes et al., 2019; Oliveira et al., 2021). To ensure an equally distributed class distribution, the F-measure was used to calculate the arithmetic mean of precision and recall. The proportion of correctly predicted positive observations to all observations in the 14 classes was classified using recall, also known as sensitivity. Besides, Precision was used to calculate the percentage of correctly predicted positive patterns in a positive class out of all predicted positive patterns. The proportion of correct predictions relative to the total number of instances evaluated was calculated using the accuracy metrics. As a result of the random initialization of weights and biases, shuffling of training data, and random sampling of validation and test data, the performance of machine learning models can vary significantly. To obtain a more reliable estimate of a model's true performance, we run each model five times and calculate the average and standard deviation of its performance metrics. The average provides us with a measure of the central tendency of the model's performance, while the standard deviation tells us about the amount of variability or uncertainty in the estimates. This approach allows us to make more informed decisions about the performance of the model and assess its suitability (Ismail et al., 2021b) for classifying fermented cocoa beans.

3.8. Dataset

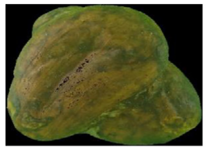
In this study, a standardized image dataset of cocoa beans classified using the Cut Test method and their corresponding classifications was created with the assistance of experts from the Centro de Inovaço do Cacau (CIC), an institution affiliated with the Universidade Estadual de Santa Cruz (UESC) (Santos et al., 2019). The database comprises 100 images per class, one image per bean, and each bean was evaluated by two specialists. If the specialists reached a consensus, the bean was included in the dataset. The classes are categorized according to the following standards: agglutinated, compartmentalized brown, compartmentalized white, compartmentalized partially purple, compartmentalized purple, compartmentalized slaty, plated brown, plated white, compartmentalized partially purple, plated purple, plated slaty, moldered, flattened, and brittle, totaling 14 classes and 1400 samples. Table 4 displays low-resolution samples of the fourteen images from the different classes in the dataset.

4. Results and discussion

For the basic model of the ANN, it was observed from the learning curve that while the cross-entropy of the losses decreased, the accuracies increased for some epochs until the validation loss began to increase, even though the training loss was still decreasing, as illustrated in Fig. 5. We interpret this as overfitting the training data, as the losses continue to diverge from the training. This effect does not encourage good generalization to unseen data. The test accuracy and test loss of the overfitting model were 83.93% and 1.2149, respectively. The L1L2 regularization was the first technique applied to resolve the overfitting problem, resulting in a higher accuracy of 85.36% and a low test loss of 0.7356, as shown in Fig. 6. This technique significantly decreased overfitting and improved its generalization. After training, we observe from Fig. 6 that the learning curve of the validation loss no longer diverges. Instead, the test accuracy increases while the test loss decreases and converges. This illustrates that the model is performing well and generalizing to unobserved data without overfitting. The L1L2 regularizer penalized the ANN model and controlled its complexity, leading to a better generalization of unseen data.

Furthermore, the use of early stopping minimized the validation loss and maximized accuracy in order to prevent overfitting. As shown in Fig. 7, there was no significant change in model performance after 10 epochs. This model achieved an accuracy of 83.93% and a validation loss of 0.7049, demonstrating its ability to generalize to unseen data and avoid overfitting. When the model complexity was reduced to a single fully-connected layer with ReLU activation and a SoftMax layer, we observed a model that was able to generalize well to unseen data

Table 4
Sample images from the fourteen classes.


Agglutinated


Brittle


Compartmentalized Brown

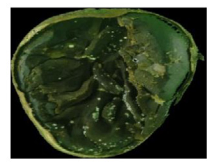

Compartmentalized Partially Purple

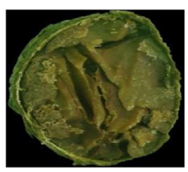

Compartmentalized Purple

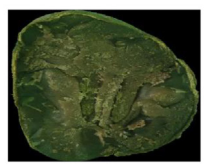

Compartmentalized Slaty

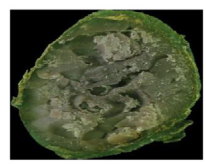

Compartmentalized White


Flattened


Moldered


Plated Brown


Plated Partially Purple


Plated Purple


Plated Slaty


Plated White

while achieving good accuracy. As shown in Fig. 8, the training and validation loss decreased steadily and the gap between the final values of the two losses was minimal, indicating a well-fit model. This model had an accuracy of 77.14% and a test loss of 0.8573. While the accuracy was lower than the previous model, it still represents a good fit and effective generalization to unseen data.

