JRSSEM 2026, Vol. 05, No. 06, January
E-ISSN: 2807 - 6311, P-ISSN: 2807 - 6494

Classification of Roasting Maturity Levels of Coffee Beans Using CNN Method Based on Mobilenetv2

Renalda Geriel Rafidan Arsyan[1]*, Arrie Kurniawardhani[2], Irving Putra Paputungan[3]
Universitas Islam Indonesia, Indonesia
Email: 21523007@uii.ac.id*, arrie.kurniawardhani@uii.ac.id, irving@uii.ac.id

Abstract. Determining the roasting maturity level of coffee beans is an important process to maintain consistency in flavor quality. However, the assessment process, which is still largely manual, tends to be subjective and highly dependent on the experience of farmers. This research develops an automatic classification model for four categories of coffee bean roasting levels—green, light, medium, and dark—using a convolutional neural network (CNN) architecture based on MobileNetV2. The dataset was divided into training, validation, and testing sets with a ratio of 75:15:10. The model was trained in two stages: initial training with a frozen base model, followed by fine-tuning of the last quarter of the layers. The experimental results show that the model achieved an accuracy of 96% with stable performance, as indicated by the loss and accuracy curves. These findings demonstrate that MobileNetV2 can serve as an effective solution for classifying coffee bean roasting levels with efficient computational time and competitive accuracy.

Keywords: Convolutional Neural Network; MobileNetV2; Image Classification; Coffee Beans; Roasting.

INTRODUCTION

The quality of coffee beans is greatly influenced by the roasting process, as this stage affects their color, aroma, and physical characteristics (Girma & Sualeh, 2022; Lu et al., 2023; Rusinek et al., 2024; Sangpimpa et al., 2025; Tarigan et al., 2022). Visual changes during roasting can indicate the maturity level, but assessments remain mostly manual, performed by business actors and farmers. This method is inconsistent and heavily influenced by each assessor's perception, as shown in the study by Anggraini et al. (2021), which demonstrated that visual assessments often produce inaccuracies in some coffee bean samples. This reliance on subjective evaluation poses a challenge, particularly for producers needing consistent quality standards across production stages (Bouramdane, 2023; Nastoska et al., 2025; Wang et al., 2024).

As digital image processing technology advances, domestic research has increasingly utilized computational methods to assess coffee bean quality (Kim et al., 2024; Lavanya et al., 2025; Li et al., 2024; Motta et al., 2024; Santoso et al., 2024). Setiawan et al. (2022) developed a classification model based on convolutional neural networks (CNN) to determine coffee bean maturity and showed that deep learning approaches achieve much higher accuracy than conventional image processing methods. Similarly, Rahmatullah and Prasetyo (2022) used Residual Networks to prove that visual features in coffee beans can be automatically recognized with modern deep learning models. These findings confirm that deep learning models offer a practical solution for enhancing the objectivity of coffee bean quality assessment in Indonesia.

Additionally, Adriyanto and Kurniawati (2023) compared CNN models using VGG16 and MobileNetV2 to classify local coffee bean types, achieving high accuracy in both—especially MobileNetV2, which excelled in processing efficiency and speed. These results highlight the potential of lightweight models for roasting maturity classifications requiring fast inference times and minimal computational resources. In contrast, studies like Wibowo et al. (2022) relied on LCH color features for quality assessment but struggled to capture texture variations and complex patterns in coffee beans at different roasting levels.

Efforts to apply computer vision also appear in Ardiansyah et al. (2021), which used color and texture features to detect quality defects in Arabica coffee beans. The study demonstrated that image-based approaches can objectively identify visual differences, though non-deep learning methods face limitations with highly variable images. Thus, deep learning is increasingly relevant for handling diverse coffee bean image datasets, including subtle differences between roasting maturity levels.

These national studies show that machine learning and deep learning enable effective visual analysis for coffee bean quality assessment and type identification. However, lightweight architectures like MobileNetV2 have rarely been applied in Indonesia specifically for classifying roasting maturity levels into four categories (green, light, medium, dark). This research gap motivates the current study, which develops an accurate, efficient classification system suitable for limited computing devices and MSME-scale production. More specifically, this study develops and evaluates a MobileNetV2-based model to automate Classification of Roasting Maturity Levels of Coffee Beans Using CNN Method Based on Mobilenetv2. Beyond academic contributions—such as testing lightweight architectures in the local coffee agro-industry—this research offers practical benefits: objective, consistent, and accessible tools for farmers, roasters, and small- to medium-scale coffee businesses to maintain product quality standards.

