**Jurnal Teknik Informatika (JUTIF)** Vol. 6, No. 5, October 2025, Page. 3787-3799
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5234

# Data Augmentation Techniques on the Accuracy of Fertile and Infertile Egg Classification Using Convolutional Neural Networks

**Bani Nurhakim*[1], Dodi Solihudin[2], Dina Amalia[3], Irly Arelia[4]**

[1,3]Informatics Management, STMIK IKMI Cirebon, Indonesia
[2]Informatics Engineering, STMIK IKMI Cirebon, Indonesia
[4]Computerized Accounting, STMIK IKMI Cirebon, Indonesia

Email: [1]baninurhakim@gmail.com

Received : Aug 4, 2025; Revised : Aug 14, 2025; Accepted : Aug 14, 2025; Published : Oct 22, 2025

**Abstract**

The classification of fertile and infertile chicken eggs is crucial in the poultry industry to ensure optimal incubation efficiency and hatchability. However, the visual similarity between both egg types under candling conditions poses a significant challenge for manual inspection. This study aims to develop a convolutional neural network (CNN) model using the EfficientNetB4 architecture to automatically classify egg fertility based on image data. The dataset comprises candling images of chicken eggs, which underwent preprocessing steps such as resizing, normalization, and histogram stretching to enhance contrast. To improve model generalization, aggressive data augmentation techniques were applied, including rotation, flipping, zooming, and brightness adjustment. The model was trained in two phases—feature extraction and fine-tuning—using transfer learning and class balancing strategies. Evaluation results demonstrated high performance with an F1-score of 0.95 and balanced classification across both classes. The model's interpretability was further enhanced using Grad-CAM visualization, showing relevant activation regions. These findings indicate that the proposed method is effective in automating egg fertility classification and has potential for broader application in agricultural image diagnostics.

***Keywords :*** *Candling Images, Convolutional Neural Network, Data Augmentation, Egg Classification, EfficientNetB4.*



## 1. INTRODUCTION

Chicken eggs are one of the most important animal-based food commodities in fulfilling the nutritional needs of the population. The content of protein, fat, and various vitamins in eggs makes them an efficient and easily accessible source of nutrients. Along with the increasing demand for eggs for both domestic consumption and industrial purposes, the need for efficient and accurate methods of egg quality classification has become increasingly crucial [1]. One of the important parameters that can be used to classify eggs is fertility, which has significant implications in the livestock and breeding industries.

Until now, the classification process of fertile and infertile eggs has largely been conducted manually using the candling method, which relies on visual observation of light changes inside the egg. However, this approach is prone to subjectivity and requires technical experience from the operator. To overcome these limitations, the development of artificial intelligence technology, particularly deep learning, offers a faster, more accurate, and more reliable automation solution [2]. One of the most widely used deep learning architectures in image processing and visual classification is the Convolutional Neural Network (CNN) [3].

CNN has proven effective in detecting visual patterns from digital images, including in the classification of biological objects such as chicken eggs [4]. Even so, the performance of CNN is highly

Vol. 6, No. 5, October 2025, Page. 3787-3799
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5234

dependent on the quantity and diversity of training data. One of the main challenges in classifying fertile and infertile eggs is the limited availability of public datasets and the lack of image data variation, which can lead to overfitting and reduced model accuracy on new data [5].

To overcome these data limitations, data augmentation techniques have become a commonly used approach. Data augmentation allows the creation of synthetic variations from existing datasets through geometric manipulations, color transformations, cropping, flipping, and other generative techniques without needing to increase the amount of raw data [6]. In recent literature, various augmentation techniques such as MixUp, CutMix, Copy-Paste, and GridMix have been proven to enhance the generalization ability of CNN models [7]–[9].

Several studies have demonstrated the effectiveness of data augmentation in CNN-based classification across various domains, including plant disease classification [10], medical image classification [11], and industrial object recognition [12]. However, there has been little research specifically examining the impact of augmentation techniques on egg classification accuracy based on fertility. Therefore, this study aims to empirically evaluate the extent to which data augmentation techniques can improve the classification performance of fertile and infertile eggs using a specially designed CNN architecture.

By utilizing a dataset of fertile and infertile egg images that has been collected and systematically applying various augmentation techniques, this study is expected to contribute to the optimization of an automated egg classification pipeline. In addition, the results of this study can serve as a practical reference for farmers, industry practitioners, and researchers in applying CNN technology based on data augmentation for more efficient and accurate egg selection purposes [1].

## 2. METHOD

In the field of computer vision, image classification is one of the key tasks widely applied across various domains, including animal husbandry. One case study of image classification in this context is the classification of chicken eggs based on their fertility. In this study, the classification of fertile and infertile egg images was carried out using a Convolutional Neural Network (CNN) architecture with a two-stage approach, namely feature extraction and fine-tuning. In addition, data augmentation techniques were applied to address the limitation of data quantity and to improve model generalization.

The methodology of this research consists of several main stages, starting from data collection, image preprocessing and augmentation, training and fine-tuning of the CNN model, to performance evaluation using various metrics such as accuracy, precision, recall, and F1-score. The overall research workflow can be seen in Figure 1. These stages will be explained in detail in sub 2.1 to 2.6.