Finally, when the number of epochs was increased mainly to 800 in the simple models, i.e., maintaining the single fully connected layer with relu activation and a SoftMax layer. We observed in Fig. 9 that the model was non-overfitting. The plot of validation loss decreases to the point of stability and has a small gap with the training loss, and the accuracy of the model increased by 4% compared with the previous simple model with the least number of epochs. The model achieved an accuracy of 81.43% and a test loss of 0.9447.

In comparison, the regularized model (with L1L2 regularization) achieved the highest accuracy of 85.36% and the lowest test loss of 0.7356, making it the best model with the best generalization in the experiments. The model using the early stopping method had a slightly lower accuracy of 83.93% and a test loss of 0.7049. However, when a larger training epoch of 800 was used, the model had an accuracy of 81.43% and a test loss of 0.9447, which was significantly better than the overfitting model. In practice, both the early stopping technique and the regularized model (with L1L2 regularization) effectively avoided overfitting and had good generalizations in the experiments. These two models could also serve as robust classifiers for grading beans in an industry. The following section presents the detailed results of the regularized model (with L1L2 regularization) for Precision, f-measure, and Recall of the individual classes (see Table 5).

4.1. Result of regularize model

The results classification report for the optimized model i.e., the regularized model (with the L1L2 regularization) is presented in Table 6.

It has been observed that the Flattened and Compartmentalized Slaty classes had equal values for the evaluation metrics of precision, recall, and f-measure, with 94 and 96 respectively, as shown in Table 6. This indicates the effectiveness and harmony in classification for these classes. The Brittle class had a precision of 100%, the Moldered class had a precision of 60%, and the Platted Partially Purple class had a precision of 62%. The Agglutinated class had the highest scores for precision, recall, and F1-measure, with 97, 100, and 98 respectively. The Plated Partially Purple class had a precision of 62%, a recall of 76%, and an F1 measure of 68%, making it the class with the worst classification out of the 14 classes. In general, the regularized model with the L1L2 regularization (i.e., elastic net) classifier approach had weighted averages of 85.36%, 83%, and 83% for precision, recall, and F1-measure respectively, indicating an efficient model for the classification task.

Table 6 presents the experimental results for precision, recall, and f1-measure, while Fig. 10 presents the accuracy and the weighted average for Precision, Recall, and F1-measure for the five machine

Table 5
Summary of a classifier trained on manually extracted features.

Model	Accuracy		Precision		Recall		F1-Measure		Test loss
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	
SVM	0.7905	0.003	0.8123	0.004	0.8104	0.001	0.7912	0.000	
RF	0.7214	0.004	0.6932	0.006	0.6823	0.006	0.6831	0.005	
NB	0.4613	0.007	0.5945	0.000	0.4621	0.004	0.4532	0.006	
DT	0.5734	0.012	0.5743	0.007	0.5534	0.004	0.5521	0.004	
Model_(overfitting)	0.8393	0.008	0.8483	0.010	0.8354	0.009	0.8124	0.008	1.2149
Model_(L1L2regularization)	**0.8536**	**0.001**	**0.8532**	**0.002**	**0.8321**	**0.001**	**0.8331**	**0.000**	**0.7356**
Model_(Early stopping)	0.8393	0.002	0.8125	0.003	0.8392	0.002	0.8214	0.001	0.7049
Model_(reduced complexity)	0.7714	0.001	0.7612	0.003	0.7714	0.003	0.7945	0.002	0.8573
Model(bigger training epoch)	0.8143	0.002	0.8211	0.010	0.8043	0.004	0.7945	0.003	0.9447

SD = Standard Deviation.

Table 6
Results for Precision, Recall, and F1-Measure for Model_(L1L2 regularization).

Class	Precision		Recall		F1-Measure	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
Agglutinated	0.983	0.002	1.00	0.000	0.986	0.001
Brittle	1.00	0.000	0.934	0.003	0.978	0.001
Compartmentalized Brown	0.862	0.003	0.953	0.002	0.902	0.001
Compartmentalized Partially Purple	0.765	0.004	0.804	0.004	0.783	0.003
Compartmentalized Purple	0.954	0.002	0.724	0.006	0.824	0.007
Compartmentalized Slaty	0.962	0.001	0.962	0.001	0.963	0.001
Compartmentalized White	0.713	0.014	0.772	0.010	0.744	0.013
Flattened	0.942	0.005	0.945	0.005	0.942	0.005
Moldered	0.604	0.020	0.864	0.017	0.714	0.014
Plated Brown	0.893	0.003	0.633	0.021	0.742	0.013
Plated Partially Purple	0.626	0.021	0.765	0.012	0.684	0.008
Plated Purple	0.854	0.018	0.893	0.014	0.875	0.007
Purple Slaty	0.955	0.001	0.913	0.013	0.931	0.003
Plated White	0.826	0.004	0.824	0.012	0.823	0.007

