

Comparative Evaluation of Convolutional Neural Network Full Learning Model with Transfer Learning (VGG-16) for Coffee Bean Roasting Level Classification

Mradipta Nindya Tama¹, Amanat Bintang Saptomo², Afrido³,
Dawi Karomati Baroroh⁴, Achmad Pratama Rifai⁵, Nguyen Huu Tho⁶

^{1,2,3,4,5}Department of Mechanical and Industrial Engineering, Universitas Gadjah Mada, Yogyakarta, Indonesia

⁶Faculty of Engineering and Technology, Nguyen Tat Thanh University, Ho Chi Minh City, Vietnam

Article Info

Article history:

Received Jan 01, 2025

Revised Apr 28, 2025

Accepted Jul 12, 2025

Keywords:

Machine Learning

CNN

Roasting

Classification

VGG-16

ABSTRACT

Indonesia is the 3rd largest coffee producing country in the world in 2022-2023 with coffee production reaching 11.85 million bags per 60 kg of coffee. One of the important processes in coffee production is roasting because the roasting level of coffee beans can affect the taste and aroma of coffee. The problem faced is that the process of assessing the level of coffee roasting is traditionally carried out through visual observation by an expert (roaster). This method produces a subjective level of assessment and requires high skills and experience, making the assessment of the level of coffee roasting less efficient and prone to human error. Therefore, in this study the author aims to develop a Convolutional Neural Network (CNN) model for the classification of the level of coffee bean roasting that can achieve better and faster accuracy. In this study, the author compared two CNN architecture approaches for the classification of the level of coffee bean roasting. The first approach is full learning with an architecture consisting of three convolution layers. The second approach is transfer learning based on the VGG-16 model. From the results of the analysis, it is known that the full learning model has a better level of accuracy and a faster running time than the VGG-16 transfer learning. The CNN full learning model for coffee bean roasting level classification is able to classify the coffee bean roasting level, with an accuracy of 98.75% and a running time of 856 ms per step. The application of CNN for coffee roasting level classification can provide benefits such as improving quality control and reducing the level of subjectivity of a roaster in assessing the roasting level of coffee beans.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Achmad Pratama Rifai,
Department of Mechanical and Industrial Engineering,
Universitas Gadjah Mada
Jalan Grafika No. 2, Yogyakarta 55284, Indonesia.
Email: achmad.p.rifai@ugm.ac.id

1. INTRODUCTION

The coffee shop industry in Indonesia has experienced quite rapid development in recent years. According to data released by the United States Department of Agriculture (USDA), global coffee production reached 170 million bags per 60 kg of coffee in the 2022/2023 period. Of that amount, Indonesia managed to produce 11.85 million bags of coffee and became the 3rd largest coffee producing country in the world in 2022/2023 [18]. One of the important processes in coffee production is roasting because it can affect the taste and aroma of coffee. Roasting is the process of

heating coffee beans until they reach a certain level of ripeness. During the roasting process, regular checks are carried out on the temperature, color, and aroma of the coffee by an experienced roaster.

The traditional coffee roasting level assessment process is carried out through visual observation by an expert (roaster). This method produces a subjective level of assessment and requires high skills and experience, making the assessment of the coffee roasting level less efficient and prone to human error. In addition to the problem of high subjectivity, the speed of a roaster in assessing the level of coffee roasting is also a problem that must be resolved immediately. Along with technological advances, the use of machine learning in the coffee industry, especially in roasting level classification, has opened up new opportunities to improve accuracy and efficiency.

One method that can be used to help classify the level of coffee ripeness in the roasting process is deep learning technology, namely Convolutional Neural Network (CNN) which can be used to analyze and classify images. The application of CNN to analyze and classify images has proven successful in several domains, one of which is the classification of plant diseases [19]. Significant progress in the field of CNN is marked by the emergence of various transfer learning architectures such as Residual network (ResNet) 50, Inception V3, GoogleNet, AlexNet, and VGG-16. Although these models have proven effective, a full learning approach that allows customization of the number of convolutional layers can also be considered considering the amount of data and expected classes is not too large. Research conducted by Ontoum et al., (2022) on Coffee Roast Intelligence created a coffee bean roasting level classification model using a transfer learning model, namely mobilenet. The model has produced an accuracy level of 82%.

Therefore, in this study the author aims to develop a CNN model that can achieve better accuracy for coffee bean roasting level classification. In this study, the author compares two CNN architecture approaches for coffee bean roasting level classification. The first approach is full learning with an architecture consisting of three convolution layers. The second approach is transfer learning based on the VGG-16 model. The output of this study is expected to be able to develop a more accurate and faster classification model in determining the level of ripeness of coffee beans, thus providing objective recommendations to roasters in the decision-making process.

