

Application of Vision AI for Assessing the Usability of Packaged Household Cooking Oil Based on Visual Images

Toshiro Frederick^a, James Owen Wikson^b, Amir Mahmud Husein^c,

^{a,b}*Universitas Prima Indonesia*, ^c*PUI-PT Teknologi dan Sistem Cerdas*

Corresponding Author:

^afredericktoshiro922@gmail.com

ABSTRACT

The quality of cooking oil plays an important role in maintaining public health; therefore, practical and accessible methods are needed to assess its feasibility. This study aims to implement a Vision Artificial Intelligence (Vision AI) approach based on Convolutional Neural Networks (CNN), specifically MobileNetV2 with transfer learning, to classify the feasibility of household packaged cooking oil using visual images. The dataset consists of cooking oil images categorized into two classes: usable and non-usable. The research stages include image acquisition, data preprocessing, model training, and performance evaluation using accuracy, precision, recall, and F1-score metrics. The experimental results show that the proposed model achieves an accuracy of 83.33% on the test dataset, with a precision of 78.57%, recall of 91.67%, and F1-score of 84.62%. Furthermore, evaluation using 5-fold cross-validation yields an average accuracy of 86.6% in the best scenario, indicating good model generalization, although a lower average accuracy of 58.1% is observed under more challenging data distributions. The training process demonstrates a stable learning pattern, characterized by increasing accuracy and decreasing loss values across epochs. Overall, these findings indicate that the Vision AI approach has strong potential to be utilized as a non-destructive decision-support system for assessing cooking oil feasibility based on visual image analysis. However, further improvements are required, particularly in enhancing model robustness, reducing false positive errors, and improving performance consistency across varying data distributions.

Keywords: Vision AI, Convolutional Neural Network, Cooking Oil, Image Classification, Deep Learning

INTRODUCTION

Cooking oil is one of the essential food ingredients that is almost always used in household cooking activities in Indonesia. It plays a crucial role in determining the flavor, texture, and overall quality of food, both in frying and sautéing processes. Along with the rising prices of food commodities, particularly cooking oil, households tend to reduce expenses by repeatedly reusing cooking oil (Mucti et al., 2023; Lilin, 2020). The repeated use of cooking oil without considering its usability limits can lead to a significant deterioration in oil quality and pose potential health risks. Repeated heating processes trigger oxidation, hydrolysis, and polymerization reactions, resulting in changes such as darker oil color, increased turbidity, and the formation of foam and sediment (Agung & Rismaya, 2024). The consumption of

degraded cooking oil has been associated with an increased risk of health disorders, highlighting the need for special attention to its suitability for use (Valantina, 2021).

Scientifically, the quality of cooking oil is commonly assessed using chemical parameters such as Total Polar Compounds (TPC), Free Fatty Acids (FFA), and Peroxide Value (PV) (Oktaviana et al., 2025). Although these methods provide accurate results, they require laboratory equipment, time, and relatively high costs, making them impractical for household-scale application (Safitri et al., 2021). Consequently, most people evaluate cooking oil quality based solely on subjective visual observation and odor, which often do not accurately reflect the actual quality condition of the oil (Ghifari & Utaminingrum, 2022). Advancements in Vision Artificial Intelligence (Vision AI), particularly in the field of computer vision, offer new opportunities to address this issue. Vision AI is capable of analyzing digital images and extracting visual patterns that are difficult to accurately recognize through human observation. This technology has been widely applied in various domains, such as product quality inspection, object classification, and image-based diagnosis in the agricultural and healthcare sectors (Pakiding et al., 2025; Putri & Rakasiwi, 2025).

In the context of cooking oil, visual characteristics such as color brightness, clarity, and the presence of particles or sediment can serve as preliminary indicators for assessing oil usability. By leveraging deep learning approaches, particularly Convolutional Neural Networks (CNNs), Vision AI systems can be trained to recognize these visual patterns and automatically classify the suitability of cooking oil (Kustijono et al., 2022). Several previous studies have investigated the classification of cooking oil quality using various methods, including approaches based on color and clarity features (Ghifari & Utaminingrum, 2022), the Naïve Bayes algorithm (Marofi et al., 2017), and Support Vector Machine (SVM) classifiers (Ramadan et al., 2019). Although these studies reported relatively good accuracy, many still rely on additional sensors or non-visual methods. The application of CNN-based Vision AI for directly assessing the usability of household cooking oil through visual images remains relatively limited.

Based on this background, this study aims to develop and evaluate a Vision AI system based on Convolutional Neural Networks (CNNs) to classify the usability of household cooking oil using visual images. This research is expected to provide an initial contribution to the development of practical, non-laboratory-based methods and to enhance public awareness of the importance of food safety in household environments.

METHODS

Research Design

This study is an applied research employing a quantitative experimental approach aimed at developing and evaluating a Vision Artificial Intelligence (Vision AI) system to classify the usability of household packaged cooking oil based on visual images. A deep learning approach, specifically Convolutional Neural Networks (CNN), is utilized due to its capability to automatically extract visual features without requiring manual feature engineering (Xie et al., 2022). The classification scheme used in this study is binary, consisting of usable and non-usable cooking oil, determined based on observable visual characteristics. The research

workflow includes data acquisition, labeling, preprocessing, model training, and performance evaluation, as illustrated in Figure 1.

