RESEARCH ARTICLE

Comparing performance of classification algorithms to use for grading coffee's raw quality by using image processing techniques

Muktar Bedaso, Million Meshesha, Chala Diriba*

Bedaso M, Meshesha M, Diriba C. Comparing performance of classification algorithms to use for grading coffee's raw quality by using image processing techniques. AGBIR.2023; 39(2):491-495.

This study tries to apply digital image processing techniques towards sample Coffee raw quality value grading. More specifically, this study emphases on comparing performance of classification algorithms to use for grading coffee raw quality by using image processing methods. To ease experimentation image processing phases are followed, including image acquisition, image preprocessing (image filtering and attribute selection), image analysis (segmentation, feature extraction and classification), and image understanding for raw quality image grading. Artificial Neural Network, support vector machine and K-Nearest neighbor classifiers on each classification parameter of morphology, color and the mixture of the two has been made. Experimental outcomes confirm that Artificial Neural Network classifier generated the highest performance of 89.45% accuracy as compared to support vector machine (with 83.75%) and K-Nearest neighbor classifier (with 77.85%). Thus, suitable selection of image processing and classification techniques paves the way for higher accuracy in the higher-level process for decision making.

Key Words: *Artificial neural networks; Coffee raw quality; Support vector machine; K-Nearest neighbor; Image processing*

INTRODUCTION

Image classification is an essential task in several areas such as biometry, remote sensing, medical images and grading of agricultural products [1]. In a classic classification method, an image is captured *via* digital camera and subsequently processed. In supervised classification, first of all training took place through labeled image data set. The trained classifier prototype is used to classify new images. The unsupervised classification uses the possessions of the images to cluster them by similarity measures and these groups are known as cluster and the practice is called clustering [1]. The cluster size is decided by experts in the area. When the labeled training data set does not exist the unsupervised classification is used for labeling them. Popular classification algorithms include Bayesian, Artificial Neural Network (ANN), K-Nearest Neighbor (KNN), Conventional Neural Network (CNN) and Support Vector Machine (SVM) [2].

Computer vision is the science that advances the hypothetical and algorithmic foundation by which valuable information can be automatically mined and examined from an experimental image by means of computation made by computers [2]. The application of the computer vision to carry out duty for value examination, sorting, and automatic processes are growing in the agricultural business. This is due to compensations such as economy, accuracy, and impartiality in terms of their ability to offer numerical data with features such as size, shape, color, and texture [2]. However, these features are not characterized by an exceptional mathematical function for many agricultural products. The natural erraticism of these products makes the task of identification and classification tremendously puzzling and computationally exhaustive because of the need to have a large number of classification features [3]. The implementation of artificial neural network as arobotics decision algorithm in computer vision has apparent advantages in the classification process. There are several computer visions in classification of agriculture products by using artificial neural network [3].

The implementation of image classification technology in the sector has a several consequence to reasonable commercial activities by swelling effectiveness, to sustain trustworthiness of customer favorites and to promote the global market. In recent decades, image processing has become an inevitable area in the agricultural sector as it acts as an expert system with decision support system [4]. Input image taken in real-time is processed and altered into useful information as an output to support farmers. In short, the main purpose of image processing is to enhance the image quality for human perception and to analyze the image for autonomous machine perception [4]. Major modules in image processing are classified as image acquisition, image preprocessing, image segmentation, feature extraction and classification. The focus of this study is on the final one which is a classification algorithm performance comparison. The aim of this study is to compare the performance of classification algorithms on grading Ethiopian Coffee.

Statement of the problem

Classification algorithms play a most important role in image processing techniques. It is used to classify the features that are extracted from the image into several classes based on diverse features. But there are challenges when selecting the best performing algorithms. The comparison of ANN, SVM and KNN algorithms on the basis of efficiency largely depends on the type of data they are being used for. The two major areas of scene classification problem are: prospects and learning prototypes for semantic classes. If the images are affected due to noise, reduced quality, obstruction or background disorder, it becomes relatively a dare to detect an object in an image. This challenge gets increased whenever an image consists of several objects. There has been a steady growth in new classification procedures, techniques in recent years. Hence, there came the requisite of such an experimentation, which will benefit in picking a suitable classification process for a specific study. Accurate identification of the features present in an image is the major objective of image classification.

