Improving Chicken Disease Classification Based on Vision Transformer and Combine with Integrated Gradients Explanation

Huong Hoang Luong, Triet Minh Nguyen
Information Technology Department
FPT University, Can Tho, Viet Nam

***Abstract*—Chicken diseases are an important problem in the livestock industry, affecting the health and production performance of chicken flocks worldwide. These diseases can seriously damage the health of chickens, reduce egg production, or increase mortality, causing great economic losses to farmers. Therefore, detecting and preventing diseases in chickens is a top concern in the livestock industry, to ensure the health and sustainable production of chicken flocks. In recent years, advances in machine learning techniques have shown promise in solving challenges related to image diagnosis and classification. Leveraging the power of machine learning models, we propose the ViT16 model for disease classification in chickens, demonstrating its potential in assisting healthcare professionals to diagnose chicken flocks more effectively. In this study, ViT16 demonstrated its potential and strengths when compared with 5 models in the CNN architecture and ViT32 in the ViT architecture in the task of classifying chicken disease images with an accuracy of 99.25% - 99.75% - 100% - 98.25% in four experimental scenarios with our enhanced dataset and fine-tuning. These results were generated from transfer learning and model tuning on an augmented dataset consisting of 8067 images classified into four classes: Coccidiosis, New Castle Disease, Salmonella, and Healthy. Furthermore, the Integrated Gradients explanation has an important role in increasing the transparency and understanding of the image classification model, thereby improving and optimizing model performance. The performance evaluation of each model is done through in-depth analysis, including metrics such as precision, recall, F1 score, accuracy, and confusion matrix.**

Keywords*—*Vision Transformer; ViT16; classification chicken disease; transfer learning; fine-tuning; image classification; integrated gradients explanation

I. Introduction

Chicken diseases are one of the important and worrying problems in the poultry industry. Chickens are considered one of the most important types of livestock, providing the main source of food for humans worldwide [1]. However, chickens often encounter a variety of infectious and non-infectious diseases that affect their health and production performance. The importance of chickens in livestock farming comes not only from the aspect of providing meat and eggs to consumers but also from the ability to create stable income for farmers [2]. Chickens provide an important source of income for farmers and breeders around the world, especially in rural and agriculturally developed areas [3]. However, when chickens get sick, the consequences can reduce productivity and product quality, causing great economic loss for farmers [4] [5]. Chicken diseases can spread quickly in chicken flocks, leading to mass deaths and reducing the economic value of chicken flocks. Using artificial intelligence and machine learning methods, technology can be used to diagnose chicken diseases early as an important method to control and prevent the spread of infectious diseases in chicken flocks [6] [7] [8]. Technology helps detect disease symptoms early, thereby allowing farmers to implement timely control and treatment measures, helping to minimize losses and increase livestock performance.

The Food and Agriculture Organization (FAO) projects that worldwide chicken meat output will be 103.5 million tons in 2012, accounting for approximately 34.3% of global meat production. A survey in the [9] study was carried out with households affected by severe aspergillosis. Data were collected from February 2018 to July 2019 from 183 households. The average risk of disease and mortality is 39% and 26% in chickens, 42% and 22% in turkeys, respectively, with young birds having a higher risk of disease and mortality than young birds. adult poultry. This loss causes economic losses in the chicken industry due to increased mortality, reduced meat and egg production, and poor growth. Viral infections, including Newcastle disease, infectious bursal disease, and avian influenza lead to economic costs in poultry farms, including bird losses, lower productivity, and employment losses [10]. Poultry farming is impacted by several viral, bacterial, parasitic, and fungal infections, which result in reduced appetite, weight loss, lower egg production, and greater mortality, leading to significant economic losses [11]. The highly pathogenic avian influenza epidemic in the United States led to fewer jobs, poorer productivity, and decreased tax collections, hurting both infected and non-infected farms [12]. The resurgence of Newcastle disease in village chicken populations in Tanzania caused major economic losses for small and medium farmers, affecting their predicted earnings and resulting in considerable economic burdens [13].

Advancements in machine learning approaches, such transfer learning [14] and fine-tuning [15], are increasingly being used to improve detection and classification accuracy. Transfer learning enables models pre-trained on big datasets to be applied to new tasks with less labeled data, making it ideal for medical imaging applications where annotated datasets are rare. Fine-tuning pre-trained models by modifying their parameters to better match the unique characteristics of the target task, resulting in improved performance. Huong Hoang Luong et al. [16] [17] [18] used fine-tuned transfer learning to categorize human skin, monkeypox, and brain tumors in the international medical field. By combining these machine learning algorithms into picture categorization, doctors may increase diagnosis accuracy, minimize screening time, and provide better patient care.

Vision Transformers (ViT) [19] represent a recent advancement in the realm of computer vision, offering a novel approach to image recognition tasks [20]. Unlike traditional Convolutional Neural Networks (CNNs), ViT utilizes a self-attention mechanism to capture long-range dependencies in images, enabling effective feature extraction and representation. This architecture consists of multiple Transformer blocks, each containing self-attention layers and feed-forward neural networks. By leveraging self-attention, ViT can effectively process images without relying on spatial hierarchies, making it suitable for tasks such as image classification, object detection, and segmentation [21]. The ViT16 model, in particular, is a variant of the Vision Transformer architecture that has demonstrated impressive performance in various computer vision tasks. It consists of 16 Transformer blocks, each with self-attention layers and feed-forward neural networks. Through extensive training on large-scale datasets, ViT16 has learned to extract informative features from input images and make accurate predictions. Moreover, ViT16 has shown robustness to variations in image content and background noise, making it suitable for real-world applications in medical image analysis.

In this study, we will use the Integrated Gradients explanation. Integrated gradients were proposed by Sundararajan et al. [22] to explain the predictions of our machine learning model. Interpretation is an important part of understanding and enhancing model transparency, especially in medical applications like ours. By explaining, we will have a more detailed look at how the model makes decisions, helping us better understand predictions and increasing the model's reliability in diagnosing and treating diseases. disability.

In this study, we suggested the ViT16 model of the Vision Transformer (ViT) to detect and classify diseases in chickens. In addition, we have deployed five well-known accumulated neural network (CNN) models (EfficiencyNetB3, ResNet50, VGG16, MobileNet, and InceptionV3) and ViT32 models of the ViT architecture to evaluate and compare with the model we have proposed.

The contributions of the work are:

- This study presents four scenarios to evaluate the classification efficiency of three common diseases. The classification of healthy and coccidiosis is carried out according to the first scenario. In the second scenario, we classify Healthy and New Castle Disease. The classification of healthy and salmonella is carried out in the third scenario. In the final scenario, we classify all four classes, including healthy, coccididiosis, new castle disease, and salmonella. The purpose of implementing these four scenarios is to determine whether the ViT model is effective when classifying each class individually or multiple classes at the same time. We propose a Vision Transformer transfer learning model based on the pre-trained ViT16 architecture for chicken disease image classification. By fine-tuning the model, we achieve promising results, surpassing other CNN architectures with accuracies up to 99.75% - 99.75% - 100% - 98.25% in four scenarios. This demonstrates the effectiveness of the ViT16 model for image classification, even without task-specific fine-tuning, achieving a transfer learning accuracy of 93%. Furthermore, fine-tuning the ViT models leads to a significant improvement in performance, increasing the accuracy from 93% to 98.25%.
- Demonstrate the effectiveness of the proposed model (ViT16) by implementing comparisons with five famous convolutional neural network architectures (EfficientNetB3, ResNet50, VGG16, MobileNet, InceptionV3) and a ViT32 model of the Vision architecture Transformer in the same setting.
- Experimental results show that the Integrated Gradients explanation is useful in helping to better understand how the model makes decisions by explaining how each pixel or feature affects the final prediction. Thanks to that, we can identify important areas in the image that the model pays attention to to make predictions. Integrated Gradients create transparent and easy-to-understand explanations, helping to increase confidence and trust in the model's predictions. This is especially important in medical and security applications, where transparency is extremely important.
- Early diagnosis will allow timely intervention and treatment measures, thereby minimizing mortality and loss in the herd. At the same time, preventing the spread of infectious diseases in chicken flocks will also play an important role in protecting health and improving food product quality. In addition, early diagnosis will also help optimize resource management and use in livestock operations, reduce waste, and increase production efficiency. This will have a positive impact on the food economy, helping to maintain and develop the livestock industry in a sustainable way.

There are five primary components to our study report. This section gives some general information about the research and discusses the approach to addressing the given difficulty. Section II includes references to related research, and the approach follows the relevant research section. Section III discusses each of the methodologies used in this paper. Section IV will discuss the experiments, including how we execute them and evaluate the deep learning model's correctness. Finally, in Section VI, we synthesize our findings and discuss the most essential parts of the research.

II. Related Works

The utilization of thermal-image processing and machine learning techniques for the detection and classification of avian diseases in chickens has gained significant attention in recent research. One notable study by M Sadeghi et al. [23] explored the application of support vector machines (SVM) and artificial neural networks (ANN) for disease classification in 14-day-old Ross 308 broilers infected with Newcastle Disease (ND) and Avian Influenza (AI), with two additional control groups included in the study. The paper demonstrated promising results, achieving high accuracy rates within 24 hours of virus infection. Specifically, SVM achieved an impressive accuracy of 97.2% for classifying AI and 100% accuracy for classifying

ND. These findings highlight the potential of machine learning algorithms for early disease detection and classification in poultry farming, contributing to improved disease management and prevention strategies.

