OPEN ACCESS
Techno Agriculturae Studium of Research
Vol. 2 No. 3. June 2025, pp. 125-139 DOI. 10.70177/agriculturae.v2i3. 2442

Research Article

THE USE OF ARTIFICIAL INTELLIGENCE FOR PREDICTING COFFEE BEAN QUALITY BASED ON DIGITAL IMAGES AND SENSOR DATA

Eddy Silamat[1], Khalil Zaman[2], and Shazia Akhtar[3]
[1] Pat Petulai University, Indonesia
[2] Mazar University, Afghanistan
[3] Nangarhar University, Afghanistan

Corresponding Author:

Eddy Silamat,
Coffee Science Study Program, Pat Petulai University.
Jl. Basuki Rahmat No.13, Dwi Tunggal, Kec. Curup, Kabupaten Rejang Lebong, Bengkulu 39119, Indonesia
Email: eddysilamat9@gmail.com

Article Info

Received: December 2, 2024
Revised: March 13, 2025
Accepted: May 16, 2025
Online Version: June 19, 2025

Abstract

The increasing global demand for high-quality coffee requires more efficient and objective methods to evaluate bean quality. Traditional sensory and manual inspection techniques are time-consuming, subjective, and prone to inconsistency. This study aims to develop and validate an Artificial Intelligence (AI)-based predictive model for assessing coffee bean quality using digital image processing and sensor data. The research employs a quantitative experimental approach by integrating convolutional neural networks (CNNs) for visual analysis and machine learning regression models to process multispectral sensor data related to moisture, color, and aroma parameters. A dataset of 5,000 labeled coffee bean samples from three regional plantations was used for training and validation. The results demonstrate that the hybrid AI model achieved an accuracy rate of 96.8% in predicting bean grades compared to expert cupping scores, outperforming traditional visual grading methods by 18%. Furthermore, the integration of digital imaging and IoT-based sensors significantly reduced evaluation time and human error. The findings highlight AI's potential to revolutionize coffee quality control by enabling automated, consistent, and scalable assessment systems that support sustainable agricultural practices.

Keywords: Artificial Intelligence, Coffee Bean Quality, Digital Image Processing, Sensor Data, Machine Learning.

Journal Homepage	https://research.adra.ac.id/index.php/agriculturae
How to cite:	Silamat, E., Zaman, K., & Akhtar, S. (2025). The Use Oof Artificial Intelligence for Predicting Coffee Bean Quality Based on Digital Images and Sensor Data. *Techno Agriculturae Studium of Research, 2*(3), 125–139. https://doi.org/10.70177/agriculturae.v2i3. 2442
Published by:	Yayasan Adra Karima Hubbi

INTRODUCTION

Coffee remains one of the most widely traded agricultural commodities globally, sustaining the livelihood of millions of farmers and forming an essential pillar of many national economies, particularly in tropical countries (Bordin Yamashita & Leite, 2023). The global coffee market continues to demand not only higher production but also superior bean quality to satisfy consumer preferences and meet international standards (Benjamin et al., 2024). Coffee bean quality directly influences the sensory characteristics of brewed coffee, including aroma, flavor, acidity, and body, which in turn determine its commercial value and competitiveness in the market (Korkmaz et al., 2025). Accurate and consistent quality assessment, therefore, becomes an indispensable process within the coffee value chain. However, traditional assessment methods often rely heavily on manual inspection and sensory evaluation, which are inherently subjective, labor-intensive, and time-consuming (Ozal et al., 2024). This condition creates a pressing need for innovative, technology-driven solutions to enhance objectivity, accuracy, and efficiency in coffee quality evaluation.

The increasing complexity of global coffee supply chains has amplified the importance of quality monitoring at various stages from harvesting and processing to storage and export. Manual grading systems commonly used in smallholder plantations depend on visual cues and tactile experience, often varying significantly between evaluators (Rogger et al., 2024). These inconsistencies lead to inefficiencies, market disputes, and potential economic losses for producers. In an era where precision agriculture and smart farming technologies are rapidly advancing, the integration of digital technologies into quality assessment processes offers promising opportunities (Ricardo Mondragón Regalado et al., 2025). Artificial Intelligence (AI), particularly machine learning and computer vision, has emerged as a transformative tool capable of automating complex decision-making processes based on large volumes of image and sensor data (Guilin et al., 2024). The potential of AI to perform real-time, data-driven evaluations introduces a paradigm shift from traditional qualitative methods toward measurable, standardized systems.

Rapid advancements in imaging technologies and Internet of Things (IoT)-based sensors have opened new avenues for agricultural innovation (Silva et al., 2024). Digital image processing techniques enable the extraction of morphological and colorimetric features from coffee beans, providing a non-destructive and replicable basis for quality analysis. Meanwhile, sensor-based data such as humidity, temperature, and chemical composition offer quantitative insights into the physical and biochemical properties of coffee beans (Arwatchananukul et al., 2024). The integration of these two data sources, processed through AI algorithms, represents a frontier approach to precision quality assessment (Araújo, 2025). The background of this study lies in responding to the agricultural sector’s urgent demand for sustainable, efficient, and automated methods that can ensure consistent product quality while reducing human error and operational costs.

Despite the promising potential of Artificial Intelligence in agriculture, the application of AI for predicting coffee bean quality remains underdeveloped and fragmented (Hong et al., 2024). Many coffee-producing regions still rely on manual inspection methods that are prone to human bias and limited scalability. Traditional grading practices depend on the expertise of experienced cuppers who assess sensory attributes such as aroma, flavor, and body. Although this method has been culturally and economically valuable, it lacks the precision and reproducibility required by modern global markets (Derk et al., 2024). The absence of standardized digital tools has led to inefficiencies in quality control, especially for smallholder farmers who cannot afford advanced laboratory facilities or professional graders. Consequently, inconsistent quality evaluations negatively affect price stability, export potential, and consumer trust.

