



Coffee Bean Defect Detection with Deep Learning: Khawlani Coffee Case Study from Jizan, Saudi Arabia

Mohammed Imran Basheer Ahmed¹, Atta-ur Rahman², Farhan Ali^{3*}, Mustafa Youldash¹, Sara Althubaiti¹,
Azzah S. Alghamdi¹, Deena S. Alqahtani¹, Rahaf Al-Shammari¹, Rawan K. Alharbi¹, Razan A. Alomari¹,
Zahraa N. Alqatari¹, Maqsood Mahmud⁴, Jamal Alhiyafi⁵, Mohammed Gollapalli⁶

¹ Department of Computer Engineering, College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, Dammam 31441, Saudi Arabia

² Department of Computer Science, College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, Dammam 31441, Saudi Arabia

³ College of Electronics and Information Engineering, Shenzhen University, Shenzhen 518060, China

⁴ School of Computing, Ulster University, Belfast BT15, United Kingdom

⁵ Department of Computer Science, Kettering University, Flint MI 48504, USA

⁶ Department of Information Technology & Engineering, Sydney Metropolitan Institute of Technology, Sydney 2000, Australia

Corresponding Author Email: farhanali@szu.edu.cn

Copyright: ©2025 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.420515>

ABSTRACT

Received: 11 March 2025

Revised: 28 May 2025

Accepted: 18 September 2025

Available online: 31 October 2025

Keywords:

coffee bean defect detection, Khawlani coffee, computer vision, real-time detection, deep learning

Quality coffee is essential to both consumer happiness and business success, since it is one of the most consumed drinks on a global scale. Saudi Arabia recognizes this importance and has invested in the coffee industry, which includes many consumers and producers. The Jizan region is well-known for its potential in coffee cultivation, featuring numerous farmers who utilize traditional farming techniques. These approaches are tedious, error-prone, and labor-intensive. To address these challenges, this paper introduces a real-time coffee bean defect detection system that leverages artificial intelligence. The proposed system employs computer vision algorithms to automate the defect inspection process. In this study, four types of coffee beans—normal, burnt, immature, and broken—were captured using an advanced camera. The images were then annotated and exported in various label formats. Additionally, the system utilized four deep learning models: Faster R-CNN, YOLOv5, YOLOv8, and YOLOv9. After being trained on a comprehensive dataset of annotated coffee beans, the results showed successful defect detection, with the highest mean Average Precision at 50% (mAP50) achieved at 99.1% for YOLOv5 and 99.5% for YOLOv8. A connected webcam was also employed to facilitate real-time detection, allowing for immediate inference of results and optimization of the coffee inspection process.

1. INTRODUCTION

The coffee industry is enormous and complex, encompassing numerous countries and cultures. It plays a vital role in daily consumer habits and the agricultural economy. The entire journey of the coffee bean, from cultivation and production to consumption, is significantly important in the coffee sector. The beans used to make the coffee are very significant, as it directly influences the flavor, aroma, and taste of the coffee, as well as its market price.

In the Kingdom of Saudi Arabia (KSA), coffee has a deep-rooted tradition and is a key component of social interactions, symbolizing hospitality and generosity within Saudi society. The kingdom is one of the top coffee-consuming nations in the world, with annual consumption exceeding 80,000 tons. To address this demand, the Ministry of Environment, Water, and Agriculture (MEWA) is currently working to enhance domestic coffee bean production and achieve greater self-sufficiency [1]. The Jizan region is a coffee-rich area known

for its Khawlani coffee, producing approximately 270 tons each year, which accounts for 85% of Saudi Arabia's overall coffee output [2]. The region has announced ongoing efforts to establish a geographical indication for Jizan's coffee. This initiative aims to help farmers increase their profitability, improve production efficiency, and enhance product quality [3]. Nevertheless, the reliance on conventional agricultural techniques for coffee beans poses challenges for farmers who depend on manual labor, including inspecting coffee beans for defects. This dependence can lead to reduced production efficiency and lower product quality [4]. To effectively address these challenges, it is essential to implement advanced technology that enhances product quality and expedites the inspection process. The proposed approach will help increase production capacity, positively impacting the economy by providing efficient and consistent solutions for the quality evaluation of coffee beans. In conjunction with Aramco's Micro-Industries Initiative, which aims to equip farmers with the necessary tools to boost their production, including smart

systems, Saudi Aramco has committed to building the National Coffee Development Center (NCDC). The center's goal is to support farmers at every stage from cultivation to production to achieve optimal practices [5]. It is also a step forward for smart agriculture in KSA.

Machine learning (ML) and deep learning (DL) have gained significant importance as promising candidates for Artificial Intelligence (AI). They have proven helpful in numerous plant disease detection problems and similar issues [6, 7]. For instance, plant leaf-based disease detection, fruit classification [6], and seed quality inspection [7] are among the various applications of ML and DL technologies in agriculture. Through the integration of AI technologies like computer vision (CV) and DL, to classify the quality of coffee beans and identify defects in them in real-time, we can develop a real-time system. This technology accurately detects and classifies a range of flaws, including broken, burned, and immature beans, enabling their exclusion from the packaging process and accelerating production. Additionally, the current study aims to considerably increase the body of knowledge in the field of CV and DL with the primary objective of developing systems that have the potential to change the coffee industry. The current study investigates the effectiveness of the Yolo Only Look Once (YOLO) family of deep learning models over a diverse and augmented dataset obtained from various sources.

The creation and growth of the Saudi coffee industry have important implications for the agricultural sector, which aligns with the Vision 2030 objectives. As the Kingdom seeks to diversify its economy and minimize its reliance on oil, the development of non-oil sectors, especially agriculture, becomes critical and the coffee industry one of the essential parts of it as the last few years Saudi Arabia has focused on farmers and how to increase the domestic production in coffee, which considers digital transformation as an essential strategy to improve numerous stages of the preparation process.

The rest of the paper is structured in the following order: Section 2 provides the related work in coffee bean defect detection. The methodology is presented in Section 3. Section 4 highlights results and discussion, while Section 5 concludes the study.

2. LITERATURE REVIEW

This section synthesizes the related work, highlighting key findings and discussing how our study contributes to this evolving research landscape. In a study by de Oliveira et al. [8], the aims of this research were to develop a computer vision system which performs Commission Internationale de l'Éclairage measurements of green coffee beans and categorizes them based on color. The sample included commercial arabica coffee in green bean form supplied by Brazilian coffee producers. They used artificial neural networks (ANN). They utilized these in the transformation model for classifying coffee beans, employing Bayes classifiers to categorize them into four groups: bluish green, green, cane green, and whitish. The Bayes classifier successfully classified all data into their anticipated categories with 100% precision. The neural network models achieved a generalization error of approximately 1.15. A paper by Portugal-Zambrano et al. [9] aimed at combining a software module with a hardware prototype for the development of a computer vision system for the physical quality assessment of green coffee beans. The

dataset comprises 1,930 images of green coffee beans. For image pre-processing, the White-Patch algorithm was employed as an image enhancement procedure, and color histograms were used as feature extractors. The model then utilized SVM classification to identify the physical defects in green coffee beans. To conclude, SVM classification achieved an overall detection accuracy of 98.8%.

Research was performed by Pinto et al. [10] to create an automated system for sorting defective coffee beans. The dataset comprises 6,500 beans and over 13,000 images of 6 different categories of coffee beans. Image pre-processing techniques were employed to isolate each bean. From that, the image samples were split into three different categories: training, validation, and testing. Each type of coffee bean defect was then categorized using the CNN model. Black beans had the highest classification accuracy at 98.75%.

