International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-4, February 2020

Developing Ethiopian Yirgacheffe Coffee Grading Model using a Deep Learning Classifier

J. R. Arunkumar, Tagele berihun Mengist

Abstract: - Coffee grading is the main procedure in producing homogenous local commercial fair system of pricing in the market and export. Grading coffee is a difficult task during the inspection, because it requires training and experience of the experts. In order to tackle grading difficulties in coffee producing industries and corporates have been employed and trained experts. Even if, those experts do not work effectively due to tiredness, costly, time consuming, inconsistency, bias and other factors. Digital image processing techniques based on automatically extracted features have been explored to classify Ethiopian coffee to corresponding quality grade labels. Samples of those coffee beans were taken from Yirgacheffe Coffee Farmers' Cooperative Union. On average, 228 images were taken from each of three grade values or levels (grade 1, grade 2 and grade 3). The total number of images taken was 684 containing 6138 coffee beans. To extract coffee bean features and build a classification model for grading coffee, the state of art deep learning algorithm called convolutional Neural Network was used. Base on the experimental results classification accuracy obtained with testing coffee bean images for grade 1, grade 2 and grade 3 coffee beans was 99.51%, 97.56%, and 98.04%, respectively with the overall classification accuracy of 98.38%. This shows a promising result, even if, images are captured under the challenging condition without laboratory setup, such as illumination, different resolution, shadow and orientation which affects greatly the performance of the classifier and hence they are the future research direction that needs further investigations of noise removal techniques.

Keywords: Coffee Grade, Image processing, deep learning, CNN, Yirgacheffe coffee

I. INTRODUCTION

Agriculture is the main economic source for most of the African countries. The leading industry of Ethiopia is agriculture. Much of the agricultural production is traditional by nature, but still provides a significant portion of cash crop exports [1]. This economic sector makes up 46.6% of the Growth Domestic Production (GDP) of the country and about 85% of the population gains their livelihood directly or indirectly from agricultural production [2].

Revised Manuscript Received on February 28, 2020.
***** Correspondence Author
Dr.J.R.Arunkumar*, FCSE, Arbaminch University, Ethiopia. Email: arunnote@yahoo.com
Mr. Tagele berihun, FCSE, Arbaminch University, Ethiopia. Email: barich33@gmail.com

Coffee is a beverage obtained from cherry, the fruit of the coffee plant. Coffea arabica and Coffea robusta are the two most commonly cultivated species of coffee plant having economic significance. Arabica accounts for about 70% of the world's coffee production. Robusta coffee trees represent about 30 percent of the world's market [3].

Coffee-grading and quality control are very useful procedures in encouraging, as well as enforcing good-quality coffee production, market promotion, and provision. Sorting and grading serve as a device for controlling the quality of an agricultural commodity so that buyers and sellers can do business without personally examining every lot sold. The criteria commonly used to evaluate the quality of coffee beans include bean size, color, shape, roast potential, processing method, storage period, flavor or cup quality, and the presence of defects. The size and shape of coffee berries and beans (as well as their other properties) depend on many factors such as coffee variety, planting conditions and geo-graphical zone [1].Humans use their eyes and their brains to see and visually sense the world around them. Image classification plays an important role in computer vision [2].Computer vision is concerned with the automatic extraction, analysis, and understanding of useful information from a single image or a sequence of images [3]. A deep learning algorithm which is a convolutional neural network (CNN) emerged as the model of choice for multiple reasons. CNN's make manual feature extraction unnecessary. So, there is no need to identify features that are used to classify images. The CNN works by extracting features directly from images. The relevant features are not pre-trained but learned while the network is trained on a set of images. This automated feature extraction makes deep learning models ideal for Computer Vision tasks such as object recognition and classification [4]. Therefore, the development of coffee classification system is quite useful. In this study, automatic coffee classification model is developed by using deep learning algorithm specifically Convolutional Neural Network

II. STATEMENT OF THE PROBLEM

Currently, there is no universal or indeed global coffee grading system. Different countries export different quality varieties of beans and grade them as desired using manual and mechanical procedures and approaches [12]. The existing coffee producing and processing corporates use a manual system for coffee grading or sorting, domain experts try to classify, sort or grade the coffee using some physical and chemical characteristics of coffee beans. This expert uses some of the major entities of coffee beans like appearance, texture, shape, size, and color of coffee beans which exposes the quality assessment to inconsistent results and subjectivism.