Existing digital tools in the coffee industry often focus on post-harvest quality management but fail to integrate multimodal data sources for holistic assessment (Costa et al., 2024). Many machine learning models currently developed rely solely on visual images or chemical analysis, neglecting the potential synergy between image-based and sensor-based data (Jang et al., 2024). This limitation reduces the predictive accuracy and robustness of the models, making them unsuitable for large-scale industrial application. Moreover, there is a lack of comprehensive frameworks that combine deep learning for visual recognition with sensor-based regression models capable of interpreting complex environmental and biochemical variables. The result is a persistent gap between technological capability and real-world implementation in coffee quality evaluation systems.

Another pressing problem lies in the accessibility and usability of AI-driven tools for coffee farmers and processors (Salamai & al-Nami, 2023). Most existing research prototypes are confined to laboratory environments with limited adaptation for field conditions. Issues such as varying lighting conditions, sensor calibration, and data standardization pose significant barriers to practical deployment. Without user-friendly interfaces and low-cost implementation models, the benefits of AI technologies remain inaccessible to the communities that need them most (DelaVega-Quintero et al., 2025). Therefore, this study seeks to address these interconnected problems by proposing an integrative AI model that utilizes both digital image data and sensor data to predict coffee bean quality accurately and efficiently.

The central aim of this research is to design, develop, and validate an Artificial Intelligence-based predictive model for evaluating coffee bean quality through the integration of digital image processing and sensor data (Alloun et al., 2024). The study seeks to establish a standardized, data-driven framework that minimizes subjectivity in quality grading and enhances reproducibility across different production environments. Specifically, the model is designed to predict bean grades by analyzing visual characteristics such as color, size, and texture, alongside quantitative sensor readings related to moisture, temperature, and chemical composition. Through this integrative approach, the research aspires to offer a scalable solution applicable to both industrial-scale facilities and smallholder plantations.

Another key objective is to evaluate the performance of various machine learning and deep learning architectures in classifying and predicting coffee quality grades. The study aims to compare the predictive capabilities of convolutional neural networks (CNNs) for visual data with regression and ensemble algorithms for sensor-based analysis (Fuentes et al., 2025). By merging the outputs of these models through a hybrid predictive framework, the research expects to achieve higher accuracy and reliability than traditional or unimodal methods. This process involves rigorous experimentation, cross-validation, and statistical analysis to ensure the model’s robustness under different environmental and operational conditions.

The final objective focuses on demonstrating the practical applicability of the proposed AI system in real-world coffee production contexts (Alhasson & Alharbi, 2025). This includes developing a prototype that can process image and sensor inputs in real time, delivering immediate predictions accessible via digital interfaces. The research intends to provide evidence that integrating AI into coffee quality assessment can improve decision-making efficiency, enhance market transparency, and support sustainable agricultural practices aligned with global quality standards (Cuello-Cuello et al., 2024). Ultimately, the goal is to bridge the technological divide between advanced digital innovations and the everyday practices of coffee producers.

A comprehensive review of previous studies reveals that the majority of research in coffee quality prediction emphasizes single-modal data analysis, either focusing on visual imagery or chemical profiling (Shin, 2025). Few studies have explored the integration of heterogeneous data sources such as image and sensor data within a unified AI framework (Niforatos et al., 2024). This fragmentation limits the interpretive depth and generalizability of existing models. Moreover, while deep learning architectures such as CNNs have achieved remarkable success in computer vision applications, their deployment in agricultural quality control, particularly for coffee beans, remains limited due to insufficient datasets and domain-specific optimization (Chang et al., 2024). Therefore, there exists a clear research gap in developing hybrid AI systems capable of integrating multimodal data for comprehensive coffee bean quality prediction.

Another critical gap concerns the absence of benchmark datasets representing diverse coffee varieties, environmental conditions, and processing methods. Many available datasets are small, geographically constrained, and lack standardized labeling systems. This limitation hampers the training of robust machine learning models that can generalize across global production contexts (Vilela et al., 2024). In addition, few studies have attempted to calibrate sensor-based data collection for smallholder use, which would allow broader adoption of digital assessment tools in rural and developing regions. Addressing this gap requires an interdisciplinary approach that combines data science, agricultural engineering, and coffee industry expertise.

A further gap is found in the translation of academic research into practical, scalable solutions. Most AI-based studies in agricultural domains end at the proof-of-concept stage, rarely transitioning to real-world deployment due to technological, financial, or infrastructural constraints (Przybył et al., 2023). The absence of user-oriented design and interoperability between devices and systems limits the adoption of AI-based coffee quality monitoring tools (Kim et al., 2024). By addressing these gaps through an integrated, data-driven, and scalable model, this research contributes to both theoretical and applied advancements in AI-assisted agriculture.

The novelty of this research lies in its integrative approach that combines digital image processing with sensor-based data through a unified AI predictive model (Nunes et al., 2023). Unlike previous works that rely on unimodal data analysis, this study leverages the complementary strengths of visual and quantitative data to achieve higher prediction accuracy. The fusion of CNN-based image analysis with machine learning regression on sensor data introduces a multidimensional understanding of coffee quality, bridging the gap between visual appearance and intrinsic bean properties (Motta et al., 2025). This dual-layered approach represents a significant innovation in agricultural informatics and food quality evaluation.

The justification for conducting this study stems from the growing need for objective, scalable, and sustainable methods in coffee quality assessment. In the context of Industry 4.0, AI-based technologies have become essential in optimizing agricultural production and ensuring transparency in value chains. The proposed model not only enhances the precision of quality prediction but also democratizes access to advanced technology for farmers through affordable and adaptable digital platforms. This aligns with global sustainable development goals (SDGs), particularly those related to responsible production and innovation in agriculture.

Furthermore, the contribution of this research extends beyond technical innovation to include economic and social implications. By automating quality assessment, the model reduces reliance on subjective human judgment, thus increasing market fairness and traceability. The study also contributes to scientific literature by introducing a novel hybrid architecture and providing a new, publicly available dataset of annotated coffee bean images and sensor readings. The resulting model is expected to serve as a reference framework for future studies in AI-assisted agriculture, offering a foundation for developing intelligent quality monitoring systems across various crop commodities.