2. LITERATURE REVIEW

The coffee bean roasting process is an important stage in coffee production that affects its chemical and organoleptic properties. The coffee roasting process is the activity of roasting raw coffee beans (green beans) to become ripe coffee beans (roast beans) to create a unique distinctive taste. In this process, there are several important factors that need to be considered, such as roasting technique, roasting temperature, roasting duration, moisture level of raw coffee beans, amount of coffee beans roasted and cooling time [10]. The level of ripeness of ripe coffee beans (roast beans) is divided into three main categories: light, medium, and dark. Light roast is the roasting stage with the lowest level of ripeness. The characteristics of this level of ripeness are the color of the coffee beans which are still bright, the taste tends to be sweet, sour, and the aroma like nuts, with a high caffeine content. Medium roast has a darker color than light roast. The medium roasting level has a balanced taste characteristic between sour and sweet, and a lower caffeine content than light roast. Meanwhile, dark roast is the roasting stage with the highest level of ripeness. Dark roast coffee beans have the darkest color, tend to be bitter, and are often accompanied by a chocolate-like aroma.

Based on research conducted by Sagita Diang (2024), the use of Internet of Things (IoT)-based spectral analysis can provide a more objective, efficient, and non-invasive assessment of the quality of roasted coffee beans. By utilizing spectral data from multiple channels, this technique allows for a more accurate classification of roasting levels, which in turn supports a more controlled roasting process. Roasting triggers changes in the chemical content of coffee beans, which affects the color, taste, and aroma formed, with the final result being highly dependent on parameters such as temperature and roasting time.

Freitas V.V et al., (2024) in their research explained that coffee quality is influenced by a number of factors, such as the environment, harvesting methods, and post-harvest processing processes, including fermentation and roasting. All of these factors contribute to the formation of

the flavor profile, aroma, and chemical content of coffee. Freitas also emphasized the importance of a holistic approach to coffee production to maximize its quality. Meanwhile, research by Sevenrin et al., (2024) discusses changes in the mechanical properties and elasticity of coffee beans during the roasting process, which are influenced by changes in temperature and the cellular structure of the coffee beans. This study provides deeper insight into how roasting control can be done by measuring the elasticity of coffee beans, in addition to relying on visual observations of color changes made by the roaster.

Research by Muhammad Zainun Na'im et al., (2024) shows that roasting duration affects the organoleptic properties of Excelsa coffee, such as taste, aroma, and color. The results of their study indicate that roasting for 15 minutes at a temperature of 200°C produces the best quality. This shows how important timing is in the roasting process to obtain an optimal flavor profile. Other research by Almayda et al., (2024) also revealed that differences in roasting methods, both conventional and modern, can affect the chemical content of Liberica Coffee, such as water content, ash, pH, and caffeine content. This shows that the technique used can have an impact on the quality and chemical composition of coffee beans. The results of previous studies show the complexity of the process needed to produce quality coffee. One of the difficult parts of coffee processing is determining the level of ripeness.

In a study conducted by Anu et al., (2024). Where the roasting level classification process for coffee beans in micro, small and medium enterprises (MSMEs) uses a manual sorting technique by taking samples of roasted coffee beans. The manual roasting technique has a weakness, namely that this manual method takes a long time to select coffee beans based on the level of ripeness, so that coffee roasters often lose focus, which results in mixing coffee beans in inappropriate packaging. To overcome this problem, an artificial neural network-based system is used to assist the classification process. In this study, 60 coffee beans with three roasting levels (light roasting, medium roasting, and dark roasting) were used. The classification process involves two stages: training and testing, with 30 coffee bean samples in the testing stage. So that the results of the study showed a training accuracy of 90%, while the testing accuracy reached 66.67%.

Convolutional Neural Network (CNN) has the potential to be applied in various fields, one of which is in the coffee processing industry, especially in the coffee bean roasting process. Monitoring and supervising the level of ripeness of coffee beans during the roasting process is important to achieve the best taste. CNN can be used to recognize visual patterns in roasted coffee beans, in the form of changes in color and texture as important indicators of the ripeness of the coffee beans. By applying a CNN-based system, this technology can help identify stages in the roasting process and provide recommendations or warnings to operators to ensure that the coffee beans are roasted perfectly automatically.

Convolutional Neural Network is a deep learning method with the ability to recognize and detect objects in digital images. The ability of Convolutional Neural Network is claimed to be the best method for detecting and recognizing objects, but like other deep learning methods, it has weaknesses in the training process which takes a long time [6]. The CNN model is an artificial neural network architecture used to analyze and recognize patterns in images. The process begins with a convolution layer that extracts important features such as edges, textures, and shapes through filters. After that, the data is processed through a pooling layer to reduce dimensions and retain key features. Finally, the processed data is passed to the fully connected layer to perform the final classification [11]. CNN acts as an effective feature extractor in recognizing objects in images and has proven to be very effective in various applications. CNN capabilities have been effectively applied in various applications, including fresh fruit production, medical diagnostics, and facial skin identification, all of which utilize object detection and pattern recognition for efficiency and accuracy.

Research by Wang et al., (2022), discusses the application of CNN technology in fresh fruit production. With the increasing need for automation in the agricultural sector, this technology accelerates mechanization and increases efficiency. Meanwhile, Tripathi (2021) tested various CNN architectures to improve fruit recognition in the industry. With the DenseNet model showing 100% accuracy of the results, it provides a practical solution in industries that require high

efficiency. Meanwhile, research by Reghunandan (2024), using a CNN model to detect and count reticulocytes in human blood smears showed an accuracy of more than 90%, making it a potentially effective diagnostic tool.