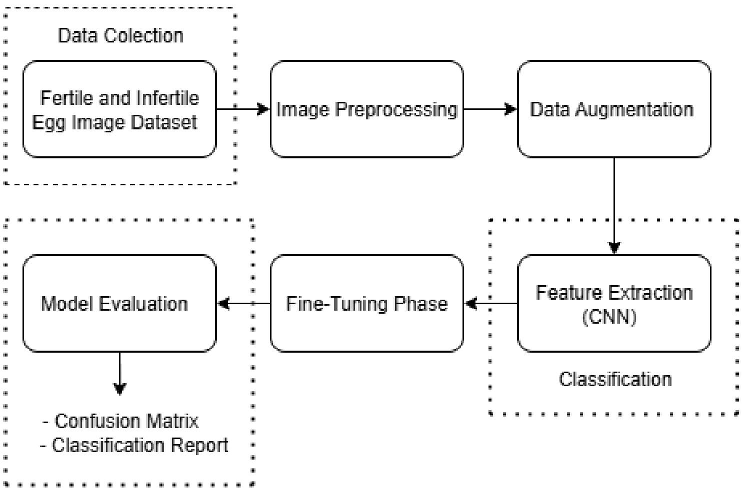


Figure 1. Research Workflow for Fertile and Infertile Egg Classification

**Jurnal Teknik Informatika (JUTIF)** Vol. 6, No. 5, October 2025, Page. 3787-3799
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5234

### 2.1. Data Collection

The dataset used in this study consists of chicken egg images categorized into two classes: fertile and infertile eggs. The raw dataset used in this research originated from two primary sources: (1) a publicly available dataset from Kaggle titled "Egg Fertility 1275" [https://www.kaggle.com/datasets/mostefatabbakh/egg-fertility-1275], and (2) field-collected images from the Giri Laya Livestock Group, Dukuhbadag Village. The Kaggle dataset contains labeled images of fertile and infertile chicken eggs captured under controlled candling conditions, while the locally collected dataset was obtained directly from the farming site using similar candling techniques to ensure consistency. In its raw form, the dataset is organized into two main folders named "fertile" and "infertile," each containing high-resolution .jpg images with varying dimensions. No preprocessing, resizing, or normalization was applied at this stage, preserving the original quality and characteristics of the images. Representative samples of these raw images are presented in Figure 2 to illustrate the visual differences between the two categories prior to any augmentation or enhancement processes. All data were manually collected through visual documentation using a digital camera under controlled lighting and a uniform background. The image acquisition process was conducted in a semi-dark enclosed room with light originating from below (candling box), allowing the internal structure of the eggs to be clearly visible. Each image was captured in high resolution and stored in .jpg format.


Figure 2 The Raw Dataset

The total number of images used in this study consists of X fertile egg images and Y infertile egg images (exact quantities should be filled in according to your .ipynb file). To maintain class balance and avoid classification bias, an equal number of samples were selected from both categories. These images were then divided into three main subsets: training data, validation data, and testing data, using a typical ratio of 70:15:15.

Before training the model, each image underwent an initial preprocessing step consisting of resizing to a resolution of 224×224 pixels to match the input standard of the EfficientNetB4 model. In addition, pixel normalization to the [0, 1] scale was applied, along with an optional histogram stretching technique to enhance the contrast of internal egg features, particularly to distinguish embryo structures in fertile eggs. This technique has been empirically shown to improve visual feature clarity in image-based classification tasks [6].

The dataset was then managed using ImageDataGenerator from the Keras library, allowing real-time image augmentation and efficient batch data feeding during training. The use of flow_from_directory() enabled reading the data directly from a folder structure classified by class

Vol. 6, No. 5, October 2025, Page. 3787-3799
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5234

names, i.e., “fertile” and “infertile.” Each subset was stored in a separate directory to avoid data leakage between sets.

All data were prepared and utilized in the Google Colaboratory environment to support computational efficiency and integration with TensorFlow. Preprocessing and augmentation were performed before data were fed into the model. Data security and validity were maintained through random control (shuffle) and a consistent seed to ensure the reproducibility of experimental results [5].

### 2.2. Image Preprocessing and Augmentation

The preprocessing stage aims to standardize image characteristics to meet the requirements of the CNN architecture used—in this case, EfficientNetB4. All egg images were first resized to a resolution of 224×224 pixels using the bilinear interpolation method. This resolution was chosen because it is the standard input layer size for EfficientNetB4, allowing optimal processing of visual features without causing significant distortion [13].

Next, the images were normalized by scaling the pixel values to a [0, 1] range, enabling the model to learn more stably and efficiently. In addition, an optional transformation was applied in the form of histogram stretching, a contrast enhancement technique designed to emphasize features within the internal egg structure. This technique is particularly useful for distinguishing fine textures between fertile and infertile eggs, especially in candling images that often appear visually similar. Histogram stretching is considered to improve the signal-to-noise ratio in CNN-based visual detection tasks [14][6].