Classification algorithms

Recognition of the looks of things in an image from an exact set of dignified values of features of the object enables the stratification of an image into several modules with analogous features. This is the essential business of strategy of classifiers, which employs stated features of an object as its inputs, thereby producing a grouping label or value showing the correct class allocation of the object [5]. Pattern classification is an area of discipline concerned with insightful objects on the basis of information accessible about these objects. The objective is to identify objects in the image from a set of proportions of the objects. Each object is a pattern and the measured values are the features of the pattern. A set of a like entities having more or less matching features are said to go to a certain pattern class [5].

Department of Information Science Program, Jimma University, Jimma, Oromia, Ethiopia

Correspondence: *Chala Diriba, Department of Information Science Program, Jimma University, Jimma, Oromia, Ethiopia, E-mail: chala.diriba@ju.edu.et*

Received: *06-Feb-2023, Manuscript No. AGBIR-23-88993*; **Editor assigned:** *10-Feb-2023, Pre QC No. AGBIR-23-88993 (PQ)*; **Reviewed:** *24-Feb-2023, QC No. AGBIR-23-88993*; **Revised:** *03-Mar-2023, Manuscript No. AGBIR-23-88993 (R)*; **Published:** *10-Mar-2023, DOI:10.35248/0970-1907.23.39.491-495*

OPEN ACCESS

This open-access article is distributed under the terms of the Creative Commons Attribution Non-Commercial License (CC BY-NC) (http://creativecommons.org/licenses/by-nc/4.0/), which permits reuse, distribution and reproduction of the article, provided that the original work is properly cited and the reuse is restricted to noncommercial purposes. For commercial reuse, contact reprints@pulsus.com

Hence, the purpose of pattern recognition is the scheme of a classifier, an instrument which proceeds the features of objects as its input and which results in a classification or a tag or value representing to which class the object belongs. This is through on the basis of the learning set; that is, a set of objects with a known labeling. The classifier performance is typically verified by means of a set of objects free of the learning set, called the test set [6]. A number of classification approaches have been used for the recognition of patterns. Classification methods are mainly based on two types. They are supervised learning and unsupervised learning. In supervised classification, the classifier is taught with a large set of labeled training pattern trials. The term labeled pattern trials means that the set of patterns whose class associations are known in advance [6]. In unsupervised case, the classification partitions the whole data set based on some likeness principles. This results in a set of clusters, where each cluster of patterns goes to a specific class. Below both a statistical classifier and neural network classifier are described.

METHODOLOGY

Coffee beans image raw quality classification

In computer vision system image classification is a last phase in which each weird novel pattern is gave to a class. Pattern recognition is the study of in what way machines can perceive the environment, learn to differentiate patterns of attention, make sound and rational result about the classes of the patterns [7]. The ultimate mission of grouping prototype is the selection of pattern classifier or recognizer. A pattern recognizer is a method that is used to train, test and analyze a tricky based on the training and testing prototypical of the classification procedure. The classification problem that desired to discourse offers whole information about the number of grades and their tags. Hence, it is supervised. The three classifiers selected for the purpose of this study were K-Nearest Neighbor (KNN), Support Vector Machine (SVM), and Artificial Neural Network (ANN) classifiers.

The main motives for using these algorithms were their usual compatibility with homogenous data and many scholars recommended the use of these algorithms for agricultural merchandise classification. Collective characteristics of morphological and color features were used as contribution patterns to construct the prototypes. Assortment of a set of suitable input feature variables is a significant concern while trying to develop classification models by using the most fitting respective classifiers. The determination of feature variable selection is to bargain the smallest set of features that can result in adequate model performance. About 80% of the data set was allocated to the training set and 20% to the test set. Training data is the portion of the data working to actually train the network. This is normally the largest portion of the data. Test data were used to authenticate the results of a trained network by using new unknown data Coffee bean images. The analytical computations of Coffee bean feature attributes were conducted using classification algorithms for training set which is iteratively evaluated.