The proposed paper [24] addresses the critical issue of early detection and classification of poultry diseases using deep learning techniques and image analysis of chicken fecal images. The model achieved an impressive accuracy of 97% utilizing the DenseNet method, showcasing its potential for practical poultry diagnostic applications. The dataset used for training and evaluation consists of 6812 images belonging to four different classes: healthy chicken, Coccidiosis, Salmonella, and Newcastle. Despite the significant progress in the field, there is a need for further research to explore the robustness and scalability of the proposed model across different poultry disease datasets and real-world scenarios.

Hoang Ngoc Tran et al. [25] present a novel approach utilizing the autoencoder and YOLOv6 model for the classification and detection of diseases in chicken flocks. The method achieved remarkable results, with an average accuracy of 99.15% and over 90% accuracy on the test dataset. The proposed approach demonstrates its versatility by being suitable for different chicken breeds from various countries and regions. This innovative method holds promise for improving the efficiency and accuracy of disease detection and classification in poultry farming, contributing to better management practices and disease control strategies. However, further studies are needed to validate the robustness and scalability of the proposed method across diverse poultry farming environments and disease scenarios.

The paper of Eduardo Carvalho Lira et al. [26] introduces a novel deep learning-based system designed for the early detection and classification of chicken diseases, including Salmonella, Coccidiosis, Healthy, and New Castle Disease. The study explores various convolutional neural network (CNN) models for categorical classification, with a focus on identifying the most efficient model based on the ratio of Maximum Validation Accuracy (MVA) to Least Validation Loss (LVL). Among the models evaluated, the ChicNetV6 model emerged as the best performer, achieving an efficiency score of 2.8198 and an impressive accuracy score of 94.49%. Notably, the total training time for the ChicNetV6 model was recorded at 1125 seconds, demonstrating its efficiency and computational feasibility. This research contributes to the advancement of automated disease detection and classification systems in poultry farming, with potential implications for enhancing disease management and prevention strategies in the industry. Further investigations may be warranted to assess the scalability and generalizability of the proposed system across different poultry farming environments and disease scenarios.

Nianpeng He et al. [27] presents a novel solution aimed at predicting diseases in chickens through the analysis of fecal images using deep Convolutional Neural Networks (CNN). Leveraging the XceptionNet deep learning framework, the proposed model demonstrates superior performance compared to other models, achieving an impressive accuracy rate of 94%. By leveraging pre-trained models and tailoring them to address the specific challenges of poultry disease prediction, the study contributes to the development of effective tools for early disease detection in poultry farming. The proposed model holds promise for application in real-world scenarios, offering a valuable resource for poultry disease detection and management. Further research may explore the scalability and applicability of the model across different poultry breeds and disease types, with potential implications for enhancing disease surveillance and control in the poultry industry.

Although the article [28] focuses on classifying poultry eggs using deep learning techniques, it does not go into the classification of diseases in chickens. The study proposes the use of Convolutional Neural Networks (CNN) in an unwashed egg classification system, classifying them into classes such as intact, cracked, soiled and soiled. The study compares the performance of three popular CNN architectures, ResNet34, ResNet50, and VGG19, using two different batch sizes (32 and 64) during training. Among the evaluated models, the VGG19 architecture achieved the highest accuracy of 97.33% when trained with a batch size of 64. While this study provides valuable insights into egg classification using deep learning, but its focus remains different from chicken disease classification, providing an additional perspective on the application of deep learning in poultry farming. Utilizing Convolutional Neural Networks (CNN) implemented in Keras/TensorFlow, Study of Moch. Kholil et al. [29] achieved an impressive accuracy rate of 95.28% in accurately predicting the classification of infectious diseases suffered by chickens. The research highlights the effectiveness of CNNs in disease classification tasks, particularly in the context of poultry farming. This work contributes valuable insights into leveraging deep learning techniques for disease diagnosis and monitoring in poultry, demonstrating the potential of advanced technology in enhancing animal health management practices.

Huong Hoang Luong et al. study in machine learning diagnostics is significant for us. In [30], Huong Hoang Luong and colleagues proposed a strategy using a model we propose as Vision Transformer (ViT), which has recently been improved by applying transfer learning methods. delivered to create strawberry disease.. The research's objective is to train this model to detect certain illnesses while fine-tuning the outcomes to attain high accuracy. The strawberry photographs in the collection are organized into seven categories, with a focus on strawberry leaf, fruit, and flower illnesses. The ViT model outperformed a comparable strategy for strawberry illness classification, with 92.7% accuracy on the Strawberry Disease Detection dataset.

The paper [31] presents a novel system for detecting and classifying poultry diseases, integrating two core algorithms: YOLO-V3 for object detection and ResNet50 for image classification. YOLO-V3 segments regions of interest (ROIs) from fecal images, while ResNet50 classifies the segmented images into four health conditions: Healthy, Coccidiosis, Salmonella, and New Castle Disease. Training is performed on a dataset of 10,500 chicken fecal images from the Zenodo open database, with oversampling and image augmentation techniques used to address class imbalance. YOLO-V3 achieves a mean average precision of 87.48% for detecting ROIs, and ResNet50 achieves a classification accuracy of 98.7%. Experimental results demonstrate the system's ability to accurately identify prevalent poultry diseases, offering potential support to poultry farmers and veterinarians in farm settings.

III. Methodology

A. The Research Implementation Procedure

Overall, in this study, we used a combination of 11 processes to produce the results, the main processes are shown in Fig. 1. Details of the steps are given below:

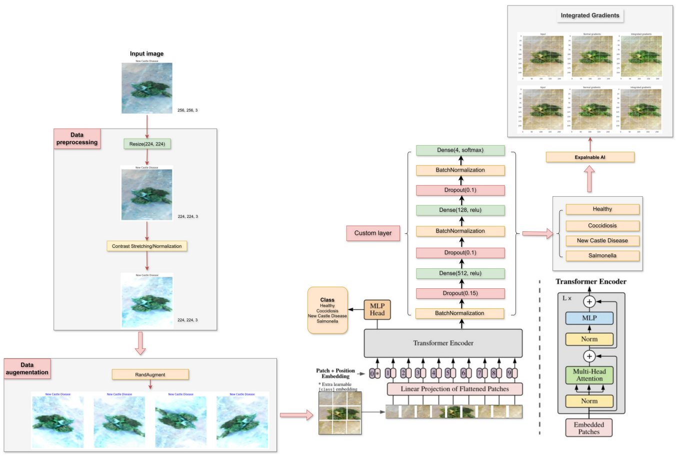

Fig. 1. Proposed architecture.

1) *Data collection:* Choosing an appropriate dataset is critical in the field of machine learning since it has a direct impact on model performance and generalizability. A high-quality dataset guarantees that the model is trained on a wide and representative sample set, allowing it to discover strong patterns and make accurate predictions in real-world circumstances. In the case of chicken illness detection, using the correct dataset enables researchers to create models that can accurately diagnose chicken diseases from photos, allowing for prompt treatment and resolution.
2) *Pre-processing Data:* Use image preprocessing techniques, for example, adjusting brightness and contrast to increase the quality and visibility of images, making them more suitable for future classification tasks. Change the input size to 224x224x3 to ensure synchronization, and use the random function in the Keras library for further processing.
3) *Dividing dataset into three categories train, validation, and test:* To ensure robust model training, validation, and evaluation, the dataset is divided into three subsets: training, validation, and test. The training set, which comprises 80% of the data, is utilized for model training. Meanwhile, the validation set, consisting of 10% of the data, is employed to monitor training progress and fine-tune the model to prevent overfitting. Finally, the test set, also comprising 10% of the data, is reserved to assess the final model's performance on previously unseen data. Stratified splitting is employed to maintain a balanced representation of classes across all subsets, facilitating effective model training and evaluation.
4) *Data Augmentation:* To augment the dataset, enhance its diversity, establish credibility, and mitigate overfitting, a range of data augmentation techniques are employed. The objective is to multiply the number of images fourfold, thus expanding the dataset from the initial 2000 images to a later count of 8000. This augmentation process encompasses the creation of new training samples through the application of diverse transformations and modifications to the existing images. This methodology effectively expands the dataset without necessitating the collection of additional data.
5) *Building the model:* To carry out the experiment, we utilized five convolutional neural network (CNN) architecture models and two Vision Transformer models. We keep the model's fundamental processing layers while making the required changes to improve its performance for our unique goal. This personalized strategy enabled us to obtain outstanding outcomes when training and testing with Keras' model library.
6) *Applying Transfer learning*: Transfer learning enables the utilization of pre-trained models that have been trained on similar tasks, such as general image classification. These models have already acquired fundamental features from large datasets, thereby saving time and effort that would otherwise be required to train a model from scratch. By leveraging pre-trained models, the amount of engineering work and resources needed to deploy the model across health systems is significantly reduced.
7) *Retrain the model with Fine-Tuning:* Fine-tuning involves tweaking the parameters of a pre-trained model to better align with a specific task. Nonetheless, to implement these adjustments and enhance the model's performance, re-training is essential. The model is fine-tuned to optimize performance for the targeted task. Re-training facilitates further learning from additional data, enhancing the model's ability to generalize and make accurate predictions on unseen data.
8) *Validate and collect metrics to evaluate the model:* By analyzing metrics like accuracy, precision, recall, and F1-score, it is possible to assess how well our model performs on data it has not been trained on. This assessment process helps in evaluating the model's effectiveness on the test dataset and offers insights into its performance under different circumstances. The conclusions drawn from the evaluation phase guide modifications to the model's hyperparameters, such as the learning rate, number of epochs, batch size, and neural network structure. These modifications are suggested based on the evaluation outcomes, with the goal of improving the model's performance and ensuring its ability to generalize to new data.
9) *Visual explanation by Integrated Gradients:* Integrated Gradients create easy-to-understand explanations for machine learning model predictions by calculating the influence of each input feature on the final prediction. By providing information about how each input feature contributes to the final prediction, Integrated Gradients help increase the transparency of the machine learning model. This can be useful in medical, financial and legal applications where transparency is important.
10) *Comparison with other advanced methods:* Comparing with other advanced techniques aids in assessing

model performance and gauging the effectiveness and novelty of the proposed approach relative to already explored and recognized methods. This enables the evaluation of which aspects of your strategy are more efficacious than others and which ones necessitate modification.