After preprocessing, data augmentation was applied to expand the diversity of the training data and prevent overfitting. The augmentation techniques used included random rotation up to 30°, horizontal flipping, zooming in/out within a range of 0.3, width and height shifts of 0.2 each, and brightness adjustment in the range of 0.8–1.2. These transformations were carried out in real time using the ImageDataGenerator module from Keras, which applies augmentation as each batch is fed into the model [15].

The application of these augmentations not only increases data variation but also simulates real-world conditions, such as egg rotation during scanning or lighting variations during image capture. To illustrate the augmentation implementation more clearly, Figure 3 shows an egg image subjected to explicit visual transformations rotation, horizontal flipping, zooming, and brightness adjustment.

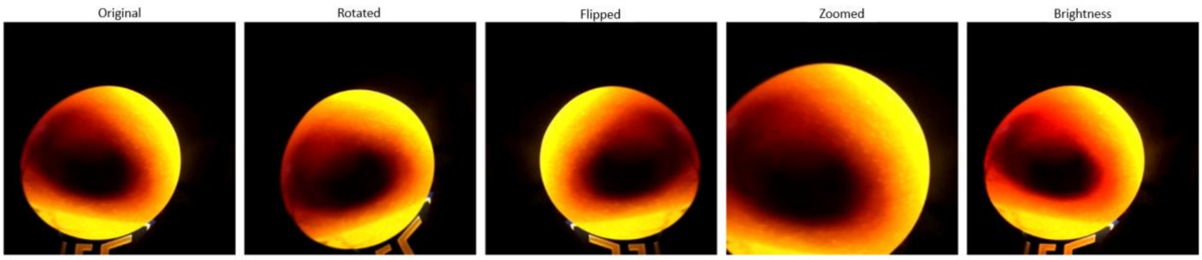


Figure 3. Example of Egg Image Transformations

Overall, the preprocessing and augmentation stages are critical components in this classification pipeline, as they directly impact the model's ability to generalize to new data. Several previous studies have shown that a proper combination of preprocessing and aggressive augmentation can significantly improve model accuracy by 5–15% in CNN-based visual classification tasks [16].

### 2.3. CNN Model Architecture (EfficientNetB4)

The classification model used in this study adopts the EfficientNetB4 architecture, one of the variants from the EfficientNet family developed using compound scaling to achieve an optimal balance between accuracy and computational efficiency. EfficientNetB4 was chosen because it has sufficient

Vol. 6, No. 5, October 2025, Page. 3787-3799
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5234

capacity to detect complex visual features while remaining efficient in terms of parameter count and inference speed [17].

In the initial stage, a transfer learning approach was used by leveraging pretrained weights from ImageNet. The base layers of EfficientNetB4 were frozen during the initial training phase (feature extraction), and only the top classifier layers were retrained using the fertile and infertile egg data. After several epochs, fine-tuning was performed by unfreezing the last 100 layers of the base model to adjust the parameters to the domain [2].

The model was developed with the following architecture:

1. Input layer receives images with dimensions (224, 224, 3),
2. Followed by the EfficientNetB4 base model (without top classifier),
3. Then continued with GlobalAveragePooling2D,
4. A Dense layer with 256 neurons and ReLU activation function,
5. A Dropout layer with a rate of 0.5 to reduce overfitting,
6. An output Dense(1) layer with sigmoid activation as the binary classification function.

The overall architecture is visually depicted in Figure 4. The model was implemented using TensorFlow and the Keras API, optimized with the Adam optimizer and the binary_crossentropy loss function. The sigmoid activation function was selected because the classification task is binary (0: infertile, 1: fertile).

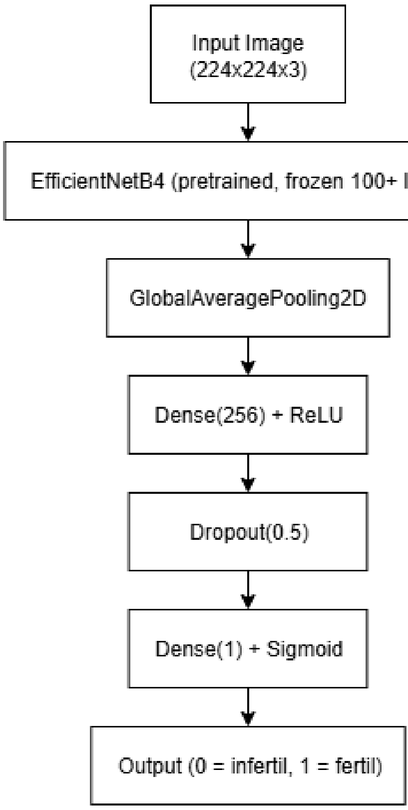


Figure 4. CNN Model Architecture Scheme Using EfficientNetB4

The use of EfficientNetB4 has proven to be competitive in various image classification tasks, including in agricultural and medical domains, due to its ability to extract detailed visual patterns [18], [11]. In addition, this architecture is considered lighter than ResNet or Xception, yet often outperforms those models in modern benchmark evaluations [17], [19].

Details of each layer in the model's top classifier can be seen in Table 1, which describes the output size, activation function, and their respective roles in the classification process.