The study by Arboleda et al. [11] focused on regulating the quality of coffee beans used in image processing methods. The researchers extracted RGB color components from training images to recognize normal beans, while removing black beans using an image processing method that included RGB values. The study used a robust variety of coffee from the *Coffea canephora* plant from Cavite State University, obtained from the National Coffee Research Development and Extension Center. Samples were taken from different coffee-growing towns and edited on a laptop running the trial version of MATLAB R2018a on Microsoft Windows 10 Professional. The researchers successfully identified and differentiated between normal and black coffee beans using color ranges, eliminating low-area bean values, and leaving only white marks representing normal bean positions. The proposed image processing method could correct other color quality flaws but needs more training and testing due to its limited testing on 180 samples. Huang et al. [12] established a Convolutional Neural Network (CNN) that can classify green coffee beans in real time. They selected green coffee beans and used image processing and data augmentation to get the images prepared for inclusion in a dataset. The trained CNN model was able to distinguish the difference between excellent and poor coffee beans, getting 93% of them correctly and only 0.1007% false positive rate. While the study did not mention specific limitations, future research could explore the generalizability of the findings to different coffee bean varieties and sources. Overall, their work demonstrated the potential of CNNs for automating the selection of coffee beans, improving the efficiency of specialty coffee production.

A paper published by Sarino et al. [13] describes research that uses image processing and an artificial neural network to figure out how well-roasted coffee beans are. A human inspector uses the color and texture of the coffee bean to figure out how well it is roasted. The study used four steps as research methods. The first step involved acquiring images using a camera, which would then be used to train and test the system. The second step is processing the images or generating and reading images using MATLAB. The third step is the extraction of the color feature from coffee beans. The fourth step is the ANN-based classification of the coffee beans. In the final analysis, the results indicate that the ANN classifier had a correct classification of 100%.

García et al. [14] conducted research to create an effective machine vision system for assessing the quality and identifying problems in green coffee beans. The researchers used image processing and machine learning (ML) techniques that were built onto an Arduino Mega board, but they didn't

provide many specifics about the dataset. They used the k-nearest neighbor technique to sort beans by quality and find out what kinds of defects they had. The algorithm was quite good at sorting beans by their looks. The method is more than 90% accurate overall. However, limitations include the focus on visual attributes and a lack of dataset information. Overall, the study presents a promising approach to improving coffee bean quality assessment and has implications for enhancing production and quality control.

Research by Kuo et al. [15] proposes an innovative Hough circle-assisting deep network inspection scheme (HCADIS) for detecting defects in dense coffee beans. The methodology uses data from a deep network and a feature engineering method called the Hough circle transform to get the best of both worlds when it comes to examining beans. Combines data from a deep network with the Hough circle transform, a feature engineering method, to get the best of both worlds. Due to its stability and the similarity of bean shapes to circles, the Hough circle transform was chosen. The deep network and the Hough circle transform work together through key mechanisms to inspect defective beans precisely and accurately. The HCADIS prototype is put into use, and tests on the system are carried out. The findings demonstrate that HCADIS performs superiorly across a range of metrics when inspecting defects in dense beans. This work gives industrial participants beneficial experience in developing deep-learning solutions for coffee bean products. The outcomes of the experiment demonstrate that combining deep networks with vision-based techniques can produce promising results for challenging tasks in agricultural applications. This work offers developers practical knowledge for building deep-learning solutions for coffee industry bean products.

In compliance with the researchers, Kuo et al. [16], they suggest a unique labor-efficient deep learning-based model generation approach that aims to offer effective automation, and a real-time model linked to a robotic gripper. It aims to overcome the following issues: first, to decrease the required time for labelling and removing defective coffee beans. Second, the complete classes of categorized beans defected by the SCAA (Specialty Coffee Association of America) must be inspected simultaneously, and the automation of the coffee industry should be enhanced. The dataset that was utilized is a 40-image sample photographed by a digital camera for 13 various classes of defective coffee beans, 30 images used for training, and 10 images for testing. Besides, they manipulate Data Augmentation using generative adversarial networks (GAN) techniques to solve the small size of the dataset and to label generative images to classify. Moreover, applying two further methodologies for their scheme, which are defect-sensitive inspection, observing defects, evaluating quality, and checking the model built, is effective. As a result, the offered scheme achieves a high-accuracy model for defect inspection with little labelling of sampled beans by humans. Also, a prototype scheme should be applied to ensure the effectiveness of the robotic arm system.

In a study by Akbar et al. [17] aimed to build a system to measure the quality of arabica green coffee beans using a CV technique, they put them into five classes depending on how many defects they had: specialized grade, premium grade, exchange grade, below grade, and off grade. The authors captured 900 RGB images with 5 classes of defect. The researchers applied a color histogram and local binary pattern (LBP) to determine both the texture and color characteristics of the green coffee beans. Subsequently, machine learning

classification techniques, namely Random Forest and K-Nearest Neighbor (KNN), were utilized to categorize the beans. The outcome demonstrated an accuracy of 87.87% and 80.47%, using Random Forest and KNN, respectively.

In a study by Eustaquio and Dioses [18] aimed to classify immature and mature coffee beans using image processing to extract RGB color. The authors used 100 samples of coffee bean images that were taken by an A4tech PK-835 G webcam, situated directly above the coffee beans at a height of 13.5 cm. In this study, 23 ML classifiers of MATLAB's Classification Learner App (CLA) were used, and the dataset was divided into 80 beans for training and 20 beans for testing. The result demonstrated that the Quadratic Support Vector Machine (SVM) algorithm achieved 94% accuracy with a training speed of 0.62 seconds.

Gope and Fukai [19] conducted yet another investigation with the objective of designing a system that is capable of mechanically and inexpensively identifying and sorting pea berries for coffee farmers in nations that have inadequate infrastructure. There were a total of 1900 normal coffee beans and 1438 pea berry coffee beans in the sample. Then, image pre-processing methods were used to separate each bean by making every image the same size. After that, both CNN and a standard linear SVM were utilized on desktop computers to make two distinct kinds of binary classifications. Due to this, CNN did better, with an accuracy of almost 97%.

Santos et al. [20] did another investigation to look at the form and color of coffee beans using several machine-learning methods. They got the dataset from an Epson L210 Flatbed Scanner. The collection included 635 images in all, each of a different variety of bean. The investigators utilized R and Python to write the code for the classification algorithms Random Forest (RF), Support Vector Machine (SVM), and Deep Neural Network (DNN). The research found that Deep Neural Network (DNN), Random Forest (RF), and Support Vector Machine (SVM) all had accuracies of 94.8%, 88.5%, and 94.7%, respectively.

Belan et al. [21] have developed a machine vision system (MVS) for visual quality inspection of beans. The system is hardware- and software-based, using inexpensive electromechanical materials for hardware and software. The system is designed for segmentation, classification, and defect detection. The experiments were conducted offline and online, using a database of 270 images of popular Brazilian beans. The MVS achieved success rates of 99.6%, 99.6%, and 90.0% in offline experiments, and 98.5%, 97.8%, and 85.0% in online mode. The study suggests computational methods for segmentation, classification, and identification of bean defects, such as cracked, bored, and moldy ones. The MVS represents significant advances in science, particularly in segmentation and defect detection. The success rates in offline experiments were 99.6% and 90.0%.

Janandi and Cenggoro [22] conducted a study to develop a mobile application that utilizes deep learning for automatic classification of coffee beans' quality. They collected and labeled a dataset of 160 coffee bean images based on the Indonesian National Standard grading. The study employed CNN architectures, VGG16 and ResNet-152, using transfer learning. The ResNet-152 model has the greatest accuracy of 73.3% and was used to create a working mobile app named Cafeon. The study's drawback experienced the tiny size of the dataset. Future study need to contemplate the augmentation of the dataset and the incorporation of a broader spectrum of coffee bean grades to enhance the precision and applicability

of the classification algorithm. In a nutshell the research showed that deep learning works well for classifying the quality of coffee beans. It gave small to medium-sized coffee businesses a useful way to keep an eye on and protect their buying process.

