International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Hybrid VGG16-Abstract Neural Network Model based Bird Detection and Classification of Images

Ramya Krishna Dhulipalla[1], Sibi Chakkaravarthy Sethuraman*[1]

Submitted: 05/05/2023 **Revised**: 16/07/2023 **Accepted**: 07/08/2023

Abstract: The classification and detection of bird images has become crucial for ecological monitoring and recognition in recent years. Classifying and detecting birds in images is a challenging task due to the presence of both high and low interclass variations. To overcome these challenges, a hybrid approach is proposed in this research that combines the VGG16 architecture with Abstract Neural Network (ANN). This approach aims to overcome the challenges that arise from both high and low interclass variations and improved the accuracy. Then, BiRD and CUB 200 datasets were used to classify birds using the suggested method. Data normalization is implemented to enhance data integrity and efficiency during the database design process, which involved restructuring the data to eliminate redundancy and enhance the overall consistency of the dataset. This research utilized bounding boxes to improve detection accuracy and provide detailed spatial information about birds, leading to more precise identification. Additionally, a hybrid VGG16-ANN model is developed to address the difficulty of detecting abnormal bird activity in real- time. From the results, it clearly shows that the hybrid VGG16-ANN achieved an accuracy of 98.50% in CUB200 dataset, which is comparatively higher than the existing methods such as Compound Model Scaling with Efficient Attention (CMSEA), YOLO-V5 and Hierarchical feature fusion value of 95.37%, 97.63% and 87.02% respectively.

Keywords: *Abstract Neural Network, BiRD, CUB 200, Data Augmentation, Normalization, Hybrid VGG16- ANN=*

1. Introduction

Monitoring birds' species is an important part of ecosystem conservation and management. It helps us understand the health of the ecosystem and the impact of human activities on wildlife populations [1,2]. Monitoring is also essential to determine the activities that are affecting attempts to conserve bird species. Bird identification and recognition in aerial images serve a crucial and major part in varied uses such as monitoring bird populations, tracking birds, and behavior of birds, etc [3]. This is due to the increasing availability of Unmanned Aerial Vehicles (UAVs). High-resolution automated identification of birds through images has become increasingly important in digital image processing [4]. However, the challenge in wildlife monitoring using aerial images can capture precise information quickly and cover large areas. The analysis of data can be time-consuming and resource intensive [5,6]. As a result, researchers have been trying to build automated bird monitoring approaches to increase the classifiers' accuracy and dependability. To overcome these challenges, researchers are exploring different techniques for improving the performance of bird recognition algorithms in real–world scenarios. The suggested approach is to develop environmental variations by training a diverse range of images in different lighting and weather conditions. [7].

Some standard metrics are normally utilized to evaluate the characteristics of bird recognition classifiers. These metrics evaluates a quantitative analysis of classifiers performance in terms of their ability to correctly identify birds. To solve this issue, commercially available sensors such as radar, acoustic, and passive Radio Frequency (RF) sensors have been investigated. Radar can operate at night and in bad weather, and it does not require any signal emission from the target. As a result, a radar sensor is a strong contender for any drone detection and classification system [8,9]. It has advantages such as long-range detection, tracking, and all-weather operational capabilities. Drones are often constructed of lightweight materials such as plastic or carbon fiber, which can make them difficult to detect by radar systems. [10]. As a result, bird's species have been examined using a variety of counting methods, including total, line-transect count, ground count, aerial count, and dropping count. These approaches are based on using human opinion to directly count the birds in particular regions, and then using this knowledge to calculate the population number in a whole area [11,12]. Finally, the suggested method has several advantages including rapidly analysing large numbers of images and the potential for continuous learning. However, it also has limitations including the need large, diverse training dataset and potential errors. This process allows the efficient correlation of fine-grained object components as well as autonomous bird recognition from collected images, and it can provide considerable, important information about bird species [13,14]. As a result, the above-mentioned practical issues highlight the need of recognizing bird species. It focuses on the automatic detection of bird species using machine learning techniques [15] The contribution of the study are as

[1] *Department of Computer Science Engineering, VIT-AP University, India*
** Corresponding Author Email: Krishna.19phd7022@vitap.ac.in*

follows,

- In this research, a hybrid VGG16- Abstract Neural Network (ANN) is suggested to improve the efficiency of the bird's image classification and detection.
- The CUB 200 and BiRD datasets were used to classify the birds using the specified procedures.
- Normalization is utilized in this research to eliminate data redundancy and increased data integrity.

The rest of this paper is organized as follows. Section II illustrated the related works. Section III illustrated the suggested VGG16-ANN, Section IV illustrated the experimental results Section V illustrated about conclusion.

2. Literature Survey

In this research, literature survey was utilized to examine the problems related to predicting bird species in the existing methods. The survey involved gathering and analysing information from various published works to identify the challenges, and areas for improvement in the field of bird's prediction.

