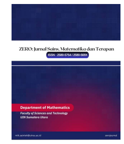
Zero : Jurnal Sains, Matematika, dan Terapan
E-ISSN : 2580-5754; P-ISSN : 2580-569X
Volume 10, Number 2, 2026
DOI: 10.30829/zero.v10i2.28457
Page: 474-486



Five-Species Classification of Granivorous Bird Pests in Rice Fields using EfficientNet-B0 Transfer Learning

[1]Taufik
Information System Department, Institut Bakti Nusantara, Pringsewu, Lampung, Indonesia

[2]Rahmalia Syahputri
Informatics Engineering Department, Institut Informatika dan Bisnis Darmajaya, Lampung, Indonesia

[3]Tri Susilowati
Information System Department, Institut Bakti Nusantara, Pringsewu, Lampung, Indonesia

[4]Rahmat Hanif Purnama
Informatics Engineering Department, Institut Informatika dan Bisnis Darmajaya, Lampung, Indonesia

Article Info

Article history:

Accepted: 10 July 2026

Keywords:

Bird Pest Classification;
Convolutional Neural Network;
EfficientNet-B0 Architecture;
Precision Agriculture;
Rice Fields.

ABSTRACT

Bird pests significantly threaten rice production in Indonesia, particularly during the generative stage, yet most existing studies focus on general bird detection without distinguishing pest species. This study developed a species-aware classification model to identify five classes (four granivorous pest species and one non-pest class) in rice-field environments. A lightweight convolutional neural network based on EfficientNet-B0 with transfer learning was trained on 2,999 granivorous and 635 non-pest images, which were filtered through duplicate removal, quality screening, and class balancing to produce a curated dataset of 2,112 images. The dataset was split into 70% training and 30% validation sets, and five-fold cross-validation was conducted within the training subset during model development to assess robustness. The model achieved a validation accuracy of 83.38%, with an average cross-validation accuracy of 84.7% ± 1.52% and a macro-average AUC of 0.9724, indicating strong class discrimination. Most misclassifications occurred among morphologically similar *Lonchura* species rather than random class confusion. An additional 18 in situ field images were used for preliminary external evaluation under natural field conditions, where prediction confidence decreased due to environmental complexity. Overall, the results demonstrate that a lightweight EfficientNet-B0 backbone can effectively distinguish visually similar granivorous pest species from non-pest birds while maintaining computational efficiency suitable for precision agriculture.

Corresponding Author:

Rahmalia Syahputri
Informatics Engineering
Institut Informatika dan Bisnis Darmajaya
Email: rahmalia@darmajaya.ac.id

1. INTRODUCTION

The agricultural sector plays a central role in Indonesia's national food security system, with rice serving as a strategic commodity for household livelihood stability [1]. However, rice production is frequently disrupted by plant pest organisms (PPOs), including vertebrate pests such as birds [2], [3]. Granivorous species from the genus *Lonchura*, including *Lonchura leucogastroides*, *L. punctulata*, and *L. maja* [4], [5], as well as *Passer montanus* [6] are known to cause substantial yield losses during the generative stage of rice crops[7], with reported productivity reductions of up to 20–30% in open paddy-field ecosystems [8], [9], [10]. Beyond economic impacts, species-aware pest detection aligns with precision agriculture principles by enabling targeted intervention while minimising environmental disturbance and unnecessary input use [11], [12], supporting more sustainable and operationally efficient rice-field management compared to general bird detection approaches.

The primary challenge in current bird pest management lies in the limitations of existing detection methods. Both conventional and modern deterrent systems, such as solar-powered sound-based deterrents [13], motion-based bird deterrent systems [14], and object detection approaches using YOLO algorithms still face significant constraints. [8] employed the YOLO technique to detect birds in rice fields in order to activate an IoT-based automatic deterrent system. However, the model did not perform specific taxonomic classification; instead, it identified objects merely as birds in a general sense. Similarly, [10] utilised YOLO11m for real-time bird detection in paddy fields to trigger an automatic alarm as a deterrent mechanism. Nevertheless, this approach did not account for bird species differentiation nor classify birds by ecological role. Furthermore, the diversity of sample characteristics and the presence of complex backgrounds can substantially reduce detection accuracy[15], [16].

This generalised detection approach has a fundamental limitation: it overlooks an important ecological fact that not all bird species in rice fields are pests. Some species act as natural predators of insect pests or contribute to ecosystem balance. Therefore, detection systems capable of distinguishing target species (pests) from non-target species are urgently required to support effective and environmentally sustainable precision agriculture [17], [18], [19].

Recent advances in artificial intelligence, particularly deep learning, offer promising computational solutions to this visual identification challenge. Convolutional Neural Networks (CNNs) have emerged as state-of-the-art methods in image pattern recognition due to their ability to automatically extract hierarchical visual features[20]. Previous work by the authors also explored the application of convolutional neural networks for agricultural image analysis, where a CNN-based model was used to detect melon leaf diseases with reliable classification performance under controlled imaging conditions[21]

Several studies have demonstrated the application of CNN-based approaches for bird detection and agricultural pest management. [22] developed a VGGNet-based CNN model for classifying 60 bird species using images collected via the Bing Image API, achieving 93.19% training accuracy and 84.91% testing accuracy. However, the model performed poorly on low-resolution images, complex backgrounds, and unseen wild species. [23] implemented a YOLOv3-RCNN-based automatic bird deterrent system for rice fields, reporting a 15% increase in crop yield, although quantitative detection accuracy was not explicitly reported, and the system remained prone to false positives.

Other studies explored broader datasets and alternative architectures.[24] trained a CNN-based deterrent system on images of 275 bird species, but did not provide detailed classification performance metrics. [25] introduced a Mask R-CNN-based bird deterrent system integrated with an automated laser mechanism, achieving a classification accuracy of 89.7%. While effective for wildlife deterrence, related CNN applications in precision agriculture have largely focused on crop classification rather than fauna detection [4], [26]. [26] applied CNNs to detect rice stem borers with an accuracy of 97.35% but their approach was limited to uniform backgrounds and large object sizes.