MATERIALS AND METHOD

This study applied a deep learning-based experimental approach to develop a classification system for the maturity level of roasting coffee beans. The model used is a Convolutional Neural Network (CNN) with the MobileNetV2 architecture, which was chosen for its computing efficiency and ability to extract visual features automatically. In general, the research flow includes the stages of data collection and structuring, dataset sharing, image preprocessing and augmentation, model development, gradual training, performance evaluation, and storage of the final model for implementation purposes.

The dataset used consists of four classes of maturity levels, namely green beans as the condition of raw coffee beans before the roasting process, light roast as a light roast level, medium roast as a medium roast level, and dark roast as dark roast level. All images are collected in PNG format and placed in separate directories based on their respective class labels. The data source comes from internet documentation that is appropriate for the research needs. To ensure uniformity of the input to the model, the entire image was resized to 224 × 224 pixels according to the MobileNetV2 input shape standard.

The dataset was divided using a train_test_split function with random seed 42 to maintain reproducibility. The proportion of data is divided into 75% training data, 15% validation data, and 10% test data. This division is done randomly but maintains a balance of class distributions so that the representation of each category remains consistent across each subset of data.

In the preprocessing stage, each image that is loaded through the process is resized to a standard size, converted into a numerical array, and then normalized using the preprocess_input function of MobileNetV2. This normalization is necessary to match the input scale to the data characteristics used in the pre-training of the ImageNet model. To improve generalization capabilities and reduce the risk of overfitting, image augmentation was carried out on the training data, including random rotation, length and width shifts, shear, zoom, horizontal flip, vertical flip, and brightness level variations. All augmentation techniques are applied only to training data, while validation and testing data are left unmodified to keep performance evaluation objective.

The classification model is built using the MobileNetV2 architecture that has been pre-trained on the ImageNet dataset. In the initial stage, include_top parameter is set to false so that only the feature extractor part is used. All parameters on the base model are locked so that they cannot be updated at the initial training stage. On top of the extractor feature, a series of classification layers in the form of GlobalAveragePooling2D, a Dropout layer, Dense measuring 128 neurons with ReLU activation, BatchNormalization, and a softmax-activated Dense output layer that produces four classes. This series serves to adapt the extraction capabilities of the MobileNetV2 feature to the needs of the roasting classification task.

The training process is carried out in two stages. The first stage is feature extraction, where the base model remains locked and only the classification layer is trained. At this stage, the learning rate of 1×10^{-4} and the number of epochs are 50. The training process is equipped with callbacks in the form of EarlyStopping, ReduceLROnPlateau, and ModelCheckpoint to improve training efficiency and stability. The second stage is a fine-tuning process by opening up about 25% of the last layer of MobileNetV2 so that most of the features can be adjusted to the image characteristics of the coffee beans. At this stage, the learning rate is lowered to 1×10^{-5} with a total of 25 epochs. Callback settings are maintained to maintain training quality and prevent overfitting.

Model evaluation was conducted using test data with consistent batch sizes. The model's performance was measured through a trend graph of accuracy and test loss, and further analyzed using a confusion matrix to see the prediction distribution in each class. In addition, a classification report containing precision values, recalls, and F1-scores is used as an additional indicator to assess the quality of model performance. The training was also analyzed to monitor stability and detect possible overfitting.

The final model is saved in .h5 format so that it can be used at the implementation stage. To test the model's performance in real scenarios, testing is carried out by uploading new images, running a pre-process process according to training procedures, and then predicting the roasting rate. The prediction results are displayed in the form of a class label along with a confidence score value to ensure the model's confidence level in the resulting decision.

RESULTS AND DISCUSSION

Model training is carried out through two main stages, namely feature extraction and fine-tuning. In the first stage, the entire base layer of the MobileNetV2 model is locked so that only the top classification layer is updated. This stage aims to stabilize the initial learning process by focusing adjustments on newly added layers. Over the course of 25 epochs, the accuracy of the training showed a consistent trend of improvement, indicating that the model managed to learn the basic features of the coffee bean imagery such as dominant colors, brightness levels, as well as general texture patterns that distinguish the roasting classes. The validation accuracy curve in the middle of the epoch began to show stability, while the validation loss gradually decreased, indicating that the learning process at this stage was optimal and did not show the initial symptoms of overfitting.

