

Detecting Coffee Bean Types using YOLOv11 Architecture

Muhammad Farhan¹, Esi Putri Silmina²

Abstract

Indonesia ranks fourth as the world's largest coffee producer with production reaching 789,000 tons annually, dominated by Arabica and Robusta varieties. The classification of these two varieties at the industrial level currently relies on highly subjective manual visual inspection. Conventional CNN methods are unable to provide object coordinate localization information essential for industrial actuator systems to perform automated physical separation on high-speed production lines. Thus, this study implements the YOLOv11m architecture to detect and classify coffee bean types using 3,705 images. The model demonstrates highly precise detection performance, achieving a Mean Average Precision (mAP@0.5) of 98.5%, F1-Score of 97.8%, precision of 97.9%, and recall of 97.8%, with an inference speed of 19.6 milliseconds per image enabling processing of more than 38 frames per second. The YOLOv11m-based classification system delivers an accurate, and efficient detection solution suitable for direct integration into automated conveyor sorting machines within the smart agriculture industry.

Keywords:

Deep Learning, Object Detection, Coffee Bean, Classification, YOLOv11

This is an open-access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license



1. Introduction

Advancements in artificial intelligence and digital image processing technology have brought significant transformations to various industrial sectors, including agriculture and plantations. In the context of the global industry, Indonesia ranks fourth as the largest coffee producer in the world after Brazil, Vietnam, and Colombia. The Ministry of Agriculture recorded that national coffee production reached approximately 789,000 tons per year for the 2022–2025 period. This total production consists of 150,000 tons of Arabica coffee and 600,000 tons of Robusta coffee [1]. These two superior varieties dominate the domestic market with significantly different physical characteristics, flavors, and economic values [2][3].

The identification and classification process of coffee bean types at the industrial level currently still relies on conventional methods through visual inspection by humans [4]. This assessment relies heavily on the subjectivity and experience of the inspector, resulting in inconsistent classifications. The manual method has limited accuracy and requires a relatively long inspection time, averaging 45 minutes per kilogram [5][6]. Errors in distinguishing between Arabica and Robusta varieties have serious implications, leading to product quality degradation, pricing errors, and a decline in consumer trust [5].

The need for an automated system capable of performing fast, accurate, and consistent

Corresponding Author: Muhammad Farhan (2211501036@student.unisayogya.ac.id)

¹ Muhammad Farhan, Universitas 'Aisyiyah Yogyakarta, 2211501036@student.unisayogya.ac.id

² Esi Putri Silmina, Universitas 'Aisyiyah Yogyakarta, esiputrisilmina@unisayogya.ac.id

identification is highly crucial. The visual morphological differences between Arabica and Robusta varieties are extremely subtle, thus requiring feature extraction precision that is difficult to consistently achieve by the human eye. Object detection-based Deep Learning algorithms play a crucial role in this context because they not only facilitate global image classification but are also capable of precisely localizing the position of each individual coffee bean within a pile. This coordinate localization capability is an absolute prerequisite for conveyor machine actuators to physically separate the coffee beans in the field.

This study proposes the implementation of the YOLOv11m architecture as an innovative solution for real-time detection [7]. The selection of YOLOv11m is based on three main reasons that specifically address the needs of the coffee processing industry. This architecture provides precise coordinate localization capabilities for each coffee bean spatial information essential for conveyor machine actuators to physically separate beans, a requirement that cannot be fulfilled by pure CNN classification methods which only produce labels without object positions. YOLOv11m offers an optimal balance between inference speed and high accuracy required on high-speed production lines, with a lighter architecture compared to previous generations while maintaining superior performance metrics [8][9]. The anchor-free architecture in YOLOv11m is specifically designed to detect small objects with shape variations such as coffee beans, overcoming the limitations of anchor-based methods that require extensive manual configuration and are less adaptive to morphological variations in agricultural objects.

Conventional CNN architectures are often hindered by slow inference times when simultaneously handling dozens of objects in a single frame and are unable to provide the object coordinate localization required for automated physical separation. The YOLOv11m variant solves this issue with a significantly lighter feature extraction system and computation optimized for industrial hardware, while simultaneously delivering precise bounding box coordinate outputs for each coffee bean a critical capability that serves as an absolute prerequisite for conveyor machine actuators to physically separate beans in the field. The novelty of this research is emphasized in three main aspects: (1) adaptation of the cutting-edge YOLOv11m architecture which has not been extensively tested in the agricultural commodity domain, particularly coffee; (2) efficiency testing of real-time detection for separating coffee commodity varieties under fast-moving conveyor simulation conditions with an inference speed of 19.6 milliseconds; and (3) validation of precision localization capabilities for direct integration with industrial actuator systems an area of automation implementation that has not been deeply explored in previous studies.

2. Related Works

The application of the You Only Look Once (YOLO) architecture has proven effective in resolving various challenges across the agricultural, food, and industrial automation sectors. This model continuously undergoes evolution in its detection performance with each iteration. Previous research demonstrates that YOLOv5 is capable of detecting the ripeness level of tomatoes with an accuracy of 73% [10]. The implementation of mosaic data augmentation on YOLOv5sM successfully improved the detection accuracy of beef freshness [11]. An analysis of YOLOv5 performance proved that variations in the number of classes within a dataset significantly affect the final outcomes of evaluation metrics, rendering object category configuration a crucial factor in determining model accuracy [12]. Furthermore, the YOLOv8 architecture has been successfully applied to classify rice leaf diseases, achieving an accuracy of up to 92% [13].