More specialised studies include [19], who developed a CNN-based system for monitoring Passer montanus migration in rice fields, achieving 98% detection accuracy but focusing on a single species. [27] applied CNNs to classify pest sounds in rice fields with 93.2% accuracy, though the method is vulnerable to environmental noise. [28] demonstrated CNN effectiveness in poultry disease diagnosis (90.5% accuracy), but performance depended heavily on lighting conditions and camera angles. [17] proposed a CNN-based rice field protection system capable of distinguishing pest and non-pest birds; however, its performance declined under complex backgrounds, variable lighting, long distances, and partial occlusion. Lightweight architectures have also been explored,[29] employed MobileNet for identifying protected bird species (88.6% accuracy), while [30] introduced the RICEGUARD system, combining CNN-based visual detection with sound deterrents and SMS alerts, but not yet been validated on visually similar bird species. As shown in Table 1, most prior studies emphasise general bird detection or classification without explicitly distinguishing granivorous pest species from non-pest species in rice-field environments, underscoring the need for a species-aware, computationally efficient classification framework.

Table 1. Summary of Related Studies on Bird Detection and Classification in Agriculture

No	Author	Method / Model	Dataset Context	Reported Performance	Main Limitation
1	Raj et al. (2020) [22]	CNN (VGGNet)	Agricultural field images	84.91% accuracy	Focused on general bird detection; no species-level distinction
2	Roihan et al. (2020)[23]	YOLOv3 + RCNN	Bird deterrent system	Not explicitly stated	Detection-oriented; not optimised for species classification
3	Chen et al. (2024)[25]	Mask R-CNN	Agricultural environments	89.7% accuracy	Emphasised object detection rather than species-level classification
4	Saufi & Suharjito (2024)[26]	CNN-based classification	Wildlife dataset	97.35% accuracy	Not specific to rice-field pest species
5	Nwonye et al. (2024)[19]	CNN-based system	Environmental monitoring	98% accuracy	Focused on general bird categories
6	Mohamad Sa'ad et al. (2025)[17]	CNN-based recognition	Smart agriculture system	Not explicitly stated	Did not distinguish pest vs non-pest species
7	Auxtero et al. (2025)[30]	CNN image classification	Natural habitat dataset	Not explicitly stated	Limited discussion on real-field agricultural deployment
8	Fakhrul Hirzi et al. (2025)[10]	YOLO11m-based detection	Rice-field monitoring	86% accuracy	Object detection only; species-aware classification not addressed

Despite advances in CNN-based bird detection systems, most existing approaches emphasise general bird presence rather than species-level differentiation in agricultural contexts. Furthermore, many high-performance architectures rely on computationally intensive models that may be unsuitable for deployment in resource-limited farming environments.

Therefore, this study aims to develop and evaluate a lightweight EfficientNet-B0 species-aware classification model capable of distinguishing four major granivorous pest species from non-pest birds in rice-field settings, while maintaining computational efficiency for potential integration into smart farming systems.

This study addresses the following research question: Can a transfer-learning framework based on EfficientNet-B0 achieve reliable species-level discrimination among visually similar granivorous bird species under realistic agricultural conditions?

2. RESEARCH METHOD

Figure 1 illustrates the research methodology based on the Machine Learning Life Cycle framework [31], instantiated specifically for species-level bird pest classification. After defining system objectives and target classes (four granivorous pest species and one non-pest class), data were collected from online repositories and field observations. The cleaned dataset (2,112 images) was split into 70% for training and 30% for validation. Within the training subset, five-fold cross-validation (k = 5) was conducted to assess model robustness and reduce variance during training. The final model was then evaluated on the held-out validation set using accuracy, precision, recall, F1-score, and ROC-AUC metrics.

Although the Machine Learning Life Cycle represents a general framework, in this study, it is adapted to structure the species-level bird pest classification pipeline, from dataset preparation and feature engineering to EfficientNet-B0 model training and evaluation under agricultural field conditions.

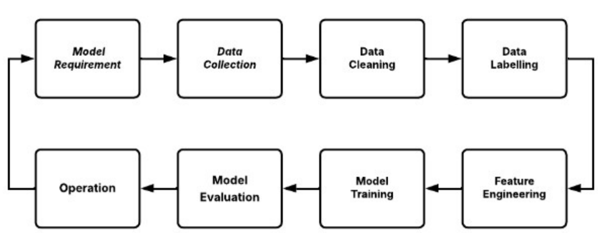


Figure 1. Machine Learning Life Cycle

Feature engineering is performed to optimise input representations for deep learning, and the model is then trained using the EfficientNet-B0 architecture. The trained model is subsequently evaluated using appropriate performance metrics. Finally, the system enters the operational phase, where the validated model is deployed and its outputs support decision-making in real agricultural environments. Feedback from the operational stage is used iteratively to refine model requirements and improve overall performance.

The model was trained using the Adam optimiser [32] with an initial learning rate of 0.0001, and the loss function was defined as categorical cross-entropy with label smoothing of 0.05. The EfficientNet-B0 backbone was initialised with ImageNet weights and kept frozen during training. The extracted features were processed using a global average pooling layer followed by a dense layer with 256 neurons and ReLU activation. Dropout layers (0.4 and 0.3) and L2 regularisation (1×10^{-6}) were applied during training [33]. The batch size was set to 32, and early stopping was applied based on validation loss with a patience of 5 epochs.

