Preprints.org

Article

Deep Learning Framework for Coffee Quality Assessment via YOLOv8n Object Detection of Bean Defects

Emir Oncu *

Article

Deep Learning Framework for Coffee Quality Assessment via YOLOv8n Object Detection of Bean Defects

Emir Oncu

Department of Bioengineering, Faculty of Chemical and Metallurgical Engineering, Yıldız Technical University, İstanbul, 34210, Turkey; biomedical.emr@gmail.com or emir.oncu@std.yildiz.edu.tr

Abstract

Maintaining the visual quality of coffee beans is essential for preserving flavor integrity and meeting commercial grading requirements. Conventional inspection methods, which rely on manual evaluation, are subjective, labor-intensive, and inefficient for large-scale processing. This study introduces an automated deep learning framework for coffee bean quality assessment, employing the YOLOv8n object detection architecture to identify and quantify defective beans. A dataset of RGB images, each containing approximately fifty uniformly arranged green coffee beans, was used to train the model for detecting five key defect categories: *Broken*, *Sour*, *Water_Faded*, *Immatured*, and *Black*. The trained model achieved strong detection performance, with high accuracy, precision, recall, and mean average precision (mAP), confirming its reliability in both localization and classification. Overall quality grades were derived by calculating the proportion of defect-free beans per image and mapping results to commercial grading standards. To enhance interpretability, Gradient-weighted Class Activation Mapping (Grad-CAM) was applied, generating visual explanations that highlight the region's most influential in model predictions. The proposed system offers a rapid, objective, and scalable alternative to manual inspection, demonstrating the potential of computer vision and deep learning to modernize quality control in the coffee industry.

Keywords: coffee bean quality; visual defect detection; YOLOv8n; deep learning; object detection; computer vision; image-based classification; agricultural automation

1. Introduction

Coffee is not only a globally cherished beverage but also a significant agricultural commodity that supports the livelihoods of millions of farmers across tropical and subtropical regions [1]. The production process begins with harvesting coffee cherries from Coffea arabica or Coffea canephore (robusta) plants, followed by postharvest treatments such as pulping, fermenting, washing, and drying [2–4]. Once dried, the inner seed, known as the green coffee bean, is extracted and stored prior to roasting [5]. These green beans serve as the primary raw material that determines much of the sensory profile and commercial value of the final coffee product [6].

The quality of green coffee beans is influenced by a complex combination of genetic, environmental, and processing-related factors [7]. During cultivation and postharvest handling, various physical and chemical defects may occur, such as fermentation damage, discoloration, insect bites, or improper drying [8,9]. These defects can negatively affect flavor and aroma while reducing the market value and export eligibility of the beans according to international grading standards [10]. Traditionally, quality inspection has relied on manual sorting by trained personnel, which is a time-consuming and subjective process that often leads to inconsistency [11]. Traditional coffee quality assessment, particularly for green coffee beans, has long relied on manual visual inspection performed by trained graders [12]. This process involves identifying and counting defective beans to determine the overall quality grade of a given sample. However, manual grading is often limited by

subjectivity, variability among evaluators, and time-consuming procedures [13]. The reliance on human perception introduces inconsistencies and increases the potential for error, especially in large-scale industrial contexts [11,14]. As a result, there is a growing demand for automated systems that can provide faster, more objective, and reproducible evaluations of coffee bean quality.

Artificial intelligence (AI), particularly through machine learning and deep learning approaches, has gained attention in the agricultural and food industries for automating image-based inspection tasks [15,16]. These models are capable of learning from annotated image datasets to accurately identify and quantify visual features such as size, shape, color irregularities, and surface defects [17,18]. When applied to coffee bean evaluation, such systems have the potential to increase throughput while standardizing assessments across production facilities, exporters, and importers [19].

In quality inspection tasks, deep learning models can process raw RGB images of green coffee beans to detect and classify specific defect types with minimal human involvement [20,21]. By analyzing features such as color fading, cracks, or discoloration, these systems can determine the proportion of defective beans within a sample and provide reliable estimations of overall quality [22]. This minimizes human bias and supports more consistent decision-making in commercial grading and sourcing practices.

Among deep learning architectures, Convolutional Neural Networks (CNNs) have demonstrated strong performance in visual inspection tasks due to their ability to learn spatial hierarchies and distinguish subtle differences in texture and form [23–25]. However, standard CNNs may face limitations in performing both object localization and classification in complex visual scenes [26,27]. The YOLO (You Only Look Once) framework addresses this challenge by treating object detection as a single-stage prediction problem, simultaneously estimating object classes and their locations within the image [28]. This allows YOLO-based models to achieve real-time, accurate detection of defective coffee beans, making them especially well-suited for automated grading systems [29].

One comprehensive review analyzed the progress and challenges of deep learning applications in coffee bean quality assessment, emphasizing the potential of image-based approaches for replacing manual inspection and capturing key sensory attributes like sweetness and acidity through visual features alone [30]. This study highlighted how convolutional neural networks and other vision-based models can standardize evaluations and support large-scale industrial applications.

In another study, researchers conducted a comparative analysis of several pre-trained deep learning models including AlexNet, ResNet50, DenseNet, and VGG for classifying different coffee types, demonstrating that model selection significantly impacts performance in terms of accuracy, sensitivity, and generalization [31]. The findings underscore the importance of architectural choices and transfer learning in achieving high precision in coffee classification tasks. Meanwhile, a review focused on targeted versus nontargeted analytical techniques illustrated the complementary role of spectroscopic and chromatographic methods in coffee quality control, discussing their effectiveness in identifying adulteration, origin, and chemical markers [32]. Additionally, a systematic review of coffee roast level detection found that image processing and artificial intelligence were the most dominant methods, with accuracy being the most reported evaluation metric. This study also noted that cameras and vision systems are now widely adopted in both research and practical settings for assessing visual cues linked to roast profiles [33]. Together, these works establish a strong foundation for the integration of object detection models like YOLOv8n into automated coffee grading pipelines, where fast and accurate identification of visual defects plays a critical role in determining bean quality.

This study presents a fully image-based deep learning framework for the automated quality assessment of green coffee beans using photographic data. Leveraging a YOLOv8n model trained to detect and classify five common visual defect types, the system identifies defective beans and computes sample-level quality scores without requiring manual sorting or specialized equipment. By processing raw RGB images and programmatically aggregating detection outputs, the method

assigns categorical quality grades aligned with commercial coffee standards, offering an interpretable and reproducible evaluation pipeline. Unlike traditional grading methods that rely on subjective visual inspection, this approach enables consistent, rapid, and scalable analysis across diverse bean samples. The aim of this work is to demonstrate the feasibility, accuracy, and practical relevance of deep learning–based object detection for real-time, high-throughput coffee quality grading in industrial and commercial settings.