The utilization of the YOLO architecture has expanded into real-time object detection

systems that require a high degree of computational precision. The YOLOv11 model demonstrates efficiency in optimizing the classification of organic and inorganic waste by balancing computational load and accuracy performance [14]. The implementation of this architecture has proven robust in precisely estimating metric distances during the dynamic movements of wheeled soccer robots [15]. Furthermore, the application of the YOLO architecture has been optimized in the MOTOGUARD-AI system to detect indications of motorcycle theft in real-time within urban environments [16]. The superiority of YOLOv11 over YOLOv10 is evident in its ability to detect complex objects such as smoke and fire, demonstrating significant improvements in precision, recall, and mAP values [17].

Focusing on coffee plant commodities, Deep Learning-based approaches such as Convolutional Neural Networks (CNN) dominate the efforts to automate the classification of bean quality and types. The implementation of classic CNNs has been able to identify the physical characteristics of Temanggung coffee with optimal accuracy [18]. A comparative evaluation between conventional CNNs and Faster Region CNN (Faster R-CNN) demonstrated the effectiveness of visual feature extraction in determining the quality standards (grades) of Arabica coffee beans [19]. The use of transfer learning techniques was adapted to accelerate model convergence, including through the application of VGG16 and MobileNetV2 architectures [20]. The combination of the VGG16 model with the GLCM texture feature extraction method exhibited more stable identification results on coffee bean images [21]. An alternative approach based on the Xception architecture was tested to detect Arabica, Robusta, and Liberica varieties from conventional camera image captures [22].

The implementation of object detection algorithms such as YOLO has begun to be adapted to overcome the inference time limitations of conventional image classification methods. An Android-based YOLO model yielded an accuracy of 75% in detecting coffee bean quality [23]. Optimization of the YOLOv5 architecture was capable of achieving an accuracy of up to 90% when specifically used to detect the ripeness level of coffee cherries [24]. The use of YOLOv7 to detect diseases in coffee leaves successfully obtained a precision value of 0.926 and an mAP@0.5 of 0.956 [25]. Efforts to distinguish Arabica and Robusta bean varieties using the YOLOv5 architecture recorded precision and mAP@0.5 metrics reaching 0.99 [26].

The aforementioned previous studies predominantly focused on pure image classification tasks or relied on older-generation object detection architectures whose computational efficiency does not necessarily meet large-scale industrial standards. This study employs YOLOv11m to bridge this gap, particularly in the real-time detection application of coffee bean varieties under static image conditions that simulate an industrial conveyor environment.

3. Proposed Method

The research methodology is designed to develop an objective and efficient coffee bean classification system utilizing a deep learning approach. A systematic research workflow is necessary for the YOLOv11 model to accurately recognize the unique morphological features of Arabica and Robusta coffee beans. The comprehensive stages of this research are illustrated in the flowchart in Fig. 1.

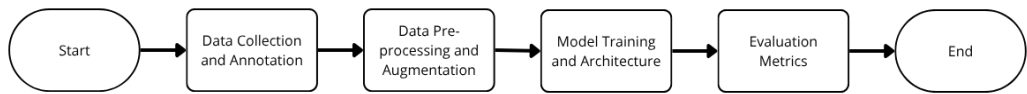


Fig. 1. Research Stage Flow for Coffee Bean Classification

3.1 Data Collection and Annotation

In this study, we utilized the dataset comprises 2,179 original images. The data sources were acquired through primary data collection via direct photography, alongside secondary data from public repositories such as Roboflow and Kaggle. Visually, Arabica coffee beans generally exhibit a flatter and more elongated shape, whereas Robusta beans tend to be rounder and more convex. These physical morphological characteristics, color, and texture are essential visual features that significantly influence the success rate of the deep learning model in accurately learning and distinguishing coffee varieties [27]. Examples of visual representations of both varieties within this research dataset are illustrated in Fig. 2 and 3.



Fig. 2. Arabica Coffee Bean



Fig. 3. Robusta Coffee Bean

The annotation process was conducted manually using labeling software to determine the bounding box for each individual coffee bean to generate precise location coordinates and class labels for the model. Accuracy and consistency in determining bounding box coordinates in the initial dataset is a highly crucial stage to ensure the CNN architecture model can recognize the target and properly separate the region of interest from the background [28]. The total annotations generated across the entire dataset amounted to 5,887 bounding boxes, consisting of 2,904 annotations for the Arabica class and 2,983 annotations for the Robusta class.

3.2 Data Pre-processing and Augmentation

Image quality was enhanced through a pre-processing stage to ensure stability during model training. All images were uniformly resized to 640×640 pixels, then subjected to Auto-Orient and contrast stretching (Auto-Contrast). This contrast adjustment process is crucial for clarifying the intensity differences between the dark and light areas in the images, thereby making the features and shape boundaries of the coffee bean objects more distinct; this has been significantly proven to improve the results and accuracy of object recognition in deep learning architectures [21][22].

For objective model evaluation purposes, the dataset was randomly divided into a proportion of 70% training data, 20% validation data, and 10% testing data. This 70:20:10 split ratio is a customary and optimal distribution standard in deep learning development,

as it provides a sufficiently large portion of data for the model to learn patterns comprehensively, while simultaneously maintaining the availability of entirely unseen data to measure the objectivity of the final evaluation [31]. The detailed distribution of the dataset split is presented in Table 1.