Artificial neural network classifier

An Artificial Neural Network (ANN) is a commonly used classification procedure which yields an effective computer-supported raw quality value classification model. There is crucial significance behind using ANN classifier that makes it prevailing systematic tools. That importance is a flexible learning algorithm, various network topology, fast learning capability and high error tolerance [8]. MATLAB R2018a which was the latest version software tool supports this ANN classification task, with the dataset improved in MATLAB format.

Statistical raw quality values mined from Coffee beans were transformed to nominal scale which made possible raw quality values as first rank (Grade I), second rank (Grade II), third rank (Grade III) and fourth rank (Grade IV), representing the data preparation method. These nominal values were used as an output column label name in the excel spreadsheet together with the associated input attribute values. The values for these nominal raw quality value output columns was then filled with the use of binary numbers, reflecting the presence or absence of the specific nominal value that represents the specific set of record in the actual dataset.

A supervised feed forward Multiple Layer Preceptors (MLP), a universal pattern classifier letting the discriminant purpose to take any shape, aided to model the classifier with 2 hidden layers. The back propagation learning rule was incorporated to calculate the shares of the errors in model building and to modify connection weight. MLP is also appropriate as the wanted reply of the outputs is known beforehand. It is one of the most commonly implemented neural network topologies. The trained results are automatically tested for the neural networks, providing an instantaneous of the network performances. The training and testing simulations in this classifier focus on confusion matrices, percent precise and performing matrices [9].

Support vector machine classifier

Support Vector Machine (SVM) is a supervised machine learning appropriate procedure which can be used for classification challenges. For this study SVM which was mostly used in classification difficulties is selected for better accuracy than the other classifiers founded on scholar's recommendation. It can be used to find an optimal hyper plane to separate different classes of input data into higher dimension features space. Also SVM has rewards of fast training techniques, even with large number of input data. In these algorithms, the plot each data item as a point in n-dimensional space (where n is the number of features used) with the value of each feature (attribute) being the value of a particular coordinates [10]. SVM classifier was used on MATLAB R2018a to categorize Coffee bean based on the quality required.

K-Nearest neighbor classifier

K-nearest neighbor is a simple algorithm that stores all obtainable cases and classifies new cases based on a correspondence measure. KNN has been used in statistical estimation and pattern recognition already in the beginning of 1970's as a non-parametric technique

A promising approach on the way to image content recognition is the practice of classification techniques to associate images with classes according to their feature's value. KNN classification algorithms adopt about the class of an image by penetrating for the K images of the training sets most analogous to the image to be classified, and by performing a class weighted frequency analysis. The k closest images are identified trusting upon a resemblance degree between images. As an alternative approach, in this study a new KNN based classification method, depend on images characterized by means of local features generated over interest points was proposed. With the use of local features and interest points, KNN classification algorithms were reviewed to consider similarity between local features of the images in the training set rather than similarity between images, opening up new occasions to examine more efficient and effective strategies. In fact, direct use of similarity between local features is generally easier to be controlled than sets of local features.

The proposed G-KNN classifier is applied for classification and related k-neighbors are chosen at each repetition for classification by using GA, the test samples are classified with these neighbors and the accuracy is calculated for different number of K values to obtain high accuracy. Hence, the computation time of K-NN is reduced from the gotten results in this method. The MATLAB image processing toolbox based implementation is done on the Coffee bean images and the classifications of these images were carried out. The k value, execution time and accuracy were calculated and tabulated.

RESULTS AND DISCUSSION

Coffee bean images that passed through image preprocessing techniques were used in the training phase. In addition, to image preprocessing Coffee bean images were segmented using histogram-based thresholding method to identify the Region of Interest (ROI). After identifying the region of interest in the Coffee bean images, useful features were extracted in order to reduce the difficulty of the computational cost of the model. Sixteen (16) different features, ten (10) morphological and six (6) color features were extracted from training the classification model.

Necessary statistical computations of each Coffee beans parameters and features were done to produces treamlined and demonstrative data. Appropriate classification techniques were applied to distinguish a given Coffee bean sample to the appropriate category using the generated data from image as input.