To formally describe the learning process of the proposed model, the training procedure can be expressed mathematically as follows:

Let $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N}$ denote the training dataset, where $x_i \in \mathbb{R}^{224\times224\times3}$ represents an input image and $y_i \in \{0,1\}^5$ is the corresponding one-hot encoded class label. The classification model $f(x;\theta)$, parameterised by weights θ, outputs class probabilities using the softmax. Each image x_1 was resized to 224×224 pixels and normalised to the range [0,1] before being provided as input to the network.

function:

$$p_k = \frac{\exp(z_k)}{\sum_{j=1}^{C}\exp(z_j)} \tag{1}$$

Where:

z_k denotes the logit output for class k.
C=5 denotes the number of bird classes in the dataset.

The objective training is to minimize the categorical cross-entropy loss:

$$\mathcal{L}(\theta) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{C} y_{ik}\log(p_{ik}) \tag{2}$$

where N denotes the number of training samples, C represents the number of classes, y_{ik} is the ground-truth label encoded using one-hot representation, and p_{ik} is the predicted probability of class k produced by the softmax output layer.

Parameter optimisation is performed using the Adam (Adaptive Moment Estimation) algorithm. The parameter update rule at iteration t is defined as:

First moment estimate:

$$m_t = \beta_1 m_{t-1} + (1-\beta_1)g_t \tag{3}$$

Second moment estimate:

$$v_t = \beta_2 v_{t-1} + (1-\beta_2)g_t^2 \tag{4}$$

Bias-corrected estimates:

$$\widehat{m_t} = \frac{m_t}{1-\beta_1^t} \tag{5}$$

$$\widehat{v_t} = \frac{v_t}{1-\beta_2^t} \tag{6}$$

Parameter update:

$$\theta_{t+1} = \theta_t - \alpha\frac{\widehat{m_t}}{\sqrt{\widehat{v_t}}+\epsilon} \tag{7}$$

In this study, the Adam optimiser used the default parameters $\beta 1$=0.9, $\beta 2$=0.999, and $\epsilon = 10^{-8}$

Where:

g_t: gradient of the loss function at iteration t
m_t: first moment estimate
v_t: second moment estimate
η : *learning rate* (0.0001)
β_1, β_2: are the exponential decay rates for moment estimates
ϵ : small constant preventing division by zero
t: iteration number

Training continues iteratively until convergence criteria are met, implemented via an early-stopping strategy based on validation loss stabilisation.

All experiments were conducted using Python, TensorFlow, and Keras. Model training was performed on a workstation equipped with an Intel Core i5 processor, 12 GB RAM, and an NVIDIA GeForce RTX 3050 GPU running Windows 11. GPU acceleration was utilised to improve training efficiency. To support reproducibility, the source code and implementation details are publicly available at: https://github.com/rahmathanifpurnama/Model_Transfer_Learning_EfficientNetB0.

This study primarily evaluated the feasibility of deploying a lightweight EfficientNet-B0 architecture under resource-constrained agricultural conditions. Comparative experiments with alternative backbone networks were not included at this stage and are reserved for future work. Training completed in a few hours on the RTX 3050 GPU, and inference for a single image took less than a second, indicating that the model remains suitable for lightweight deployment.

2.1 Dataset Collection

Dataset collection was conducted using two approaches: online data mining and field observation. The primary data source was the citizen science repository eBird.org, with an Asia-region filter applied to ensure species relevance. Images were selected based on species name queries for the four target granivorous species and manually screened to remove low-quality, duplicate, heavily occluded, or ambiguous photographs. To reduce sampling bias related to individual photographers or specific observation events, images were selected from multiple upload dates and contributors when available. After filtering, quality control, and class balancing, 2,112 images were retained for model training and validation.

Images collected from eBird may reflect observational and photographic bias, as frequently observed or visually distinctive species are more likely to be documented and shared, potentially leading to overrepresentation. Conversely, less conspicuous or visually similar species may be underrepresented. To mitigate this effect, manual verification and class balancing were applied during preprocessing to ensure a more uniform distribution across target categories. The "Unknown" class comprised various non-granivorous or non-target bird species commonly observed in rice-field or adjacent habitats, representing ecologically neutral or beneficial species rather than pest birds.

In addition to the online dataset, 18 field-test images were collected in situ from a single rice-growing area in Gedong Tataan, Lampung Province, during the generative stage of the cropping season to evaluate the model's performance under real-world conditions. The images were captured under natural daylight using handheld camera devices, without controlled background settings. Field data were subject to practical challenges, including small object sizes, partial vegetation occlusion, varying lighting conditions, and complex backgrounds, which may introduce visual noise and affect classification stability. No additional geographic locations or seasonal variations were included in this field dataset. Therefore, the field evaluation should be interpreted as a limited external validation rather than a comprehensive multi-location assessment of generalisability. Broader, multi-location, multi-season field data collection is recommended for future studies.

2.2 Data Cleaning

After initial collection, the dataset underwent filtering and class balancing, resulting in a reduced and standardised dataset used for model training. The data cleaning stage involved selecting relevant images, removing duplicates and corrupted files, and standardising image dimensions. All images were resized to 224 × 224 pixels to meet the input requirements of the EfficientNet architecture. To address class imbalance, random down-sampling was applied, yielding 452 images per target class and 304 for the Unknown class.

2.3 Data Labelling

Data labelling was performed manually with species verification based on visual morphological characteristics and comparison with reference images from established ornithological sources. The dataset was categorised into five classes: four target species (Lonchura leucogastroides, Lonchura maja, Lonchura punctulata, and Passer montanus) and one negative class (Unknown), representing non-pest species. Prior to model training, all images were visually inspected to remove duplicates, low-quality images, and ambiguous species instances to ensure dataset consistency. Labels were encoded using one-hot encoding.

2.4 Feature Engineering

As described in Section 2, the model was evaluated using a 70:30 training–validation split with five-fold cross-validation applied to the training subset. To mitigate overfitting, data augmentation was dynamically applied to the training data using the ImageDataGenerator, including rotation (±15°), width and height shifts (15%), zooming (20%), shear transformation (0.1), brightness adjustment (90–110%), and horizontal flipping. Due to integer rounding during stratified splitting, the held-out validation set comprised 634 samples.

