JOURNAL LA MULTIAPP

VOL. 06, ISSUE 04 (921-933), 2025
DOI: 10.37899/journallamultiapp.v6i4.2269

Comparison of AlexNet and ResNet50 Model Performance in Classifying Images of Indonesian Traditional Food

Muhammad Randy Kurniawan[1], Yulison Herry Christanto[1], Gunawan Abdillah[1]

[1]Universitas Jenderal Achmad Yani, Indonesia

**Corresponding Author: Muhammad Randy Kurniawan*
Email: mrandykurniawan21@if.unjani.ac.id

Article Info

Article history:
Received 14 June 2025
Received in revised form 15 July 2025
Accepted 4 August 2025

Keywords:
Image Classification
AlexNet
ResNet50
Traditional Food
Deep Learning

Abstract

Image classification using deep learning has become an effective approach in various fields, including visual object recognition such as food identification. This study aims to compare the performance of two well-known Convolutional Neural Network (CNN) architectures, AlexNet and ResNet50, in classifying traditional Indonesian food images. The dataset used in this research is a combination of two sources: a traditional Indonesian cake dataset from Kaggle and an additional set of images of Cirebon's traditional dishes. The final dataset consists of 24 food categories with more than 4,000 images in total. Each image was preprocessed through several steps including resizing to 224x224 pixels, applying data augmentation to training samples to enhance variation, and normalization based on standard input formats of the models. The training process was carried out using the 5-Fold Cross Validation method, while performance was evaluated using accuracy, precision, recall, and F1-score metrics. Experimental results show that ResNet50 consistently outperformed AlexNet across all evaluation metrics. ResNet50 achieved an average accuracy of 92%, compared to 86% obtained by AlexNet. Additionally, ResNet50 demonstrated superior performance in terms of precision, recall, and F1-score. This difference indicates that deeper and more complex architectures like ResNet50 are more effective in learning visual patterns in diverse traditional food images. The study concludes that ResNet50 is a more optimal choice for the task of traditional Indonesian food image classification. These findings serve as a basis for future development of image-based food recognition systems and support the preservation of culinary heritage through artificial intelligence technology.

Introduction

Traditional Indonesian food is part of the country's rich culinary culture, reflecting its diversity and national identity. Each region has various types of specialities with unique shapes, textures and flavours. Unfortunately, amid the development of the times and changes in people's lifestyles, traditional foods are beginning to be replaced by modern food trends that are often considered more practical and appealing. This situation threatens the sustainability of traditional culinary heritage, especially for the younger generation who are beginning to lose knowledge about the diversity and cultural values behind traditional foods (Vishwakarma et al., 2025; Pugra et al., 2025; Bihari, 2023; Chukwurah et al., 2025; Apriyanto et al., 2024).

Technically, CNN is a deep learning method specifically designed to process two-dimensional image data, and it has demonstrated high effectiveness in classification tasks, especially in

ISSN: 2716-3865 (Print), 2721-1290 (Online)

pattern and feature recognition (Danendra et al., 2023; Iqbal et al., 2023; Yuan et al., 2022). CNN can automatically learn complex features in images without requiring manual feature engineering (Rohim et al., 2019; Liu et al., 2021; Demirci et al., 2021). Compared to conventional machine learning approaches, CNN is considered superior due to its ability to perform both feature extraction and classification in an end-to-end process (Lasniari et al., 2022). These advantages have made CNN the primary method in many image processing studies, including in the classification of traditional foods from various regions (Zidni & Akbar, 2024; Elngar et al., 2021; Marpaung et al., 2023). CNNs can also be implemented on various programming platforms, including MATLAB, for offline classification tasks (Irfansyah et al., 2021; Watanobe et al., 2023; Hamzah & Hreshee, 2025).

Two commonly used CNN architectures and the focus of this study are AlexNet and ResNet50. AlexNet was the first CNN architecture to gain worldwide attention after winning the ImageNet competition in 2012 with significantly better accuracy than previous models (Hermanto et al., 2024; Jena et al., 2021; Sharma & Guleria, 2022; Qawasmeh et al., 2025). Featuring five convolutional layers and three fully connected layers, AlexNet pioneered the use of GPUs in training deep networks and has been widely adopted for image classification tasks (DLY et al., 2023; Ullah et al., 2022; Tang et al., 2023). On the other hand, ResNet50 is a much deeper architecture with 50 layers and utilizes residual learning mechanisms to address the vanishing gradient problem (Sulistia & Vatresia, 2024). This mechanism enables deep network training without accuracy degradation and has been shown to improve model generalization (Permana et al., 2024; Gusti et al., 2024; Jiang et al., 2024).