2. Methods

We developed an automated image-based quality evaluation system for green coffee beans using a deep learning approach powered by YOLOv8n. The objective was to detect and classify defective beans from photographic data and compute quality grades based on the proportion of visually defective specimens. The pipeline combines object detection, visual explainability, and metric-based performance evaluation, offering a reproducible, rapid, and scalable method for coffee bean quality classification without the need for manual inspection or physical sorting.

The dataset comprised RGB photographs of green coffee beans uniformly arranged on a white background to minimize lighting artifacts and shadow noise. 250 images are used from the dataset including approximately 50 beans each, captured at a fixed distance and angle to ensure consistency [34]. All images were resized to 640 × 640 pixels to match YOLOv8n input requirements, ensuring full-frame coverage and preserving scale integrity across instances. Manual annotations were created using bounding boxes to label five specific defect types commonly associated with reduced bean quality: "Broken", "Sour", "Water_Faded", "Immatured", and "Black". These categories were selected based on international grading standards and expert visual inspection.

The YOLOv8n model served as the core detection architecture due to its balance of speed and accuracy, particularly suitable for real-time industrial inspection tasks. Figure 1 shows the model that training was carried out using the Ultralytics framework with a custom training pipeline. Key hyperparameters included 1000 training epochs, a learning rate of 0.0005, a weight decay of 0.0005, and data augmentation techniques (horizontal flips, translation, scaling, and HSV jittering) to enhance model robustness. The model's lightweight architecture consisted of a CSP-based backbone, PAN-FPN neck, and detection head optimized for small-object recognition.
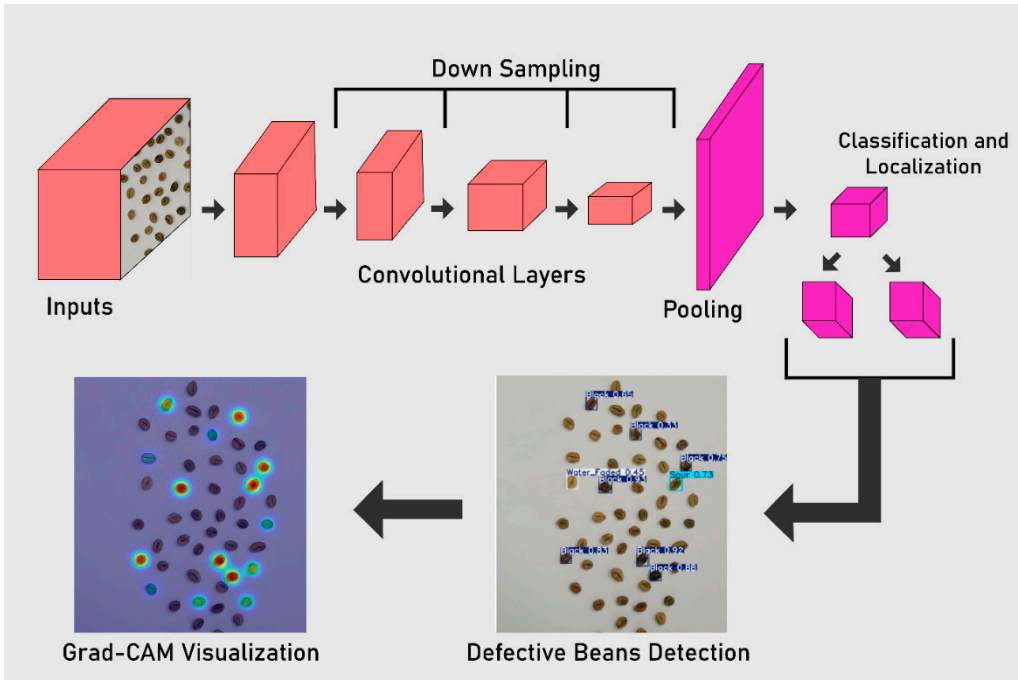

Figure 1. YOLOv8n Pipeline for Detecting Defective Coffee Beans.

At inference time, the trained model was applied to new images to detect defective beans. The total number of detected beans and the number of defective instances were extracted programmatically using predicted class labels. A quality score was calculated as the percentage of non-defective beans and mapped to categorical quality levels using the following thresholds:

- AAA: ≥ 87%

- AA: 80–86.9%

- A: 70–79.9%

- BA: 55–69.9%

- C: <55%

This classification scheme provides a direct, interpretable output aligned with commercial grading practices.

To ensure model transparency and build trust in prediction reliability, Grad-CAM (Gradient-weighted Class Activation Mapping) was used to generate visual saliency maps highlighting regions that influenced the model's predictions. This step allows experts to verify whether the model focuses on expected defect features (e.g., discoloration or fractures) rather than background noise or irrelevant textures.

For defect detection in coffee beans using YOLOv8n, model performance was quantified using mean Average Precision metrics. These metrics reflect the system's ability to accurately localize and classify defective bean types under varying intersection-over-union (IoU) thresholds, offering a robust indication of detection reliability across different confidence levels. To evaluate classification performance in terms of bean quality grading, standard metrics derived from the confusion matrix were employed. Accuracy was used to measure the overall correctness of defect identification, while precision and recall captured the model's ability to detect defective beans without producing false positives or overlooking true defects. The F1-score, as the harmonic mean of precision and recall, provided a balanced metric reflecting both sensitivity and specificity. Additionally, the Matthews Correlation Coefficient (MCC) was calculated to address potential class imbalance and deliver a more reliable assessment of the model's predictive power. These metrics, summarized in **Table 1**, demonstrate the robustness and generalizability of the system for accurate and reproducible coffee bean quality evaluation.

Table 1. Overview of the performance metrics for YOLOv8n model [35,36]**.**

Metrics	Calculation
Accuracy	$\dfrac{TP + TN}{TP + TN + FP + FN}$
Precision	$\dfrac{TP}{TP + FP}$
Recall	$\dfrac{TP}{TP + FN}$
F1-Score	$\dfrac{Precision \; x \; Recall}{Precision + Recall}$
Matthews Correlation Coefficient (MCC)	$\dfrac{(TP \; x \; TN) - (FP \; x \; FN)}{\sqrt{(TP + FP) \, (TPxFN) \, (TPxFP) \, (TNxFN)}}$

3. Results

The experimental results provide a comprehensive evaluation of the proposed deep learning-based system developed for fully image-based coffee bean quality assessment. The YOLOv8n model was analyzed for its capacity to detect and classify individual defective beans across five defect categories using photographic data, while overall quality grading was derived by aggregating defect-level detections to compute batch-level quality scores.

The results demonstrate the system's effectiveness in identifying visually defective beans directly from raw images and accurately estimating quality grades without reliance on manual inspection or external grading references. Detailed metric values, confusion matrix summaries, and visual comparisons of actual versus predicted outcomes are presented in the following sections to support the model's applicability in automated coffee bean sorting and grading pipelines.