Chen et al. [23] used a push-broom infrared (VIS-NIR) hyperspectral sensor in their investigation. They established a hyperspectral algorithm that can automatically detect insect damage in coffee beans that are already damaged. Using a push-broom infrared (VIS-NIR) hyperspectral sensor, the authors gathered 1139 beans and twenty images of different kinds of coffee beans. The research used multiple selection approaches to identify the most effective band for detecting the insect-damaged green bean. Following that, the researchers utilized two classifiers: one integrated CEM with a support vector machine (SVM), while the other employed convolutional neural networks (CNN). The study's findings indicated that just three bands are enough to get 95% accuracy and a 90% kappa coefficient.

Whereas the researchers Wang et al. [24] proposed an explainable, intelligent, and lightweight coffee bean quality inspection system that utilized CV and DL techniques. The dataset is open source and was used in their research. It contains 4626 samples of green beans, of which 2150 are images of good beans and 2476 are images of defective beans. The approaches manipulated are a compound of the KD method and the ResNet (residual neural network). The KD method is adapted to abridge redundant operations of CNN; besides, ResNet is improved based on Visual Geometry Group Network 19 (VGG19). Additionally, they use the LIME Method (Local Explainable Model-Agnostic Explanation) to estimate the regionally selected black-box model. As a result, the combination of CV and DL technologies for coffee bean inspection quality is classified precisely. Also, it enables the identification of the model's attribute marks in the sample, which turn the information into transparent and open form.

As specified by Chang and Huang [25], they seek to overcome the usual methods to detect defective beans and sort the bean types that all utilize human resources and cost time. Furthermore, they proposed a solution by using Deep Learning algorithms for coffee bean defect detection. The authors used the original dataset, which consisted of 3621 images, for this research. After implementing the rotation technique for Data Augmentation in the dataset, there were 7203 images for eight classes of coffee beans. The methodology that developed AlexNet is fundamental to the CNN (Convolutional Neural Network) framework used to detect defective beans. The results achieved for AlexNet across various phases show 95.5% accuracy for classifying beans and 100% accuracy for detecting defective beans.

A study by Hsia et al. [26] developed a lightweight deep convolutional neural network for detecting green coffee bean quality, which combined several frameworks, including the DSC, the SE block, and the skip block. The dataset was provided by a small optical sorter which included 4626 images, 2149 good images and 2477 bad images. Before feeding the photos to DCNN, image pre-processing was performed, such as feature extraction, noise reduction, and image resizing. For ensuring that every input image was the same size during training and validation. The study then combined RA, LA, and GC models to avoid random predictions. Based on the study, the model achieved an accuracy of 98.38%.

The work by Kim [27] from Johns Hopkins University aims to forecast coffee bean quality via the random forest machine

learning model. The dataset used is the Arabica coffee beans dataset from the Coffee Quality Institute database, which incorporates variables such as total cup points, altitude, color, moisture, processing method, quakers, number of bags, bag weight, category one defects, category two defects, country of origin, and variety. The random forest model, known for its ability to handle non-linear data and require minimal data preparation, achieves an F1-score of 61.7% for coffee quality prediction. Although limitations are not explicitly mentioned, potential ones include the reliance on a specific machine learning algorithm and the generalizability of findings to different coffee bean varieties and regions.

Adiwijaya et al. [28] assessed how effectively an instrument can directly measure coffee quality. They utilized a dataset of 90 coffee beans, which is 30 beans for each of the three classifications. The waterfall model was utilized as a research technique. It is a sequential and methodical way to design software. So, the accuracy score remained at 83% for the testing scheme, which used distinct k values for all 18 data points: 3, 5, and 7. One problem they ran across was that the tool has to be improved on smartphones with better cameras to achieve better results.

In research, Arboleda [29] suggested a categorization of immature and mature coffee beans using image processing and texture features, namely entropy, contrast, energy, and homogeneity. The coffee beans were picked in the Philippines in 2019 and kept in hucks for a year. Then, 100 beans, both ripe and immature, were utilized as samples. There are a number of machine-learning techniques that have been utilized to sort the coffee beans. The study indicated that texture characteristics are better at finding out how mature coffee beans are than color features RGB or HSV. A medium KNN algorithm got 97% accuracy.

Buonocore et al. [30] conducted a study on using deep learning to classify coffee bean varieties and detect fraud. The study aimed to distinguish between Arabica and Robusta coffee beans in real-time industrial scenarios. The authors created a dataset of over 2500 coffee beans and employed object detection techniques and a YOLO-based CNN. The results showed successful classification and fraud detection with 98% correct detection name. While limitations were not explicitly mentioned, challenges may arise from altered visual appearances due to roasting or grinding. Overall, this research describes the potential of deep learning and object detection in ensuring coffee quality.

In the latest work, Nawaz et al. [31] developed CoffeeNet, which is a deep learning model for localizing and classifying coffee plant leaf disorders. By using an improved CenterNet approach with a spatial-channel attention strategy-based ResNet-50 model, the deep learning model successfully addressed the issues of identifying coffee plant leaf diseases induced by picture distortions. Using the Arabica coffee leaf resource, the model achieved a high classification accuracy of 98.54% with a mAP of 0.97. CoffeeNet shows potential in terms of early identification and management of coffee plant diseases, which benefits crop productivity and quality.

Further research is required for validation and broader application in diverse coffee-growing regions. Based on the literature review, it is apparent that the detection of coffee defects is among the hottest areas of research in artificial intelligence applications in the agriculture sector. Moreover, these technologies have been promising in terms of accuracy and other figures of merit.

3. PROPOSED METHODOLOGY

Using deep learning and computer vision to understand and process digital data (i.e., images or videos) has a significantly important area of research with a high and direct impact on the domain of visual data in real-time detection and classification [32–34]. Inspection of defects is one area where computer vision systems have the potential to replace human inspectors for many activities, including traditional visual quality inspection. Therefore, we aim to develop a computer vision deep learning system for real-time coffee bean defect detection and classification.

As a result, we aim to build a computer vision deep learning system for real-time coffee bean defect detection and classification. This system will detect and classify defects accurately and quickly, efficiently, and effectively. The methodological procedure will contain detailed stages, as shown in Figure 1. The first stage of the process is data collection, which will be explained in detail in the next section, followed by the dataset pre-processing stages, which include many steps (i.e., image labeling and data augmentation). Once the data preprocessing step is complete, deep learning techniques and object detection algorithms are implemented.