On the other hand, Nissa et al., (2021) explored the use of CNN in classifying women's facial skin types, demonstrating the effectiveness of CNN in assisting the selection of appropriate skincare products. Finally, a study by Li et al., (2024) discussed the use of CNN in a computer vision system for livestock management, which not only improves production efficiency but also supports animal welfare with early disease detection. All of these studies demonstrate the broad application of CNN in various sectors, from agriculture to health, and confirm the potential of this technology in improving detection efficiency and accuracy in various fields. With this evidence, CNN also has the potential to be used in the coffee production process, especially in the roasting results checking process. That way, the checking process will be easier, and the roasting results obtained will be more consistent.

Significant progress in the field of CNN is marked by the emergence of various transfer learning architectures such as Residual network (ResNet) 50, Inception V3, GoogleNet, AlexNet, and VGG-16 Aquil (2021). In their research, Brahmani et al., (2018) evaluated the performance of various modern CNN architectures such as VGG-13, SqueezeNet, ResNet, Inception, DenseNet, and AlexNet for plant disease classification. This study compared three main approaches: full learning or developing and training a model from scratch and two transfer learning models. The dataset used was divided into train data (80%) and test data (20%). The result is that the transfer learning model can produce a better level of accuracy than full learning with an accuracy level of 99.76%. Wang et al., (2021) proposed the application of deep learning (DL) for disease severity estimation. The results showed that the VGG-16 architecture trained with transfer learning achieved the highest accuracy of 90.4%. Furthermore, in the study of Wang et al., (2022) combined transfer learning and CNN for plant disease image classification, The result is that the model is able to produce an accuracy level of 90.84%. In a study conducted by Anvesh et al., (2024), a comparative analysis of the model between custom Deep CNN using 2 convolution layers with VGG-16 to detect anomalies was conducted. The result is that custom CNN shows better performance compared to VGG-16. Therefore, in this study, the author intends to develop a CNN model for classifying the roasting level of coffee beans and to compare models between full learning and transfer learning (VGG-16) to determine the best model for classifying the roasting level of coffee beans.

3. RESEARCH METHOD

In this study, a CNN model will be developed that can achieve high accuracy in classifying the roasting level of coffee beans. In addition, this study will also compare two CNN architecture approaches for classifying the roasting level of coffee beans. The first approach is full learning with an architecture consisting of three convolution layers. The second approach is transfer learning based on the VGG-16 model. The flow of this research is shown in Figure 1.

1. Collecting Data

The dataset used in this study is real secondary data obtained from public datasets. This dataset comes from coffee bean documentation in Thailand. The dataset consists of 1,600 images divided into 2 parts, namely 1,200 train data images and 400 test data images. The images in each part are grouped based on the coffee bean roasting level, namely dark, green, light medium.

The dataset used in this study is balanced across the different roasting levels, ensuring that no category is underrepresented. Each roasting level—Dark, Green, Light, and Medium—contains 300 images for training and 100 images for testing, maintaining an equal distribution of samples across all classes. This balanced dataset minimizes the risk of model bias toward any particular class and ensures that the classification performance is not skewed due to an imbalance in data representation. As a result, the model's ability to generalize across different roasting levels is enhanced, leading to more reliable and consistent classification results.

