Jurnal Teknik Informatika Vol. 18 No. 2, October 2025 (292-302)
ISSN: p-ISSN 1979-9160 (Print)| e-ISSN 2549-7901 (Online)
DOI: https://doi.org/10.15408/jti.v18i2.46449

JURNAL TEKNIK INFORMATIKA

Homepage : http://journal.uinjkt.ac.id/index.php/ti

Food Image Classification and Recipe Recommendation for South Sumatran Cuisine Using EfficientNetB1

Farhah Salsabillah[1], Ade Silvia Handayani[2], Nurhajar Anugraha[3*]

[1,2,3]Department of Electrical Engineering, Telecommunication Engineering, Sriwijaya State Polytechnic
[1,2,3]Jl. Srijaya Negara, Bukit Besar, Palembang, Sumatera Selatan, Indonesia

Article:
Accepted: July 22, 2025
Revised: April 12, 2025
Issued: October 30, 2025

***Correspondence Address:**
nurhajar.anugraha@polsri.ac.id

ABSTRACT

Visual-based food classification and recipe recommendation systems remain underexplored in the context of local culinary traditions. To address this gap, a system was developed using the EfficientNetB1 architecture of Convolutional Neural Networks (CNN), integrated with a Large Language Model (LLM) to generate South Sumatran recipes from food images, adapting suggestions to classification results. The model was trained using transfer learning on eight food ingredient classes selected for their prevalence in local cuisine. It achieved a validation accuracy of 98.2% and a test accuracy of 98%, with average precision, recall, and F1-score all exceeding 98%, indicating consistent and reliable performance. The system was deployed as a web-based application, DapoerKito, allowing users to upload food images, receive classification results, and obtain generated recipe suggestions. LLM-generated recipes are produced instantly, matched to ingredients, and shown in a clear format. These findings demonstrate the value of integrating computer vision and language generation in an AI-based platform that supports usability and cultural relevance. In addition to its technical capabilities, the system contributes to the digital preservation of regional culinary heritage through interactive AI. This CNN–LLM integration offers a novel approach for advancing food AI with diverse ingredients, personalized nutrition, and multilingual support.

Keywords : *convolutional neural network (CNN); EfficientNetB; food image classification; transfer learning; recipe recommendation system; large language model.*

1. INTRODUCTION

Technological advancement has catalyzed transformation across multiple sectors, including the culinary business. Culinary activities have become integral to a lifestyle. Nonetheless, most users depend on text-based recipe searches, necessitating manual entry of food ingredients, rendering them less efficient for rapid and practical culinary requirements[1]. Users should be able to find recipes effortlessly by submitting pictures of ingredients. Some systems have begun to explore automated recipe generation using food images with promising results, although this technology remains far from widespread adoption-particularly in everyday culinary practices[2]. The RecipeIS framework demonstrates a potential application by utilizing food image inputs to recognize components and acquire recipe information instantaneously[3]. This is especially evident in regions like Indonesia, where traditional methods for obtaining recipes continue to dominate.This is especially evident in regions like Indonesia, where traditional methods of obtaining recipes still dominate. Based on a survey by Jakpat, as many as 73% of Indonesians still rely on family recipes when cooking. As many as 56% of respondents seek cooking inspiration online, while 54% use social media to try new, more practical recipes. In addition, 21% of respondents get recipes from friends, and the same percentage choose to follow the recipe instructions listed on the food product packaging. Meanwhile, 19% of respondents choose to read cookbooks as a reference, and only 4% get recipes through cooking classes[4]. Presently, the majority of recipe recommendation algorithms are generic and fail to consider the distinctive culinary attributes of certain locations, such as South Sumatra, which boasts unique flavors and ingredients. This region is renowned for its rich and diversified traditional cuisine, including pempek, pindang patin, and tekwan [5], [6]. Nonetheless, the lack of recipe recommendation systems tailored to the unique South Sumatran specialties is a significant gap. Therefore, a system that can identify these food items from photos and suggest recipes autonomously is not just a necessity, but a potential game-changer.

The rapid advancement of artificial intelligence (AI) in sectors such as the culinary industry has opened new opportunities for the development of intelligent food systems[7]. One practical example is the use of Convolutional Neural Networks (CNN) to classify individual food ingredients from images, as demonstrated in recent segmentation-based models[8]. Among the various models developed for this task, one notable contribution is NutriFoodNeta CNN-based approach that achieves high accuracy in food image classification and nutritional estimation using a modified Inception-V3 architecture, showing robust performance on the Food101 dataset [9]. These CNN models are a form of machine learning (ML), where computer systems are trained to recognize patterns, group information, and make predictions based on previously learned data[10]. Deep learning techniques are accelerating the advancement of machine learning in digital image processing. The Convolutional Neural Network (CNN) is a prominent and widely employed Deep Learning model explicitly designed for efficiently and precisely classifying objects in images [11], [12]. CNN extracts visual information from images through multiple computational levels, including basic element detection and complex object recognition. This approach has shown considerable effectiveness in various applications, including automatic food classification[13], [14]. A recent approach enhances food image recognition by combining multi-level feature fusion with an attention mechanism, which significantly improves accuracy on complex food datasets[15]. Due to its capacity for visual pattern recognition, CNN can identify different food types from photos[16], thereby assisting users in locating recipes that match their available ingredients.