The existing coffee producing and processing corporates use a manual system for coffee grading. Domain experts try to grade the coffee using some physical and chemical characteristics of coffee beans. This expert uses some of the major entities of coffee beans like appearance, texture, shape, size, and color of coffee beans which exposes the quality assessment to inconsistent results and subjectivism.

This frequently results in expensive, time-consuming, laborious, inefficiencies, inconsistences and also prone to error, which is a serious problem in quality control, productivity, and variation of coffee export and market prices. Since there is no automated coffee classification and grading across the country, the same sample of coffee will be manually inspected and graded in different centers for the same activity with the same procedures which is costly in terms of human resource, labor, budget, and others. Attempts are made to design an automatic coffee classification to the corresponding growing region to facilitate the sorting, classification and grading of different botanical coffee bean growing areas [5, 6, 7]. These researchers were used traditional machine learning algorithms to develop and design the classification model using a limited size of datasets. Based on their findings, they suggested future research studies need to be done using state of the art technology and increasing the sample size and grade labels. Most of the classification and grading researches are done outside Ethiopia; for local and evolving grade levels, new models should be researched and even classification accuracy of existing algorithms should be verified and optimized. As to the researcher's knowledge, no studies are conducted to find a way to design such a model in the area of Ethiopian existing and evolving coffee grading values using a convolutional neural network

III. SIGNIFICANCE OF THE STUDY

Quality of coffee is determined through manual inspection which leads to error, open to confusion and misinterpretation on grading. The model of coffee bean grading can be applied in Ethiopian coffee processing industries as a base to support the domain experts during coffee bean grading. And also, it enables the expertise to save time, save effort, give a more accurate grade, and increase the quality of coffee and market prices with the support of the automatic grading model. This research study will enable agriculture experts to appreciate the importance of image processing and computer vision technologies in the field of agriculture. The research will help to achieve a high quality of coffee export due to the fact that coffee beans will be graded automatically with higher accuracy to maintain uniformity of coffee quality level and with-out finding additional cost for agricultural experts. It will reduce the cost of experts for continues manual classification of coffee beans.

IV. PROPOSED CLASSIFICATION MODEL

As shown in figure 1 the proposed CNN model contains two Layers feature extraction and classification. This model consists of convolution, pooling and classification layers combined together. The feature extractor consists of four convolution layers (C1, C2, C3, C4) with ReLU activation in be-tween them except between C3 and C4 and max pooling between each convolution operations. The classifier also consists of three fully connected layers (FC1, FC2 and F3) and dropout is included after the first two fully connected layers to prevent the problem of overfitting. Flattening layer is also included to convert the output of feature extractor in to 1D feature vectors for the classifiers. The output of the convolution and max-pooling operations are assembled into feature maps and outputs different sizes of feature maps in each layer with an input of image of size $64 \times 64 \times 3$ as shown in Table. Due to the reduction of trained classes, hardware resources that we have, and a number of images, we scaled down the number of neurons, parameters, and filters sizes.

A) Input layer: the input layer of the proposed CNN model accepts RGB images of size $64 \times 64 \times 3$ with three different classes (grade 1, grade 2 and grade 3). This layer pass raw image as input to the first convolution layer without any computation or change.