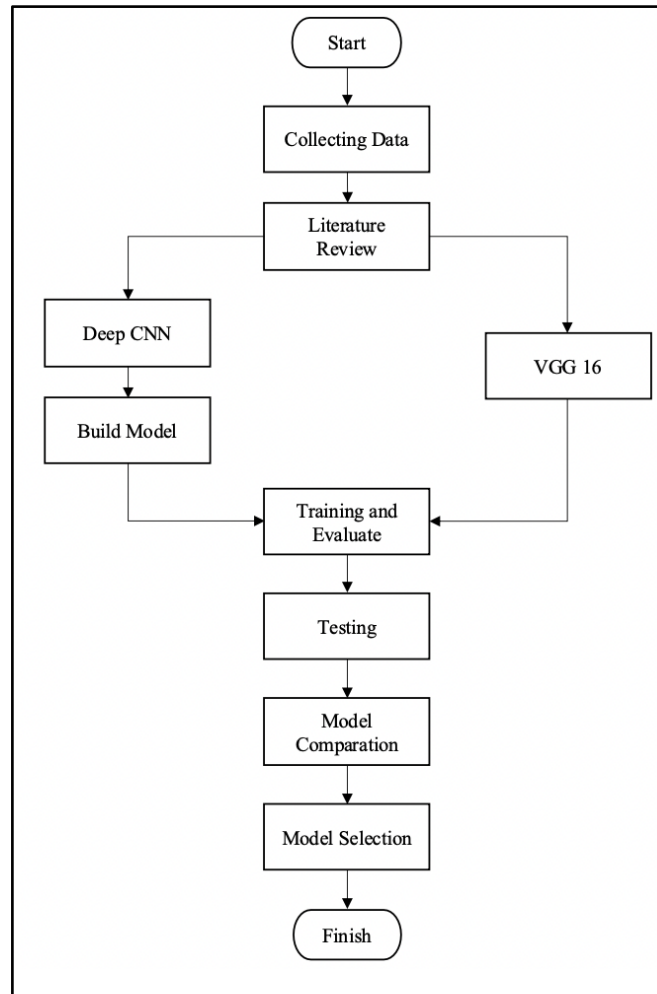


Figure 1. Research Flow

2. Literature Review

In this study, a literature review was conducted to determine the CNN transfer learning model that will be used as a comparison for the CNN full learning model. From the results of the literature review, it is known that there are differences in model performance results where research conducted by Wang et al., (2021) showed that VGG-16 has very good performance reaching 90.4%. Meanwhile, research conducted by Anvesh et al. (2024) who conducted a model comparison between custom Deep CNN using 2 convolution layers with VGG-16 showed that CNN showed better performance compared to VGG-16. Therefore, in this study, the author will compare two CNN architecture approaches for classifying coffee bean roasting levels. The first approach is full learning with an architecture consisting of three convolution layers. The second approach is transfer learning based on the VGG-16 model.

a. Transfer Learning

The VGG-16 model structure consists of 13 convolutional layers and 3 fully connected layers. The visualization of VGG-16 architecture is presented in Figure 2.

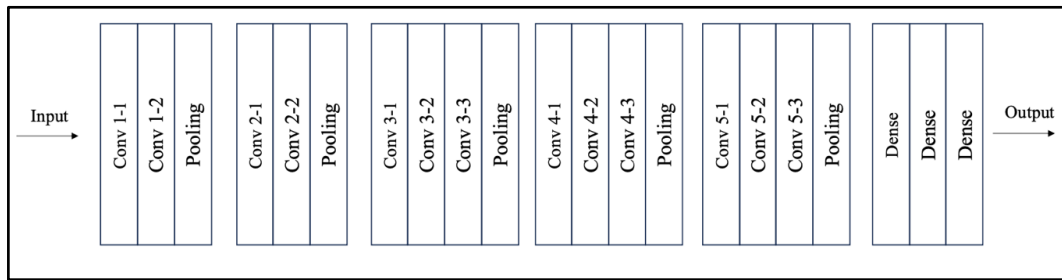


Figure 2. VGG-16 model architecture

b. Full Learning

The full learning CNN model used in this study uses 3 convolutional layers and 2 fully connected layers. The epoch used for training and validation is 5 iterations. The full learning CNN architecture used is presented in Figure 3.

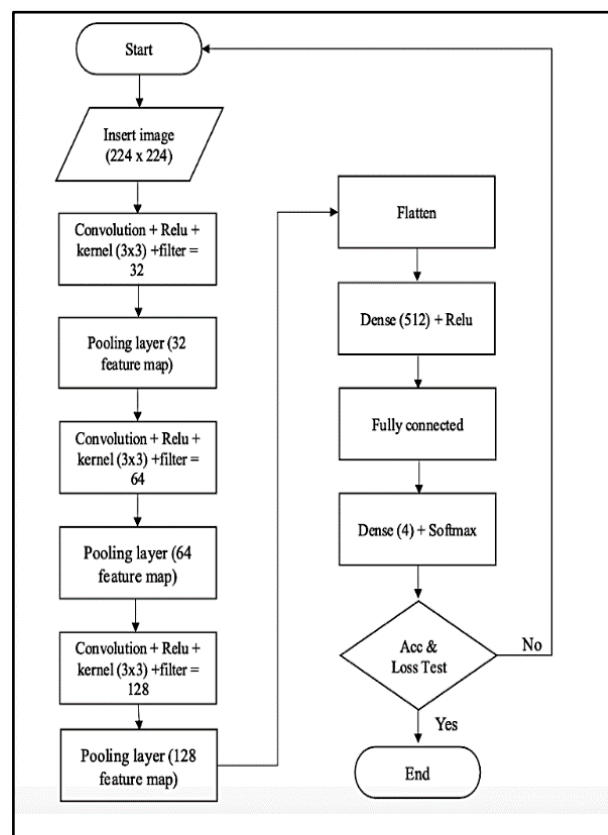


Figure 3. Full learning CNN model architecture

3. Training and Evaluate

During the training process, model parameters are adjusted by repeating the training set repeatedly over several periods. To maximize model performance, hyperparameters such as learning rate, batch size, and number of epochs are carefully adjusted. Then the level of accuracy and loss both during training and validation are analyzed to assess the model. In the training and evaluation process for both the VGG-16 and CNN full learning models, 5 iterations are carried out.

4. Testing

After the training process is complete, the model's performance in identifying anomalies is evaluated in depth using different data, namely images in the test data. The assessment is carried

out by looking at the level of accuracy, testing and validation graphs, running time, and confusion matrix for each model

5. Model Comparison

The performance of both models, both full learning and VGG-16, will be compared based on their accuracy level parameters. In addition, both models will be compared with the accuracy level of the CNN model from previous researchers [17].

6. Model Selection

The selection of the model is determined based on the highest level of accuracy, and the faster running time in classifying the level of roasting maturity of coffee beans.

