Green Coffee Bean Quality Inspection Using Computer Vision Systems

Jameel John M. Reales
Department of Electronics Engineering, Caraga State University

Ronieto N. Mendoza
Department of Electronics Engineering, Caraga State University

Von Romer D. Labas
Department of Electronics Engineering, Caraga State University

https://doi.org/10.5109/7323390

出版情報：Proceedings of International Exchange and Innovation Conference on Engineering & Sciences (IEICES). 10, pp.1065-1070, 2024-10-17. International Exchange and Innovation Conference on Engineering & Sciences
バージョン：

Green Coffee Bean Quality Inspection Using Computer Vision Systems

Jameel John M. Reales[1], Ronieto N. Mendoza[1, 2], Von Romer D. Labas[1]
[1]Department of Electronics Engineering, Caraga State University
[2]Center for Renewable Energy, Automation and Fabrication Technologies (CSU CRAFT)
Corresponding Author: jameeljohn.reales@carsu.edu.ph

Abstract:
In the 19th century, the Philippines led global coffee exports until the 1890s, boasting ideal conditions for growing Arabica, Robusta, Excelsa, and Liberica beans. However, traditional methods of quality inspection were slow and error-prone, impacting industry reputation and profitability. This study introduces a method for green coffee bean grading and quality inspection using computer vision technology. Specifically, the study presents a software designed to improve the accuracy and efficiency of bean quality assessment. The software was trained on over 9000 coffee beans, with 65% of the data allocated for training, 25% for validation, and 15% for testing. The YOLOv8n Final Model achieved over 90% accuracy within 75 epochs. ACER NITRO 5 and Lenovo IdeaPad Gaming laptops, equipped with Nvidia graphics, delivered stable visual performance at 28-33 frames per second (FPS). The 48MP microscope camera provided clear and reliable images during sampling and testing, ensuring precise and dependable quality evaluations.

Keywords: Green coffee bean, Computer Vision, YoloV8, CNN, Robusta.

1. INTRODUCTION

Coffee has played a considerably more vital role in the globe for generations than most people think. Coffee is also a valuable commodity and an important component of many countries' economy; in fact, being the principal agricultural export in dozens of locations across the Equatorial Belt, coffee employs millions of people involved in its cultivation, processing, and distribution. In the 19th century, the Philippines was the world's largest coffee exporter through to the 1890s. The Philippines has a lot of potential capable of producing the four major varieties of beans, Arabica, Robusta, Excelsa and Liberica [1]. Many factors influence the flavor of coffee, with the quality of coffee beans being the most important. Thanks to the kind of climate and soil conditions that we have, coffee growing in the Philippines, from the lowland to mountainous regions, became suitable. Processing coffee is the act of removing the layers of skin, pulp, mucilage, and parchment that surround a coffee bean [2]. In this study it focuses on using computer vision. Computer vision is a technique used to automate the process of sorting and classifying objects based on their visual properties. It involves the use of cameras, sensors, and algorithms to analyze images of objects or any other distinguishing features. Detection using computer vision involves capturing images of objects, preprocessing the images, extracting features using computer vision techniques, and applying the trained model to new images to identify objects and assign them to the appropriate category. It is a rapidly growing field that has the potential to transform many industries and improve efficiency and productivity. Moreover, the integration of computer vision with machine learning models allows for continuous improvement in defect detection, adapting to new patterns and variations in coffee bean quality over time.