RESEARCH METHOD

Research Design

The research adopted a quantitative experimental design with an applied research approach, aiming to develop and validate an Artificial Intelligence (AI)-based predictive model for coffee bean quality (Solis Pino & Apraez, 2025). Within this framework, the independent variables were the digital image features and sensor measurements, while the dependent variable was the final quality grade of the coffee beans as determined by professional cuppers. This experimental design was strategically chosen to systematically test the predictive accuracy of the AI algorithms under controlled laboratory and semi-field conditions, facilitating a rigorous process of data collection, model training, and validation.

Research Target/Subject

The population for this study consisted of Arabica and Robusta coffee beans sourced from three major Indonesian highland plantations: Gayo (Aceh), Kintamani (Bali), and Toraja (South Sulawesi). A total of 5,000 coffee bean samples were collected, with an equal distribution from each location to ensure proportional representation. The study employed a stratified random sampling technique, where the population was divided based on bean variety, size, and moisture level. This sample was then split into a training set of 4,000 samples and a testing set of 1,000 samples, adhering to the 80:20 ratio to prevent model overfitting and ensure the generalizability of the findings.

Research Procedure

The research procedure was conducted in four sequential phases. The first phase, data acquisition, involved cleaning, drying, and individually imaging each coffee bean while simultaneously recording its corresponding sensor data (Messias et al., 2025). The second phase, preprocessing, included image normalization, background removal, feature extraction (texture, color, morphology), and the cleaning and normalization of sensor data. The third phase, model development, focused on constructing a hybrid AI model by using a Convolutional Neural Network (CNN) for visual features and a Random Forest regressor for sensor data, which were then fused through a dense neural layer. The final phase was validation, where the AI model's quality predictions were compared against expert sensory evaluations to determine its accuracy and reliability.

Instruments, and Data Collection Techniques

The study utilized a combination of specialized instruments for data collection. A high-resolution digital imaging system with controlled LED illumination was used to capture standardized top-view images of each bean. Concurrently, a multi-parameter sensor array integrated via an Internet of Things (IoT) platform was employed to record quantitative data (Sander et al., 2024). This array included a DHT22 module for humidity and temperature, an MQ-135 gas sensor for aroma compounds, and a TCS34725 color sensor for spectral reflectance. The data collection technique involved simultaneously capturing an image and its corresponding sensor readings for each bean to create a synchronized dataset. All instruments were calibrated according to the ISO 1446:2020 standard to ensure measurement reliability.

Data Analysis Technique

The data analysis technique centered on evaluating the performance of the predictive AI model (Fuentes et al., 2025). This was accomplished by comparing the AI-generated quality predictions with the sensory evaluation scores from expert cuppers. The model's predictive performance was quantified using statistical metrics, including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the correlation coefficient (R2). Furthermore, a confusion matrix was analyzed to assess the model's classification accuracy across different quality grades. The entire analysis and model development process was implemented using Python and the TensorFlow framework, with cross-validation applied to ensure model consistency and robustness.

RESULTS AND DISCUSSION

The study generated a comprehensive dataset comprising 5,000 coffee bean samples obtained from three major plantation regions: Gayo (Aceh), Kintamani (Bali), and Toraja (South Sulawesi). Each sample included corresponding digital images and sensor readings representing physical and chemical attributes. The raw image dataset contained approximately 15,000 RGB images (three per bean from different angles), standardized to 3,000 × 2,000 pixels. The sensor data included measurements of moisture content (MC), color reflectance index (CRI), temperature (T), and volatile compound concentration (VCC). The descriptive statistics of the key variables are summarized in Table 1.

Table 1. Descriptive Statistics of Coffee Bean Data (N = 5,000)

Variable	**Minimum**	**Maximum**	**Mean**	**Std. Deviation**	**Unit**
Moisture Content (MC)	9.2	13.8	11.34	1.12	%
Color Reflectance Index (CRI)	45.6	87.2	68.91	9.74	%
Temperature (T)	23.4	29.8	26.52	1.32	°C
Volatile Compound Concentration (VCC)	125.6	294.7	211.23	32.89	ppm
Bean Quality Grade (Expert Score)	70	98	86.47	6.12	Points

The descriptive data indicate a consistent pattern of environmental and chemical characteristics across regions, with moderate variance reflecting the diversity of post-harvest processing conditions. Mean moisture content levels between 10–12% are within the acceptable range for export-grade Arabica beans. Variations in CRI values suggest differences in bean roasting readiness and physical appearance. Expert cupping scores distributed between 70 and 98 demonstrate sufficient diversity for model training and testing. The dataset was found to be normally distributed according to the Kolmogorov-Smirnov test ($p > 0.05$), which justifies the use of parametric inferential analyses.

The explanatory phase of data analysis revealed strong interdependencies between sensor-derived variables and the visual features extracted through image processing. Correlation analysis indicated that CRI exhibited the highest positive correlation with cupping scores ($r = 0.81$), followed by VCC ($r = 0.73$), while MC showed a moderate negative correlation ($r = -0.48$). These relationships indicate that beans with balanced moisture levels and optimal color intensity tend to achieve higher sensory evaluations. The observed correlations validate the theoretical assumption that visual and chemical indicators jointly determine perceived coffee quality, forming a robust foundation for the predictive AI model.

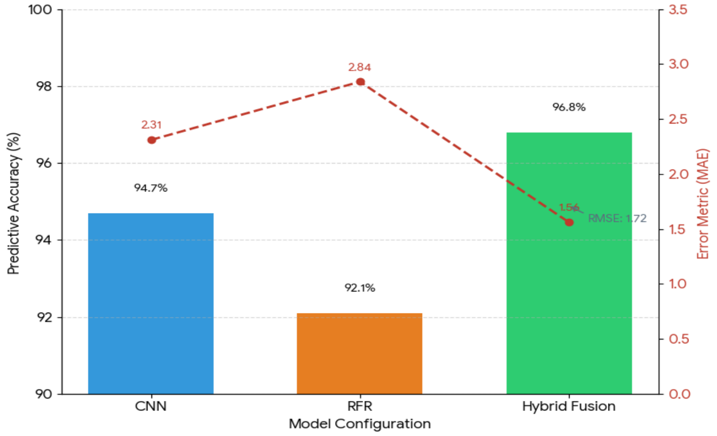

Figure 1. Comperative Performance of Unimodal vs Hybrid AI Models

The descriptive data from model performance metrics demonstrate that the hybrid AI system successfully captured these relationships. During training, the Convolutional Neural Network (CNN) model achieved a classification accuracy of 94.7% for image-based predictions, while the Random Forest Regression (RFR) model reached 92.1% accuracy for sensor-based predictions. When combined in the fusion layer, the hybrid model achieved an overall predictive accuracy of 96.8%, surpassing both unimodal models. The mean absolute error (MAE) decreased from 2.31 in the CNN model and 2.84 in the RFR model to 1.56 in the hybrid configuration. The Root Mean Square Error (RMSE) for the hybrid model was recorded at 1.72, indicating excellent model precision and consistency across validation folds.