SD = Standard Deviation.

learning algorithms. It was discovered that the ANN model performed relatively better than the other models when precision values were poor. While the Decision Tree and Naïve Bayes had the lowest precision values, the Agglutinated class had a precision value of 100% and 96% respectively for both models, which is nearly the same as the other models. As shown in Table 6, the Compartmentalized Partially Purple and Moldered classes have the lowest precision values for all the machine learning algorithms used. In these two classes, the Decision Tree and Naïve Bayes performed similarly, while the ANN performed slightly better. Additionally, it was observed that the SVM predicted 100% for the Plated White class and the Artificial Neural Network was second in terms of precision efficiency for that class. When evaluating the ML models based solely on precision, it was observed that the proposed ANN model performed better on the extracted features.

SVM and ANN outperformed the other algorithms in terms of the proportion of correctly classified positive patterns, as shown in Table 6. For the majority of the classes, the recall values for the Decision Tree and Naïve Bayes are similar to their precision values. The Compartmentalized Partially Purple class had a recall value of 0.5%, while the Plated class had a recall value of 12%. The Flattened class had the highest recall value of 94% for the Naïve Bayes algorithm. In certain cases, the ANN performed exceptionally well, with recall values of 100% for the Agglutinated and Flattened classes and the lowest recall value of 56% for the Plated Brown class. Overall, it was generally observed that the ANN performed extremely well on the extracted features of the cocoa cut-test dataset when evaluated using the recall metric.

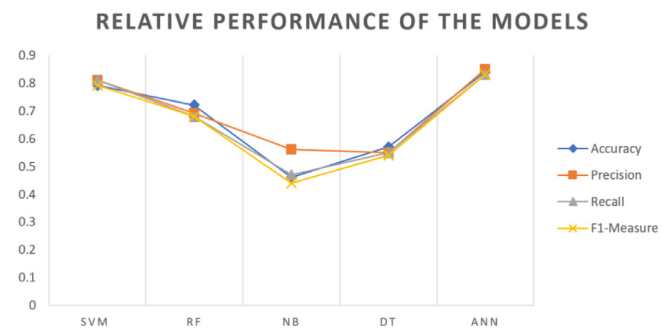


Fig. 10. The relative performance of the models.

The arithmetic mean of precision and recall, or F1 score, was also considered in the evaluation of model performance. In Table 6, the ANN and SVM algorithms appear to be better, with the ANN performing slightly better than the SVM. For almost all the classes, the Decision Tree and Naïve Bayes algorithms perform poorly. For the Compartmentalized Partially Purple class, the F1 score for the Naïve Bayes algorithm was 80.0%, while the highest F1 score was 81%. Generally, for the F1 score metric, none of the machine learning algorithms achieved a score of 100% for the 14 classes. However, it was observed that the ANN reported a score of 98% for the Agglutinated class, 97% for the Brittle class, and the lowest score of 65% for the Partially Purple class. SVM also predicted good scores for the Agglutinated and Brittle classes, with F1 scores of 95% and 93% respectively. Its lowest score was 59% for the Plated Partially Purple class. Based on this metric, the ANN and SVM models are relatively better compared to the others, with the ANN performing slightly better than the SVM on the extracted features.

The weighted average of precision, recall, and F1 score for each algorithm's results were computed and plotted with the algorithm accuracies to provide a clearer illustration of the results. As shown in Fig. 10, the ANN produced the best performance, while the Naive Bayes and Decision Tree algorithms performed poorly. The neural network and Support Vector Machine model accounted for two of the five highest-performing algorithms. In Fig. 10, the best-performing machine learning model in this study was the ANN, with an accuracy of 85%, a weighted average precision of 85%, a recall value of 83%, and an F1 score of 83%. The SVM was found to be statistically superior to the ANN when the results of the evaluation metrics were compared, with an accuracy of 79% and a weighted average of 81% for precision, 81% for recall, and 79% for F1-measure. It can also be seen in Fig. 10 that the Random Forest algorithm performed averagely in all the model evaluation metrics. Overall, it was observed that the proposed generalized ANN performed better than all four machine learning algorithms based on the performance metrics as a discriminator.