Table 1. Coffee Image Dataset Distribution

Category	Percentage	Number of Images
Training Set	70%	1.526
Validation Set	20%	436
Testing Set	10%	217
Total Overall	100%	2.179

Extensive image augmentation techniques were applied exclusively to the Training Set to enhance visual variation and minimize the risk of overfitting. The augmentations utilized include horizontal and vertical flipping (flip), static rotation (90° Clockwise, Counter-Clockwise, and 180°), dynamic rotation ($\pm 15^\circ$), and brightness level adjustments ($\pm 15\%$). The implementation of this data augmentation aims to diversify the training image samples, thereby strengthening the model's robustness in recognizing general object patterns across various layout variations and lighting conditions when encountering new data in the field [32]. Through the application of this data augmentation, the number of images in the training set was successfully increased to 3,052 images. Consequently, the total number of images utilized throughout the entire study reached 3,705 images.

3.3 Model Training and Architecture

The implementation of the model utilizes the YOLOv11m variant, which adopts an anchor-free architecture to enhance detection efficiency for small objects such as coffee beans. This architecture consists of three main components: the Backbone for feature extraction, the Neck for multi-scale feature fusion, and the Head for bounding box prediction and classification. The detailed image processing mechanism within the YOLOv11m architecture is explained as follows:

- 1) **Basic Convolution and Activation Operations:** Each layer in YOLOv11m performs two-dimensional convolution operations (*Conv2D*) to extract spatial features such as edges, texture, and the shape of the coffee beans. Mathematically, the output from a single convolution layer (Y) is defined in Equation (1):

$$Y_{i,j,k} = \sigma \left(\sum_{l,m,n} W_{l,m,n,k} \cdot X_{i+l,j+m,n} + b_k \right) \quad (1)$$

Description: W represents the kernel weights, X is the input matrix, b is the bias, and σ is the Sigmoid Linear Unit (SiLU) activation function. The SiLU equation is expressed as follows:

$$\sigma(x) = x \cdot \text{sigmoid}(x) = \frac{x}{1 + e^{-x}} \quad (2)$$

- 2) **Cross Stage Partial Bottleneck (C3k2) Structure:** The primary component in YOLOv11 feature extraction is the C3k2 block, which divides the feature map into two paths to reduce computational redundancy. This structure enables the model to learn the surface texture differences of Arabica and Robusta beans in a layered manner. The processing logic for this block is formulated in Equation (3):

$$X_{out} = \text{Conv} \left(\text{Concat} \left(X_{part1}, \text{BottleneckSeq}(X_{part2}) \right) \right) \quad (3)$$

Where X_{part1} and X_{part2} are the results of the division of the main input, while

BottleneckSeq is a sequence of bottleneck layers that deepens the feature representation.

- 3) **Spatial Pyramid Pooling - Fast (SPPF)**: This layer functions to handle variations in coffee bean sizes within a single frame by performing serial feature pooling with a fixed kernel size of 5×5 . The feature fusion operation in SPPF is formulated as follows:

$$Y_{SPPF} = \text{Concat}\left(X, P_1(X), P_2(P_1(X)), P_3(P_2(P_1(X)))\right) \quad (4)$$

Description: P represents the max pooling operation. The result of this fusion produces an efficient multi-scale feature representation.

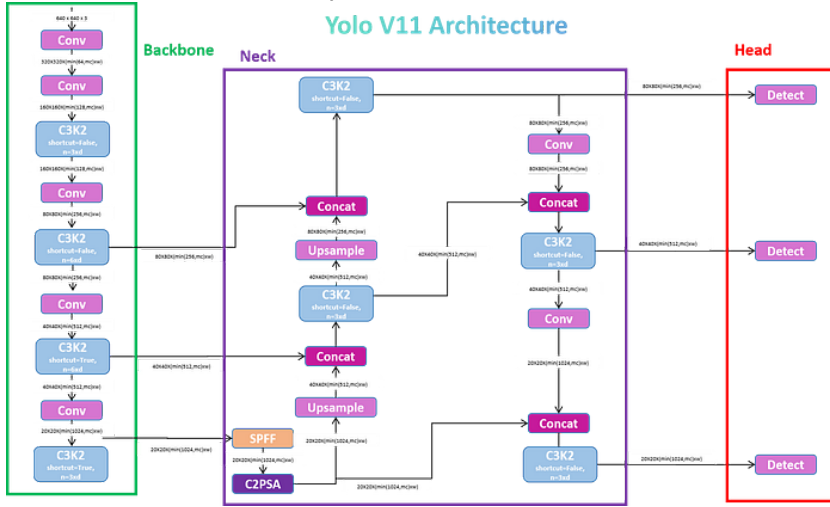


Fig. 4. YOLOv11 Model Architecture [33]

The training process was conducted by specifically configuring the hyperparameters to achieve stable convergence. The selection of the AdamW optimizer was based on its ability to perform weight decay separately, which has proven more effective for object detection models compared to the standard Adam or SGD optimizers. The details of the training parameter configurations used in this study are presented in Table 2.

Table 2. Model Training Hyperparameter Configuration

Parameter	Value	Reference/Justification
Model Version	YOLOv11m	Balance of accuracy and speed
Input Size	640 x 640	Standard YOLOv11 resolution
Epochs	100	Convergence stability
Batch Size	16	Tesla T4 GPU memory optimization
Optimizer	AdamW	Weight decay efficiency
Initial Learning Rate	0.001	Standard AdamW learning rate

3.4 Evaluation Metrics

Model performance evaluation was conducted using four primary metrics to provide a comprehensive overview of the accuracy, sensitivity, and reliability of the detection system. The use of these various metrics ensures that the developed YOLOv11 model possesses high validity, both statistically and operationally in the field.