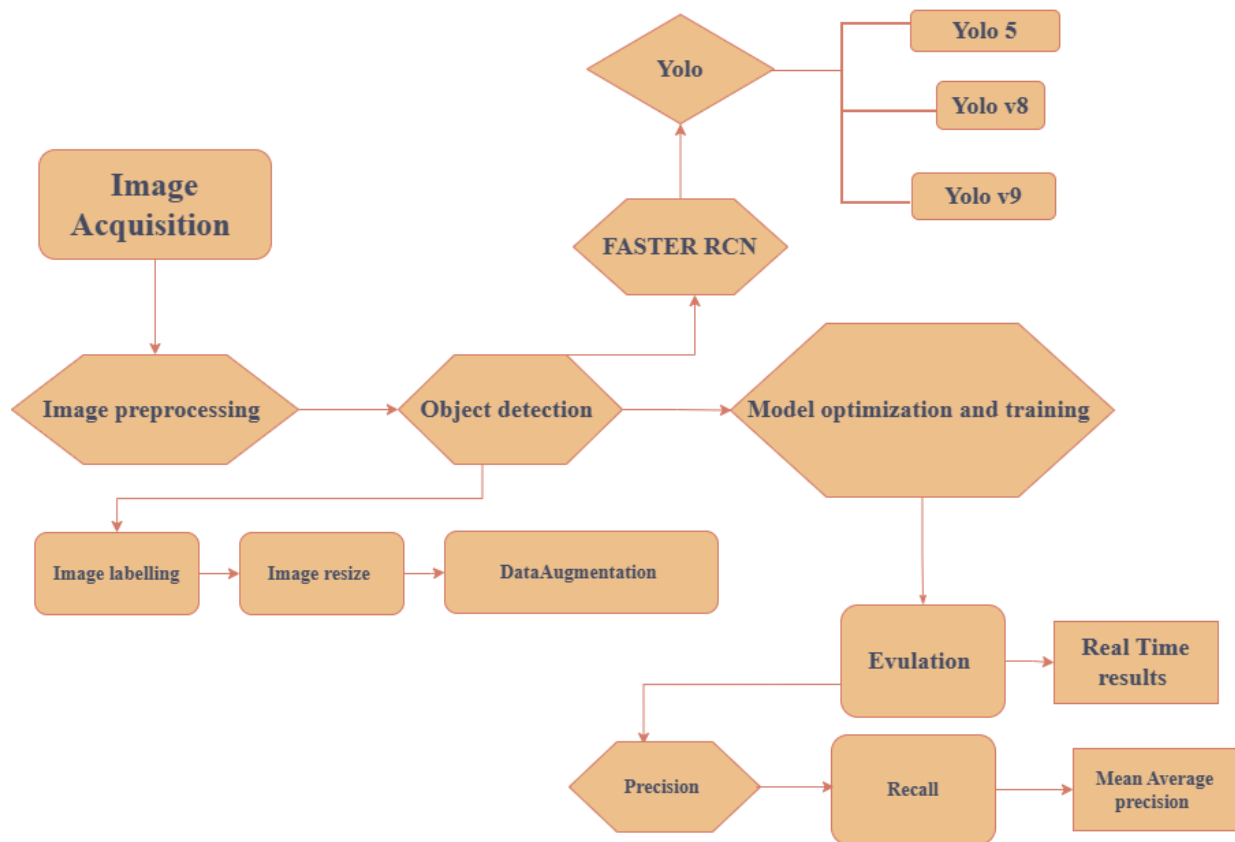


Figure 1. Methodology

3.1 Data collection

The dataset we acquired involved the procurement of Brazilian coffee bean, Colombia Pantera Fuerte and Turkish coffee bags that encompassed a range of defects, including burnt beans, under-roasted or green beans, and intentionally broken beans.

Additionally, we included a set of well-roasted beans as a reference for comparison purposes. To maintain accuracy and consistency, we diligently captured high-quality photographs of each type using a Canon 1100D camera, and others were captured using an iPhone 11 Pro. We also included some of the images in an online dataset on Kaggle. Moreover, the photography was conducted against a white background to uphold uniformity throughout the dataset and to simplify the pre-processing phase.

3.2 Data description

The dataset comprises a wide array of coffee bean images, encompassing various types of defects that may be present.

These defects include burnt beans, under-roasted or immature beans, and broken beans. In addition, a collection of well-roasted beans has also been included. A meticulous selection process has resulted in a balanced four-class dataset. Table 1 shows the number of images for each class. Furthermore, the dataset has been meticulously organized, with each specific coffee bean type allocated to its dedicated folder. This meticulous categorization ensures that the dataset is primed for any future processing or analysis. By implementing this systematic approach, the processing of the dataset has been streamlined, guaranteeing its seamless integration into subsequent stages of research or analysis pipelines.

Table 1. Dataset distribution

Class	Instances
Normal	423
Green	449
Burnt	452
Broken	449
Total	1773

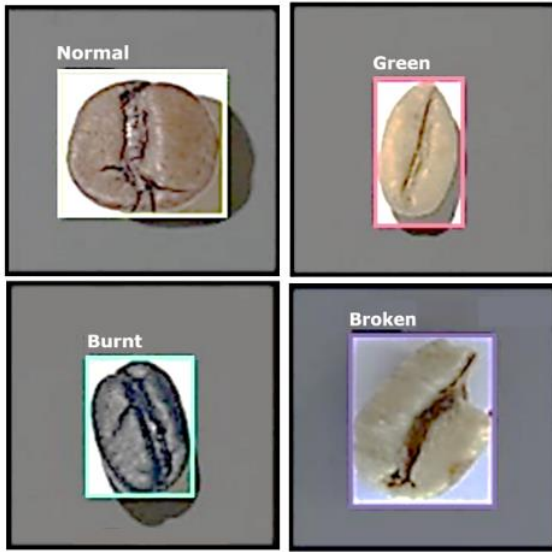


Figure 2. Sample of annotated dataset images

Figure 2 shows the sample annotated images from the qualified dataset.

3.3 Data preprocessing

Data pre-processing is crucial in computer vision applications, particularly in the proposed advanced deep learning-based system for the detection and classification of defects in coffee beans. The following steps have been carried out in terms of preprocessing of the dataset.

(1) Data Labeling: Labeling the images is a crucial step in an object detection project. It provides the foundation for the model to learn, recognize, and accurately detect the desired objects within images. It involves defining the precise objects that the model should detect in the images, which includes drawing bounding boxes around the objects and assigning class labels. The annotation process itself has been executed with precision, leveraging the capabilities of a computer vision platform known as Roboflow [33].

(2) Image Resizing: To facilitate the training phase, all the images were cropped to a standardized size through Roboflow as part of the pre-processing stages. The original dimensions of the images were 4272×2848; however, after the cropping process, they were resized to 416×416. This resizing ensures a more refined and focused view of the coffee bean.

(3) Data Augmentation: To expand the variety of images in our dataset and increase its diversity, we utilize data augmentation methods using Roboflow. Table 2 shows the number of images for each class after augmentation. It is worth noting that introducing data augmentation inherently increases the generalizability of the proposed models. In terms of augmentation, we have employed rotation, scaling, flipping, and hue, saturation, and brightness (HSB) operations using Roboflow. Consequently, to help develop more robust and generalized models for real-time coffee bean defect detection.

Table 2. Dataset after augmentation

Class	Instances
Normal	848
Green	932
Burnt	923
Broken	923
Total	3626

(4) Data Split and Validation: For the training and testing the data was split into 80:20 ratio. In addition, to perform validation, a 10-fold cross validation has been employed.

3.4 Model evaluation

The study has been evaluated by the well-known figures of merit used in literature.

Precision: Precision is a necessary measurement to be concluded, extending comprehension into the model's trustworthiness in determining images that exhibit defects particularly [34]. It can be calculated by the given Eq. (1):

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

Here, TP and TF refer to true positive and false positive rates, respectively, which can be calculated from the confusion matrix. mAP50: Mean Average Precision (mAP) is a metric utilized to assess object detection models such as Fast R-CNN and the YOLO family. These values are computed over recall values from 0 to 1. The confusion matrix, intersection over union (IoU), recall, and precision can be used to compute the mAP. The precision formula is given in Eq. (1), while the recall can be expressed as [34].

Recall: It measures how well you can find the TP value out of all predictions, including TP and false negative (FN) in Eq. (2).

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2)$$

4. RESULTS AND DISCUSSION

4.1 Experimental setup

Using Roboflow [35], we partitioned our dataset into two distinct subsets: an 80% portion designated for training the proposed models and 20% reserved for testing. To achieve robust coffee bean defect detection, we employed four CNN-based Object Detection Models Known for their effectiveness in the field: Faster R-CNN, YOLOv5, YOLOv8, and YOLOv9 [36-40].

Following dataset annotation and data split (80% training, 20% testing), we initially trained the models on the original dataset to assess their ability to learn from existing data without external adjustments or augmentation steps.