Most previous works on cocoa bean classification (Aubain et al., 2020a; Astika et al., 2010; Barbin et al., 2016; Bhagawati et al., 2016;

Lawi and Adhitya, 2018; Sukartiko et al., 2019) used fewer than seven classes or categories of beans for their experiments, making it difficult to generalize the models developed for the classification of cocoa beans. Additionally, many of these works only report the accuracy of their model, which presents an evaluation paradox on both balanced and unbalanced datasets (León-Roque et al., 2016; Antunes et al., 2019). In comparison, this work presents a thorough analysis of Precision, Recall, and F-Measure to properly evaluate the model on a comprehensive dataset. The novel feature extraction technique also demonstrated that color and texture features can be used as predictors for classification models, unlike previous results with accuracies less than 80% based on experiments with seven or fewer grades of cocoa bean fermentation (León-Roque et al., 2016; Majumdar et al., 1996; Parra et al., 2018).

The study emphasizes the importance of devising a feature extraction approach from color images and identifying the most pivotal features to accurately predict the fermentation grade of cocoa beans. The utilization of RGB images to classify cocoa samples based on their fermentation grade was deemed a valuable predictor for classification models. This approach holds the potential to be seamlessly incorporated into a computer vision system, thereby providing the cocoa industry with cutting-edge technology at an affordable cost.

5. Conclusions and future work

The cocoa industry and its associated field of food research need automated and reliable methods for evaluating cocoa bean quality. Currently, results indicate that image analysis combined with machine learning algorithms show promise as a system for classifying cocoa beans according to their grade of fermentation. In this study, we propose the use of Machine Learning techniques to facilitate the automation of the process of categorizing cocoa beans into 14 classes. A unique feature extraction method was utilized at the data preprocessing stage, which is capable of extracting both the color and texture features of the cocoa beans and using them as predictors for classification. To evaluate the effectiveness of these techniques, we employed Support Vector Machines, Decision Trees, Naive Bayes, Random Forest Machine Learning algorithms, and a proposed Artificial Neural Network (ANN) model on a cut-test cocoa dataset. Experimental results on the standard, publicly available cut-test dataset show that the novel extraction method combined with the developed Artificial Neural Network provides a more homogeneous classification rate for all grades, obtaining 85.36%, 85%, 83%, and 83% for accuracy, precision, recall, and F1 measure, respectively.

Our experiments demonstrated that the proposed color and texture extraction features, in conjunction with the developed ANN model, were optimized and able to generalize to unseen data with a high degree of accuracy. The ANN model outperformed the other machine learning models in terms of accuracy, precision, recall, and F1-measure. This is especially noteworthy given that the dataset included 14 different classes of cocoa beans, indicating that the proposed model could be applied to a variety of cocoa industries and laboratories.

Additionally, the feature extraction method we employed primarily focused on the use of the color and texture of RGB images as inputs for model training. It would be interesting to explore the effects of incorporating other features, such as bean size, in future work.

Furthermore, increasing the number of data samples would likely have a positive impact on the model. Overall, the proposed techniques demonstrated strong performance on the cut-test dataset and have the potential to serve as a reliable Computer-Aided Diagnostic tool that requires minimal computational resources and human intervention for the classification of graded fermented cocoa beans.

CRediT authorship contribution statement

Opoku Eric: Concept, Design, Analysis, Writing, or revision of the manuscript. **Rose-Mary Owusuaa Mensah Gyening:** Concept, Design, Analysis, Writing, or revision of the manuscript. **Obed Appiah:** Concept, Design, Analysis, Writing, or revision of the manuscript. **Kate Takyi:** Concept, Design, Analysis, Writing, or revision of the manuscript. **Peter Appiahene:** Concept, Design, Analysis, Writing, or revision of the manuscript.

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.

Data availability

The authors do not have permission to share data.

References

Adhitya, Y., Prakosa, S.W., Köppen, M., Leu, J.S., 2020. Feature extraction for cocoa bean digital image classification prediction for smart farming application. Agronomy 10 (11), 1642.

Afoakwa, E.O., Budu, A.S., Mensah-Brown, H., Felix, J., 2014. Changes in biochemical and physico-chemical qualities during drying of pulp preconditioned and fermented cocoa (*Theobroma cacao*) beans. J. Nutr. Health Food Sci. 2 (1), 1–8.