4. RESULTS AND DISCUSSION

4.1. Results of full learning model

The Full Learning CNN architecture model is designed from scratch without using pretrained weights. This causes all model parameters to be trained from scratch. Using this approach will provide high flexibility but requires more data to obtain optimal training. This model is lighter than pretrained models such as VGG-16. Figure 4 presents the training progress of full learning model, while Figure 5 details the training results of each epoch.

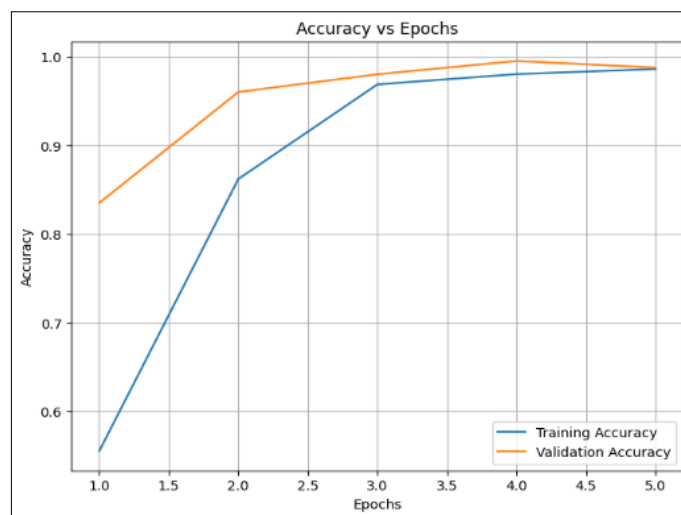


Figure 4. Training progress of full learning model

Epoch 1/5									
38/38	-	506s	-	13s/step	-	accuracy: 0.5554	-	loss: 1.3211	-
								val_accuracy: 0.8350	-
								val_loss: 0.4089	
Epoch 2/5									
38/38	-	5s	-	139ms/step	-	accuracy: 0.8620	-	loss: 0.4329	-
								val_accuracy: 0.9600	-
								val_loss: 0.1391	
Epoch 3/5									
38/38	-	10s	-	272ms/step	-	accuracy: 0.9686	-	loss: 0.0934	-
								val_accuracy: 0.9800	-
								val_loss: 0.0520	
Epoch 4/5									
38/38	-	7s	-	172ms/step	-	accuracy: 0.9802	-	loss: 0.0514	-
								val_accuracy: 0.9950	-
								val_loss: 0.0252	
Epoch 5/5									
38/38	-	9s	-	234ms/step	-	accuracy: 0.9860	-	loss: 0.0434	-
								val_accuracy: 0.9875	-
								val_loss: 0.0251	

Figure 5. Epoch results of the full learning model

The Accuracy vs Epochs graph shows the training capability of the Full Learning Deep CNN model over five epochs. In the first epoch, the training accuracy was only at 55.54%, while the validation accuracy was already high at 83.50%. Finally, in the fifth epoch, the training and validation accuracies reached 98.60% and 98.75%, respectively. The consistency between training

and validation accuracies shows that the model does not experience overfitting and is able to generalize well on the validation data.

The decrease in loss during training further clarifies how efficient this Full Learning Deep CNN model is. In the first epoch, the training loss was recorded at 1.3211, while the validation loss was at 0.4089. Until the fifth epoch, the training loss decreased to 0.0434, while the validation loss decreased to 0.0251. This graph shows a healthy trend, where both training and validation losses decrease consistently, indicating an increase in model performance without losing generalization ability.

Additionally, the accuracy vs. epoch graph showed early convergence, with stable training and validation performance, indicating that the model had already learned the key features necessary for classification. Extending the training further did not yield additional benefits and only increased computational time without meaningful accuracy gains. Therefore, five epochs were deemed sufficient to achieve optimal model performance.

In addition, the model training time per epoch decreases after the first epoch. The Accuracy vs Epochs graph confirms the stability and high performance of the Full Learning Deep CNN model. In a short time (around 3-4 epochs), the model was able to achieve convergence with stable validation accuracy. This shows that the Full Learning Deep CNN model is able to learn quickly and work optimally in classifying data.

