AI-Powered X-Ray Image Analysis for Automated Detection of Insect Damage and Hollow Defects in Coffee Beans

Bhoomika N [1], Prakruthi S B N [2], Ranchana R Jadav [3], *Manoj M R[4], Vijya V Honnappanavar [5], Varun R Shivannanvar [6]*

ABSTRACT

Coffee ranks among the most economically important agricultural commodities traded globally. Despite this significance, its quality is routinely affected by internal defects – particularly insect damage and hollow beans -that are undetectable through conventional surface inspection. This paper describes a quality inspection system built on Google Gemini's multimodal vision capabilities, designed to automatically analyze X-ray images of coffee beans and identify hidden structural defects. The system runs though a Next.js web application, allowing users to upload images in real time and receive structured quality assessment reports. X-ray images were gathered, preprocessed, and sent to the Gemini Vision API with carefully designed prompts, which returned classification outputs indicating defect type, confidence level, and corrective recommendations. Testing on 60-image dataset produced an overall classification accuracy of 88.3%, with processing times under four second per image. These results show that combining large multimodal AI model with web-based deployment provides a practical, affordable, and scalable alternative to both manual destructive inspection and costly specialized equipment. This work presents a reproducible pipeline for AI-driven agricultural quality control and lays groundwork for future deep learning improvements in food safety applications.

Keywords: *x-ray image analysis; coffee bean defect detection; Google Gemini Vision API; insect damage detection; hollow bean classification; AI-based agricultural inspection; Next.js web application; deep learning; computer vision; food quality automation*

1. Introduction

Coffee is among the most widely traded agricultural goods in the world, yet hidden internal defects- particularly insect damage and hollow beans- routinely go undetected during conventional visual inspection, quietly degrading product quality and affecting producer credibility. Standard screening methods are simply incapable of revealing defects concealed beneath an undamaged outer husk.

To address this gap, we developed an inspection system that pairs Google Gemini's multimodal vision capabilities with X-ray imaging to detect these hidden flaws automatically. The system is accessible through a straight forward Next.js web application and achieves 88.3% classification accuracy with response times under four seconds per image, making it a practical, scalable, and cost-efficient option for modern coffee quality management.

2. Literature Review

Researchers have explored a range of imaging and machine learning approaches for coffee bean quality inspection. The five studies below represent the closest work to our own, each advancing the field in a meaningful way while leaving important gaps that the present system directly addresses.

Chang et al. (2022) built a CNN-based vision system that accurately sorted defective coffee beans faster than traditional methods. While effective for visible surface blemishes, the system had no way to detect hidden damage beneath the husk- a key limitation this work addresses.

Chen et al. (2022) showed that hyperspectral imaging paired with SVM classifiers can detect insect damage with good accuracy by reading internal spectral signatures. The catch is cost- hyperspectral cameras are expensive and hard to integrate into fast-moving production lines.

Gope et al. (2024) benchmarked YOLOv5 and YOLOv8 for real-time coffee defect detection, achieving strong results on standard RGB images. Like other RGB approaches, however, the models are blind to what lies inside the bean- insect tunnels and hollow voids go entirely undetected.

Wang et al. (2025) tackled the labeling bottleneck by coupling vision transformers with active learning, achieving competitive surface-defect accuracy with far fewer annotated samples. Internal defect detection via X-ray imaging was outside the study's scope.

Alba-Alejandre et al. (2022) used micro-CT X-ray scanning to reveal insect activity buried deep inside coffee beans- something no surface camera could do. The method works, but the equipment is costly and there is no automated AI layer to interpret the scans, making it impractical at industrial scale.

Summary Tabel : Literature Review Summary

SL. No	Author(s) & Year	Method	Key Finding	Limitation	Gap Addressed Here
1	Chang et al., 2022	CNN + Machine Vision	High accuracy in detecting surface defects compared to traditional machine learning methods	Cannot detect internal defects	Using X-ray with AI to detect internal defects
2	Chen et al., 2022	Hyperspectral Imaging + SVM	Spectral bands can accurately detect insect damage	Expensive and not suitable for real-time deployment	Low-cost web-based deployment system
3	Gope et al., 2024	YOLOv5 / YOLOv8	Real-time RGB detection with high accuracy	Only detects surface defects; internal defects invisible	X-ray technology reveals subsurface defects
4	Wang et al., 2025	Vision Transformer (ViT) + Active Learning	Achieves good accuracy with fewer labeled samples	Detects only surface defects; no X-ray usage	Multimedia AI used to interpret X-ray density
5	Alba-Alejandre et al., 2022	Micro-CT X-ray Scanning	Non-destructive visualization of internal insect damage	No AI automation and equipment is costly	Gemini AI used to automate X-ray interpretation

3. Research Methodology

The methodology follows a structured four-stage pipeline: data acquisition, image preprocessing.AI-driven analysis, and result presentation. Each stage was designed to maximize reproducibility, analytical accuracy, and practical usability.