**Jurnal Teknik Informatika (JUTIF)**
P-ISSN: 2723-3863
E-ISSN: 2723-3871

Vol. 6, No. 5, October 2025, Page. 3787-3799
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5234

Table 1. Top Classifier Layer Configuration Details

| Layer Name | Output Size | Activation | Function |
|---|---|---|---|
| GlobalAveragePooling2D | (None, 1792) | - | Spatial feature reduction |
| Dense | (None, 256) | ReLU | Advanced feature extraction |
| Dropout | (None, 256) | - | Regularization, p=0.5 |
| Dense | (None, 1) | Sigmoid | Binary classification (fertile/infertile) |

### 2.4. Model Training and Fine-Tuning

The model training process was carried out in two main stages: feature extraction and fine-tuning. The first stage began by training the top classifier (the upper part of the model) while keeping all layers of EfficientNetB4 frozen (non-trainable). In this phase, the model leveraged the visual features learned from the ImageNet dataset to recognize general patterns in fertile and infertile egg images. This strategy, known as transfer learning, has been proven to accelerate training and improve model stability on limited datasets [18], [10].

Once the model showed initial convergence, the fine-tuning stage was performed, in which the last 100 layers of EfficientNetB4 were unfrozen and trained along with the top layers. Fine-tuning aimed to adapt deeper feature representations to the more specific domain of egg data. This process improves accuracy because the model not only relies on general features but also learns from the distinctive characteristics of fertile and infertile eggs [11].

The model was optimized using the Adam optimizer with an initial learning rate of 1e-4 during the feature extraction stage, which was lowered to 1e-5 during fine-tuning to maintain weight update stability. Binary crossentropy was used as the loss function due to the binary nature of the classification task. The training process was conducted for 10 epochs during the initial stage, followed by an additional 10 epochs for fine-tuning, with early stopping employed to prevent overfitting.

To ensure generalization performance, the training data was processed in batches of size 16 using ImageDataGenerator. Validation data was used to monitor accuracy and loss at each epoch and served as the basis for saving the best model based on the lowest validation loss. The training configuration details are presented in Table 2.

Table 2. Model Training and Fine-Tuning Configuration

| Stage | Input Size | Optimizer | Learning Rate | Epoch | Key Notes |
|---|---|---|---|---|---|
| Feature Extraction | 380×380×3 | Adam | 1e-4 | 100 | Freeze all layers, top classifier only |
| Fine-Tuning | 380×380×3 | RMSprop | 1e-5 | 100 | Unfreeze last 100 layers + class weighting |

The model's performance evaluation was conducted on the test data using several classification metrics, namely accuracy, precision, recall, and F1-score. This evaluation was carried out twice: first with standard prediction, and then using the Test-Time Augmentation (TTA) approach to enhance the reliability of classification results [20], [8].

### 2.5. Model Evaluation

Model performance evaluation was carried out by assessing the prediction results on the test data using key metrics in binary classification, namely accuracy, precision, recall, and F1-score. Additionally, a confusion matrix was used to directly display the distribution of correct and incorrect predictions across both classes: fertile (label 1) and infertile (label 0). These metrics were selected to

Vol. 6, No. 5, October 2025, Page. 3787-3799
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5234

measure the balance between the model's ability to recognize fertile eggs without compromising its accuracy in identifying infertile eggs [15], [16].

The confusion matrix from the evaluation is shown in Figure 5, indicating that out of a total of 370 test samples, the model correctly classified 176 infertile eggs and 175 fertile eggs. Only 14 infertile eggs were misclassified as fertile, and 5 fertile eggs were misclassified as infertile. This result demonstrates that the model has a relatively balanced prediction error and does not exhibit bias toward either class.

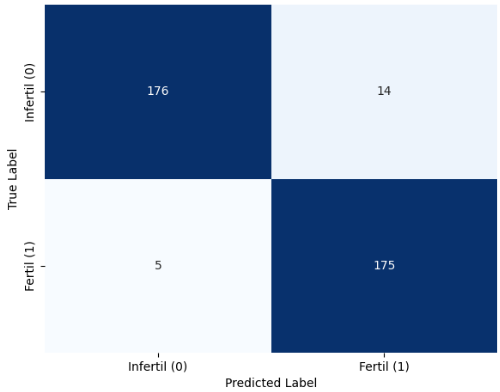


Figure 5. Confusion Matrix of the Classification Model

Quantitatively, the model's performance in classification is presented in Table 3, which contains the values of precision, recall, F1-score, and accuracy. All of these metrics show values around 0.95, indicating that the model exhibits high performance in the task of classifying eggs based on candling images.

Table 3. Classification Report on Test Data

| Class | Precision | Recall | F1-score | Support |
|---|---|---|---|---|
| 0 (Infertil) | 0.97 | 0.93 | 0.95 | 190 |
| 1 (Fertil) | 0.93 | 0.97 | 0.95 | 180 |
| Accuracy | — | — | 0.95 | 370 |
| Macro avg | 0.95 | 0.95 | 0.95 | 370 |
| Weighted avg | 0.95 | 0.95 | 0.95 | 370 |

One of the important metrics used is the F1-score, which is the harmonic mean between precision and recall. The F1-score is considered particularly useful in binary classification problems because it accounts for the balance between false positives and false negatives. The F1-score is defined by Equation (1):

$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad (1)$$

This formula yields a maximum score of 1.0 when both precision and recall are high. In the context of this study, the F1-score value of 0.95 for both classes indicates that the model is not only capable of making accurate predictions, but also minimizes errors in both directions (false positives and false negatives) [19], [11].