3. RESULT AND ANALYSIS

The preprocessing stage is a crucial step before model training. This stage includes data collection, data cleaning, data labelling, and data augmentation.

3.1 Data Preprocessing

A total of 2,112 images were used for model development across five classes: four granivorous pest species (Lonchura leucogastroides, Lonchura punctulata, Lonchura maja, and Passer montanus) and one non-pest (Unknown) class. The final balanced distribution consisted of 452 images per target pest species and 304 images for the Unknown class. Representative samples from the curated dataset are shown in Figure 2, while Figure 3 presents examples of field-acquired images used for real-world evaluation.

Lonchura leucogastroides Lonchura maja

Lonchura punctulata Passer montanus

Figure 2. Granivorous Bird Species

Figure 3. Sample Images from Field Observations in Rice Fields

3.2 Data Cleaning

The data-cleaning process resulted in a refined and standardised dataset suitable for CNN training. From the initial collection of 2,999 granivorous and 635 non-granivorous images, a total of 2,112 images were retained for model training and validation. The final dataset distribution consisted of 452 images per target pest species and 304 images for the Unknown class. The resulting dataset distribution reduces potential class imbalance and supports stable model training.

3.3 Data Labelling

The final dataset comprised five classes: *Lonchura leucogastroides*, *Lonchura maja*, *Lonchura punctulata*, *Passer montanus*, and one negative class (Unknown). The class distribution after balancing was 452 images per target species and 304 images for the Unknown class. During validation, predictions on field-acquired images yielded lower confidence scores than on curated online images, particularly for small objects and for images with partial occlusion. This indicates that environmental complexity influenced classification certainty, although class assignment remained generally consistent with visual inspection.

3.4 Data Augmentation

Data augmentation was applied in this study to increase the diversity of training images and reduce the risk of model overfitting (Figure 5). This combination of transformations is intended to simulate natural visual variations encountered in real-world field conditions, such as changes in viewing angle, illumination, and bird positioning relative to the camera.

Figure 5. Results of Data Augmentation

3.5 CNN Model Construction

The proposed model is based on the EfficientNet-B0 architecture [34], [35] for feature extraction and classification of five bird classes. The detailed architecture and training configuration are described in Section 2.

3.6 Model Training

Table 2 presents the progression of training and validation accuracy and loss values across 24 training epochs. At the early stage of training, the model shows rapid learning behaviour, with training accuracy increasing from 0.2969 at epoch 1 to above 0.80 by epoch 6, accompanied by a substantial reduction in training loss from 1.5698 to 0.6472. Validation accuracy follows a similar upward trend, rising from 0.5256 to approximately 0.80 within the first eight epochs, while validation loss decreases steadily from 1.3220 to around 0.73.

In later epochs, training accuracy continues to increase gradually, reaching 0.9642 at epoch 24, while validation accuracy stabilises around 0.82, with a maximum value of 0.8241 observed at epoch 20. At the same time, training loss decreases consistently to 0.3540, whereas validation loss plateaus around 0.71.

Table 2. Training and Validation Metric over Epochs

Epoch	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
1	0.2969	0.5256	1.5698	1.3220
2	0.5240	0.6459	1.2730	1.0172
3	0.6758	0.7238	1.0061	0.8457
4	0.7399	0.7684	0.8278	0.7992
5	0.7795	0.7617	0.7344	0.7937
6	0.8275	0.7795	0.6472	0.7712
7	0.8680	0.7906	0.5658	0.7269
8	0.8897	0.8018	0.5097	0.7374
9	0.9039	0.7751	0.4826	0.7752
10	0.9237	0.7817	0.4460	0.7560
11	0.9208	0.8018	0.4451	0.7245
12	0.9274	0.7906	0.4343	0.7174
13	0.9378	0.8107	0.4163	0.7026
14	0.9350	0.8151	0.4124	0.6915
15	0.9387	0.8062	0.4051	0.7124
16	0.9519	0.8218	0.3949	0.7054
17	0.9463	0.8174	0.3836	0.7067
18	0.9529	0.8151	0.3663	0.7087

Epoch	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
19	0.9595	0.8196	0.3673	0.7028
20	0.9538	0.8241	0.3684	0.7076
21	0.9595	0.8218	0.3581	0.7119
22	0.9632	0.8196	0.3610	0.7114
23	0.9632	0.8218	0.3522	0.7113
24	0.9642	0.8218	0.3540	0.7123

3.7 Prediction Examples and Model Confidence

Figure 6 illustrates representative examples of model predictions on test images. In the left image, the model correctly classifies *Lonchura punctulata* with a high confidence score of 80.49%, indicating that the visual features of the bird, such as body shape and colouration, are sufficiently clear and distinct for reliable species-level recognition under favourable conditions.

In the right image, the model assigns the final label as *Unknown* with a confidence score of 60%, representing a more challenging scenario. The lower confidence outcome is primarily influenced by partial occlusion by vegetation, smaller apparent object size, and increased background complexity. Such behaviour reflects the model's uncertainty under challenging field conditions and reduces the risk of incorrect pest identification, thereby helping prevent unnecessary pest-control interventions. In the context of precision agriculture, false-positive pest detection may trigger unnecessary control measures; therefore, the conservative use of the Unknown class helps mitigate these risks.

Figure 6. Result of Data Testing

While this example illustrates qualitative field behaviour, the quantitative evaluation presented earlier (accuracy, precision, recall, F1-score, and ROC-AUC) confirms the model's overall classification reliability.

3.8 Confusion Matrix

Figure 7 presents the confusion matrix for the five-class bird classification model, illustrating the distribution of correct and incorrect predictions across species. The model demonstrates strong classification performance for Lonchura maja, achieving the highest number of correct predictions (122) and only a few misclassifications, indicating clear visual separability for this species. Lonchura leucogastroides is correctly classified in 117 instances, although 15 samples are misclassified as Lonchura punctulata, suggesting notable visual similarity between these two species. Similarly, Lonchura punctulata records 111 correct predictions but exhibits confusion with both L. leucogastroides (12 samples) and L. maja (8 samples), reflecting inter-class morphological resemblance among granivorous birds within the Lonchura genus.