2. METHODOLOGY

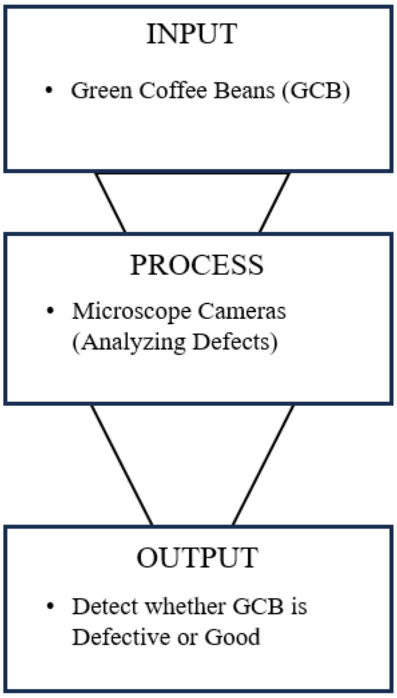

Figure 1. General Framework

In the green coffee bean quality inspection system, as depicted in Figure 1, the Input-Process-Output (IPO) framework guided the efficient inspection of green coffee beans. The overall process involved using computer vision to detect and classify defective green coffee beans. In the input phase, the system received fresh green coffee beans. The process stage defined how the computer vision system, utilizing a high-resolution camera, effectively analyzed each bean for defects. This process involved capturing images of the green coffee beans. The camera was programmed to detect defects such as insect bites, shape irregularities, or color variations. When defects were identified, the system identified the respective beans as defective. On the other hand, beans without detected defects were classified as good-quality.

After the analysis, the output provided a clear categorization of the green coffee beans, distinguishing between good and defective beans, thus ensuring a reliable quality inspection process.

2.1 System Overview

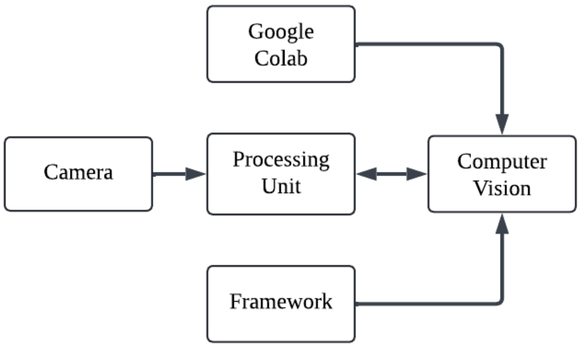

Figure 2. Block Diagram of the System

In the presented block diagram, a central processing unit acts as the main controller for the system, managing the entire process of green coffee bean quality inspection. The camera captures images of the green coffee beans and sends them to the processing unit. This processing unit is integrated with a computer vision system, which analyzes the images to detect any defects in the beans. The framework, which includes different machine learning tools, operates within the computer vision system to ensure accurate and efficient defect detection. Google Colab is used for training and developing these machine learning models, providing the necessary computational power and resources. The processed data from the computer vision system is then used to categorize the beans as either good or defective. This setup optimizes the inspection process, making it faster and more reliable compared to manual quality inspection methods.

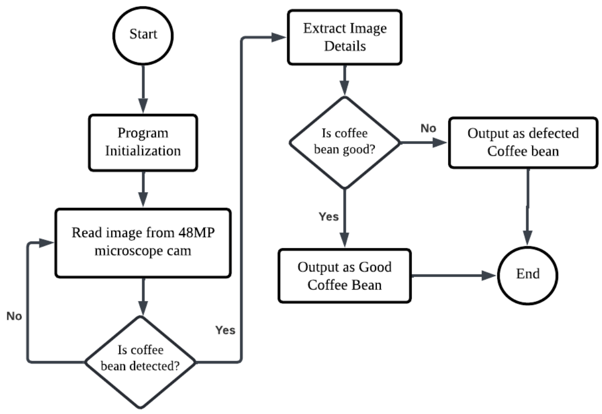

Figure 3. Flowchart of the system

This flowchart illustrates the process of green coffee bean quality inspection using a computer vision system. The process begins with program initialization, followed by reading images from a 48MP microscope camera. If a coffee bean is detected in the image, the system extracts image details for analysis. The bean is then evaluated for quality: if it meets the quality criteria, it is classified as a good coffee bean; otherwise, it is classified as a defective coffee bean. Defective beans are further classified into specific categories such as cracked, destroyed, discolored, or insect-bitten. The process loops back to reading new images until no more coffee beans are detected, at which point the system ends the inspection.

