JURNAL INFORMATIKA
PROXIES

Vol.10, No.1, July, 2026

Evaluating CNN Models for Chicken Disease Identification Using Feces Image

Dwita Avrilly[1], Yulianto Tejo Putranto[2*]

[1] Teknik Informatika, Fakultas Ilmu Komputer, Universitas Katolik Soegijapranata, 22k30003@student.unika.ac.id
[2] Teknik Informatika, Fakultas Ilmu Komputer, Universitas Katolik Soegijapranata, yulianto@unika.ac.id

Corresponding Author Email: yulianto@unika.ac.id

https://doi.org/10.18280/proxies.v10i1.14953

Received: 2026-01-27
Revised: 2026-07-20
Accepted: 2026-07-20
Available online: 2026-07-23

Keywords:
Chicken Disease Classification, Convolutional Neural Network, Transfer Learning, Fine Tuning, Fecal Images, Deep Learning.

ABSTRACT

Poultry farming faces significant challenges due to infectious diseases such as Newcastle Disease, coccidiosis, and salmonellosis, which lead to high mortality rates and economic losses. The purpose of this study is to evaluate and compare the performance of multiple Convolutional Neural Network (CNN) architectures for automatic chicken disease classification using feces images. The methodology employs a transfer learning approach using five pretrained CNN models: VGG16, ResNet50, InceptionV3, MobileNetV2, and EfficientNetB0. The dataset consists of 8,067 images from four classes: Healthy, Newcastle Disease, Coccidiosis, and Salmonella. Image preprocessing includes resizing, normalization, and data augmentation, while model performance is evaluated using accuracy, macro precision, macro recall, and macro F1-score. The results show that all models achieve strong performance with accuracies above 96%. VGG16 and ResNet50 obtain the highest accuracy of 97.52%, while EfficientNetB0 achieves 97.19% with faster inference speed. The conclusions indicate that ResNet50 provides the most optimal balance between accuracy and computational efficiency. The implications of this study suggest that CNN-based transfer learning can be effectively applied as an early disease detection tool in poultry farming, especially for real-time and resource-limited applications.

1. INTRODUCTION

Poultry farming plays an important role in supplying animal protein, particularly in developing countries such as Indonesia. However, this sector faces serious challenges caused by infectious diseases, including Newcastle Disease, coccidiosis, and salmonellosis, which can lead to high mortality rates, decreased productivity, and significant economic losses [1], [2]. Early detection of these diseases is therefore essential to prevent large-scale outbreaks and to improve farm management efficiency.

Conventional diagnostic methods are generally based on visual observation or laboratory testing. Although laboratory analysis provides accurate results, it is often costly, time-consuming, and impractical for rapid or large-scale deployment in the field [3], [4]. These limitations reduce the effectiveness of traditional approaches in supporting real-time disease monitoring in poultry farms.

Recent advances in artificial intelligence, particularly deep learning, have provided promising solutions for automated disease detection. Convolutional Neural Networks (CNNs) have demonstrated strong performance in image-based classification tasks due to their ability to automatically extract meaningful visual features [5], [6]. Several studies have applied CNN architectures such as VGG16, InceptionV3, MobileNet, and EfficientNet for poultry disease detection and achieved high accuracy [1], [7], [8]. However, most previous research focuses on a limited number of models and uses different experimental settings, making objective comparison difficult [3], [9].

To address this research gap, this study conducts a comparative evaluation of five CNN architectures—VGG16, ResNet50, InceptionV3, MobileNetV2, and EfficientNetB0—using a consistent dataset and experimental framework. Fecal images are used as the input data because they are non-invasive and provide valuable indicators of chicken health conditions [8], [4]. The models are evaluated based on accuracy, macro precision, macro recall, macro F1-score, and inference runtime, with the aim of identifying the most optimal architecture for practical poultry disease detection systems.

Therefore, the objectives of this study are:

(1) To compare the performance of five Convolutional Neural Network (CNN) architectures, namely VGG16, ResNet50, InceptionV3, MobileNetV2, and EfficientNetB0, for chicken disease classification using fecal images; and

(2) To identify the most optimal model based on classification accuracy and inference runtime under identical experimental conditions.