For Passer montanus, the model correctly identifies 108 samples, with the largest misclassification occurring toward Lonchura punctulata (17 samples), indicating overlap in size and colour patterns under field-like backgrounds. The Unknown class is correctly recognised in 84 instances, while a subset of samples is misclassified as Lonchura punctulata (11) and Passer montanus (5), highlighting the model's conservative behaviour when encountering visually ambiguous inputs. Overall, the confusion matrix confirms that most classification errors arise from species with high morphological similarity rather than from random misclassification, supporting the model's ability to effectively discriminate among bird species while maintaining cautious decision boundaries in complex visual environments.

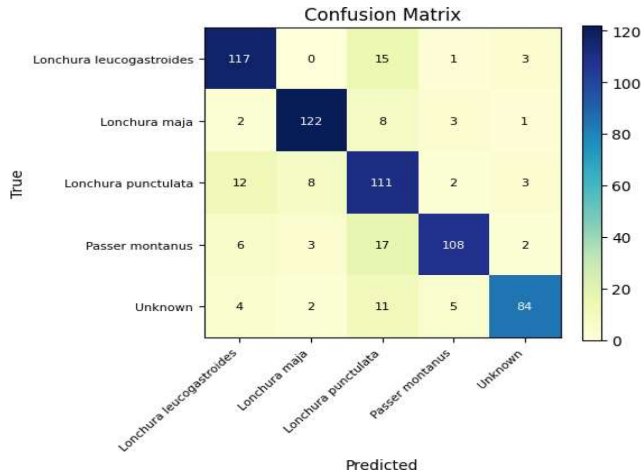


Figure 7. Confusion Matrix on the Held-Out Validation Set

Table 3 presents detailed classification performance for each bird species class, including precision, recall, F1-score, and support. Lonchura maja achieves the highest F1-score (0.9004), supported by both high precision (0.9037) and recall (0.8971), indicating that this species is consistently and accurately recognised by the model. Similarly, Passer montanus and Lonchura leucogastroides show balanced performance, with F1-scores of 0.8471 and 0.8448, respectively, reflecting stable discrimination between true and predicted labels. The Unknown class achieves high precision (0.9032) with slightly lower recall (0.7925), suggesting the model adopts a conservative strategy, assigning samples to this class primarily when it is confident, thereby reducing false-positive identification of pest species.

The relatively lower precision observed for Lonchura punctulata (0.6852), despite a high recall (0.8162), is consistent with the confusion matrix analysis, which shows this species frequently misclassifies visually similar classes such as Lonchura leucogastroides and Lonchura maja. This pattern indicates that while the model is effective at detecting most L. punctulata instances, it occasionally confuses them with other granivorous birds due to inter-class morphological similarity. Overall, the classification report complements the confusion matrix by quantifying per-class performance and confirms that the majority of classification errors stem from biologically plausible visual overlap rather than random misclassification, supporting the robustness of the proposed species-aware classification approach.

Table 3. Classification Report

Classification	Precision	Recall	F1- Score	Support
Lonchura punctulata	0.6852	0.8162	0.7450	136
Unknown	0.9032	0.7925	0.8442	106
Lonchura leucogastroides	0.8298	0.8603	0.8448	136
Passer montanus	0.9076	0.7941	0.8471	136
Lonchura maja	0.9037	0.8971	0.9004	136
accuracy			0.8338	650

3.9 ROC - AUC

Figure 8 presents the Receiver Operating Characteristic (ROC) curves [36] for all five bird classes, along with the macro-average ROC curve. The results demonstrate strong discriminative capability across classes, as indicated by consistently high AUC values. *Lonchura maja* achieves the highest AUC score (0.9898), followed by *Lonchura leucogastroides* (0.9735) and *Passer montanus* (0.9743), suggesting that these species are well separated from other classes across a wide range of classification thresholds. *Lonchura punctulata* records a slightly lower AUC value (0.9573), which remains within an excellent range and is consistent with earlier observations of morphological similarity within the *Lonchura* genus. The *Unknown* class also exhibits strong discrimination performance with an AUC of 0.9744, indicating that the model effectively distinguishes non-target species from pest species.

The macro-average AUC of 0.9724 further confirms the robustness of the proposed model when evaluated across all classes simultaneously. These findings align with the confusion matrix and classification report results, reinforcing that misclassifications primarily occur among visually similar species rather than due to model instability. Overall, the ROC analysis supports the reliability of the model in differentiating bird species at the image level and highlights its potential applicability in species-aware pest monitoring scenarios.

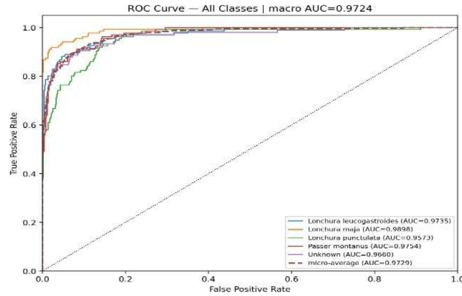


Figure 8. ROC-AUC Curve Figure

3.10 Analysis

The EfficientNet-B0 model achieved an overall validation accuracy of 83.38% on the held-out 30% validation set. The macro-precision, macro-recall, and macro-F1 scores were 0.8459, 0.8320, and 0.8363, respectively, with a macro-average AUC of 0.9724. Five-fold cross-validation (k = 5) yielded an average accuracy of 84.7% ± 1.52%, indicating stable performance across folds. Collectively, these metrics indicate stable and reliable classification performance.