B) Convolution Layer: In the proposed model four convolutional layers are used. The first convolution (C1) filters $64 \times 64 \times 3$ input image by using 32 kernels with a size of $5 \times 5 \times 3$ and stride of 2 pixels. The second (C2) convolutional layer takes as input the output of fist convolution and pooling (if exist) of the first convolutional and filters it by using 32 kernels of size$3 \times 3 \times 32$. The third (C3) and fourth (C4) convolutional layers are connected to each other without using pooling layer. The third convolutional layer takes as an input the output of the second pooled convolutional layer and filters with 64 kernels of size$3 \times 3 \times 64$. The fourth convolutional layer has 64 kernels of size $5 \times 5 \times 64$. All the convolutional layers of the proposed model use ReLU nonlinearity as activation functions which are used to change negative values to zeros.

C) Pooling Layer: In this proposed model three max pooling operation are used after convolution layers (C1, C2, and C4). The first max pooling layer filters the output of the first convolutional layer with a filter of size 3×3 and stride1. The second max pooling layer takes as an input the output of second convolutional layer and pools by using 2×2 filters of stride1. The third max pooling layer has a filter of size 2×2 with stride2.

D) Fully Connected (FC) layer: Before this layer Flatten is used to convert the output of features extractor called feature maps in to 1D feature vectors for the classifier. In the proposed model three fully connected layers are used. The first two fully connected layers (FC1 and FC2) each with 64 nodes and ReLU activation function after both layers and use dropout of 0.5 to prevent overfitting during tanning's final dense layer (FC3) with 3(number of class labels) nodes is used with Softmax activation function for categorical, multi class classification.

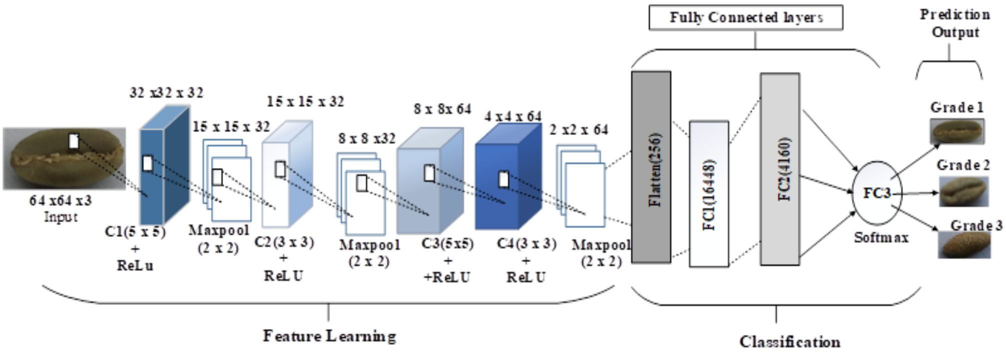

Figure 1: Proposed Coffee Grade CNN Model Architecture

E) Feature Extraction: Features extraction is a type of dimensionality reduction that represents an interesting part of the image which is given as a compact feature matrix. Image features are pattern which is seen on the raw image. For the classification of coffee beans, the color feature is one of the important parameters. During training CNN, the network learns what types of features to extract from input images given to the network model. During feature extraction convolution and polling, processes are used.

F) Convolution process: It is the process of combining two numbers or values to produces the third number. The implementation is in the form of an array matrix. As a features are extracted by convolutional layers of CNN and feature extraction is the main purpose of this layer. At the input, the image has a pixel size of 64x64x3, this shows that the pixel height and width of the image is 64 and that the image has 3 channels, namely red, green, and blue or commonly called RGB. Each pixel channel has different matrix values. The input will be convolved with the specified filter value. Filters are other blocks or cubes with smaller height and width but the same depth that is swept above the basic image or original image. The filter is used to determine what pattern to detect which is then convoluted or multiplied by the value in the input matrix, the value of each column and row in the matrix depends on the type of pattern to be detected. The filters in the convolution layer slide from left to right across the entirety of the input image to detect features.