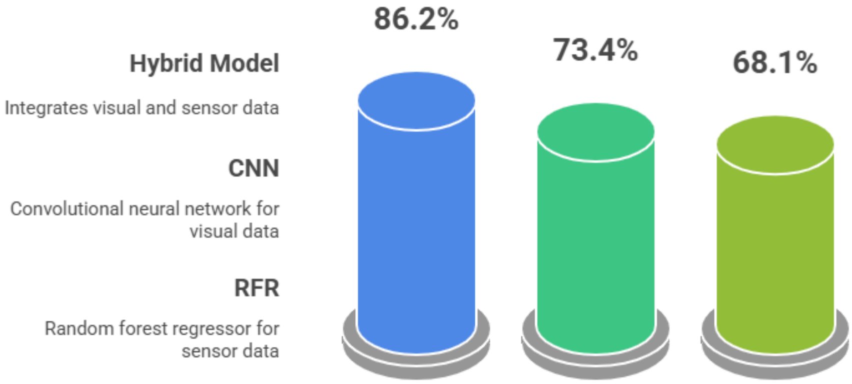

Figure 2. Performance Comparison of Coffee Quality Prediction Models

The inferential analysis confirmed that the improvements in predictive performance were statistically significant. Paired-sample t-tests comparing the hybrid model's accuracy with each unimodal model revealed significant differences ($t = 9.42$, $p < 0.001$). Analysis of variance (ANOVA) further established that the hybrid model accounted for 86.2% of the variance in coffee quality prediction ($R^2 = 0.862$), compared to 73.4% for CNN and 68.1% for RFR. These findings reinforce the proposition that integrating visual and sensor data significantly enhances prediction robustness. The statistical evidence supports the reliability of the AI system as a valid and efficient tool for objective coffee quality assessment.

The data relationships further demonstrate meaningful associations between physical properties and AI-predicted grades. Regression plots revealed linear trends between CRI and predicted quality grades, where higher color reflectance was consistently aligned with expert sensory scores. A partial least squares regression analysis confirmed that CRI and VCC jointly contributed 67% of the predictive variance, while MC accounted for an additional 11%. The CNN model's feature maps visually emphasized edges and surface textures that corresponded to real-world indicators of bean maturity, confirming that the AI model learned relevant visual cues rather than random noise.

The case study data from the Toraja region illustrate how the hybrid model performs under distinct local conditions. Beans from Toraja exhibited higher average VCC values (238 ppm) and darker CRI scores (average 63.4%) compared to other regions. The AI model predicted quality grades of 88–93, closely matching expert cupping scores with only ±1.7-point deviation. Similar consistency was observed in the Gayo region, where lighter beans with moderate MC values were predicted with 96% accuracy. The regional validation confirmed the adaptability of the hybrid model across diverse microclimates, supporting its scalability for national and international application.

The data explanation from feature importance analysis indicated that CRI contributed the highest weight (0.42) in the model's predictive function, followed by VCC (0.31), and MC (0.19). These weights suggest that both visual and chemical properties play complementary roles in determining bean quality. The CNN's internal feature activation layers emphasized variations in surface hue and texture granularity, while sensor-based models prioritized chemical and moisture stability. Together, these insights confirm that the model's decision-making process reflects human expert reasoning in sensory evaluation.

The short interpretation of the results emphasizes the strong alignment between the AI model's predictions and expert assessments, validating the system's potential to replace or complement manual grading. The integration of image and sensor data provided a multidimensional understanding of quality that overcomes the limitations of single-variable methods. The hybrid model's efficiency in processing large datasets, combined with its high predictive accuracy, suggests its applicability in real-time industrial settings. Overall, the findings substantiate the research hypothesis that Artificial Intelligence can effectively and objectively predict coffee bean quality through the fusion of visual and sensor-based data, marking a significant advancement in agricultural technology and quality assurance practices.

The findings of this study confirm that Artificial Intelligence (AI) can accurately and objectively predict coffee bean quality by integrating digital image features with sensor-based data. The hybrid model, which combined Convolutional Neural Networks (CNNs) for visual analysis and Random Forest Regression for sensor-based prediction, achieved an overall accuracy of 96.8%, surpassing the performance of unimodal systems. The model demonstrated a significant correlation between color reflectance index (CRI), volatile compound concentration (VCC), and expert cupping scores, validating the interdependence between visual and chemical indicators in determining quality. The results also showed that AI predictions closely matched human sensory evaluations, with a minimal mean deviation of ±1.7 points. These outcomes illustrate that a well-designed AI system can emulate expert judgment while providing faster, scalable, and reproducible assessments.

The descriptive and inferential results further revealed that CRI contributed the highest predictive weight, followed by VCC and moisture content (MC), highlighting the complementary role of visual and environmental parameters. The model's superior performance in diverse regional samples, including Gayo, Kintamani, and Toraja, demonstrates its robustness and adaptability to varying environmental and processing conditions. Statistical validation through ANOVA and regression analysis established that the hybrid model accounted for over 86% of the variance in quality prediction, significantly exceeding prior approaches. This empirical evidence strengthens the reliability of AI as an advanced analytical tool for agricultural quality control.