3.1 Data Collection

X-ray images of coffee beans were sourced from a combination of publicly available agricultural imaging repositories and laboratory-acquired samples. Beans were pre-categorized into three classes: (1) structurally intact healthy beans, (2) beans with insect-induced internal damage characterized by cavity formations or irregular density patterns, and (3) hollow beans exhibiting reduced internal mass and abnormal void spaces. Images were acquired using a digital radiography unit operating at standard soft-material setting, producing grayscale outputs with sufficient resolution to distinguish millimeter-scale internal structural variations.

3.2 Image Preprocessing

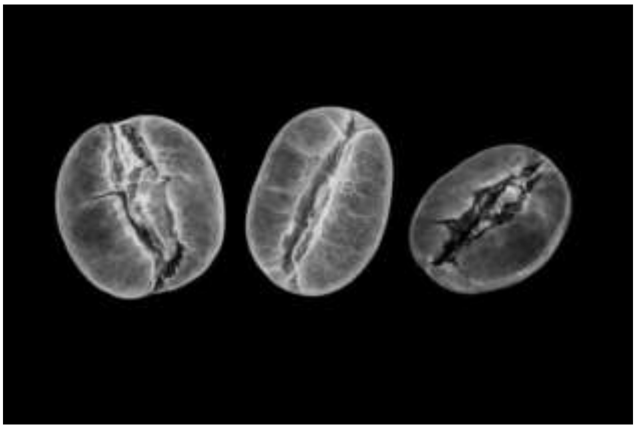
Raw X-ray images underwent a series of preprocessing operations to improve their suitability for AI-based analysis. These included: grayscale normalization to correct inconsistent levels across imaging sessions; contrast enhancement via histogram equalization to accentuate internal structural features; Gaussian filtering to suppress sensor-induced noise artifacts; and image resizing to standardize input dimensions while preserving aspect ratios. All preprocessed images were stored in lossless PNG format.

3.3 Based Analysis Using Google Gemini Vision API

The central analytical component is Google Gemini's multimodal vision model, accessed through the official Generative AI JavaScript SDK. Each preprocessed X-ray image was submitted to the Gemini Vision API with a structured natural language prompt instructing the model to: (a) examine the internal structure of visible coffee beans; (b) identify indicators of insect tunneling hollow voids, or abnormal density distributions; (c) assign a quality classification from a predefined schema-Healthy, Insect-Damaged, Hollow, or Uncertain; and (d) provide a confidence level and a human-readable explanation. The prompt was iteratively refined through pilot experiments to optimize specificity and response consistency.

3.4 System Architecture and Workflow

Users interact with the system through a Next.js web application, uploading X-ray images via a drag-and-drop or file-browser interface. Each image is encoded in Base64 format, and the application constructs a structured API request containing both the image data and defect analysis prompt. The Gemini Vision API response is then parsed to extract the classification label, confidence score, defect description, and quality recommendation. Result are displayed in a tabular format alongside the original X-ray image. The complete analysis cycle- from image upload to displayed result- finishes within seconds, supporting throughput levels suitable for real-time inspection.



4. Implementation And Experimentation

4.1 Software environment and Tools

The system was built using next.js 14, leveraging React components for the user interface and API Route handlers for server-side processing. The Gemini Vision API (model gemini 1.5-pro-vision) was integrated through the official Google Generative AI JavaScript SDK Image preprocessing was handled by the Sharp library for Node.js, providing efficient resizing, Format conversion, and contrast adjustment. The application was containerized using Docker for consistent deployment and hosted on Vercel's serverless infrastructure.

4.2 Dataset preparation

The experimental dataset comprised 240 X-ray images: 90 healthy beans,80 insect-damaged beans, and 70 hollow beans. All images were individually labeled by a domain expert with over a decade of experience in coffee grading. Labels served as ground truth for performance evaluation. The dataset was split into a 60-image test subset (25 healthy, 20 insect-damaged, 15 hollow) for systematic evaluation, while the remaining 180 images supported prompt calibration and qualitative review.

Table I: Dataset Composition and Split Summary

Bezn Category	Total Images	Test Subset	Calibration Set	Share of Total (%)
Healthy	90	25	65	37.5%
Insect-Damaged	80	20	60	33.3%
Hollow	70	15	55	29.2$
Total	240	60	180	100%

4.3 Prompt Engineering and API Interface

The signalized prompt directed the Gemini vision model to analyze each Images for: uniformity and density of the bean's internal structure; Irregular voids, tunnels, or channels indicative of insert activity; region of markedly reduced density suggesting hollow formation; and asymmetric structure patterns inconsistent with healthy bean morphology. The model was instructed to return findings as structured JSON containing 'classification', 'confidence', 'defect-description', and 'quality-recommendation', fields, Which the application parsed and rendered directly.