Experimentation

For experimentation, Artificial Neural Network classifier, Support Vector Machine and K-Nearest Neighbor classifiers were used. After distinctly testing morphological features and color features of Coffee bean images, a combination of these two features is used in this experiment. Classification was verified by using morphological features, color features and combination

of both morphological and color features. There are two basic phases of classification used in this study. Those are training and testing phases. The trained system is applied to new data to check the performance of the classification in testing phase. The classifier was designed by dividing the total dataset into training and testing dataset. From the total dataset of each grade, 80% was used for training and 20% was used for testing data. This means that, out of the total 10, 000 datasets, 116 images (with 8000 piece of Coffee beans) were used for training and 29 images (with 2000 piece of Coffee beans) were used for testing.

Artificial neural network classifier and its output

The network was trained to yield output 1000, 0100, 0010, 0001 in the correct class of the output vector for Grade I, Grade II, Grade III and Grade IV respectively. When the network was trained, the neuron number of the input layer is subject to the nominated features. The neuron numbers of hidden layers were sixteen (16) for the first hidden layer and ten (10) for the second hidden layer neurons. The neuron number of the output layer was four (4) established on the number of Coffee bean grade that were proposed for the study. When the network training was done, the network was confirmed with 20% of the total dataset. As anticipated, MATLAB version R2018a software was used as artificial neural network simulation program. There were four (4) layers in ANN classifier which are an input layer consisting of nodes/parameters for morphological and color features, the two (2) hidden layers, and an output layer node representing the nominal ideals of the raw quality value of Coffee bean images which are Grade I, Grade II, Grade III, Grade IV. The simulation was piloted on the mutual color and morphological features of sample Coffee beans. The classifier yields for the modeling dataset partitions yielded meaningfully lower values of mean square errors and higher values of correlation coefficients. Eighty eight point two percent (88.2%) of the samples are classified correctly with respect to their real raw excellence value group.

Using morphological features: In this experimentation, ten (10) morphological features of Coffee were used as input to the network and the neuron numbers of the input layers were also ten (10). The output neurons were four (4) that resembles to four (4) predefined Coffee grade well-thought-outs in this study. The network was trained by 80% of the total dataset and for computing performance of the trained network, 20% of the total dataset was used (Table 1)

Table 1

Confusion matrix of morphological features in ANN

Actual class / Predicted class	Grade-I	Grade-II	Grade-III	Grade-IV
Grade-I	391	24	0	0
Grade-II	29	460	82	18
Grade-III	10	93	518	53
Grade-IV	0	14	8	360
Total	430	531	608	431
Correctly classified (Recall)	90.93%	86.62%	85.19%	83.52
Correctly classified (Precision)	94.20%	78.20%	76.90%	94.20%

The Grade I Coffee was misclassified to the Grade II Coffee (6.7%) and Grade II Coffee was more misclassified to the Grade III Coffee (17.5%). This shows that there is a strong morphology relationship between Grade II and Grade III Coffee beans. A comparable appearance at the morphological feature of these Coffee been displays that their comparative bigger size from other beans. There is also a misclassification of Grade II and Grade III Coffee bean images to the Grade IV 2.6% and 1.4% respectively Coffees since the assembly and bean shapes of these Coffees were associated. The Grade I Coffees were not misclassified to Grade IV. Grade III and Grade IV Coffees were not misclassified to the Grade I because their morphological feature variances between them. The Grade IV Coffees were also misclassified to the Grade II and Grade III. In general, the morphological classification pattern in Artificial Neural Network classifiers was the best in performance accuracy. From the overall performance results, the overall classification accuracy was 86.45% under this experimentation. The ANN classifier yields best accuracy performances and also has compensations of compatibility with pool illumination in Coffee bean images.

Using color features: In this experimentation six (6) color features were used as input to the neural network and the neuron numbers of the input layer were also six (6). The output neurons were four (4) corresponds to the four (4) labeled Coffee grades for this study. The below Table 2 shows the summary result of artificial neural network classifier using color features. Out of the total test set of 2000 Coffee beans 59.1% were correctly classified and 40.9% were incorrectly classified (Table 2).