Figure 3 illustrates the proposed Faster R-CNN model architecture employed for the coffee bean defect detection. For YOLO models, Roboflow default optimized architecture settings have been incorporated in the proposed study.

4.2 Experimental results

As an initial attempt, YOLOv5 and YOLOv8 successfully detected coffee beans in real-time, exhibiting high mean average precision (mAP), a widely used metric for assessing object detection model accuracy. Conversely, YOLOv9 and Faster R-CNN yielded moderate mAP scores compared to YOLOv5 and YOLOv8. Detailed results and hyperparameters for each algorithm are presented in Table 3. Moreover, multiple experiments were conducted with YOLOv9 until optimal results were achieved. It is apparent that YOLOv8 outperformed YOLOv5, YOLOv9 and Faster R-CNN with a margin of 0.4%, 1.3% and 2%, respectively.

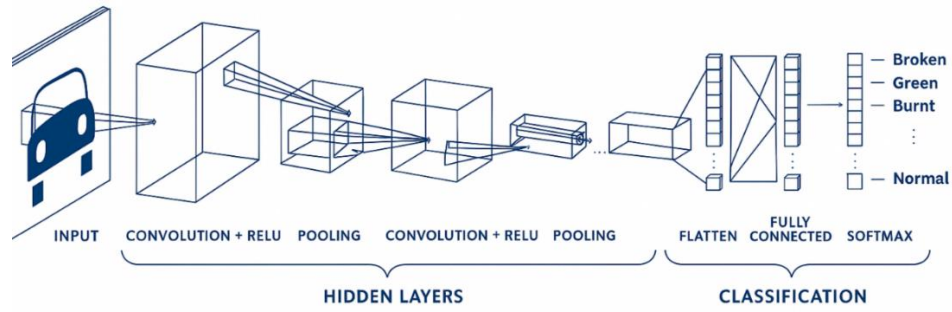


Figure 3. Faster R-CNN architecture

In the second experiment, we explored the application of data augmentation techniques employing various transformations, including rotation within the range of -15° to $+15^\circ$, horizontal and vertical flipping, and the addition of noise up to 0.1% of pixels. These techniques aim to diversify the dataset and enhance the performance and robustness of machine learning models. However, after comprehensive training on the augmented data, all models exhibited a decline in performance compared to the original dataset. This decline can be attributed to the introduction of additional complexities and variations that the models were not adequately trained to handle.

Performance analysis revealed that YOLOv8 experienced a slight decrease, followed by YOLOv5 and YOLOv9. Notably, Faster-RCNN demonstrated a more significant drop in performance as shown in Table 4, likely due to the limited number of epochs set to 25, as constrained by GPU limitations in the academic setup.

As shown in Figures 4 and 5, the confusion matrix highlights the models' performance across 'broken,' 'burnt,' 'green,' and 'normal' classes, with impeccable accuracy for most.

YOLOv5 showed slightly higher precision in 'broken' class classification (TP=1.00) compared to YOLOv8 (TP=0.98).

However, YOLOv8 demonstrated marginally improved true positive rates for 'burnt' and 'green' classes (TP=0.99), whereas YOLOv5 achieved perfect classification.

Notably, the inclusion of a 'background' class allows the model to differentiate between objects of interest and background elements, enabling it to abstain from making detections in the absence of relevant objects. These findings emphasize the models' robustness in object detection while indicating subtle differences in class-specific accuracies.

Table 3. Results of first experiment

Algorithm	No. of Epochs	No. of Batches	Optimizer	mAP50
Faster-RCNN	100	8	-	97.11%
YOLOv5	200	16	SGD	99.1%
YOLOv8	50	16	Adam	99.5%
YOLOv9	16	2	SGD	97.8%

Table 4. Results of second experiment

Algorithm	No. of Epochs	Batch Size	Optimizer	mAP50
Faster-RCNN	25	8	-	86.36%
YOLOv5	100	16	SGD	98.7%
YOLOv8	50	16	Adam	99.3%
YOLOv9	16	2	SGD	95.3%

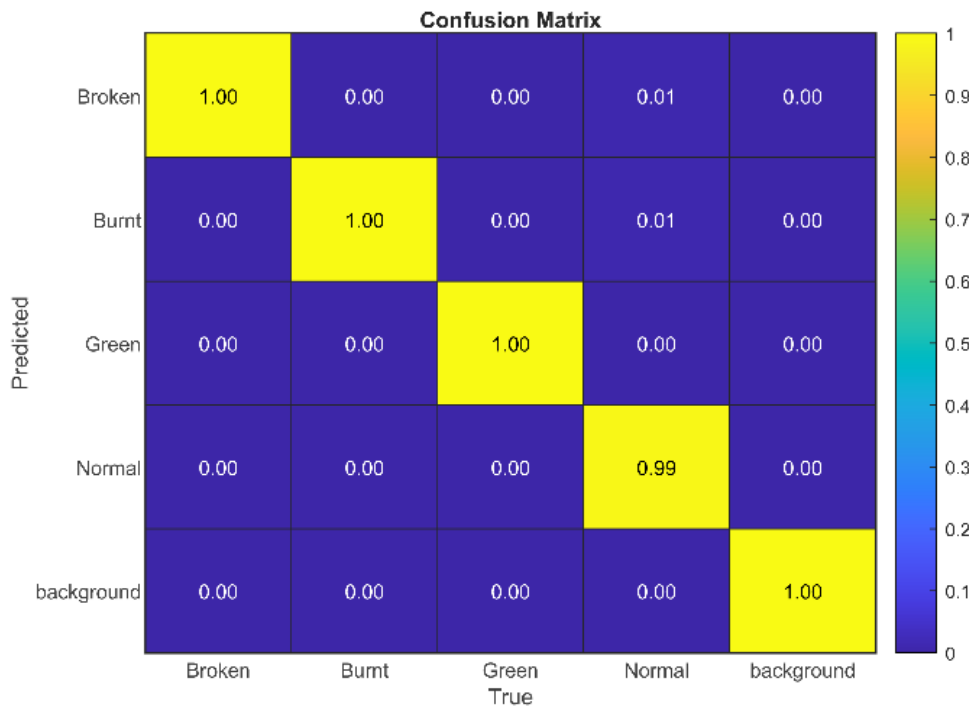


Figure 4. YOLOv5 confusion matrix

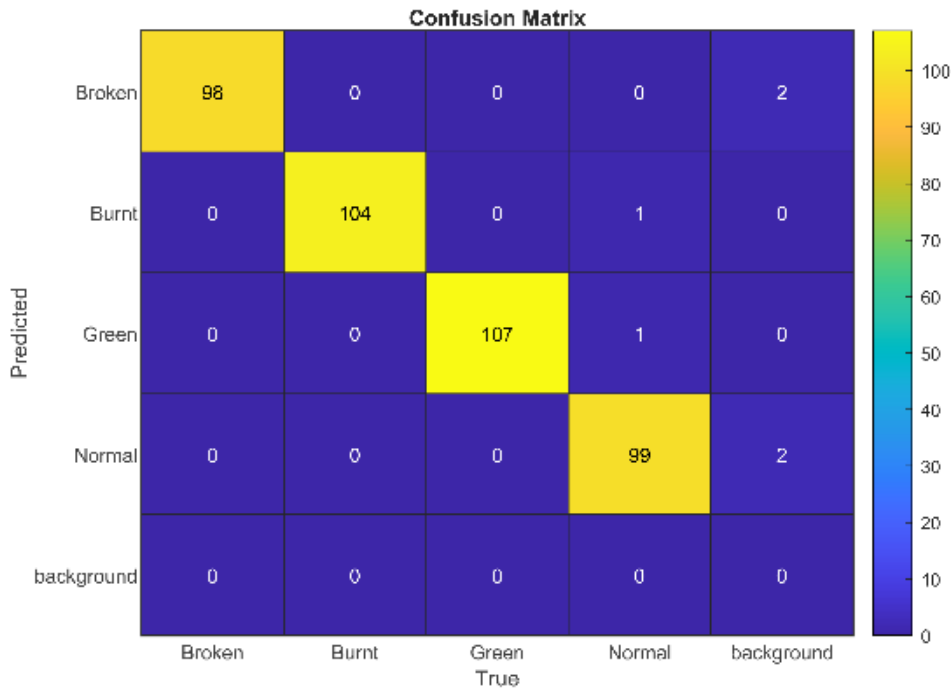


Figure 5. YOLOv8 confusion matrix



Figure 6. Sample of predictions

The experiments conducted with the four models provided promising outcomes. Particularly notable were the exceptional performances of YOLOv5 and YOLOv8, as evidenced by their high mean average precision (mAP).