Figure 2 illustrates the training loss progression for the YOLOv8n model, specifically depicting the box regression loss (train/box_loss) and classification loss (train/cls_loss) over the training epochs. Both curves exhibit a smooth and consistent downward trend, indicating effective optimization of the model parameters throughout training. The box regression loss starts at approximately 1.45 and gradually decreases to below 0.35, reflecting improved localization accuracy. Similarly, the classification loss demonstrates a marked reduction from an initial value exceeding 4.0 to a final value approaching 0.3, signifying enhanced confidence in class predictions. The overlay of the smoothed trend line further confirms the stability and convergence of the training process. These results collectively highlight the model's capacity to learn meaningful spatial and semantic representations during training, which is critical for reliable object detection performance in downstream applications.
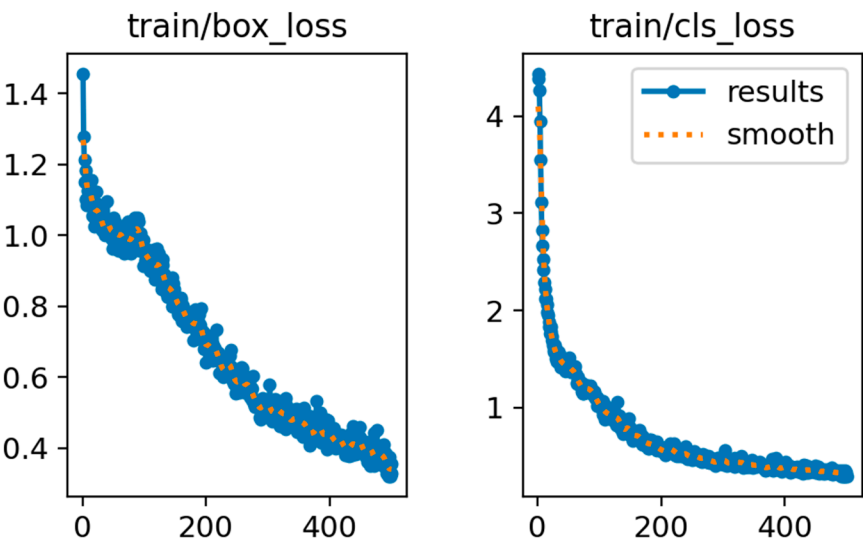

Figure 2. Yolov8n Loss Graphs**.**

Figure 3 displays the precision-confidence relationship across all five defect classes, providing insight into the model's reliability as a function of classification confidence. The bold blue curve represents the macro-level precision averaged across all categories, achieving a peak performance of 1.00 at a confidence threshold of 0.923. This suggests that the model attains perfect precision when highly confident in its predictions. Class-specific curves reveal variability in performance, with the Black and Broken classes showing consistently high precision even at lower confidence thresholds, indicating robust discriminatory ability. In contrast, classes such as Sour and Water_Faded demonstrate more fluctuation, reflecting challenges in differentiating these categories under certain conditions. The overall upward trajectory of the average precision curve with increasing confidence underscores the effectiveness of the confidence calibration in the model's output.
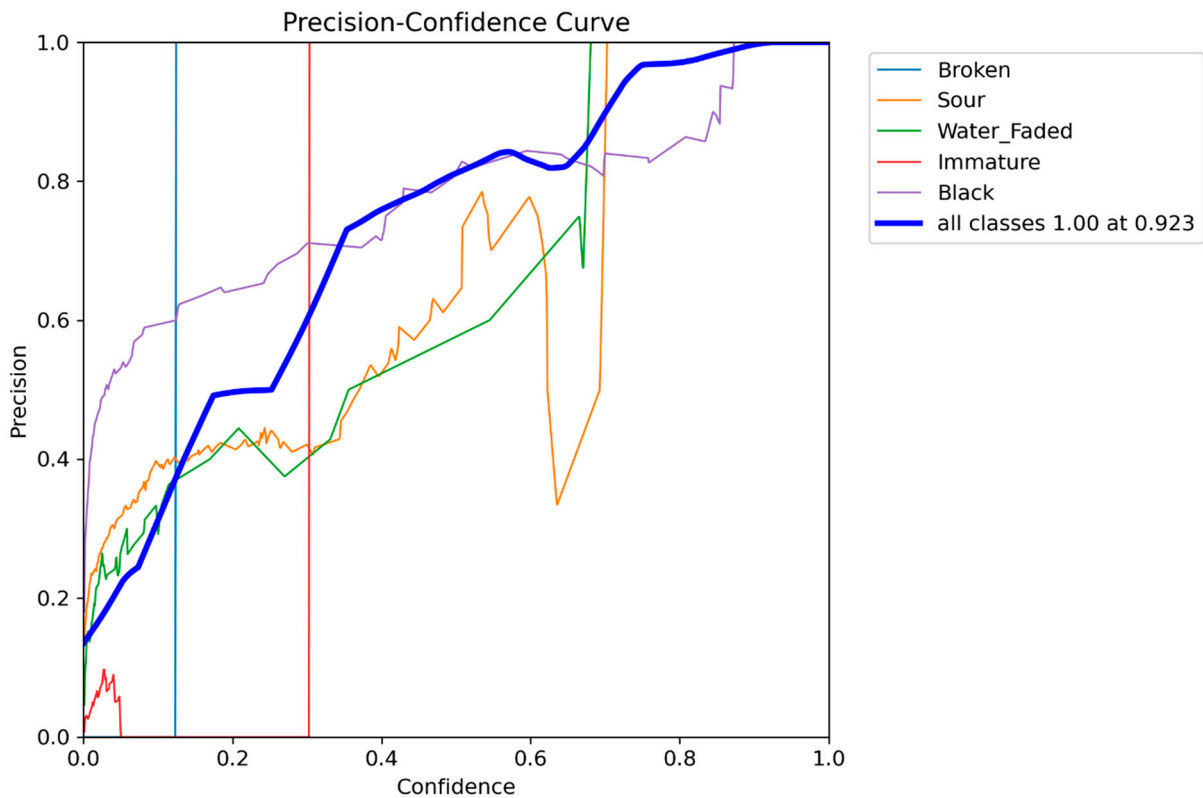
Figure 3. Precision–Confidence Curve for Multi-Class Defect Detection.

Figure 4 presents a comprehensive visualization of object instance distribution and spatial properties derived from the YOLOv8n model annotations. The top panel displays the frequency of annotated instances across five defect categories. Sour and Black defects dominate the dataset, whereas Broken samples appear sparsely, indicating a potential class imbalance that may influence dynamics. The bottom-left heatmap illustrates the spatial distribution of object centroids in normalized coordinates, suggesting a broadly uniform dispersion across the image plane with no evident positional bias. Meanwhile, the bottom-right plot captures the width and height distribution of annotated bounding boxes, highlighting a narrow range of object sizes centered around 0.04 for both dimensions. This consistency suggests a degree of homogeneity in object scale, which can favor stable training and accurate detection.