G) Convolution Process calculation: During feature extraction, the convolution layer accepts pixel values of the input image and these values are multiplied and summed with the values of filter (set of weights) then it gives the output called feature map. Feature map is the extracted feature of the input image it contains patterns that are used to distinguish the given images. The feature map (M_i) is computed as:

$$M_i = \sum_k w_{ik} * x_k + b_i$$

Where w_{ik} is a filter of the input, x_k is the kth channel of the input image, and b is the bias term.

h) Pooling Process: Pooling is a matrix of size reduction by using pooling operations. In this pooling process, the Max-poling process is used Max-pooling is one of the common methods commonly used by researchers in deep learning researches. In a study conducted by [8], it was shown that the use of the max pooling method was more effective compared to the averaging method.

H) Classification

The classifier is placed at the end of the proposed CNN model. It is simply an artificial neural network (ANN) often referred to as a dense layer. This part of the CNN model is also called a fully connected process where data dimensions are transformed so that data can be classified linearly.

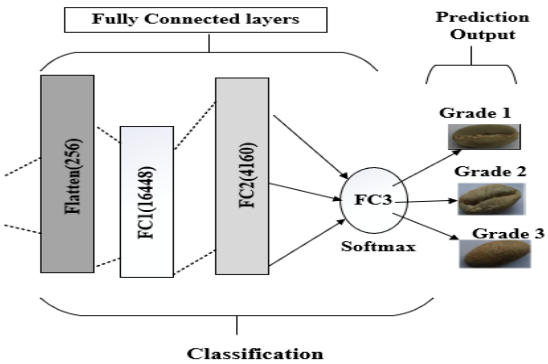

Figure 2: Classification part of proposed CNN model

As shown in Figure 2, the classifier (Softmax) requires individual features (vectors) to perform computations like any other classifier. Therefore, the output of the feature extractor (feature maps) being flattened into a 1D feature vector for the classifiers. This process is the same as the MLP (Multilayer Perceptron) Process.

This network generally uses fully connected layers where each pixel is considered a separate neuron. In this process, the dropout method is usually applied. This method aims to deactivate some edges that are connected to each neuron to avoid overfitting. After that the last process is classification. In this process, softmax function activation is used. This activation will help MLP to classify input to its target, which is into 3 coffee grade value classes (grade 1, grade 2, grade 3).
4.1

V. EXPERIMENTAL PARAMETERS AND RESULTS

A.Dataset Preparation

The original image dataset consists of three main folders (i.e., training, testing, and validation folders) and three subfolders for each main folder containing grade 1 grade 2 and grade 3 coffee bean images, respectively. A total of 6,138 coffee bean images are carefully chosen for training.

During training data augmentation methods were applied to training datasets to artificially increase the number of training sets to reduce overfitting to the model. Images of Ethiopian coffee bean with different grade values were captured from Yirgacheffe Coffee Farmers' Cooperative Union

B. Hyperparameter Choices

These hyperparameters are determined empirically according to a serious of experiments carried on the whole dataset that give the best result of classification accuracy.

Table 1: Summary of hyper parameters choices used for the proposed model

Parameter	Optimization Algorithm	Epoch	Batch size	Activation Function	Loss Function	Optimization algorithm	Learning rate
Value	SGD	150	32	Softmax	CCE	Adam	0.001

C. Experimental Result

To obtain an optimal model the experimentation was conducted in different parts for the determination of hyper parameters used for the proposed model. All parameters and hyper parameters were used by trial and error during model training in order to obtain significantly optimal results.