11) *Showing the result:* The outcomes and visual representations post comparison will be presented through confusion matrices, line graphs, and tables. These results illustrate the model's real-world performance and its effectiveness in diagnosing chicken disease.

B. Pre-processing Image

Pre-processing plays a crucial role in preparing image data for machine learning tasks as it enhances the quality, consistency, and informativeness of the images, thereby improving model performance. In our study, we conducted the following data pre-processing steps:

1) *Resize image:* A critical component of image preprocessing is ensuring uniform input size. To accomplish this, we resized all images to a consistent size of 224 pixels in width and 224 pixels in height, as determined by (Eq. 1):

$$I_{ReSize}(new_{width}, new_{height}) = I_{ReSize}(224, 224) \tag{1}$$

2) *Add Weighted:* We adjusted the brightness of the photos by -0.15 to improve the visibility of essential elements, especially in darker locations. This minor change demonstrated in (Eq. 2) improves feature visibility while preserving overall contrast.

$$B_{adjusted} = B_{original} - 0.15 \tag{2}$$

We used a contrast enhancement factor of 1.8 to highlight the differences between successive pixel intensities. This stage reveals tumor borders and structural features, facilitating precise diagnosis and categorization. In (Eq. 3), double the original contrast value by 1.5 to increase contrast. This factor can be fine-tuned based on the image content and desired amount of focus on differences in pixel intensities.

$$C_{adjusted} = C_{original} * 1.5 \tag{3}$$

3) *Data Augmentation:*
We used data augmentation techniques to scale the generated training dataset by making various changes on the input samples. First, we extracted 500 images as a subset for each class because the proportion of images of the four classes is not equal, this avoids leading to data imbalance as well as the model's learning ability. Common enhancement methods then include geometric transformations such as rotation, scaling, and flipping, as well as color and contrast adjustments. Finally, we obtained the result that the number of original images increased from 2000 to 8000 images. By expanding the data set, the model is exposed to more variables and situations, leading to better generalization and performance in real-world applications. In summary, data augmentation is a crucial method that enhances the performance and generalization capacity of machine learning models, particularly in scenarios where extensive and varied datasets are lacking.

C. Transfer Learning and Fine-Tuning of our Proposed Model

Transfer learning refers to a technique in machine learning and deep learning where a model is initially trained on a large dataset and then repurposed (transferred) to address a related or similar problem. Rather than commencing training from scratch with a small dataset, transfer learning enables leveraging the insights and knowledge acquired from prior training on extensive datasets to enhance the model's performance on a new dataset [32]. Throughout the training process, existing model parameters are reused. Consequently, transfer learning utilizes the pre-trained model layers instead of initiating training anew, thereby enhancing the model's accuracy.

After applying transfer learning, fine-tuning the model can help improve results. Fine-tuning involves adjusting and updating certain parts of the pre-trained model, such as the final layers, to better suit the specific problem at hand. To preserve the ability to extract low-level features acquired during pre-training, the initial layers of the model are often frozen during this process. By freezing these layers, the focus of adaptation shifts to the later layers, which are responsible for task-specific learning, thereby maximizing the effectiveness of training. Additionally, fine-tuning requires adjusting hyperparameters, such as the number of training epochs, batch size, hidden layer configuration, and learning rate, to optimize model performance and prevent overfitting.

We used a hyperparameter search to fine-tune the model in order to get the best results without overfitting. The many combinations of training epochs, batch sizes, hidden layer configurations, and learning rates were investigated in this search. The following hyperparameters were chosen based on the findings in order to achieve a decent balance between training efficiency and performance:

- Training epochs: 20
- Number of batches: 16
- Hidden layer set up: [512, 128]
- Rate of learning: defaut.

During fine-tuning the model, we unlock and tune the last 20 layers of the model without adjusting the rest of the model. The main reason is that the final layers often contain more complex and specific information about the data of the specific task we are training the model for. After unlocking and fine-tuning the last 20 layers, we added a custom layer set consisting of nine layers of three layer types Dense, BatchNormalization, Dropout. Our architecture is depicted in Fig. 2 By adjusting only the final layers, we can preserve the more general and abstract features learned from big data and only adjust them to suit our specific task. This helps minimize the risk of overfitting and increases the generality of the model on new data sets.

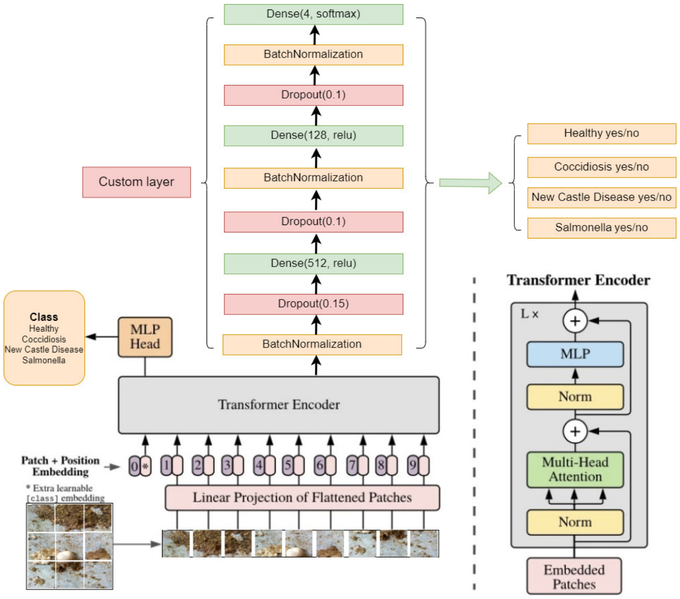

Fig. 2. Fine-tuning architecture.

D. Visual Explanation by Integrated Gradients

The need to use explanations is to better understand how the model makes decisions and makes predictions. This enhances model transparency and reliability, especially in fields such as healthcare, where a detailed explanation of the decision-making process can be extremely important for diagnosis and treat diseases.

Integrated Gradients is a method for explaining machine learning model predictions, used to understand the model's decision-making process based on inputs. This method calculates the importance of each input feature by integrating over the path from the reference point to the data point under consideration. When calculating, each feature is gradually changed from its reference value to its current value, helping to determine how each feature affects the model's final prediction.

Integrated Gradients have many advantages over other interpretation methods. First, it is computationally efficient and easy to understand, allowing to determine the importance of each feature accurately. Second, this method does not require specific information about the structure or characteristics of the model, making it flexible and applicable to many different types of machine learning models. Finally, Integrated Gradients enable both quantitative and qualitative interpretation, providing a comprehensive view of how the model makes decisions.

The use of Integrated Gradients has been used in various machine learning models, including deep neural networks, to enhance transparency and interpretability [33] [34]. The approach is suitable for both regression and classification models. When dealing with a non-scalar output, as seen in classification models or multi-target regression, the gradients are computed for a specific element of the output. In classification models, the gradient typically pertains to the output associated with the true class or the class predicted by the model.

Let's suppose we have an input instance x_1 a baseline instance x' and a model $M : X \to Y$ that operates on the feature space X and generates an output y in the output space Y. Now, let's define the function F as

- $F(x) = M(x)$ if the model output is a scalar;
- $F(x) = M_k(x)$ if the model output is a vector, with the index k denoting the k-th element of $M(x)$.

For instance, in case of a K-class classification, $M_k(x)$ is the probability of class k, which could be the true class corresponding to x or the highest probability class predicted by the model. The attributions $A_i\left(x, x'\right)$ for each feature x_i with respect to the corresponding feature x'_i in the baseline are computed as shown in Eq. 4;

$$A_i\left(x, x'\right) = \left(x_i - x'_i\right) \int_0^1 \frac{\partial F\left(x' + \alpha\left(x - x'\right)\right)}{\partial x_i} d\alpha \tag{4}$$

In summary, employing Integrated Gradients for visual explanations provides a promising method to improve the transparency, accountability, and reliability of machine learning models, thereby enhancing their utility and credibility in real-world applications. As illustrated in Fig. 3, analyzing the contribution of individual feature maps to the final decision provides valuable insights that experts and clinicians can leverage in future endeavors.

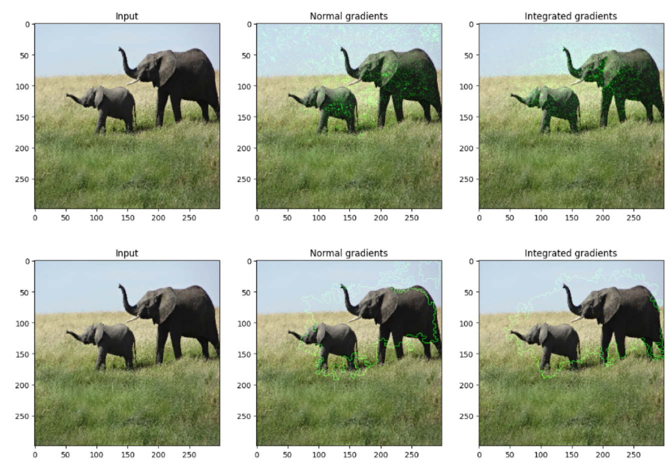

Fig. 3. The sample applying integrated gradients by keras library.

