International Journal of Scientific Research and Engineering Development-– Volume 7 Issue 1, Jan-Feb 2024

Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Approaches, Methods, and Challenges: A Comprehensive Systematic Review on Image Processing Algorithms on Coffee Beans Grade Determination

Airah S. Cahigan*, Ma. Angela Kristiana L. Pascasio**, Edwin R. Arboleda***

*(Department of Computer and Electronics Engineering, Cavite State University, Indang, 4122, Cavite, Philippines
Email: mrs5.cahiganas@gmail.com)

** (Department of Computer and Electronics Engineering, Cavite State University, Indang, 4122, Cavite, Philippines
Email: mrs17.pascasiomakl@gmail.com)

*** (Department of Computer and Electronics Engineering, Cavite State University, Indang, 4122, Cavite, Philippines
Email: edwin.r.arboleda@cvsu.edu.ph)

--************************----------------------------------

Abstract:

The purpose of this literature review is to investigate the current state of image processing algorithms for determining the grade of coffee beans. Following a thorough, comprehensive review, screening, and assessment of the available literature, it was determined that only a total of eight scholarly papers met the inclusion criteria for this study. These papers were selected based on their thorough consideration of coffee bean grading through the implementation of image processing techniques and algorithms.The review revealed that a number of image processing algorithms were utilized to determine various parameters that can determine the coffee bean grade. These algorithms include Canny Edge Algorithm, Back-Propagation Neural Network, Convolutional Neural Network (CNN), RGB Color Model, Computer Vision algorithms, Support Vector Machine-Based Digital Image Processing, Local Binary Pattern (LBP), Random Forest, and K Nearest Neighbor (KNN). However, this review also highlights the research gap regarding the use of image processing algorithms in coffee bean grading.

Keywords **—Coffee bean grading, Image processing, Grade determination, Machine learning**

--************************----------------------------------

I. INTRODUCTION

The coffee business is a cornerstone of the worldwide economic landscape, constituting one of the most crucial commercial value chains spanning continents [1]. With its roots deeply embedded in various regions, the industry is essential in shaping local economies and international trade. Coffee, revered as one of the most consumed beverages globally [2], transcends cultural boundaries and societal norms. Its widespread popularity reflects individuals' diverse tastes and preferences and underscores the economic significance of its cultivation, production, and distribution. Coffee is a highly valued and widely available commodity that acts as a unifying factor amongst communities globally.

The impact of coffee bean grading on quality evaluation and market value emphasizes the importance of this process. The Specialty Coffee Association of America (SCAA) has significantly influenced the grading and evaluation of coffee, mainly through the development of the "Q" system for arabica and robusta coffees [3]. This system has been further enhanced by computer vision technology for the physical quality evaluation of green coffee beans, achieving a high level of accuracy in defect detection [4]. This technological advancement ensures a standardized and precise approach to evaluating coffee quality. However, the consistency of sensory analysis by Q-graders, particularly about the perception of coffee attributes, remains a challenge [5]. To address this, Sittipod [6] identified specific compounds that positively impact coffee quality, providing a potential avenue for standardizing the sensory evaluation process. Together, these scholarly contributions testify to the instrumental role of coffee bean grading methodologies in upholding quality standards and enhancing market competitiveness within the dynamic coffee industry.

The need for accurate and effective grading techniques in the coffee industry is the subject of the research problem. Various markets demand distinct bean size distributions and defect contents related to appearance and, most crucially, quality as a crucial part of quality control. The demand for accuracy in grading methods stems from the intricate nature of coffee, where distinct flavours, aromas, and characteristics contribute to its diverse profile. Standardized grading procedures may improve the openness and dependability of direct trade ties since the globalization of the coffee trade highlights the importance of direct trade in fostering quality and distinction in the specialty coffee market. Addressing this research issue is critical to improving coffee's overall quality. It is essential for economic stakeholders ranging from farmers and producers to merchants and consumers who rely on correct grading for informed decision-making and long-term market practices.

A. *Purpose of the Literature Review*

A detailed systematic evaluation is necessary to investigate and comprehend the state of image processing algorithms for determining the grade of coffee beans thoroughly. This kind of review seeks to summarize the results of previous studies, point out knowledge gaps, and offer a comprehensive picture of the potential problems in the field.

The aim is to perform an in-depth analysis of several publications and related techniques, clarify the particular strategies used in each group, and measure the occurrence of these strategies in the literature. This methodical approach offers a thorough grasp of the wide range of image processing techniques available for determining the grade of coffee beans.