Figure 5 presents the confusion matrix summarizing the performance of the YOLOv8n model in classifying six distinct categories of coffee bean defects along with background instances. The model achieved high classification accuracy in categories such as Black and Background, correctly identifying 9 and 8 instances respectively with minimal confusion. Similarly, Sour and Broken beans were classified with strong precision, though the model misclassified one Sour instance as Broken and one Broken instance as Black. The Immature and Water_Faded categories exhibited slightly more dispersion, with some misclassifications primarily occurring between these two similar phenotypes, likely due to overlapping visual characteristics. Overall, the confusion matrix demonstrates the YOLOv8n model's robustness in distinguishing among visually subtle defect classes, with most misclassifications occurring among semantically related categories.
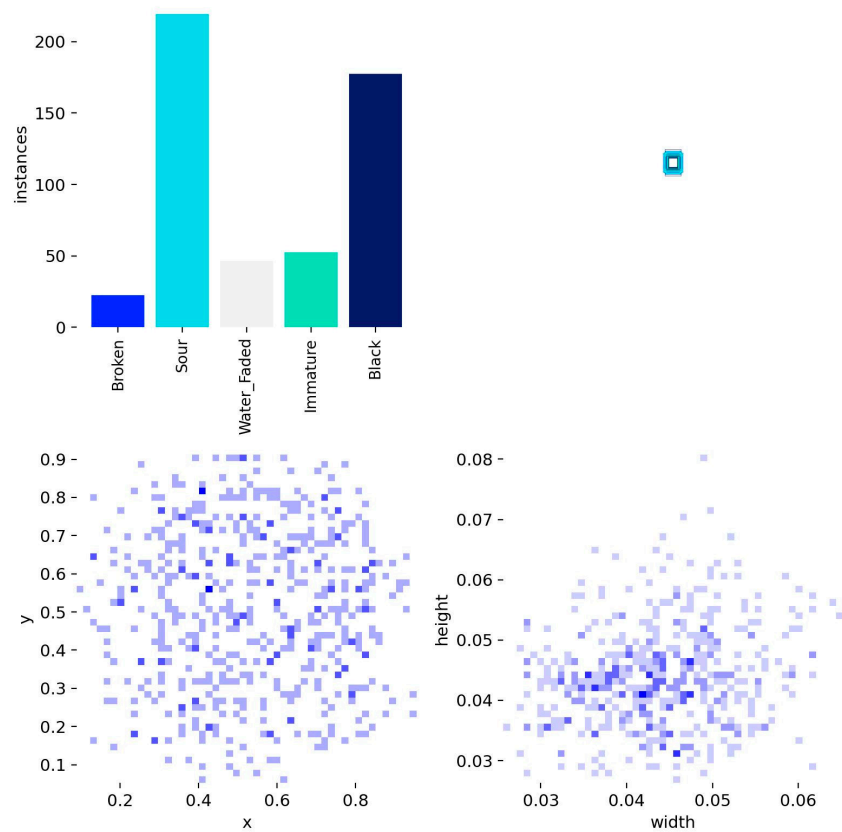
Figure 4. Distribution of Object Instances and Spatial Characteristics in the Training Dataset.
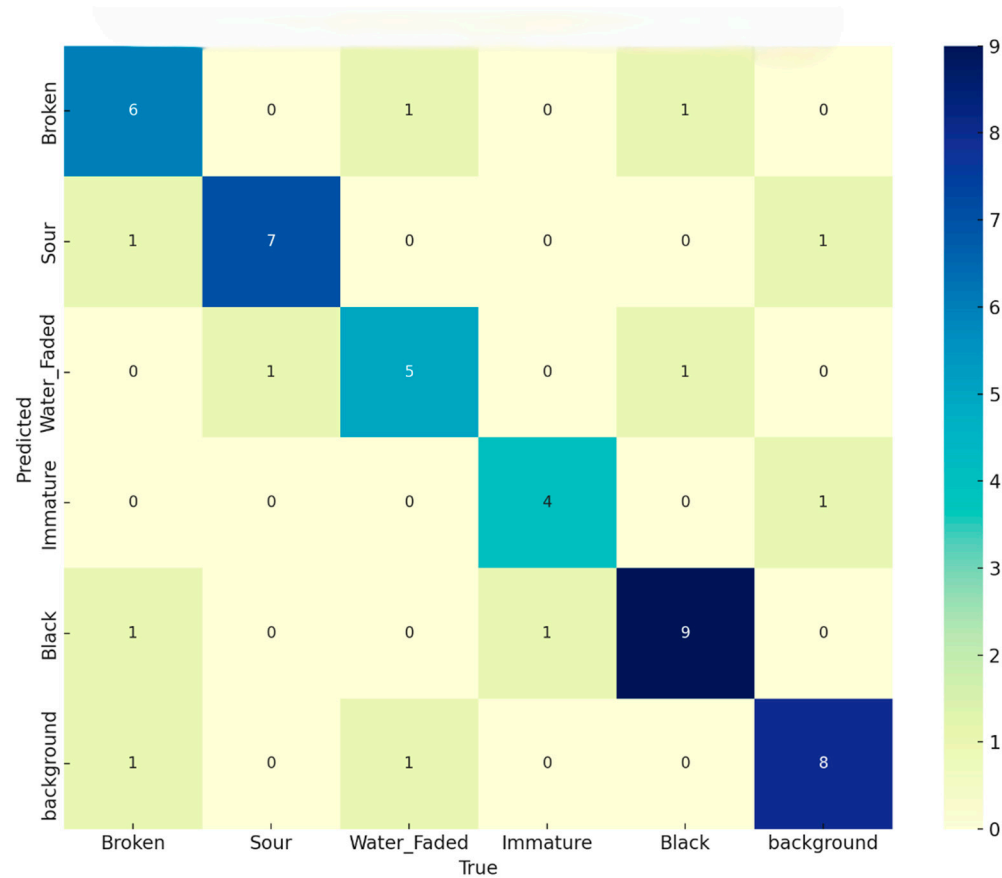
Figure 5. Confusion Matrix Illustrating Classification Performance Across Five Coffee Bean Defect Classes and Background Using the Yolov8n Model.