Xiang et al. [16] implemented a Compound Model Scaling with Efficient Attention (CMSEA) for fine-grained image classification. CMSEA balances each dimension of network width, depth, and image resolution in model scaling, which replaces the two fully connected layers with a fast one-dimensional features. Hence, CMSEA avoids dimensionality reduction and effectively describes the discriminative features. However, CMSEA requires attention mechanism modules to improve the model performance.

Qiu et al. [17] developed YOLOv4-tiny algorithm to detect transmission line faults for bird species. The YOLOv4-tiny algorithm was created and processed for the training stage, cosine annealing, mosaic data augmentation, and smoothing label. YOLOv4-tiny was developed to study the features of bird photographs by integrating ML and DL algorithms. As a result, the suggested techniques achieved better results for the precision of 92.04% respectively. However, because various factors impact model detection performance during the process, YOLOv4 model parameters must be adjusted.

Hilal et al. [18] designed fine grained YOLO v5 to solve the problems in bird's image recognition. When the model detects the detection frame of the object in the predicted image, multiple detection frames were generated for the same object, and each detection frame has a corresponding score. The suggested method used as the detector in the detection part, and the Soft-NMS mechanism was used in the detection part to enhance the detection results. However, the suggested method need to explore the application potential methods in other fields, such as motion estimation, modelling optimization and temporal prediction.

Wang et al. [19] suggested a hierarchical feature fusion module and a counterfactual feature enhancement module to efficiently guide the model in learning the fine-grained features of bird. The model comprises a feature extraction network, a hierarchical feature fusion module, and a counterfactual feature selection module. The feature extractor extracts local area image features and global image features. The suggested counterfactual feature enhancement module obtains effective attention weights by performing a counterfactual intervention on the fused attention weight. However, in the suggested method some feature information was lost during classification of image.

Lin et al. [20] presented an end-to-end bird classification based on target key points to recognize the position of birds. The components production module filters the bird mechanisms and produces the feature map for bird position identification based on the bird's important points using a shallow network. In both network depths, the suggested model has the highest recognition accuracy of 87.78% in the five positions, such as stand, crouch, feed, and flap wings. However, the proposed techniques do not have the best performance in classification for swim and fly.

The proposed method aims to address the limitations found in existing bird's classification approaches. By enhancing data quality and eliminating inconsistencies, the suggested approach aims to significantly improve the accuracy and reliability of classification models.

3. Proposed Method

This research proposed Birds classification and detection which contains the following stages such as Datasets, Pre-processing, Data augmentation, classification and performance. This section gives a general overview of bird classification and detection using various techniques. The suggested method process is illustrated in Fig. 1.

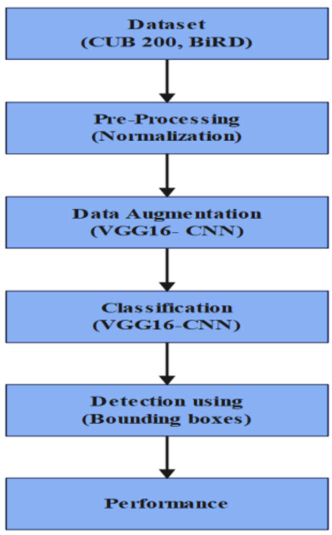

Fig. 1. Flow diagram of the suggested Hybrid VGG-16 ANN techniques

3.1. Data collection

In this research, the raw data is obtained from two publicly available datasets such as CUB 200 [16] and BiRD [17] dataset. The description of the mentioned dataset is explained as follows;

- **CUB 200 dataset**

The CUB-200 dataset is utilized for fine- grained bird species recognition, offering a real world challenge with diverse images. The CUB 200 datasets provided the necessary information and samples to train and test algorithms, facilitating accurate classification and detection of birds.

- **BiRD dataset**

The BiRD dataset helps this research, by providing labelled data. The BiRD is a collection of 2751 images and 50 small-sized flying bird species recorded under several meteorological circumstances.

3.2. Data Preprocessing

After data acquisition, the pre-processing stage is conducted to transform the data into a processed format. The process involved removing noisy and unwanted portions from the images using normalization techniques

- **Normalization**

Normalization is used to standardize data ranges, preventing skewed influence of features, enhancing model training, convergence, and performance in machine learning algorithms. Normalization is a data pre-processing technique used in statistics and machine learning to scale numeric values within a consistent range. It ensures that data features have similar magnitudes, preventing one feature from dominating others. This helps to algorithms in making accurate predictions or classifications by improving convergence and efficiency during training. The mathematical expression of the normalization is shown in (1).

$$Xnormalized = \frac{(X-X_{minimum})}{range\ of\ x} \tag{1}$$

Where, X is the original value of data point, X_{min} is the minimum value of the feature and X_{max} is the maximum value of the feature.