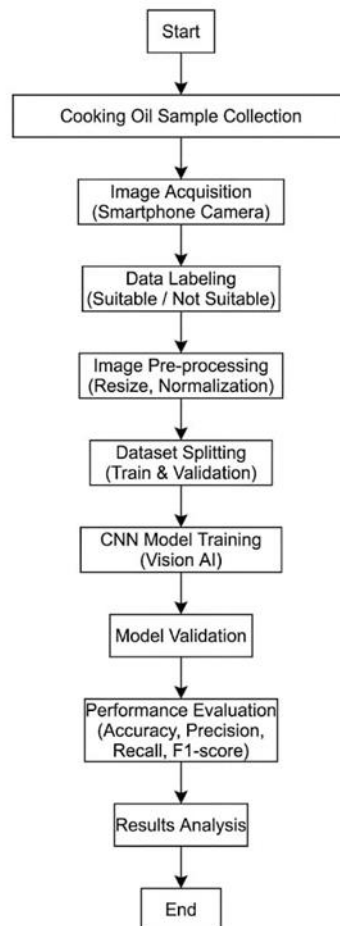


Figure 1. Research Framework

Dataset Acquisition and Characteristics

The dataset used in this study consists of 119 images of household packaged cooking oil based on palm oil, divided into 57 images of usable cooking oil and 62 images of non-usable cooking oil. All images were captured using a smartphone camera with a resolution of 3060 × 4080 pixels and stored in JPEG (.jpg) format. The image acquisition process was conducted under relatively consistent conditions in terms of distance, camera angle, and lighting to minimize irrelevant visual variations that could affect the learning process.

The dataset was divided into 119 images for training and 24 images for validation, where the validation data were used to monitor model performance during the training process.

Data Labeling Process

The data labeling process was conducted based on the visual characteristics of cooking oil, including color changes, clarity levels, and the presence of particles, foam, or sediment. To enhance labeling validity, this study employed standardized visual criteria referring to commonly recognized indicators of cooking oil degradation in household practices and food-related literature. The labeling process was performed directly by the researcher as a single

annotator, following a predefined classification guideline consistently across the entire dataset.

To minimize subjectivity, an internal validation process was carried out through repeated re-checking of all assigned labels to ensure consistency in visual interpretation across images. Although this study does not incorporate chemical parameters as ground truth, the visual-based labeling approach remains relevant as it reflects how users commonly assess cooking oil quality in real-world scenarios. Therefore, the developed system aligns with practical implementation contexts.

Image Preprocessing

Prior to model training, all images underwent preprocessing steps, including resizing to 224×224 pixels, normalization of pixel values to the range $[0,1]$, and adjustment of data format to match CNN input requirements. These steps aim to improve computational efficiency, ensure input consistency, and reduce the influence of irrelevant visual variations on model performance.

Vision AI Model Architecture and Training

The Vision AI model was developed using a Convolutional Neural Network (CNN) architecture based on a sequential model, consisting of multiple convolutional and pooling layers to extract visual features hierarchically. The main parameters used in the model include a kernel size of 3×3 , stride of 1×1 , and valid padding. The ReLU activation function was applied to convolutional layers, while the Softmax activation function was used in the output layer. To reduce overfitting, a dropout technique with a rate of 0.5 was implemented in certain layers.

The training process was conducted using two model scenarios, namely model v2 with 20 epochs and a batch size of 32, and model v4 with 40 epochs and a batch size of 16. Both models employed the Adam optimizer with a learning rate that was gradually reduced from 0.001 to 0.00005, along with the binary cross-entropy loss function.

Model Validation Strategy

Model validation was performed using a hold-out validation approach by utilizing 24 images as validation data that were not included in the main training process. This approach was chosen considering the limited dataset size, allowing for a reasonably representative evaluation of the model's generalization capability without significantly reducing the training data.

Performance Evaluation Method

Model performance was evaluated using standard classification metrics, including accuracy, precision, recall, and F1-score. The use of multiple metrics aims to provide a comprehensive assessment of the model's performance, particularly in the context of binary classification.

Development Environment

The model development was conducted using the Python programming language with the support of TensorFlow and Keras libraries for deep learning implementation, while Matplotlib and Pandas were used for data visualization and analysis.

RESULTS

Table 1. Summary of the CNN Model Architecture

Layer (Type)	Output Shape	Param #
convd (Conv2D)	(None, 150, 150, 16)	448
max_pooling2d (MaxPooling2D)	(None, 75, 75, 16)	0
conv2d_1	(None, 75, 75, 32)	4,640
max_pooling2d_1 (MaxPooling2D)	(None, 37, 37, 32)	0
conv2d_2 (Conv2D)	(None, 37, 37, 64)	18,496
max_pooling2d_2 (MaxPooling2D)	(None, 18, 18, 64)	0
flatten (Flatten)	(None, 128)	2,654,336
dense_1 (Dense)	(None, 1)	129
dropout (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 1)	129