IV. Experiments

A. Dataset and Performance Metrics

An annotated dataset on poultry disease diagnostics for small and medium-sized poultry farmers includes images of poultry feces. Images of poultry droppings were taken in the Arusha and Kilimanjaro regions of Tanzania between September 2020 and February 2021 using the Open Data Kit (ODK) mobile app. The data set contains 8067 images, divided into 4 classes in Fig. 4: Coccidiosis(30.4%), Healthy(29.5%), New Castle Disease(7.8%), Salmonella(32.3%).

Due to imbalance in the data set, we randomly selected 500 images for each class to use for training. It is important to provide a variety of representations while reducing the risk of overfitting and improving the generalizability of the model. After image preprocessing and data enhancement, we obtained

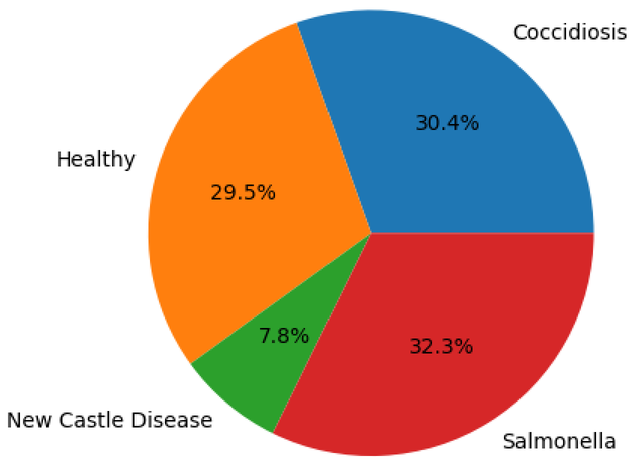

Fig. 4. Data set characteristics before processing.

a new dataset with 8000 images as shown in Fig. 5 from 2000 original images.

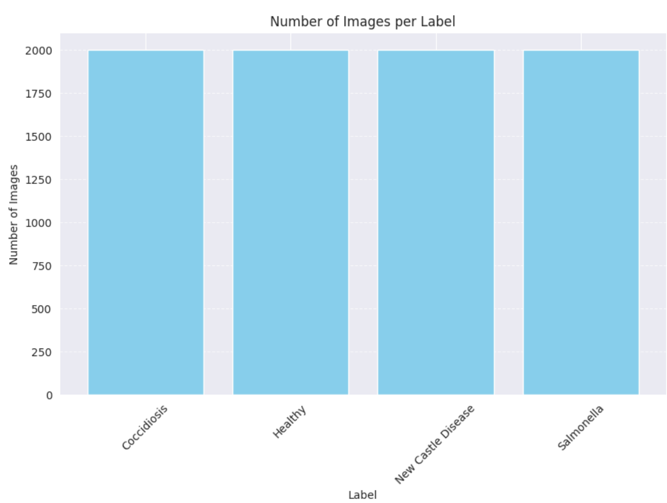

Fig. 5. Dataset characteristics after processing.

Evaluating the performance of a machine learning model is an important part of the research and implementation process. In the field of machine learning, there are many metrics used to evaluate the performance of a model, including Accuracy, Precision, Recall and F1-score.

Eq. 5 represents the ratio between the number of correct predictions and the total number of samples. Eq. 6 represents the accuracy of detecting Positive points. The higher this number, the more accurate the model receives Positive scores. Eq. 7 represents the ability to detect all positive, the higher this rate shows the lower the possibility of missing Positive points. Eq. 8 is a compromise number for Recall and Precision, used when it is necessary to consider both values, giving us a basis for choosing a model.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (6)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (7)$$

$$\text{F1} = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (8)$$

These metrics offer a holistic perspective on the effectiveness of a machine learning model, enabling users to precisely assess its capability to make predictions and identify significant instances.

B. Scenario 1: Classification of 2 Classes (Coccidiosis and Healthy)

TABLE I. THE RESULTS OF CLASSIFYING IMAGES INTO 2 CLASSES COCCIDIOSIS AND HEALTHY IN TRANSFER LEARNING

Transfer learning Without Augmentation				
Model	**ACC**	**Precision**	**Recall**	**F1**
EfficientNetB3	69,00%	69,00%	69,00%	68,99%
ResNet50	92,00%	92,06%	92,00%	91,99%
VGG16	95,00%	95,16%	95,00%	94,99%
MobileNet	100,00%	100,00%	100,00%	100,00%
InceptionV3	98,00%	98,07%	98,00%	97,99%
ViT32	98,00%	98,00%	98,00%	98,00%
ViT16 (Our Proposed)	**98,00%**	**98,00%**	**98,00%**	**98,00%**
Transfer learning With Augmentation				
Model	**ACC**	**Precision**	**Recall**	**F1**
EfficientNetB3	71,50%	71,67%	71,50%	71,44%
ResNet50	94,75%	95%%	94,75%	94,74%
VGG16	96,75%	96,75%	96,75%	96,75%
MobileNet	100,00%	100,00%	100,00%	100,00%
InceptionV3	98,50%	98,50%	98,50%	98,50%
ViT32	99,50%	99,50%	99,50%	99,50%
ViT16 (Our Proposed)	**99,25%**	**99,25%**	**99,25%**	**99,25%**

In this scenario, we apply transfer learning and fine-tuning in both cases with and without data augmentation to classify Coccidiosis and Healthy of seven different machine learning models. The results obtained in the transfer learning part in Table I show the effectiveness of the model when trained on the data set after augmentation. The accuracy of the proposed model has been improved from 98% to 99.25%. A bright spot besides the proposed model is that the MobileNet model also achieves high efficiency when achieving 100% accuracy. Regarding the fine-tuning showed in Table II, we obtain the results before and after enhancing the data set respectively as 98.00%-98.75%.

Fig. 6 and Fig. 7 depict a graphical representation of the training accuracy and loss on the augmented dataset. Throughout the training process, the two curves intersect multiple times, illustrating the model's ability to strike a balance between learning from the training data and generalizing to new data. In general, both the training accuracy and loss curves exhibit a smooth behavior without any significant disparity, thereby indicating the model's suitability and robustness in terms of generalization capability.

TABLE II. THE RESULTS OF CLASSIFYING IMAGES INTO 2 CLASSES COCCIDIOSIS AND HEALTHY IN FINE-TUNING

Fine-Tuning Without Augmentation				
Model	**ACC**	**Precision**	**Recall**	**F1**
EfficientNetB3	67,00%	70,64%	67,00%	65,47%
ResNet50	82,00%	83,33%	82,00%	81,81%
VGG16	97,00%	97,00%	97,00%	97,00%
MobileNet	97,00%	97,00%	97,00%	97,00%
InceptionV3	94,00%	94,00%	94,00%	94,00%
ViT32	97,00%	97,00%	97,00%	97,00%
ViT16 (Our Proposed)	**99,00%**	**99,00%**	**99,00%**	**99,00%**
Fine-Tuning With Augmentation				
Model	**ACC**	**Precision**	**Recall**	**F1**
EfficientNetB3	50,00%	25,00%	50,00%	33,33%
ResNet50	91,25%	91,25%	91,25%	91,25%
VGG16	97,50%	97,50%	97,50%	97,50%
MobileNet	99,50%	99,50%	99,50%	99,50%
InceptionV3	97,50%	97,50%	97,50%	97,50%
ViT32	99,00%	99,00%	99,00%	99,00%
ViT16 (Our Proposed)	**99,25%**	**99,25%**	**99,25%**	**99,25%**

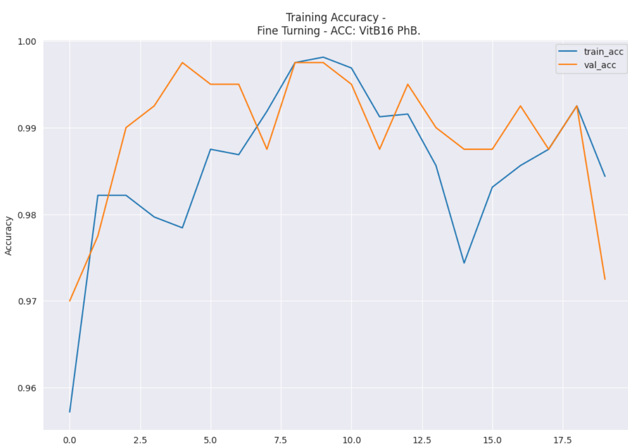

Fig. 6. Training accuracy and validation accuracy in fine-tuning of ours model (coccidiosis and healthy).

Fig. 8 presents the confusion matrix of 400 test images of Coccidiosis and Healthy. Fig. 9 is the result of the Integrated Gradients explanation. Through the two pictures above, we can see the transparency of the training process as well as overfitting does not happen.