The second stage of training, fine-tuning, is done by opening up about 25% of the last layer of MobileNetV2 so that the model can refine a more specific representation of the features of the coffee bean's visual characteristics. At this stage, the learning rate is lowered so that the weight update takes place more carefully and does not damage the weight of the pre-training results. The results of the training showed a more significant increase in accuracy compared to the previous stage. Validation accuracy increased steadily to close to the maximum value in the final epoch, while training accuracy also increased without showing significant divergence. This indicates that fine-tuning has succeeded in deepening the model's understanding of micro-features such as burn intensity, color variations in the bean surface, and the distribution of brownish hues that distinguish between roasting levels. The EarlyStopping mechanism stops training when the validation loss no longer shows a decrease, so the final model used is the model with the best performance in the validation data.

Evaluation using a test set, which was never involved in the training or validation process, provided an objective picture of the model's generalization capabilities to new data. The test accuracy and loss values show that the model has reliable classification capabilities to distinguish the four levels of roasting doneness. Performance stability between training, validation and testing confirms that the model does not suffer from significant deviations or overfitting, so the results are reliable for practical applications as well as small and medium-sized industrial use.

The confusion matrix analysis provides a more detailed picture of the prediction distribution in each class. The green bean class is the most easily recognizable class because of its visual differences that are very contrasting compared to other classes. The bright green color and texture of the seeds that have not undergone the roasting process allow the model to achieve a very high level of precision in this class. The dark roast class is also relatively easy to recognize thanks to its brownish-black hue that attracts strong feature extraction characteristics. In contrast, the light roast and medium roast classes are the source of the most prediction errors. This is due to the proximity of the two classes, so variations in lighting and differences in image quality can affect the visual perception of the model. Despite this, the error rate is still within reasonable limits, and F1 grades show that the performance of each class remains in the good category.

The results of the classification report further strengthen the analysis. The green and dark roast classes obtain high precision and recall values due to their very clear visual characteristics. Meanwhile, the F1 scores for the light roast and medium roast classes were slightly lower due to the visual proximity between the two. Overall, the macro-average and weighted-average values show a balanced performance, so the model can be said to be able to classify all classes well and is not biased towards a particular class.

The visualization of the training curve provides additional insight into the stability of the model during the learning process. The pattern of the accuracy and loss curves between training and validation did not show a large difference, indicating that the model was able to maintain good generalizations. The implementation of ReduceLROnPlateau has proven to be effective in maintaining the stability of the optimization process when the model begins to reach the saturation point, so that the training process runs smoothly without extreme fluctuations. This success is also supported by data augmentation that expands the variation of imagery in the training set, so that the model not only memorizes specific patterns but is also able to recognize variations in new data.

Trials using the new imagery show that the model can provide consistent predictions with high confidence levels, generally above 80 percent. The prediction process includes image uploads, pre-processes in the form of resize and normalization, and inference using the final model. The predicted results show that the model is able to accurately recognize variations in texture, hue and burn intensity in the test image, so that it can be declared ready for use in real conditions. This performance reinforces that the model is feasible to be applied in applications to support roasting level assessment, both in the context of manual production, roasting laboratories, and camera-based automated systems.

Overall, this study shows that MobileNetV2 architecture is the right choice for the task of classifying the maturity level of coffee beans. Its lightweight architecture and fast inference time make it ideal for mobile-based systems as well as IoT devices. Data augmentation is shown to play an important role in improving the model's generalization of exposure variations and shooting angles. The fine-tuning process contributes significantly to the improvement in accuracy, showing that the partial adjustment of the base layer of the model is able to improve the model's ability to recognize detailed features in the coffee beans. The prediction errors that occur mostly come from classes with adjacent color ranges, but the overall accuracy rate remains high. With stable and reliable performance, this model has great potential to be implemented in roasting-level classifier applications in the MSME to commercial scale coffee industry.

Figure 1. Example of coffee bean dataset images based on four roasting levels: green bean, light roast, medium roast, dark roast.

Source: Research image dataset, 2025

An example of a coffee bean dataset image based on four levels of roast maturity, namely green bean, light roast, medium roast, and dark roast. Each class features variations in the color and texture of the coffee bean surface as a result of the difference in roasting levels. Green bean coffee beans have a pale green color as a characteristic of raw beans, light roast shows a light yellowish-brown color, medium roast is medium brown, and dark roast is dark brown to blackish. This variation in visual characteristics is the basis for model learning in carrying out the roasting level classification process.