3.3. Data Augmentation

The pre-processed data is given to the data augmentation to give efficient data and information while classifying the image. In this research data augmentation is utilized by VGG16-CNN is briefly described below;

- **VGG16 –CNN**

Data augmentation techniques were utilized to analyse the data and improve the quantity of information. It functions is regularized and helps in the reduction of overfitting problems while training machine learning algorithms. In the field of data augmentation, VGG16 refers to the VGGNet model. It is a 16-layer CNN model that has been pre-trained utilizing over a million photographs from the dataset. The suggested VGG16- CNN has twelve fire modules for the convolution layer to decide the number of parameters include in the fire module.

It multiplies the number of input channels by the filter counts to generate a CNN model with a low parameter count. At the VGG16-CNN stage, the training data will be passed via a convolution layer, a pooled layer, a flattening layer, and a dense layer. CNN models have the potential to improve feature learning techniques in visualization systems. VGG16 extracts features from raw remote sensing data, and the classification stage utilizes CNN output feature maps to solve the classification problem. Transferring the well-learned parameters of VGG16, which is trained on the extensively annotated dataset, it can greatly enhance the efficiency of image recognition tasks. It is challenging to get and annotate large volumes of samples. As a result, the VGG16 architecture contains millions of variables for training and necessitates a large amount of labelled data. Fig. 2. showed the sample input and output images for data augmentation.

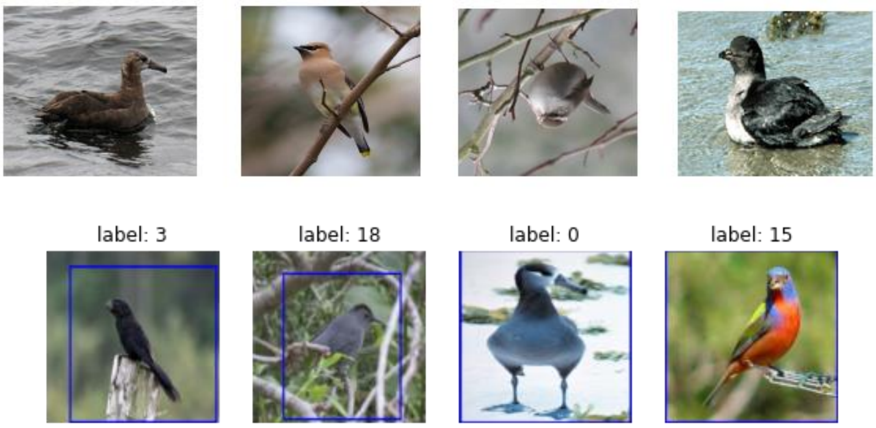

Fig. 2. (a) Sample input image for data augmentation (b) Sample output image for data Augmentation

3.4. Classification using VGG16-CNN

After data augmentation, the VGG16-CNN is utilized to the classification of birds in the images. VGG16-CNN is employed for image classification due to its deep architecture, which captures intricate features, enabling accurate recognition in tasks like object classification and detection. Each VGG16-CNN model functions as a mathematical event that stores and categorizes data information particular to the architecture. It features adaptive parallel processing capacity for many refined images such as hyperspectral images and RGB images. This is an important stage since these attributes help classifiers to enhance accuracy during the classification step. VGG16-CNN replaced its fully linked layers with customized multilayers for categorization. Between each MLP layer, the VGG16- CNN layers were initialized with weights. Fig. 3 illustrated the structure of VGG16 method.

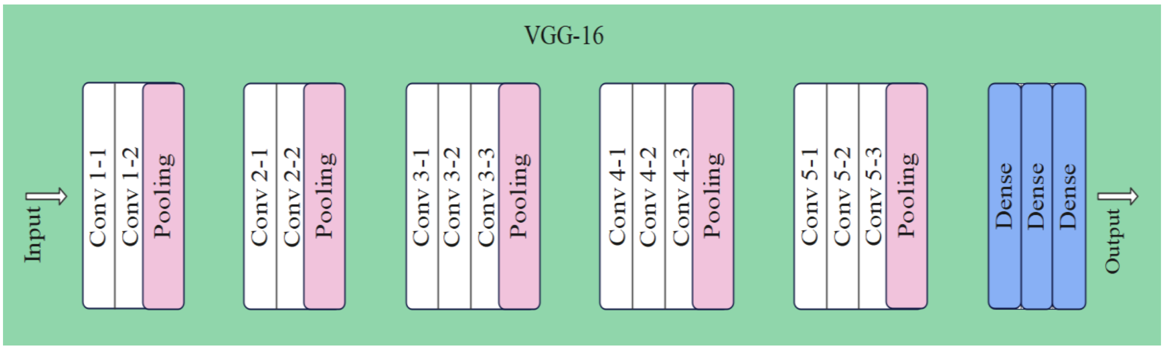

Fig. 3. The structure of VGG16 method

3.5. Birds detection using Bounding Boxes

Bounding boxes play a pivotal role in bird detection by providing accurate localization information. They enable precise identification and tracking of birds in images or videos, helping in ecological research, wildlife conservation, and behavioural studies. Bounding boxes facilitate counting bird populations, assessing migration patterns, and monitoring nesting behaviours. Moreover, bounding boxes are crucial for training and evaluating object detection models, enhancing their ability to separate birds from backgrounds and other objects, contributing to the development of advanced AI-powered bird detection systems. The bounding box is rectangular, and it is defined by the x and y synchronizes the rectangle's lower-right and upper-left corners. Another popular bounding box format coordinates the box's center, as well height and width. Fig. 4. showed the sample images for bird's detection.