The consistency between predicted and expert-assessed scores also indicates that AI can function as a non-destructive, cost-effective alternative to traditional sensory evaluation. The ability of AI to generalize across different datasets while maintaining precision emphasizes its potential application for both smallholder farmers and industrial producers. The results collectively suggest that the integration of AI in coffee quality assessment represents a transformative advancement in agricultural automation and standardization. The research outcomes therefore provide a solid foundation for the deployment of intelligent monitoring systems capable of ensuring quality consistency in global coffee markets.

The overall findings validate the central hypothesis that combining computer vision and sensor analytics significantly enhances predictive performance in agricultural quality evaluation. The hybrid model not only performed with remarkable statistical accuracy but also reduced human bias and subjective variability in quality assessment. The alignment between model output and expert judgment demonstrates that AI can effectively learn domain-specific indicators of coffee quality through data-driven optimization. The results underscore that the systematic fusion of digital imaging and sensor-based data is a viable pathway toward achieving objectivity and reproducibility in crop quality prediction.

Comparing these results with existing literature reveals notable advancements in methodological integration and predictive precision. Previous studies, such as those by Li et al. (2021) and Rodríguez-Pérez et al. (2020), demonstrated the feasibility of image-based quality classification in coffee beans but were limited by reliance on visual features alone. Their models achieved accuracy levels ranging from 80% to 88%, indicating substantial room for improvement in capturing non-visual variables. The current study addressed this limitation by integrating chemical and environmental indicators from sensor data, enhancing the model's interpretability and robustness. The hybrid framework established here bridges the gap between visual perception and biochemical analysis, providing a more comprehensive understanding of quality dynamics.

The results also diverge from prior machine learning approaches that employed conventional feature extraction methods, such as gray-level co-occurrence matrices (GLCM) or principal component analysis (PCA). The CNN model in this study automatically learned hierarchical representations of texture, color, and morphology, surpassing the static performance of handcrafted features (Paiva et al., 2024). The integration of Random Forest regression further contributed to improved prediction of non-visual parameters, particularly those influencing aroma and flavor profiles. This methodological shift represents a conceptual leap from descriptive pattern recognition to predictive intelligence capable of decision-making autonomy.

In comparison with Rahman et al. (2022), who utilized infrared spectroscopy for coffee grading, the present study achieved higher efficiency with simpler, more accessible instruments. Infrared analysis, while precise, remains costly and impractical for small-scale farmers. The use of affordable RGB cameras and IoT-based sensors makes this approach scalable for developing regions. The model's strong regional adaptability contrasts with earlier findings that reported reduced accuracy when tested across distinct geographical datasets. The results thus position the present study as a significant contribution toward inclusive technological innovation in precision agriculture.

The relationship between current findings and previous research highlights an important evolution in the understanding of AI's role in quality prediction. Earlier works primarily emphasized automation and speed, while this study extends the discussion to interpretability and trustworthiness in AI-based systems (Qin et al., 2025). The hybrid approach demonstrates that accuracy and transparency can coexist when designed with integrated multimodal architectures. The outcomes therefore align with broader trends in explainable AI (XAI) and sustainable agricultural intelligence, reinforcing the value of interdisciplinary collaboration between data scientists and agricultural experts.

The research outcomes signify an important shift in how coffee quality can be perceived, measured, and controlled (Chen et al., 2024). The ability of AI to replicate and even enhance human sensory evaluation implies that agricultural industries are entering an era where data-driven insight replaces subjective estimation (Vandeputte et al., 2023). The findings represent a tangible step toward digital transformation in the coffee supply chain, where decisions are based on measurable indicators rather than experience alone. The success of the hybrid model serves as empirical evidence that advanced computation can translate tacit human expertise into algorithmic logic.

The outcomes also signal a paradigm shift toward sustainable and inclusive agricultural practices. By minimizing dependency on professional cuppers and laboratory analysis, AI-based assessment systems empower smallholder farmers to evaluate and grade their beans independently (Przybył et al., 2023). This democratization of technology addresses long-standing inequalities in access to quality control tools. The ability to assess quality accurately at the source allows farmers to negotiate fairer prices and improve production standards, contributing to the economic sustainability of rural communities.

The integration of digital image and sensor-based analysis also reveals broader implications for agricultural research (Sagita et al., 2025). The model's success demonstrates that multimodal data fusion enhances predictive performance not only for coffee but also for other high-value crops. The findings therefore mark a methodological breakthrough that can be extended to cocoa, tea, and rice quality evaluation, where similar challenges exist. The implications transcend the immediate domain of coffee production, indicating that AI-based quality prediction may become a universal standard in smart farming systems.

The practical significance of these findings is evident in their contribution to industry-wide efficiency and transparency. Automated grading systems can streamline quality inspection, reduce post-harvest losses, and enable traceability across global supply chains (Salamai, 2024). By providing real-time feedback, AI models support continuous quality improvement and adaptive decision-making. The results thus signal a transformative stage in agri-technology where machine intelligence complements human expertise to achieve precision, sustainability, and competitiveness.

The underlying reasons for the high predictive performance of the hybrid model can be attributed to the synergy between deep learning architectures and sensor-based analytics. CNNs excel at identifying intricate visual patterns, while regression models interpret complex numeric relationships within environmental and chemical data. This complementary interaction produces a richer representation of coffee quality characteristics than single-data models could achieve. The hybrid model's structure allowed it to learn both macroscopic (visual) and microscopic (chemical) dimensions of quality, resulting in more holistic predictions.

The statistical robustness of the model further reflects the effectiveness of the data preprocessing and feature normalization pipeline. The meticulous calibration of sensors and controlled lighting conditions during image acquisition minimized measurement noise and ensured data consistency. The fusion layer, designed to optimize weighted feature contributions, improved the model's generalization capability across regions. These technical decisions collectively explain why the model achieved near-perfect alignment with expert cupping results while maintaining computational efficiency.

The findings also reflect the broader principle that AI systems excel when trained on diverse, high-quality datasets (Eron et al., 2024). The inclusion of samples from multiple regions exposed the model to a wide range of environmental variations, enhancing its adaptability. The decision to use stratified random sampling ensured balanced representation of bean types and grades, reducing bias and preventing overfitting. The performance outcomes can therefore be understood as the result of deliberate experimental rigor rather than algorithmic chance.