Alaa, H., Waleed, K., Samir, M., Mohamed, T., Sobeah, H., Salam, M.A., 2020. An intelligent approach for detecting palm tree diseases using image processing and machine learning. Int. J. Adv. Comput. Sci. Appl. 11 (7).

Anggraini, C.D., Putranto, A.W., Iqbal, Z., Firmanto, H., Al Riza, D.F., 2021. Preliminary study on the development of cocoa beans fermentation level measurement based on computer vision and artificial intelligence. In: IOP Conference Series: Earth and Environmental Science, vol. 924, (no. 1), IOP Publishing, 012019.

Antunes, F., Canuto, A., Bedregal, B., Palmeira, E., Silva, I., 2019. Supervised methods applied to the construction of a vision system for the classification of cocoa beans in the cut-test. In: Anais Do XVI Encontro Nacional de Inteligencia Artificial Computational 2019 Oct 15. SBC, pp. 72–83.

Asogwa, A.O., Orimoloye, E.O., 2021. Evaluating the performance of the artificial neural network and support vector machine for the classification of cocoa beans using colour features. J. Food Sci. Technol. 56 (3), 1.

Astika, I.W., Solahudin, M., Kurniawan, A., Wulandari, Y., 2010. Determination of cocoa bean quality with image processing and artificial neural network.

Aubain, Y., Camille, N.Z.É., Kidiyo, K., 2020a. Machine vision-based cocoa beans fermentation degree assessment. In: Advances in Smart Technologies Applications and Case Studies: Selected Papers from the First International Conference on Smart Information and Communication Technologies. SmartICT 2019, September 26-28 2019, Saidia, Morocco, Springer International Publishing, pp. 144–148.

Aubain, Y., Camille, N.Z.É., Kidiyo, K., 2020b. Machine vision-based cocoa beans fermentation degree assessment. In: Advances in Smart Technologies Applications and Case Studies: Selected Papers from the First International Conference on Smart Information and Communication Technologies. SmartICT 2019, September (2019) 26-28, Saidia, Morocco, Springer International Publishing, pp. 144–148.

Barbin, D.F., Mastelini, S.M., Jr., S.Barbon., 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.

Barbon, A.P., Barbon, Jr., S., Mantovani, R.G., Fuzyi, E.M., Peres, L.M., Bridi, A.M., 2016. Storage time prediction of pork by 38 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.

Bhagawati, K., Bhagawati, R., Jini, D., 2016. Intelligence and its application in agriculture: Techniques to deal with variations and uncertainties. Int. J. Intell. Syst. Appl. 8 (9), 56.

Board, G.C., 2018. Manual for cocoa extension in Ghana.

Campos, G.F.C., Mastelini, S.M., Aguiar, G.J., Mantovani, R.G., Melo, L.F.D., Barbon, S., 2019. Machine learning hyperparameter selection for contrast limited adaptive histogram equalization. EURASIP J. Image Video Process. 2019, 1–18.

Catsberg, C., Dommelen, K.V.G., 1990. Cocoa and chocolate. In: Food Handbook. Springer, pp. 330–334.

Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P., 2002. SMOTE: Synthetic minority over-sampling technique. J. Artificial Intelligence Res. 16, 321–357.

Dand, R., 1999. The International Cocoa Trade. Woodhead Publishing.

Dias, D.R., 2014. Agro-Industrial Uses of Cocoa By-Products. Cocoa and Coffee Fermentation. CRC Press, Boca Raton, FL.