2. RESEARCH METHODOLOGY

This study adopts an experimental comparative research design to evaluate the performance of multiple Convolutional Neural Network (CNN) architectures for chicken disease classification using fecal images. A comparative approach is suitable for assessing the effectiveness and efficiency of different models under identical experimental conditions [10], [11].

Five CNN architectures were selected for evaluation, namely VGG16, ResNet50, InceptionV3, MobileNetV2, and EfficientNetB0. These models represent a diverse range of structural characteristics, from deep sequential networks (VGG16) to lightweight mobile-oriented architectures (MobileNetV2) and compound-scaled networks (EfficientNetB0).

All experiments were conducted using the same dataset, preprocessing techniques, data splitting strategy, and evaluation metrics to ensure fair and objective performance comparison. The overall research workflow is illustrated in **Figure 1**.

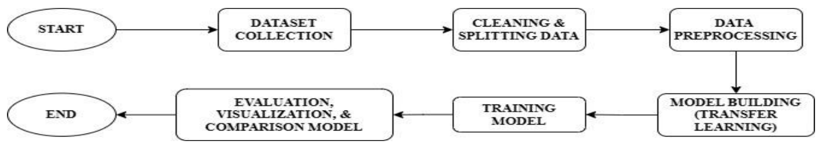

Figure 1. Research Metodology

2.1 Dataset Collection

The dataset used in this study is titled “Chicken Disease Image Classification” and was obtained from the Kaggle platform. It consists of 8,067 chicken fecal images categorized into four classes: Healthy, Newcastle Disease (ND), Coccidiosis, and Salmonella. The dataset used in this study is publicly available on Kaggle and can be accessed at: https://www.kaggle.com/datasets/allandclive/chicken-disease-1

The selection of fecal images as the diagnostic object is motivated by their non-invasive nature and their strong correlation with chicken health conditions, as validated in previous studies [8], [4]. Compared to laboratory-based testing, fecal image analysis offers faster and more practical disease screening in large-scale poultry farms.

2.2 Dataset cleaning and splitting

Before model development, a data cleaning process was conducted to ensure dataset quality. This process involved removing duplicate samples, corrupted files, and visually distorted images that could introduce noise into the training process. Previous studies have shown that low-quality image samples can negatively affect CNN generalization performance [12].

After cleaning, the dataset was divided into three subsets: training, validation, and testing with proportions of 70%, 15%, and 15%, respectively. A stratified splitting strategy was applied to maintain balanced class distributions across all subsets. This stratified strategy ensures that each class is fairly represented in all subsets, thereby reducing evaluation bias and improving result reliability.

This partitioning strategy is widely adopted in image classification studies to ensure sufficient training data while preserving unbiased model evaluation [11].

2.3 Data Preprocessing and Augmentation

All images underwent a preprocessing pipeline consisting of resizing, normalization, and data augmentation. First, all images were resized to 224 × 224 pixels, which corresponds to the standard input size of pretrained CNN architectures such as VGG16, ResNet50, and EfficientNetB0.

Next, normalization was performed using the *preprocess_input()* function specific to each model architecture. This step aligns the pixel distribution with the original ImageNet dataset, which has been shown to accelerate convergence in transfer learning scenarios [13].

To improve generalization and reduce overfitting, data augmentation was applied exclusively to the training set. The applied augmentation techniques included:

- Random rotation up to ±20° [14]
- Zooming up to 0.2 [15]
- Horizontal flipping [15]

These strategies are consistent with the findings of Shorten and Khoshgoftaar, who demonstrated that rotation, zooming, and flipping are among the most effective augmentation methods for enhancing CNN robustness [12].

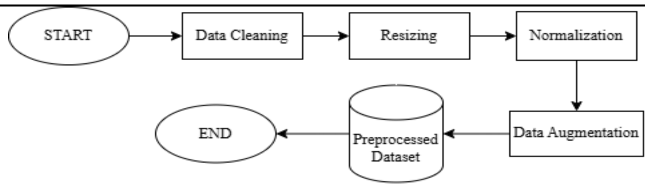

Figure 2. Image preprocessing and augmentation workflow.