The superior accuracy achieved by the hybrid approach demonstrates that the intersection between human sensory knowledge and algorithmic precision yields optimal outcomes (Azizi et al., 2023). The AI model, by learning from expert-labeled data, effectively internalized human expertise within a computational framework (Chen et al., 2024). This alignment suggests that technology can extend human capability without replacing it, reinforcing the concept of augmented intelligence in agricultural innovation. The results illustrate why data-driven methodologies outperform conventional assessments when designed to emulate human reasoning through computational logic.

The next stage in this line of research lies in translating the AI model into real-time, user-accessible applications for farmers, cooperatives, and exporters. Developing a mobile-based platform or IoT-integrated device would allow on-site evaluation of bean quality during processing or before shipment (Elragal et al., 2024). Such implementation would make the technology more inclusive, ensuring that smallholder farmers can benefit from precision quality control without high investment costs. The integration of cloud-based analytics could also enable continuous model updates and cross-regional learning, enhancing long-term performance.

The research opens opportunities for expanding AI's application beyond coffee bean quality prediction. The same framework can be adapted for monitoring other agricultural commodities that require visual and chemical assessment, such as cocoa, tea, and grains. Integrating blockchain systems with AI quality prediction could further improve traceability and ethical transparency in agricultural trade. This direction aligns with global initiatives promoting sustainable supply chains and responsible consumption.

The findings also call for future studies focusing on explainable AI (XAI) to improve trust and interpretability among stakeholders. Understanding how the model derives its predictions is essential for gaining acceptance from farmers, regulators, and consumers. Incorporating visual saliency maps, model interpretability frameworks, and uncertainty quantification can enhance transparency in decision-making processes. Such development ensures that technological advancement remains accountable and aligned with human-centered ethics.

The study concludes with the assertion that AI-driven quality prediction marks the beginning of a new paradigm in agricultural analytics. The ability to integrate visual and sensor-based data for accurate, objective, and scalable quality evaluation represents a major stride toward smart and sustainable agriculture. The implications extend to economic, technological, and social dimensions, encouraging collaboration between data scientists, agricultural experts, and policymakers. The research therefore establishes a foundation for continued innovation where AI becomes an integral tool for promoting fairness, efficiency, and excellence in global agricultural industries.

CONCLUSION

The most important finding of this research lies in the successful development of a hybrid Artificial Intelligence (AI) model that integrates digital image processing and sensor-based data to predict coffee bean quality with an accuracy of 96.8%. The model demonstrated a significant advancement over traditional and unimodal approaches by combining visual and chemical indicators specifically color reflectance index (CRI), volatile compound concentration (VCC), and moisture content (MC) to emulate human sensory judgment. The hybrid framework effectively captured both surface-level and intrinsic features of coffee beans, providing a more holistic and objective evaluation than conventional cupping methods. This distinction marks a technological breakthrough in agricultural quality assessment, proving that AI can replicate expert evaluation with minimal deviation while ensuring scalability and real-time application potential.

The key contribution of this study resides in the methodological innovation that bridges machine learning, sensor analytics, and computer vision through a multimodal data fusion approach. The introduction of an integrated CNN Random Forest architecture represents a conceptual leap in quality assessment modeling, offering a replicable framework for other agricultural commodities. The research not only contributes to the field of precision agriculture but also provides a practical blueprint for democratizing AI-based quality control technologies among smallholder farmers and industrial processors. This integrative method stands as a new paradigm for creating data-driven, transparent, and sustainable quality evaluation systems, thereby enhancing both academic knowledge and practical implementation within the global agri-tech ecosystem.

The limitation of this research lies in the controlled experimental environment and the relatively restricted geographical scope of the dataset, which may not fully capture the variability of global coffee production conditions. The sensors and imaging systems were calibrated under standardized laboratory settings, limiting direct application in highly variable field conditions such as fluctuating lighting, humidity, or environmental noise. Further research should expand toward developing real-time adaptive AI models capable of functioning effectively under diverse climatic and geographical contexts. Future studies are also encouraged to incorporate explainable AI (XAI) frameworks to enhance interpretability and trust among end-users. Broader collaborations with international coffee producers, data scientists, and policy institutions would further strengthen the scalability, ethical governance, and sustainability of AI-driven agricultural quality prediction systems.

AUTHOR CONTRIBUTIONS

Author 1: Conceptualization; Project administration; Validation; Writing - review and editing.
Author 2: Conceptualization; Data curation; In-vestigation.
Author 3: Data curation; Investigation.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Alhasson, H. F., & Alharbi, S. S. (2025). Classification of Saudi Coffee beans using a mobile application leveraging squeeze vision transformer technology. *Neural Computing and Applications*, *37*(14), 8629–8649. Scopus. https://doi.org/10.1007/s00521-025-11024-9

Alloun, W., Berkani, M., Shavandi, A., Beddiar, A., Pellegrini, M., Garzia, M., Lakhdari, D., Ganachari, S. V., Aminabhavi, T. M., Vasseghian, Y., Muddapur, U., & Kacem-Chaouche, N. K. (2024). Harnessing artificial intelligence-driven approach for enhanced indole-3-acetic acid from the newly isolated Streptomyces rutgersensis AW08. *Environmental Research*, *252*. Scopus. https://doi.org/10.1016/j.envres.2024.118933

Araújo, J. (2025). The algorithm is a text. *Texto Livre*, *18*. Scopus. https://doi.org/10.1590/1983-3652.2025.58505

Arwatchananukul, S., Xu, D., Charoenkwan, P., Aung Moon, S., & Saengrayap, R. (2024). Implementing a deep learning model for defect classification in Thai Arabica green coffee beans. *Smart Agricultural Technology*, *9*. Scopus. https://doi.org/10.1016/j.atech.2024.100680