As AI and Deep Learning technology have improved, many studies have been done in many areas. Reference [17] presents one such study. This research employed a digital picture collection comprising ginseng, ginger, and galangal. There were 300 photos, with 100 from Google for each ingredient type. The dataset was divided in an 80:20 ratio. The CNN model used produced an accuracy of 98.75% for training data and 85% for testing data. This study shows that CNN is effective in classifying spice and herb images. However, one of the problems is the limited feature engineering process. In the study [18] classifying food based on regional food, based on the accuracy, precision, recall, and f1-score values, it can be

Jurnal Teknik Informatika Vol. 18 No. 2, October 2025 (292-302)
ISSN: p-ISSN 1979-9160 (Print)| e-ISSN 2549-7901 (Online)
DOI: https://doi.org/10.15408/jti.v18i2.46449

concluded that the CNN model performs exceptionally well in classifying typical Javanese and Sumatran foods. With a precision value of 68% for the Javanese class, it shows that of all the images predicted as typical Javanese foods, 68% of them are truly typical Javanese foods. Meanwhile, a precision of 63.07% for the Sumatran class shows something similar for the Sumatran class. Although the CNN method has been widely used, this model still has limitations in recognizing food ingredients with similar shapes and textures. To overcome this problem, recent studies have begun to adopt the EfficientNet architecture, which is more efficient in image processing and can improve classification accuracy. One study [19] showed that EfficientNetB1 can produce high accuracy in lung cancer classification from CT images, with an accuracy rate reaching 99.10%. This shows that the EfficientNetB1 architecture has good generalization capabilities for image variations and is stable in the training process, even on multiclass classification tasks.

With the advancement of artificial intelligence technology, the integration of image processing and natural language comprehension has become progressively prevalent in developing interactive and context-aware systems. Large Language Models (LLM) are pivotal in this scenario. One source that discusses the potential and implementation of LLM in depth is the book *DeepSeek AI: Panduan Super Lengkap untuk Memanfaatkan AI Secara Maksimal*[20]. The book outlines that DeepSeek AI has advantages in processing efficiency, context capture capabilities, and ease of integration, making it relevant for various of AI-based applications that require dynamic natural language responses. Furthermore, Pic2Plate, a vision-language-based system, utilizes retrieval-augmented generation to match user-submitted food images with appropriate recipes, demonstrating the growing importance of multimodal systems in culinary AI[21].

Various previous studies have discussed aspects of food ingredient classification and recipe recommendation systems, but research gaps still have not been answered. In the study [22] it has been shown that the ResNet-18-based Transfer Learning method can classify food and non-food images with an accuracy of up to 98.8%. However, this study has not integrated an AI-based recommendation system that can provide recipe suggestions based on available ingredients. To address this gap, a web-based recipe system was proposed in[23], which identifies ingredients using EfficientNet and links them to suitable recipes in real time. However, certain limitations remain. As shown in[24], the ResNet-18-based Transfer Learning method can classify food and non-food images with an accuracy of up to 98.8%. Nevertheless, this approach has yet to integrate an AI-based recommendation system that provides recipe suggestions based on available ingredients. Therefore, research that combines image recognition technology and AI-based recommendation systems in one web-based platform is still needed.

Based on the description above, this study aims to develop a web-based South Sumatran food recipe recommendation system that is able to identify food ingredients from images using CNN in the EfficientNetB1 architecture. While previous studies have either focused on image classification or recipe retrieval separately, this research uniquely integrates both in a single platform tailored to the regional characteristics of South Sumatran cuisine. By combining CNN-based image recognition and LLM-powered recipe generation via the DeepSeek API, the proposed system not only preserves culinary heritage but also provides a novel, automated solution for personalized cooking assistance. This dual integration-especially in a localized culinary context-has not been extensively explored in existing literature. In addition, this study will evaluate the performance of EfficientNetB1 in ensuring accurate and efficient recognition of diverse food ingredients.

2. METHODS

This section describes the overall methodology used in this study, which consists of several sequential stages. A structured overview of the process is presented in Figure 1.