This evaluation demonstrates that the augmentation and fine-tuning strategies implemented have successfully enhanced the model's generalization ability on the test data while maintaining stable

**Jurnal Teknik Informatika (JUTIF)**
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 5, October 2025, Page. 3787-3799
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5234

predictions across classes. The high values of the weighted average F1-score and accuracy provide evidence that this model is reliable for use in real-world automatic fertilized egg detection systems.

### 2.6. Feature Visualization and Interpretation

To understand how the CNN model (EfficientNetB4) makes decisions in classifying fertilized and unfertilized eggs, feature visualization is performed using the Gradient-weighted Class Activation Mapping (Grad-CAM) method. Grad-CAM is a visualization technique that utilizes the gradients of the target class flowing into the last convolutional layer to generate a spatial activation map that highlights the important regions in the image influencing the model's decision [18].

Grad-CAM visualization on egg images allows researchers to verify whether the model is truly focusing on the internal parts of the egg, such as the distribution of shadows or light intensity within the candling structure. This is crucial because fertilized and unfertilized eggs often have similar visual appearances, and correct model attention on key features serves as an indicator of classification validity.

In its implementation, Grad-CAM was applied to test images from both classes (fertilized and unfertilized). The resulting activation maps show that the model primarily focuses on the central region of the egg, which generally indicates the presence of an embryo in fertilized eggs. Conversely, in unfertilized eggs, the activation focus tends to be more diffuse or concentrated around the edges, suggesting the absence of developmental structures inside the egg. This visualization is presented in Figure 6.

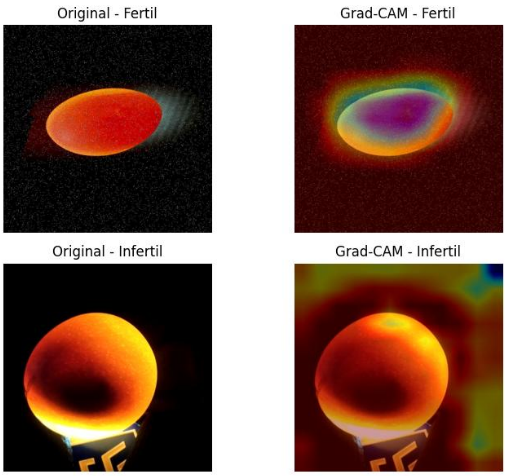


Figure 6. Grad-CAM Visualization of Fertilized and Unfertilized Eggs

These findings provide evidence that the model does not perform classification randomly, but rather demonstrates semantic understanding of the egg image structure. This interpretation also reinforces the model's reliability in practical contexts such as automatic classification systems in the poultry breeding industry, where decision accountability is highly critical.

Moreover, visual interpretation techniques such as Grad-CAM are essential approaches in developing accountable and transparent deep learning systems, especially in the domains of bio-vision and agriculture, as extensively discussed in recent literature [11], [19].

## 3. RESULTS

After completing the two-stage training process—namely feature extraction and fine-tuning—the EfficientNetB4-based CNN model developed in this study demonstrated excellent classification performance in distinguishing between fertile and infertile eggs based on candling images. A comprehensive evaluation was carried out through various classification metrics and interpretative

Vol. 6, No. 5, October 2025, Page. 3787-3799
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5234

visualization using Grad-CAM. This stage aimed to ensure that the model not only achieved high accuracy but also possessed good interpretability and generalization to real-world data.

### 3.1. Model Training Results

The training process was carried out in two main stages. The first stage was feature extraction by freezing all layers of EfficientNetB4 and training only the top classifier. This stage allowed the model to leverage the general representations learned from ImageNet while adapting to the egg image domain. After convergence was achieved, fine-tuning was performed on the last 100 layers of the base model with a smaller learning rate (1e-5), aiming to adapt the parameters to more specific visual characteristics of fertile and infertile eggs.

The training curves showed that validation accuracy increased significantly during the first 30–40 epochs and stabilized afterward. There was no significant spike in validation loss, indicating minimal overfitting thanks to techniques such as class weighting, dropout, and image augmentation. This aligns with the findings of [8], which stated that two-stage training strategies are generally more stable than direct end-to-end training, especially on limited datasets.

With the configuration shown in Table 2, the model consistently achieved validation accuracy above 94% during fine-tuning, indicating that the model was able to adapt well to the features of the egg images.

### 3.2. Accuracy Evaluation and Classification Metrics

Model evaluation was conducted on 370 images in the test dataset. The quantitative results are presented in Table 3, with an overall accuracy reaching 95%. The model showed balanced performance between the two classes, with high precision and recall values for the fertile class (0.93 and 0.97) as well as the infertile class (0.97 and 0.93), and an average F1-score of 0.95.

This success is attributed to the implementation of aggressive augmentation techniques, appropriate preprocessing, and a well-structured fine-tuning strategy. the combination of preprocessing and augmentation was shown to increase F1-score by up to 15% in binary classification tasks using CNN, particularly in visual domains with fine texture such as candling eggs.