Previous studies have shown the great potential of CNN in traditional food classification. For instance, the use of the ResNet50V2 architecture achieved 73.19% accuracy in traditional cake classification (Iskandar & Kristinto, 2023; Matarat, 2024). Meanwhile, a transfer learning approach using MobileNetV2 recorded up to 98% accuracy, outperforming other models such as VGG16 and ResNet50 (Faturrahman et al., 2023). In the context of AlexNet, a study using a traditional cake dataset reported 79% accuracy when combined with the Adam optimizer (Azizah et al., 2024; Chempak Kumar & Mubarak, 2024). Although these results are promising, there is still a lack of studies that directly compare the performance of AlexNet and ResNet50 architectures specifically in the classification of Indonesian traditional food, which features complex and highly diverse visuals.

Therefore, this study aims to evaluate and compare the performance of these two CNN architectures in classifying images of Indonesian traditional food. The dataset consists of various types of traditional dishes from different regions, with image variations designed to resemble real-world conditions. Each image undergoes preprocessing steps such as resizing, normalization, and augmentation before being used for model training. Evaluation is conducted using accuracy, precision, recall, and F1-score metrics, along with k-fold cross-validation to ensure stable and measurable results. This research is expected to contribute to the field of image processing and promote the use of artificial intelligence technology in preserving Indonesia's culinary heritage.

Methods

The research flow was systematically arranged to explain the stages of the experiment in comparing the performance of the AlexNet and ResNet50 models in classifying traditional Indonesian food images.

ISSN: 2716-3865 (Print), 2721-1290 (Online)

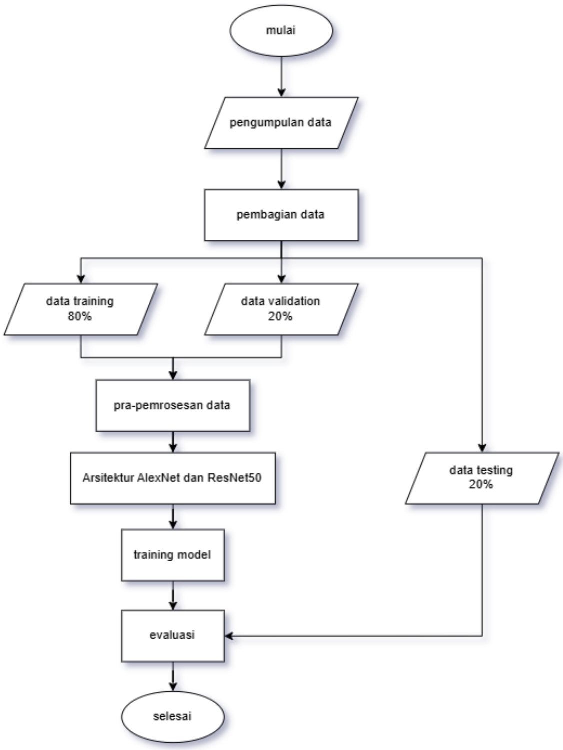


Figure 1. Research Flow

The flowchart illustrates the systematic flow of the research process, starting from the data collection stage to model evaluation. This explanation aims to provide a comprehensive overview of the methodology used in the research.

Dataset

The dataset is an important component in this research because the entire process of training, validating, and testing deep learning models depends on the quality and structure of the data used. In this research, the dataset used is a combination of two different sources, which are then compiled into a single entity, namely Indonesian Traditional Food.

Initially, the dataset used was a dataset of traditional Indonesian cakes obtained from the Kaggle platform. This dataset contains images of various types of traditional Indonesian cakes such as layered cake, klepon, putu cake, and others. However, to enrich and expand the scope of the data, the researchers added data on traditional foods from the Cirebon region, such as nasi jamblang, empal gentong, tahu gejrot, docang, and mie koclok. The images of these traditional Cirebon foods were obtained independently through searches from various online sources and personal documentation.

The combination of these two data sources resulted in a more diverse dataset, covering not only traditional cakes but also main dishes, side dishes, and street food from various regions. There are a total of 24 traditional food classes, with each class representing one type of food. Each image has been labelled according to the class name and organised into a folder structure to be compatible with the automatic labelling system using ImageFolder from PyTorch.

Data Collection

Data collection was conducted independently without directly involving participants. This process was divided into two types of sources: initial data from Kaggle and additional data from manual searches.

Data Collection Procedure

The steps taken in data collection are as follows:

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Initial Dataset Retrieval from Kaggle

The initial dataset, consisting of images of traditional Indonesian cakes, was downloaded from Kaggle. This dataset already has a classification structure per class, which facilitates the initial training process.

Addition of Cirebon Speciality Food Data

Images of Cirebon traditional foods were searched for and selected from various online sources such as culinary websites, social media, and personal documentation. The main criteria for image selection included visual clarity and the integrity of the food's appearance.