2.2 PyTorch and YoloV8 Implementation

In terms of improving the accuracy and speed of the overall system, the researchers have transitioned from the utilization of TensorFlow Lite to PascalVOC to the implementation of PyTorch and YoloV8, recognizing its enhanced capabilities in its overall performance. YoloV8, which came from the YOLO (You Only Look Once) series of object detection algorithms, marks a significant advancement in computer vision. It builds upon the strengths of earlier versions while incorporating cutting-edge techniques and optimizations, resulting in notable enhancements in accuracy and efficiency. YoloV8's superior performance has been evidenced across various datasets and practical applications, establishing it as the preferred choice for object detection tasks. Its ability to quickly and accurately identify objects within images makes it particularly well-suited for real-time applications, which is a critical factor in the context of green coffee bean quality inspection. Additionally, the integration of PyTorch provides a more flexible and efficient framework for model training and deployment, facilitating rapid development and iteration. Considering PascalVOC's outdated status in contemporary research and the compelling performance advantages offered by YoloV8, its adoption within this study is essential to ensure the use of the most effective and current methodologies in achieving the research objectives. Furthermore, this transition aligns with the broader trends in the field, where modern techniques and tools are increasingly being leveraged to push the boundaries of what is possible in machine learning and computer vision. The combination of PyTorch and YoloV8 thus represents a strategic enhancement to the system, allowing significant improvements in both the speed and accuracy of defect detection in green coffee beans.

2.3 AI Training Image Processing and Data Sampling

The efficiency of the automated green coffee bean quality inspection system relies on the quality and diversity of the training data. High-quality, diverse data ensures the system can accurately recognize both defective and non-defective beans.

2.3.1 Image Processing Systems

The study uses a Convolutional Neural Network as the Image Processing Algorithm of the model, which is crucial for detecting and classifying the quality of green coffee beans.

2.3.2 Data Sampling

Prior to the AI Training, the study had a total of 9000+ samples to train the Image Processing System of the device. This dataset was divided into 65% for training, 25% for validation, and 15% for testing. This distribution ensured an adequate amount of data for each class of coffee beans pre-determined for detection by the system. This comprehensive dataset also included various types of defects to enhance the system's ability to generalize and accurately classify beans in real-world conditions.

2.4 Accuracy Test

The accuracy test were carried out with 5 tests to determine the accuracy of the model. The table below shows the number of Good and Defective GCBs that were used in each test. The accuracy rate of the Green Coffee Bean Quality Inspection Using Computer Vision Systems were determined by the Equation 1.

Table 1. Accuracy Test

Test 1	120 Good GCB
Test 2	120 Defective GCB
Test 3	100 Good and 20 Defective GCBs
Test 4	20 Good and 100 Defective GCBs
Test 5	60 Good and 60 Defective GCBs

$$\text{Accuracy } (\%) = \frac{\text{Number of Correctly Identified beans}}{\text{Total Number of beans input}} \times 100 \quad \text{(eq. 1)}$$

3. DATA RESULTS AND ANALYSIS

This chapter presents the results of the test conducted by the researchers with its discussion in order to achieve their desired objectives.

3.1 Performance Evaluation using different processors

In this subsection, the researchers utilized two Raspberry Pi 4's with different RAMs (4GB and 8GB) and three different laptops to evaluate their performances which are the, Acer Nitro 5, Lenovo IdeaPad Gaming, and Acer TravelMate to determine which processors offered the best frames per second (FPS) and accuracy during testing.

3.1.1 Raspberry Pi 4 (4GB) vs Raspberry Pi 4(8GB)

Based on the results displayed in Table 1, the shift from Tensorflow lite to Pytorch and YoloV8 aimed to assess changes in performance and accuracy. Specifically, on the Raspberry Pi 4 platform, an improvement in accuracy was observed alongside a reduction in frames per second (fps). The researchers concluded that while the adoption of PyTorch and YOLOv8 led to enhanced accuracy, the limited computational capacity of the Raspberry Pi 4 resulted in decreased fps. Thus, it was decided that to fully capitalize on the heightened accuracy provided by PyTorch and YOLOv8, a more powerful computational device capable of achieving higher fps is necessary.