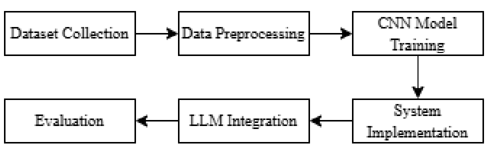

Figure 1. *Research design process*

2.1. Dataset Collection

The food ingredient data in this study were obtained from publicly available datasets on Kaggle and Roboflow. Specific keywords were used to search and retrieve relevant ingredient classes aligned with the study's objectives. Each image was manually selected to ensure label accuracy and visual consistency, thereby improving classification reliability. This study limited the classification of objects to only eight classes: *Daging Ayam, Daging Sapi, Durian, Ikan, Tahu, Telur, Tempe,* and *Udang*. These eight ingredients were chosen because they were most often used in culinary preparations typical of South Sumatra and had quite different visual characteristics so that they were easily recognized. The datasets were sourced from open-access repositories that provide images for research and educational use. Manual curation was performed to remove duplicates, blurry images, or irrelevant samples. Efforts were made to include visual diversity, such as differences in lighting, camera angles, ingredient states, and backgrounds. All datasets were confirmed to be free from copyright restrictions and ethically appropriate for academic research.

2.2. Data Preprocessing

The images underwent a data preprocessing step before being used to train the model. The first step was to normalize the pixels values to a scale of [-1, 1] using the built-in function of EfficientNetB1. Data augmentation was then applied throught rotation (up to 40%), horizontal flipping, zooming (up to 30%), shearing, and both horizontal and vertical shifting. The dataset was split into 80% training, 10% validation, and 20% testing, with a batch size of 16. Increasing data variability, decreasing overfitting, and preserving training efficiency were the goals of this procedure. Figure 2 illustrates the results of these augmentation techniques, particularly showing how variations like rotation and zoom affect the same food ingredient image (e.g., *Tahu*). The figure is not intended to represent a cropping process, but rather to demonstrate visual diversity introduced during preprocessing.

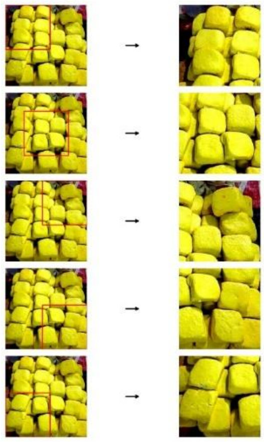

Figure 2. *Local Cropping of Ingredient Image*

2.3. CNN Model Training

This study used the EfficientNetB1 architecture, which is part of the Convolutional Neural Network (CNN) family, to build the image classification model. The architecture was chosen because it requires fewer parameters, supports various input image sizes, and has shown reliable performance in classification tasks.

The training procedure utilized the transfer learning technique, employing initial weights from the pre-trained ImageNet model. The model underwent training for 50 epochs, with a batch size of 16, a learning rate of 0.0001, and the Adam optimizer. The activation functions employed were ReLU for the hidden layer and softmax for the output layer, with categorical cross-entropy as the loss function for multiclass classification. The complete implementation was executed utilizing TensorFlow and Keras.

Modifications were implemented in the final layers of the model architecture to facilitate the classification of eight food classes. The modifications comprised incorporating a GlobalAveragePooling2D layer, a Dropout layer, two Dense layers with 1024 and 8 neurons, respectively, and a softmax activation function at the output layer. Figure 3 illustrates the model's final structure.

Jurnal Teknik Informatika Vol. 18 No. 2, October 2025 (292-302)
ISSN: p-ISSN 1979-9160 (Print)| e-ISSN 2549-7901 (Online)
DOI: https://doi.org/10.15408/jti.v18i2.46449

```
Layer name: block7b_drop
Output shape: (None, 8, 8, 320)
Layer name: block7b_add
Output shape: (None, 8, 8, 320)
Layer name: top_conv
Output shape: (None, 8, 8, 1280)
Layer name: top_bn
Output shape: (None, 8, 8, 1280)
Layer name: top_activation
Output shape: (None, 8, 8, 1280)
Layer name: global_average_pooling2d_10
Output shape: (None, 1280)
Layer name: dropout_20
Output shape: (None, 1280)
Layer name: dense_20
Output shape: (None, 1024)
Layer name: dropout_21
Output shape: (None, 1024)
Layer name: dense_21
Output shape: (None, 8)
```

Figure 3. *Adjusted EfficientNetB1 Model Architecture*

2.4. Sytem Implementation

A web application named DapoerKito was developed to implement the system in a realistic format. Streamlit was chosen for development due to its ease of integrating the model with the user interface. On the homepage, users may upload an image of a food component, and the system will attempt to identify the type of ingredient via a classifier based on EfficientNetB1.

After classification, the prediction results are used to automatically generate recipe recommendations through integration with the DeepSeek API's Large Language Model (LLM). The categorization output and the created recipes are presented directly in the application interface post-processing. The application is engineered to operate both locally and online, with interactions beginning via user-uploaded photo.