The culmination of these investigations emphasizes how important sophisticated image processing methods are to addressing the complex issues surrounding coffee bean grading. Integrating data mining, remote sensing, and modelling technologies can create a more complete and resilient method for precisely evaluating coffee quality. Researchers, business experts, and politicians will find this systematic valuable review as it offers insights into the existing state of knowledge, points out areas that require more research, and encourages innovation in the coffee bean grading industry.

B. *Research Questions*

The literature aims to identify the current approaches, methods, and challenges on image processing algorithms on coffee beans grade determination. Furthermore, it seeks to answer the following research questions:

a) What are the prevailing image processing algorithms used to determine coffee bean grades?

b) What methodologies have been proposed to handle these variations in coffee bean grading practices?

c) What gaps exist in the current body of literature on image processing algorithms for coffee bean grading, and what recommendations can be made for future research to address these gaps and advance the field further?

II. METHODOLOGY

A. *Inclusion Criteria*

The criteria for selecting relevant studies, articles, and papers on coffee bean grading via image processing take a diverse approach to ensure comprehensive and high-quality coverage of this specialist topic. The material's subject relevance is the main criterion, focusing on studies examining image-processing algorithms for coffee bean grading. This criterion ensures that the chosen literature aligns with the targeted research goal and adds significant knowledge to the comprehension of advanced grading techniques.

Respectable publication platforms are given priority during the selection process to maintain the accuracy and dependability of the data collected. Reputably rigorous and peer-reviewed research can be found on platforms like IEEE Xplore, Google Scholar, ScienceDirect, SpringerLink, Wiley Online Library, IET Digital Library, and Foundations and Trends in Machine Learning.

Temporal considerations are also important, with a preference for current publications to record the most recent developments and breakthroughs in image processing technologies for coffee bean grading. This criterion ensures that the literature review is up to date and accurately reflects the evolving state of research on this constantly changing subject.

B. Search Strategy

In January 2024, a thorough search was conducted, with the primary goal of investigating coffee bean grading using image processing algorithms. The search started with the keyword "coffee bean image processing", Figure 1 presents the approach and results visually, illustrating the methodical flow of the search procedure.

The extensive number of results from many databases that the search yielded showed the depth of the body of knowledge on the subject. To be more precise, results obtained from the IEEE Xplore website were 47 results, Google Scholar 71,400 results, ScienceDirect 5,077 results, SpringerLink 2,971 results, Wiley Online Library 7,270 results, IET Digital Library 7 results, and Foundations and Trends in Machine Learning an astounding 619,271 results.

A further selection was done to narrow the search and concentrate on open-access papers, producing a more focused set of results. More specifically, the quantity of results was more controllable when the search was restricted to open-access papers. The IEEE Xplore website had 46 open-access results. In contrast, Google Scholar had 88, ScienceDirect had 24, SpringerLink had 21, Wiley Online Library had 19, and IET Digital Library had 5, and Foundations and Trends in Machine Learning had 17.

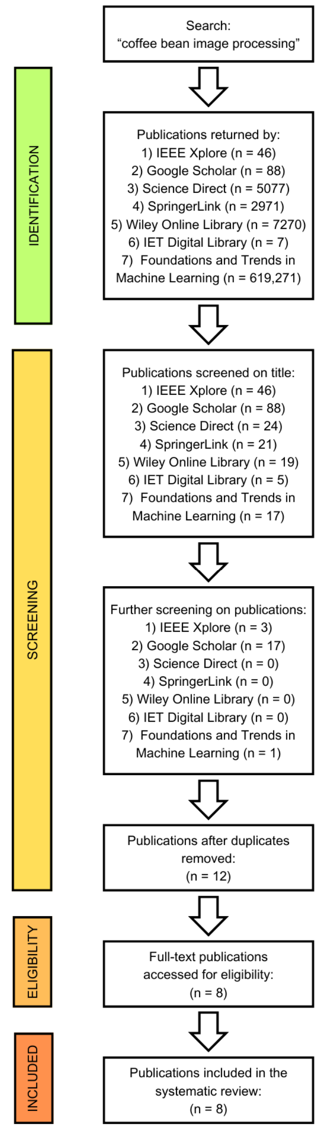

Fig. 1. Flowchart of the Publication Selection Process

C. Study Selection Process

Given the large number of suggested articles generated by our thorough searches, a meticulous screening approach was required to streamline and identify relevant material for inclusion in our systematic review. The first stage of the screening process entailed a comprehensive analysis of the publication titles. The inclusion criteria were centered around two primary elements:

a) A conspicuous connection to the overall theme of coffee and
b) A clear indication in the title is that the publication addressed the topic of coffee bean grading.