Table 2. Rpi 4 (4GB) vs Rpi 4 (8GB)

Processor	Camera	Framework	FPS	Average Accuracy
Raspberry Pi 4 (4GB RAM)	8MP Rpi CamV2	TensorFlow & Pascalvoc	7-8	0%
	5MP OKDO Cam			60-70%
	12MP HQ Cam			<50%
	5MP OKDO Cam	Pytorch & YOLOv8	0-1	80%
	12MP HQ Cam			<50%
Raspberry Pi 4 (8GB RAM)	1080p Webcm	Tensorflow & Pascalvoc	3-5	0%
	5MP OKDO Cam		6-8	60-70%
	5MP OKDO Cam	Pytorch & YOLOv8	0-1	80%

3.1.2 CUDA Integration and Performance Evaluation of Acer Nitro 5, Lenovo Ideapad Gaming, and Acer TravelMate

Table 3. Processor Performance in terms of FPS

Processor	Camera	CUDA Enabled	Fps	Framework
Acer Nitro 5	13MP CAM	NO	3-4	
	48MP CAM	YES	28-33	Pytorch & YOLOv8
Acer Travelmate	13MP CAM	NO	4-6	
	48MP CAM	NO	4-6	
Lenovo Ideapad Gaming	13MP CAM	YES	30-34	Pytorch & YOLOv8
	48M CAM	YES	30-34	

Table 4. Processor Performance in terms of Accuracy

Processor	Camera	CUDA Enable	Accuracy	FW
Acer Nitro 5	13MP CAM	NO	>90%	Pytorch & YOLOv8
	48MP CAM	YES	>90%	
Acer Travelmate	13MP CAM	NO	70%	
	48MP CAM	NO	70%	
Lenovo Ideapad Gaming	13MP CAM	YES	>90%	Pytorch & YOLOv8
	48M CAM	YES	>90%	

Acknowledging the capability of NVIDIA graphics cards to enhance computational performance through CUDA (Compute Unified Device Architecture), the researchers pursued this approach to increase the fps. Following the installation of CUDA on the Acer Nitro 5 laptop, utilizing its NVIDIA graphics card, significant enhancements were observed. Specifically, fps rates increased to 28-33 fps upon employing a tailored YOLOv8 model optimized for performance. However, the Acer Travelmate 11th Gen i3 Laptop, lacking an NVIDIA GPU, is not equipped to benefit from this enhancement. Tables 2 and 3 shows the overall fps and accuracy of the different processors, the researchers determined that without CUDA, the FPS reached only 3-5 FPS. However, with CUDA enabled, the FPS increased to 28-34 FPS. The researchers concluded that the Acer Nitro 5 and Lenovo IdeaPad Gaming were the most suitable to achieve the objectives of the study due to their compatibility with CUDA and superior performance. Additionally, both laptops were tested using 13 MP and 48 MP cameras on each YOLOv8 trained model. The results showed higher accuracies in object detection when using the Acer Nitro 5 and Lenovo IdeaPad Gaming with the best-trained model in YOLOv8. This demonstrates the significant impact that appropriate hardware selection can have on the effectiveness of machine learning models in practical applications.

3.2 13MP and 48MP Microscope Camera Evaluation

In this study, two digital microscope cameras were evaluated to optimize the use of PyTorch and YOLOv8 for green coffee bean quality inspection using computer vision systems.

3.2.1 13MP Microscope Camera with 130x Lens

Initially, the researchers tested a 13MP microscope camera with a single-faced view to evaluate its suitability for the object detection system. Despite its ability to capture detailed images of individual coffee beans, the close up nature of the camera proved impractical for the intended application. The narrow field of view made it challenging to detect multiple beans simultaneously, thereby limiting its overall usability.

3.2.2 Utilization of 48MP Camera with 130x Lens

In response to the limitations encountered with the 13MP microscope camera, the researchers opted to test a 48MP camera equipped with a 130x lens. This setup provided a wider field of view, allowing for a clearer imaging of multiple coffee beans in a single-faced view. By adjusting the camera's distance from the subject, the researchers were able to optimize the imaging parameters, resulting in improved clarity and usability for the object detection system.

Figure 4. 48MP Cam

3.3 Data Training and Model Optimization

The iterative process of data training and model optimization was conducted utilizing a combination of Roboflow and Google Colab. These platforms provided the necessary infrastructure and tools to train the AI model efficiently, enabling it to learn and adapt to varying input conditions. Through rigorous experimentation, multiple models were trained within the YOLOv8 framework, with the selection of the most accurate model tailored to the specific requirements of the final setup.