2.5. LLM Integration

The system was integrated with the LLM using the DeepSeek API, where the results of food item categorization were used as elements in the automated prompt. The prompt was constructed to simulate a natural interaction in Bahasa Indonesia, consisting of a system role defining the model's persona (e.g., a South Sumatran home cook) and a user role requesting recipes based on the classified ingredient. The request was sent via a POST method in JSON format, including model parameters (deepseek-chat), role, and prompt content. The LLM used for generating recipe text was not trained locally but accessed through a pre-trained external API. This approach simplifies system integration while still benefiting from the capabilities of a large-scale language model. The recipe outputs were returned in markdown format and streamed incrementally to simulate real-time feedback in the user interface. At this stage, the generated recipes were not validated by human experts; the goal was to demonstrate the technical feasibility of integrating LLM-based generation. Figure 4 shows the role-based prompt used in the LLM API request.

```
def get_recipes(self, jumlah, input_bahan):
    """Mengambil resep dari API Deepseek dengan streaming."""
    url = "https://api.deepseek.com/chat/completions"
    headers = {"Content-Type": "application/json",
            "Authorization": "Bearer [redacted]"}

    data = {
        "model": "deepseek-chat",
        "messages": [
            {
                "role": "system",
                "content": "Anda seorang ibu rumah tangga asal Sumatera Selatan yang hanya bisa berbahasa Indonesia, Anda
            },
            {
                "role": "user",
                "content": f"""Buatkan {jumlah} resep masakan khas Sumatera Selatan yang berbahan dasar {input_bahan}.
```

Figure 4. *Prompt structure and API call for LLM-based recipe generation*

2.6. Evaluation

The evaluation of model performance was performed utilizing four primary metrics: accuracy, precision, recall, and F1-score. These metrics assess the classification efficacy of the algorithm across eight food categories. The explanation of each metric is as follows.

$$\text{Accuracy} = \frac{\boldsymbol{TP+TN}}{\boldsymbol{TP+TN+FP+FN}} \text{ x } 100 \qquad (1)$$

Equation 1 explains that accuracy is the proportion of total data correctly classified by the model compared to all test data.

$$\text{Precision} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FP}} \qquad (2)$$

Equation 2 defines precision as the proportion of accurate optimistic predictions among all optimistic predictions.

$$\text{Recall} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FN}} \qquad (3)$$

Equation 3 showed that recall measured the model's ability to identify positive instances among all actual positive cases correctly.

$$\text{F1} - \text{Score} = \frac{\boldsymbol{2 \times Precision \times Recall}}{\boldsymbol{Precision + Recall}} \text{ x } 100\% \quad (4)$$

Equation 4 explains that the F1-Score is the harmonic mean between precision and recall and is often used as a primary metric in possibly imbalanced multiclass classification cases.

The confusion matrix is used to identify error patterns and to quantify correct and incorrect predictions for each class. It plays an important role in evaluating class imbalance, which can affect the interpretation of overall model performance.

Jurnal Teknik Informatika Vol. 18 No. 2, October 2025 (292-302)
ISSN: p-ISSN 1979-9160 (Print)| e-ISSN 2549-7901 (Online)
DOI: https://doi.org/10.15408/jti.v18i2.46449

3. RESULTS AND DISCUSSION

3.1. Image Classification

This study utilized a dataset of raw food images categorized into eight classes: *Daging Ayam, Daging Sapi, Durian, Ikan, Tahu, Telur, Tempe,* and *Udang*. This class selection is based on ingredients commonly used in traditional South Sumatran cuisine and characterized by distinct visual features.

The dataset comprised 4,000 photographs evenly distributed over eight categories, with each category containing 500 images. Each image was carefully chosen to ensure relevance and sufficient visual quality for classification.

The dataset was divided into training, validation, and testing. Table 1 presents the dataset split.

Table 1. Dataset Split

Data Type	Amount
Training Data	2.560
Validation Data	640
Tes Data	800

The list of classes used in this study can be seen in Table 2 :

Table 2. Dataset

No	Class	Example Dataset
1	Daging Ayam	
2	Daging Sapi	
3	Durian	
4	Ikan	
5	Tahu	
6	Telur	
7	Tempe	
8	Udang	

ISSN: p-ISSN 1979-9160 (Print)| e-ISSN 2549-7901 (Online)
DOI: https://doi.org/10.15408/jti.v18i2.46449

Using datasets with a balanced class distribution was intended to improve the classification process and reduce the chance of bias toward certain classes. Additionally, including different types of visuals in each class helps the model generalize better to new data.

3.2. CNN Model Training (EfficientNetB1)

We employed transfer learning and the EfficientNetB1 architecture for model training to obtain the initial weights from the previously trained ImageNet model. The model was trained over 50 epochs using the Adam optimizer, a batch size 16, and a learning rate 0.0001. ReLU was utilized as the activation function for the hidden layer, and softmax was employed for the output layer. The loss function employed was categorical cross-entropy.

The training results showed a consistent increase in accuracy and decreased loss for both training and validation data. The model training performance graph is presented in Figure 5.