Manual Classification and Folder Format Adjustment

All additional images were manually classified and placed in folders based on their respective class names, using lowercase format without spaces (e.g., tahu_gejrot, nasi_jamblang), to ensure compatibility with the PyTorch system.

Data Storage Structure

The dataset is organised into three main sections: train, validation, and test, each comprising 24 class folders. Storage is structured on Google Drive for training purposes and integration with Google Colab.

Data Collection Techniques

The data collection technique used in this study is secondary digital documentation, namely by utilising images that are publicly available on the internet and open dataset repositories such as Kaggle. The data sources used include: a) Open datasets from Kaggle (specifically the Indonesian traditional cake category); b) Google Images, for searching for specific regional food images; c) Public social media platforms like Instagram, with open access; d) Personal documentation, particularly for Cirebon-specific foods that are difficult to find in public repositories.

Each image collected is checked for quality and relevance through visual observation, then classified and prepared for use in training and evaluating the deep learning model. This combination of data collection techniques from public repositories and manual searches enables the creation of a diverse, authentic dataset that better reflects the diversity of traditional cuisine in Indonesia.

Data Division

After all images of traditional Indonesian food have been collected and classified into 24 classes, the next step is to divide the dataset into three main parts, namely training data, validation data, and testing data. The purpose of this division is to ensure that the model training, evaluation, and testing processes are conducted fairly, measurably, and without overlap. The division is carried out as follows:

Training Data

Training data is used to train the model to recognise and learn the visual patterns of each food class. The proportion of training data is set at 80% of the total dataset, which will later be rearranged during the K-Fold Cross Validation process to improve training reliability.

Validation Data

20% of the training data will be used as validation data in each fold during the model training process. This data is not used directly in the model parameter updating process, but plays a role in measuring model performance at each epoch and detecting potential overfitting.

Test Data

ISSN: 2716-3865 (Print), 2721-1290 (Online)

The testing data consists of 20% of the total dataset that is completely separate and has never been used during the training or validation process. This data is used only once, after the training process is complete, to objectively measure the final performance of the model.

Each subset (train, validation, test) maintains a balanced class proportion so that the model can learn and be tested fairly across all food types.

Additionally, the dataset storage structure is systematically organised into three main folders (train, validation, and test), each containing 24 subfolders corresponding to class names. This format follows the PyTorch ImageFolder input standard, facilitating automated model training and testing processes.

Data Pre-Processing

Data pre-processing is an important stage that aims to prepare the dataset to be compatible with the model architecture used, as well as to improve data quality and variety. The food images that have been collected vary in size and orientation, so they need to be standardised before further processing. The pre-processing process applied includes the following steps: a) Resize: All images are resized to 224x224 pixels to match the standard input of the AlexNet and ResNet50 models, which were previously trained using the ImageNet dataset; b) Augmentation (specific to training data): Augmentation techniques such as random rotation and horizontal flipping are applied to increase data variation without having to manually add new data. The aim is to improve the model's generalisation ability to data it has never seen before; c) Normalisation: All image pixel values are normalised based on the mean and standard deviation of the ImageNet dataset. This normalisation aims to speed up the convergence process during training and maintain the stability of input values.

These pre-processing steps are an important foundation in the machine learning pipeline as they can significantly affect the quality of training and the classification results of the model.

CNN Architectures

In this study, two convolutional neural network (CNN) architectures that have proven reliable in image classification tasks were used, namely AlexNet and ResNet50. The selection of these two models was intended to compare the performance of architectures that differ in terms of complexity and network depth.

AlexNet is a CNN model with 5 convolutional layers and 3 fully connected layers. AlexNet is one of the pioneers in the advancement of deep learning, and although its architecture is relatively simple compared to modern models, its performance is still quite good for medium-sized classification tasks. AlexNet is suitable for use in systems with limited computational resources.

ResNet50 is a CNN model with a deeper architecture, consisting of 50 layers and equipped with a residual connection mechanism. This mechanism allows information flow to pass through multiple layers without degradation, helping the model learn more efficiently despite having many layers. ResNet50 is designed to address the vanishing gradient problem and is suitable for complex image classification tasks with large datasets.

The selection of these two architectures represents a contrasting approach between lightweight and deep models. By comparing the two, this study aims to identify which architecture is most optimal for classifying traditional Indonesian food images in terms of accuracy, efficiency, and generalisation.

Model Training

The training process was carried out using the 5-Fold Cross Validation method. The entire dataset was trained five times, where in each fold, 80% of the data was used for training and

ISSN: 2716-3865 (Print), 2721-1290 (Online)

20% for validation. Each model (AlexNet and ResNet50) is trained and validated on different data combinations, and the results are averaged. During training, the Adam optimiser, CrossEntropyLoss loss function, and batch size 16 are used. Training is performed using Google Colab to take advantage of free GPU access and avoid local computational limitations.