Fig. 4. Sample images for bird detection

3.6. Proposed Hybrid VGG16-ANN

The classification and detection of VGG16-ANN are given to the Proposed VGG16-ANN to identify the birds in the fields. It is used to decrease limitations and prevent the device from being used for real-time video surveillance. To address the challenge of real-time identification of abnormal activities of birds, a hybrid VGG16-ANN is proposed. The VGG16-ANN might have processed images by constantly estimating the goal function and abstracting relevant feature representations. The combination of VGG16-ANN with an optimal weight decay optimizer is employed to enhance model generalization and mitigate overfitting, resulting in improved performance and robustness in identifying abnormal activities of words.

3.6.1. Optimal Weight Decay Optimizer

The proposed VGG16-ANN is given to the optimal weight decay optimizer that helps to reduce overfitting by penalizing large weights. Weight decay is a widely used regularization technique essential for training deep neural networks to achieve better generalization. It is commonly applied as L2 regularization, which can be assumed as imposing a Gaussian prior on the model weights. This relationship is exemplified by (2).

$$\theta_t = (1 - \lambda_0)\theta_{t-1} - \eta gt \tag{2}$$

Where λ_0 is the weight decay hyper parameter, θ_t is the model parameters step , η is the learning rate, and gt is the grade of minibatch loss function $L(\theta)$ at θ_{t-1}. Then, computation of empirical gradient incorporates a process of noise that observes to a specific distribution with a mean of zero. The variance of this noise process is linked to the desired level of regularization to be applied. The objective function takes into account the regularization term γ(w) in (3).

$$\mathcal{L}\,(w) = \rho(w) + \frac{\lambda}{2}||\,w\,||_2^2 \quad (3)$$

The optimization algorithm for first order illustrates the better gradient step at each iteration t, considering the coefficient λ ∈ R for the regularization term. Equation (4) described this process.

$$w^{t+1} = w^t - \eta^t(\nabla\,\rho\,(w^t) + \lambda w^t$$

$$= (1 - \eta^t\lambda)\,w^t - \eta^t\,\nabla\,\rho\,(w^t) \quad (4)$$

where $\eta^t\lambda$ is a weight-decay regularization scheme that imposes constraints on the values of unknown model parameters, denoted as w during each iteration. These constraints limit the values of w to the range of (0, 1), resulting in a shrinkage effect on the parameters. Typically, weight-decay regularization utilizes a static coefficient λ ∈ R to control the regularization strength. However, in this proposed regularization scheme, adaptive regularity is imposed on each model parameter as shown in (4).

w_j with an extra term θ_J as follows in (5).

$$\mathcal{L}\,(w) = \rho(w) + \frac{\lambda}{2}||\theta\,w\,||_2^2 \quad (5)$$

The regularization degree for various model parameter w_j^t at iteration t is illustrated by the adaptive value θ_j^t , in the following modified (6).

$$w_j^{t+1} = (1 - \eta^t\lambda\theta_j^t\,)\,w_j^t - \eta^t g_j^t \quad (6)$$

The weight-decay regularization scheme involves the gradient g_j^t, representing the data fidelity sensitivity to the model parameter w_j^t during iteration t. In this scheme, the weight-decay coefficient λ (λ ≥ 0) sets a constant level of regularity for each method parameters. Conversely, the adaptive term θ described the relative significance of each method parameter at various optimization steps. Equation (6) with $\theta_j^t = 1$ for all j and t converts as constant weight-decay. Hence, this research proposed Birds classification and detection which contains the following stages such as Datasets, Pre-processing, Data augmentation, classification and performance that are described. This section gives a general overview of bird classification using various techniques. The experimental results for the suggested techniques were compared with various methods described below.

4. Result and Discussion

The proposed approach for distinguishing images of birds from images of non-birds. The inference engine's hardware and software specifications are summarized. Preliminary testing of 100 bird shots is conducted to manually filter non-bird images supplied to the system and to analyse the efficacy of the suggested strategy. The model correctly identified the photographs as authentic bird shots.

4.1. Performance Metrics

In order to improve the qualities of the proposed approach for bird species the following metrics were employed during the testing phase. The for evaluation metrics is shown in table 1.

It displays the results of the bird detection. Fig. 5, Fig. 6, and Fig. 7 described the input and output pictures of the example images.

The multiscale slide window approach is utilized to generate images with or without birds, with the retrieved sub-window defining the target item. The hardware/software requirements of the suggested techniques are illustrated in Table 2.