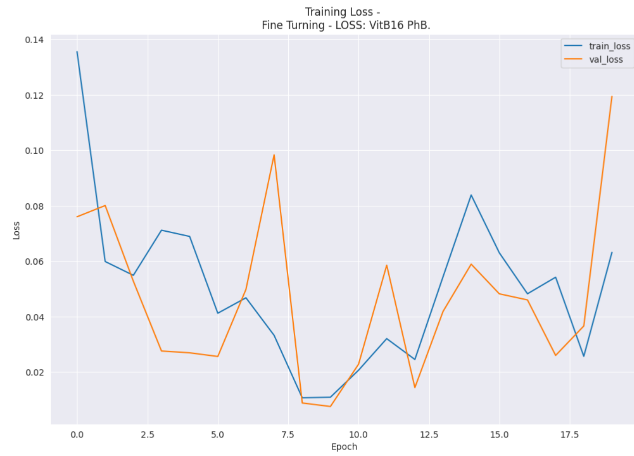

Fig. 7. Training loss and validation accuracy in fine-tuning of ours model (coccidiosis and healthy).

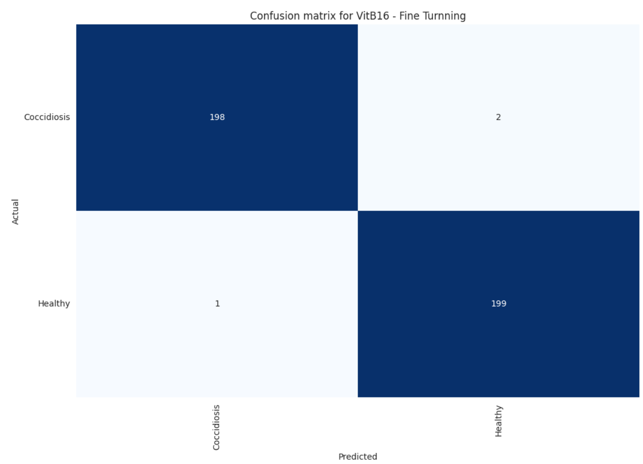

Fig. 8. Confusion matrix in fine-tuning of ours model (coccidiosis and healthy).

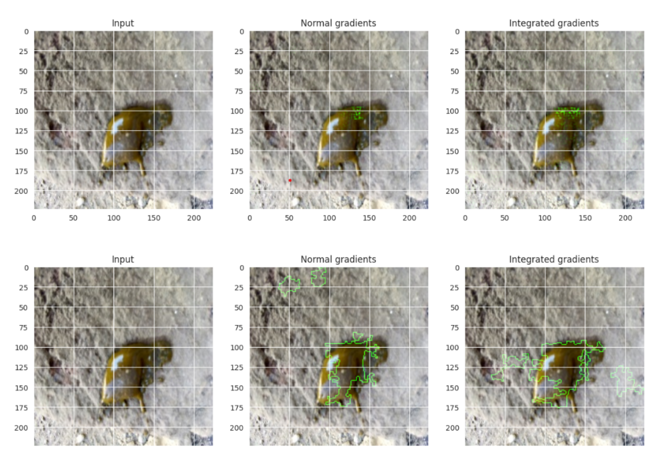

Fig. 9. Output of our model with integrated gradients explanation in scenario 1.

C. Scenario 2: Classification of 2 Classes (New Castle Disease and Healthy)

In this scenario, we classify the next two classes including New Castle Disease and Healthy. The scenario performs transfer learning and fine-tuning in both cases with and without data augmentation. The results obtained in the transfer learning part of the proposed model in Table III are 98.50% accuracy - an improvement of more than 3.5% compared to training on the original data set. Table IV also shows the effectiveness of fine-tuning when the obtained accuracy is 99.75%, which is higher than that of transfer learning.

TABLE III. THE RESULTS OF CLASSIFYING IMAGES INTO 2 CLASSES NEW CASTLE DISEASE AND HEALTHY IN TRANSFER LEARNING

Transfer learning Without Augmentation				
Model	**ACC**	**Precision**	**Recall**	**F1**
EfficientNetB3	76,00%	76,68%	76,00%	75,84%
ResNet50	86,00%	86,05%	86,00%	85,99%
VGG16	88,00%	88,06%	88,00%	87,99%
MobileNet	95,00%	95,16%	95,00%	94,99%
InceptionV3	89,00%	89,01%	89,00%	88,99%
ViT32	95,00%	95,16%	95,00%	94,99%
ViT16 (Our Proposed)	**95,00%**	**95,01%**	**95,00%**	**94,99%**
Transfer learning With Augmentation				
Model	**ACC**	**Precision**	**Recall**	**F1**
EfficientNetB3	74,75%	75,14%	74,75%	74,65%
ResNet50	88,75%	88,75%	88,75%	88,75%
VGG16	92,25%	92,49%	92,25%	92,24%
MobileNet	98,50%	98,50%	98,50%	98,50%
InceptionV3	96,50%	96,51%	96,50%	96,50%
ViT32	98,50%	98,52%	98,50%	98,50%
ViT16 (Our Proposed)	**98,50%**	**98,50%**	**98,50%**	**98,50%**

TABLE IV. THE RESULTS OF CLASSIFYING IMAGES INTO 2 CLASSES NEW CASTLE DISEASE AND HEALTHY IN FINE-TUNING

Fine-Tuning Without Augmentation				
Model	**ACC**	**Precision**	**Recall**	**F1**
EfficientNetB3	68,00%	68,11%	68,00%	68,00%
ResNet50	86%	87,50%	86%	85,85%
VGG16	95%	95%	95%	95%
MobileNet	93%	93,43%	93%	93%
InceptionV3	91,00%	91,00%	91,00%	91,00%
ViT32	97%	97%	97%	97%
ViT16 (Our Proposed)	**97%**	**97%**	**97%**	**97%**
Fine-Tuning With Augmentation				
Model	**ACC**	**Precision**	**Recall**	**F1**
EfficientNetB3	74,50%	74,62%	74,50%	74,46%
ResNet50	51,25%	75,31%	51,25%	36,05%
VGG16	94,25%	94,38%	94,25%	94,24%
MobileNet	98,75%	98,75%	98,75%	98,74%
InceptionV3	95,50%	95,54%	95,50%	95,49%
ViT32	99,50%	99,50%	99,50%	99,49%
ViT16 (Our Proposed)	**99,75%**	**99,75%**	**99,75%**	**99,74%**

Fig. 10 and Fig. 11 show the training process's accuracy and loss in the second scenario experiment. The two curves do not noticeably vary from one another during the training period and climb steadily. The model is adequate and has a great potential for generalization because the training and loss accuracy curves are generally smooth and show no discernible deviation between them.

Fig. 10. Training accuracy and validation accuracy in fine-tuning of ours model (new castle disease and healthy).

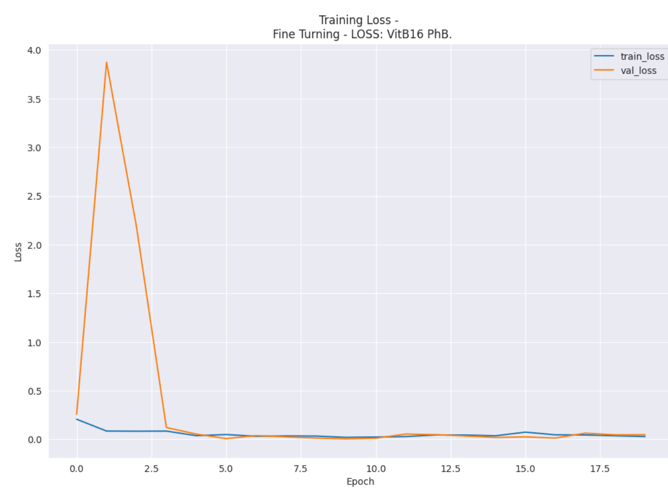

Fig. 11. Training loss and validation accuracy in fine-tuning of ours model (new castle disease and healthy).

Fig. 12 presents the confusion matrix of the 2-type classification scenario New Castle Disease and Healthy. From the matrix, we see that the model performs absolutely well when identifying the healthy class. At the same time, the New Castle Disease class has only one flaw. Fig. 13 is the result of the built-in Gradient explanation for this scenario.

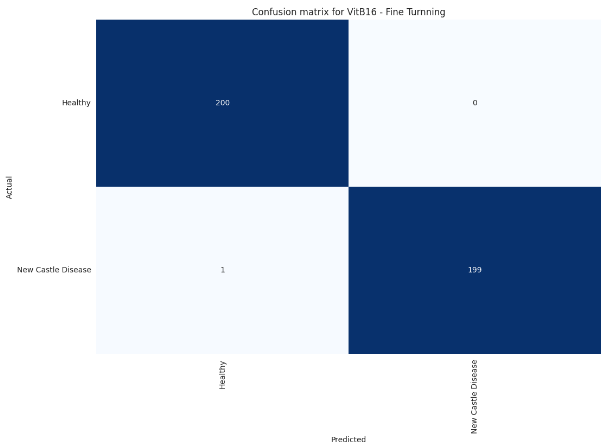

Fig. 12. Confusion matrix in fine-tuning of ours model (New castle disease and healthy).

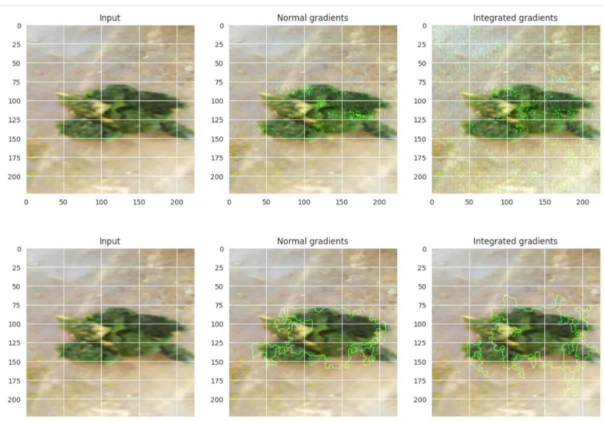

Fig. 13. Output of our model with integrated gradients explanation in scenario 2.