Figure 6 provides visual representations of sample predictions generated by the YOLOv5 model on the test images, clearly illustrating its proficiency in object detection and precise object identification. It is demonstrated that for each class, its accuracy is almost 98%.

In numerous situations, real-time coffee bean detection poses a significant challenge that calls for precise and effective object detection algorithms. The dataset used for training and evaluating the object detection models contained 1773 coffee bean images before data augmentation. By using augmentation techniques, this dataset was further increased to 3626 images. It's significant to remember that the dataset divided coffee beans into four groups: broken, green, burnt, and normal. The broken bean class, with its additional sub- categories, offered the biggest complexity. The need for data augmentation was probably influenced by this complexity because it allowed the models to learn different variations within the broken bean class, which improved their overall performance.

Four well-known object detection models were examined in this study: Faster R-CNN, YOLOv5, YOLOv8 and YOLOv9. A thorough assessment procedure was carried out to identify the best model for real-time deployment. This entails testing different algorithms iteratively to determine which one

performs best. A defined set of hyperparameters, such as the number of batches, epochs, and iterations for each experiment, were carefully adjusted during several experimental test runs. Essentially, one of the most important steps in computer vision tasks is hyperparameter tuning. We can greatly increase the performance, effectiveness, and generalizability of our models, resulting in more precise and useful computer vision applications, by carefully choosing and modifying these parameters.

Table 5. Experimental results

Algorithm	Epoch/Iteration	Batch Size	mAP50	Precision
Faster-RCNN	100	8	97.11%	85.5%
YOLOv5	200	16	99.1%	97.8%
YOLOv8	50	16	99.5%	99.8%
YOLOv9	16	2	97.8%	96.2%

Consequently, Table 5 shows the best object detection model chosen after a careful evaluation process is YOLOv8, which has the highest precision of 99.80% and the highest MAP of 99.50%. It only needs 50 training epochs and has a batch size of 16. After that is YOLOv5, whose precision is 97.80% and MAP is 99.10%. It needs 200 training epochs with a batch size of 16. Faster R-CNN has 85.5% precision and a 97.11% MAP. It needs 100 training epochs with a batch size of 8. Out of the four algorithms, YOLOv9 is the algorithm with

the lowest MAP of 97.80% and precision of 96.2%. Along with having the smallest batch size 2, it also has the fewest epochs 16 of the four.

4.3 Discussion

The present study focuses on automating quality control and sorting processes in coffee roasters and factories through real-time detection of different coffee bean types using computer vision techniques. The implementation utilized advanced deep learning models, including YOLOv5, along-side exploration evaluations with YOLOv8, YOLOv9, and Faster R-CNN. The training was conducted on a meticulously labeled dataset prepared via Roboflow, enabling robust detection of coffee bean types during dynamic processing scenarios. The developed computer vision model holds significant promise for deployment within coffee processing facilities. Real-time monitoring and classifying coffee bean types empower specialists to optimize sorting processes, ensuring product consistency and quality. The bar chart presented in Figure 7 provides a comparison of the performance of different models for coffee bean type detection, as measured by mean Average Precision (mAP50) scores. The evaluated models include the proposed YOLOv5 approach, YOLOv8, YOLOv9, and Faster R-CNN. Among these models, the proposed YOLOv8-based approach achieved the highest mAP50 score of 99.5%, demonstrating exceptional accuracy in identifying various coffee bean types, such as black, normal, immature, and broken beans, during real-time detection processes.

In addition, the YOLOv5 detection model exhibited remarkable precision, attaining an impressive mAP value of 99.1%. While both YOLOv5 and YOLOv8 demonstrated similar mAP results across all classes, we opted to utilize YOLOv5 and YOLOv8 in this study due to their superior performance in real-time practical scenarios.

In terms of misclassification, it is noteworthy that the proposed models exhibited a very low misclassification rate (less than 1% in case of YOLOv5 and YOLOv8). However, in case of YOLOv9 and Faster R-CNN the misclassification was

around 2% which is still bearable in real time applications of coffee bean defect detection. It is worth noting that introducing data augmentation inherently increases the generalizability of the proposed models. In terms of augmentation, we have employed several operations using Roboflow. Consequently, it helped develop more robust and generalized models for real time coffee bean defect detection.

Compared to existing methodologies, the proposed approach excels in granularity by effectively distinguishing between distinct coffee bean types, including black, normal, immature, and broken beans, during the roasting and packaging stages. This finer classification enhances the quality control processes, setting the study apart from previous works, which primarily focused on binary classifications, such as good versus bad. Nevertheless, challenges persist, particularly in improving the speed and adaptability of the models under varying environmental conditions, such as poor lighting and conveyor belt movement.

Nonetheless, as far as the quantitative comparison with the state-of-the-art is concerned, the study has utilized a self-curated dataset with a special focus on coffee beans for Jizan province of Saudi Arabia. It is hard to compare the results without some common grounds such as common dataset and nature/type of coffee bean under consideration.

The study is an outcome of the final year project, and a prototype for real-world application has been developed for idea testing. The system was successfully deployed as a web application where live feeds are fetched for bean detection. The system was tested in semi-real-life conditions, and it was promising in terms of detection accuracy.

The future endeavors will prioritize enhancing real-time performance and integrating alert systems to augment monitoring and decision-making capabilities using a dashboard. Further, the trained models can be implicated to a diverse range of coffee types other than the one under consideration in the study. It can potentially invite an investigation of the transfer learning capabilities of the proposed models [41, 42].