Table 1. Evaluation Metrics Formula

Performance Metrics	Equations
Sensitivity	$\frac{TP}{TP + FN}$
Accuracy	$\frac{TP + TN}{TP + FP + TN + FN}$
Specificity	$\frac{TN}{TN + FP}$

Table 2. Hardware and software requirements

Hardware/ Software	Specification
GPU	12 Intel Xeon CPUs, 32GB memory
NDK	Helps the bridge SDK and TF platforms
Tensor Flow	Executive numerical computation using data flow graphs
Phycharm	Python ID programmer interface

Fig. 5. Input Image

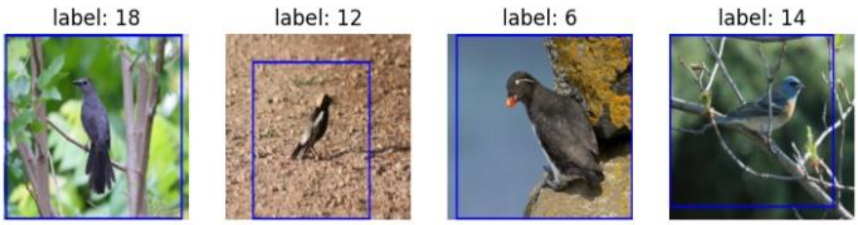

Fig. 6. Sample output Image 1

Fig. 7. Sample output image 2

K-fold cross-validation were utilized to estimate the performance of an ML method by dividing the dataset and training the model K times and K equally sized subsets. It helps mitigate overfitting problems and provides robust estimate of the model characteristics by averaging the results from multiple train-test splits. This technique is particularly useful when the dataset is limited or imbalanced. the performance of the 3- fold validation is shown in Table 3.

Table 3. Performance comparisons of α values of training and testing for 3-fold

3-Fold	Alpha = 0.3		Alpha = 0.5		Alpha = 0.7		Alpha= 0.9	
	Train	Test	Train	Test	Train	Test	Train	Test
1	100	90.99	100	94.69	100	94.17	100	95.32
2	100	93.31	100	94.83	100	95.54	100	97.27
3	100	93.51	100	95.31	100	96.10	100	97.71
4	100	93.70	100	95.49	100	95.82	100	97.00
5	100	93.90	100	95.29	100	95.97	100	97.62
6	100	94.19	100	95.67	100	96.39	100	98.24
7	100	94.48	100	95.82	100	96.12	100	97.53
8	100	94.77	100	95.88	100	95.96	100	96.97
9	100	95.55	100	96.83	100	96.91	100	98.71
10	100	96.61	100	97.76	100	97.89	100	98.92
Proposed	100	94.87	100	95.76	100	98.00	100	97.52

In Table 3, for alpha 0.3 the suggested method achieved 100% of accuracy in the training and 94.87% testing set. Similarly, when using an alpha 0.5, the method achieved 100% accuracy in training and 95.76% testing. Furthermore, with an alpha value of 0.7, the training remained at 100% accuracy, while the testing increased to 98.00%. Finally, when alpha is set to 0.9, the model achieved 100% accuracy in training and 97.52% accuracy testing set. Table 3, it demonstrates that the suggested method achieved better accuracy for training and testing when compared to the existing methodology.

Table 4. Performance comparisons of α values of Training and Testing for 5-Fold

5-Fold	Alpha = 0.3		Alpha = 0.5		Alpha = 0.7		Alpha= 0.9	
	Train	Test	Train	Test	Train	Test	Train	Test
1	100	95.02	100	96.14	100	95.61	100	95.80
2	100	95.72	100	96.29	100	97.00	100	97.76
3	100	95.92	100	96.76	100	97.57	100	98.21
4	100	96.12	100	96.95	100	97.29	100	97.49
5	100	96.32	100	96.74	100	97.44	100	98.12
6	100	96.61	100	97.13	100	97.87	100	98.73
7	100	96.91	100	97.29	100	97.59	100	98.03
8	100	97.21	100	97.35	100	97.43	100	97.46
9	100	98.01	100	98.32	100	98.40	100	99.21
10	100	99.10	100	99.25	100	99.39	100	99.42
Proposed	100	96.69	100	97.22	100	99.50	100	98.02

In Table 4, for alpha 0.3 the suggested method achieved 100% of accuracy in training and 96.69% in testing set. Similarly, when using an alpha 0.5, the method achieved 100% accuracy in training and 97.22% testing. Furthermore, with an alpha value of 0.7, the training remained at 100% accuracy, while the testing increased to 99.50%. Finally, when alpha is set to 0.9, the model achieved 100% accuracy in the training and 98.02% accuracy in the testing set. In table 4, it demonstrates that the suggested method achieved better accuracy for testing and training when compared to the existing methodology.