Essah, R., Anand, D., Singh, S., 2022. An intelligent cocoa quality testing framework based on deep learning techniques. Meas.: Sensors 24, 100466.
Indra, Z., Jusman, Y., 2021. Performance of GLCM algorithm for extracting features to differentiate normal and abnormal brain images. In: IOP Conference Series: Materials Science and Engineering. vol. 1082, (no. 1), IOP Publishing, 012011.
Ismail, A., Elpeltagy, M., Zaki, M., ElDahshan, K.A., 2021b. Deepfake video detection: YOLO-face convolution recurrent approach. PeerJ Comput. Sci. 7, e730.
Ismail, A., Elpeltagy, M., Zaki, M.S., Eldahshan, K., 2022. An integrated spatiotemporal-based methodology for deepfake detection. Neural Comput. Appl. 34 (24), 21777–21791.
Jintawatsakoon, S., Inthachot, M., 2021. Classification of cocoa bean cut-test image using stacked machine learning model. In: 2021 25th International Computer Science and Engineering Conference. ICSEC, IEEE, pp. 234–239.
Kamilaris, A., Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: A survey. Comput. Electron. Agric. 147, 70–90.
Kurniawati, Y.E., 2019. Class imbalanced learning menggunakan algoritma synthetic minority over-sampling technique–nominal (SMOTE-N) pada dataset tuberculosis anak. J. Buana Inform. 10 (2), 134–143.
Lawi, A., Adhitya, Y., 2018. Classifying physical morphology of cocoa beans digital images using multiclass ensemble least-squares support vector machine. J. Phys.: Conf. Ser. 979 (1), 012029.
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 color measurement and artificial neural networks. Talanta 161, 31–39.
Liu, Z., Zhang, L., 2020. A review of failure modes, condition monitoring, and fault diagnosis methods for large-scale wind turbine bearings. Measurement 149, 107002.
Lopes, J.F., Ludwig, L., Barbin, D.F., Grossmann, M.V.E., Barbon, Jr., S., 2019. Computer vision classification of barley flour based on spatial pyramid partition ensemble. Sensors 19 (13), 2953.
Majumdar, S., Jayas, D.S., Hehn, J.L., Bulley, N.R., 1996. Classification of various grains using optical properties. Can. Agric. Eng. 38 (2), 139–144.
Manoj, I.V., Soni, H., Narendranath, S., Mashinini, P.M., Kara, F., 2022. Examination of machining parameters and prediction of cutting velocity and surface roughness using RSM and ANN using WEDM of altemp HX. Adv. Mater. Sci. Eng. 2022, 1–9.
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 (1), 60–68.
Medar, R.A., Rajpurohit, V.S., Ambekar, A.M., 2019. Sugarcane crop yield forecasting model using supervised machine learning. Int. J. Intell. Syst. Appl. 11 (8), 11.
Ohta, Y.I., Kanade, T., Sakai, T., 1980. Color information for region segmentation. Comput. Graph. Image Process. 13 (3), 222–241.
Oliveira, M.M., Cerqueira, B.V., Barbon, Jr., S., Barbin, D.F., 2021. Classification of fermented cocoa beans (cut test) using computer vision. J. Food Compos. Anal. 97, 103771.
Owusu Ansah, G., Ofori, F., Pokuah Siaw, L., Manu, J., 2018. The stake of licence buying companies (LBCs) in the promotion of quality cocoa in Ghana. Cogent Bus. Manag. 5 (1), 1560857.
Parra, P., Negrete, T., Llaguno, J., Vega, N., 2018. Computer vision techniques were applied in the estimation of the cocoa beans fermentation grade. In: 2018 IEEE ANDESCON. IEEE, pp. 1–10.
Peres, L.M., Junior, S.B., Lopes, J.F., Fuzyi, E.M., Barbon, A.P., Armangue, J.G., Bridi, A.M., 2021. Meta-recommendation of pork 2 technological quality standards. Biosyst. Eng. 210, 13–19.
Santos, F.A., Palmeira, E.S., Jesus, G.Q., 2019. Image datasets of cocoa beans for taxonomy nuances evaluation. Data Brief 27.
Savakar, D., 2012. Identification and classification of bulk fruits images using artificial neural networks. Int. J. Eng. Innov. Technol. (IJEIT) 1 (3), 35–40.
Snoek, J., Larochelle, H., Adams, R.P., 2012. Practical Bayesian optimization of machine learning algorithms. Adv. Neural Inf. Process. Syst. 25.
Sturm, B.L., 2013. Classification accuracy is not enough: On the evaluation of music genre recognition systems. J. Intell. Inf. Syst. 41 (3), 371–406.
Sukartiko, A.C., Kurniawan, M.P., Pamudji, R.A., Saragi, G.C., Bangun, B.P., Putri, D.P., 2019. Study on physical feature extraction of fermented cocoa bean with digital image processingr. In: IOP Conference Series: Earth and Environmental Science, vol. 327, (no. 1), IOP Publishing, 012020.
Van Klompenburg, T., Kassahun, A., Catal, C., 2020. Crop yield prediction using machine learning: A systematic literature review. Comput. Electron. Agric. 177, 105709.
Voora, V., Bermúdez, S., Larrea, C., 2019. Global Market Report: Cocoa. International Institute for Sustainable Development, Winnipeg, MB, Canada, p. 12.
Yudianto, M.R.A., Agustin, T., James, R.M., Rahma, F.I., Rahim, A., Utami, E., 2021. Rainfall forecasting to recommend crops varieties using moving average and naive Bayes methods. Int. J. Mod. Educ. Comput. Sci. 13 (3).
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.