Tabel 1 menyajikan ringkasan arsitektur model Convolutional Neural Network (CNN) yang digunakan dalam penelitian ini. Model dirancang secara bertingkat (sequential) untuk mengekstraksi fitur visual dari citra minyak goreng secara hierarkis, dimulai dari fitur dasar hingga fitur yang lebih kompleks. Lapisan pertama adalah Conv2D dengan output sebesar (150, 150, 16), yang berfungsi untuk menangkap fitur dasar seperti tepi (edges) dan tekstur awal dari citra. Lapisan ini memiliki 448 parameter yang menunjukkan jumlah bobot yang dipelajari pada tahap awal ekstraksi fitur. Selanjutnya, dilakukan MaxPooling2D yang mereduksi dimensi spasial menjadi (75, 75, 16), sehingga mengurangi kompleksitas komputasi sekaligus mempertahankan fitur penting. Lapisan konvolusi kedua (Conv2D) menghasilkan output (75, 75, 32), yang menunjukkan peningkatan jumlah filter menjadi 32 untuk menangkap fitur yang lebih kompleks. Proses ini kembali diikuti oleh MaxPooling2D yang menurunkan dimensi menjadi (37, 37, 32). Pada tahap ini, model mulai mempelajari pola visual yang lebih abstrak, seperti distribusi warna dan kejernihan minyak. Lapisan konvolusi ketiga (Conv2D) menghasilkan output (37, 37, 64), dengan jumlah parameter yang meningkat signifikan menjadi 18.496. Hal ini menunjukkan bahwa model mulai menangkap fitur tingkat tinggi yang lebih kompleks. Proses ini kembali diikuti oleh pooling untuk menghasilkan representasi fitur berukuran (18, 18, 64). Selanjutnya, fitur yang telah diekstraksi diubah menjadi vektor satu dimensi melalui lapisan Flatten, dengan total parameter sebesar 2.654.336. Nilai parameter yang besar pada lapisan ini menunjukkan bahwa sebagian besar kompleksitas model terletak pada proses transformasi fitur spasial menjadi bentuk yang dapat digunakan oleh lapisan fully connected. Lapisan Dense pertama berfungsi sebagai penghubung antara fitur hasil ekstraksi dan proses klasifikasi, dengan output satu neuron. Untuk mengurangi risiko overfitting, diterapkan lapisan Dropout, yang secara acak menonaktifkan sebagian neuron selama proses pelatihan. Terakhir, lapisan Dense

digunakan sebagai lapisan output dengan satu neuron, yang sesuai dengan skema klasifikasi biner (layak dan tidak layak). Lapisan ini umumnya menggunakan fungsi aktivasi sigmoid untuk menghasilkan probabilitas kelas.

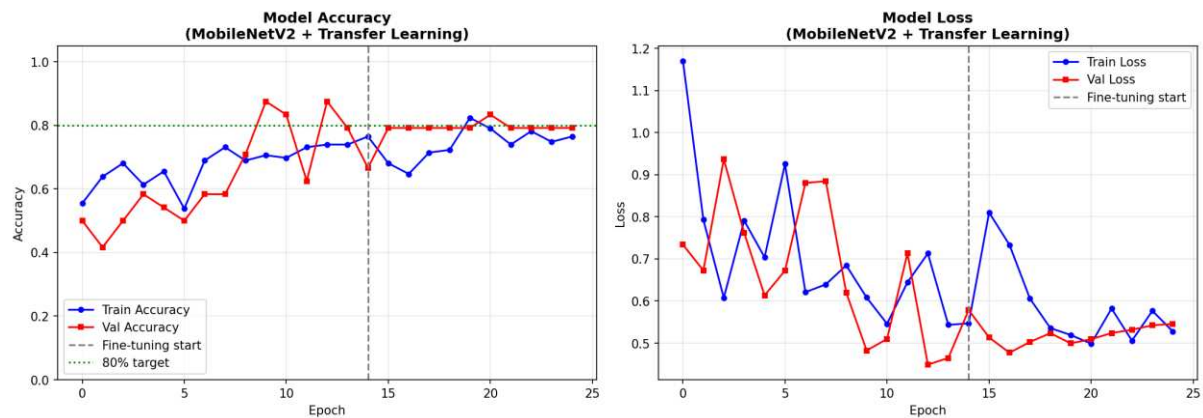


Figure 2. MobileNetV2 Model Training with Transfer Learning

Figure 2 presents the accuracy graph illustrating the performance progression of the MobileNetV2 model during the training and validation processes across each epoch, where the blue curve represents training accuracy and the red curve represents validation accuracy. In the early stage of training (epochs 0–5), the model accuracy remains relatively low and fluctuates, with training accuracy ranging from 0.55 to 0.68 and validation accuracy ranging from 0.42 to 0.58. This condition indicates that the model is still in the initial learning phase and has not yet been able to optimally capture visual feature patterns. Entering the epoch range of 6–13, there is a significant improvement in accuracy, particularly in the validation data, which reaches approximately 0.85–0.88. This indicates that the model begins to effectively extract relevant features from cooking oil images, supported by the effectiveness of the transfer learning approach in accelerating the learning process. Furthermore, the vertical dashed line at epoch 14 marks the beginning of the fine-tuning process. After this stage, both training and validation accuracy exhibit a more stable pattern, with validation accuracy ranging from 0.79 to 0.81 and training accuracy between 0.70 and 0.82. This stability suggests that the model has reached a convergent state.