Table 5. Performance comparisons of α values of Training and Testing for 7-Fold

7-Fold	Alpha = 0.3		Alpha = 0.5		Alpha = 0.7		Alpha= 0.9	
	Train	Test	Train	Test	Train	Test	Train	Test
1	100	95.21	100	96.33	100	95.80	100	95.99
2	100	95.91	100	96.48	100	97.20	100	97.96
3	100	96.11	100	96.96	100	97.77	100	98.40
4	100	96.31	100	97.15	100	97.49	100	97.69
5	100	96.51	100	96.94	100	97.64	100	98.31
6	100	96.81	100	97.33	100	98.06	100	98.93
7	100	97.11	100	97.49	100	97.79	100	98.22
8	100	97.41	100	97.55	100	97.63	100	97.66
9	100	98.20	100	98.51	100	98.59	100	99.41
10	100	99.30	100	99.45	100	99.59	100	99.62
Proposed	100	96.89	100	97.42	100	99.70	100	98.21

In Table 5, the results illustrate the performance of the suggested method utilizing a 7-fold validation approach with different alpha values. The training set achieved a perfect accuracy of 100% across all alpha values, while the corresponding accuracies in the testing set varied. Specifically, for an alpha value of 0.3, the testing accuracy is 96.89%. Increasing the alpha value to 0.5 resulted in a slightly improved testing accuracy of 97.42%. Notably, with an alpha value of 0.7, the testing accuracy further increased to 99.70%. Finally, when the alpha value is set to 0.9, the model attained a testing accuracy of 98.21%. respectively.

Table 6. Characteristics of α values in Training and Testing for 10-Fold

10-Fold	Alpha = 0.3		Alpha = 0.5		Alpha = 0.7		Alpha= 0.9	
	Train	Test	Train	Test	Train	Test	Train	Test
1	100	95.5	100	96.62	100	96.09	100	96.28
2	100	96.2	100	96.77	100	97.49	100	98.25
3	100	96.4	100	97.25	100	98.06	100	98.70
4	100	96.6	100	97.44	100	97.78	100	97.98
5	100	96.8	100	97.23	100	97.93	100	98.61
6	100	97.1	100	97.62	100	98.36	100	99.23
7	100	97.4	100	97.78	100	98.08	100	98.52
8	100	97.7	100	97.84	100	97.92	100	97.95
9	100	98.5	100	98.81	100	98.89	100	99.71
10	100	99.6	100	99.75	100	99.89	100	99.92
Proposed	100	97.18	100	97.71	100	100	100	98.51

Table 6, demonstrate that the suggested method, employing a 10-fold validation approach with an alpha value of 0.3, achieved 100% accuracy in training and 97.18% in the testing set. Similarly, when using an alpha value of 0.5, the method achieved 100% accuracy in training and 97.71% testing. Furthermore, with an alpha value of 0.7, the training remained at 100% accuracy, while the testing increased to 100%. Finally, when alpha is set to 0.9, the model achieved 100% accuracy in the training and 98.51% accuracy in the testing set respectively.

Table 7. Comparison of K-fold values for the suggested method in terms of parameters

K Folds	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
3	97.95	98.00	97.00	97.00
5	98.34	98.00	98.00	98.00
7	96.89	97.42	99.70	98.21
10	98.58	99.00	99.00	99.00

Table 7 presents a comparison of different fold validation approaches (3-fold, 5-fold, 7-fold, and 10-fold) for evaluating various parameters such as Precision, Accuracy, F1-Score and Recall. The outcomes demonstrate that the suggested 10-fold validation outperforms the other fold validation methods. It achieved, precision of 99.00%, accuracy of 98.58%, F1-Score of 99.00% and recall of 99.00% respectively.

4.2. Ablation Study

An ablation study is a scientific investigation technique used to understand the impact of removing or disabling specific components or factors from a system or experiment. It helps assess the contribution or significance of individual elements by comparing the system's performance before and after their removal. This method to identify critical components and understand their influence on overall outcomes. By incorporating migration learning, it observed a remarkable improvement in the accuracy of the student method, increasing it from 35.26% to 64.98% and decreasing issues related to overfitting during training. Under the condition of image augmentation, the accuracy experienced a significant achievement of 72.97%. Furthermore, the utilization of key part identification and image augmentation techniques resulted in a notable increase in the model's accuracy, reaching 78.24%. This improvement demonstrated a significant enhancement of

13.26% compared to the previous performance. These findings strongly support the conclusion that attention-guided data augmentation greatly enhances the effectiveness of training the fine-grained bird classification model. By incorporating object image prediction, the negative impact of background features is mitigated, and the potential for crucial object details to be overwhelmed by irrelevant information is minimized. Through the implementation of localization-recognition techniques, the model's accuracy experienced a significant boost, reaching an impressive 84.05% with a remarkable improvement of 5.81%. Moreover, the precision rate showed a notable increase, reaching 84.85% with a substantial outcome of 4.82%. Similarly, the recall rate achieved a significant enhancement, reaching 84.16% with a development of 5.75%. Additionally, the F1 score achieved better outcomes of 5.52%, reaching 84.03%. Furthermore, the assistance of the suggested method experienced a significant accuracy improvement, reaching 87.02%, which is a noteworthy enhancement of nearly 3%. It achieved a precision of 88.31%, accuracy of 87.63%, F1 score of 87.74% and recall of 87.78%. Moreover, the proposed VGG16-ANN data augmentation model for bird image classification demonstrated outstanding performance, achieving a maximum accuracy of 98.50% along with perfect precision, recall, and F1 score of 99% is illustrated in Table 8.