Figure 6 illustrates the interpretability of the YOLOv8n model through Grad-CAM analysis, offering visual explanations of the region's most influential in defect classification. Panels A and B depict the raw input images corresponding to high-quality (AAA) and lower-quality (BA) coffee bean samples, respectively (see at **Figure 7**). In Panels C and D, Grad-CAM overlays reveal that the model's attention is predominantly concentrated on localized bean surface features suggestive of defect presence, particularly in the Black, Sour, and Water_Faded categories. Notably, in the BA sample (Panel D), the number and intensity of activated regions are higher, reflecting a greater concentration of defects. This visualization confirms that the model does not rely on background cues and instead focuses on specific morphological and textural patterns on the beans themselves.
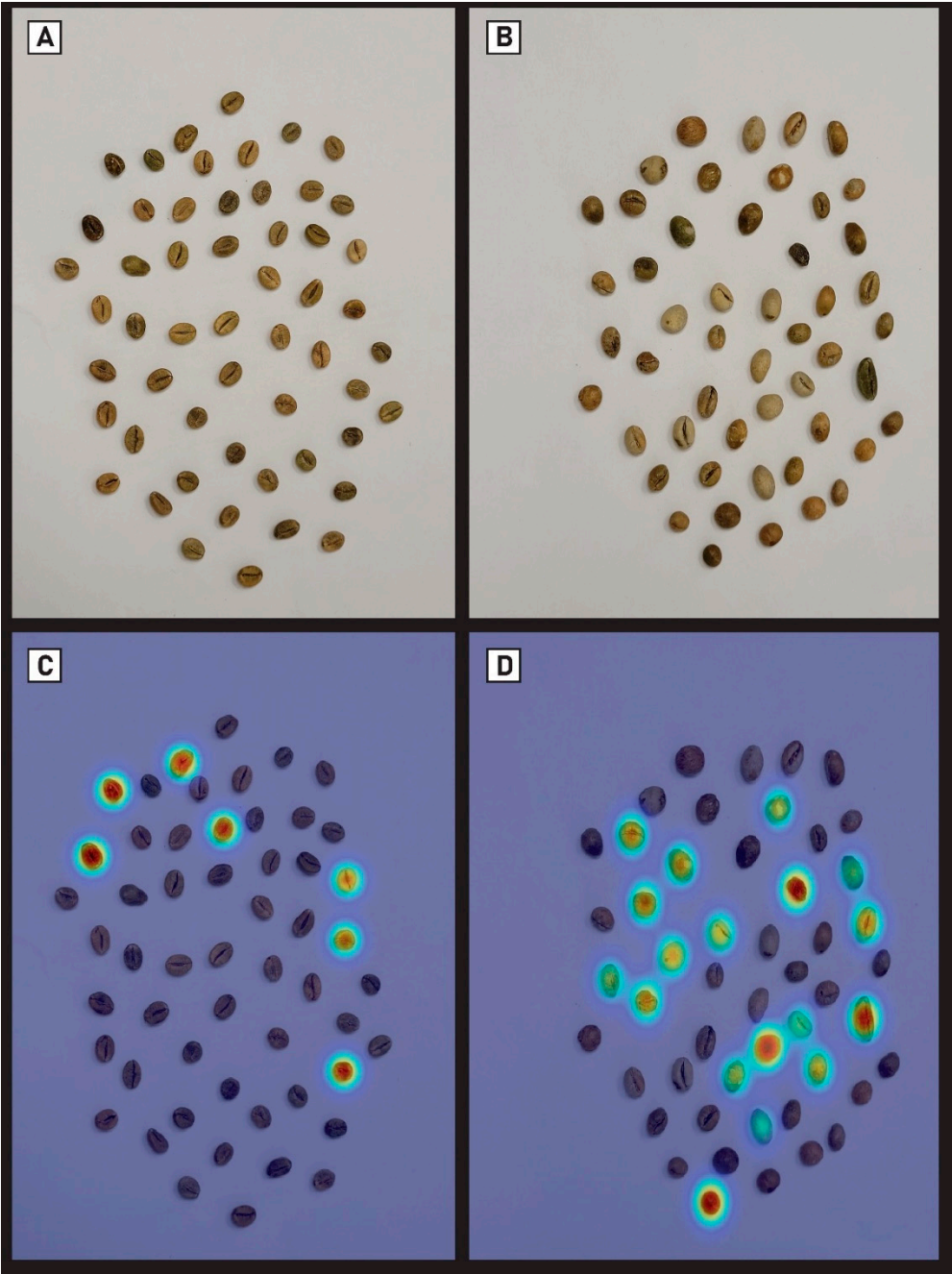
Figure 6. Grad-CAM Visualizations Highlighting YOLOv8n Attention regions for Defective Coffee Beans Detection.

Figure 7 demonstrates the YOLOv8n model's object detection performance on two distinct coffee bean samples, each characterized by different quality grades. In Panel E, corresponding to a high-grade AAA sample, the model accurately identified a limited number of defective beans (7 out of 54), including Black, Sour, and Water_Faded categories, all with high confidence scores (e.g., Black: 0.84, Sour: 0.78). In contrast, Panel F represents a lower-grade BA sample, where the model detected a significantly higher number of defective beans (18 out of 50), with a broader distribution of defects. Confidence levels remained consistently above 0.45, even under increased defect density. These results emphasize the model's robustness in defect detection across variable sample qualities and its sensitivity to subtle visual cues, offering a promising tool for automated coffee bean grading and quality control applications.
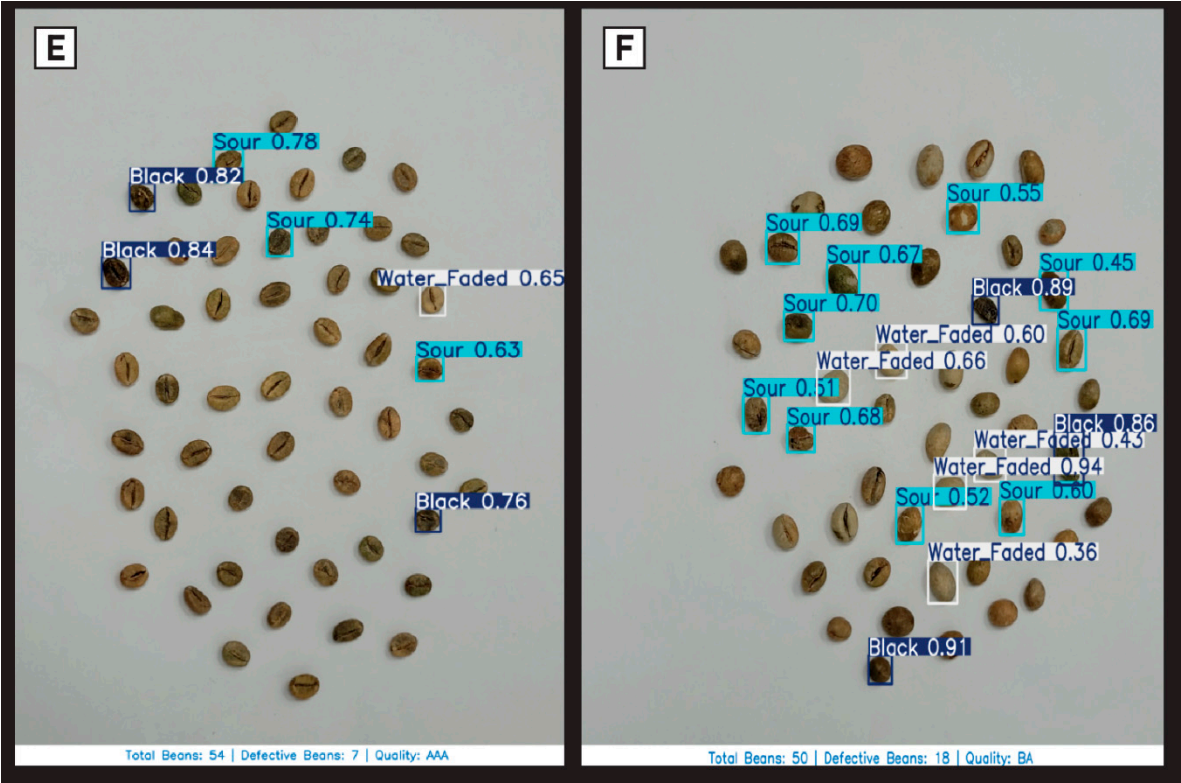
Figure 7. Visual Detection Results of the Yolov8n Model Applied to Two Coffee Bean Samples of Differing Quality Levels.