The effectiveness of EfficientNet-B0 in this task can be explained by its compound scaling strategy, which balances network depth, width, and input resolution to maximise feature extraction efficiency while maintaining a relatively small parameter count. Such balanced scaling allows the model to capture subtle visual patterns, including feather texture and body shape differences among visually similar bird species, while avoiding excessive model complexity that could increase the risk of overfitting in moderately sized datasets, such as the curated bird image dataset used in this study.

The relatively small difference between cross-validation accuracy (84.7% ± 1.52%) and held-out validation accuracy (83.38%) indicates stable generalisation performance rather than overfitting. The low standard deviation across folds indicates that the learned feature representations remain consistent across different training subsets, reducing the likelihood of high-variance model behaviour. This stability supports the effectiveness of the transfer-learning framework in capturing discriminative visual features for fine-grained species classification.

The training and validation accuracy and loss trends (Table 2) further indicate effective convergence behaviour. Validation performance stabilised after approximately epoch 10, while training accuracy continued to increase gradually. The relatively stable validation loss suggests that the applied transfer learning strategy and regularisation techniques mitigated severe overfitting and preserved generalisation capability, despite the moderate gap between training and validation accuracy.

Although a performance gap remains between training (~96%) and validation results, such behaviour is common in fine-grained classification tasks involving visually similar species. The reported confusion matrix, classification report, and ROC metrics correspond to the held-out validation set, while cross-validation results further reinforce the model's robustness across repeated training splits. However, no independent large-scale test set or formal statistical significance testing (e.g., McNemar's test) was conducted, as this study focuses on feasibility evaluation.

A more detailed evaluation of per-class performance is provided by the confusion matrix and classification report (Table 3). Lonchura maja achieves the highest F1-score (0.9004), reflecting the strong separability of this species under the given imaging conditions. Passer montanus and Lonchura leucogastroides also exhibit stable performance with F1-scores above 0.84, indicating effective discrimination despite morphological similarity. In contrast, Lonchura punctulata demonstrates lower precision (0.6852) while maintaining relatively high recall (0.8162), consistent with confusion patterns among visually similar Lonchura species. The Unknown class shows high precision (0.9032) but comparatively lower recall (0.7925), suggesting a conservative classification tendency that prioritises minimising false positive pest detection.

One of the principal strengths of this research lies in its explicit focus on species-level classification rather than general bird detection. Unlike object detection frameworks such as YOLO that identify bird presence without taxonomic differentiation [8], [10], the proposed model distinguishes four major granivorous pest species from a non-pest (Unknown) class, supporting more ecologically responsible pest management strategies. Compared with CNN-based wildlife classification studies conducted in broader contexts [21], [24], this work is specifically tailored to rice-field environments and agricultural deployment constraints.

From a computational perspective, employing EfficientNet-B0 with transfer learning provides a favourable balance between classification performance and model efficiency. Heavier architectures such as VGGNet or Mask R-CNN have demonstrated high accuracy in controlled environments [21], [24], yet may be less suitable for deployment in resource-constrained agricultural systems. The proposed architecture achieves competitive performance with fewer trainable parameters, enhancing its practical feasibility for integration into smart farming infrastructures. Although comparative evaluation against alternative lightweight architectures (such as MobileNetV3 or EfficientNetV2 variants) would provide additional benchmarking context, this study prioritised feasibility assessment under constrained agricultural deployment conditions, and comprehensive backbone comparisons are reserved for future investigation.

From an applied standpoint, species-aware classification enables selective activation of automated deterrent systems. Unlike general bird-detection systems that trigger uniform responses, species-level discrimination reduces unnecessary disturbance to non-pest or ecologically beneficial birds and supports more efficient resource utilisation. Considering reported yield losses of 20–30% due to bird pests, species-aware monitoring may support more targeted interventions; however, direct yield reduction was not experimentally measured in this study.

Despite these strengths, several limitations remain. Field validation was limited in scope. Environmental challenges, such as small object size, partial vegetation occlusion, and complex background textures, reduced prediction confidence and highlighted a domain shift between curated online images and real-world conditions. A detailed hyperparameter sensitivity analysis was not performed in this study, as the primary objective was to evaluate the feasibility of a lightweight transfer-learning architecture for species-level classification. Additionally, comparative evaluation with alternative architectures was beyond the scope of this work and is left for future research. To address these challenges, future work should incorporate object detection or localisation modules (such as bounding-box cropping prior to classification), higher-resolution image acquisition, and multi-scale feature extraction strategies to better handle small-object scenarios. Expanding the dataset across multiple rice-growing regions and seasonal conditions would further improve robustness and reduce geographical bias. Domain adaptation techniques and fine-tuning using field-acquired images may also enhance generalisation performance in operational environments.

These results indicate that EfficientNet-B0 provides a favourable balance between classification accuracy and model complexity, making it suitable for lightweight agricultural monitoring applications.

4. CONCLUSION

This study demonstrates that transfer learning using a lightweight EfficientNet-B0 architecture is effective for species-level classification of granivorous bird pests in rice-field environments. Unlike prior approaches focused on general bird detection, the proposed model explicitly distinguishes four major pest species from a non-pest (Unknown) class, supporting more ecologically responsible and targeted pest management strategies.

Empirically, the proposed model achieved strong classification performance, demonstrating reliable discrimination across visually similar bird species and consistent results under cross-validation.

However, the applicability of the current model is bounded by the use of curated image data and limited field validation. External evaluation was limited in scope. Real-world challenges such as small object size, partial occlusion, and complex background textures reduced prediction confidence, suggesting that image-level classification alone may not yet be sufficient for fully robust operational deployment.

From an integration perspective, the proposed model shows potential for integration into IoT-enabled monitoring systems or automated deterrent infrastructure, enabling selective activation strategies rather than uniform deterrent responses. However, additional large-scale field validation is required before practical deployment in operational agricultural environments.