Table 8. Experiments result over the companion models

Models [19]	Accuracy (%)	Precision (%)	Recall (%)	Fl Score (%)
ShuffleNetV2	35.26	37.54	35.40	35.87
V2(P)	64.98	65.66	65.19	64.51
V2(P,Os)	72.97	74.45	73.18	73.32
V2(P,Os,L)	76.48	77.40	76.70	76.54
V2(P,.Os,Ps)	78.24	80.03	78.41	78.51
V2(P,Os,Ps,L)	84.05	84.85	84.16	84.03
V2(P,0s,Ps.L,D)	87.02	87.61	87.16	87.01
V2(P,0s,Ps,L,D,Ot,Pt)	87.63	88.31	87.78	87.74
Proposed Method	98.50	99.00	99.00	99.00

4.3. Comparative analysis

Comparative analysis is a method of examining and evaluating two or more methods, between them. It involves assessing and contrasting the similarities and differences between data, or phenomena to derive patterns. Through this analysis, it can identify the strengths, weaknesses, opportunities, and threats of the subjects being compared. It provides a basis for making informed decisions and understanding the relative advantages or disadvantages of each option. Table 9 illustrated the comparison of the suggested method with the existing methods.

Table 9. Comparative analysis of the suggested method for CUB-200 dataset

Methods	Accuracy (%)
CMSEA [16]	95.37
YOLO V5 [18]	97.63
Hierarchical fusion [19]	87.02
Proposed Method	98.50

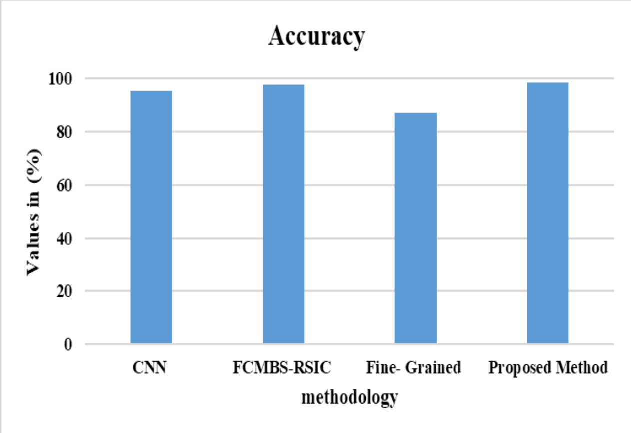

Fig. 8. Graphical representation of the suggested method in terms of accuracy

The proposed hybrid VGG16-ANN utilized CUB-200 dataset to classify the birds. The CUB200-2011 dataset has 200 bird categories, including 5994 training images and 5794 test data, each category contains about 30 training data. According to the result from the table 9 and Figure 8 shows that the proposed classification approach achieved better results in the performance metrics. For example, the accuracy is considered as a common performance metric to evaluate the efficacy of the suggested approach. The accuracy of the proposed approach is 98.50% which is comparatively higher than the existing Compound Model Scaling with Efficient Attention (CMSEA) [16], YOLO-V5 [18] and Hierarchical feature fusion [19] value of 95.37%, 97.63% and 87.02% respectively.

5. Conclusion

In this research, a hybrid VGG16- Abstract Neural Network (ANN) is designed to improve the accuracy of the bird's image classification and detection. The BiRD and CUB 200 datasets were used to classify birds using the suggested method. Data normalization is implemented to enhance data integrity and efficiency during the database design process, which involved restructuring the data to eliminate redundancy and enhance the overall consistency of the dataset. This research utilized bounding boxes to improve detection accuracy and provide detailed spatial information about birds, leading to more precise identification. From the results, it clearly shows that the hybrid VGG16-ANN achieved an accuracy of 98.50% in CUB200 dataset, which is comparatively higher than the existing methods such as CMSEA, YOLO-V5 and Hierarchical feature fusion value of 95.37%, 97.63% and 87.02% respectively. In future, it plans to create mechanisms for various bird species with interclass and intra-class variations.

Notation

S. No.	Variables	Definition
1	λ_0	Weight decay of the hyper parameter
2	θ_t	Model parameter
3	η	Learning rate
4	gt	Grade of minibatch loss function
5	$\gamma(w)$	Regularized term
6	$\eta^t\lambda$	Weight decay
7	w	Unknown model parameter
8	w_j^t	The regularization degree for various model parameter
9	w_j^t	Iteration value
10	θ_j^t	Adaptive value
11	g_j^t	Gradient weight-decay regularization
12	$\lambda\,(\lambda \geq 0)$	constant level of regularity for each method parameters
13	TP	True Positive
14	TN	True Negative
15	FP	False Positive
16	FN	False Negative
17	θ	Weight decay in the rescaled coordinates
18	$1-\eta^t\lambda$	Increase the learning rate iteration
19	$1-\lambda$	Rescaling rate of the dynamical system
20	$\gamma\,(w)$	Regularization
21	j	Constant weight decay optimization

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

For this research work all authors' have equally contributed in Conceptualization, methodology, validation, resources, writing—original draft preparation, writing—review and editing.