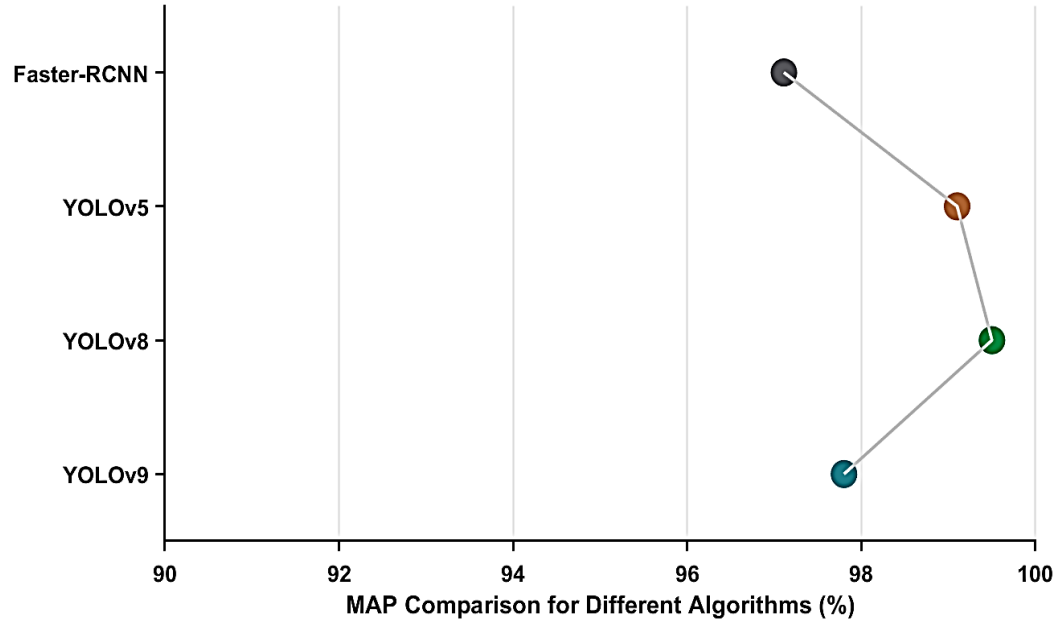


Figure 7. Coffee bean defect detection comparison

5. CONCLUSIONS

The proposed study represents a significant advancement in utilizing artificial intelligence to enhance the efficiency and quality of coffee manufacturing. Using sophisticated computer vision methods and deep learning models, we have built a real-time system that can find and sort different varieties of coffee beans. Our findings demonstrate exceptional performance, with the YOLOv8-based approach achieving the highest accuracy of 99.50%, surpassing other models evaluated in our study, including YOLOv8, YOLOv9, and Faster R-CNN. Furthermore, our study highlights the potential for practical applications of these technologies within coffee processing facilities. The ability to accurately detect and classify coffee bean defects in real-time empowers specialists to optimize sorting processes, ensuring product consistency and quality. This not only contributes to improving customer satisfaction but also holds economic implications by minimizing production losses and maximizing market value. Importantly, our research aligns closely with the goals of Vision 2030 in the Kingdom of Saudi Arabia, which aims to diversify the economy and promote innovation in key sectors such as agriculture. As Saudi Arabia continues to invest in agricultural self-reliance and digital transformation, our study offers a tangible example of how technology can drive progress and sustainability in the coffee industry. Looking ahead, future research endeavors will focus on addressing challenges such as optimizing model speed and adaptability under varying environmental conditions. Additionally, the integration of alert systems and dashboard interfaces could further enhance monitoring and decision-making capabilities in coffee processing facilities. By continuing to advance the capabilities of artificial intelligence in coffee production, we aim to contribute to the ongoing transformation of the industry and support the broader objectives of Vision 2030.

REFERENCES

- [1] Arab News. (2023). Saudi coffee industry to join top table of global producers. <https://www.arabnews.com/node/2277846/business-economy>, accessed on Sep. 29, 2024.
- [2] Mehrez, K.H., Khemira, H., Medabesh, A.M. (2023). Marketing strategies for value chain development: Case of Khawlani coffee-Jazan region, Saudi Arabia. *Journal of the Saudi Society of Agricultural Sciences*, 22(7): 449-460. <https://doi.org/10.1016/j.jssas.2023.04.004>
- [3] MEWA to establish geographical indication for coffee in Jizan region. (2024). <https://www.mewa.gov.sa/en/MediaCenter/News/Pages/News912020.aspx>, accessed on Jun. 3, 2024.
- [4] Al monitor Saudi Arabia. (2020). How Saudi Arabian farmers are trying to preserve traditional coffee production-al-monitor.com. <https://www.al-monitor.com/originals/2020/02/saudi-arabias-coffee-culture-shifts.html>, accessed on Jun. 3, 2024.
- [5] Microindustries. <https://www.aramco.com/en/sustainability/responsible-business/supporting-communities/economic-and-community-projects/microindustries>, accessed on Jun. 3, 2024.
- [6] Ibrahim, N.M., Gabr, D.G.I., Rahman, A.U., Dash, S., Nayyar, A. (2022). A deep learning approach to intelligent fruit identification and family classification. *Multimedia Tools and Applications*, 81(19): 27783-27798. <https://doi.org/10.1007/s11042-022-12942-9>
- [7] Ibrahim, N.M., Gabr, D.G., Rahman, A., Musleh, D., AlKhulaifi, D., AlKharraa, M. (2023). Transfer learning approach to seed taxonomy: A wild plant case study. *Big Data and Cognitive Computing*, 7(3): 128. <https://doi.org/10.3390/bdcc7030128>
- [8] de Oliveira, E.M., Leme, D.S., Barbosa, B.H.G., Rodarte, M.P., Pereira, R.G.F.A. (2016). A computer vision system for coffee beans classification based on computational intelligence techniques. *Journal of Food Engineering*, 171: 22-27. <https://doi.org/10.1016/j.jfoodeng.2015.10.009>
- [9] Portugal-Zambrano, C.E., Gutiérrez-Cáceres, J.C., Ramirez-Ticona, J., Beltran-Castañón, C.A. (2016). Computer vision grading system for physical quality evaluation of green coffee beans. In 2016 XLII Latin American Computing Conference (CLEI), Valparaíso, Chile, pp. 1-11. <https://doi.org/10.1109/CLEI.2016.7833383>
- [10] Pinto, C., Furukawa, J., Fukai, H., Tamura, S. (2017). Classification of green coffee bean images base on defect types using convolutional neural network (CNN). In 2017 International Conference on Advanced Informatics, Concepts, Theory, and Applications (ICAICTA), Denpasar, Indonesia, pp. 1-5. <https://doi.org/10.1109/ICAICTA.2017.8090980>
- [11] Arboleda, E.R., Fajardo, A.C., Medina, R.P. (2018). An image processing technique for coffee black beans identification. In 2018 IEEE International Conference on Innovative Research and Development (ICIRD), Bangkok, Thailand, pp. 1-5. <https://doi.org/10.1109/ICIRD.2018.8376325>
- [12] Huang, N.F., Chou, D.L., Lee, C.A. (2019). Real-time classification of green coffee beans by using a convolutional neural network. In 2019 3rd International Conference on Imaging, Signal Processing and Communication (ICISPC), Singapore, pp. 107-111. <https://doi.org/10.1109/ICISPC.2019.8935644>
- [13] Sarino, J.N.C., Bayas, M.M., Arboleda, E.R., Guevarra, E.C., Delloso, R.M. (2019). Classification of coffee bean degree of roast using image processing and neural network. *International Journal of Scientific & Technology Research*, 8(10): 3231-3233.
- [14] García, M., Candelo-Becerra, J.E., Hoyos, F.E. (2019). Quality and defect inspection of green coffee beans using a computer vision system. *Applied Sciences*, 9(19): 4195. <https://doi.org/10.3390/app9194195>
- [15] Kuo, C.J., Wang, D.C., Chen, T.T., Chou, Y.C., Pai, M.Y., Horng, G.J., Hung, M.H., Lin, Y.C., Hsu, T.H., Chen, C.C. (2019). Improving defect inspection quality of deep-learning network in dense beans by using hough circle transform for coffee industry. In 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), Bari, Italy, pp. 798-805. <https://doi.org/10.1109/smc.2019.8914175>
- [16] Kuo, C.J., Chen, C.C., Chen, T.T., Tsai, Z., Hung, M.H., Lin, Y.C., Chen, Y.C., Wang, D.C., Homg, G.J., Su, W.T. (2019). A labor-efficient GAN-based model generation scheme for deep-learning defect inspection among dense beans in coffee industry. In 2019 IEEE 15th International Conference on Automation Science and Engineering (CASE), Vancouver, Canada, pp. 263-270.