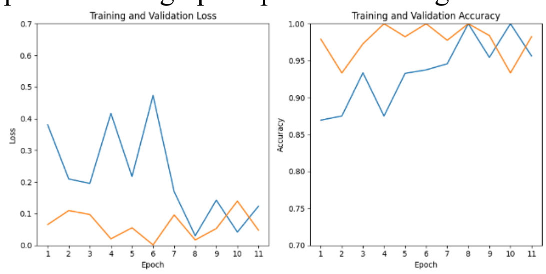

Figure 5. Model Training Accuracy and Loss Graph

To optimize the model performance, fine-tuning was performed with variations in learning rate values: 0.1, 0.01, 0.001, and 0.0001. The experimental results showed that a learning rate of 0.001 produced the best performance, with a validation accuracy of 98.2% and a validation loss of 0.048. Table 3 compares fine-tuning result.

Table 3. Fine-Tuning Results of the EfficientNetB1 Model

Learning rate	Epoch	Loss	Accuracy	Val Loss	Vall Accuracy
0,1	7	46,115	0,891	12,867	0,963
0,01	9	1,459	0,911	0,312	0,971
0,001	**11**	**0,124**	**0,956**	**0,048**	**0,982**
0,0001	13	0,153	0,950	0,069	0,979

Based on these results, the configuration with a learning rate of 0.001 was chosen as the final model because it provided the best training and validation results without indicating overfitting.

When considering the choice of EfficientNetB1, it helps to reflect on how baseline convolutional models have performed in similar classification contexts. In research by Li et al. (2020), several convolutional architectures were evaluated, including AlexNet, VGG, VGG-M, and VGG-M-BN, in a classification task involving vegetable images. Their findings showed that VGG-M-BN performed best, achieving 96.5% accuracy, followed by VGG-M at 95.8%, VGG at 92.1%, and AlexNet at 86.3%[25]. Although that dataset centered on vegetables, the core objective-image classification-remains comparable, which makes their results relevant as a point of reference.

In the present study, EfficientNetB1 reached 98.2% on validation and 98% on test accuracy, indicating a notably stronger performance. Admittedly, the models were not evaluated under the same conditions, but even so, these results hint at the efficiency and capability of EfficientNetB1 for food-related image tasks. A broader experimental comparison is encouraged in future studies to validate and generalize the observed performance.

3.3. Model Evalution

The image classification model performance evaluation results were carried out based on the test results on 800 images of eight food classes. The classification results are shown in the confusion matrix in Figure 6 below.

Jurnal Teknik Informatika Vol. 18 No. 2, October 2025 (292-302)
ISSN: p-ISSN 1979-9160 (Print)| e-ISSN 2549-7901 (Online)
DOI: https://doi.org/10.15408/jti.v18i2.46449

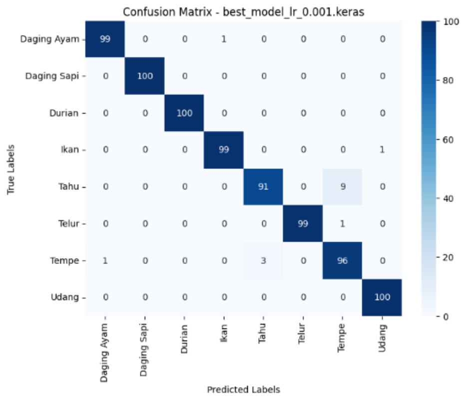

Figure 6. Confusion Matrix

Based on the confusion matrix, the model produces 784 correct predictions from 800 data, so the model accuracy can be calculated as follows:

$$\text{Acurracy} = \frac{784}{800} \text{ x } 100\% = 98\%$$

To assess each class's performance, precision, recall, and F1-score metrics were calculated using the formulas described in Section 2.6. An example calculation is provided for the *Daging Ayam* class:

- True Positive (TP) = 99
- False Positive (FP) = 1
- False Negative (FN) = 1

$$\text{Precision} = \frac{99}{99+1} = 0{,}99 = 99\% \quad (5)$$

$$\text{Recall} = \frac{99}{99+1} = 0{,}99 = 99\% \quad (6)$$

$$\text{F1} - \text{Score} = \frac{2 \times 0{,}99 \times 0{,}99}{0{,}99 + 0{,}99} = 0{,}99 = 99\% \quad (7)$$

The same approach was applied to the remaining classes. The complete evaluation results are shown in Table 4.

Table 4. Per-Class Evaluation Metrics

Class	*Precision*	Recall	F1 – Score
Daging ayam	99	99	99
Daging Sapi	100	100	100
Durian	100	100	100
Ikan	99	99	99
Tahu	91	91	92
Telur	99	99	99
Tempe	96	96	96
Udang	100	100	100
Average	98	98	98

With a test accuracy of 98% and average precision, recall, and F1-score all exceeding 98%, the EfficientNetB1 model demonstrated reliable and consistent classification performance across all food classes. These results support the system's effectiveness in identifying food ingredients as a foundation for automated recipe recommendations. Among the eight classes, the *Tahu* category showed slightly lower classification metrics. This may be attributed to its low visual contrast, uniform texture, and visual similarity to ingredients such as *Tempe* or *Daging Ayam*. Intra-class variation in *Tahu* appearance (e.g., different preparation forms or lighting conditions) may have also contributed to this dip in performance.