Model Evaluation

Model evaluation is one of the most important stages in research because it determines how well a trained model can accurately classify images. Evaluation is carried out using testing data that is completely independent from the training and validation processes. Thus, the evaluation results reflect the model's performance when faced with new, previously unknown data. The evaluation process is carried out based on several standard metrics in multi-class classification, namely: a) Accuracy: Measures the proportion of correct predictions out of the total predictions. The higher the accuracy value, the better the model is at classification in general; b) Precision: Measures the accuracy of positive predictions that are truly relevant. High precision means that the model does not produce many false positives; c) Recall: Measures the completeness or sensitivity, i.e., how much relevant data is correctly predicted. A high recall indicates that the model successfully captures all target categoriesl; d) F1-score: The harmonic mean between precision and recall, providing a balanced assessment if there is an imbalance between classes.

The evaluation is performed for each fold in the cross-validation scheme and averaged to obtain an overview of the model's performance. In addition, a confusion matrix is also used to evaluate the model's ability to distinguish between classes.

The evaluation results are presented in tables and bar graphs to clarify the comparison between models. All of these evaluation results form the basis for performance analysis and the final conclusions of the study.

Results and Discussion

Results per Fold

Model performance testing was conducted using the 5-Fold Cross Validation method to ensure that the results obtained were stable, independent of specific data subsets, and capable of describing the model's overall performance across the entire dataset. In this method, the training data was divided into five parts, and the training process was carried out five times.

AlexNet

The AlexNet model was tested using the 5-Fold Cross Validation method to assess the consistency and stability of classification performance on a dataset of traditional Indonesian food images. Each fold produced four main evaluation metrics, namely Precision, Recall, F1-Score, and Accuracy. The evaluation results per fold are shown in Table 3 below:

Table 1. Per-Fold Results from AlexNet

Fold	Precision	Recall	F1-Score	Accuracy
1	86.00%	86.00%	86.00%	86.00%
2	86.00%	86.00%	86.00%	86.00%
3	85.00%	85.00%	84.00%	85.00%
4	88.00%	86.00%	86.00%	86.00%
5	87.00%	87.00%	86.00%	87.00%

From the table, it can be seen that the results in Fold 1 and Fold 2 show identical performance with precision, recall, F1-score, and accuracy values of 86.00%. In Fold 3, there was a slight decrease, especially in the F1-score, which reached 84.00%, while precision and recall remained at 85.00%. Fold 4 shows an increase in precision to 88.00%, while recall, F1-score,

ISSN: 2716-3865 (Print), 2721-1290 (Online)

and accuracy remain at 86.00%. Then, in Fold 5, the model produces precision and recall of 87.00%, and F1-score and accuracy at 86.00%.

As a visual supplement, Figure 5 below presents a bar graph comparing the evaluation metrics of the AlexNet model on each fold. This graph shows that the distribution of metrics is fairly balanced across folds, making it easier for readers to observe evaluation trends between validation iterations.

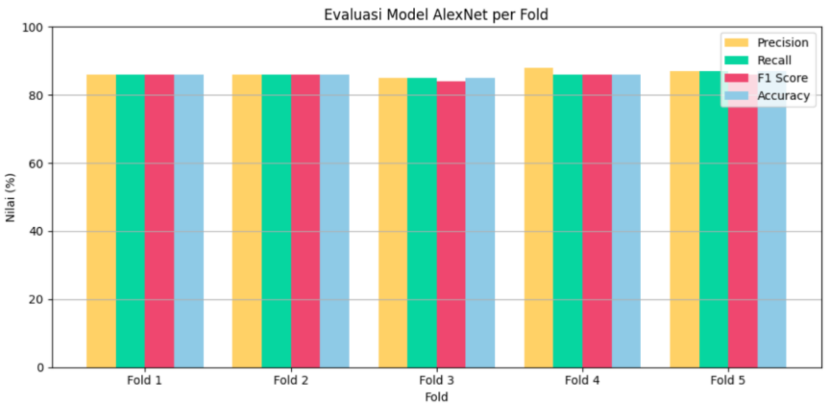


Figure 2. Per-Fold Evaluation of the AlexNet Model

From these results, it can be concluded that the AlexNet model is capable of producing stable performance in traditional food image classification, even though it uses a simpler architecture compared to modern models. The consistency of the results across each fold shows that AlexNet is quite effective in handling validation data variations during the training and evaluation process.

ResNet50

The ResNet50 model was implemented using the 5-Fold Cross Validation method to evaluate classification performance on a dataset of traditional Indonesian food images. This model is known for its deep residual network architecture, which enables deeper and more stable training without performance degradation.