- 1) **Precision**, or the precision value, measures the model's ability to provide correct

positive predictions compared to the total positive predictions generated by the object recognition system [34]. This metric is highly crucial in coffee bean classification to minimize false positives, which is the condition where the Robusta variety is incorrectly detected as Arabika or vice versa. Equation (5) is used to calculate the precision value:

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

- 2) **Recall**, or sensitivity, measures the success ratio of the model in redetecting all target objects actually present in the input image (ground truth) [35]. A high recall value indicates that the model is capable of recognizing most of the coffee beans in the sample without many objects being missed during the detection process. The recall calculation is performed using Equation (6):

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

Description for Equations (5) and (6):

- **TP (True Positive)**: The number of coffee bean targets correctly detected and classified by the model according to their original class.
 - **FP (False Positive)**: The number of incorrect predictions, specifically when the model detects an object that does not match its original class (misclassification).
 - **FN (False Negative)**: The number of actual coffee bean targets that the model missed or failed to detect entirely.
- 3) **Mean Average Precision (mAP)** is a comprehensive evaluation metric that globally combines precision and sensitivity aspects. [36]. This value is calculated from the average of the Average Precision (AP) of each object class. This study utilizes mAP@0.5 (detection accuracy at an Intersection over Union or IoU threshold of 0.5) and mAP@0.5:0.95. The mAP@0.5:0.95 indicator is a stricter metric calculated from the average mAP values across an IoU range of 0.5 to 0.95 with an increment interval of 0.05 to evaluate the level of bounding box precision against the ground truth. The general calculation of mAP is expressed through Equation (7):

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i \quad (7)$$

Description for Equation (7):

- n : The total number of categories or object classes evaluated (in this case, $n = 2$, namely Arabika and Robusta).
 - AP_i : The Average Precision value obtained for the i -th class.
- 4) **F1-Score** is the harmonic mean that connects precision and recall to reflect the overall balance and stability of the model's performance [37]. This metric serves as a highly reliable parameter to ensure that the model is not biased toward only one metric, but instead maintains stable performance in recognizing objects accurately and completely. The F1-Score value is calculated using Equation (8):

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (8)$$

4. Result and Analysis

Model testing was conducted using 217 testing set images that were entirely unseen during the training process. The evaluation results in Table 3 demonstrate that the YOLOv11m model is highly reliable in distinguishing between Arabika and Robusta

varieties. This is evidenced by the Mean Average Precision (mAP@0.5) reaching 98.5% and a balance score (F1-Score) of 97.8%. In more detail, the model recorded a precision rate of 97.9% and an object detection success rate (recall) of 97.8%. These figures indicate that the model almost never misidentifies the coffee types and rarely misses any coffee beans present in the images.

Table 3. Model Performance Evaluation Results on Testing Set

Class / Category	Precision	Recall	mAP@0.5	mAP@0.5:0.95	F1-Score
Arabica	96,4%	98,9%	98,8%	90,5%	97,6%
Robusta	99,4%	96,7%	98,2%	86,9%	98,0%
Total (All)	97,9%	97,8%	98,5%	88,7%	97,8%

This stable detection quality is also clearly visible in the Precision-Recall curve in Fig 5. The curve shows results that are very close to perfect values for all classes. The large area under the curve indicates that the model maintains a high level of accuracy even as the number of detected objects increases, meaning the model is highly consistent in recognizing coffee beans.

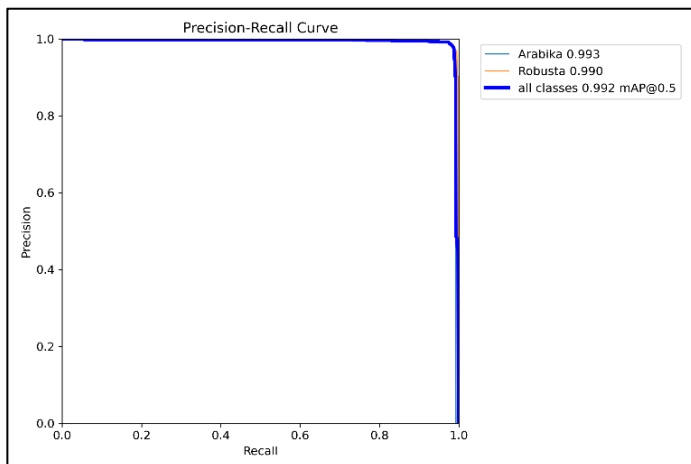


Fig. 5. Precision-Recall Curve for Model Testing

When examining the prediction results more closely through the Confusion Matrix in Fig. 6, the model demonstrates high consistency in separating the two varieties. Arabica coffee beans were correctly predicted 99% of the time, while Robusta reached 98%. Prediction errors between the coffee types were minimal, occurring only at approximately 1% to 2%. This demonstrates that the model is exceptionally capable of distinguishing the very similar physical features between the two types of beans.