Table 2

Confusion matrix of color features in ANN

Actual class / Predicted class	Grade-I	Grade-II	Grade-III	Grade-IV
Grade-I	296	33	51	25
Grade-II	68	280	199	33
Grade-III	48	192	318	85
Grade-IV	18	26	40	288
Total	430	531	608	431
Correctly classified (Recall)	68.80%	52.74%	52.30%	0.668
Correctly classified (Precision)	73.10%	50.00%	50.10%	77.40%

The Grade I Coffee was misclassified to all Grade II, III and IV Coffee, but it was classified more to Grade II Coffees (16%). Grade II Coffee was more misclassified to the Grade III Coffee (36%) and Grade Coffee III is also more misclassified to the Grade II Coffee (33%). Grade IV Coffee was classified to all other grades and more misclassified to the Grade III Coffee (18%). In addition, there is a significant misclassification among each Grade using color features. There is a better classification pattern than SVM and KNN color feature classification and also a better classification performance was obtained in most Grades than the others though there is a slight color difference in each Coffee grades. This shows that there is a slight difference in color between each grade Coffee.

Using aggregated features: In this experimentation all of the three algorithms produce increased classification accuracy in terms of recall and it reflects the property of both features. The analysis result of this experiment yields good performance accuracy. But here also it can be concluded there is a great feature relation between Grade II and Grade III Coffee. The color contributes its part of misclassified Grade I to Grade IV Coffee and vice versa, which not happened in morphological feature experimentation. It was found that aggregated features of morphological and color features are good to use for developing models (Table 3).

Table 3

Confusion matrix of aggregated features in ANN

Actual class / Predicted class	Grade-I	Grade-II	Grade-III	Grade-IV
Grade-I	403	9	6	0
Grade-II	12	480	61	13
Grade-III	14	40	529	41
Grade-IV	1	2	12	377
Total	430	531	608	431
Correctly classified (Recall)	93.70%	9039.00%	87.00%	87.47%
Correctly classified (Precision)	96.40%	85.00%	84.70%	9610.00%

Support vector machine classifier

The result of training and testing SVM classifier using morphological, color and aggregated features is presented below.

Using morphological features: The same to the ANN of morphology feature experimentation, Grade II Coffee was misclassified more to Grade III Coffee (20%) and Grade III Coffee was more misclassified to Grade II Coffee (22%). Here also SVM is telling us there is a strong morphology relationship between Grade II and Grade III Coffee beans. There is also misclassification of Grade I and Grade IV Coffee bean images to Grade II (7% and 7.5% respectively). Grade IV Coffees were not misclassified to Grade I Coffees as in ANN, but Grade I was misclassified to Grade IV. In general, the morphological

classification pattern of SVM classifiers was less in performance accuracy than ANN. From the performance results, the overall classification accuracy of SVM using morphological features was 81.9% (Table 4).

Table 4

Confusion matrix of morphological features in SVM

Actual class / Predicted class	Grade-I	Grade-II	Grade-III	Grade-IV
Grade-I	367	4	3	0
Grade-II	29	430	94	32
Grade-III	27	83	488	46
Grade-IV	7	14	23	353
Total	430	531	608	431
Correctly classified (Recall)	85.34%	80.97%	80.26%	81.90%
Correctly classified (Precision)	98.12%	74.50%	76.70%	88.90%

Using color features: All grade Coffees were misclassified to each other. Here also Grade I Coffee yields better accuracy performances than other grades and Grade II Coffee attain the least accuracy performance. Grade I Coffee was misclassified to the Grade II Coffee in 15.3%, Grade III in 8% and Grade IV in 10%. Grade II Coffees were more misclassified to the Grade III Coffee in 30.13%. Grade III Coffee was also misclassified to all others Grade Coffee. Grade III Coffee was more misclassified to Grade II Coffee by 27.96% and also misclassified to the Grade I and Grade IV Coffee were 6.74% and 13.65% respectively. Grade IV Coffee is also misclassified to the Grade I, II, III and more classified to the grade III by 16.93%. In addition, there is a significant misclassification among each grade as shown in ANN color feature experimentation. There is a better classification performance was obtained in most regions, but less accuracy performance than ANN using color features (Table 5).