In addition, the presence of a horizontal line representing the target accuracy of 80% indicates that the model is able to achieve and maintain the expected performance in the final epochs, particularly on the validation data. Although slight fluctuations are observed between training and validation accuracy, the difference is not significant, indicating no strong signs of overfitting. The close alignment of both curves in the final training phase suggests that the model has good generalization capability. Meanwhile, the loss graph illustrates the model's error rate during the training process. In the early stage (epochs 0–5), both training loss and validation loss are relatively high and unstable, with training loss exceeding 1.1 and validation loss ranging from 0.7 to 0.9. As the number of epochs increases, particularly within the range of 6–13, the loss values gradually decrease, indicating that the model becomes more effective in minimizing prediction errors and learning from the training data. After entering the fine-tuning phase (epoch ≥ 14), the loss values for both training and validation data become more stable, ranging approximately between 0.50 and 0.60. Although

there is a slight increase in training loss after epoch 15, the overall trend remains stable. Interestingly, validation loss tends to be more stable than training loss in the final epochs, indicating that the model has good generalization capability toward unseen data. Moreover, there is no significant indication of overfitting, as there is no sharp increase in validation loss when training loss decreases. Overall, the analysis of both graphs indicates that the model successfully learns relevant visual patterns for classifying the feasibility of cooking oil. The transfer learning approach combined with fine-tuning proves to contribute positively to improving model performance. The model also reaches a convergent state after approximately epoch 14 and is able to meet the target accuracy of 80% in the final training phase. With no significant indication of overfitting, it can be concluded that the model has good generalization capability.

These findings demonstrate that a Vision AI-based approach using a Convolutional Neural Network (CNN) architecture, particularly MobileNetV2 with transfer learning, is effective in classifying cooking oil feasibility based on visual images. The stability between accuracy and loss on both training and validation data indicates that the model has the potential to be implemented in real-world scenarios with a reliable level of performance.

Table 2. classification report

Algoritma	Label	Precision	Recall	F1-score	Support	average accuracy	Amount of data
CNN	MLP	0.77	0.83	0.80	12	0.79	24
	MLTP	0.82	0.75	0.78	12		

Table 2 presents the performance evaluation results of the MobileNetV2 model in classifying two classes, namely usable cooking oil (MLP) and non-usable cooking oil (MLTP), based on precision, recall, F1-score, and the number of samples (support). Overall, the model demonstrates a fairly balanced performance across both classes, achieving an overall accuracy of 79% from a total of 24 test samples. For the MLP class, the model achieves a precision of 0.77, a recall of 0.83, and an F1-score of 0.80 with a total of 12 image samples. The higher recall compared to precision indicates that the model is effective in identifying most of the truly usable cooking oil instances. However, the lower precision suggests the presence of false positives, where samples from the MLTP class are incorrectly predicted as MLP. Nevertheless, the relatively high F1-score reflects a good balance between detection capability and prediction accuracy for this class.

Meanwhile, for the MLTP class, the model achieves a precision of 0.82, a recall of 0.75, and an F1-score of 0.78, also with 12 image samples. The higher precision indicates that the model's predictions for non-usable cooking oil tend to be accurate. However, the lower recall suggests the presence of false negatives, where non-usable oil is incorrectly classified as usable. This issue is particularly important as it is directly related to consumption safety. A comparison between the two classes shows that the model exhibits different tendencies in handling each category. For the MLP class, the model is more sensitive in detecting usable oil, whereas for the MLTP class, the model is more accurate in confirming non-usable oil

predictions. This condition indicates that the model tends to be more conservative when classifying non-usable oil, but relatively more permissive when classifying usable oil. Overall, the model demonstrates fairly good performance without significant bias toward either class, supported by a balanced data distribution and relatively similar F1-score values. However, further improvement is still needed, particularly in increasing the recall value for the MLTP class to minimize false negatives. Therefore, although the model has achieved an accuracy of 79%, there remains room for improvement to enhance sensitivity toward non-usable cooking oil and to improve the overall reliability of the system.

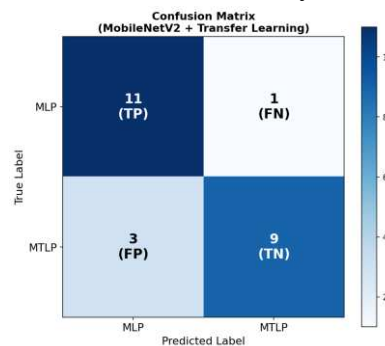


Figure 3. confusion matrix

Figure 3 presents the confusion matrix illustrating the performance of the MobileNetV2 model based on transfer learning in classifying cooking oil images into two categories, namely usable cooking oil (MLP) and non-usable cooking oil (MLTP). Based on the test results, the model correctly classified 11 samples as True Positives (TP), representing usable cooking oil images that were accurately identified, and 9 samples as True Negatives (TN), indicating that non-usable cooking oil images were also correctly recognized. These results suggest that the model has a good capability in distinguishing between the two main classes. However, several misclassifications are still observed. There is 1 False Negative (FN), where a usable cooking oil image is incorrectly classified as non-usable, indicating that in one case the model failed to recognize the characteristics of oil that is still suitable for use. In addition, there are 3 False Positives (FP), where non-usable cooking oil images are incorrectly classified as usable. This condition indicates that the model tends to be slightly permissive in classifying oil as usable.