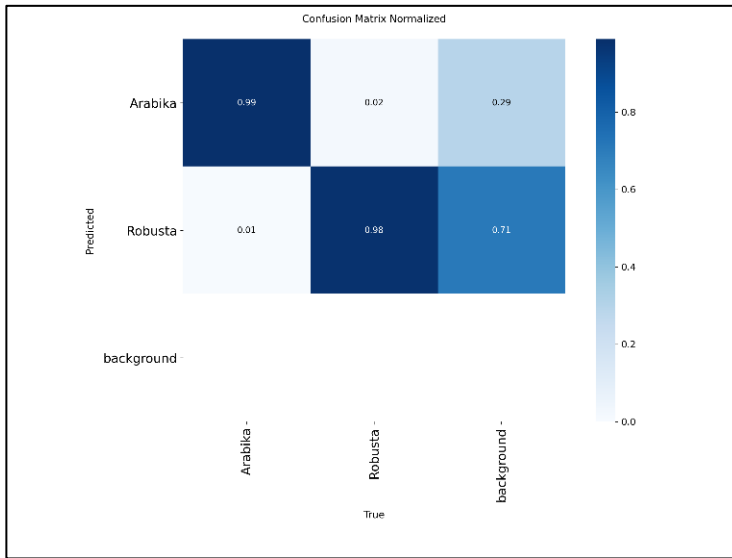


Fig. 6. Normalized Confusion Matrix of Prediction Results

This success is supported by a training process that proceeded stably, as demonstrated by the metric graphs in Fig. 7. The error rate (loss) consistently decreased until it reached its lowest point at the end of the training session at epoch 100. This trend indicates that the model successfully learned patterns from the training data effectively without encountering issues of overfitting or learning failure.

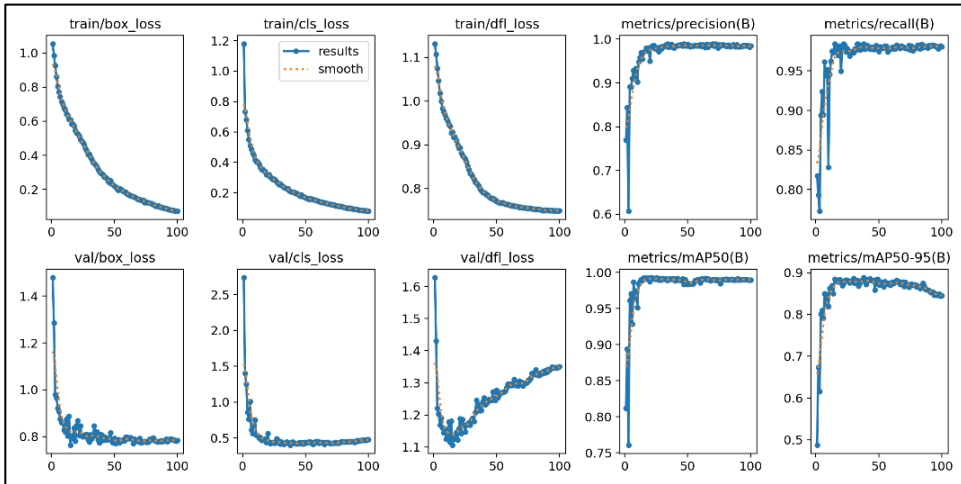


Fig. 7. Training and Validation Metric Curve Graphs for YOLOv11m

In addition to statistical figures, the model's capabilities were also directly tested through visual detection results in Fig. 8. The model is able to detect each coffee bean with a high degree of confidence, with average confidence scores reaching 0.9. This proves that the image contrast adjustment stage performed previously significantly aided the model in recognizing the physical boundaries of the coffee beans more distinctly.

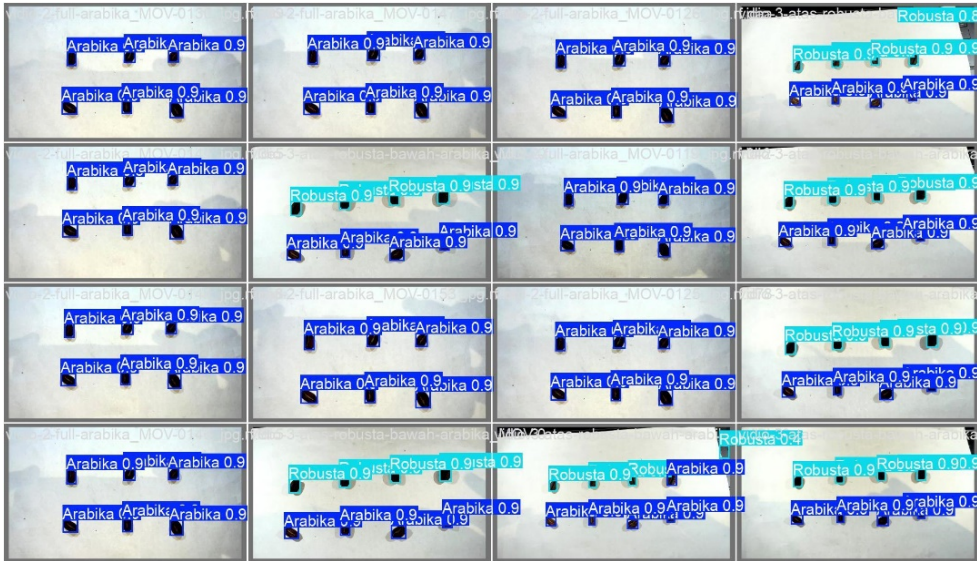


Fig. 8. Visual Representation of Coffee Bean Detection and Classification Results

According to the experimental results, this study provide a significant improvement compared to previous methods. In comparison, the YOLOv5 model used in previous research achieved approximately 90% accuracy for coffee ripeness detection. Meanwhile, conventional methods such as pure CNN or the use of VGG16 and MobileNetV2 architectures require heavier computational processes to achieve stable results. The primary advantage of YOLOv11m here is its ability not only to determine the coffee variety but also to precisely locate the coordinates of each individual coffee bean. This localization feature is crucial for automated sorting machines and is not possessed by standard image classification methods used in some previous studies.