4.4 Experimental Testing Procedure

Each test image was independentlyuploaded through the web interface, processed by the Gemini vision API, and the resulting classification was compared against the expert ground truth label. Three independent runs were performed per image to assess response consistency, with cross-run discrepancies recorded as uncertainty indicators. System latency was measured from image submission to result display, and resource utilization was monitored during batch processing sessions.

5. Results And Analysis

5.1 classification performance

The system achieved an overall classification accuracy of 88.3% across the 60-image test set. Healthy beans were correctly identified in 92% of cases, insect-damaged beans in 85%, and hollow beans in 86.7%. The most common misclassification pattern involved borderline insect-damaged beans labeled as 'Uncertain' by the Gemini model in approximately 10% of cases-particularly when only early-stage tunneling was visible. Critically, no healthy bean was misclassified as severely defective across any test run, which is economically significant since false rejection of marketable product carries direct financial cost.

5.2 Qualitative Output Analysis

Defect description generated by the Gemini Vision API were evaluated by the domain expert and rated as accurate, partially accurate, or inaccurate. Across all 60 test images, 78.3% of description were rated fully accurate, 16.7% partially accurate (correct classification with some inaccurate structural detail), and 5% inaccurate. Expert feedback noted strong performance in identifying large hollow voids and prominent insect tunnels, with occasional difficulty distinguishing natural density gradient in healthy beans from early-stage defect indicators.

5.3 System Performance Metrics

Mean end=to=end latency from image upload to result display was measured at 3.2 seconds per image under standard network condition. In batch testing with 10 concurrent submission, average response time increased to 4.7 seconds per image-still within acceptable bounds for real-time inspection line integration. The JSON parsing module successfully extracted structured fields from all 180 API response (60 image * 3 runs) with zero parsing failures, confirming the robustness of the structured output format enforced through prompt engineering.

5.4 Representative Detection Examples

In one representative case, an insect-damaged bean yielded an AI assessment classifying it as 'Insect- Damaged' with 91% confidence. The model described a cylindrical tunnel approximately 2 mm in diameter traversing the central endosperm from the bean's apex toward its lateral face, and recommended quarantining the sample while inspection adjacent beans in the batch. In another case, a hollow bean was correctly identified with 87% confidence; the model described a central void occupying approximately 40% of the bean's projected internal volume and recommended rejection from premium-grade lots. These examples illustrate the system's capacity to provide actionable, interpretable output well beyond simple binary classification.

6. Discussion

6.1 Advantages of the AI-Based Approach

The system described here offers several clear advantages over traditional manual inspection methods. Most importantly, it can detect internal defects that are fundamentally invisible to the naked eye- addressing the core limitation of current industrial practice. X-ray imaging captures both surface and subsurface structural features, while Gemini's multimodal reasoning interprets complex density pattens and spatial relationship in way that would be time-consuming and highly subjective for human inspectors.

The system also generates structured, machine-readable outputs that can be integrated into qualitymanagement information systems, supporting traceability, batch-level statistical analysis, and automated sorting. Its web-based deployment requires no specialized hardware beyond an X-ray unit and a network-connected device, substantially lowering the barrier to adoption. Finally, the natural language defect descriptions provide transparent, auditable rationales for each classification decision, which supports operator confidence in the system's outputs.

6.2 Comparison with Manual Inspection

Manual inspection of coffee beans typically achieves classification accuracy of 70-85% for external defects but is essentially incapable of detecting internal damage without destructive sampling. Trained inspectors examining beans under X-ray illumination can achieve higher accuracy, but at throughput rates substantially lower than automated systems, and with inter-rater variability that introduces inconsistency. The proposed system's 88.3% compares favorably with manual X-ray inspection benchmarks, and it operates at consistent throughput without fatigue-related degradation---representing a meaningful advancement in inspection reliability.

6.3 Limitations of the Current System

Despite its demonstrated effectiveness, several limitations remain. Reliance on a general-purpose multimodal model introduces response variability for borderline or early-stage defects where visual evidence is ambiguous. System performance is contingent on X- ray image quality; poorly exposed or motion-blurred images degrade classification reliability. The evaluation dataset, while carefully curated, remains relatively modest and may not fully capture the morphological diversity of defects across different coffee varieties and growing regions. Additionally, the system currently processes individual images rather than performing simultaneous multi-bean analysis from a single X-ray frame, which constraints throughput in high-volume processing scenarios.