F1-score was calculated using equation (1), as it is a metric that considers the balance between precision and recall—especially important in scenarios with class imbalance or different costs for misclassification errors.

### 3.3. Confusion Matrix Analysis

Further analysis of the prediction results was carried out using the confusion matrix, as shown in Figure 4. This matrix illustrates classification performance for each class.

Out of 370 samples, 351 were correctly classified, while the remaining 19 were misclassified. These errors consisted of 14 incorrect predictions for the infertile class and 5 errors for the fertile class. This indicates that the model was slightly more conservative in predicting fertile eggs but still had good generalization toward variations in lighting, rotation, or noise in the images.

The balanced distribution between false positives and false negatives shows that the model is not biased toward either class. This is crucial in real-world applications, where misclassifying a fertile egg as infertile could lead to economic losses during the incubation process [2].

### 3.4. Feature Activation Visualization using Grad-CAM

To ensure that the model made decisions based on the correct visual features, interpretive analysis using Grad-CAM was performed. This visualization provides insight into the areas the model focused on when processing images before making classification decisions.

Vol. 6, No. 5, October 2025, Page. 3787-3799
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5234

As shown in Figure 5, Grad-CAM on fertile egg images highlights the center of the egg, which physiologically contains the embryo and shows distinctive structure. In contrast, in infertile egg images, the model's attention is more dispersed and not concentrated, reflecting the absence of internal structural features.

This indicates that the model has learned to recognize important discriminative features from candling images, rather than merely relying on noise or color distribution.

## 4. DISCUSSION

The classification results obtained indicate that the CNN architecture based on EfficientNetB4 is capable of distinguishing between fertile and infertile eggs with high accuracy, even when using candling images that are often visually difficult to differentiate manually. This success is mainly influenced by the use of transfer learning and fine-tuning strategies, as well as effective preprocessing and data augmentation techniques. The use of a pretrained model from ImageNet offers an initial advantage since the model has already learned to recognize general patterns from various objects, which are then specifically adapted through domain adaptation to egg images [21].

In general, achieving an F1-score of 0.95 places the model in the category of excellent performance for binary classification tasks. The balanced F1-score between fertile and infertile classes indicates that the model is not biased toward either class, which is crucial in the context of poultry production, where misclassification of fertile eggs can result in economic losses. This finding aligns with the importance of classification stability in computer vision systems for agricultural products [2].

First, our results align with previous findings using deep learning for egg fertility detection. For instance, a study employing CNN-based transfer learning (including architectures like VGG16, ResNet50, InceptionNet, and MobileNet) achieved an outstanding accuracy of 98 %, with InceptionNet delivering an accuracy of 0.98, sensitivity of 1.0 for fertile eggs, and specificity of 0.96 for infertile eggs [22]. This performance is comparable to our F1-score of 0.95, reinforcing the effectiveness of transfer learning in egg classification.

Second, alternative approaches using hyperspectral and CCD imaging with advanced signal processing techniques have also demonstrated high accuracy. One method utilizing ROI-based imaging, CLAHE preprocessing, binarization, and BPNN classification on 5-day incubated eggs yielded an accuracy of 98.25 % [23]. Another study integrating hyperspectral imaging, PCA, and K-means classification for embryo detection achieved over 97 % accuracy [24]. These outcomes further support the potential of combining imaging enhancements with machine learning to reach classification levels comparable to our model.

Third, traditional machine learning techniques such as first-order statistical feature extraction combined with SVM have been explored as well. In that approach, features like mean, entropy, variance, skewness, and kurtosis were extracted from 100 egg images, and the SVM classifier achieved a success rate of 84.57 % [25]. Although this is markedly lower than our CNN-based method, it underscores the substantial performance boost offered by modern deep learning architectures.

Fourth, object detection and segmentation models have also been leveraged in the fertility context. A Mask R-CNN–based system used incubator imagery and correctly identified fertile eggs with high precision, leveraging both segmentation and classification in one model [26]. While our focus remains on classification via EfficientNetB4, the integration of segmentation and localization, as demonstrated, represents an interesting direction for enhancing model interpretability and robustness.

Finally, as noted in broader surveys on precision agriculture and hyperspectral applications, ensemble methodologies and transfer learning have repeatedly proven effective. One review reports transfer-learning–based CNN methods achieving up to 99.5 % accuracy on small-scale incubated egg datasets, and multi-feature fusion techniques achieving 98.4 % accuracy [27]. These findings validate

Vol. 6, No. 5, October 2025, Page. 3787-3799
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5234

our proposal to explore ensemble and attention-based models in future work, as mentioned in our discussion.

The effectiveness of image augmentation as part of the training process is evident in this study. Techniques such as rotation, flipping, zooming, and brightness adjustment have proven to increase the diversity of training data, enhance model generalization, and reduce overfitting. This demonstrates that aggressive augmentation can significantly improve CNN model accuracy in domains with limited data availability [16].