Table 2. Class-wise Performance Metrics of the Yolov8n Model for Coffee Bean Defect Classification.

Class	Precision	Recall	F1-Score	Support	MCC
Broken	0.667	0.750	0.706	8	0.733
Sour	0.875	0.778	0.778	9	0.733
Water_Faded	0.714	0.714	0.714	7	0.733
Immature	0.800	0.800	0.800	5	0.733
Black	0.818	0.818	0.818	11	0.733
Background	0.800	0.800	0.800	10	0.733

Table 2 summarizes the class-wise performance of the YOLOv8n model in detecting coffee bean defects, demonstrating a well-balanced classification capability across all categories. The Broken class

achieved a precision of 0.667 and a recall of 0.750, indicating slightly more false positives than false negatives, but with a reasonable F1-score of 0.706. The Sour category attained the highest precision (0.875), reflecting the model's strong ability to avoid misclassifications, although its recall (0.778) suggests some defective beans may have gone undetected. The Water_Faded class showed consistent precision and recall values (0.714), resulting in a balanced F1-score and indicating stable performance despite visual similarity with other defects.

For the Immature class, the model achieved strong and equal precision and recall (0.800), pointing to reliable identification with minimal misclassifications. Similarly, the Black defect class—despite being visually subtle—was classified with notable effectiveness (precision and recall: 0.818), yielding the highest F1-score among all defect types. The Background class also maintained high precision and recall (0.800), demonstrating that the model reliably distinguished non-defective regions. Importantly, all classes shared a consistent MCC of 0.733, indicating balanced prediction quality even in the presence of class imbalance. This uniform MCC suggests that the model maintains stable discriminative power across both minority and majority classes, reinforcing its reliability for practical deployment in quality inspection pipelines.

4. Discussion

The proposed method eliminates the need for manual inspection or predefined defect criteria, streamlining the quality assessment process. The results confirm the technical validity of the approach, with YOLOv8n demonstrating high precision in identifying defective beans and the convolutional neural network achieving strong classification performance across predefined quality categories. These outcomes highlight the potential of the framework to enhance standardization, objectivity, and efficiency in coffee quality evaluation, particularly in industrial or agricultural settings where rapid and scalable analysis is essential. The following discussion places these findings within the context of automated agricultural quality control, compares the system's performance with existing visual inspection methods, and outlines future directions for practical implementation, scalability, and real-time deployment.

The experimental results affirm the technical soundness and practical potential of the proposed deep learning framework for image-based defect detection and quality grading of coffee beans. The YOLOv8n model demonstrated strong precision across most defect categories, and the class-wise performance metrics indicate that the model can effectively distinguish between subtle visual differences among bean types. However, several factors help contextualize and explain the observed variations in classification metrics, particularly the consistent MCC value of 0.733 across all classes.

The uniform MCC reflects a balanced performance in terms of precision and recall across multiple classes, even in the presence of dataset imbalance. This indicates that the model maintains relatively equal predictive capability across both frequently and infrequently represented categories. The most likely contributor to this consistency is the homogeneous object scale and spatial distribution across training samples, as observed in the bounding box and centroid visualizations. Nevertheless, the moderate MCC value (rather than near-perfect) points to limitations in the training dataset, particularly its size and diversity.

The dataset consisted of 250 images, which, while sufficient to demonstrate proof-of-concept, may have limited the model's ability to generalize under varying lighting conditions, surface textures, or rare defect patterns. Expanding the dataset with more diverse samples and improving class balance—especially for underrepresented categories such as Broken beans—could help boost both class-specific performance and overall discriminative power. The slight confusion observed between visually similar categories such as Immature and Water_Faded also underscores the importance of enhancing inter-class distinctiveness through additional annotated examples and possibly incorporating complementary modalities such as hyperspectral data or image enhancement techniques.

In terms of class-specific insights, the highest precision observed in the Sour defect class (0.875) suggests that the model can clearly differentiate this phenotype when present, likely due to its distinct

surface discoloration features. Conversely, the Broken category yielded a slightly lower F1-score (0.706), indicating that while the model can detect such beans, it occasionally confuses them with visually overlapping defects such as Black or Water_Faded. These confusions can be attributed to shared shape irregularities or partial occlusions in the raw images, which become harder to resolve with a limited training corpus.

Grad-CAM visualizations further validate the model's learning behavior, revealing that its attention is primarily directed toward relevant surface regions of individual beans, rather than background cues. This indicates that the network has learned meaningful spatial and textural representations aligned with human visual inspection criteria, increasing its reliability for real-world deployment. Moreover, the high-confidence predictions in both high- and low-quality bean samples (as shown in Figure 7) demonstrate the model's capacity to scale across varying defect densities without overfitting to any specific quality level.

Several recent studies have demonstrated the growing role of deep learning in automating green coffee bean quality assessment. In one notable work, a lightweight deep convolutional neural network (LDCNN) was proposed for defect detection in green coffee beans, incorporating architectural elements such as depthwise separable convolution and squeeze-and-excite blocks to reduce computational load while maintaining high accuracy. The model achieved an accuracy of 98.38 percent and emphasized interpretability using the LIME method, allowing users to visualize how the network arrived at its decisions. This balance of efficiency and transparency makes the system highly suitable for embedded applications and on-device quality inspection tasks [37].

Another study explored the use of transfer learning with pre-trained convolutional neural networks to classify different coffee bean species based on visual features. Models such as SqueezeNet, Inception V3, VGG16, and VGG19 were evaluated on a dataset of images representing three commercial coffee types. Among them, SqueezeNet showed the highest classification performance, with an average success rate of 87.3 percent. This work highlighted the value of compact architectures and transfer learning strategies for effective and low-resource bean type discrimination [38].