3.4. Web Application Result

The web-based application developed in this study is called DapoerKito. It is designed to help users discober traditional South Sumatran recipes by uploading images of food ingredients. This system integrates an image classification process using the EfficientNetB1 model and automatic recipe generation through the Large Language Model (LLM) DeepSeek API.

On the application's main page, users are guided by simple instructions. They can upload a picture of a food ingredient, pick how many recipes they want, and click the predict button. The model analyzes the image and shows the predicted ingredient. Based on this result, the system suggests recipes that match the identified ingredient. Figure 7 shows the main interface of the DapoerKito web application.

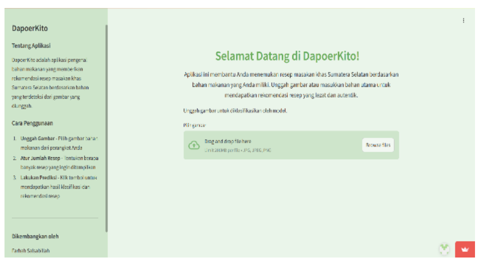

Figure 7. DapoerKito Web Interface

The application was developed using the Streamlit framework, which facilitated the integration of machine learning models with the user interface. The entire recipe classification and recommendation process was fully automated and executed with a short response time, resulting in a fast and responsive user experience.

Jurnal Teknik Informatika Vol. 18 No. 2, October 2025 (292-302)
ISSN: p-ISSN 1979-9160 (Print)| e-ISSN 2549-7901 (Online)
DOI: https://doi.org/10.15408/jti.v18i2.46449

3.5. LLM Recipe

Once the food ingredient has been classified, the system automatically produces recipe suggestions for traditional South Sumatran dishes using a Large Language Model (LLM) integrated via the DeepSeek API. The LLM is accessed through an external API and was not trained locally, allowing seamless integration without requiring additional model training or fine-tuning. The output is aligned with the classification results and can be adjusted according to the user's chosen number of recipes.

For example,a real-world image of raw *Udang* was taken under natural lighting conditions, with a non-uniform background and visible human interaction. The uploaded picture, which was not part of the training dataset, was properly labeled as *Udang* with a confidence score of 93.64%. Based on this prediction, the system recommended *Pindang Udang*, a traditional *Palembang* dish. Figure 8 illustrates this prediction process and the resulting recipe suggestion within the application interface.

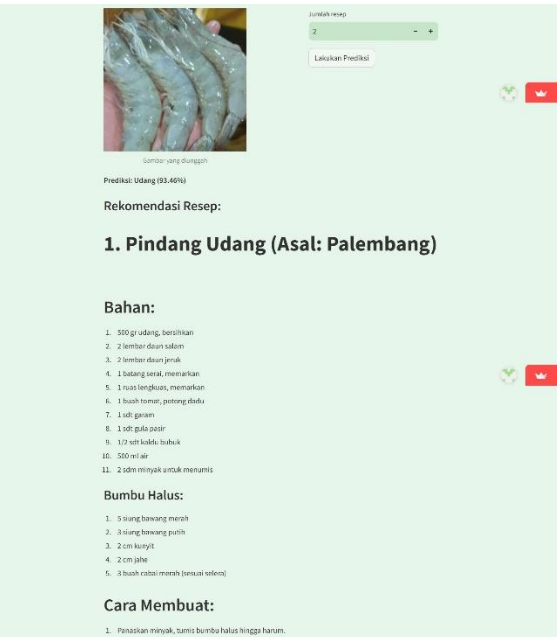

Figure 8. Recipe Recommendation Display

The generated recipe output includes:

a. The recipe name and its region of origin
b. A list of ingredients, covering both main components and spices
c. Clearly structured cooking instructions

Everything is arranged in a simple format that users can easily understand. This shows that the LLM works well in generating useful content and supports the app's goal of introducing local dishes based on the ingredients in the image. This example demonstrates that the system can accurately categorize items even when the input conditions are suboptimal and beyond the curated dataset. Although this singular case does not constitute a comprehensive generalization study, it provides preliminary evidence that the model can perform effectively with a broader spectrum of real-world image inputs. While this section focuses on the LLM-generated outputs, it is also important to acknowledge that the system's classification backbone performed reliably on a real-world image input. This observation suggests initial robustness beyond the curated dataset, although further testing with more diverse and noisy images is necessary to validate this capability.