The evaluation results per fold show consistent and high performance across all metrics. In Fold 1, precision, recall, F1-score, and accuracy all reached 92.00%, reflecting the model's initial stability. In Fold 2, all metrics improved to 93.00%, making this fold the highest overall. In Fold 3, precision and recall remained at 92.00%, but F1-score and accuracy slightly decreased to 91.00%. Fold 4 again shows solid performance with all metrics at 92.00%, the same as Fold 1. Meanwhile, in Fold 5, precision remains at 92.00%, but recall, F1-score, and accuracy decrease slightly to 91.00%, which is still in the very good category.

The complete evaluation results for each fold are shown in Table 4 below:

Table 2. Per-Fold Results from ResNet50

Fold	Precision	Recall	F1-Score	Accuracy
1	92.00%	92.00%	92.00%	92.00%
2	93.00%	93.00%	93.00%	93.00%
3	92.00%	92.00%	91.00%	92.00%
4	92.00%	92.00%	92.00%	92.00%
5	92.00%	91.00%	91.00%	91.00%

To clarify the distribution of values per metric in each fold, Figure 6 below presents a bar chart visualization of the ResNet50 model evaluation. This chart shows a stable metric

ISSN: 2716-3865 (Print), 2721-1290 (Online)

distribution pattern, with relatively small differences between folds, indicating that ResNet50 performs very well in traditional food classification tasks.

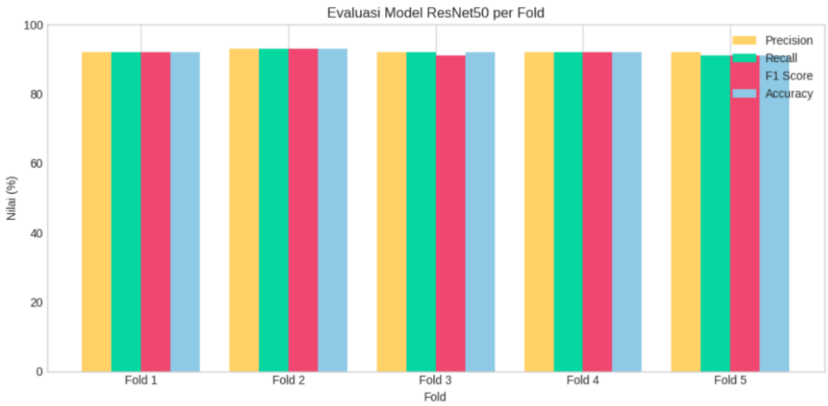


Figure 3. Per-Fold Evaluation of the ResNet50 Model

From these fold results, it can be seen that ResNet50 consistently delivers high performance in each validation iteration, with very minimal variation, reinforcing its reputation as one of the most effective deep learning architectures for image classification.

Results and Comparison of Average Evaluation Metrics

The analysis was conducted by comparing the performance of two CNN architectures: AlexNet and ResNet50 on the same dataset using the 5-Fold Cross Validation method. The evaluation was based on four main metrics: accuracy, precision, recall, and F1-score. Each model was evaluated on each fold, and the results were averaged to obtain the final score.

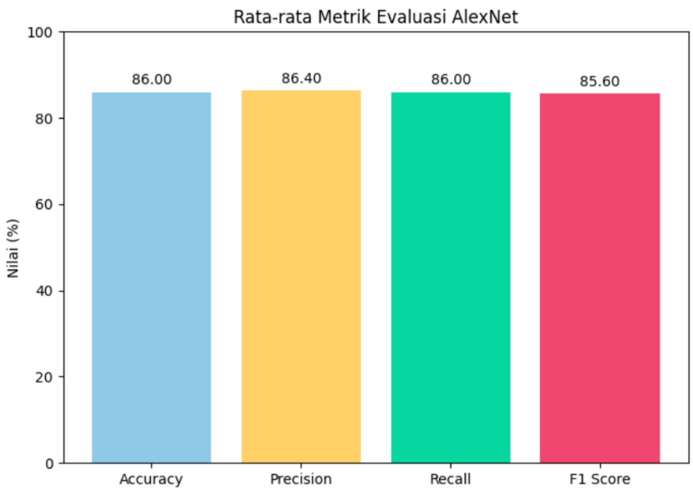
The following table shows the average metrics from the five folds for both models:

Table 3. Average Metrics from 5 Folds

Model	Acuracy	Precision	Recall	F1-Score
AlexNet	86.00%	86.40%	86.00%	85.60%
ResNet50	92.00%	92.20%	92.00%	91.80%

Based on these results, it is evident that the ResNet50 model consistently produces higher performance on all four metrics compared to AlexNet. Additionally, a confusion matrix was used to evaluate the details of predictions for each class, and the results indicate that ResNet50 is more capable of distinguishing between classes with visually similar appearances.