Table 5

Confusion matrix of color features in SVM

Actual class / Predicted class	Grade-I	Grade-II	Grade-III	Grade-IV
Grade-I	287	40	41	36
Grade-II	66	269	170	45
Grade-III	34	160	314	73
Grade-IV	43	54	83	277
Total	430	531	608	431
Correctly classified (Recall)	66.74%	50.65%	51.64%	64.26%
Correctly classified (Precision)	71.03%	49.90%	54.04%	60.61%

Using aggregated features: The Grade I Coffee accuracy performances were greater than others and Grade III Coffee yields the least accuracy performance under this experimentation. As compared to morphological and color features individually, their combination features of Coffee bean yield better performance under this experimentation. The result shows that, there was no misclassification between Grade I and Grade IV Coffee in morphological feature experimentation of ANN but here color value had contributed to change that value (Table 6).

Table 6

Confusion matrix of aggregated features in SVM

Actual class / Predicted class	Grade-I	Grade-II	Grade-III	Grade-IV
Grade-I	372	7	5	1
Grade-II	35	444	91	29
Grade-III	22	60	493	35
Grade-IV	1	20	19	366
Total	430	531	608	431
Correctly classified (Recall)	86.50%	83.60%	81.10%	84.90%
Correctly classified (Precision)	96.60%	74.12%	80.81%	90.14%

K-Nearest neighbor classifier and its output

Using morphological features: These Coffees were relatively less misclassified to Grade II because the size of these Coffee beans is small. Grade I Coffee wasn't misclassified to Grade IV Coffee and Grade II Coffee misclassified to the Grade I Coffee. Grade III and Grade IV Coffee also misclassified to Grade I. This shows that the nonexistence of the strong morphology relationship between Grade I and the left three grades (Grade II, Grade III and Grade IV) Coffee according to the result obtained from this experimentation. So KNN was perfect on classifying Grade I Coffee and also didn't misclassified other Coffee Grade to Grade I. KNN was better at classifying Grade I Coffee than SVM even if the overall accuracy of this classifier was less than that of SVM using morphological features. From the performance results, the overall grading accuracy of KNN using morphological features was 74.2%. The KNN algorithm yields the least accuracy performances as compared to ANN and SVM classifier (Table 7).

Table 7

K-Nearest Neighbor (KNN) classifier and its output

Actual class / Predicted class	Grade-I	Grade-II	Grade-III	Grade-IV
Grade-I	321	0	0	0
Grade-II	63	384	110	24
Grade-III	46	112	456	84
Grade-IV	0	35	42	323
Total	430	531	608	431
Correctly classified (Recall)	74.65%	72.31%	75.00%	74.94%
Correctly classified (Precision)	100.00%	66.10%	65.32%	80.75%

Using color features: The Grade I Coffee was misclassified more to the Grade II Coffee (13%) and Grade II Coffees were more misclassified to the Grade III Coffee (12%). Grade III Coffee was more misclassified to Grade II Coffee (14.3%) and Grade IV Coffee is also more misclassified to the Grade II Coffee (13%). All grades were misclassified to each other because there is a slight difference in color feature between each Coffee grade and there is no regular pattern regarding color feature classification. The Grade III Coffee was better in accuracy than the others and Grade IV was the least under this experimentation (Table 8).

Table 8

Confusion matrix of color features in KNN

Actual class / Predicted class	Grade-I	Grade-II	Grade-III	Grade-IV
Grade-I	302	31	16	32
Grade-II	55	376	87	56
Grade-III	46	63	436	45
Grade-IV	27	61	69	298
Total	430	531	608	431
Correctly classified (Recall)	70.23%	70.80%	71.71%	69.14%
Correctly classified (Precision)	79.26%	65.50%	73.90%	65.50%

Using aggregated features: Under this experimentation most of the images are classified under their respective classes. Under all classifiers, aggregated features produce better accuracy performances over the separate features. But, under this experimentation, one thing that can be concluded is color contribute its part on misclassifying other grade to grade I which happened in ANN. Here also aggregated features were selected as per its accuracy performance over other features (Table 9).