Model: "functional_1"

Layer (type)	Output Shape	Param #
input_layer_3 (InputLayer)	(None, 224, 224, 3)	0
mobilenetv2_1.00_224 (Functional)	(None, 7, 7, 1280)	2,257,984
global_average_pooling2d_1 (GlobalAveragePooling2D)	(None, 1280)	0
dropout_2 (Dropout)	(None, 1280)	0
dense_2 (Dense)	(None, 128)	163,968
batch_normalization_1 (BatchNormalization)	(None, 128)	512
dropout_3 (Dropout)	(None, 128)	0
dense_3 (Dense)	(None, 4)	516

Total params: 2,422,980 (9.24 MB)
Trainable params: 164,740 (643.52 KB)
Non-trainable params: 2,258,240 (8.61 MB)

Figure 2. Architectural diagram of the MobileNetV2-based classification model.
Source: MobileNetV2 model output, 2025

A summary of the architecture of the MobileNetV2-based classification model used in this study. The model receives 224 × 224 pixels of input images with three color channels (RGB), then utilizes MobileNetV2 as a feature extractor. The output of the base model was processed using Global Average Pooling and several classification layers consisting of Dense, Batch Normalization, and Dropout before producing four classes of coffee bean roasting maturity level outputs. The total model parameters are 2,422,980, with most of the parameters on the base model being non-trainable in the early stages of training.

Figure 3. Training and validation accuracy graph of the model.
Source: Model training results, 2025

A comparison graph of training accuracy and validation accuracy during the training process and fine-tuning the MobileNetV2 model. The graph shows a rapid increase in accuracy in the early stages of training and then tends to stabilize until the end of the epoch. The proximity of the training accuracy and validation values showed that the model had good generalization capabilities and did not suffer from significant overfitting.

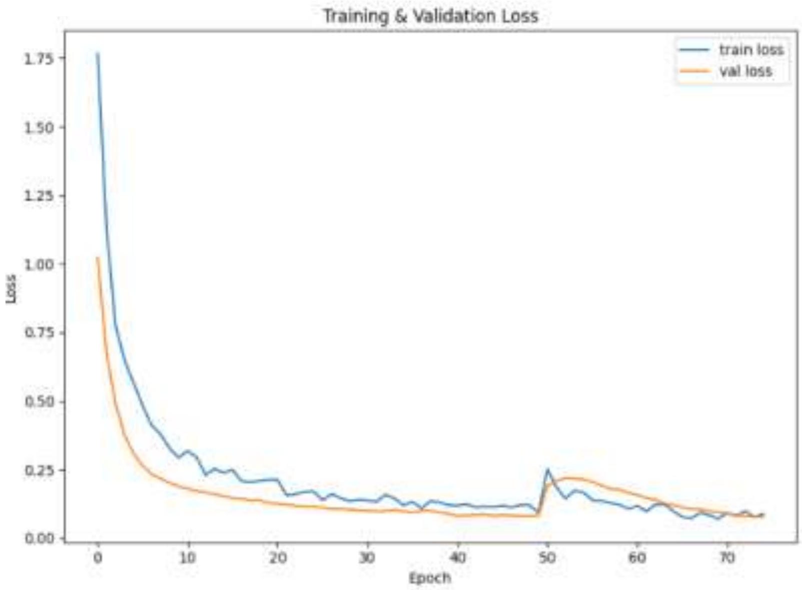

Figure 4. Training and validation loss graph of the model.
Source: Model training results, 2025

A comparison graph of training loss values and validation losses during the training process and fine-tuning the MobileNetV2 model. The chart shows a significant decrease in losses in the early stages of training and tends to stabilize in the following epoch. The proximity of the training loss and validation loss curves indicates that the learning process is stable and does not show significant symptoms of overfitting.