The results are also visualised in bar graphs to clarify the performance comparison between models:



ISSN: 2716-3865 (Print), 2721-1290 (Online)

Figure 4. AlexNet Evaluation Metrics Average Graph

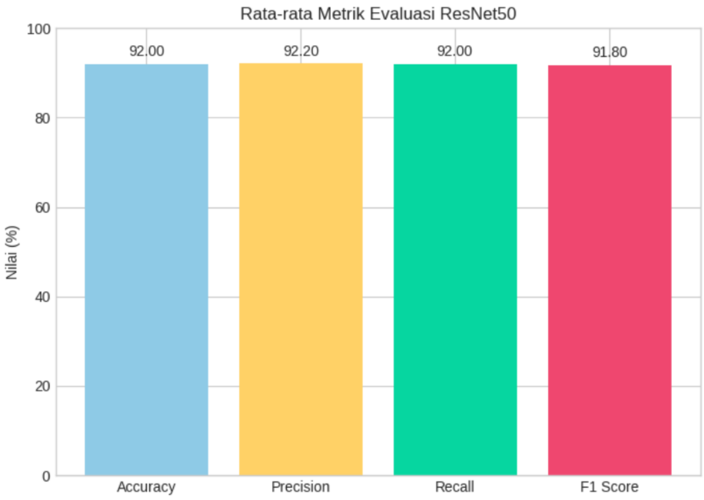


Figure 5. ResNet50 Evaluation Metrics Average Graph

The visualisation in the figure above shows a comparison of the average evaluation metric values for the AlexNet and ResNet50 models based on the results of 5-fold cross-validation. The metrics compared include accuracy, precision, recall, and F1-score, which are key indicators in measuring classification performance.

From the first graph, it can be seen that the AlexNet model produces an accuracy value of 86%, precision of 86.40%, recall of 86%, and F1-score of 85.60%. Meanwhile, in the second graph, the ResNet50 model shows higher performance, with an accuracy of 92%, precision of 92.20%, recall of 92%, and F1-score of 91.80%.

This comparison indicates that ResNet50 consistently outperforms AlexNet across all metrics. This reinforces the finding that the ResNet50 architecture, which has greater network depth and residual connection features, is able to capture visual patterns more accurately and effectively in the traditional Indonesian food dataset.

Conclusion

This study aims to compare the performance of two convolutional neural network (CNN) architectures, namely AlexNet and ResNet50, in classifying images of traditional Indonesian food. Based on the results of experiments and evaluations using the K-Fold Cross Validation method, it was concluded that ResNet50 showed superior performance compared to AlexNet in all aspects of evaluation. The average accuracy, precision, recall, and F1-score values achieved by ResNet50 were higher, indicating that this model is capable of recognising visual patterns more accurately and consistently.

The superiority of ResNet50 is supported by its deeper architecture structure and the use of residual blocks, which enable more efficient feature learning. Conversely, although AlexNet is quite good as a baseline model, its performance tends to be lower when faced with complex visual variations. The use of a combined dataset from Kaggle and Cirebon's traditional foods also successfully enriched the diversity of the data, allowing the system to be tested more comprehensively across many food classes.

This study also underscores the importance of selecting an architecture that aligns with the characteristics of the dataset. Additionally, cross-validation methods have proven effective in maintaining consistency in results and providing a more objective evaluation of model performance.

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Suggestion

Based on the findings and processes that have been carried out in this study, there are several suggestions that can be made for further development. Further research is recommended to expand the scope of the dataset, both in terms of the number and variety of foods from other regions in Indonesia, so that the classification system becomes more nationally representative. In addition, testing with real image data from user cameras is also important to measure the model's ability in more dynamic conditions.

The development of an application-based system is also a potential direction, either in the form of a web or mobile application, so that the results of this research can be directly utilised by the general public, culinary businesses, or educational institutions. Conceptually, this research supports the notion that deep learning models with deeper and more complex architectures are capable of providing better classification results on diverse images, and confirms that integrating data from various sources is an effective strategy for building adaptive and accurate systems.

Thus, the results of this research not only contribute to the academic field but also have the potential to be applied in real-world practice, particularly in supporting the preservation and promotion of traditional Indonesian food through artificial intelligence technology

References

Apriyanto, Y., Md Sharif, M. S., Mohd Shahril, A., Ishak, N., & Hashim, N. F. (2024). Exploring the concept of Indonesia traditional food sustainability. *Journal of Tourism, Hospitality and Culinary Arts*, *16*(1), 351-377.