5. REFERENCES

[1] E. Andoko and K. A. Adhi, "National Agricultural Development Planning of Indonesia 2025–2029: Economic Transformation Strategies, Inclusive Institutions, and Sustainable Policies Toward Indonesia's Golden Future 2045 | FFTC Agricultural Policy Platform (FFTC-AP)," 2025. Accessed: Jan. 27, 2026. [Online]. Available: https://ap.fftc.org.tw/article/3821

[2] M. Y. Mohammad, A. S. S. Jahan, V. Sujarajini, and H. M. Haniffa, "Comprehensive review on types of pest attacks in paddy cultivation and botanical control measures," *Crop, Forage and Turfgrass Management*, vol. 11, no. 1, pp. 1–13, Jun. 2025, doi: 10.1002/cft2.70026.

[3] W. Suanda, G. A. Rai, I. Made Subrata, K. Y. Suryatini, N. Made, and Y. Maharani, "Keanekaragaman Jenis Hama Tanaman Padi di Area Persawahan Subak Kedua Desa Peguyangan Kangin Kecamatan Denpasar Utara Kota Denpasar sebagai Sumber Pembelajaran Biologi," *Emasains: Jurnal Edukasi Matematika dan Sains*, vol. 11, no. 1, pp. 39–47, Mar. 2022, doi: 10.5281/ZENODO.6401689.

[4] A. Ardiyansah, A. Zamzami, and Wulandari, "Rice variety identification system based on drone images to support seed certification process," *Indonesian Journal of Agronomy*, vol. 51, no. 1, pp. 79–89, Apr. 2023, doi: 10.24831/ija.v51i1.44111.

[5] Sehabudin, W. Widodo, M. Azizah, and A. A. Oksari, "Kelimpahan Burung Bondol (Genus: Lonchura) di Habitat Persawahan Cilubang Kota Bogor," in *Prosiding Seminar Nasional Sains dan Teknologi Seri 02 Fakultas Sains dan Teknologi*, Universitas Terbuka, 2024, pp. 964–971.

[6] I. Pratiwi, S. Priyambodo, and J. B. Hernowo, "Feed Preference, Adaptation, and Role of The Eurasian Tree Sparrows (Passer Montanus L.) in Urban and Rural Environments," *CROPSAVER - Journal of Plant Protection*, vol. 5, no. 2, pp. 45–53, Dec. 2022, doi: 10.24198/cropsaver.v5i2.41425.

[7] A. Kurniawan *et al.*, "Pengendalian Hama Burung Pipit Menggunakan Gelombang Ultrasonik Pada Lahan Sawah Musim Kemarau di Tasikmalaya," *Jurnal Pengabdian Cendikia*, vol. 2, no. 8, pp. 28–33, 2023, doi: 10.5281/zenodo.10085123.

[8] I. Lucia Kharisma, G. Purnama Insany, A. Rizki Firdaus, and D. Nasrulloh, "Overcoming the Impact of Bird Pests on Rice Yields Using Internet of Things Based YOLO Method," *International Journal of Engineering and Applied Technology*, vol. 7, no. 2, pp. 18–32, Nov. 2024, doi: doi.org/10.52005/ijeat.v7i2.111.

[9] N. W. I. Valerience, N. A. Norazlimi, and N. B. M. Abdullah, "Global Perspectives on Pest Bird Research and Management Methods: Implications for Agricultural Areas in Malaysia," in *IOP Conference Series: Earth and Environmental Science*, Institute of Physics, 2025, pp. 1–11. doi: 10.1088/1755-1315/1470/1/012012.

[10] M. Fakhrul Hirzi, H. Arahman, and Y. Wibisono, "Smart Ricefield: Development of an Automated Bird Pest Repellent System in Rice Fields Based on IOT and Artificial Intelligence," *Jurnal Ilmiah Multidisiplin Indonesia (JIM-ID)*, vol. 4, no. 10, pp. 1517–1529, 2025, doi: 10.58471/esaprom.v4i10.

[11] S. Lupica, S. Privitera, A. Trusso Sfrazzetto, E. Cerruto, and G. Manetto, "Toward Modern Pesticide Use Reduction Strategies in Advancing Precision Agriculture: A Bibliometric Review," Oct. 01, 2025, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/agriengineering7100346.

[12] S. Saha, O. D. Kucher, A. O. Utkina, and N. Y. Rebouh, "Precision agriculture for improving crop yield predictions: a literature review," Jul. 21, 2025, *Frontiers Media SA*. doi: 10.3389/fagro.2025.1566201.

[13] M. Habil Cahaya Gusti *et al.*, "Diseminasi Teknologi Tepat Guna Berbasis Lingkungan Untuk Mitigasi Hama Burung dan Ketahanan Pertanian di Desa Pakuluran Kabupaten Pandeglang," *Jurnal Pengabdian Ilmu Pendidikan, Sosial, dan Humaniora*, vol. 5, no. 1, pp. 229–239, 2025, [Online]. Available: http://journal.unram.ac.id/index.php/darmadiksani

[14] K. Yanel, "Alat Pengusir Hama Burung," *Manutech: Jurnal Teknologi Manufaktur*, vol. 15, no. 1, pp. 98–104, 2023, doi: https://doi.org/10.33504/manutech.v15i01.287.

[15] M. Dalal and P. Mittal, "A Systematic Review of Deep Learning-Based Object Detection in Agriculture: Methods, Challenges, and Future Directions," *Computers, Materials and Continua*, vol. 84, no. 1, pp. 57–91, 2025, doi: 10.32604/cmc.2025.066056.

[16] W. Yang and X. Qiu, "A Novel Crop Pest Detection Model Based on YOLOv5," *Agriculture (Switzerland)*, vol. 14, no. 2, pp. 1–23, Feb. 2024, doi: 10.3390/agriculture14020275.