D. Scenario 3: Classification of 2 Classes (Salmonella and Healthy)

This scenario presents the results of classifying the two classes Salmonella and Healthy before and after data fortification. In Table V, we note the improvement in the accuracy of the proposed model by up to 8% when it reaches 98%. Furthermore, after fine-tuning, the proposed model has achieved 100% absolute accuracy with data augmentation and most other models in the Table VI have also improved.

Fig. 14 and Fig. 15 show the training process' accuracy and loss in the scenario 3 experiment. The two curves continuously rise and do not considerably diverge from one another during the training phase, demonstrating the transparency and dependability of the suggested model.

Fig. 16 presents the confusion matrix of the Salmonella and Healthy class classification test. From the matrix, we see that the model performs best when classifying chicken diseases with an accuracy rate of 100%. Fig. 17 is the result of the explanation of Integrated Gradients for Salmonella and Healthy classification.

TABLE V. The Results of Classifying Images Into 2 Classes Salmonella and Healthy in Transfer Learning

Transfer learning Without Augmentation				
Model	**ACC**	**Precision**	**Recall**	**F1**
EfficientNetB3	68,00%	68,26%	68,00%	67,88%
ResNet50	81,00%	81,61%	81,00%	80,90%
VGG16	88,00%	88,06%	88,00%	87,99%
MobileNet	93,00%	93,85%	93,00%	92,96%
InceptionV3	92,00%	92,61%	92,00%	91,97%
ViT32	92,00%	92,00%	92,00%	92,00%
ViT16 (Our Proposed)	**90,00%**	**90,06%**	**90,00%**	**89,99%**
Transfer learning With Augmentation				
Model	**ACC**	**Precision**	**Recall**	**F1**
EfficientNetB3	76%	76,87%	76%	75,80%
ResNet50	92,25%	92,33%	92,25%	92,24%
VGG16	94,25%	94,26%	94,25%	94,25%
MobileNet	96,50%	96,50%	96,50%	96,50%
InceptionV3	94,75%	94,76%	94,75%	94,75%
ViT32	97,75%	97,75%	97,75%	97,75%
ViT16 (Our Proposed)	**98%**	**98,02%**	**98%**	**98%**

TABLE VI. The Results of Classifying Images Into 2 Classes Salmonella and Healthy in Fine-Tuning

Fine-Tuning Without Augmentation				
Model	**ACC**	**Precision**	**Recall**	**F1**
EfficientNetB3	72%	72,32%	72%	71,89%
ResNet50	92%	92%	92%	92%
VGG16	95%	95%	95%	95%
MobileNet	95%	95,45%	95,00%	95%
InceptionV3	94%	94,28%	94%	93,99%
ViT32	97%	97,17%	97%	97%
ViT16 (Our Proposed)	**98%**	**98,07%**	**98%**	**98%**
Fine-Tuning With Augmentation				
Model	**ACC**	**Precision**	**Recall**	**F1**
EfficientNetB3	59,00%	63,33%	59,00%	55,37%
ResNet50	91,25%	91,55%	91,25%	91,23%
VGG16	92,25%	92,81%	92,25%	92,22%
MobileNet	95,25%	95,44%	95,25%	95,24%
InceptionV3	95,50%	95,50%	95,50%	95,50%
ViT32	98,75%	98,78%	98,75%	98,74%
ViT16 (Our Proposed)	**100,00%**	**100,00%**	**100,00%**	**100,00%**

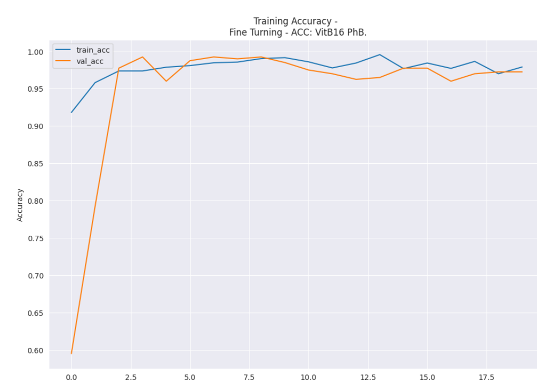

Fig. 14. Training accuracy and validation accuracy in fine-tuning of ours model (salmonella and healthy).

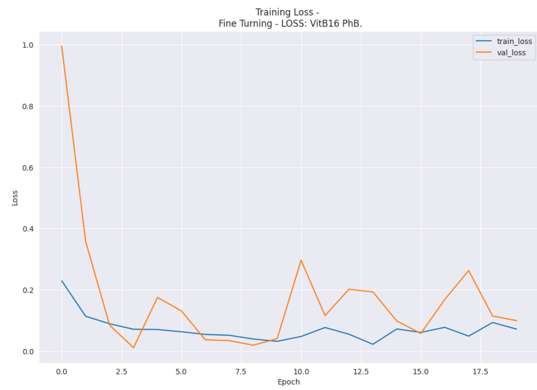

Fig. 15. Training loss and validation accuracy in fine-tuning of ours model (salmonella and healthy).

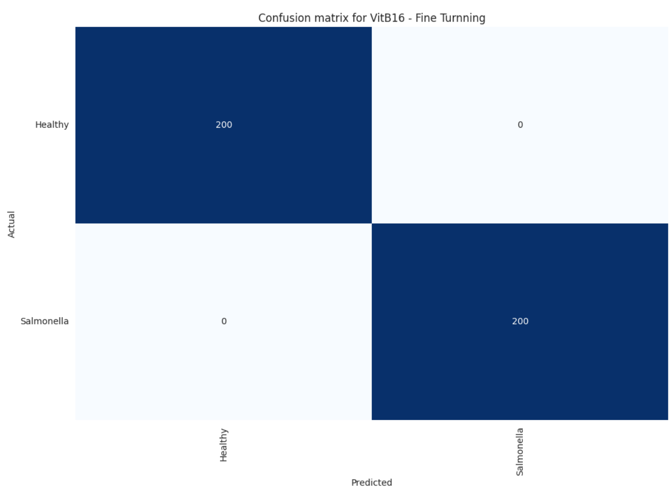

Fig. 16. Confusion matrix in fine-tuning of ours model (salmonella and healthy).

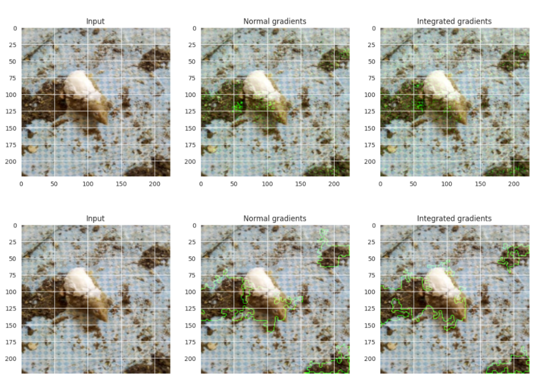

Fig. 17. Output of our model with integrated gradients explanation in scenario 3.

E. Scenario 4: Classification of 4 Classes (Coccidiosis, New Castle Disease, Salmonella and Healthy)

This is an important scenario that shows the strong performance of the proposed model when the classification problem has up to 4 classes. It can be seen that after the transfer learning process in Table VII, the proposed model achieved 97.75% accuracy when trained on the augmented data set. In contrast, the model only achieved 89.50% accuracy when trained on the original data set. After the stage of fine-tuning the proposed model with the augmented data set, the final result obtained in Table VIII has an accuracy of 98.25%.

TABLE VII. THE RESULTS OF CLASSIFYING IMAGES INTO 4 CLASSES COCCIDIOSIS, NEW CASTLE DISEASE, SALMONELLA AND HEALTHY IN TRANSFER LEARNING

Transfer learning Without Augmentation				
Model	**ACC**	**Precision**	**Recall**	**F1**
EfficientNetB3	40,50%	30,50%	40,50%	33,92%
ResNet50	70,50%	72,04%	70,50%	70,75%
VGG16	73,50%	74,83%	73,50%	73,84%
MobileNet	89,00%	89,22%	89,00%	89,03%
InceptionV3	84,00%	84,16%	84,00%	83,90%
ViT32	92,00%	92,13%	92,00%	92,00%
ViT16 (Our Proposed)	**89,50%**	**89,59%**	**89,50%**	**89,50%**
Transfer learning With Augmentation				
Model	**ACC**	**Precision**	**Recall**	**F1**
EfficientNetB3	41,37%	41,48%	41,37%	39,38%
ResNet50	73,50%	74,19%	73,50%	73,63%
VGG16	85,37%	85,51%	85,37%	85,40%
MobileNet	93,87%	93,86%	93,87%	93,86%
InceptionV3	91,87%	91,88%	91,87%	91,87%
ViT32	96,25%	96,24%	96,25%	96,24%
ViT16 (Our Proposed)	**97,75%**	**97,76%**	**97,75%**	**97,74%**

TABLE VIII. THE RESULTS OF CLASSIFYING IMAGES INTO 4 CLASSES COCCIDIOSIS, NEW CASTLE DISEASE, SALMONELLA AND HEALTHY IN FINE-TUNING