Azizah, A. N., Budiman, I., Indriani, F., Faisal, M. R., & Herteno, R. (2024). Image classification of traditional Indonesian cakes using convolutional neural network (CNN). *Computer Engineering and Applications*, *13*, 38–49. http://dx.doi.org/10.18495/comengapp.v13i2.469

Bihari, S. (2023). Cultural heritage and indigenous knowledge: Reviving traditions for future generations. *Sustainable Development Goals in SAARC Countries: Key Issues, Opportunities and Challenges*, *1*, 24-32. http://dx.doi.org/10.5281/zenodo.8249545

Chempak Kumar, A., & Mubarak, D. M. N. (2024). Ensembled CNN with artificial bee colony optimization method for esophageal cancer stage classification using SVM classifier. *Journal of X-Ray Science and Technology*, *32*(1), 31-51. https://doi.org/10.3233/XST-230111

Chukwurah, G. O., Okeke, F. O., Isimah, M. O., Enoguanbhor, E. C., Awe, F. C., Nnaemeka-Okeke, R. C., ... & Okeke, C. A. (2025). Cultural Influence of Local Food Heritage on Sustainable Development. *World*, *6*(1), 10. https://doi.org/10.3390/world6010010

Danendra, F. A., Anggraeny, F. T., & Maulana, H. (2023). Image classification of chili leaf disease using CNN architecture AlexNet and SqueezeNet. *Syntax: Journal of Informatics*, *12*, 50–61. http://dx.doi.org/10.1016/j.ecoinf.2022.101663

Demirci, M. Y., Beşli, N., & Gümüşçü, A. (2021). Efficient deep feature extraction and classification for identifying defective photovoltaic module cells in Electroluminescence images. *Expert Systems with Applications*, *175*, 114810. https://doi.org/10.1016/j.eswa.2021.114810

DLY, I. A., Jasril, Sanjaya, S., Handayani, L., & Yanto, F. (2023). Image classification of beef and pork using CNN AlexNet and data augmentation. *Journal of Information System Research (JOSH)*, *4*, 1176–1185. https://doi.org/10.47065/josh.v4i4.3702

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Elngar, A. A., Arafa, M., Fathy, A., Moustafa, B., Mahmoud, O., Shaban, M., & Fawzy, N. (2021). Image classification based on CNN: a survey. *Journal of Cybersecurity and Information Management*, *6*(1), 18-50. https://doi.org/10.54216/JCIM.060102

Faturrahman, R., Hariyani, Y. S., & Hadiyoso, S. (2023). Classification of Indonesian traditional snacks based on deep learning and transfer learning methods. *ELKOMIKA: Journal of Electrical Energy Engineering, Telecommunication Engineering, & Electronics Engineering*, 945–957.

Gusti, H. P., Haerani, E., Syafria, F., Yanto, F., & Gusti, S. K. (2024). Implementation of convolutional neural network (ResNet-50) algorithm for classification of benign and malignant skin cancer. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *4*, 984–992. https://doi.org/10.57152/malcom.v4i3.1398

Hamzah, A. S., & Hreshee, S. S. (2025). A Scalable Face Recognition Framework: Integrating HOG Features, PCA, and CNN for Offline and Web-Based Systems. *International Journal of Intelligent Engineering & Systems*, *18*(2). https://doi.org/10.22266/ijies2025.0331.28

Hermanto, A. R., Aziz, A., & Sudianto. (2024). Comparison of MobileNetV2 and ResNet50 architectures for date fruit classification. *Journal of Information Systems and Technology*, *12*, 630–637. https://doi.org/10.26418/justin.v12i4.80358

Iqbal, S., N. Qureshi, A., Li, J., & Mahmood, T. (2023). On the analyses of medical images using traditional machine learning techniques and convolutional neural networks. *Archives of Computational Methods in Engineering*, *30*(5), 3173-3233. https://doi.org/10.1007/s11831-023-09899-9

Irfansyah, D., Mustikasari, M., & Suroso, A. (2021). AlexNet convolutional neural network (CNN) architecture for pest classification on coffee plant leaf images. *Informatics Journal: IT Development Journal (JPIT)*, *6*, 87–92. http://dx.doi.org/10.30591/jpit.v6i2.2802

Iskandar, J. S., & Kristinto, R. P. (2023). Recognition and classification of various Indonesian cakes using ResNet50V2 architecture in convolutional neural network (CNN). *[Journal name missing]*, 81–92.