2.4 Transfer Learning and Model Architecture

This study employs a transfer learning strategy using pretrained CNN models trained on the ImageNet dataset. Transfer learning enables the reuse of generic visual features such as edges, textures, and object shapes, which significantly improves performance when working with limited domain-specific datasets [10], [16].

In the fine-tuning phase, most layers of each pretrained model were frozen, and only the final 20 layers were unfrozen and retrained. This approach allows the models to adapt to domain-specific patterns in chicken feces images while maintaining training stability.

The original classification layers were removed and replaced with:

- A Global Average Pooling layer to reduce dimensionality and overfitting [10],
- A fully connected layer with 128 neurons and ReLU activation,
- An output layer with four neurons and Softmax activation.

2.5 Model Training

Three optimization algorithms were evaluated in this study: Adam, Stochastic Gradient Descent (SGD), and RMSprop. Learning rates ranged from 1×10^{-3} to 1×10^{-5}, batch sizes varied between 16, 32, and 64, and training epochs ranged from 10 to 20.

Early stopping and model checkpoint mechanisms were implemented to prevent overfitting and preserve the best-performing model weights. This training strategy has been widely adopted in deep learning-based agricultural image classification studies [8], [15].

2.6 Hyperparameter Configuration

Multiple hyperparameter combinations were explored to identify the optimal configuration for each CNN architecture. The tested hyperparameters are summarized in **Table 1**.

Table 1. Hyperparameter settings.

Hyperparameter	Tested Values
Learning rate	1E-3, 1E-4, 1E-5
Batch size	16, 32, 64
Epochs	10, 15, 20
Optimizer	Adam, SGD, RMSprop
Activation	ReLU, ELU, LeakyReLU
Training ratio	50%, 60%, 70%, 80%

The final analysis focuses exclusively on the best-performing configuration for each model to maintain clarity and avoid redundant comparisons.

2.7 Inference Runtime Measurement

Inference runtime was measured to evaluate computational efficiency during the prediction phase. Inference time is defined as the average duration required by a trained model to generate a prediction for a single test image.

All runtime measurements were conducted using identical computing environments on the Kaggle platform with NVIDIA Tesla T4 GPUs. A warm-up phase was applied prior to measurement to stabilize GPU operations.

This metric is essential for assessing model suitability in real-time disease detection systems, where rapid response and limited computational resources are critical [8], [12].

3. RESULTS AND DISCUSSION

3.1 Classification Performance

This section presents the experimental results obtained from the five CNN architectures evaluated in this study. The performance comparison is summarized in **Table 2**.

All models achieved strong classification performance, with accuracy values exceeding 96%, indicating that CNN-based transfer learning is highly effective for chicken disease classification using fecal images.

Figure 3 presents the confusion matrix of the ResNet50 model. The model correctly classified most samples in all four classes, with particularly high accuracy for Salmonella (261 correct predictions) and Coccidiosis (240 correct predictions).

Minor misclassifications mainly occurred between Coccidiosis and Healthy classes, as well as between Healthy and Newcastle Disease, which may be caused by visual similarities in fecal texture and color patterns. Overall, the confusion matrix confirms the robustness and stability of ResNet50 in distinguishing chicken disease categories. Similar misclassification patterns were also observed in other models, particularly between Healthy and Coccidiosis classes, indicating inherent visual similarity in fecal textures.