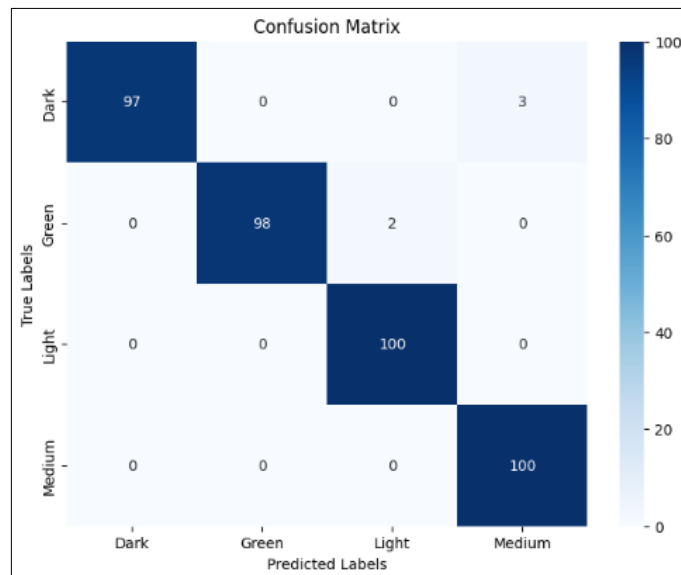


Figure 6. Confusion matrix of full learning CNN model

Afterward, the model is evaluated using test datasets. From the evaluation results with confusion matrix in Figure 6 shows that the Full Learning Deep CNN model has very good classification ability. In the Dark class, there are 97 samples that are correctly classified, while 3 samples are misclassified as Medium. For the Green class, 98 samples are correctly classified, and only 2 samples are misclassified as Light. While the Light and Medium classes are perfectly classified. This confusion matrix explains the high accuracy of the Full Learning Deep CNN model with all classes with only 3 total misclassifications. From this figure, it can be concluded that no systematic error patterns were found. This shows that the model is able to distinguish classes well.

```
13/13 - 1s - 93ms/step - accuracy: 0.9875 - loss: 0.0251
Test Accuracy: 98.75%
```

Figure 7. Accuracy test of full learning CNN model

Further, Figure 7 presents the detail information of the test result. In the accuracy evaluation, the Full Learning Deep CNN model achieved a Test Accuracy of 98.75% with a Test Loss of 0.0251.

4.2. Results of VGG-16 transfer learning model

Meanwhile, the VGG-16 model uses a transfer learning approach. This model utilizes weights that have been previously trained on large datasets such as ImageNet. This approach should make VGG-16 more effective, especially for small datasets, since the model already has a basic understanding of common visual features. The architecture of VGG-16 is deeper than other models, with many convolutional layers capable of capturing more complex feature representations of images. Figure 8 presents the training progress of full learning model, while Figure 9 details the training results of each epoch.

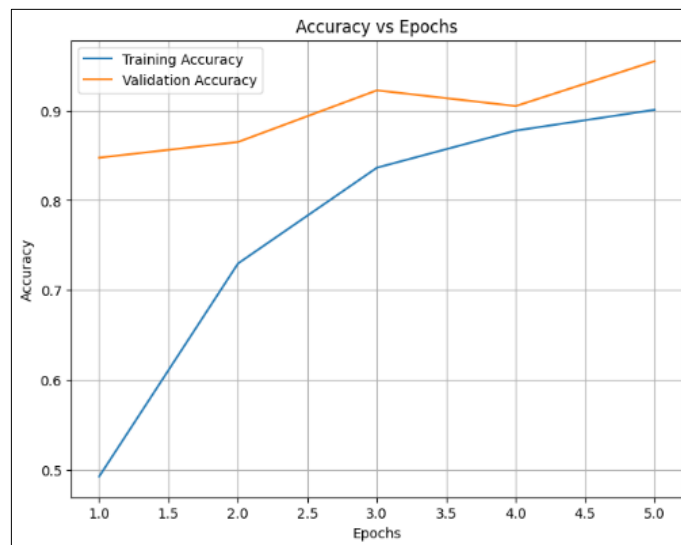


Figure 8. Training progress of the VGG-16 model

Epoch 1/5
38/38 - 1073s - 28s/step - accuracy: 0.4926 - loss: 1.1424 - val_accuracy: 0.8475 - val_loss: 0.6342
Epoch 2/5
38/38 - 1041s - 27s/step - accuracy: 0.7298 - loss: 0.6846 - val_accuracy: 0.8650 - val_loss: 0.4023
Epoch 3/5
38/38 - 1007s - 26s/step - accuracy: 0.8364 - loss: 0.4499 - val_accuracy: 0.9225 - val_loss: 0.2557
Epoch 4/5
38/38 - 1010s - 27s/step - accuracy: 0.8777 - loss: 0.3538 - val_accuracy: 0.9050 - val_loss: 0.2277
Epoch 5/5
38/38 - 1050s - 28s/step - accuracy: 0.9008 - loss: 0.3116 - val_accuracy: 0.9550 - val_loss: 0.1633

Figure 9. Epoch results of VGG-16 model

The Accuracy vs Epochs graph shows a significant increase in training accuracy on the VGG-16 model from 49.26% in the first epoch to 90.08% in the fifth epoch. The validation accuracy also shows an increase from 84.75% in the first epoch to 95.50% in the fifth epoch. In this case, the validation accuracy is higher than the training accuracy on all epochs in the VGG-16 model which may indicate that data augmentation is effective in helping the model perform better generalization.

The training loss also decreases over time. In the first epoch of the VGG-16 model, the training loss was 1.1424, and the validation loss was 0.6342. In the fifth epoch, the training loss decreased to 0.3116, and the validation loss decreased more significantly to 0.1633, indicating that the model continues to improve its performance on the validation data.

The training time per epoch is relatively stable due to the frozen layers and the use of a predetermined architecture, as well as data augmentation. This allows the model to continue to function efficiently even though more layers cannot be trained.