In a focused comparison of object detection approaches, a study evaluated multiple YOLO variants for detecting and classifying green coffee beans and their defects. Among the versions tested, a custom-tuned YOLOv8n model demonstrated superior performance in terms of precision, recall, and mean average precision. The study stressed the importance of tailoring model parameters and labeling strategies to maximize detection sensitivity, especially for subtle defects, and confirmed the practicality of YOLO-based models for real-time quality control in coffee processing environments [39].

A fourth work presented a real-time deep learning system for automatic coffee bean classification, integrating YOLOv8 with cloud-based infrastructure and mobile applications. The model was deployed using image-streaming technologies and open-source libraries such as OpenCV, achieving detection speeds of 1 to 3 seconds per sample. The study focused on practical deployment in large-scale coffee facilities and demonstrated the feasibility of mobile-compatible deep learning pipelines for end-to-end classification across multiple coffee bean varieties [40].

Building on these advances, the current study proposes a fully image-based coffee bean quality evaluation framework using a custom-trained YOLOv8n model. While prior research has explored either defect detection, species classification, or system deployment, our work integrates object detection, interpretability, and sample-level quality scoring into a unified pipeline. Unlike systems focused solely on classification, the proposed method detects five common defect types and calculates a quality score that aligns with commercial grading standards. Additionally, visual explanations generated by Grad-CAM enhance model transparency, addressing concerns about reliability in automated inspection. Through its emphasis on both technical performance and interpretability, this study contributes a scalable and reproducible approach for real-time coffee quality grading suitable for industrial implementation.

The results of this study demonstrate the effectiveness of the proposed YOLOv8n-based framework in detecting and classifying defective green coffee beans with high accuracy and operational efficiency. The model achieved strong performance across standard object detection metrics, including precision, recall, and mean average precision, indicating its robustness in identifying five distinct defect types under variable imaging conditions. The system's ability to generate interpretable quality scores aligned with commercial grading standards provides a practical and objective alternative to manual sorting. Moreover, the integration of Grad-CAM visualizations enhanced transparency, allowing domain experts to verify the decision-making process. Despite these promising outcomes, certain limitations remain. The model's performance may be influenced by variations in lighting, bean orientation, or overlapping instances in densely packed images. Additionally, the dataset size and controlled imaging environment limit generalizability to more complex real-world settings, such as conveyor-based inspection systems or field-level analysis.

To address these limitations and expand the utility of the system, future work will focus on several directions. First, the dataset will be augmented with more diverse samples, including beans from different geographical origins, post-harvest conditions, and imaging setups, to improve generalization and robustness. Second, domain adaptation techniques will be explored to allow the model to perform consistently across varying operational environments. Incorporating multispectral or hyperspectral imaging data could further enhance defect detection sensitivity, especially for internal or subtle quality indicators not visible in standard RGB images. In addition, expanding the system into a real-time, mobile-enabled platform integrated with edge computing could enable on-site quality assessment for smallholder farmers and processing facilities. Finally, the development of a feedback-driven training loop with human-in-the-loop correction mechanisms may ensure continuous improvement and adaptability of the model in real-world deployment scenarios.

5. Conclusion

In this study, we developed a lightweight, fully image-based deep learning framework for the automated detection of green coffee bean defects and the prediction of sample-level quality grades. By employing a YOLOv8n model trained on annotated photographic data, the system accurately identified five common defect types and computed quality scores based on the proportion of non-defective beans. The pipeline eliminates the need for manual sorting or expert evaluation, offering a rapid, consistent, and reproducible approach to coffee quality assessment. The model demonstrated strong performance across key detection and classification metrics, while visual explainability through Grad-CAM enhanced transparency and interpretability. This end-to-end solution provides a practical alternative to traditional grading methods and holds strong potential for adoption in industrial and commercial coffee production environments. Future work will focus on expanding dataset diversity, incorporating advanced imaging techniques, and enabling mobile or edge-based deployment to support real-time, field-level applications in quality monitoring and traceability.

Author Contributions: E.O. solely contributed to the conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft, writing—review & editing, visualization, and supervision of this study. The author has read and agreed to the published version of the manuscript.

Conflicts of competing interests: In this article, the author states that they have no competing financial interests or personal affiliations.

References

1. S. Jha, C. M. Bacon, S. M. Philpott, R. A. Rice, V. E. Méndez, and P. Läderach, "A review of ecosystem services, farmer livelihoods, and value chains in shade coffee agroecosystems," *Integr. Agric. Conserv. ecotourism examples from F.*, pp. 141–208, 2011.

2. R. J. Guimarães, F. M. Borém, J. Shuler, A. Farah, and J. C. Peres Romero, "Coffee growing and post-harvest processing," 2019.

3. M. Tesfa, "Review on post-harvest processing operations affecting coffee (Coffea Arabica L.) quality in Ethiopia," *J. Environ. Earth Sci*, vol. 9, no. 12, pp. 30–39, 2019.

4. M. Haile and W. H. Kang, "The harvest and post-harvest management practices' impact on coffee quality," in *Coffee-Production and Research*, IntechOpen, 2019.

5. M. Kleinwächter, G. Bytof, and D. Selmar, "Coffee beans and processing," in *Coffee in health and disease prevention*, Elsevier, 2025, pp. 105–114.

6. M. de S. G. Barbosa, M. B. dos Santos Scholz, C. S. G. Kitzberger, and M. de Toledo Benassi, "Correlation between the composition of green Arabica coffee beans and the sensory quality of coffee brews," *Food Chem.*, vol. 292, pp. 275–280, 2019.

7. D. Selmar, M. Kleinwächter, and G. Bytof, "Metabolic responses of coffee beans during processing and their impact on coffee flavor," *Cocoa coffee Ferment.*, pp. 431–476, 2014.

8. D. Singh, R. R. Sharma, and A. K. Kesharwani, "Postharvest losses of horticultural produce," in *Postharvest handling and diseases of horticultural produce*, CRC Press, 2021, pp. 1–24.

9. G. Farrell, R. J. Hodges, P. W. Wareing, A. N. Meyer, and S. R. Belmain, "Biological Factors in Post-Harvest Quality," *Crop Post-Harvest Sci. Technol. Princ. Pract.*, vol. 1, pp. 93–140, 2002.

10. D. A. Sukha, "The Grading and Quality of Dried Cocoa Beans," in *Drying and Roasting of Cocoa and Coffee*, CRC Press, 2019, pp. 89–139.

11. L. Agnusdei, P. P. Miglietta, and G. P. Agnusdei, "Quality in beans: tracking and tracing coffee through automation and machine learning," *EuroMed J. Bus.*, 2024.