In parallel, although the generated recipe results appear promising, this study did not include a dedicated evaluation of the LLM-produced descriptions through human judgment or established linguistic metrics such as BLEU or readability scoring. The main emphasis of the research was to analyze the classification capabilities of the EfficientNetB1 model and to build a functioning end-to-end system prototype. For that reason, the textual outputs were used mainly to demonstrate integration potential, rather than being tested for qualitative effectiveness. Future studies are encouraged to include human-centered evaluations to better assess the clarity, relevance, and usefulness of the generated content.

CONCLUSION

This study successfully developed a food image classification system using a Convolutional Neural Network (CNN) with EfficientNetB1 architecture, which showed good performance. The model achieved a validation accuracy of 98.2% and a testing accuracy of 98%, and generalized well to unseen data. The optimization of the learning rate parameters enhanced the model's accuracy and stability throughout training. The classification system is executed through an interactive web application named DapoerKito, allowing users to upload food images and

ISSN: p-ISSN 1979-9160 (Print)| e-ISSN 2549-7901 (Online)
DOI: https://doi.org/10.15408/jti.v18i2.46449

receive automatic classification results. Integration with a large language model (LLM) via the DeepSeek API facilitates the creation of traditional cooking recipes based on classification outcomes. This study emphasizes the capability of AI technology to aid in the preservation and promotion of local culinary culture, alongside its technical performance. The integration of CNN and LLM in one system is a multidisciplinary and user-centric approach. This study highlights its potential to support AI-based food recommendation systems that consider ingredient diversity, user preferences, and cultural relevance. Nevertheless, the current study has certain limitations, including a restricted number of ingredient classes, static LLM outputs without contextual adaptation, and the absence of human-centered evaluation. These constraints may affect the system's generalizability and user relevance. Future work is expected to address these challenges by expanding the dataset to include a wider range of food categories, incorporating user feedback mechanisms, and evaluating the quality of generated recipes through structured human studies. Personalization features such as nutrition-aware recommendations, time-based filtering, and contextual prompts (e.g., based on local weather) are also proposed to enhance user experience and practical impact.

REFERENCES

[1] K. A. Nair, S. G. Chavhan, and T. N. Pawar, "Recipe-fusion : Multimodal Food Recipe Recommendation System," *J. Artif. Intell. Res. Adv.*, vol. 11, no. 3, pp. 82–91, 2024.

[2] P. V. G. Bharane, C. S. Ashish, D. V. Subhash, and Z. K. Bharat, "RecipeReveal – Ingredients and Recipe Generation from Food Image," *Int. J. Adv. Comput. Eng. Commun. Technol.*, vol. 14, no. 1, pp. 36–40, 2025.

[3] M. S. Rodrigues, F. Fidalgo, and Â. Oliveira, "RecipeIS—Recipe Recommendation System Based on Recognition of Food Ingredients," *Appl. Sci.*, vol. 13, no. 13, 2023, doi: 10.3390/app13137880.

[4] B. Reynaldy, "73% Masyarakat Indonesia Memasak dengan Resep Asal Keluarga," *GoodStats*, 2024. [Online]. Available: https://data.goodstats.id/statistic/73-masyarakat-indonesia-memasak-dengan-resep-asal-keluarga-pZf3S.

[5] I. Fikriansyah, "10 Makanan Khas Sumatera Selatan Ini Wajib Dicoba, Enak Banget," *detikcom*, 2023. [Online]. Available: https://www.detik.com/sumbagsel/kuliner/d-6763244/10-makanan-khas-sumatera-selatan-ini-wajib-dicoba-enak-banget.

[6] Kallysta Wijaya, "Mengenal Lezatnya Makanan Khas Palembang: Dari Pempek hingga Pindang Patin," *Good News From Indonesia*, 2025. [Online]. Available: https://www.goodnewsfromindonesia.id/2025/02/20/mengenal-lezatnya-makanan-khas-palembang-dari-pempek-hingga-pindang-patin.

[7] M. Addanki, P. Patra, and P. Kandra, "Recent advances and applications of artificial intelligence and related technologies in the food industry," *Appl. Food Res.*, vol. 2, no. 2, p. 100126, 2022, doi: 10.1016/j.afres.2022.100126.

[8] Z. Zhu and Y. Dai, "A New CNN-Based Single-Ingredient Classification Model and Its Application in Food Image Segmentation," *J. Imaging*, vol. 9, no. 10, p. 205, 2023, doi: 10.3390/jimaging9100205.

[9] S. E. Sreedharan, G. N. Sundar, and D. Narmadha, "NutriFoodNet: A High-Accuracy Convolutional Neural Network for Automated Food Image Recognition and Nutrient Estimation," *Trait. du Signal*, vol. 41, no. 4, pp. 1953–1965, 2024, doi: 10.18280/ts.410425.