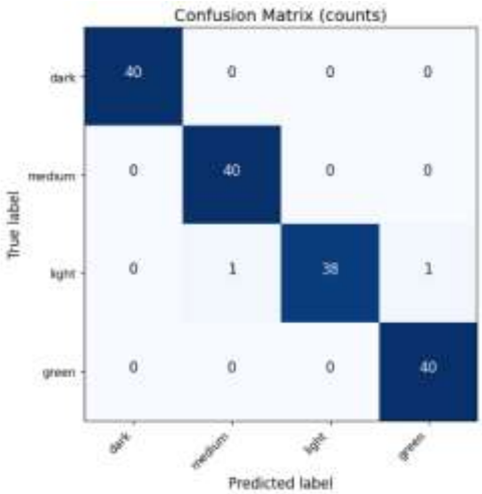

Figure 5. Confusion matrix of prediction counts on test data.
Source: Model evaluation results, 2025

The confusion matrix is based on the number of predictions (counts) on the test data for the classification of the maturity level of roasting coffee beans. The test results showed that the dark, medium, and green classes were all correctly classified with no prediction errors. Misclassification occurs only in light classes, where a small fraction of the sample is incorrectly predicted as medium and green. These findings indicate that the model has excellent classification accuracy, with errors limited to classes that have similarities in visual characteristics.

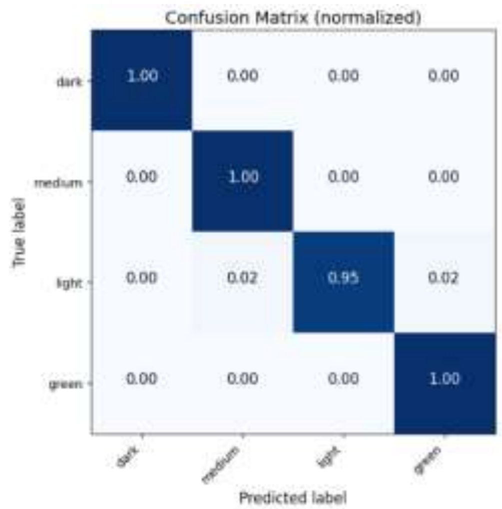

Figure 6. Normalized confusion matrix on test data.

Source: Model evaluation results, 2025

The confusion matrix was normalized by the test results of the MobileNetV2 model on the test data. The matrix shows the proportion of true and false predictions for each class of roasting coffee bean maturity level. Diagonal values close to 1.00 in the dark, medium, and green classes indicate a very high level of accuracy, while the light class has a small error rate with some false predictions to the medium and green classes. These results confirm that the model can classify most samples well, with the main error occurring in classes that have similar visual characteristics.

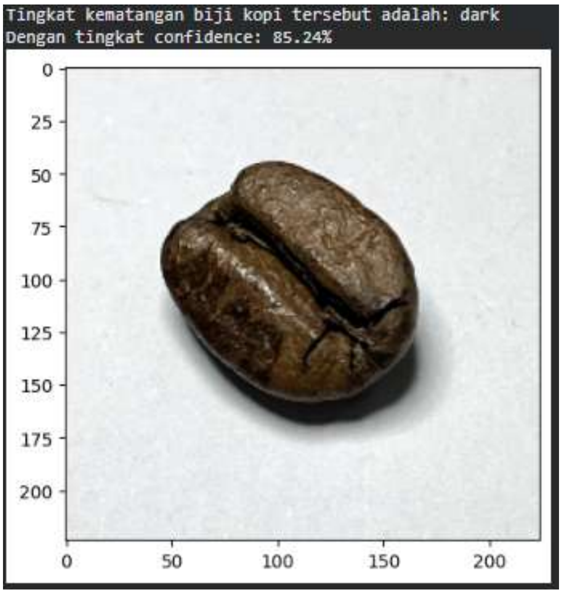

Figure 7. Prediction of new test image as dark roast (confidence 85.24%).
Source: Model prediction, 2025

An example of the prediction of the maturity level of roasting coffee beans using the MobileNetV2 model in the new test image. The model classifies the image as a dark roast with a confidence level of 85.24%. These results show that the model is able to provide consistent class predictions with high confidence scores, so it has the potential to be used in real-world application scenarios for automatic assessment of coffee bean roasting rates.