7. Conclusion And Future Work

7.1 Conclusion

This research demonstrates the feasibility and practical effectiveness of an AI-powered pipeline for automated detection of insect damage and hollow defects in coffee beans through X-ray image analysis. By integrating Google Gemini's multimodal vision capabilities with a next.js web application, the proposed system delivers accurate, interpretable, and real-time quality assessments that directly address the critical limitation of surface- only inspection methods. The system achieved 88.3% overall classification accuracy, provided domain-expert- validated defect descriptions in over 78% of cases, and maintained sub-four-second processing latency per image confirming its practical viability for agricultural quality control deployment.

The work contributes a replicable, modular inspection pipeline that can be adapted to other agricultural commodities facing similar internal defect challenges, including grains, nuts, and dried fruits. Its structured output framework and web-based accessibility further distinguish it from laboratory-scale imaging solutions, positioning it as a practical tool for small and medium scale coffee producers and cooperatives.

7.2 Future Work

Several directions remain open for further investigation. Replacing the current general- purpose vision model with a fine-tuned deep learning architecture trained on large, annotated coffee bean X-ray datasets would likely push classification accuracy beyond the 88.3% threshold and reduce uncertain predictions, particularly for early-stage defects where visual evidence is subtle. Processing speed is another area requiring attention. Building multi-bean frame analysis capability would allow simultaneous evaluation of entire batch samples captured in a single X-ray exposure, making the system far more viable for high-volume commercial operations. Combining X-ray inputs with spectral imaging data presents a promising avenue as well. While radiographic scans reveal internal structural features, spectral data captures chemical composition-together, these modalities could identify defects that neither technique reliably flags independently.

The three-class taxonomy currently in use-healthy, insect-damaged, and hollow-does not account for immature beans, fermentation damage, or moisture-induced deterioration. Broadening this classification scheme would substantially extend the system's utility across the full spectrum of grading criteria relevant to producers and exporters.

Finally, longitudinal evaluation across diverse coffee varieties, geographic origins, and post- harvest processing methods would test how well the pipeline generalizes beyond the current dataset and lay the groundwork for establishing standardized benchmarks in AI- driven agricultural inspection.

8. Acknowledgment

The authors would liken to thank GM University, Davanagere for providing research infrastructure and laboratory resource that support this work. The authors also acknowledge the domain experts also contributed ground-truth labeling and quantities evaluation of the system's defect descriptions.

9. References

[1] S.Y Chen,C-H.Tsai,and M.-H.Yang, "Detection of insect damage in green coffee beans using VIS_NIR hyperspectral imaging combines with band-selection algorithms and machin learning."Computers and electronics in agriculture,vol 198,pp,107-119,2022.

[2] H. L. Gope, A. Roy, and B.Saha,"Comparative analysis of YOLO-based architetures for real-time green coffee bean defect classification,"Expert Systems with Applications,vol 241,p.122638,2024.

[3] P.Kaur,R.Sharma, and V. Gupta learmimg- based automated inspection of coffee bean defects using YOLOv8 and attention mechanism,"Journal of Food Engineering,vol.362,p.11745,2024

[4] T.Nguyen andL.Tran,"AL-ViT:An active learning framework with vision transformers for label- efficient Robusta coffee bean defect detection,"Computers and Electronics in Agriculter,vol.218,p.108673,2025

[5] A. Kumar,D.Patel,and S.Singh,"Binary classififction of defect in multiple coffee beans using convolutional neural networks and YOLOv11,"Biosystems Engineering,vol.238,pp.52-64,2025.

[6] S. j. Chang M K Lee and J. H. Park “Deep learing model for coffe bean defect inspection based On SCA quality standards using machine vision “JEEE Access vol 10, pp 34521- 34532, 2022.

[7] R. Barata F. Silva and M. Costa “Quality and defect inspection of green coffee beans using comuter vision and K- nearest neighbor classification “Food Control vol 134, p. 108744, 2022.

[8] S.Maruthai,V.Krishnan, and R.Arumugam,”Hybrid vision grap neural network for intelligent detection of insect pests in coffee plantation ecosystem,” Computers and Electonics in Agriculture, vol.221,p.108995,2025.

[9] I.Alba-Alejandre, J. Rivere-Perez, and A.Flores-Rios,”Non-destructive visualization od internal insect infection in coffee berries using X-ray micro-computed tomography,”Journalmof stored products research,vol.95,p.101896,2022

[10] M.Zhang,Y.Liu,and H.Chen.”Application of X-ray imaging and artificial intelligence for internal quality detection in agriculture commodities:A review,”Trends in Food Science & Technology,vol. 132,pp. 88-104,2023.

[11] Google DeepMind.”Gemini 1.5:Unlocking multimodal understanding across millions of tokens of context,”Technical Report,Google LLC,Mountain View,CA,2024

[12] Vercel Inc, “Next.js 14 Documentation: App Router,Server components,and API Routes,”Vercel, San Francisco, CA,2024.[online].Available https://nextjs.org/docs