Experiment 1: Selection of the number of convolution Layers

Table 2: Training result using a different number of convolution Layers

Total Convolution	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
2	90.0%	93.4%	0.270	0.172
3	93.14%	94.6%	0.186	0.14
4	**94.14%**	**94.54%**	**0.158**	**0.151**
5	94.52%	90.32%	0.098	0.251

Experiment 2: Selection of Kernel size

Table 3: Training result using different kernel size (filter size)

Kernel size	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
3 x 3	91.21%	94.9%	0.228	0.173
5 x 5	92.94%	93.80%	0.184	0.168
7 x 7	84.7%	88.47%	0.325	0.278
3x3(C2&3)+5x5(C1&C4)	**94.14%**	**94.54%**	**0.158**	**0.151**

Experiment 3: Selection of training and validation dataset ratio

Table 4: Training result using different training and validation dataset ratio

Training Split	Validation Split	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
90%	10%	94.24%	96.05%	0.156	0.113
80%	**20%**	**94.14%**	**94.54%**	**0.158**	**0.151**
60%	40%	93.15%	82.78%	0.177	1.010
50%	50%	92.81%	92.47%	0.189	0.212

Experiment 4: Selection of Input Image Sizes

Table 5: Training results using different input image sizes.

Image Size	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
32 x 32	93.5%	94.11%	0.173	0.163
64 x 64	**94.14%**	**94.54%**	**0.158**	**0.151**
127 x 127	94.05%	94.5%	0.175	0.144
150 x 150	92.6%	94.2%	0.199	0.152

Experiment 5: Effect of Number of Epochs

Table 6: Training result using a different number of epochs

Bach Size	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
10	86.6%	93.3%	0.384	0.201
20	91.1%	95.4%	0.261	0.148

International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-4, February 2020

30	92.9%	94.2%	0.202	0.145
50	95.8%	96.1%	0.123	0.109
100	96.2%	98.0%	0.10	0.080
140	95.3%	93.7%	0.072	0.059
150	**96.0%**	**96.9%**	**0.134**	**0.090**

Experiment 6: Selection of the number of batch size

Table 7: Training of result using a different batch size

Bach Size	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
10	87.01%	90.02%	0.320	0.247
16	93.7%	94.2%	0.170	0.169
32	**94.14%**	**94.54%**	**0.158**	**0.151**
64	92.8%	92.4	0.1889	0.212

Experiment 7: Changing learning rate values

Table 8: Training Result using different learning rate

Learning Rate	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
0.01	33.33%	33.33%	10.746	10.742
0.001	**94.14%**	**94.54%**	**0.158**	**0.151**
0.0001	90.92%	93.65%	0.246	0.170

Experiment 8: Selection of Image Color Channel

Table 9: Training result using different color channels

Color Chanel	Training Accuracy	Validation Accuracy	Test Accuracy	Training Loss	Validation Loss	Test Loss
RGB	**94.14%**	**94.54%**	**98.38%**	**0.158**	**0.151**	**0.0499**
Gray Scale	87.13%	85.98%	91.43%	0.321	0.3706	0.281

After many experimental scenarios, the final results obtained were with training loss 0.158, training accuracy 94.14.0%, validation loss 0.151 and validation accuracy of 94.54% using parameters and hyperparameters selected during the above experimental scenarios. As shown in Figure 3 and Figure 4 below, the validation accuracy and validation loss both are in sync with training accuracy and training loss respectively. Even if the validation accuracy and loss lines are not perfectly learner, but the proposed model is not overfitting. Because the validation loss is decreasing and three is no significant gap between training and validation accuracy. .