Beyond numerical accuracy, model interpretability is also a key focus of this study. Visualization using Grad-CAM shows that the model does not rely on random patterns but focuses on physiologically relevant areas. This strengthens the model's credibility for application in real-world automatic detection systems. Grad-CAM interpretation is also considered essential in enhancing user trust in AI-based systems [19].

However, there are several limitations that should be noted. First, although augmentation was applied, the number of original images remains limited, so there is still a risk of overfitting if the model becomes too complex. Second, the fine-tuning process requires additional training time and involves sensitive hyperparameter selection, such as the number of layers to unfreeze and the learning rate used. Third, model evaluation in this study is limited to images from a single imaging source, so generalization to other imaging devices needs to be validated in future studies.

For future development, it is recommended to use ensemble methods across CNN architectures or apply semi-supervised learning strategies to utilize unlabeled data. Another alternative is to explore Vision Transformer architectures or integrate attention mechanisms to improve spatial interpretation of important features [11], [17].

Considering all these findings, the study demonstrates that the use of EfficientNetB4 supported by preprocessing, augmentation, and visual interpretation is a reliable and scalable approach for other visual classification tasks in the domain of livestock and precision agriculture.

## 5. CONCLUSION

This study successfully demonstrated that the CNN architecture based on EfficientNetB4 is effective in classifying fertile and infertile eggs from candling images. With appropriate transfer learning and fine-tuning approaches, the model achieved high accuracy and F1-score, while producing balanced predictions across both classes.

Preprocessing steps such as resizing to 380×380 pixels, normalization, and histogram stretching were proven to enhance the quality of images learned by the model. Aggressive image augmentation processes, including rotation, flipping, zooming, and brightness adjustment, played a crucial role in enriching the diversity of training data and reducing the risk of overfitting.

The model was trained in two stages: feature extraction and fine-tuning, utilizing initial weights from ImageNet. This training strategy effectively improved performance and accelerated convergence, even with a limited amount of data. The evaluation process showed that the model was not only accurate but also capable of identifying important features through Grad-CAM visualization, which highlighted the model's focus areas during classification.

These findings reinforce the potential of CNN utilization in automated egg classification systems, especially for the poultry industry, which demands fast and accurate detection. Moreover, this approach can be extended to other domains facing similar visual challenges, such as biological object classification or medical imaging.

Although the results are highly satisfactory, there remains room for development, such as expanding the dataset, conducting validation tests across different imaging devices, and exploring alternative architectures. This study opens up great opportunities for developing smarter, more efficient, and easily integrable automatic classification systems in real-world production environments.

## ACKNOWLEDGEMENT

This research was made possible through funding support from the Regular Beginner Lecturer Research Program (PDP) organized by Kemendiktisaintek. We extend our highest appreciation to STMIK IKMI Cirebon for their support in providing facilities and an inspiring academic environment, as well as to the administrative staff and faculty members who have provided the opportunity and venue for the implementation of this research. The results of this study are expected to support SDG No. 2 (Zero Hunger) by increasing the efficiency of poultry production through more accurate prediction of egg quality, and Asta Cita No. 7 (An Economy Based on Advanced Technology and Innovation) by offering a Convolutional Neural Network (CNN)-based solution for optimizing color feature extraction, thereby contributing to the sustainability of the poultry farming sector and supporting food security in Indonesia.

## REFERENCES

[1] F. Nurdiyansyah, S. Fatriana Kadir, I. Akbar, and L. Ursaputra, "Penerapan Convolutional Neural Network Untuk Deteksi Kualitas Telur Ayam Ras Berdasarkan Warna Cangkang," *J. Mnemon.*, vol. 7, no. 1, pp. 40–47, 2024, doi: 10.36040/mnemonic.v7i1.8767.

[2] S. Ghosal and R. Cucchiara, "AI-Based Egg Classification: A Survey of Approaches and Trends," *Comput. Electron. Agric.*, vol. 190, 2021, doi: 10.1016/j.compag.2021.106394.

[3] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-December, pp. 770–778, 2016, doi: 10.1109/CVPR.2016.90.

[4] M. F. A. Pratama, "Klasifikasi Kualitas Telur Ayam Menggunakan CNN," *eProceedings Eng.*, vol. 10, no. 1, 2023.

[5] M. FIRDAUS, KUSRINI, and M. Rudyanto Arief, "Impact of Data Augmentation Techniques on the Implementation of a Combination Model of Convolutional Neural Network (CNN) and Multilayer Perceptron (MLP) for the Detection of Diseases in Rice Plants," *J. Sci. Res. Educ. Technol.*, vol. 2, no. 2, pp. 453–465, 2023, doi: 10.58526/jsret.v2i2.94.

[6] Z. Xu, A. Meng, Z. Shi, W. Yang, Z. Chen, and L. Huang, "Continuous Copy-Paste for Multi-Object Tracking and Segmentation," *ICCV Proc.*, 2021.

[7] K. Baek, D. Bang, and H. Shim, "GridMix: Strong regularization through local context mapping," *Pattern Recognit.*, vol. 109, 2021, doi: 10.1016/j.patcog.2020.107594.

[8] A. Dabouei, S. Soleymani, F. Taherkhani, and N. M. Nasrabadi, "SuperMix: Supervising the mixing data augmentation," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 13789–13798, 2021, doi: 10.1109/CVPR46437.2021.01358.