Fine-Tuning Without Augmentation				
Model	**ACC**	**Precision**	**Recall**	**F1**
EfficientNetB3	30,50%	18,57%	30,50%	20,95%
ResNet50	38,00%	20,89%	38,00%	25,48%
VGG16	56,50%	53,28%	56,50%	51,11%
MobileNet	86,00%	88,22%	86,00%	85,59%
InceptionV3	83,00%	83,78%	83,00%	83,02%
ViT32	95,00%	95,14%	95,00%	95,01%
ViT16 (Our Proposed)	**98,00%**	**98,03%**	**98,00%**	**97,99%**
Fine-Tuning With Augmentation				
Model	**ACC**	**Precision**	**Recall**	**F1**
EfficientNetB3	49,75%	39,66%	49,75%	43,24%
ResNet50	26,25%	31,32%	26,25%	12,48%
VGG16	90,25%	91,12%	90,25%	90,16%
MobileNet	92,25%	92,56%	92,25%	92,25%
InceptionV3	90,25%	90,25%	90,25%	90,24%
ViT32	97,25%	97,25%	97,25%	97,24%
ViT16 (Our Proposed)	**98,25%**	**98,25%**	**98,25%**	**98,24%**

Fig. 18 and Fig. 19 illustrate the training accuracy and loss in the test of scenario 4. Fig. 20 presents the confusion matrix of 800 test images of four classes including Coccidiosis, New Castle Disease, Salmonella and Healthy. Fig. 21 is the result of the explanation of the Integrated Gradient to classify the above four classes.

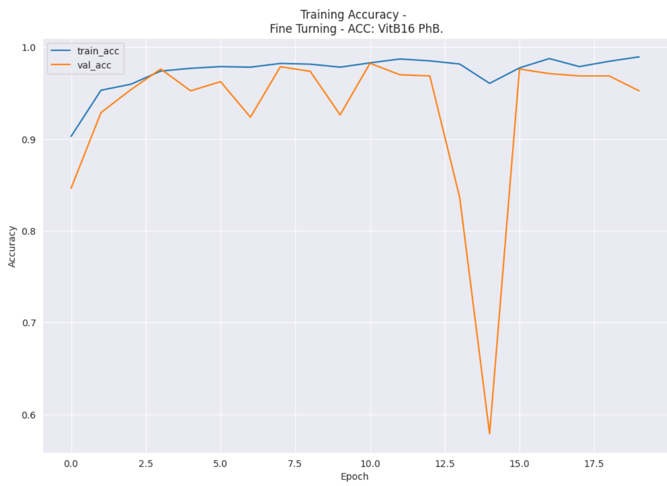

Fig. 18. Training accuracy and validation accuracy in fine-tuning of ours model (coccidiosis, new castle disease, salmonella and healthy).

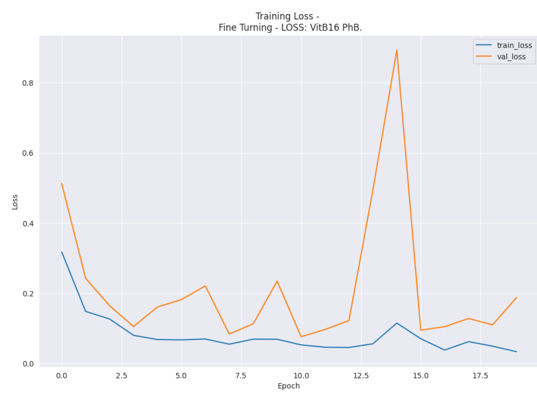

Fig. 19. Training loss and validation accuracy in fine-tuning of ours model (Coccidiosis, new castle disease, salmonella and healthy).

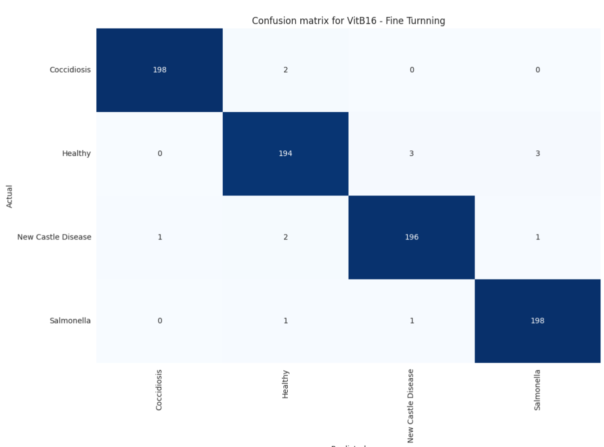

Fig. 20. Confusion matrix in fine-tuning of ours model (Coccidiosis, new castle disease, salmonella and healthy).

F. Comparison with other State-of-the-art Methods

This section completely compares our proposed method with several existing state-of-the-art classification methods.

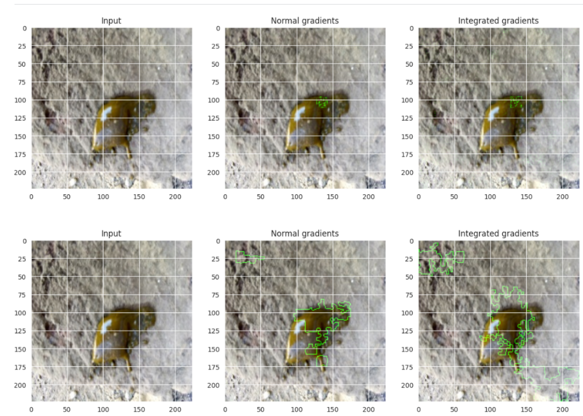

Fig. 21. Output of our model with Integrated Gradients Explanation in Scenario 4.

Table IX compares chicken disease classification methods using ViT and CNN architectures. It can be seen that the proposed model (fine-tuned ViT16) performs better than most recently published works on disease detection in chickens.

TABLE IX. Comparison with Others State-of-The-Art Methods

Ref.	Architecture	ACC
Moch. Kholil et al. [29]	CNN	95,28%
Mizanu Zelalem Degu et al. [31]	YOLOv3	98,70%
Dina Machuve et al. [35]	Xception	98,24%
Our Proposed Model (ViT16)		**98,25%**

Our proposed model outperforms most state-of-the-art methods in terms of accuracy and other evaluation metrics. Specifically, our model achieves higher accuracy than CNN, Xception and approximates the YOLOv3 model. The outstanding performance of our model can be attributed to its ability to effectively capture and classify features associated with poultry diseases, leveraging the strengths of deep learning techniques and innovative architectural design.

V. Discussion

Our study aimed to develop a robust model for poultry disease image classification, leveraging transfer learning and fine-tuning techniques. The results demonstrated the efficacy of this approach, with the proposed model (ViT16) achieving high accuracy rates across various scenarios. This outcome underscores the importance of utilizing pre-trained models and optimizing them for specific tasks, highlighting the potential of transfer learning in medical image analysis.

Furthermore, by meticulously fine-tuning hyperparameters and incorporating dense layers, dropout layers, and BatchNormalization, we successfully mitigated overfitting and improved classification accuracy. These findings underscore the importance of meticulous model development and optimization to achieve superior performance in medical image classification tasks.

Moreover, the integration of Integrated Gradients for visual explanation provided valuable insights into the model's decision-making process. This transparency not only enhances trust in the model's predictions but also facilitates error recognition and model improvement. The discussion also emphasizes the broader implications of the research, particularly in advancing diagnostics and anomaly detection in livestock farming.

In summary, through rigorous experimentation and optimization, our study contributes to the growing body of literature on transfer learning and deep learning applications in medical imaging, paving the way for future advancements in disease detection and diagnosis.

VI. CONCLUSION

In this work, we used transfer learning, a powerful machine learning method, to improve the model's performance in four-class classification. Transfer learning uses information obtained from a model that has been pre-trained on a large data set for a given task and applies it to another activity. In our study, we started with 5 models (EfficientNetB3, ResNet50, VGG16, MobileNet, InceptionV3) of CNN architecture and two models (ViT32, ViT16) of Vision Transformer architecture. These models have been trained on large amounts of data, which allows our model to inherit knowledge about common image aspects, allowing it to focus on the complexity of image classification.

Fine-tuning was important in developing a pre-trained model for our medical image classification application. To improve the performance of the model, we added dense layers, dropout layers, and BatchNormalization, as well as modified many hyperparameters. This combination of layers enables the network to learn complex patterns and relationships in the data, leading to efficient classification performance while minimizing overfitting, resulting in improved accuracy higher. After fine-tuning and training on the augmented data set, the proposed model (ViT16) achieved accuracy in 4 scenarios of 99.25% - 99.75% - 100% - 98.25% respectively. Compared to the situation without data augmentation, the model has achieved accuracy in 4 scenarios of 99% - 97% - 98% - 98% respectively.

To provide transparency and to better understand how the model makes decisions and makes predictions during training. We used Integrated Gradients for visual explanation. This helps experts understand the model's predictions to recognize errors and easily improve the model.

Our results demonstrate the robustness of the ViT16 model in the image classification problem compared to other popular models. Strong precision, accuracy, recall and F1 score demonstrate their usefulness in livestock production. This research actually makes a significant contribution as we tackle the rather rare problem of disease diagnosis in chickens. This research paves the way for future developments in image processing and diagnostics in livestock.

In the future, we will to continue to fine-tuning the model, expand the challenge, use other advanced visualization tools, and improve the dataset. In addition, evaluating different preprocessing techniques on different chicken disease images is also an issue that needs research. By undertaking this action, our objective is to enhance the Accuracy of the model, solidifying its role as a fundamental solution within the realm of avian image categorization.. Our ongoing efforts highlight the importance of artificial intelligence in improving diagnostics and anomaly detection.