3.3.1 YoloV8n Model Training Data Summary

Table 5. Version 1 to 8 Data Summary

Training Verson	Epoch	Box_loss	Cls loss	Dfl loss
1	25	0.1736	0.3035	0.7891
2	25	0.1688	0.2909	0.7897
3	25	0.1811	0.2585	0.794
4	25	0.211	0.3511	0.7855
5	100	0.1788	0.1407	0.7767
6	70	0.181	0.1512	0.7784
7	75	0.1845	0.149	0.7769
8	75	0.1892	0.1579	0.7763

Among the various trainings conducted, Training version 5 stands out as the best training results with 0.1788 Boxloss, 0.1407 Cls loss, and 0.7767 Dfl loss. Epoch 75 was observed to be the most optimal total number of iterations of all the training data.

3.4 Final YoloV8n Model Evaluation

In this section, the researchers present the comprehensive evaluation of the Training Version 8 as the final YOLOv8n model used for green coffee bean quality inspection. The model's performance was assessed based on several key metrics, including accuracy, precision, recall, and F1 score. Among all the other training versions, Training version 8 exhibits the highest robustness and potent performance in the identification of the quality of each green coffee bean that were tested. This is due to the following factors: the highest variability and uniqueness of each coffee bean in the different classes sampled, the varying lighting conditions during data collection, and the largest number of data sets compared to the others.

3.4.1 Accuracy Percentage

Table 6. Final Model Accuracy Test

Final Model Test	Accuracy (%)
120 Good GCB	**90.83%**
120 Defective GCB	**97.5%**
100 Good and 20 Defective GCB	**96.67%**
20 Good and 100 Defective GCB	**95%**
60 Good and 60 Defective GCB	**92.5%**

From the 48MP digital camera with 130x lens with single-faced view, the results of these tests demonstrated the model's high accuracy levels for each faces of the coffee beans captured in the camera in detection for any presence of defects in each coffee bean. In Test 1, which included 120 good GCBs, the final model achieved an accuracy of 90.83%, with 11 incorrect classifications. Test 2, involving 120 defective GCBs, showed a higher accuracy of 97.5%, with only 3 misclassifications. In Test 3, with a mix of 100 good and 20 defective GCBs, the accuracy was 96.67%, with 4 errors. Test 4, comprising 20 good and 100 defective GCBs, resulted in an accuracy of 95%, with 6 incorrect classifications. Finally, Test 5, which had an equal mix of 60 good and 60 defective GCBs, yielded an accuracy of 92.5%, with 9 misclassifications. These results indicate that the model is highly effective in distinguishing between good and defective green coffee beans across various test conditions. The accuracy test results were based from the coffee beans classified with their respected category only on the specific face captured by the camera and the green coffee bean classification that the model has recognized for each coffee bean.

3.4.2 Confusion Matrix

In this study, the confusion matrix was utilized to evaluate the final model's performance in classifying coffee beans into their respective defect categories. The matrix includes counts of true positives, true negatives, false positives, and false negatives, which are essential for calculating other performance metrics such as accuracy, precision, recall, and F1 score. By analyzing the confusion matrix, the researchers gained insights into the model's strengths and weaknesses in distinguishing between different types of defects. This detailed evaluation helps in fine-tuning the model for improved accuracy and reliability in real-world applications.