5. Conclusion

This paper successfully demonstrates that the implementation of the YOLOv11m architecture provides an effective and reliable solution for automated real-time detection and classification of Arabika and Robusta coffee beans. We utilize deep learning–based object detection to overcome the major limitations of manual visual inspection that still dominate the Indonesian coffee industry, including subjective judgment, inconsistent classification results, and long inspection durations reaching up to 45 minutes per kilogram. Our proposed method can produce simultaneous classification and coordinate localization information, allowing the system to support automated sorting mechanisms on conveyor-based production lines. The integration of object detection and localization capabilities makes this study highly relevant for modern smart agriculture and industrial automation applications, particularly for improving productivity and maintaining coffee bean quality consistency during large-scale processing activities.

This study can produce highly accurate and robust classification performance through the application of the YOLOv11m model. Our proposed method achieves a Mean Average Precision (mAP@0.5) of 98.5%, precision of 97.9%, recall of 97.8%, and an aggregate F1-Score of 97.8%, indicating excellent capability in distinguishing subtle morphological differences between Arabika and Robusta coffee beans. We also utilize mAP@0.5:0.95 evaluation to validate geometric localization accuracy, where the obtained value of 88.7% confirms precise bounding-box prediction performance. The confusion matrix analysis

reveals very low classification errors of only 1–2%, while the training curves show stable convergence without significant overfitting until epoch 100. In addition, this paper demonstrates that the proposed YOLOv11m architecture provides substantial advantages compared to conventional CNN classification approaches such as VGG16 and MobileNetV2, because the proposed system performs both classification and object localization simultaneously in real time.

Furthermore, our proposed method can harvest significant computational efficiency that fulfills industrial real-time processing requirements. The system records an inference time of only 19.6 milliseconds per image with a total processing cycle below 26 milliseconds, enabling throughput exceeding 38 FPS without creating operational bottlenecks in high-speed production environments. This paper confirms that YOLOv11m achieves a balanced combination of accuracy, speed, and localization precision compared to older detection frameworks such as Faster R-CNN, which generally require substantially higher computational latency. However, this study still has several limitations, including the restricted dataset containing only Arabica and Robusta varieties and the absence of testing under motion blur, vibration, and extreme lighting conditions. Therefore, future research can expand the dataset with additional coffee varieties and defect categories, validate the model in dynamic industrial environments, and integrate the system with IoT devices, robotic arms, or pneumatic sorting mechanisms to realize a fully automated end-to-end coffee bean separation platform for the national coffee processing industry.

References

- [1] “Berita BRMP Perkebunan - Tren 2025: Peluang dan Daya Saing Kopi.” Accessed: Apr. 19, 2026. [Online]. Available: <https://perkebunan.brmp.pertanian.go.id/berita/tren-2025-peluang-dan-daya-saing-kopi>
- [2] K. Muiz, N. Nilahayati, N. Nasruddin, J. Jamidi, and H. Haffah, “Analisis Kualitas Fisik Dan Cita Rasa Kopi Arabika (*Coffea arabica* L.) Pada Dua Rentang Ketinggian,” *J. Agrium*, vol. 21, no. 3, p. 258, Sep. 2024, doi: 10.29103/agrium.v21i3.18994.
- [3] D. A. Mahardhika, A. H. Antonius, and B. Dwiloka, “Perbedaan Sifat Fisikokimia dan Organoleptik Produk Kopi Rempah dari Kopi Arabika (*Coffea arabica*) dan Kopi Robusta (*Coffea robusta*),” *J. Apl. Teknol. Pangan*, vol. 11, no. 4, Nov. 2022, doi: 10.17728/jatp.13827.
- [4] M. Farhan, A. Fauzi, and H. Khair, “Identifikasi Kualitas Biji Kopi pada Pengolahan Citra Menggunakan Metode Backpropagation,” *Kohesi J. Sains Dan Teknol.*, vol. 10, no. 5, pp. 71–80, Sep. 2025, doi: 10.2238/f3anjc67.
- [5] L. H. Rambe, Y. Manza, A. Ashari, and W. S. Negoro, “Klasifikasi Biji Kopi Menggunakan Teknik Kombinasi Random Forest dan Inception V3 untuk Ekstraksi Fitur,” *J. Sci. Soc. Res.*, vol. 8, no. 3, pp. 3845–3851, Aug. 2025, doi: 10.54314/jssr.v8i3.3976.
- [6] A. Aripriharta *et al.*, “Integrasi Smart Camera dan AI untuk Mendukung Agribisnis Kopi Berkelanjutan di Wilayah Pedesaan Indonesia,” *Abditeknika J. Pengabd. Masy.*, vol. 5, no. 2, pp. 88–97, Oct. 2025, doi: 10.31294/abditeknika.v5i2.9602.
- [7] A. I. Pradana, H. Harsanto, and W. Wijyanto, “Deteksi Rambu Lalu Lintas Real-Time di Indonesia dengan Penerapan YOLOv11: Solusi Untuk Keamanan Berkendara,” *J. Algoritma*, vol. 21, no. 2, pp. 145–155, Nov. 2024, doi: 10.33364/algoritma/v.21-2.2106.
- [8] E. P. Silmina, S. Sunardi, and A. Yudhana, “Comparative Analysis of YOLO Deep Learning Model for Image-Based Beef Freshness Detection,” *JITK J. Ilmu Pengetah. Dan Teknol. Komput.*, vol. 11, no. 1, pp. 250–265, Aug. 2025, doi: 10.33480/jitk.v11i1.6784.
- [9] T. Takyudin, M. S. Rais, J. A. Putra, and A. Hamsar, “Metode Deteksi dan Estimasi Luas Lubang Jalan Menggunakan Deep Learning Berbasis YOLOv11,” *RIGGS J. Artif. Intell. Digit. Bus.*, vol. 4, no. 2, pp. 6786–6790, Jul. 2025, doi: 10.31004/riggs.v4i2.1760.
- [10] S. Aras, P. Tanra, and M. Bazhar, “Deteksi Tingkat Kematangan Buah Tomat Menggunakan YOLOv5: Detection of Tomato Fruit Ripeness Level Using YOLOv5,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 2, pp. 623–628, Mar. 2024, doi: 10.57152/malcom.v4i2.1270.
- [11] A. Yudhana, E. P. Silmina, and Sunardi, “Deteksi Kesegaran Daging Sapi Menggunakan Augmentasi Data Mosaic pada Model YOLOv5sM,” *JRST J. Ris. Sains Dan Teknol.*, pp. 63–71, Apr. 2025, doi: 10.30595/jrst.v9i1.24990.