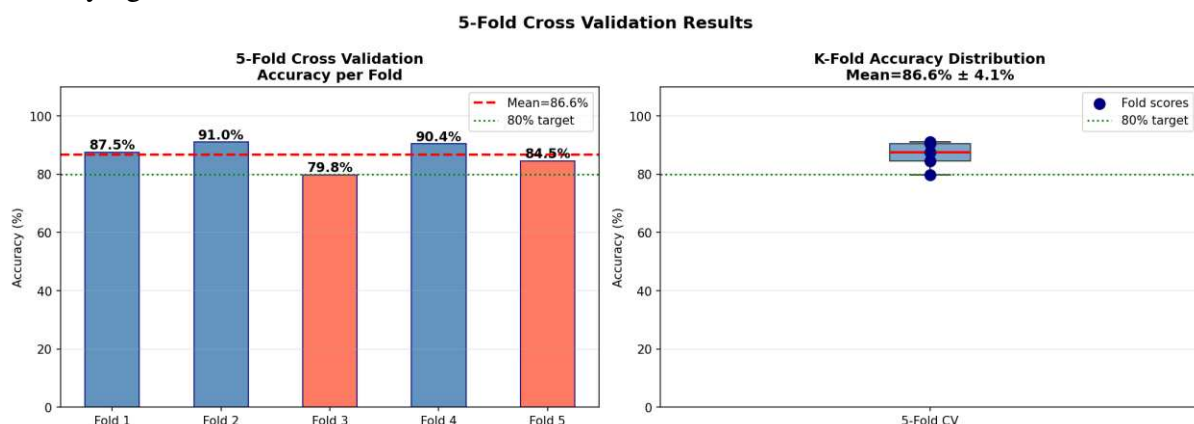


Figure 4. 5-Fold Cross Validation Results

Figure 4 presents the evaluation results of the MobileNetV2 model using the 5-fold cross-validation method, which aims to assess the model's stability and generalization capability across data variations. In this approach, the dataset is divided into five subsets (folds), where each subset is alternately used as testing data while the remaining subsets are used for training. This method enables a more comprehensive evaluation of model performance compared to using a single data split. Based on the experimental results, the accuracy values for each fold are as follows: 87.5% for Fold 1, 91.0% for Fold 2, 79.8% for Fold 3, 90.4% for Fold 4, and 84.5% for Fold 5. The average accuracy across all folds is 86.6%, indicating that the model generally achieves good performance. Most accuracy values exceed the target threshold of 80%, although one fold (Fold 3) falls slightly below this threshold. This variation suggests that the model performance is still influenced by the data distribution within each fold, particularly when there are significant differences in visual characteristics.

The analysis of accuracy distribution using a boxplot visualization shows a standard deviation of approximately $\pm 4.1\%$. This relatively low standard deviation indicates that the model performance is fairly stable and does not exhibit extreme fluctuations across folds. Most accuracy values fall within the range of 84% to 91%, reflecting the model's consistency in performing classification across different data subsets. The absence of significant outliers further indicates that the model does not experience drastic performance degradation under specific conditions. Overall, the cross-validation results demonstrate that the model has good generalization capability, as evidenced by consistent performance across multiple folds and high stability reflected by low variation in accuracy values. Although there is a slight performance drop in one fold, the model still shows relatively robust performance against data variability.

The observed variation in accuracy across folds may be attributed to several factors, such as uneven data distribution, variations in visual conditions of the images (e.g., lighting, color, and clarity), and the relatively limited dataset size. The fold with lower accuracy likely contains data with more complex or ambiguous visual characteristics, making it more challenging for the model to perform accurate classification. Thus, the results of the 5-fold cross-validation indicate that the model demonstrates stable and reliable performance, with an average accuracy of 86.6%. These findings reinforce the previous evaluation results and show that the model not only performs well on a specific data subset but also maintains consistent performance across various data partitioning scenarios.

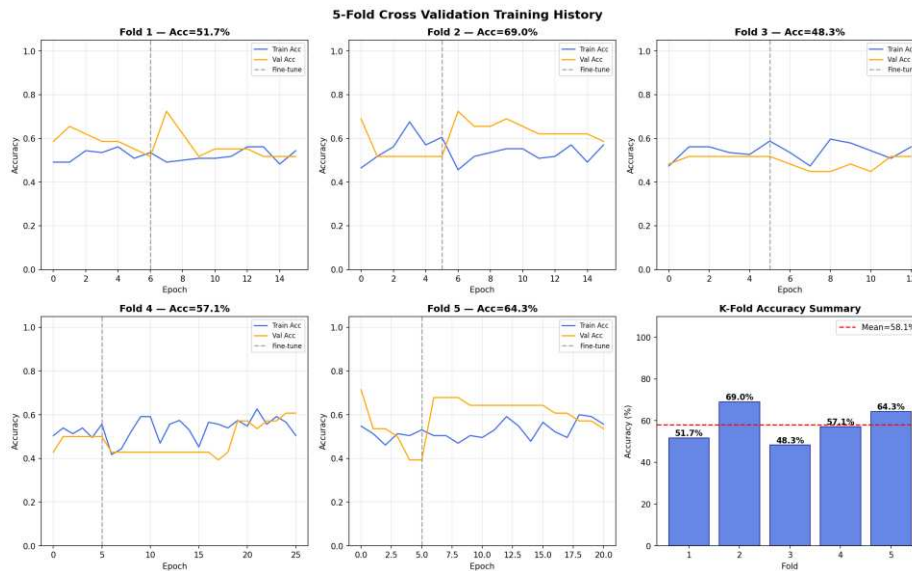


Figure 5. 5-Fold Cross Validation Training History

Figure 5 presents the performance evaluation results of the MobileNetV2 model using the 5-fold cross-validation method to assess its stability and generalization capability across data variations. Based on the experimental results, the accuracy values for each fold are 51.7% (Fold 1), 69.0% (Fold 2), 48.3% (Fold 3), 57.1% (Fold 4), and 64.3% (Fold 5), with an overall average accuracy of 58.1%. This value indicates that the model's performance is still in the moderate category and has not yet demonstrated optimal consistency across all data subsets. In more detail, Fold 2 shows the best performance with an accuracy of 69.0%, followed by Fold 5 at 64.3%. In these folds, validation accuracy exhibits a relatively stable improvement after the fine-tuning phase, with a small gap between training and validation accuracy. This indicates that the model has fairly good generalization capability without significant signs of overfitting. In contrast, Fold 3 shows the lowest performance with an accuracy of 48.3%, where validation accuracy tends to stagnate and remains lower than training accuracy. This condition suggests that the model struggles to learn the data patterns, likely due to uneven data distribution or high variability in visual characteristics within that subset.