Then, a more focused examination was carried out on a subset of search pages across many platforms, such as IEEE, Google Scholar, Science Direct, SpringerLink, Wiley Online Library, IET Digital Library, and Foundations and Trends in Machine Learning. Using this method, we found and saved three IEEE articles, seventeen Google Scholar publications, and one in Foundations and Trends in Machine Learning publication. However, none of the papers from Science Direct, SpringerLink, Wiley Online Library, or IET Digital Library fit our selection criteria.After consolidating papers from these sources, duplicate records were meticulously removed to maintain the dataset's integrity. Subsequently, the full-text publications were examined in detail by three inclusion criteria:

a) The publication was written in English.
b) It dealt with digital image processing.
c) It highlighted the quality elements of sorting and grading coffee beans.

The systematic literature review only included articles that satisfied these requirements; papers that did not meet these requirements were excluded.

D. Data Extraction

Careful consideration was given to four crucial factors to guarantee a thorough examination of pertinent data when performing this systematic literature review. The main areas of interest for data extraction included:

a) the quality evaluation methodology,
b) test image features,
c) information about the ground-truth used in the studies, and
d) details about the number of benchmark methods used.

Using a methodical approach, relevant information was extracted from each publication, including essential details like:

a) The names of the contributing authors,
b) The year of publication,
c) The particular quality assessment method or methods used,
d) Specifics about the ground truth used as a reference, and
e) A breakdown of the number of benchmark techniques used in the evaluated research

This structured data extraction approach was implemented to ensure that a thorough and systematic literature examination was made possible.

III. RESULTS AND DISCUSSION

The following section is comprised of four subsections. Subsection III-A provides an overview of the selected studies, including their publication years, utilized algorithms, models, image processing techniques, and coffee bean quality classifications. Subsection III-B covers the image acquisition process, while Subsection III-C delves into the various image processing techniques, feature extraction methods, and algorithms utilized. Section III-D presents the classification of coffee beans in various studies, and finally, Section III-E showcases the coffee bean grading classifications used.

A. Overview of Selected Studies

Upon eliminating duplicates, we pinpointed 12 distinct publications (n = 12) sourced from IEEE, Google Scholar, and Foundations and Trends in Machine Learning to serve as candidates for our literature review. After subjecting these publications to a thorough screening process to assess their eligibility, we identified 8 publications (n = 8) that met the inclusion criteria. A visual representation of this selection process can be found in Figure 1. Table 1 shows the Summary of Algorithms and Coffee Bean Grading Methods for 8 Publications.

TABLE I
SUMMARY OF ALGORITHMS AND COFFEE BEAN GRADING METHODS FOR 8 PUBLICATIONS

Authors	Publication	Algorithms	General Classification	Used Classification for Coffee Bean Grading
Balbin et al.	Grading and Profiling of Coffee Beans for International Standards Using Integrated Image Processing Algorithms and Back-Propagation Neural Network	Image Processing Algorithms and Back-Propagation Neural Network	Size Classification: Shell, Small, Medium, Large Roast Level: Cinnamon Roast, Light Roast, City Roast, Full City Roast, Vienna Roast, French Roast Type of Defects: Broken Bean, Husk Fragment, Husk Bean	Grade 1: Two or less defective bean Grade 2: Three defective beans Grade 3: Four or more defective beans
Pinto et al.	Classification of Green Coffee Bean Images Based on Defect Types Using Convolutional Neural Network (CNN)	Convolutional Neural Network (CNN)	Type of Defects: Fade, Black, Sour, Broken, Peaberry, Normal	Not indicated
Faridah et al.	Coffee Bean Grade Determination Based on Image Parameter	RGB Color Model, Statistical test by ANOVA	Not indicated	Grade I, II, III, IVA, IVB, V, and VI
García et al.	Quality and Defect Inspection of Green Coffee Beans Using a Computer Vision System	Computer vision algorithms	Type of Defects: Sour, very long-berry, Black, Broken, Small, High-quality coffee bean Quality Evaluation and Defect Type Evaluation	Not indicated
Przybył et al.	Application of Machine Learning to Assess the Quality of Food Products—Case Study: Coffee Bean	Convolutional neural network (CNN) and RGB color space model	Class of Coffee: Underdeveloped, Overdeveloped, Standard	Not indicated
Nasuli et al.	Arabica Coffee Bean Quality Identification Using Support Vector Machine Based Digital Image Processing	Support Vector Machine-Based Digital Image Processing	Good quality and Poor Quality	Not indicated
Akbar et al.	Visual Feature and Machine Learning Approach for Arabica Green Coffee Beans Grade Determination	Local Binary Pattern (LBP), Random Forest and K-Nearest Neighbor (KNN)	Not indicated	Specialty, Premium, Exchange, Below Grade, Off Grade
Thongnop et al.	Quality Sorting of Green Coffee Beans from Wet Processing by Using the Principle of Machine Learning	Not indicated	Green Cluster: A-grade; Blue Cluster: X-grade; Black Cluster: Y-grade	Grade A, X, Y

B. Image Acquisition

Image acquisition is the initial step in identifying coffee bean features [8]. In this phase, electronic signals are sent from a sensor to a device like a camera to convert them into numerical data. The quality of the image obtained depends on the lighting conditions during the acquisition process.