- [12] E. P. Silmina, T. Hardiani, and S. L. Mahfida, "The Effect of the Number of Classes on the Values Resulting from Evaluation Metrics in the YOLOv5 Model," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 15, no. 2, pp. 419–425, Apr. 2025, doi: 10.18517/ijaseit.15.2.20495.
- [13] A. A. Firdaus, A. Rosadi, and A. H. Tantri, "Penerapan Algoritma Yolov8 Untuk Mendeteksi Penyakit Pada Daun Padi," 2024.
- [14] B. A. Wisesa, V. M. Putri, E. Faristasari, S. A. J. Duli, and R. Lionza, "Optimizing Waste Classification Model using YOLOv11 Architecture," *Int. J. Inform. Comput.*, vol. 7, no. 2, pp. 735–743, Dec. 2025, doi: 10.35842/ijicom.v7i2.212.
- [15] S. M. Irsyad, A. Khumaidi, R. Y. Adhitya, A. R. Ramadhan, D. A. Pratama, and M. J. Saputra, "Predicting Real Distance for Wheeled Soccer Robot using YOLO Architecture," *Int. J. Inform. Comput.*, vol. 7, no. 2, pp. 806–821, Dec. 2025, doi: 10.35842/ijicom.v7i2.200.
- [16] N. Isnaeni, B. A. Wisesa, A. D. P. Wicaksono, and S. A. Darma, "MOTOGUARD-AI: Real-Time Motorcycle Theft Detection Using YOLO Architecture," *Int. J. Inform. Comput.*, vol. 8, no. 1, pp. 18–29, Jan. 2026, doi: 10.35842/ijicom.v8i1.215.
- [17] E. H. Alkhamash, "A Comparative Analysis of YOLOv9, YOLOv10, YOLOv11 for Smoke and Fire Detection," *Fire*, vol. 8, no. 1, p. 26, Jan. 2025, doi: 10.3390/fire8010026.
- [18] P. A. Maharani and M. Akbar, "Implementasi Convolutional Neural Network dalam Klasifikasi Jenis Kopi Temanggung," *JATI J. Mhs. Tek. Inform.*, vol. 8, no. 3, pp. 3030–3037, May 2024, doi: 10.36040/jati.v8i3.9582.
- [19] G. A. Pratama, E. Y. Puspiningrum, and H. Maulana, "Convolutional Neural Network dan Faster Region Convolutional Neural Network untuk Klasifikasi Kualitas Biji Kopi Arabika," *J. Inform. Dan Tek. Elektro Terap.*, vol. 12, no. 3, Aug. 2024, doi: 10.23960/jitet.v12i3.4887.
- [20] M. Murinto, M. Rosyda, and M. Melany, "Klasifikasi Jenis Biji Kopi Menggunakan Convolutional Neural Network dan Transfer Learning pada Model VGG16 dan MobileNetV2," *JRST J. Ris. Sains Dan Teknol.*, vol. 7, no. 2, p. 183, Sep. 2023, doi: 10.30595/jrst.v7i2.16788.
- [21] M. Y. A. Hasibuan, W. Marleta, A. Y. Sibuea, M. Sultan, and M. Ula, "Klasifikasi Jenis Biji Kopi Menggunakan Convolutional Neural Network dan Transfer Learning pada Model VGG16 dan GLCM 47E954," *Pros. Semin. Nas. Teknol. Dan Tek. Inform. SENASTIKA*, vol. 1, no. 1, Nov. 2024, Accessed: May 15, 2026. [Online]. Available: <https://proceedings.unimal.ac.id/senastika/article/view/834>
- [22] M. A. Pratama and T. Y. Hadiwandura, "Implementasi Deep Learning Untuk Identifikasi Jenis Biji Kopi Menggunakan Metode Convolutional Neural Network," *CSRID Comput. Sci. Res. Its Dev. J.*, vol. 17, no. 3, pp. 387–398, Oct. 2025, doi: 10.22303/csrid-17.3.2025.387-398.
- [23] F. Kamil, "Pengolahan Citra Digital Menggunakan Metode Yolo Untuk Mendeteksi Kualitas Dari Biji Kopi Berbasis Android," vol. 1, no. 1, 2023.
- [24] Arfan Astaraja, Bilal Shandyarta Syamsudin, Muhammad Diaz Maulana Dhafin, Fatkhul Hidayah, and Niccola Jody Setiawan, "Optimizing Coffee Ripeness Classification Using Yolov5 for Automated Detection and Sorting," *J. Tek. Mesin Ind. Elektro Dan Inform.*, vol. 4, no. 1, pp. 278–294, Feb. 2025, doi: 10.55606/jtmei.v4i1.4821.
- [25] A. Ardiansyah and N. F. Hasan, "Deteksi dan Klasifikasi Penyakit Pada Daun Kopi Menggunakan Yolov7," *J. Sisfokom Sist. Inf. Dan Komput.*, vol. 12, no. 1, pp. 30–35, Mar. 2023, doi: 10.32736/sisfokom.v12i1.1545.
- [26] D. Buonocore, M. Carratù, and M. Lamberti, "Classification of coffee bean varieties based on a deep learning approach," in *Proceedings of the 18th IMEKO TC10 Conference on Measurement for Diagnostics, Optimisation and Control*, Warsaw, POLAND: IMEKO, 2023, pp. 14–19. doi: 10.21014/tc10-2022.002.
- [27] M. R. Al Firdaus, R. Mardhiyyah, and F. I. Sanjaya, "Klasifikasi Citra Biji Kopi Sangrai Arabika dan Robusta Menggunakan Convolutional Neural Network," *TIN Terap. Inform. Nusant.*, vol. 6, no. 7, pp. 969–980, Dec. 2025, doi: 10.47065/tin.v6i7.8695.
- [28] T. Pascal Munthe and M. Akbar, "Klasifikasi Citra Biji Kopi Temanggung Menggunakan Residual Network (ResNet-50)," *J. Pustaka Data Pus. Akses Kaji. Database Anal. Teknol. Dan Arsit. Komput.*, vol. 5, no. 1, pp. 94–102, Jun. 2025, doi: 10.55382/jurnalpustakadata.v5i1.1028.
- [29] R. Kurniawan, K. I. Apriandy, T. Badriyah, and I. Syarif, "The Impact of Image Pre-processing for Tuberculosis Prediction System Based on Chest X-ray Images," *J. Inform.*, vol. 10, no. 4, 2025.