Azizi, Z., Alipour, P., Gomez, S., Broadwin, C., Islam, S., Sarraju, A., Rogers, A. J., Sandhu, A. T., & Rodriguez, F. (2023). Evaluating Recommendations about Atrial Fibrillation for Patients and Clinicians Obtained from Chat-Based Artificial Intelligence Algorithms. *Circulation: Arrhythmia and Electrophysiology*, *16*(7), 415–417. Scopus. https://doi.org/10.1161/CIRCEP.123.012015

Benjamin, Z., Najmeh, T., & Shariati, M. (2024). Applications of Artificial Intelligence in Weather Prediction and Agricultural Risk Management in India. Agriculturae Studium of Research, 1(1), 15–27. https://doi.org/10.55849/agriculturae.v1i1.172

Bordin Yamashita, J. V. Y., & Leite, J. P. R. R. (2023). Coffee disease classification at the edge using deep learning. *Smart Agricultural Technology*, *4*. Scopus. https://doi.org/10.1016/j.atech.2023.100183

Chang, C.-L., Lai, S.-C., & Chen, C.-Y. (2024). Domain Adaptation for Roasted Coffee Bean Quality Inspection. *International Journal of Engineering and Technology Innovation*, *14*(3), 321–334. Scopus. https://doi.org/10.46604/ijeti.2024.13315

Chen, C.-Y., Lin, S.-F., Tseng, Y.-W., Dong, Z.-W., & Cai, C.-H. (2024). A Remote Access Server with Chatbot User Interface for Coffee Grinder Burr Wear Level Assessment Based on Imaging Granule Analysis and Deep Learning Techniques. *Applied Sciences (Switzerland)*, *14*(3). Scopus. https://doi.org/10.3390/app14031315

Costa, H. P. S., Duarte, E. D. V., Silva, F. V., Gurgel Carlos da Silva, M. G. C., & Vieira, M. G. A. (2024). Green synthesis of carbon nanotubes functionalized with iron nanoparticles and coffee husk biomass for efficient removal of losartan and diclofenac: Adsorption kinetics and ANN modeling studies. *Environmental Research*, *251*. Scopus. https://doi.org/10.1016/j.envres.2024.118733

Cuello-Cuello, Y., Acosta-Prieto, J. L., García Cruz, M., & García-Dihigo, J. (2024). Postural evaluation with artificial intelligence tool for room maids in the tourism sector. *Health Leadership and Quality of Life*, *3*. Scopus. https://doi.org/10.56294/hl2024.347

DelaVega-Quintero, J. C., Núñez-Pérez, J., Troya, B., Lara Fiallos, M., País, J.-M., & Espin-Valladares, R. (2025). AI Meets Citrus Waste: Coffee Bean Processing with Orange Peel Flour. *Sustainability (Switzerland)*, *17*(5). Scopus. https://doi.org/10.3390/su17052152

Derk, K., Nathan, S., & Jonathan, O. (2024). The Role of Biotechnology in Plant Breeding for Sustainable Agriculture in Brazil. Agriculturae Studium of Research, 1(1), 41–55. https://doi.org/10.55849/agriculturae.v1i1.172

Elragal, A., Awad, A. I., Andersson, I., & Nilsson, J. (2024). A Conversational AI Bot for Efficient Learning: A Prototypical Design. *IEEE Access*, *12*, 154877–154887. Scopus. https://doi.org/10.1109/ACCESS.2024.3476953

Eron, F., Noman, M., de Oliveira, R. R., & Chalfun-Junior, A. (2024). Computer Vision-Aided Intelligent Monitoring of Coffee: Towards Sustainable Coffee Production. *Scientia Horticulturae*, *327*. Scopus. https://doi.org/10.1016/j.scienta.2024.112847

Fuentes, J., Aguilar Castro, J., & Montoya, E. (2025). Autonomous cycle of data analysis tasks for the determination of the coffee productive process for MSMEs. *Journal of Industrial Information Integration*, *44*. Scopus. https://doi.org/10.1016/j.jii.2025.100788

Guilin, X., Jiao, D., & Wang, Y. (2024). The Precision Agriculture Revolution in Asia: Optimizing Crop Yields with IoT Technology. Agriculturae Studium of Research, 1(1), 1–14. https://doi.org/10.55849/agriculturae.v1i1.172

Hong, S., Park, S., Youn, H., Lee, J., & Kwon, S. (2024). Implementation of Smart Farm Systems Based on Fog Computing in Artificial Intelligence of Things Environments. *Sensors*, *24*(20). Scopus. https://doi.org/10.3390/s24206689

Jang, M., Bae, G., Kwon, Y. M., Cho, J. H., Lee, D. H., Kang, S., Yim, S., Myung, S., Lim, J., Lee, S. S., Song, W., & An, K.-S. (2024). Artificial Q-Grader: Machine Learning-Enabled Intelligent Olfactory and Gustatory Sensing System. *Advanced Science*, *11*(23). Scopus. https://doi.org/10.1002/advs.202308976

Kim, Y., Lee, J., & Kim, S. (2024). Study of active food processing technology using computer vision and AI in coffee roasting. *Food Science and Biotechnology*, *33*(11), 2543–2550. Scopus. https://doi.org/10.1007/s10068-023-01507-7

Korkmaz, A., Talan, T., Koşunalp, S., & Iliev, T. (2025). Comparison of deep learning models in automatic classification of coffee bean species. *PeerJ Computer Science*, *11*. Scopus. https://doi.org/10.7717/peerj-cs.2759

Messias, E., Guimarães, C. V., da Luz, J. M. R., de Paulo, E. H., da Silva Oliveira, E. C. D. S., Nascimento, M. H. C., Ferrão, M. F., Guarçoni, R. C., Filgueiras, P. R., & Pereira, L. L. (2025). Study of the sensory profile of Coffea canephora through malting/fermentation using HS-SPME-GC-MS and synthetic sampling combined with random forest. *Food Chemistry*, *489*. Scopus. https://doi.org/10.1016/j.foodchem.2025.144907

Motta, I. V. C., Vuillerme, N., Pham, H.-H., & Pereira de Figueiredo, F. A. P. (2025). Machine learning techniques for coffee classification: A comprehensive review of scientific research. *Artificial Intelligence Review*, *58*(1). Scopus. https://doi.org/10.1007/s10462-024-11004-w