Table 2. Performance comparison of CNN models

Model	Train:Val:Test	Learning Rate, Batch Size, Epoch	Optimizer, Activation	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Runtime Inference (second)
VGG16	80:10:10	0.00001, 64, 20	RMSprop, Relu	97.52	97.03	97.34	97.32	0.0706
ResNet50	80:10:10	0.0001, 32, 20	Adam, Relu	97.52	97.62	96.59	97.09	0.0721
EfficientNetB0	70:15:15	0.001, 16, 15	RMSprop, Relu	97.19	97.25	96.58	96.09	0.0618
MobileNetV2	80:10:10	0.001, 16, 15	Adam, Leaky	96.28	97.01	94.93	95.89	0.0638
InceptionV3	80:10:10	0.001, 16, 15	Adam, Leaky	96.16	95.29	93.09	94.09	0.0751

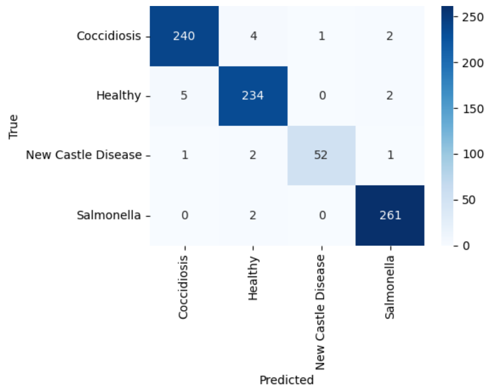

Figure 3. Confusion matrix of ResNet50 classification results.

3.2 Model-wise Performance Analysis

Among all evaluated models, ResNet50 was selected as the primary reference model for further analysis because it provides the best balance between accuracy and computational efficiency. In addition, ResNet50 demonstrates stable performance across all evaluation metrics and shows robust generalization capability as indicated by its confusion matrix results.

VGG16 achieved the highest recall (97.34%) and F1-score (97.32%), indicating its strong ability to correctly identify disease classes. The deep and sequential structure of VGG16 enables effective hierarchical feature extraction, which is consistent with findings reported in previous poultry disease studies [5], [10].

ResNet50 obtained the highest precision (97.62%) and matched the highest accuracy (97.52%). The use of residual connections allows stable gradient propagation and prevents performance degradation in deeper networks, which explains its robust and consistent learning behavior [10], [16].

EfficientNetB0 demonstrated competitive accuracy (97.19%) while achieving the fastest inference runtime (0.0618 seconds). This confirms the effectiveness of compound scaling in producing lightweight models with high performance, making EfficientNetB0 particularly suitable for real-time and mobile-based applications [17].

MobileNetV2 achieved lower accuracy compared to VGG16 and ResNet50 but maintained relatively fast inference speed. Its depthwise separable convolution significantly reduces computational complexity, which aligns with previous studies focusing on mobile-friendly deep learning models [8], [4].

InceptionV3 produced the lowest performance among the tested models. Although its multi-scale convolution structure is powerful for complex image patterns, it appears less effective for the relatively simple texture-based features present in fecal images [4].

3.3 Accuracy vs Efficiency Trade-off

The experimental results reveal a clear trade-off between classification accuracy and computational efficiency. VGG16 and ResNet50 offer the highest accuracy, but require longer inference time. In contrast, EfficientNetB0 provides slightly lower accuracy with significantly faster inference speed.

This trade-off is critical in real-world poultry monitoring systems, where both prediction accuracy and response time must be considered. In resource-limited environments, such as small-scale farms or edge devices, EfficientNetB0 becomes a more practical choice [17], [12].

4. CONCLUSION

This study conducted a comprehensive comparative evaluation of five CNN architectures—VGG16, ResNet50, InceptionV3, MobileNetV2, and EfficientNetB0—for automatic chicken disease classification using fecal images.

Experimental results demonstrate that all models achieve high performance, with accuracy exceeding 96%. VGG16 and ResNet50 provide the highest classification accuracy (97.52%), while EfficientNetB0 offers the fastest inference runtime (0.0618 seconds).

Based on both accuracy and computational efficiency, ResNet50 is identified as the most balanced model, making it the most suitable architecture for practical poultry disease detection systems. EfficientNetB0 is recommended for real-time and mobile applications, whereas VGG16 is more appropriate for high-performance environments with sufficient computational resources.

Limitations. Although the results are promising, the dataset is limited to fecal images from a single public source, which may not fully represent real-world variations. Moreover, only four disease categories are considered.

Future Work. Future studies may integrate multimodal data such as sound and behavioral information, and explore model compression techniques for deployment on embedded and edge computing platforms.