- [30] T. Wang, G. T. Kim, M. Kim, and J. Jang, "Contrast Enhancement-Based Preprocessing Process to Improve Deep Learning Object Task Performance and Results," *Appl. Sci.*, vol. 13, no. 19, p. 10760, Sep. 2023, doi: 10.3390/app131910760.
- [31] F. Saputro and A. Witanti, "Implementasi Algoritma YOLOv11 untuk Sistem Klasifikasi Kelayakan Setor Sampah Anorganik dalam Pengelolaan Bank Sampah," *J. Inf. Syst. Res. JOSH*, vol. 6, no. 4, pp. 1789–1798, Jul. 2025, doi: 10.47065/josh.v6i4.7486.
- [32] A. S. Ardiansyah and A. Nugroho, "Klasifikasi Penyakit Daun Kopi Dengan Arsitektur MobileNetV2," *J. Ilmu Komput. Dan Bisnis*, vol. 14, no. 1, pp. 66–73, May 2023, doi: 10.47927/jikb.v14i1.622.
- [33] S. N. Rao, "YOLOv11 Explained: Next-Level Object Detection with Enhanced Speed and Accuracy," Medium. Accessed: May 16, 2026. [Online]. Available: <https://medium.com/@nikhil-rao-20/yolov11-explained-next-level-object-detection-with-enhanced-speed-and-accuracy-2dbe2d376f71>
- [34] A. A. Mustapha, S. 'Atifah Saruchi, H. Supriyono, and M. I. Solihin, "A Hybrid Deep Learning Model for Waste Detection and Classification Utilizing YOLOv8 and CNN," in *The 8th Mechanical Engineering, Science and Technology International Conference*, MDPI, Mar. 2025, p. 82. doi: 10.3390/engproc2025084082.
- [35] R. R. Achmad, M. Abil, M. R. Fadhilah, and Sandi, "Comparative Evaluation of YOLOv5, YOLOv7, and YOLOv8 for Outdoor Traffic Object Detection," *Int. J. Appl. Math. Comput.*, vol. 3, no. 2, pp. 32–39, Apr. 2026, doi: 10.62951/ijamc.v3i2.291.
- [36] L. A. Nguyen, M. D. Tran, and Y. Son, "Empirical Evaluation and Analysis of YOLO Models in Smart Transportation," *AI*, vol. 5, no. 4, pp. 2518–2537, Nov. 2024, doi: 10.3390/ai5040122.
- [37] O. F. Odeyinka, W. A. Raheem, C. O. Folorunso, and A. Fajinmi, "Performance Evaluation of Yolo Object Detection Models for Automated License Plate Recognition," *ARID ZONE J. Eng. Technol. Environ.*, vol. 21, no. 2, pp. 449–459, 2025.