The conclusion from the graph and training results in Figure 9 shows that the VGG-16 model shows no signs of overfitting. The validation accuracy is higher than the training accuracy in the first few epochs, indicating that the regularization techniques used, such as frozen layers and data augmentation, have successfully helped the model to generalize better. However, VGG-16 tends to have more errors in certain classes such as Light, indicating that it has difficulty in distinguishing classes that have similar visual characteristics.

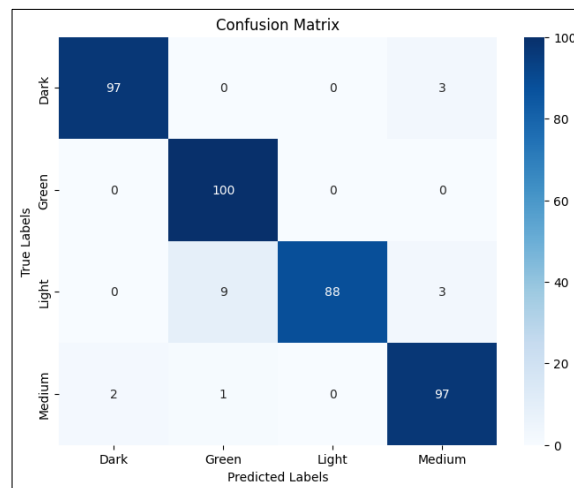


Figure 10. Confusion matrix of VGG-16 model

The evaluation results using the confusion matrix show how VGG-16 has a fairly good classification performance, although there are some errors. For the Dark class, there are 97 correct samples, while 3 samples are misclassified as Medium. The Green class is perfectly classified (100 correct), and the Light class has some errors and there are only 88 correct samples, where 9 samples are misclassified as Green and 3 samples as Medium. The Medium class shows good results like the Dark class, with 97 correct samples, and prediction errors occur in 2 samples classified as Dark and 1 as Green. Overall, the Green class has the highest accuracy, while the Light class shows more errors, indicating that the visual characteristics of this class are similar to other classes.

13/13 - 249s - 19s/step - accuracy: 0.9550 - loss: 0.1633
Test Accuracy: 95.50%

Figure 11. Test Accuracy model VGG-16

Further, Figure 11 presents the detail information of the test result. In the accuracy evaluation, the VGG-16 model obtained a Test Accuracy of 95.50% with a Test Loss of 0.1633.

4.3. Results of VGG-16 transfer learning model

Model comparison is done by comparing 3 models including full learning, transfer learning (VGG-16) and Mobilenet [17]. The parameter used as a comparison is the level of accuracy. Table 1 presents the is comparison table of the accuracy levels of the three models.

Table 1. Comparison of Accuracy		
No	Model	Accuracy (%)
1	Full learning CNN with 3 Convolution Layer	98,75 %
2	VGG-16	95,5 %
3	Mobilenet (Ontoum et al., 2022)	82 %

From the trials of the three models, it was found that the Full Learning Deep CNN model excels in terms of accuracy, time efficiency, and model stability on this dataset. The VGG-16 model, although slightly lower in accuracy, can also be relied on for small datasets due to transfer learning, but requires further optimization.

Therefore, the Full Learning Deep CNN model will be used further in detecting coffee bean images. The model was further tested by providing images outside the dataset using images of coffee beans that have been roasted to Medium. Figure 12 presents an example of full learning model trial using image taken from the researcher's documentation.



Figure 12. Full learning CNN model trial

Based on the prediction results, the model recognizes this coffee bean image as a “Medium” class, which is in accordance with its visual characteristics. This prediction is consistent with the visual characteristics of the coffee beans in the image which have a medium brown color, reflecting the characteristics of a medium roast. The prediction process is very fast, which is only 23 ms per step. This shows high efficiency in image processing. Overall, the model shows relevant and accurate capabilities for the task of classifying coffee beans based on images.

5. CONCLUSION AND SUGGESTIONS

Based on the research conducted, it was found that the full learning CNN model demonstrated superior accuracy and faster processing time compared to the VGG-16 transfer learning model. The results indicate that the CNN full learning approach effectively classifies coffee bean roasting levels, achieving a test accuracy of 98.75% with a running time of 856 ms per step. The application of CNN in classifying roasting levels enables a more efficient, consistent, and objective assessment process. In the context of the instant coffee industry, this approach offers significant advantages, particularly in improving quality control and minimizing the subjectivity associated with traditional roasting assessments conducted by expert roasters.

For future research, it is recommended to enhance the dataset used in training by incorporating data augmentation techniques as part of the preprocessing stage in the deep learning model. Additionally, expanding the classification system by introducing two additional class outputs for coffee bean roasting levels could further refine the model's performance and applicability in real-world scenarios. Furthermore, hyperparameter tuning, including optimization of the learning rate, batch size, and dropout rates, could be explored to further improve model accuracy and generalization. Fine-tuning these parameters may enhance the efficiency and robustness of the CNN model, ensuring its reliability across different datasets and practical implementations.