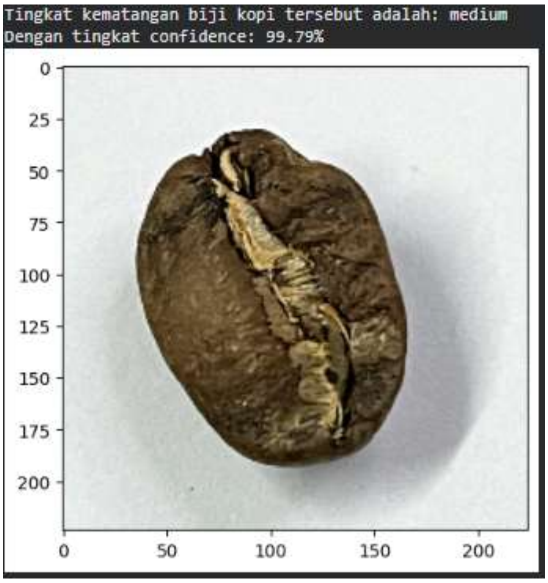

Figure 8. Prediction of new test image as medium roast (confidence 99.79%).
Source: Model prediction, 2025

An example of the prediction of the maturity level of roasting coffee beans using the MobileNetV2 model in the new test image. The model classifies the image as a medium roast with a confidence level of 99.79%. These results show that the model can identify the visual characteristics of coffee beans very accurately and provide a high confidence score value in the classification process.

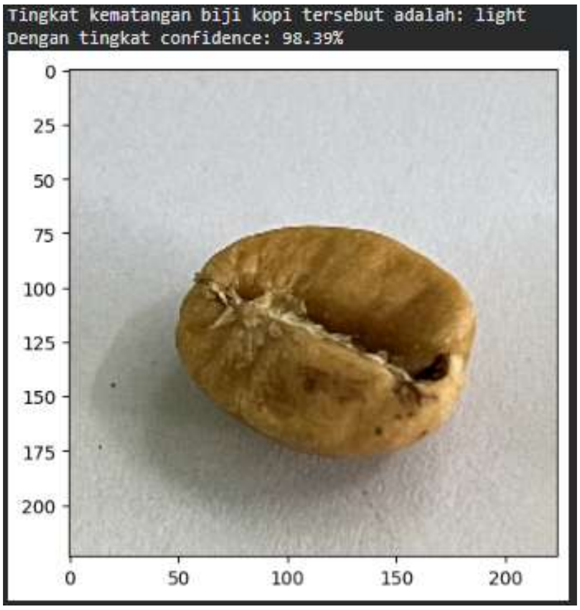

Figure 9. Prediction of new test image as light roast (confidence 98.39%).
Source: Model prediction, 2025

An example of the prediction of the maturity level of roasting coffee beans using the MobileNetV2 model in the new test image. The model classifies the image as a light roast with a confidence level of 98.39%. These results show that the model can accurately distinguish the visual characteristics of coffee beans with light roasting levels and provide a high confidence score.

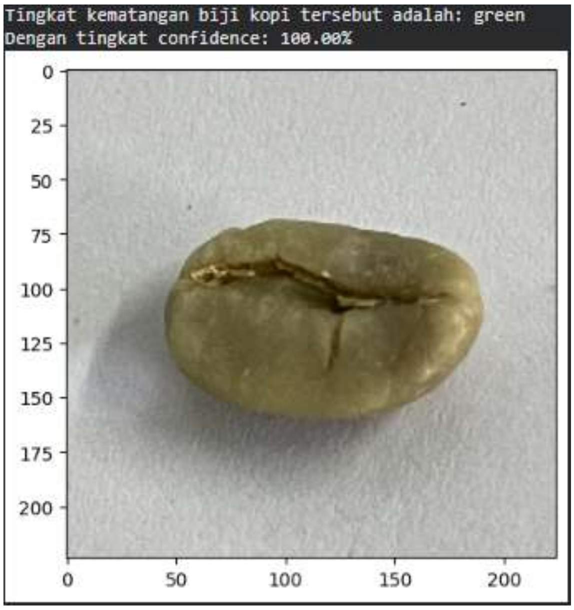

Figure 10. Prediction of new test image as green bean (confidence 100.00%).
Source: Model prediction, 2025

An example of the prediction of the maturity level of coffee beans using the

MobileNetV2 model in the new test image. The model classifies the image as a green bean with a confidence level of 100.00%. These results show that differences in the visual characteristics of raw coffee beans can be recognized very well by the model, thus resulting in predictions with a very high degree of confidence.

CONCLUSIONS

This study successfully developed a MobileNetV2-based classification model for roasting maturity levels in coffee beans, achieving 96% test accuracy alongside a lightweight design and efficient computation time, making it suitable for mobile or IoT deployment in coffee industry quality control. For future research, expanding the dataset to include diverse bean types, lighting conditions, and backgrounds would enhance robustness and generalization; additionally, testing other lightweight architectures like EfficientNet-Lite or custom CNNs, integrating explainable AI for better interpretability, and creating user-friendly mobile/web apps with IoT and cloud integration could drive broader practical adoption among farmers, roasters, and small-to-medium coffee businesses.