12. M. García, J. E. Candelo-Becerra, and F. E. Hoyos, "Quality and defect inspection of green coffee beans using a computer vision system," *Appl. Sci.*, vol. 9, no. 19, p. 4195, 2019.

13. B. Chandu, R. Surendran, and R. Selvanarayanan, "To Instant of Coffee Beans using K-nearest Algorithm Over Clustering for Quality and Sorting Process," in *2024 International Conference on IT Innovation and Knowledge Discovery (ITIKD)*, IEEE, 2025, pp. 1–7.

14. R. W. Thurston, J. Morris, and S. Steiman, *Coffee: A comprehensive guide to the bean, the beverage, and the industry*. Bloomsbury Publishing PLC, 2013.

15. T. Bidyalakshmi *et al.*, "Application of artificial intelligence in food processing: Current status and future prospects," *Food Eng. Rev.*, pp. 1–28, 2024.

16. R. Singh, C. Nickhil, R. Nisha, K. Upendar, B. Jithender, and S. C. Deka, "A comprehensive review of advanced deep learning approaches for food freshness detection," *Food Eng. Rev.*, vol. 17, no. 1, pp. 127–160, 2025.

17. R. Archana and P. S. E. Jeevaraj, "Deep learning models for digital image processing: a review," *Artif. Intell. Rev.*, vol. 57, no. 1, p. 11, 2024.

18. M. M. Adnan, M. S. M. Rahim, A. Rehman, Z. Mehmood, T. Saba, and R. A. Naqvi, "Automatic image annotation based on deep learning models: a systematic review and future challenges," *IEEe Access*, vol. 9, pp. 50253–50264, 2021.

19. P. Poltronieri and F. Rossi, "Challenges in specialty coffee processing and quality assurance," *Challenges*, vol. 7, no. 2, p. 19, 2016.

20. V. Kumar, P. S. S. Aydav, and S. Minz, "Crop Seeds Classification Using Traditional Machine Learning and Deep Learning Techniques: A Comprehensive Survey," *SN Comput. Sci.*, vol. 5, no. 8, p. 1031, 2024.

21. X. Huang *et al.*, "Application of Image Computing in Non-Destructive Detection of Chinese Cuisine," *Foods*, vol. 14, no. 14, p. 2488, 2025.

22. J. Ma *et al.*, "Applications of computer vision for assessing quality of agri-food products: a review of recent research advances," *Crit. Rev. Food Sci. Nutr.*, vol. 56, no. 1, pp. 113–127, 2016.

23. M. M. Taye, "Theoretical understanding of convolutional neural network: Concepts, architectures, applications, future directions," *Computation*, vol. 11, no. 3, p. 52, 2023.

24. L. A. Gatys, A. S. Ecker, and M. Bethge, "Texture and art with deep neural networks," *Curr. Opin. Neurobiol.*, vol. 46, pp. 178–186, 2017.

25. L. Aziz, M. S. B. H. Salam, U. U. Sheikh, and S. Ayub, "Exploring deep learning-based architecture, strategies, applications and current trends in generic object detection: A comprehensive review," *Ieee Access*, vol. 8, pp. 170461–170495, 2020.

26. G. Cheng, X. Xie, J. Han, L. Guo, and G.-S. Xia, "Remote sensing image scene classification meets deep learning: Challenges, methods, benchmarks, and opportunities," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 3735–3756, 2020.

27. Z.-Q. Zhao, P. Zheng, S. Xu, and X. Wu, "Object detection with deep learning: A review," *IEEE Trans. neural networks Learn. Syst.*, vol. 30, no. 11, pp. 3212–3232, 2019.

28. T. Diwan, G. Anirudh, and J. V Tembhurne, "Object detection using YOLO: challenges, architectural successors, datasets and applications," *Multimed. Tools Appl.*, vol. 82, no. 6, pp. 9243–9275, 2023.

29. K. Kanna S, K. Ramalingam, and P. PC, "YOLO deep learning algorithm for object detection in agriculture: a review.," *J. Agric. Eng.*, vol. 55, no. 4, 2024.

30. A. Lavanya, M. Arakeri, and B. J. Ambika, "Advancements in Coffee Bean Quality Assessment Using Computer Vision and Deep Learning Techniques," in *2025 International Conference on Next Generation Communication & Information Processing (INCIP)*, 2025, pp. 758–763. doi: 10.1109/INCIP64058.2025.11019029.

31. E. Hassan, "Enhancing coffee bean classification: a comparative analysis of pre-trained deep learning models," *Neural Comput. Appl.*, vol. 36, no. 16, pp. 9023–9052, 2024, doi: 10.1007/s00521-024-09623-z.

32. Y. Chen, B. Gao, and W. Lu, "Recent Research Advancements of Coffee Quality Detection: Targeted Analyses vs. Nontargeted Fingerprinting and Related Issues," *J. Food Qual.*, vol. 2023, no. 1, p. 6156247, 2023, doi: https://doi.org/10.1155/2023/6156247.

33. I. A. F. Anto, A. Munandar, J. W. Wibowo, T. I. Salim, and O. Mahendra, "Coffee Bean Roasting Levels Detection: A Systematic Review," in *2023 IEEE 7th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, 2023, pp. 146–151. doi: 10.1109/ICITISEE58992.2023.10404775.

34. B. Nair, "CBD_Coffee Bean Dataset," 2024, *Mendeley Data*. doi: 10.17632/52877z55vr.1.

35. G. Naidu, T. Zuva, and E. M. Sibanda, "A review of evaluation metrics in machine learning algorithms," in *Computer science on-line conference*, Springer, 2023, pp. 15–25.

36. Ž. Vujović, "Classification model evaluation metrics," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 6, pp. 599–606, 2021.

37. C.-H. Hsia, Y.-H. Lee, and C.-F. Lai, "An Explainable and Lightweight Deep Convolutional Neural Network for Quality Detection of Green Coffee Beans," *Appl. Sci.*, vol. 12, no. 21, 2022, doi: 10.3390/app122110966.

38. Y. Unal, Y. S. Taspinar, I. Cinar, R. Kursun, and M. Koklu, "Application of Pre-Trained Deep Convolutional Neural Networks for Coffee Beans Species Detection," *Food Anal. Methods*, vol. 15, no. 12, pp. 3232–3243,

2022, doi: 10.1007/s12161-022-02362-8.

39. H. L. Gope, H. Fukai, F. M. Ruhad, and S. Barman, "Comparative analysis of YOLO models for green coffee bean detection and defect classification," *Sci. Rep.*, vol. 14, no. 1, p. 28946, 2024, doi: 10.1038/s41598-024-78598-7.

40. H.-D. Thai, H.-J. Ko, and J.-H. Huh, "Coffee Bean Defects Automatic Classification Realtime Application Adopting Deep Learning," *IEEE Access*, vol. 12, pp. 126503–126517, 2024, doi: 10.1109/ACCESS.2024.3452552.