- <https://doi.org/10.1109/COASE.2019.8843259>
- [17] Akbar, M.N.S., Rachmawati, E., Sthevanie, F. (2020). Visual feature and machine learning approach for arabica green coffee beans grade determination. In Proceedings of the 6th International Conference on Communication and Information Processing, New York, United States, pp. 97-104. <https://doi.org/10.1145/3442555.3442571>
 - [18] Eustaquio, W.R., Dioses Jr, J.L. (2020). Classification of immature and mature coffee beans using RGB values and machine learning algorithms. *International Journal of Emerging Trends in Engineering Research*, 8(7): 3016-3022. <https://doi.org/10.30534/ijeter/2020/22872020>
 - [19] Gope, H.L., Fukai, H. (2020). Normal and peaberry coffee beans classification from green coffee bean images using convolutional neural networks and support vector machine. *International Journal of Computer and Information Engineering*, 14(6): 189-196.
 - [20] Santos, F.F.L.D., Rosas, J.T.F., Martins, R.N., Araújo, G.D.M., Viana, L.D.A., Gonçalves, J.D.P. (2020). Quality assessment of coffee beans through computer vision and machine learning algorithms. *Coffee Science*, Lavras, 15: 1-9. <https://doi.org/10.25186/v15i.1752>
 - [21] Belan, P.A., de Macedo, R.A.G., Alves, W.A.L., Santana, J.C.C., Araujo, S.A. (2020). Machine vision system for quality inspection of beans. *The International Journal of Advanced Manufacturing Technology*, 111(11): 3421-3435. <https://doi.org/10.1007/s00170-020-06226-5>
 - [22] Janandi, R., Cenggoro, T.W. (2020). An implementation of convolutional neural network for coffee beans quality classification in a mobile information system. In 2020 International Conference on Information Management and Technology (ICIMTech), Bandung, Indonesia, pp. 218-222. <https://doi.org/10.1109/ICIMTech50083.2020.9211257>
 - [23] Chen, S.Y., Chang, C.Y., Ou, C.S., Lien, C.T. (2020). Detection of insect damage in green coffee beans using VIS-NIR hyperspectral imaging. *Remote Sensing*, 12(15): 2348. <https://doi.org/10.3390/rs12152348>
 - [24] Wang, P., Tseng, H.W., Chen, T.C., Hsia, C.H. (2021). Deep convolutional neural network for coffee bean inspection. *Sensors & Materials*, 33(1): 2299-2310. <https://doi.org/10.18494/SAM.2021.3277>
 - [25] Chang, S.J., Huang, C.Y. (2021). Deep learning model for the inspection of coffee bean defects. *Applied Sciences*, 11(17): 8226. <https://doi.org/10.3390/app11178226>
 - [26] Hsia, C.H., Lee, Y.H., Lai, C.F. (2022). An explainable and lightweight deep convolutional neural network for quality detection of green coffee beans. *Applied Sciences*, 12(21): 10966. <https://doi.org/10.3390/app122110966>
 - [27] Kim, J.Y. (2022). Coffee beans quality prediction using machine learning. SSRN. <https://doi.org/10.2139/ssrn.4024785>
 - [28] Adiwijaya, N.O., Romadhon, H.I., Putra, J.A., Kuswanto, D.P. (2022). The quality of coffee bean classification system based on color by using k-nearest neighbor method. *Journal of Physics: Conference Series*, 2157(1): 012034. <https://doi.org/10.1088/1742-6596/2157/1/012034>
 - [29] Arboleda, E. (2023). Classification of immature and mature coffee beans using texture features and medium K nearest neighbor. *Journal of Artificial Intelligence and Technology*, 3(3): 114-118. <https://doi.org/10.37965/jait.2023.0203>
 - [30] Buonocore, D., Carratu, M., Lamberti, M. (2022). Classification of coffee beans varieties based on deep learning approach. In 18th IMEKO TC10 Conference "Measurement for Diagnostics, Optimization and Control to Support Sustainability and Resilience, Warsaw, Poland, pp. 14-19.
 - [31] Nawaz, M., Nazir, T., Javed, A., Amin, S.T., Jeribi, F., Tahir, A. (2024). CoffeeNet: A deep learning approach for coffee plant leaves diseases recognition. *Expert Systems with Applications*, 237: 121481. <https://doi.org/10.1016/j.eswa.2023.121481>
 - [32] Ismail, N., Malik, O.A. (2022). Real-time visual inspection system for grading fruits using computer vision and deep learning techniques. *Information Processing in Agriculture*, 9(1): 24-37. <https://doi.org/10.1016/j.inpa.2021.01.005>
 - [33] Rahman, A. (2025). Solar panel surface defect and dust detection: Deep learning approach. *Journal of Imaging*, 11(9): 287. <https://doi.org/10.3390/jimaging11090287>
 - [34] Atta, R., Mustafa, Y., Ghaida, A., Abrar, S., Joury, A., Manar, A., Mona, A., Noor, A. (2024). Diabetic retinopathy detection: A hybrid intelligent approach. *Computers, Materials & Continua*, 80(3): 4561-4576. <https://doi.org/10.32604/cmc.2024.055106>
 - [35] Jamil, M., Sudjud, S., Asriany, S., Said, M. (2025). X-CoffeeNet: A novel framework for coffee bean image classification using explainable artificial intelligence (XAI) and MobileNet. *Journal of Advances in Information Technology*, 16(10): 1423-1429. <https://doi.org/10.12720/jait.16.10.1423-1429>
 - [36] Mustafa, Y., Atta, R., Manar, A., Abrar, S., Joury, A., Noor, A., Ghaida, A., Mona, A. (2024). Early detection and classification of diabetic retinopathy: A deep learning approach. *AI*, 5(4): 2586-2617. <https://doi.org/10.3390/ai5040125>
 - [37] Basheer Ahmed, M.I., Zaghdoud, R., Ahmed, M.S., Sendi, R., Alsharif, S., Alabdulkarim, J., Saad, B.A.A., Alsabt, R., Rahman, A., Krishnasamy, G. (2023). A real-time computer vision based approach to detection and classification of traffic incidents. *Big Data and Cognitive Computing*, 7(1): 22. <https://doi.org/10.3390/bdcc7010022>
 - [38] Alnuaimi, M.N., Alqahtani, N.S., Gollapalli, M., Rahman, A., Alahmadi, A., Bakry, A., Youldash, M., Alkhulaifi, D., Ahmed, R., Al-Musallam, H. (2024). Transfer learning empowered skin diseases detection in children. *CMES-Computer Modeling in Engineering & Sciences*, 141(3): 2609-2623. <https://doi.org/10.32604/cmcs.2024.055303>
 - [39] Farooqui, M., Rahman, A., Alsuliman, L., Alsaif, Z., Albaik, F., Alshammari, C., Sharaf, R., Olatunji, S., Althubaiti, S.W., Gull, H. (2024). A deep learning approach to industrial corrosion detection. *Computers, Materials & Continua*, 81(2): 2587-2605. <http://doi.org/10.32604/cmc.2024.055262>
 - [40] Jan, F., Rahman, A., Busaleh, R., Alwarthan, H., Aljaser, S., Al-Towailib, S., Alshammari, S., Alhindi, K.R., Almogbil, A., Bubshait, D.A., Ahmed, M.I.B. (2023). Assessing acetabular index angle in infants: A deep learning-based novel approach. *Journal of Imaging*, 9(11): 242. <https://doi.org/10.3390/jimaging9110242>
 - [41] Rahman, A.U., Alqahtani, A., Aldhafferi, N., Nasir, M.U., Khan, M.F., Khan, M.A., Mosavi, A. (2022). Histopathologic oral cancer prediction using oral

- squamous cell carcinoma biopsy empowered with transfer learning. *Sensors*, 22(10): 3833. <https://doi.org/10.3390/s22103833>
- [42] Khan, T.A., Fatima, A., Shahzad, T., Alissa, K., Ghazal, T.M., Al-Sakhnini, M.M., Abbas, S., Khan, M.A., Ahmed, A. (2023). Secure IoMT for disease prediction empowered with transfer learning in healthcare 5.0, the concept and case study. *IEEE Access*, 11: 39418-39430. <https://doi.org/10.1109/ACCESS.2023.3266156>