ACKNOWLEDGMENT

The author would like to express sincere gratitude to her parents for their continuous support and encouragement throughout the completion of this research. The author also extends her appreciation to the lecturers of the Informatics Engineering Program, Soegijapranata Catholic University, for their valuable guidance, insights, and academic support during the research process. Finally, the author would like to thank her friends and colleagues for their constructive discussions, motivation, and collaboration that contributed to the successful completion of this study.

REFERENCES

[1] Dwicahyo A, Mufandi I, Nurfadila AR, et al. Early Detection of Disease in Chicks Using CNN on Bangkok Chicken Health. *Buletin Ilmiah Sarjana Teknik Elektro* 2024; 6: 126–141.

[2] R MsL, P MsH, M MsR. Classification and Detection of Chicken Disease Using CNN with Image Classification Technique. *Int J Eng Res Appl* 2024; 14: 126–129.

[3] Al Qatrawi M, Abu-Naser SS. Classification of Chicken Diseases Using Deep Learning, www.ijeais.org/ijaisr.

[4] Kholil M, Priya Waspada H, Akhsani R, et al. Klasifikasi Penyakit Infeksi Pada Ayam Berdasarkan Gambar Feses Menggunakan Convolutional Neural Network, https://doi.org/10.31598.

[5] Schmidhuber J. Deep Learning in Neural Networks: An Overview. Epub ahead of print 8 October 2014. DOI: 10.1016/j.neunet.2014.09.003.

[6] Kim KG. Book Review: Deep Learning. *Healthc Inform Res* 2016; 22: 351.

[7] Muhammad Salimy Ahsan, Kusrini Kusrini, Dhani Ariatmanto. Penerapan model InceptionV3 dalam klasifikasi penyakit ayam. *JNANALOKA* 2023; 55–62.

[8] Mustopa A, Sasongko A, Mahmud Nawawi H, et al. *SISTEMASI: Jurnal Sistem Informasi Deteksi Penyakit Ayam berdasarkan Citra Feses dengan Model EfficientNetV2L Chicken Disease Detection based on Fases Image using EfficientNetV2L Model*, http://sistemasi.ftik.unisi.ac.id.

[9] Hasan MohA, Riyanto Y, Riana D. Grape leaf image disease classification using CNN-VGG16 model. *Jurnal Teknologi dan Sistem Komputer* 2021; 9: 218–223.

[10] Shin HC, Roth HR, Gao M, et al. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Trans Med Imaging* 2016; 35: 1285–1298.

[11] Du X, Sun Y, Song Y, et al. A Comparative Study of Different CNN Models and Transfer Learning Effect for Underwater Object Classification in Side-Scan Sonar Images. *Remote Sens (Basel)*; 15. Epub ahead of print 1 February 2023. DOI: 10.3390/rs15030593.

[12] Shorten C, Khoshgoftaar TM. A survey on Image Data Augmentation for Deep Learning. *J Big Data*; 6. Epub ahead of print 1 December 2019. DOI: 10.1186/s40537-019-0197-0.

[13] Perez L, Wang J. The Effectiveness of Data Augmentation in Image Classification using Deep Learning, http://arxiv.org/abs/1712.04621 (2017).

[14] Boulent J, Foucher S, Théau J, et al. Convolutional Neural Networks for the Automatic Identification of Plant Diseases. *Frontiers in Plant Science*; 10. Epub ahead of print 23 July 2019. DOI: 10.3389/fpls.2019.00941.

[15] Mafukidze HD, Owomugisha G, Otim D, et al. Adaptive Thresholding of CNN Features for Maize Leaf Disease Classification and Severity Estimation. *Applied Sciences (Switzerland)*; 12. Epub ahead of print 1 September 2022. DOI: 10.3390/app12178412.

[16] Li C, Zhen T, Li Z. Image Classification of Pests with Residual Neural Network Based on Transfer Learning. *Applied Sciences (Switzerland)*; 12. Epub ahead of print 1 May 2022. DOI: 10.3390/app12094356.

[17] Tan M, Le Q V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, http://arxiv.org/abs/1905.11946 (2020).