In a study conducted by Balbin et al. [10], a prototype was created to capture Arabica coffee beans based on their size, grade, and roast level. The prototype consisted of a wooden box with slots for each camera and lights. It captured 18 coffee beans per batch.

For their study, Pinto et al. [11] used a digital camera that was set to automatic mode with F/16, exposure time 1/60 s, ISO 200, exposure compensation 1.3, auto-focus, and placed 1 m above the bean surface. They took pictures of both the front and back sides of 6,500 beans, resulting in approximately 13,000 full-color images.

In their study, Faridah et al. [12] developed hardware that positioned the light source to produce illumination of 414.5±2.9 lux. They positioned the digital camera 5 cm parallel to the light source.

García et al. [13] obtained a total of 544 Arabica coffee grains using a Canon PowerShot SX420 camera with CDD technology. They captured RGB images with a size of 20 megapixels. They implemented uniform illumination using a matrix of diffuse LEDs.

In their study, Przybył et al. [14] used a NIKON D5100 camera equipped with a 16.2-megapixel sensor that captures digital photos at a speed of about 4 frames/second. They took 160 digital images of a selected type of coffee bean, with each image containing a single object representing a coffee bean. They calibrated the device and set the image parameters, including ISO sensitivity at 200, aperture at f/6.3, exposure time of 160 s, and focal length of 28 mm. They did not use flash during the image acquisition process to maintain illumination uniformity. They obtained 480 digital images with a resolution of 4928 × 3264 (300 dpi) and 24-bit image depth saved in .JPG format.

Another study by Nasuli et al. [15] used a digital camera 14 cm away from the sample, with the sample on a white background (a short bond paper). They used 60 samples in the process and stored the images in JPEG (Joint Photographic Expert Group) format with a size of 4160x2336.

Akbar et al. [16] used a Fujifilm X-A3 camera to capture images of green coffee beans. They took the images on a white screen in a mini photo studio with a light source, with the sampled green coffee beans captured from the top 40 cm away from the camera. The researchers captured a series of images of 900 RGB, which were categorized into five distinct defect levels, namely specialty, premium, exchange, below-grade, and off-grade. Each class consisted of 180 images, with every image containing precisely 300 grams of green coffee beans.

Thongnop et al. [17] did not mention any image acquisition process by expressing the configurations of the camera used and the setup to where the samples were obtained. However, Thongnop et al. [17] mentioned that a total of 20 datasets of green coffee beans, including size, color, and characteristics were used in their study.

C. Image Processing, Feature Extraction and Algorithms

The prototype developed by Balbin et al. [10] efficiently processes 18 beans simultaneously, utilizing image masking to isolate subjects and set the background. This process was divided into three features: size, grade, and roast level, which were extracted and analyzed by the Back-Propagation Neural Network with training data as the base. Cropping was also utilized to manage the high volume of objects processed. The study implemented the Canny Edge Algorithm to remove noise and minimize information loss, while the K-Mean shift technique was used to eliminate unwanted RGB colors. This process ensured greater accuracy and consistency in RGB analysis. The Back-Propagation Neural Network served as the "brain" of the prototype, comparing extracted attributes to the training data and determining the closest match for size, grade, and roast level classification of the coffee bean batch.

In the study performed by Pinto et al. [11], the images were manually labeled into six categories: fade, black, sour, broken, peaberry, and no defect. To ensure consistency, all images were resized to 256 × 256 pixels. To analyze the features of the images, a Convolutional Neural Network (CNN) was utilized, with the convolution layer responsible for extracting spatial filters.

Faridah et al. [12] opted to use an RGB model for the color model in image processing techniques, and the mean of each color component intensity of pixels in each coffee bean sample image was sorted and called Rmean, Gmean, and Bmean. Texture features such as entropy, energy, contrast, and homogeneity were used to obtain the textural aspect of the images, which can be used for segmentation, classification, and image interpretation.