[9] G. Ghiasi *et al.*, "Simple Copy-Paste is a Strong Data Augmentation Method for Instance Segmentation," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 2917–2927, 2021, doi: 10.1109/CVPR46437.2021.00294.

[10] S. Y. Prasetyo, "Overcoming Overfitting in CNN Models for Potato Disease Classification Using Data Augmentation," *Eng. Math. Comput. Sci. J.*, vol. 6, no. 3, pp. 179–184, 2024, doi: 10.21512/emacsjournal.v6i3.11840.

[11] K. Alomar, H. I. Aysel, and X. Cai, "Data Augmentation in Classification and Segmentation: A Survey and New Strategies," *J. Imaging*, vol. 9, no. 2, 2023, doi: 10.3390/jimaging9020046.

[12] N. Cauli and D. Reforgiato Recupero, "Survey on Videos Data Augmentation for Deep Learning Models," *Futur. Internet*, vol. 14, no. 3, 2022, doi: 10.3390/fi14030093.

[13] Z. Liu *et al.*, "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows," *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 9992–10002, 2021, doi: 10.1109/ICCV48922.2021.00986.

[14] B. Nurhakim, A. Rifai, D. A. Kurnia, D. Sudrajat, and U. Supriatna, "Smart Attendance Tracking System Employing Deep Learning for Face Anti-Spoofing Protection," *JITK (Jurnal Ilmu*

Vol. 6, No. 5, October 2025, Page. 3787-3799
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5234

*Pengetah. dan Teknol. Komputer)*, vol. 10, no. 3, pp. 496–505, 2025, doi: 10.33480/jitk.v10i3.5992.

[15] S. R. Yang, H. C. Yang, F. R. Shen, and J. Zhao, "Image Data Augmentation for Deep Learning: A Survey," *Ruan Jian Xue Bao/Journal Softw.*, vol. 36, no. 3, pp. 1390–1412, 2025, doi: 10.13328/j.cnki.jos.007263.

[16] A. Mumuni, F. Mumuni, and N. K. Gerrar, "A Survey of Synthetic Data Augmentation Methods in Machine Vision," *Mach. Intell. Res.*, vol. 21, no. 5, pp. 831–869, 2024, doi: 10.1007/s11633-022-1411-7.

[17] Z. Liu, H. Mao, C. Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A ConvNet for the 2020s," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2022-June, pp. 11966–11976, 2022, doi: 10.1109/CVPR52688.2022.01167.

[18] M. Xu, S. Yoon, A. Fuentes, and D. S. Park, "A Comprehensive Survey of Image Augmentation Techniques for Deep Learning," *Pattern Recognit.*, vol. 137, 2023, doi: 10.1016/j.patcog.2023.109347.

[19] D. Lewy and J. Mańdziuk, "An overview of mixing augmentation methods and augmentation strategies," *Artif. Intell. Rev.*, vol. 56, no. 3, pp. 2111–2169, 2023, doi: 10.1007/s10462-022-10227-z.

[20] Rachel Kim and Emily White, "Convolutional neural network for data augmentation," *World J. Adv. Eng. Technol. Sci.*, vol. 13, no. 2, pp. 870–886, 2024, doi: 10.30574/wjaets.2024.13.2.0528.

[21] M. Tan and Q. V. Le, "EfficientNetV2: Smaller Models and Faster Training," in *Proceedings of Machine Learning Research*, 2021, vol. 139, pp. 10096–10106, [Online]. Available: https://proceedings.mlr.press/v139/tan21a.html.

[22] S. Song, Q. Li, and J. Xu, "Deep Learning-Based Classification of Broiler Chicken Egg Fertility Using Convolutional Neural Networks," *Appl. Sci.*, vol. 11, no. 14, p. 6403, 2021, doi: 10.3390/app11146403.

[23] L. Huang, A. He, M. Zhai, Y. Wang, R. Bai, and X. Nie, "A multi-feature fusion based on transfer learning for chicken embryo eggs classification," *Symmetry (Basel).*, vol. 11, no. 5, p, 2019, doi: 10.3390/sym11050606.

[24] "A Multi-Feature Fusion Based on Transfer Learning with Hyperspectral Imaging and PCA," *Symmetry (Basel).*, 2021, doi: [DOI].

[25] S. Saifullah and A. P. Suryotomo, "Identification of chicken egg fertility using SVM classifier based on first-order statistical feature extraction," *Ilk. J. Ilm.*, vol. 13, no. 3, pp. 285–293, 2021, doi: 10.33096/ilkom.v13i3.937.285-293.

[26] V. Gupta, A. Mandloi, S. Pawar, T. V. Aravinda, and K. R. Krishnareddy, "Deep learning based object detection using mask R-CNN," *AI IoT-based Intell. Heal. Care Sanit.*, pp. 207–221, 2023, doi: 10.2174/9789815136531123010016.

[27] A. O. Adegbenjo, L. Liu, and M. O. Ngadi, "Non-destructive assessment of chicken egg fertility," *Sensors (Switzerland)*, vol. 20, no. 19, pp. 1–23, 2020, doi: 10.3390/s20195546.