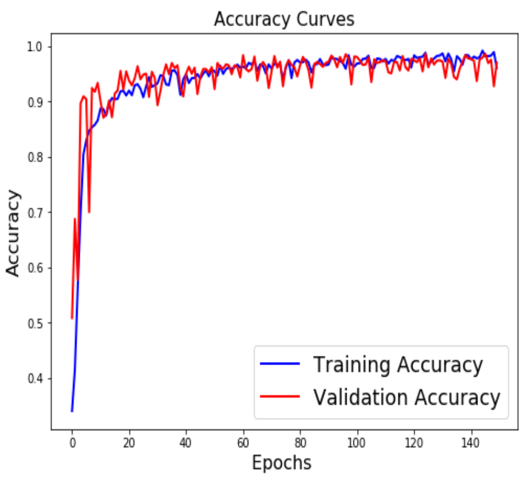

Figure 3: Training and Validation accuracy of the proposed model

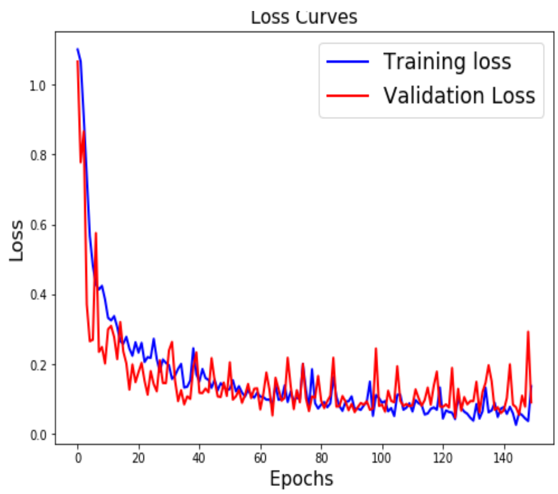

Figure 4: Training and Validation Loss of the proposed model

Retrieval Number: D1823029420/2020©BEIESP
DOI: 10.35940/ijitee.D1823.029420
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

D.Overall Classification Accuracy of the proposed Model

The overall performance of the classification model was evaluated by using a total of 1227 (409 for each grade value) unseen images samples of the test samples as summarized in below.

Table 09: Summery result of CCM

Matrices		Predicted Class		
		Grade 1	**Grade 2**	**Grade3**
Actual Class	**Grade 1**	407	2	0
	Grade 2	6	399	4
	Grade 3	1	7	401
Total		413	401	405
Correctly classified (%)		**99.51**	**97.56**	**98.04**

Generally, from a total of 1227 coffee bean test images 1209(98.38%) are correctly classified to corresponding grade value and 20 coffee images (1.62%) are incorrectly classified. Therefore, the overall accuracy produced by the proposed model with 64x64 RGB pixel image input was the accuracy of 98.38%.

The Proposed CNN model in this study was obtained using 64 x 64 x 3 input shape, learning rate value 0.001,fiiter size 3 x3 on convolution layers two and three and 5 x 5 on convolution layers one and four, and get optimal model on epoch number 150, from total of 6,138, 3684 for tranning,1,227 for validation and 1,227 for testing, optimization algorithm Adam(Adaptive Moment Estimation), pooling operation Max poling, loss function CCE(categorical cross entropy), activation function Softmax used for classification and batch size 32. Using the above hyper parameters, the classification accuracy of CNN for grade 1, grade 2 and grade 3 coffee grade values are 99.51%, 97.56%, and 98.04%, respectively. In general, from the total test sets that are used to test the model 98.38 % are correctly classified to its specific grade value and the remaining 1.62% are misclassified. Most deep learning algorithms especially computer vision for image classification problems are trained by using high performance computing machines with faster GPU, a huge number of images (in millions),and tens of millions of parameters.

Table 10: Summary of Performance & classification accuracy of the proposed model

Metrics	**Training Accuracy**	**Validation Accuracy**	**Test Accuracy**	**Training Loss**	**Validation Loss**	**Test Loss**
Value	94.14 %	94.54 %	98.38 %	0.158	0.151	0.0499

Even if, these limitations were challenges during experimentation, the modified proposed model trained from scratch has been obtained promising results with small sized networks with fewer parameters, less hardware consumption, and fewer datasets. Additionally, improved performance will be obtained if the images of the dataset are captured in a stable (laboratory setup) environmental, with proper light and proper focus.