Entropy: Entropy feature was utilized to measure the degree of randomness in the intensity distribution.

$$Entropy = -\sum_{i1}\sum_{i2} p(i_1, i_2) \log p(i_1, i_2) \quad (1)$$

Energy: Energy is used to determine the concentration of intensity pairs in a co-occurrence matrix.

$$Energy = \sum_{i1}\sum_{i2} p^2 (i_1, i_2) \quad (2)$$

Contrast: Contrast is used to measure the difference in the strength of intensity between two or more areas in an image.

$$Contrast = \sum_{i1}\sum_{i2} (i_1 - i_2)^2 p(i_1, i_2) \quad (3)$$

Homogenity: Homogenity is a measure of the uniformity of the intensity variation within an image, and is the opposite of contrast.

$$Homogenity = \sum_{i1}\sum_{i2} \frac{p(i_1, i_2)}{1 + |i_1 - i_2|} \quad (4)$$

The symbol p represents probability in the given equations. Its value ranges from zero to one, indicating the element's significance within the co-occurrence matrix, while i_1 and i_2 represent the neighbouring intensity pair's strength. In the co-occurrence matrix, this neighbour pair serves as the row and column numbers, respectively. Additionally, ANOVA was utilized as a statistical test to determine whether there are any variations in the mean value of each image parameter.

García et al. [13] divided the image processing stage into two stages: image pre-processing and image processing. In the pre-processing stage, every photo underwent a Gaussian filter to eliminate or reduce any Gaussian noise present in the photos. Image segmentation was also employed, which involves dividing an image into several parts to extract specific objects of interest from the background. Thresholding methods were used to locate the coffee beans in the digital images. The study used HSV (Hue, Saturation, and Value) and LUV (Luminance and UV chromaticity coordinates) color models to identify external defects related to the color, such as sour, black, and partially black coffee beans. The feature extraction process determined four quantities to be used for developing the classification algorithm: surface area, roundness of the coffee bean, area relation or color feature, and the eccentricity of the coffee bean. The quality and defects classification were obtained using the K-Nearest Neighbour (KNN) algorithm. A software to determine the quality grade for coffee beans was developed based on the basic algorithm of Back Propagation in Artificial Neural Network (ANN).

The study done by Przybył et al. [14] used the Python libraries in preprocessing their acquired data. Specifically, they used the "rembg" library module in Python to remove image backgrounds and save the files in PNG format. Afterward, image cropping was carried out to create square images with a resolution of 678 x 678 pixels.

A convolutional neural network (CNN) was developed in the Python environment to classify coffee beans into three categories: overdeveloped, underdeveloped, and standard. The CNN consisted of input layers, such as Conv2D, MaxPooling, and GlobalAveragePooling2D, and an output layer (Dense) with three neurons, each representing a class of coffee beans. The RGB color model's digital image was reduced by calculating the average of pixel values on each channel. This method prevented overfitting and gradient fading, resulting in better CNN performance. A statistical method called ANOVA was utilized to assess if the coffee beans had consistent properties across different research categories. Afterwards, Tukey's test was utilized to compare the average values between the various coffee categories. The mean values were calculated for each coffee category and the differences between them were compared. The significance level for the Tukey test was set at $\alpha = 0.05$.

For image processing, Nasuli et al. [15] employed pre-processing methods such as image denoising, image filtering, and picture enhancement. The study revealed that CNN-based architectures like U-Net and Mask R-CNN were used for image segmentation, and both CNNs and RNNs showed promising results. To pre-process and extract features from coffee bean sample images, Nasuli et al. [15] utilized MATLAB (version 2015) to develop a computer routine algorithm. In addition, morphological features such as area, perimeter, major axis length, minor axis length, aspect ratio, circularity, roundness, and ferret diameter were also extracted in this study. Various SVM variations were investigated, including linear SVM, polynomial SVM, and radial basis function (RBF) SVM, with the benefits and drawbacks of each highlighted.

Akbar et al. [16] conducted a study where they used a technique to extract features from the color histogram. This technique searched for different values such as the mean (red, green, and blue), the median (red, green, and blue), and the modes (red, green, and blue). To analyze grayscale and invariant rotation textures, researchers used a method called Local Binary Pattern (LBP). This method encoded the difference between a central pixel and its neighbor to characterize the spatial structure of the image patch. To classify objects, they used a Random Forest Classifier and a K Nearest Neighbors (KNN) Classifier based on the objects' distance from the training examples in the feature space.

Thongnop et al. [17] did not provide details of the image processing techniques used to evaluate the samples. It only indicated that Python was the main principle in machine learning.