REFERENCES

- [1] Almayda, N., & Najih, R. R. (2024). Analysis of Caffeine Content of Liberica Coffee (*Coffea liberica*) with Variations of Roasting Methods at Kalipuro Research Center, Banyuwangi District. *Journal of Coffee and Sustainability*, 1(1), 17-21.
- [2] Anu, T. A., Rosnelly, R., Irawan, D., Hasibuan, U., & Bulolo, P. (2023). KLASIFIKASI FITUR WARNA LEVEL ROASTING BIJI KOPI MENGGUNAKAN ARTIFICIAL NEURAL NETWORK. *Device*, 13(1), 8-13.
- [3] Freitas, V. V., Borges, L. L. R., Vidigal, M. C. T. R., dos Santos, M. H., & Stringheta, P. C. (2024). Coffee: A comprehensive overview of origin, market, and the quality process. *Trends in Food Science & Technology*, 104411.
- [4] Li, G., Huang, Y., Chen, Z., Gary D. Chesser Jr, Purswell, J. L., Linhoss, J., & Zhao, Y. (2024). Practices and Applications of Convolutional Neural Network-Based Computer Vision Systems in Animal Farming: AReview. *Sensors* 24(23), 1-44.
- [5] Na'im, M. Z., Chusnah, M., & Hartanti, D. A. S. (2024). Effect of Roasting Time on The Organoleptic Characteristics of Excelsa Wonosalam Jombang Coffee. *AGARICUS: Advances Agriculture Science & Farming*, 4(1), 19-22.
- [6] Nissa, N. F., Janiati, A., Cahya, N., Anton, & Astuti, P. (2021). Application of Deep Learning Using Convolutional Neural Network (CNN) Method for Women's Skin Classification. *Scientific Journal of Informatics* 8(1), 144-153.
- [7] Reghunandanan, K., Lakshmi, V., Raj, R., Viswanath, K., Davis, C., & Chandramohanadas, R. (2024). A Convolutional Neural Network- Based Deep Learning To Detect Reticulocytes From Human Peripheral Blood. *Intelligence-Based Medicine* 10, 1-10.
- [8] Sagita, D., Mardjan, S. S., Purwandoko, P. B., & Widodo, S. (2024). Low-cost IoT-based multichannel spectral acquisition systems for roasted coffee beans evaluation: Case study of roasting degree classification using machine learning. *Journal of Food Composition and Analysis*, 133, 106478.
- [9] Severin, H. G., & Lindemann, B. (2024). Elasticity of coffee beans: A novel approach to understanding the roasting process. *Journal of Food Engineering*, 383, 112212.
- [10] Solikhin, Wicaksono, P. A., & Santoso, A. W. (2023). PENERAPAN TEKNOLOGI TEPAT GUNA MESIN ROASTING KOPI PADA UKM KOPI PINANGGIH. *JURNAL PASOPATI* 5(3), 138-143.
- [11] Tripathi, M. (2021). Analysis of Convolutional Neural Network based Image Classification Techniques. *Journal of Innovative Image Processing* 3(2), 100-117.
- [12] Wang, C., Liu, S., Wang, Y., Xiong, J., Zhang, Z., Zhao, B., . . . He, P. (2022). Application of Convolutional Neural Network-Based [12] Detection Methods in Fresh Fruit Production: A Comprehensive Review. *Frontiers in Plant Science* 13, 1-28.
- [13] M. Brahimi, M. Arsenovic, S. Laraba, S. Sladojevic, K. Boukhalfa, and A. Moussaoui, "Deep learning for plant diseases: detection and saliency map visualisation," In Human and machine learning, pp. 93-117, Springer, Cham, 2018. https://doi.org/10.1007/978-3-319-90403-0_6.
- [14] G. Wang, Y. Sun, and J. Wang, "Automatic image-based plant disease severity estimation using deep learning," *Computational Intelligence and Neuroscience*, 2017, <https://doi.org/10.1155/2017/2917536>.
- [15] J. Wang, L. Chen, J. Zhang, Y. Yuan, M. Li, and W. Zeng, "CNN Transfer Learning for Automatic Image-Based Classification of Crop," In Image and Graphics Technologies and Applications: 13th Conference on Image and Graphics Technologies and Applications, IGTA 2018, Beijing, China, April 8-10, 2018, vol. 875, pp. 319-392, 2018, https://doi.org/10.1007/978-981-13-1702-6_32.
- [16] Anvesh, N., Yuvaraj, G., Lakshmi, A. V., & Reddy, S. S. (2024, April). Comparative Analysis of Deep CNN and VGG-16 for Anomaly Detection. In *2024 Third International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)* (pp. 1-7). IEEE.
- [17] Ontoum, S., Khemanantakul, T., Sroison, P., Triyason, T., & Watanapa, B. (2022). Coffee roast intelligence. *arXiv preprint arXiv:2206.01841*.
- [18] Shihab, M., & Perdana, A. S. (2024). PENDAMPINGAN KOMUNIKASI PEMASARAN KEDAI KOPI JEVERA COFFEE AND EATERY CIKARANG. *Jurnal Abdi Panca Marga*, 5(1), 20-26.
- [19] Aquil, M. A. I., & Wan Ishak, W. H. W. (2021). Evaluation of scratch and pre-trained convolutional neural networks for the classification of tomato plant diseases. *IAES International Journal of Artificial Intelligence (IJ-AI)*, 10(2), 467-475.