Additionally, Fold 1 and Fold 4 also exhibit relatively low performance, with accuracies of 51.7% and 57.1%, respectively. In these folds, training accuracy tends to be higher than validation accuracy, indicating a tendency toward underfitting, where the model has not yet been able to capture complex patterns in the data effectively. Meanwhile, Fold 5 shows a more balanced pattern between training and validation accuracy, suggesting better generalization capability compared to several other folds. The significant variation in performance across folds, with a difference of more than 20% between the highest and lowest values, indicates that the model is still sensitive to data splitting. This condition commonly occurs in datasets with limited size or uneven feature distribution. Folds with lower performance likely contain data with more complex or ambiguous visual characteristics, making accurate classification more challenging for the model.

The analysis of the fine-tuning process shows that this stage contributes variably to performance improvement. In some folds, particularly Fold 2 and Fold 5, fine-tuning

significantly enhances accuracy, while in other folds the improvement is less pronounced, although it still helps maintain performance stability. This suggests that the effectiveness of fine-tuning is highly influenced by the characteristics of the data in each fold. Overall, the cross-validation results indicate that the model has a moderate level of generalization capability, but it is not yet optimal. There is no strong indication of overfitting, as the difference between training and validation accuracy is relatively small in most folds. However, the considerable variation in performance suggests that the model is still affected by limitations in dataset size and diversity. The implications of these findings indicate that although the transfer learning approach using MobileNetV2 is capable of classifying cooking oil images, the model performance still needs improvement. Therefore, expanding the dataset with more diverse samples is necessary to enable the model to learn more representative features. In addition, techniques such as data augmentation and ensemble learning approaches may be considered to improve accuracy and model stability in future research.



Figure 6. Failure Cases

Figure 6 presents four examples of misclassified images by the MobileNetV2 model based on transfer learning. Out of a total of 24 validation samples, 4 images were incorrectly classified, contributing to the model's overall accuracy of 83.33%. An analysis of these cases is conducted to identify error patterns and to understand the limitations of the model in distinguishing the visual characteristics of cooking oil. One identified error is a false negative (MLP → MLTP), specifically in the image MLP069.jpg, where oil that is actually classified as usable is predicted as non-usable with a confidence level of 89.0%. Visually, the oil in this image appears darker and slightly reddish compared to typical usable oil. This suggests that the model has high sensitivity to color attributes, where darker tones are associated with non-usable oil. As a result, the model fails to correctly recognize that the image still belongs to the usable category.

In addition, there are three false positive cases (MLTP → MLP), namely OleinAble-Oppo.jpg, OleinMM-Oppo.jpg, and OleinRFA-Oppo.jpg. In these cases, oil that is actually non-usable is classified as usable with confidence levels ranging from 80.8% to 95.3%. Visually, these images exhibit relatively bright yellow color, acceptable clarity, and minimal visible particles or sediment. These characteristics closely resemble those of usable oil, making it difficult for the model to distinguish them accurately. This finding indicates that

the model tends to rely heavily on surface-level visual features such as color and clarity, but has not yet captured more complex indicators of oil degradation. Based on the overall error cases, several general patterns can be identified, including the model's strong dependence on color as the primary feature, difficulty in distinguishing borderline conditions, and a tendency toward overconfidence in incorrect predictions. Some misclassifications even show confidence levels above 90%, indicating that the model is not yet well-calibrated in representing prediction uncertainty.

In real-world implementation, errors dominated by false positives have more critical implications, as non-usable oil may be classified as usable, potentially posing risks to users. In contrast, false negatives are relatively safer, although they may still reduce trust in the system. Several factors are suspected to contribute to these errors, including the limited size and variability of the dataset, visual similarity between classes, and the absence of non-visual information such as chemical parameters of the oil. To improve model performance, several strategies can be considered, including expanding the dataset with greater diversity, applying data augmentation techniques such as brightness, contrast, and blur, integrating additional texture-based features, and performing probability calibration to reduce overconfidence. Overall, the failure case analysis shows that although the model has achieved fairly good performance, it still has limitations in distinguishing images with similar visual characteristics. Therefore, improvements in data quality and training strategies are essential to enhance the accuracy and reliability of the system in future research.

DISCUSSION

Performance of Vision AI Model in Cooking Oil Feasibility Classification

The results of this study indicate that the Vision AI approach based on Convolutional Neural Networks (CNN), particularly using the MobileNetV2 architecture with transfer learning, is capable of classifying the feasibility of cooking oil based on visual images with a fairly good level of performance. The model achieved an accuracy of up to 83.33% on test data and an average accuracy of 86.6% under the best cross-validation scenario, demonstrating its ability to effectively distinguish between usable cooking oil (MLP) and non-usable cooking oil (MLTP). This performance suggests that visual characteristics such as color changes, clarity, and the presence of particles contribute significantly to the classification process. Although the visual differences between classes can be subtle in some cases, the model is still able to produce relatively consistent predictions across most test samples. This finding is consistent with previous studies, where CNN-based models tend to rely heavily on visual attributes in image classification tasks.