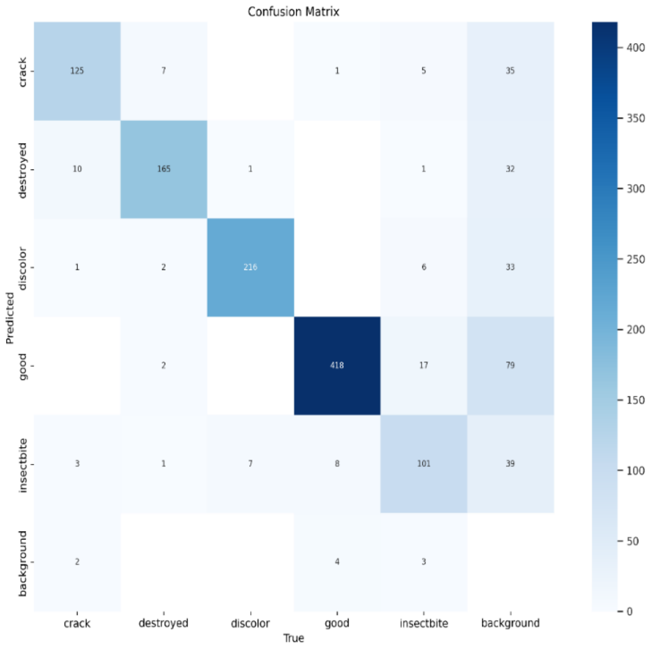

Figure 5. Confusion Matrix

3.4.3 Training and Validation Graph

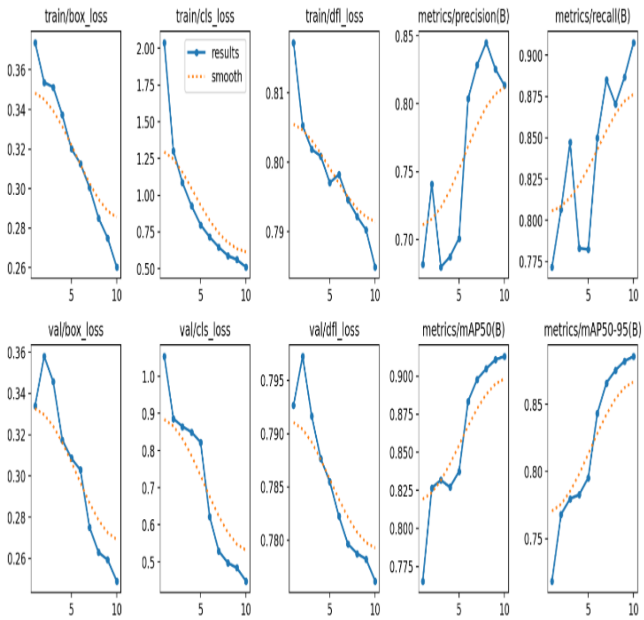

Figure 6. Loss Graph

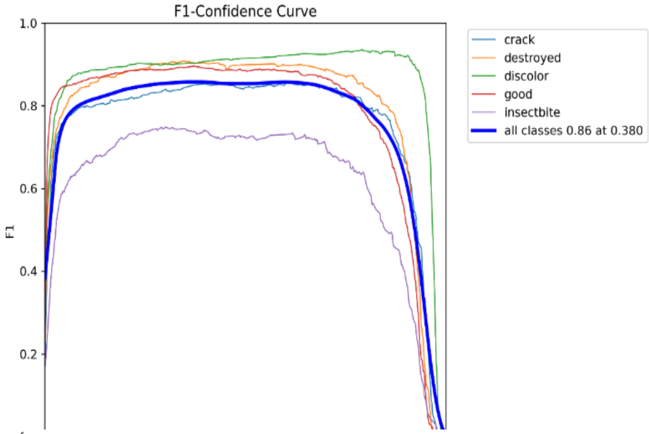

Fig. 7. F1 Confidence Curve

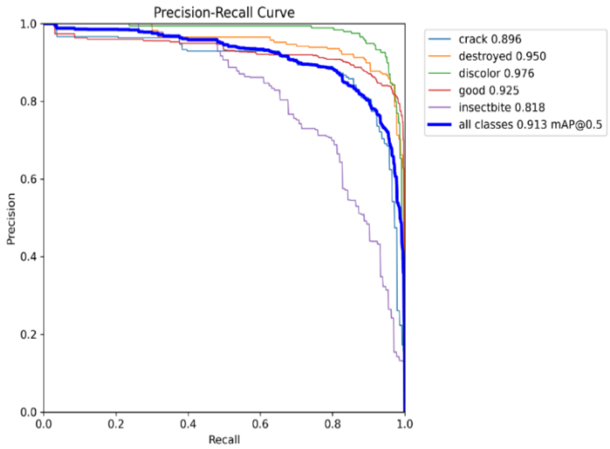

Fig. 8. PR Curve

Including the precision-recall (PR) curve and F1 curve in the evaluation provides essential insights into the classification model's performance, particularly in scenarios where precision and recall are critical. The PR curve visualizes the trade-off between precision and recall across different classification thresholds, offering a comprehensive view of the model's ability to make accurate positive predictions while capturing all positive instances. Similarly, the F1 curve illustrates how the harmonic mean of precision and recall (F1 score) varies with threshold settings, aiding in the selection of an optimal threshold that balances precision and recall effectively. These curves complement the information provided by the confusion matrix and other metrics, enriching the assessment of the model's performance across varying decision thresholds.

4. CONCLUSION

The results shown from the tables that using YOLOv8n Training version 8 as the CNN for Image Processing Algorithm shows the most positive results with no less than 90% of Accuracy under 75 epochs. Moreover, ACER NITRO 5 and Lenovo IdeaPad Gaming that are both equipped with Nvidia graphics exhibit the most reliable and stable visual output performance in accommodating the YOLOv8 Algorithm with the 28-34FPs. Furthermore, cameras that shows the most potent performance in sampling and in testing of the accuracy of the CNN are the 48MP microscope camera due to its clarity and the reliability of captures images both in the sampling phase and the testing phase of the study. For future iterations of the study, the researchers recommend the utilization of a higher-end processor. The enhanced fps rate enhances the accuracy and efficiency of the inspection process, leading to improved outcomes. Additionally, it is recommended to use a more robust and higher performance digital camera for an improve data sampling and testing of the CNN.

5. REFERENCES

Reference to a journal publication:

[1] Philippine Coffee Board Inc. (2022). Philippine Coffee Board. Retrieved February 9, 2023, from https://philcoffeeboard.com/.

[2] Philippine Coffee Board, Inc. (2023). OUR COFFEE HERITAGE: Coffee's Rich History in the Philippines. Retrieved November 10, 2023, from https://philcoffeeboard.com/philippinecoffee/#:~:text=The%20Philippines%20is%20one%20of,distinctively%20robust%20and%20powerful%20cup.

[3] Lualhati, A. J. N., Mariano, J. B., Al Eugene, L. T., & Fenol, S. D. (2022). Development and Testing of Green Coffee Bean Quality Sorter using Image Processing and Artificial Neural Network. Mindanao Journal of Science and Technology, 20(1)

[4] Gao, J., Dai, S., Huang, J., Xiao, X., Liu, L., Wang, L., & Li, M. (2022). Kiwifruit detection method in orchard via an improved light-weight YOLOv4. Agronomy, 12(9), 2081.Engineering, vol. 38, no. 3, pp. 233-240, 2021, doi: 10.7736/KSPE.2021.38.3.233.

[5] Dili, T. L. (2018). Classification of green coffee beans by convolutional neural network and its implementation on raspberry Pi and Camera Module.

[6] Pinto, C., Furukawa, J., Fukai, H., & Tamura, S. (2017). Classification of Green coffee bean images based on defect types using convolutional neural network (CNN). In 2017 International Conference on Advanced Informatics, Concepts, Theory, and Applications (ICAICTA) (pp. 1-5). doi: 10.1109/ICAICTA.2017.8090980.

[7] Balbin, J. R., Del Valle, C. D., Van Julius Leander, G. L., & Quiambao, R. F. (2020, December). Grading and Profiling of Coffee Beans for International Standards Using Integrated Image Processing Algorithms and Back-Propagation Neural Network. In 2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM) (pp. 1-6). IEEE.

[8] Zhao, G., Quan, L., Li, H., Feng, H., Li, S., Zhang, S., & Liu, R. (0000). Real-time recognition system of soybean seeds full-surface defects based on deep learning. Computers and Electronics in Agriculture, 187, 106230. doi: 10.1016/j.compag.2021.106230

[9] Marshall, A. D., & Martin, R. R. (1992). Computer vision, models and inspection. In World Scientific series in robotics and intelligent systems. https://doi.org/10.1142/1492

[10] Szeliski, R. (2011). Computer Vision. In Texts in computer science. https://doi.org/10.1007/978-1-84882-935-0