VI. CONCLUSION

Coffee grading is the main procedure before marketing. The development of coffee classification model can support the marketing and grading process of coffee processing industries, coffee producing farmers and other stockholders and makes Ethiopia more competitive with other coffee producing countries in the market place. Therefore, in this study, an effort has been made to design an optimal model for classification of Ethiopian coffee bean grades. Images of Ethiopian coffee bean with different grade values (grade 1, grade 2 and grade 3) were captured from Yirgacheffe Coffee Farmers Cooperative Union. The images were preprocessed and augmented and followed by automatic feature extraction. Since feature extraction and classification are in the same from work in the CNN model, the features of coffee beans are extracted automatically without any human intervention. Because of this, at present the researchers can not exactly assume which coffee bean feature contributed to the higher and lower model performance. CNN has some advantages for feature extraction on the shapes of the image such as parameters of spatial filters. However, as presented on result discussion the RGB and gray scale color channels of the coffee bean have shown significant impact on classification accuracy of the model. The result was showed that the classification accuracy of the optimal model was 98.38% and 91.43% for RGB color image and gray scale images respectively. The test result showed that using test datasets the three categories of coffee grades, grade 1 grade2 and grade 3 was identified classified 99.51%, 97.56%, and 98.04%, respectively, and the overall performance was 98.38%. In the future works, the researcher will increase the number of data and also add more grade values of coffee beans and compare the proposed model with other pretrained CNN architecture using high computational resources.

REFERENCES

1. M. Tesfa, "Study on Physical Properties of Beans and Cup Quality of Selected Cultivars of Arabica Coffee (coffea arabica l.) Grown in South West Ethiopia," *MSc Thesis,* Nov.2018.
2. J. D. ,. L. Y. G. Tianmei Guo, "Simple Convolutional Neural Network on Image Classification," *IEEE 2nd International Conference on Big Data Analysis,* 2017.
3. "What is Computer Vision?," The British Machine Vision Association and Society for Pattern Recognition, [Online].Available:http://www.bmva.org/visionoverview. [Accessed 20 12 2018].
4. "Deep Learning Three Things you Should know," Math works,[Online].Available:https://ch.mathworks.com/de/discovery/deep-learning.html. [Accessed 21 12 2018].
5. A.R. Baleker, "Raw Quality Value Classification of Ethiopian CoffeeUsing Image Processing Techniques:In the case of Wollega region," MSc Thesis, Addis Ababa University,, 2011.
6. H. Minassie, "Image Analysis for Ethiopian Coffee Classification," Addis Abeba University,Ethiopia, 2008.
7. C. Pinto, J. Furukawa, H. Fukai and S. Tamura, "Classification of Green coffee bean images based on defect types using convolutional neural network (CNN)," Denpasar, Indonesia, 2017.
8. N. Sharma, V. Jain and A. Mishra, "An analysis of convolutional neural networks for image classification," *Procedia computer science,* vol. 132, 2018.

9. "The Economy of Ethiopia," [Online]. Available: https://www.worldatlas.com/articles/the-economy-of-ethiopia.html. [Accessed 9 12 2018].
10. Federal Democratic Republic of Ethiopia Central Statistical Agency (FDRECSA), "Key Findings of the 2014/2015 (2007 E.C.),Agricultural Sample Surveys," Federal Democratic Republic of Ethiopia Central Statistical Agency (FDRECSA), Adiss Abeba, 2015.
11. "Ethiopian Commodity Exchange," [Online]. Available: http://www.ecx.com.et/commodities.aspx#COFFEE. [Accessed 10 12 2018].
12. "Exporter's, The Coffee," International Trade Centre (ITC), Geneva, 2011.

AUTHORS PROFILE

Dr. J. R. Arunkumar, has excellent contribution in higher technical education in India and Abroad. He has B.E, M.Tech, and Ph.D. in the field of Computer Science and Engineering, Currently, He is working more than 7 years as an Assistant professor in faculty of Computing and Software Engineering, Institute of Technology, Arbaminch University, under the MOEFDRE, UNDP projects in Ethiopia. He published more than fifteen International and National journals like Scopus and UGC approved journals, He also obtained one patent for his work

Mr. Tagele Berihun Mengist, MSc Scholar at Arba Minch university, He has Bsc Degree from Adama science and technology university. His Research Interest is Image processing, Computer vision, Cloud computing, he has excellent contribution in higher technical education.