Training Process Analysis Based on Accuracy and Loss

During the training process, both training and validation accuracy showed a consistent increasing trend as the number of epochs increased, while the loss values gradually decreased. This pattern reflects a well-optimized learning process, where prediction errors are minimized as the network weights are updated. After entering the fine-tuning phase, the model performance became more stable, with a relatively small gap between training and validation accuracy. This indicates that the model has reached a convergent state without

significant overfitting. The stability also confirms that the transfer learning approach effectively accelerates the learning process while maintaining balanced performance on unseen data.

Discussion of K-Fold Cross Validation Results

The evaluation using K-Fold Cross Validation shows that the model performance is relatively stable across most folds, with average accuracy values exceeding the 80% threshold in the best scenario. This indicates that the model has a good generalization capability when dealing with data variations. However, in another scenario, a significant variation in performance was observed, with an average accuracy of 58.1%. This variation suggests that the model is still sensitive to data distribution, especially when the dataset is limited or contains uneven visual characteristics. Folds with lower accuracy likely contain more complex or ambiguous visual patterns, making classification more challenging. Nevertheless, the absence of drastic performance degradation indicates that the model does not rely on a single data split and maintains a reasonable level of robustness.

Error Model and Failure Case Analysis

The analysis of the confusion matrix reveals that classification errors are predominantly dominated by false positives (MLTP \rightarrow MLP) rather than false negatives. This indicates that the model tends to be more permissive in classifying oil as usable, which has more critical implications in real-world applications as it may lead to unsafe recommendations. Further analysis of failure cases identifies several key error patterns. In false negative cases, the model fails to recognize usable oil that appears darker or slightly reddish, suggesting a strong dependence on color features as the primary classification indicator. On the other hand, false positive cases occur when non-usable oil is visually similar to usable oil, characterized by bright yellow color, acceptable clarity, and minimal visible particles. This indicates that the model mainly relies on surface-level visual features and has not yet captured more complex degradation indicators.

Additionally, the model exhibits a tendency toward overconfidence in incorrect predictions, with some misclassified samples showing confidence levels above 90%. This limitation indicates that the model is not yet well-calibrated in representing prediction uncertainty. Overall, this error analysis provides important insights into the limitations of the model and highlights the need for improvement in feature representation and prediction reliability.

Suitability of CNN Architecture to Data Complexity

The MobileNetV2 architecture used in this study demonstrates a good balance between model complexity and computational efficiency. With a relatively lightweight structure, the model is still capable of extracting important visual features without significantly increasing the risk of overfitting. This suitability is reflected in the achieved accuracy and the stability of performance during both training and evaluation phases. However, the error analysis indicates that the architecture still has limitations in capturing more complex and ambiguous visual features, particularly in borderline cases. This limitation suggests that further

exploration of deeper architectures or the integration of additional feature representations may enhance model performance.

Quantitative Implications for System Implementation

From a quantitative perspective, the model demonstrates promising potential for implementation as a practical decision-support system in evaluating cooking oil feasibility. The achieved accuracy indicates that the system can provide reasonably reliable recommendations in real-world scenarios. However, the dominance of false positive errors remains a critical concern, as it directly relates to user safety. Variations in performance across folds and fluctuations during training also indicate that there is still room for improvement.

To enhance model performance, several strategies can be considered, including expanding the dataset with more diverse samples, applying data augmentation techniques such as brightness, contrast, and blur, integrating additional texture-based or non-visual features, performing probability calibration to reduce overconfidence, and exploring ensemble learning approaches to improve stability and accuracy. Therefore, although the model has demonstrated fairly good performance, improvements in data quality and training strategies remain essential to increase the reliability of the system for real-world applications.

CONCLUSION

This study successfully implemented a Vision Artificial Intelligence (Vision AI) approach based on Convolutional Neural Networks (CNN) to classify the feasibility of household packaged cooking oil using visual images. The developed model, utilizing the MobileNetV2 architecture with transfer learning, is capable of extracting and leveraging visual features such as color variation and clarity level as key indicators in determining oil usability. The experimental results demonstrate that the model achieves a fairly good level of classification performance, with validation accuracy ranging from moderate to high levels and reaching up to 83.33% on the test dataset. The training process exhibits a stable learning pattern, characterized by increasing accuracy and decreasing loss values as the number of epochs progresses, indicating effective model optimization. Furthermore, evaluation using the K-Fold Cross Validation method shows that the model maintains relatively consistent performance across different data partitions, suggesting an adequate level of generalization capability. However, variations in accuracy across folds also reveal that the model remains sensitive to data distribution and dataset limitations. In addition, the analysis of the confusion matrix and failure cases highlights several important limitations of the model. The classification errors are predominantly dominated by false positives, indicating that the model tends to be permissive in classifying non-usable oil as usable, which poses a critical concern in real-world applications. The failure case analysis further reveals that the model heavily relies on surface-level visual features, particularly color and clarity, and struggles to distinguish borderline or visually ambiguous samples. Moreover, the presence of overconfident incorrect predictions suggests that the model is not yet well-calibrated in representing prediction uncertainty. Overall, these findings indicate that while the Vision AI approach shows strong potential as a non-destructive decision-support system for assessing

cooking oil feasibility based on visual analysis, further improvements are still necessary. Future work should focus on expanding the dataset with greater diversity, applying advanced data augmentation techniques, integrating additional feature representations (such as texture or non-visual attributes), and exploring more sophisticated model architectures or ensemble methods. These enhancements are expected to improve classification accuracy, robustness, and the overall reliability of the system under varying real-world conditions.