[10] X. Zhao, L. Wang, Y. Zhang, X. Han, M. Deveci, and M. Parmar, "A review of convolutional neural networks in computer vision," *Artif. Intell. Rev.*, vol. 57, no. 4, pp. 1–43, 2024, doi: 10.1007/s10462-024-10721-6.

[11] W. Hua, C. Li, and X. Wang, "Review of Convolutional Neural Network Models and Image Classification," *Acad. J. Sci. Technol.*, vol. 10, no. 3, pp. 178–184, 2024, doi: 10.54097/644jqv20.

[12] L. Alzubaidi *et al.*, "Review of deep learning : concepts , CNN architectures , challenges , applications , future directions," *J. Big Data*, vol. 8, p. Art. no. 53, 2021, doi: 10.1186/s40537-021-

ISSN: p-ISSN 1979-9160 (Print)| e-ISSN 2549-7901 (Online)
DOI: https://doi.org/10.15408/jti.v18i2.46449

00444-8.

[13] M. Dandi Darojat, Y. A. Sari, and R. C. Wihandika, "Convolutional Neural Network untuk Klasifikasi Citra Makanan Khas Indonesia," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 11, pp. 4764–4769, 2021.

[14] R. Abiyev and J. Adepoju, "Automatic Food Recognition Using Deep Convolutional Neural Networks with Self - attention Mechanism," *Human-Centric Intell. Syst.*, vol. 4, no. 1, pp. 171–186, 2024.

[15] Z. Chen, J. Wang, and Y. Wang, "Enhancing Food Image Recognition by Multi-Level Fusion and the Attention Mechanism," *Foods*, vol. 14, no. 3, pp. 1–21, 2025, doi: 10.3390/foods14030461.

[16] C. Mahaputri, Y. Kristian, and E. Setyati, "Pengenalan Makanan Tradisional Indonesia Beserta Bahan-bahannya dengan Memanfaatkan DCNN Transfer Learning," *INSYST Inf. Syst. Technol. J.*, vol. 04, no. 02, pp. 94–101, 2022, doi: 10.52985/insyst.v4i2.252.

[17] R. A. Boimau and Y. R. Kaesmetan, "Klasifikasi Citra Digital Bumbu dan Rempah Dengan Algoritma Convolutional Neural Network (CNN)," *Repeater Publ. Tek. Inform. dan Jar.*, vol. 2, no. 3, pp. 26–34, 2024, doi: 10.62951/repeater.v2i3.81.

[18] W. Prayogo Kusumo and C. Sri Kusuma Aditya, "Klasifikasi Citra Makanan Berdasarkan Asal Daerah Menggunakan Convolutional Neural Network," *Techno.COM*, vol. 23, no. 1, pp. 87–95, 2024, doi: 10.62411/tc.v23i1.9735.

[19] R. Raza *et al.*, "Lung-EffNet: Lung cancer classification using EfficientNet from CT-scan images," *Eng. Appl. Artif. Intell.*, vol. 126, no. PB, p. 106902, 2023, doi: 10.1016/j.engappai.2023.106902.

[20] B. A. Y. K. W. Putra *et al.*, *DeepSeek AI: Panduan Super Lengkap untuk Memanfaatkan AI Secara Maksimal*. SIEGA Publisher, 2025.

[21] Y. S. Soekamto, A. Lim, L. C. Limanjaya, Y. K. Purwanto, S. H. Lee, and D. K. Kang, "Pic2Plate: A Vision-Language and Retrieval-Augmented Framework for Personalized Recipe Recommendations," *Sensors*, vol. 25, no. 2, pp. 1–37, 2025, doi: 10.3390/s25020449.

[22] G. Thiodorus, A. Prasetia, L. A. Ardhani, and N. Yudistira, "Klasifikasi citra makanan/non makanan menggunakan metode Transfer Learning dengan model Residual Network," *Teknologi*, vol. 11, no. 2, pp. 74–83, 2021, doi: 10.26594/teknologi.v11i2.2402.

[23] M. R. M. Razali, H. F. Almarzuki, and N. A. S. Bahar, Rabiah Adawiyah Abdullah, "Food Recipe Recommendation Based on Ingredients Detection Using Deep Learning," *ACM Int. Conf. Proceeding Ser.*, no. InvENT, pp. 191–198, 2022, doi: 10.1145/3542954.3542983.

[24] K. A. Nfor, T. P. Theodore Armand, K. P. Ismaylovna, M. Il Joo, and H. C. Kim, "An Explainable CNN and Vision Transformer-Based Approach for Real-Time Food Recognition," *Nutr.* , vol. 17, no. 2, pp. 1–24, 2025, doi: 10.3390/nu17020362.

[25] Z. Li, F. Li, L. Zhu, and J. Yue, "Vegetable recognition and classification based on improved VGG deep learning network model," *Int. J. Comput. Intell. Syst.*, vol. 13, no. 1, pp. 559–564, 2020, doi; 10.2991/ijcis.d.200425.001.