Table 9

Confusion matrix of aggregated features in KNN

Actual class / Predicted class	Grade-I	Grade-II	Grade-III	Grade-IV
Grade-I	340	4	9	16
Grade-II	53	413	77	27
Grade-III	34	86	473	57

Grade-IV	3	28	49	331
Total	430	531	608	431
Correctly classified (Recall)	79.10%	77.77%	77.79%	76.79%
Correctly classified (Precision)	92.10%	72.40%	72.70%	80.53%

CONCLUSION

In this study an effort is made to smear image processing for comparing classifiers used for raw Coffee bean classification. To this end, morphological and color features were extracted from a Coffee bean images by using image analysis techniques. These features are tested individually and by merging them with respect to Artificial Neural Network (ANN), Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) classifiers.

The experimental results show that morphological features have more discriminating power to classify Coffee based on their quality than color features in all of the classification algorithms used. But the classification accuracy of Coffee increases when the morphological and color features were used together. The result of the experimentation also showed that different grade of Coffees has been classified more accurately by Artificial Neural Network (ANN) than Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) classifiers. It is concluded that there is a possibility of applying classification of raw quality images of Coffee beans using computer inspection system. The major challenges during conducting this study were keeping the best quality control environment when acquiring images, extracting best features of HSB color feature and the homogeneity of Coffee bean color features. There is a great similarity in color between each of four Coffee grade samples.

RECOMMENDATIONS

The current study investigates the application of image processing for grading Ethiopian Coffee. Based on the findings of this study, the following recommendations are forwarded.

• The performance of the classification model is highly affected by Coffee image quality. To reduce the effect of noises in image, we propose conducting further experiment in applying advanced image filtering techniques.

• In this study the most widely used classification algorithms are applied to constructed classification model. Further study needs to compare the performance with other classification algorithms, such as deep learning which can have the advantage of improving both feature extraction and classification.

ACKNOWLEDGEMENT

The authors acknowledge Jimma University for allowing budget for this research as well as writing a letter to Ethiopian Commodity Exchange (ECX) where the study was conducted. Moreover, the authors are grateful to the management of the Ethiopian Commodity Exchange (ECX) in this study for providing us with the data (sample coffee beans).

REFERENCES

1. Al-Doski J, Mansorl SB, Shafri HZ. Image classification in remote sensing. Department of Civil Engineering, University Putra, Malaysia. 2013;3(10).
2. Su TL, Kung FC, Kuo YL. Application of back-propagation neural network fuzzy clustering in textile texture automatic recognition system. In2008 international conference on wavelet analysis and pattern recognition 2008; 46-49. IEEE.
3. Markou M, Singh M, Singh S. Colour Texture Analysis of Natural Scene using Neural Networks. In Proceedings of the International IEEE/INNS Joint Conference on Neural Networks, Hawaii, 2002.
4. Deenan SP, SatheeshKumar J. Image processing methods and its role in agricultural sector: A study. Int J Business Intelligents. 2014;3:366-373.
5. Acharya T, Ray AK. Image processing: principles and applications. John Wiley & Sons; 2005.
6. Bhanu B, Lin Y, Krawiec K. Evolutionary synthesis of pattern recognition systems. New York: Springer; 2005.
7. Baleker AR. Raw Quality Value Classification of Ethiopian Coffee Using Image Processing Techniques: In the case of Wollega region (Doctoral dissertation, Addis Ababa University).
8. Narendra VG, Hareesh KS. Prospects of computer vision automated grading and sorting systems in agricultural and food products for quality evaluation. Int J Comput Appl. 2010;1(4):1-9.
9. Bhavsar H, Panchal MH. A review on support vector machine for data classification. IJARCET. 2012;1(10):185-189.
10. Tobias OJ, Seara R. Image segmentation by histogram thresholding using fuzzy sets. IEEE transactions on Image Processing. 2002;11(12):1457-1465.