D. Classification of Coffee Beans

Various pre-parameters were employed to evaluate the grade of coffee bean samples before classifying them according to their grade.

In one study conducted by Balbin et al. [10], coffee beans were classified according to size, roast level, and defects based on the Bureau of Agriculture and Fisheries Standards and the Philippine Coffee Board Inc. The size classifications included Shell, Small, Medium, and Large. The roast levels included Cinnamon Roast, Light Roast, City Roast, Full City Roast, Vienna Roast, and French Roast. The types of defects were Broken Bean, Husk Fragment, and Husk Bean. These parameters were used to obtain the Classification of Grade in the study (Grade 1, Grade 2, and Grade 3). The confusion matrix accuracy showed that the values were more than 87% accurate.

In another study performed by Pinto et al. [11], parameters such as Fade, Black, Sour, Broken, and Peaberry were used as main classifications in coffee beans. The Black bean had the highest accuracy of 98.75%, followed by the Sour bean of 92.93% accuracy. The broken bean had the lowest accuracy of 78.07%. The high accuracy of the Black and Sour bean indicated that the developed CNN model was able to extract the image feature for those labels effectively.

A study done by García et al. [13] inspected 946 coffee beans to test the accuracy of machine vision when carrying out quality and defect inspection. The results of the study indicate that the machine vision system performs with a high degree of accuracy in classifying defect types in black coffee beans. Specifically, the k-nearest neighbour (KNN) algorithm is found to be effective, with an accuracy rate of 97.04%. The color feature of black coffee beans is determined to be particularly distinctive, contributing to the system's high accuracy in classification. However, the system's accuracy in identifying the sour defect type is found to be relatively lower, at 92.12%. The study also demonstrates a high degree of accuracy in identifying very long berry coffee beans, with a rate of 98.05%. This is attributed to the use of two quantities, eccentricity and roundness, which distinguish these beans from all others in the plot. Other defect types are identified using only one quantity. Finally, it is worth noting that these results were obtained with a k-value of 10.

Another study conducted by Przybył et al. [14] classified coffee beans into three classes: Underdeveloped, Overdeveloped, and Standard. The study used two color models, the RGB color space model, and the Lab* color model, to distinguish three distinct categories of coffee beans. The results of the study demonstrated that utilizing ANOVA for color statistical analysis aided in identifying three categories of coffee beans in the RGB color space model, which was beneficial in recognizing coffee defects acquired during the coffee roasting process. However, distinguishing between the quality classes of under-roasted and appropriately roasted coffee beans using the L*a*b* model posed a challenge, although it effectively differentiated between the over-roasted coffee bean class.

In another study done by Nasuli et al. [15], a method was suggested to determine the value quality of Arabica coffee beans using an SVM classifier which categorizes the coffee beans based on the extracted features in order to verify the efficacy of the suggested strategy. The findings of this study show that with the image processing method, 85% were classified as good quality while 15% were classified as poor quality, garnering an overall 95% accuracy rate.

Finally, in a study done by Thongnop et al. [17], Thai coffee beans grading standards were used to grade the coffee beans, where only three grades, Grade A, X, and Y, were used to classify the coffee beans' grade. The research did not specify the accurate findings of each parameter, but it indicated that an accuracy of 85% and 15% error were obtained.

E. Coffee Bean Grade Determination

Out of the eight literature and studies collected and reviewed, only two studies followed the grading parameters of the Specialty Coffee Association. The remaining studies, Faridah et al. [12] and Akbar et al. [16], did not mention a general classification or pre-parameters in grading the coffee beans. Instead, they directly tested and graded the sample coffee beans.

One of the studies, conducted by Balbin et al. [10], used three grades of classification, where Grade 1 indicates that there are two or fewer defective beans, Grade 2 indicates that there are three defective beans, and Grade 3 indicates that there are four or more defective beans. However, the study did not include the succeeding grades established by the Specialty Coffee Association.

The study conducted by Faridah et al. [12] employed a visual sensor to determine the quality of the coffee beans based on image parameters such as energy, entropy, contrast, homogeneity, Rmean, Gmean, and Bmean. A statistical test by the ANOVA method was used for image parameters in each grade, which obtained a confidence level of 95%. The accuracy of system testing for the coffee beans of grade I, II, III, IVA, IVB, V, and VI had the value of 100%, 80%, 60%, 40%, 100%, 40%, and 100%, respectively.