Jena, B., Nayak, G. K., & Saxena, S. (2022). Convolutional neural network and its pretrained models for image classification and object detection: A survey. *Concurrency and Computation: Practice and Experience*, *34*(6), e6767. https://doi.org/10.1002/cpe.6767

Jiang, P., Neri, F., Xue, Y., & Maulik, U. (2024). A generalised attention mechanism to enhance the accuracy performance of neural networks. *International journal of neural systems*, 2450063. https://doi.org/10.1142/s0129065724500631

Lasniari, S., Jasril, Sanjaya, S., Yanto, F., & Affandes, M. (2022). Effect of ResNet-50 convolutional neural network hyperparameters on classification of beef and pork images. *National Journal of Computing and Information Technology*, *5*, 474–481. https://doi.org/10.32672/jnkti.v5i3.4424

Liu, Y., Pu, H., & Sun, D. W. (2021). Efficient extraction of deep image features using convolutional neural network (CNN) for applications in detecting and analysing complex food matrices. *Trends in Food Science & Technology*, *113*, 193-204. https://doi.org/10.1016/j.tifs.2021.04.042

Marpaung, F., Koemadji, Z. A., Hidayat, M., Widianto, A., & Aulia, F. (2023). Selection of Food Identification System Features Using Convolutional Neural Network (CNN)

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Method. *Scientific Journal of Informatics*, *10*(2), 205-2016. http://dx.doi.org/10.15294/sji.v10i2

Matarat, K. (2024). Enhancing Thai Food Classification: A CNN-Based Approach with Transfer Learning. *Mathematical Modelling of Engineering Problems*, *11*(6). http://dx.doi.org/10.18280/mmep.110625

Permana, F. Y., Aditya, C. S. K., & Chandranegara, D. R. (2024). Segmentation and image classification of skin cancer using convolutional neural network (CNN) with ResNet-50 architecture. *REPOSITOR*, *6*, 391–404.

Pugra, I. W., Kencanawati, A. A. A. M., & Kurniawan, I. G. W. A. (2025). The cultural significance of traditional foods in shaping Indonesian social identity: Challenges and preservation strategies. *Journal of Language, Literature, Social and Cultural Studies*, *3*(1), 21-31. https://doi.org/10.58881/jllscs.v3i1.318

Qawasmeh, B., Oh, J. S., & Kwigizile, V. (2025). Comparative Analysis of AlexNet, ResNet-50, and VGG-19 Performance for Automated Feature Recognition in Pedestrian Crash Diagrams. *Applied Sciences*, *15*(6), 2928. https://doi.org/10.3390/app15062928

Rohim, A., Sari, Y. A., & Tibyani. (2019). Convolutional neural network (CNN) for classification of traditional food images. *Journal of Information Technology and Computer Science Development*, *3*, 7037–7042.

Sharma, S., & Guleria, K. (2022, April). Deep learning models for image classification: comparison and applications. In *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)* (pp. 1733-1738). IEEE. http://dx.doi.org/10.1109/ICACITE53722.2022.9823516

Sulistia, F. Y., & Vatresia, A. (2024). Implementation of deep learning using convolutional neural network (CNN) for chicken meat classification with ResNet-50 architecture. *Journal of Information Technology and Computer Science (INTECOMS)*, *7*, 685–691. http://dx.doi.org/10.31539/intecoms.v7i3.10087

Tang, W., Sun, J., Wang, S., & Zhang, Y. (2023). Review of AlexNet for medical image classification. *arXiv preprint arXiv:2311.08655*. https://doi.org/10.48550/arXiv.2311.08655

Ullah, A., Elahi, H., Sun, Z., Khatoon, A., & Ahmad, I. (2022). Comparative analysis of AlexNet, ResNet18 and SqueezeNet with diverse modification and arduous implementation. *Arabian journal for science and engineering*, *47*(2), 2397-2417. http://dx.doi.org/10.1007/s13369-021-06182-6

Vishwakarma, N., Jaiswal, A., & Verma, M. (2025). Introduction to Culinary Heritage. In *Quality Assurance of Ethno-Herbals: Cultivating Confidence in Alternative Medicine* (pp. 1-15). Bentham Science Publishers. http://dx.doi.org/10.2174/9789815274554125010003

Watanobe, Y., Rahman, M. M., Amin, M. F. I., & Kabir, R. (2023). Identifying algorithm in program code based on structural features using CNN classification model. *Applied Intelligence*, *53*(10), 12210-12236. http://dx.doi.org/10.1007/s10489-022-04078-y

Yuan, S., Song, G., Huang, G., & Wang, Q. (2022). Reshaping hyperspectral data into a two-dimensional image for a CNN model to classify plant species from reflectance. *Remote Sensing*, *14*(16), 3972. https://doi.org/10.3390/rs14163972

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Zidni, E., & Akbar, M. (2024). Classification of Pasuruan City's traditional food images using convolutional neural network. *Informatics and Artificial Intelligence Journal*, *1*, 65–72. http://dx.doi.org/10.1109/ICICoS51170.2020.9299039

ISSN: 2716-3865 (Print), 2721-1290 (Online)