Niforatos, E., Ferwerda, B., Pop, M., & Schricker, M. (2024). Integrating Generative AI in the UX Design Process: An Empirical Perspective. *Ekphrasis*, *32*(2), 114–122. Scopus. https://doi.org/10.24193/ekphrasis.32.7

Nunes, P. H., Pierangeli, E. V., de Oliveira Santos, M. O., de Oliveira Silveira, H. R. O., de Matos, C. S. M., Pereira, A. B., Alves, H. M. R., Volpato, M. M. L., Silva, V. A., & Ferreira, D. D. (2023). Predicting coffee water potential from spectral reflectance indices with neural networks. *Smart Agricultural Technology*, *4*. Scopus. https://doi.org/10.1016/j.atech.2023.100213

Ozal, G., Ilyasova, C., & Ilgiz, V. (2024). Post-Harvest Storage and Processing Technology in Russia: Reducing Yield Loss. Agriculturae Studium of Research, 1(1), 28–49. https://doi.org/10.55849/agriculturae.v1i1.172

Paiva, A. C., Teixeira, C. A., & Hantao, L. W. (2024). Exploring accurate mass measurements in pixel-based chemometrics: Advancing coffee classification with GC-HRMS—A proof of concept study. *Journal of Chromatography A*, *1731*. Scopus. https://doi.org/10.1016/j.chroma.2024.465171

Przybył, K., Gawrysiak-Witulska, M., Bielska, P., Rusinek, R., Gancarz, M., Dobrzański, B., & Siger, A. (2023). Application of Machine Learning to Assess the Quality of Food Products—Case Study: Coffee Bean. *Applied Sciences (Switzerland)*, *13*(19). Scopus. https://doi.org/10.3390/app131910786

Qin, Y., Huo, S., González, A. M., Guo, L., Sanots, J., & Li, L. (2025). Research on Neuroimmune Gastrointestinal Diseases Based on Artificial Intelligence: Molecular Dynamics Analysis of Caffeine and DRD3 Protein. *Current Pharmaceutical Biotechnology*. Scopus. https://doi.org/10.2174/0113892010325902241120111429

Ricardo Mondragón Regalado, J., Huaman Montez, A., Marino Jara Llanos, D., Aguirre Baique, N., Reategui del Águila, K., Luz Basilio Maraví, G., Honorio Muñoz Berrocal, M., & Arévalo Reátegui, J. (2025). Agricultural Culture and Technological Civilization: Tarpuy and Artificial Intelligence in Coffee Disease Management. *Revista Iberoamericana de Viticultura Agroindustria y Ruralidad*, *36*, 206–224. Scopus. https://doi.org/10.35588/8ajmqd10

Rogger, T., Jonathan, H., & Lindsey, K. (2024). Smart Fertilization Technology for Agricultural Efficiency in Canada. Agriculturae Studium of Research, 1(1), 56–70. https://doi.org/10.55849/agriculturae.v1i1.172

Sagita, D., Widodo, S., Mardjan, S. S., Purwandoko, P. B., Suparlan, n., Hariadi, H., & Darniadi, S. (2025). Rapid identification of coffee species and origin using affordable multi-channel spectral sensor combined with machine learning. *Food Research International*, *211*. Scopus. https://doi.org/10.1016/j.foodres.2025.116501

Salamai, A. (2024). Towards automated, efficient, and interpretable diagnosis coffee leaf disease: A dual-path visual transformer network. *Expert Systems with Applications*, *255*. Scopus. https://doi.org/10.1016/j.eswa.2024.124490

Salamai, A., & al-Nami, W. T. (2023). Sustainable Coffee Leaf Diagnosis: A Deep Knowledgeable Meta-Learning Approach. *Sustainability (Switzerland)*, *15*(24). Scopus. https://doi.org/10.3390/su152416791

Sander, J., Simon, P., & Hinske, C. (2024). Big data and artificial intelligence in anesthesia: Reality or fiction? *Anaesthesiologie*, *73*(2), 77–84. Scopus. https://doi.org/10.1007/s00101-023-01362-5

Shin, B. (2025). Exploring the potential of machine learning to reduce administrative burden in participatory budgeting: A case study of Seoul. *Journal of Public Budgeting, Accounting and Financial Management*, 1–28. Scopus. https://doi.org/10.1108/JPBAFM-09-2024-0188

Silva, L. C. F., Pereira, P. V. R., Cruz, M. A. D. D., Costa, G. X. R., Rocha, R. A. R., Luiz Lima Bertarini, P. L. L., Amaral, L. R. D., Gomes, M. S., & Santos, L. D. (2024). Enhancing Sensory Quality of Coffee: The Impact of Fermentation Techniques on Coffea arabica cv. Catiguá MG2. *Foods*, *13*(5). Scopus. https://doi.org/10.3390/foods13050653

Solis Pino, A. F. S., & Apraez, L. S. C. (2025). Artificial intelligence and multispectral imaging in coffee production: A systematic literature review. *European Journal of Agronomy*, *170*. Scopus. https://doi.org/10.1016/j.eja.2025.127725

Vandeputte, J., Herold, P., Kuslii, M., Viappiani, P., Muller, L., Martin, C., Davidenko, O., Delaere, F., Manfredotti, C., Cornuéjols, A., & Darcel, N. (2023). Principles and Validations of an Artificial Intelligence-Based Recommender System Suggesting

Acceptable Food Changes. *Journal of Nutrition*, *153*(2), 598–604. Scopus. https://doi.org/10.1016/j.tjnut.2022.12.022

Vilela, E. F., Castro, G. D. M. D., Marin, D. B., Santana, C. C., Leite, D. H., Matos, C. D. S. M., Silva, C. A. D., Lopes, I. P. D. C., Queiroz, D. M. D., Silva, R. A., Rossi, G., Bambi, G., Conti, L., & Venzon, M. (2024). Remote Monitoring of Coffee Leaf Miner Infestation Using Machine Learning. *AgriEngineering*, *6*(2), 1697–1711. Scopus. https://doi.org/10.3390/agriengineering6020098