ACKNOWLEDGMENT

Luong Hoang Huong was funded by the Vingroup Innovation Foundation (VINIF) 's Master, Ph.D. Scholarship Programme, code VINIF.2023.TS.049.

We would like to express my sincere gratitude to Duy Khanh Nguyen, and Bang Huu Do Dang for their invaluable support and assistance throughout the course of this research. Their expertise, guidance, and encouragement have been instrumental in the successful completion of this study. We are truly grateful for their dedication and commitment, which have greatly contributed to the quality and depth of this research endeavor.

REFERENCES

[1] A. Mottet and G. Tempio, "Global poultry production: current state and future outlook and challenges," *World's poultry science journal*, vol. 73, no. 2, pp. 245–256, 2017.

[2] H. Al-Khalaifah and A. Al-Nasser, "Importance of poultry industry during global crisis with special reference to covid-19 crisis," *International Multidisciplinary Scientific GeoConference: SGEM*, vol. 22, no. 6.1, pp. 89–96, 2022.

[3] M. K. Padhi *et al.*, "Importance of indigenous breeds of chicken for rural economy and their improvements for higher production performance," *Scientifica*, vol. 2016, 2016.

[4] W. Landman and J. Van Eck, "The incidence and economic impact of the escherichia coli peritonitis syndrome in dutch poultry farming," *Avian Pathology*, vol. 44, no. 5, pp. 370–378, 2015.

[5] S. M. Kinung'hi, G. Tilahun, H. M. Hafez, M. Woldemeskel, M. Kyule, M. Grainer, and M. P. Baumann, "Assessment of economic impact caused by poultry coccidiosis in small and large scale poultry farms in debre zeit, ethiopia," *International Journal of Poultry Science*, vol. 3, no. 11, pp. 715–718, 2004.

[6] H. Mbelwa, "Image-based poultry disease detection using deep convolutional neural network," Ph.D. dissertation, NM-AIST, 2021.

[7] S. Suthagar, G. Mageshkumar, M. Ayyadurai, C. Snegha, M. Sureka, and S. Velmurugan, "Faecal image-based chicken disease classification using deep learning techniques," in *Inventive Computation and Information Technologies: Proceedings of ICICIT 2022*. Springer, 2023, pp. 903–917.

[8] T. Shibanuma, Y. Nunomura, M. Oba, F. Kawahara, T. Mizutani, and H. Takemae, "Development of a one-run real-time pcr detection system for pathogens associated with poultry infectious diseases," *Journal of Veterinary Medical Science*, vol. 85, no. 4, pp. 407–411, 2023.

[9] D. Kalkayeva, A. Maulanov, P. Sobiech, M. Michalski, G. Kuzembekova, A. Dzhangabulova, N. Nurkhojayev, and N. Aldayarov, "Epidemiological characteristics and financial losses due to avian aspergillosis in households in the almaty region, republic of kazakhstan," *Frontiers in Veterinary Science*, vol. 10, p. 1141456, 2023.

[10] M. Sadiq and B. Mohammed, "The economic impact of some important viral diseases affecting the poultry industry in abuja, nigeria," *Sokoto Journal of Veterinary Sciences*, vol. 15, no. 2, pp. 7–17, 2017.

[11] A. Nawab, Y. Nawab, S. Tang, J. Wu, W. Liu, G. Li, M. Xiao, L. An *et al.*, "A pictorial guidebook on poultry diseases; diagnostic techniques and their effective treatment," *Animal Review*, vol. 5, no. 2, pp. 34–50, 2018.

[12] K. K. Johnson, R. M. Seeger, and T. L. Marsh, "Local economies and highly pathogenic avian influenza," *Choices*, vol. 31, no. 2, pp. 1–9, 2016.

[13] F. Chuma, "Modeling the dynamics, control and economic loss of newcastle disease in village chicken: a case of pwani region in tanzania (doctoral dissertation)," 2019.

[14] H. H. Luong, L. T. T. Le, H. T. Nguyen, V. Q. Hua, K. V. Nguyen, T. N. P. Bach, T. N. A. Nguyen, and H. T. Q. Nguyen, "Transfer learning with fine-tuning on mobilenet and grad-cam for bones abnormalities diagnosis," *Complex, Intelligent and Software Intensive Systems*, pp. 171–179, 2022.

[15] H. H. Luong, N. T. L. Phan, T. C. Dinh, T. M. Dang, T. T. Duong, T. D. Nguyen, and H. T. Nguyen, "Fine-tuning mobilenet for breast cancer diagnosis," *Inventive Computation and Information Technologies*, pp. 841–856, 2023.

[16] H. T. Nguyen, H. H. Luong, P. T. Phan, H. H. D. Nguyen, D. Ly, D. M. Phan, and T. T. Do, "Hs-unet-id: An approach for human skin classification integrating between unet and improved dense convolutional network," *International Journal of Imaging Systems and Technology*, vol. 32, no. 6, pp. 1832–1845, Jun. 2022.

[17] L. H. Huong, N. H. Khang, L. N. Quynh, L. H. Thang, D. M. Canh, and H. P. Sang, "A proposed approach for monkeypox classification," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 8, 2023.

[18] H. T. Nguyen, H. H. Luong, T. H. N. Kien, N. T. L. Phan, T. M. Dang, T. T. Duong, T. D. Nguyen, and T. C. Dinh, "Brain tumors detection on mri images with k-means clustering and residual networks," *Advances in Computational Collective Intelligence*, pp. 317–329, 2022.

[19] H. Wu, B. Xiao, N. Codella, M. Liu, X. Dai, L. Yuan, and L. Zhang, "Cvt: Introducing convolutions to vision transformers," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 22–31.

[20] H. Xu, Q. Xu, F. Cong, J. Kang, C. Han, Z. Liu, A. Madabhushi, and C. Lu, "Vision transformers for computational histopathology," *IEEE Reviews in Biomedical Engineering*, 2023.

[21] A. Hatamizadeh, G. Heinrich, H. Yin, A. Tao, J. M. Alvarez, J. Kautz, and P. Molchanov, "Fastervit: Fast vision transformers with hierarchical attention," *arXiv preprint arXiv:2306.06189*, 2023.

[22] M. Sundararajan, A. Taly, and Q. Yan, "Axiomatic attribution for deep networks," in *International conference on machine learning*. PMLR, 2017, pp. 3319–3328.

[23] M. Sadeghi, A. Banakar, S. Minaei, M. Orooji, A. Shoushtari, and G. Li, "Early detection of avian diseases based on thermography and artificial intelligence," *Animals*, vol. 13, no. 14, p. 2348, 2023.

[24] S. Suthagar, G. Mageshkumar, M. Ayyadurai, C. Snegha, M. Sureka, and S. Velmurugan, "Faecal image-based chicken disease classification using deep learning techniques," in *Inventive Computation and Information Technologies: Proceedings of ICICIT 2022*. Springer, 2023, pp. 903–917.

[25] K. H. Nguyen, H. V. N. Nguyen, H. N. Tran, and Q. Luyl-Da, "Combining autoencoder and yolov6 model for classification and disease detection in chickens," in *Proceedings of the 2023 8th International Conference on Intelligent Information Technology*, 2023, pp. 132–138.

[26] M. K. Gourisaria, A. Arora, S. Bilgaiyan, and M. Sahni, "Chicken disease multiclass classification using deep learning," in *Proceedings of International Conference on Information Technology and Applications: ICITA 2022*. Springer, 2023, pp. 225–238.

[27] "Image - based poultry disease detection using deep convolutional neural network," 2023.

[28] J. M. P. Sánchez, L. A. O. Moreno, J. L. R. Rodríguez, and I. D. M. Corro, "Poultry egg classification system using deep learning," in *2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*. IEEE, 2023, pp. 01–06.

[29] M. Kholil, H. P. Waspada, and R. Akhsani, "Classification of infectious diseases in chickens based on feces images using deep learning," in *2022 International Conference on Electrical and Information Technology (IEIT)*. IEEE, 2022, pp. 362–365.

[30] H. T. Nguyen, T. D. Tran, T. T. Nguyen, N. M. Pham, P. H. N. Ly, and H. H. Luong, "Strawberry disease identification with vision transformer-based models," *Multimedia Tools and Applications*, Feb. 2024.

[31] M. Z. Degu and G. L. Simegn, "Smartphone based detection and classification of poultry diseases from chicken fecal images using deep learning techniques," *Smart Agricultural Technology*, vol. 4, p. 100221, 2023.

[32] J. A. Wani and N. Sharma, "Comparative analysis of transfer learning models in classification of histopathological whole slide images," in *Proceedings of International Conference on Recent Innovations in Computing: ICRIC 2022, Volume 1*. Springer, 2023, pp. 351–369.

[33] M. Schwegler, C. Müller, and A. Reiterer, "Integrated gradients for feature assessment in point cloud-based data sets," *Algorithms*, vol. 16, no. 7, p. 316, 2023.

[34] D. Lundstrom and M. Razaviyayn, "Four axiomatic characterizations of the integrated gradients attribution method," *arXiv preprint arXiv:2306.13753*, 2023.

[35] D. Machuve, E. Nwankwo, N. Mduma, and J. Mbelwa, "Poultry diseases diagnostics models using deep learning," *Frontiers in Artificial Intelligence*, vol. 5, 08 2022.