On the other hand, Akbar et al. [16] opted to perform various models and algorithms s uch as Local Binary Pattern (LBP), Random Forest, and K-Nearest Neighbor (KNN), to obtain the grade determination process for the Arabica coffee beans. The Random Forest grading method has a concatenate feature that produces five different grades, namely specialty, premium, exchange, below grade, and off grade. The accuracy testing for each of these grades resulted in 97%, 71%, 79%, 96%, and 100% accuracy, respectively. The study conducted experiments using Random Forest and KNN, which produced average accuracy rates of 86.56% and 80.8%, respectively.

IV. CONCLUSIONS

This literature review aimed to thoroughly investigate the current state of image processing algorithms for determining the grade of coffee beans. The initial search for sources included IEEE Xplore, Google Scholar, ScienceDirect, SpringerLink, Wiley Online Library, IET Digital Library, and Foundations and Trends in Machine Learning. This resulted in a total of 46, 87, 5077, 2971, 7270, 7, and 619,271 publication results, respectively. After a thorough review, screening, and assessment of the publication results, only 8 publications that fully considered coffee bean grading using various image processing techniques and algorithms were included in the literature review.

The 8 publications included in the literature review revealed that various image processing algorithms such as Canny Edge Algorithm, Back-Propagation Neural Network, Convolutional Neural Network (CNN), RGB Color Model, Computer vision algorithms, Support Vector Machine-Based Digital Image Processing, Local Binary Pattern (LBP), Random Forest, and K Nearest Neighbor (KNN) were utilized to determine various parameters that can determine the coffee bean grade. The literature review identified that 6 publications developed their parameters in identifying the grade and quality of the coffee beans. Parameters such as roast levels, defect types, and coffee classes were formed to determine the coffee beans' quality. The remaining 2 publications, specifically by Faridah et al. [12] and Akbar et al. [16], directly used image processing algorithms to determine the coffee beans' grade. Faridah et al. [12] used RGB Color Models and Statistical Test by ANOVA to determine the coffee beans' grade (Grade I, II, III, IVA, IVB, V, and VI), while Akbar et al. [16] used Local Binary Pattern (LBP), Random Forest, and K Nearest Neighbor (KNN) to classify coffee beans in the same grading system. This grading system was acquired from the Specialty Coffee Association of America (SCAA).

This literature review highlights the research gap regarding the use of image processing algorithms in coffee beans grading. Out of all the initial results obtained, only 8 publications were relevant to this study, and only 2 publications directly used image processing algorithms in determining the coffee beans' grade. In order to advance the field of coffee bean grading, it is imperative that researchers explore the application of various image processing techniques. Given the current lack of studies on the use of image processing algorithms for coffee bean grading, there is a pressing need to investigate this area further. As such, future research endeavors should focus on developing and testing novel image processing methods that can accurately and efficiently grade coffee beans.

ACKNOWLEDGMENT

The authors gratefully acknowledge the professors and faculty members of DCEE for evaluating and proofreading this paper.

REFERENCES

[1] J. A. Barreto Peixoto, J. F. Silva, M. B. P. P. Oliveira, and R. C. Alves, "Sustainability issues along the coffee chain: From the field to the cup," *Comprehensive Reviews in Food Science and Food Safety*, vol. 22, no. 1, Dec. 2022, doi: https://doi.org/10.1111/1541-4337.13069.

[2] S. I. Mussatto, E. M. S. Machado, S. Martins, and J. A. Teixeira, "Production, Composition, and Application of Coffee and Its Industrial Residues," *Food and Bioprocess Technology*, vol. 4, no. 5, pp. 661–672, Mar. 2011, doi: https://doi.org/10.1007/s11947-011-0565-z.

[3] T. R. Lingle and S. N. Menon, "Cupping and Grading—Discovering Character and Quality," *The Craft and Science of Coffee*, pp. 181–203, 2017, doi: https://doi.org/10.1016/b978-0-12-803520-7.00008-6.

[4] C. E. Portugal-Zambrano, J. C. Gutierrez-Caceres, J. Ramirez-Ticona, and C. A. Beltran-Castanon, "Computer vision grading system for physical quality evaluation of green coffee beans," Oct. 2016, doi: https://doi.org/10.1109/clei.2016.7833383.

[5] L. L. Pereira, W. S. Cardoso, R. C. Guarçoni, A. F. A. da Fonseca, T. R. Moreira, and C. S. ten Caten, "The consistency in the sensory analysis of coffees using Q-graders," *European Food Research and Technology*, vol. 243, no. 9, pp. 1545–1554, Mar. 2017, doi: https://doi.org/10.1007/s00217-017-2863-9.

[6] SichayaSittipod, E. L. Schwartz, LaurianneParavisini, and D. G. Peterson, "Identification of flavor modulating compounds that positively impact coffee quality," *Food Chemistry*, vol. 301, pp. 125250–125250, Dec. 2019, doi: https://doi.org/10.1016/j.foodchem.2019.125250.