REFERENCES

Adriyanto, R., & Kurniawati, N. (2023). Klasifikasi jenis biji kopi menggunakan VGG16 dan MobileNetV2. Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), 7(1), 120–128.

Anggraini, R., Sari, D. P., & Putra, A. R. (2021). Analisis kualitas biji kopi menggunakan metode pengolahan citra berdasarkan warna dan tekstur. Jurnal Teknologi Informasi dan Ilmu Komputer, 8(3), 451–458.

Ardiansyah, R., Saputra, D., & Lestari, S. (2021). Identifikasi kerusakan mutu biji kopi arabika menggunakan fitur warna dan tekstur. Jurnal Sistem Informasi dan Teknologi, 5(3), 211–218.

Bouramdane, A.-A. (2023). Cyberattacks in smart grids: challenges and solving the multi-criteria decision-making for cybersecurity options, including ones that incorporate artificial intelligence, using an analytical hierarchy process. *Journal of Cybersecurity and Privacy*, *3*(4), 662–705.

Girma, B., & Sualeh, A. (2022). A review of coffee processing methods and their influence on aroma. *Int. J. Food Eng. Technol*, *6*(7), 7–16.

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.

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

Li, T., Yu, X., Zu, W., Cai, Y., & Hong, W. (2024). Application and research progress of near infrared spectroscopy for multidimensional assessment of coffee beans. *Journal of Near Infrared Spectroscopy*, *32*(6), 179–191.

Lu, T., Sun, Y., Huang, Y., & Chen, X. (2023). Effects of roasting on the chemical compositions, color, aroma, microstructure, and the kinetics of changes in coffee pulp. *Journal of Food Science*, *88*(4), 1430–1444.

Motta, I. V. C., Vuillerme, N., Pham, H.-H., & de Figueiredo, F. A. P. (2024). Machine learning techniques for coffee classification: a comprehensive review of scientific research. *Artificial Intelligence Review*, *58*(1), 15.

Nastoska, A., Jancheska, B., Rizinski, M., & Trajanov, D. (2025). Evaluating trustworthiness in AI: Risks, metrics, and applications across industries. *Electronics*, *14*(13), 2717.

Rahmatullah, M., & Prasetyo, E. (2022). Klasifikasi citra biji kopi Temanggung menggunakan residual network (ResNet-50). Jurnal Informatika dan Sistem Informasi, 8(2), 95–104.

Rusinek, R., Dobrzański Jr, B., Gawrysiak-Witulska, M., Siger, A., Żytek, A., Karami, H., Umar, A., Lipa, T., & Gancarz, M. (2024). Effect of the roasting level on the content of bioactive and aromatic compounds in Arabica coffee beans. *International Agrophysics*, *38*(1), 31–42.

Sangpimpa, W., Srichan, P., Jannu, T., Renaldi, G., Setthaya, P., & Samakradhamrongthai, R. S. (2025). Influence of roasting technique on physicochemical properties, volatile compounds, and flavor profile of robusta (Coffea canephora) coffee beans. *European Food Research and Technology*, *251*(11), 3575–3589.

Santoso, I., Yuanita, E. A., & Karomah, R. S. (2024). Application of digital image processing method for roasted coffee bean quality identification: a systematic literature review. *African Journal of Food, Agriculture, Nutrition and Development*, *24*(1), 25264–25287.

Setiawan, A., Nugroho, Y. S., & Wicaksono, A. (2022). Klasifikasi kematangan biji kopi menggunakan convolutional neural network (CNN). Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, 6(5), 2301–2308.

Tarigan, E. B., Wardiana, E., Hilmi, Y. S., & Komarudin, N. A. (2022). The changes in chemical properties of coffee during roasting: A review. *IOP Conference Series: Earth and Environmental Science*, *974*(1), 12115.

Wang, C., Yang, Z., Li, Z. S., Damian, D., & Lo, D. (2024). Quality assurance for artificial intelligence: A study of industrial concerns, challenges and best practices. *ArXiv Preprint ArXiv:2402.16391*.

Wibowo, A., Santoso, B., & Hidayat, R. (2022). Klasifikasi kualitas biji kopi menggunakan multilayer perceptron berbasis fitur warna LCH. Jurnal Ilmiah Komputer dan Informatika, 10(2), 67–74.