LIMITATION

This study has several limitations that should be considered when interpreting the obtained results. First, the size and diversity of the cooking oil image dataset remain limited and may not fully represent all household cooking oil usage conditions. Variations in oil conditions, such as different levels of reuse, types of cooking oil, and differences in packaging brands, may influence visual characteristics that are not fully captured in this study.

Second, image acquisition was conducted under relatively uniform lighting conditions, camera angles, and container types. While this approach was intended to reduce visual noise, it limits the model's ability to handle more diverse environmental variations. In real-world applications, differences in lighting and image acquisition angles may affect the model's classification performance.

Third, the CNN architecture employed in this study is relatively simple. Although it achieved satisfactory performance, exploring more complex architectures or applying advanced optimization techniques could potentially improve model performance. In addition, training parameter settings such as the number of epochs, batch size, and learning rate were not extensively explored.

Finally, this study relies solely on visual information extracted from cooking oil images without integrating additional data sources, such as chemical parameters or supplementary sensors. A multimodal approach could potentially yield more accurate and robust classification results; however, this was beyond the scope of the present study.

REFERENCES

- Agung, G. S., & Rismaya, R. (2024). Pengaruh suhu pemanasan terhadap karakteristik mutu minyak goreng bekas pakai pedagang gorengan. *Agritekno: Jurnal Teknologi Pertanian*, 13(1), 15–23.
- Ghifari, H. S., & Utaminingrum, F. (2022). Klasifikasi kualitas minyak goreng berdasarkan fitur warna dan kejernihan dengan metode K-Nearest Neighbour berbasis Arduino Uno. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 6(7), 3269–3274.
- Kustijono, J. C., Utaminingrum, F., & Prasetio, B. H. (2022). Rancang bangun sistem klasifikasi kualitas minyak goreng berdasarkan warna dan kejernihan menggunakan metode Naïve Bayes berbasis Arduino Uno. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 6(8), 3761–3766.
- Lilin, P. S. C. D. A. N. (2020). Pemanfaatan minyak jelantah sebagai bahan dasar.

- Marofi, M. N., Syauqy, D., & Fitriyah, H. (2017). Rancang bangun sistem klasifikasi frekuensi penggunaan minyak goreng menggunakan metode Bayes. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 1(11), 1169–1177.
- Mucti, S., Purwasih, R., & Destiana, I. D. (2023). Analisis mutu minyak goreng yang dipakai oleh pedagang gorengan di Pasar Pujasera Subang. *Edufortech*, 8(1), 1–10.
- Oktaviana, L., Nugroho, W. A., & Al Riza, D. F. A. R. (2025). Evaluation of peroxide value and free fatty acid content in coconut oil and palm oil under different heating temperature treatments. *Jurnal Pangan dan Agroindustri*, 13(2), 101–111.
- Pakiding, D., Selao, A., & Wahyuddin, W. (2025). Implementasi computer vision dalam mendeteksi penyakit pada tanaman cabai dan tomat menggunakan convolutional neural networks. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 5(3), 841–850.
- Putri, C. A., & Rakasiwi, S. (2025). Diagnosis dini penyakit mata menggunakan convolutional neural network VGG-16. *Edumatic: Jurnal Pendidikan Informatika*, 9(1), 208–216.
- Ramadan, M. Y., Syauqy, D., & Tibyani, T. (2019). Implementasi metode support vector machine (SVM) terhadap pemakaian minyak goreng. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 3(2), 1669–1677.
- Safitri, I., Kushadiwijayanto, A. A., Sofiana, M. S. J., Yuliono, A., Warsidah, W., & Apriansyah, A. (2021). Penerapan IPTEK melalui pelatihan pemanfaatan limbah minyak jelantah sebagai sabun cuci piring. *Journal of Community Engagement in Health*, 4(2), 313–318.
- Valantina, S. R. (2021). Measurement of dielectric constant: A recent trend in quality analysis of vegetable oil—A review. *Trends in Food Science & Technology*, 113, 1–11.
- Miller, C., Portlock, T., Nyaga, D. M., & O’Sullivan, J. M. (2021). A review of model evaluation metrics for machine learning in genetics and genomics. *Frontiers in Bioinformatics*, 1, 1–14.
- Smith, L. A., Oakden-Rayner, L., Bird, A., Zeng, M., To, M.-S., Mukherjee, S., & Palmer, L. J. (2023). Machine learning and deep learning predictive models for long-term prognosis in patients with chronic obstructive pulmonary disease: A systematic review and meta-analysis. *The Lancet Digital Health*, 5(12), e872–e881.
- Xie, Y., Zaccagna, F., Rundo, L., Testa, C., Agati, R., Lodi, R., Manners, D. N., & Tonon, C. (2022). Convolutional neural network techniques for brain tumor classification (from 2015 to 2022): Review, challenges, and future perspectives. *Diagnostics*, 12(8), 1850.