[17] M. I. F. Mohamad Sa'ad, R. Che Ismail, S. Z. M. Naziri, M. N. M. Isa, and A. H. Mohd Shapri, "FPGA-accelerated bird detection algorithm for paddy farmers," *Majlesi Journal of Electrical Engineering*, vol. 19, no. 3, pp. 1–11, Sep. 2025, doi: 10.57647/j.mjee.2025.17276.

[18] I. N. Yulita, M. F. R. Rambe, A. Sholahuddin, and A. S. Prabuwono, "A Convolutional Neural Network Algorithm for Pest Detection Using GoogleNet," *AgriEngineering*, vol. 5, no. 4, pp. 2366–2380, Dec. 2023, doi: 10.3390/agriengineering5040145.

[19] C. A. Nwonye, K. A. Akpado, T. L. Alumona, O. S. Oguejiofor, and D. O. Amaefule, "Monitoring of Migration of Harmful Sparrow Birds in a Rice Farm Using Convolutional Neural Network (CNN) Model," *UNIZIK Journal of Engineering and Applied Sciences*, vol. 3, no. 4, pp. 1144–1162, 2024.

[20] S. Yuliany and A. Nur Rachman, "Implementasi Deep Learning pada Sistem Klasifikasi Hama Tanaman Padi Menggunakan Metode Convolutional Neural Network (CNN)," *Jurnal Buana Informatika*, vol. 13, no. 1, pp. 54–65, Apr. 2022, doi: doi.org/10.24002/jbi.v13i1.5022.

[21] R. Syahputri, W. Winarto, S. Trisnawati, and T. Taufik, "Android-based Detection of Melon Leaf Diseases Using Convolutional Neural Network and TensorFlow," vol. 9, no. 1, pp. 48–59, Jul. 2025, doi: 10.31289/jite.v9i1.14542.

[22] S. Raj, S. Garyali, S. Kumar, B. E. Scholar, and S. Shidnal, "Image based Bird Species Identification using Convolutional Neural Network," *International Journal of Engineering Research & Technology (IJERT)*, vol. 9, no. 6, pp. 346–351, Jun. 2020, [Online]. Available: www.ijert.org

[23] A. Roihan, M. Hasanudin, and E. Sunandar, "Evaluation Methods of Bird Repellent Devices in Optimizing Crop Production in Agriculture," in *Journal of Physics: Conference Series*, Institute of Physics Publishing, 2020. doi: 10.1088/1742-6596/1477/3/032012.

[24] M. O. Arowolo, F. T. Fayose, J. A. Ade-Omowaye, A. A. Adekunle, and S. O. Akindele, "Design and Development of an Energy-efficient Audio-based Repellent System for Rice Fields," *International Journal of Emerging Technology and Advanced Engineering*, vol. 12, no. 10, pp. 82–94, Oct. 2022, doi: 10.46338/ijetae1022_10.

[25] Y. C. Chen, J. F. Chu, K. W. Hsieh, T. H. Lin, P. Z. Chang, and Y. C. Tsai, "Automatic wild bird repellent system that is based on deep-learning-based wild bird detection and integrated with a laser rotation mechanism," *Sci. Rep.*, vol. 14, no. 1, pp. 1–15, Dec. 2024, doi: 10.1038/s41598-024-66920-2.

[26] A. Saufi and Suharjito, "White rice stem borer pest detection system using image-based convolution neural network," in *Procedia Computer Science*, Elsevier B.V., 2024, pp. 518–527. doi: 10.1016/j.procs.2024.10.278.

[27] C. B. Balingbing *et al.*, "Application of a multi-layer convolutional neural network model to classify major insect pests in stored rice detected by an acoustic device," *Comput. Electron. Agric.*, vol. 225, pp. 1–13, Oct. 2024, doi: 10.1016/j.compag.2024.109297.

[28] A. Dwicahyo, I. Mufandi, A. R. Nurfadila, M. T. Ardani, and U. Dzilhilmi, "Early Detection of Disease in Chicks Using CNN on Bangkok Chicken Health," *Buletin Ilmiah Sarjana Teknik Elektro*, vol. 6, no. 2, pp. 126–141, Jun. 2024, doi: 10.12928/BISTE.V6I2.10245.

[29] K. Eskioğlu and D. Özdemir, "Transforming Poultry Health Management with AI-Powered Disease Detection," *Turkish Journal of Agriculture - Food Science and Technology*, vol. 13, no. s3, pp. 4022–4028, Dec. 2025, doi: 10.24925/turjaf.v13is3.4022-4028.8196.

[30] A. Auxtero *et al.*, "RICEGUARD: IoT Noise-Enabled Scarecrow with Image Processing through Machine Learning and SMS Notification for Rice Crop Protection Enhancement," *International Journal of Computing Sciences Research*, vol. 9, pp. 3739–3758, 2025, doi: 10.25147/ijcsr.2017.001.1.247.

[31] M. Gärtler, V. Khaydarov, B. Klöpper, and L. Urbas, "The Machine Learning Life Cycle in Chemical Operations – Status and Open Challenges," Dec. 01, 2021, *John Wiley and Sons Inc*. doi: 10.1002/cite.202100134.

[32] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, International Conference on Learning Representations, ICLR, 2015, pp. 1–15. Accessed: Mar. 13, 2026. [Online]. Available: https://arxiv.org/pdf/1412.6980

[33] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15, pp. 1929–1958, 2014, Accessed: Mar. 13, 2026. [Online]. Available: https://www.jmlr.org/papers/volume15/srivastava14a/srivastava14a.pdf?utm_content=buffer79b43&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer,

[34] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in *Proceedings of the 36th International Conference on Machine Learning,* Long Beach, California: PMLR, Sep. 2019, pp. 6105–6114.

[35] F. Zhuang *et al.*, "A Comprehensive Survey on Transfer Learning," in *Proceedings of the IEEE*, Jul. 2020, pp. 43–76.

[36] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, pp. 861–874, 2006, doi: 10.1016/j.patrec.2005.10.010.