[7] C. Henrique, "CHAPTER 6. Coffee Grading and Marketing," *Royal Society of Chemistry eBooks*, pp. 202–212, Jan. 2019, doi: https://doi.org/10.1039/9781782622437-00202.

[8] B. W. Ligar, "Computer Vision for Identifying and Classifying Green Coffee Beans: A Review," *JurnalCoreIT: Jurnal Hasil PenelitianIlmuKomputer dan TeknologiInformasi*, vol. 8, no. 1, p. 11, Jun. 2022, doi: https://doi.org/10.24014/coreit.v8i1.17450.

[9] D. I. Patrício and R. Rieder, "Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review," *Computers and Electronics in Agriculture*, vol. 153, pp. 69–81, Oct. 2018, doi: https://doi.org/10.1016/j.compag.2018.08.001.

[10] J. R. Balbin, C. D. Del, V. Julius, and R. F. Quiambao, "Grading and Profiling of Coffee Beans for International Standards Using Integrated Image Processing Algorithms and Back-Propagation Neural Network," Dec. 2020, doi: https://doi.org/10.1109/hnicem51456.2020.9400086.

[11] C. Pinto, J. Furukawa, H. Fukai, and S. Tamura, "Classification of Green Coffee Bean Images Based on Defect Types Using Convolutional Neural Network (CNN)," *Dimensions*, 2017. https://app.dimensions.ai/details/publication/pub.1093594562

[12] F. Faridah, G. O. F. Parikesit, and F. Ferdiansjah, "Coffee Bean Grade Determination Based on Image Parameter," *TELKOMNIKA Telecommunication Computing Electronics and Control*, vol. 9, no. 3, pp. 547–547, Dec. 2011, doi: https://doi.org/10.12928/telkomnika.v9i3.747.

[13] M. García, J. E. Candelo-Becerra, and F. E. Hoyos, "Quality and Defect Inspection of Green Coffee Beans Using a Computer Vision System," *Applied Sciences*, vol. 9, no. 19, p. 4195, Oct. 2019, doi: https://doi.org/10.3390/app9194195.

[14] K. Przybył*et al.*, "Application of Machine Learning to Assess the Quality of Food Products—Case Study: Coffee Bean," *Applied Sciences*, vol. 13, no. 19, p. 10786, Jan. 2023, doi: https://doi.org/10.3390/app131910786.

[15] J. A. Nasuli, J. P. Lumbis, and E. R. Arboleda, "Arabica Coffee Bean Quality Identification Using Support Vector Machine-Based Digital Image Processing," *International Journal of Advanced Research and Publications*, Jun. 2023, doi: http://dx.doi.org/10.13140/RG.2.2.19015.47522.

[16] M. Akbar, E. Rachmawati, and F. Sthevanie, "Visual Feature and Machine Learning Approach for Arabica Green Coffee Beans Grade Determination," *Proceedings of the 6th International Conference on Communication and Information Processing*, Nov. 2020, doi: https://doi.org/10.1145/3442555.3442571.

[17] T. Thongnop, T. Perpaman, P. Kansiri, W. Nuchda, and S. Peungsungwan, "Quality Sorting of Green Coffee Beans from Wet Processing by Using The Principle of Machine Learning," *ASEAN Journal of Science and Engineering*, Dec. 2021, Accessed: Feb. 03, 2024. [Online]. Available: https://download.garuda.kemdikbud.go.id/article.php?article=2431053&val=23245&title=Quality%20Sorting%20of%20Green%20Coffee%20Beans%20from%20Wet%20Processing%20by%20Using%20The%20Principle%20of%20Machine%20Learning

[18] W. Man, Y. Ji, and Z.-Y. Zhang, "Image classification based on improved random forest algorithm," *2018 IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*, Apr. 2018, doi: https://doi.org/10.1109/icccbda.2018.8386540.

[19] E. R. Arboleda, A. C. Fajardo, and R. P. Medina, "Classification of coffee bean species using image processing, artificial neural network and K nearest neighbors," *IEEE Xplore*, May 01, 2018. https://ieeexplore.ieee.org/abstract/document/8376326 (accessed Dec. 23, 2023).

[20] J. T. Yamada *et al.*, "Use of factorial design with significant interaction and Tukey's Test in agricultural and environmental experiments," *Revista de Gestão e Secretariado*, vol. 14, no. 6, pp. 9815–9828, Jun. 2023, doi: https://doi.org/10.7769/gesec.v14i6.2335