References

[1] C. Li, B. Zhang, H. Hu, and J. Dai, "Enhanced bird detection from low-resolution aerial image using deep neural networks," *Neural Process. Lett.*, vol. 49, no. 3, pp. 1021–1039, 2019. https://doi.org/10.1007/s11063-018-9871-z

[2] J. Fan, X. Liu, X. Wang, D. Wang, and M. Han, "Multi-Background Island Bird Detection Based on Faster R-CNN," *Cybern. Syst.*, vol. 52, no. 1, pp. 26–35, 2021. https://doi.org/10.1080/01969722.2020.1827799

[3] S. -J. Hong, Y. Han, S. -Y. Kim, A. -Y. Lee, and G. Kim, "Application of Deep-Learning Methods to Bird Detection Using Unmanned Aerial Vehicle Imagery," *Sensors*, vol. 19, no. 7, p. 1651, 2019. https://doi.org/10.3390/s19071651

[4] L. B. Boudaoud, F. Maussang, R. Garello, and A. Chevallier, "Marine Bird Detection Based on Deep Learning using High-Resolution Aerial Images," in OCEANS 2019 - Marseille, Marseille, France, IEEE, 2019, pp. 1–7, doi: 10.1109/OCEANSE.2019.8867242.

[5] L. J. Hernández-González, J. Frausto-Solís, J. J. González-Barbosa, J. P. Sánchez-Hernández, D. L. Hernández-Rabadán, and E. Román-Rangel, "PSEV-BF Methodology for Object Recognition of Birds in Uncontrolled Environments," *Axioms*, vol. 12, no. 2, p. 197, 2023. https://doi.org/10.3390/axioms12020197

[6] K. Yang and Z. Song, "Deep Learning-Based Object Detection Improvement for Fine-Grained Birds," *IEEE Access*, vol. 9, pp. 67901-67915, 2021, doi: 10.1109/ACCESS.2021.3076429

[7] A. Alghamdi, T. Mehtab, R. Iqbal, M. Leeza, N. Islam, M. Hamdi, and A. Shaikh, "Automatic Classification of Monosyllabic and Multisyllabic Birds Using PDHF," *Electronics*, vol. 10, no. 5, p. 624, 2021, doi: 10.3390/electronics10050624.

[8] S. Rahman and D. A. Robertson, "Classification of drones and birds using convolutional neural networks applied to radar micro-Doppler spectrogram images," *IET Radar Sonar Navig.*, vol. 14, no. 5, pp. 653–661, 2020. https://doi.org/10.1049/iet-rsn.2019.0493

[9] B. Ghani and S. Hallerberg, "A Randomized Bag-of-

Birds Approach to Study Robustness of Automated Audio Based Bird Species Classification," *Appl. Sci.*, vol. 11, no. 19, p. 9226, 2021. https://doi.org/10.3390/app11199226

[10]J. Liu, Q. Y. Xu, and W. S. Chen, "Classification of Bird and Drone Targets Based on Motion Characteristics and Random Forest Model Using Surveillance Radar Data," *IEEE Access*, vol. 9, pp. 160135–160144, 2021, doi: 10.1109/ACCESS.2021.3130231

[11]B. R. Devassy and J. K. Antony, "Histopathological image classification using CNN with squeeze and excitation networks based on hybrid squeezing," *SIViP*, vol. 17, no. 7, pp. 3613–3621, 2023. https://doi.org/10.1007/s11760-023-02587-y

[12]M. Surender, K. C. Shekar, K. Ravikanth, and R. Saidulu, "Automatic Identification of Bird Species from the Image Through the Approaches of Segmentation," in Innovations in Computer Science and Engineering, Lecture Notes in Networks and Systems, vol. 74, Singapore: Springer, 2019, pp. 203–214. https://doi.org/10.1007/978-981-13-7082-3_25

[13]J. Xie and M. Zhu, "Acoustic Classification of Bird Species Using an Early Fusion of Deep Features," *Birds*, vol. 4, no. 1, pp. 138–147, 2023. https://doi.org/10.3390/birds4010011

[14]J. A. J. Mortimer, T. C. Greene, and S. P. Mortimer, "Assessing bird vocalisation identification accuracy using a computer-based quiz," *N. Z. J. Zool.*, vol. 46, no. 3, pp. 201–224, 2019. https://doi.org/10.1080/03014223.2018.1536068

[15]Y. -P. Huang and H. Basanta, "Bird Image Retrieval and Recognition Using a Deep Learning Platform," *IEEE Access*, vol. 7, pp. 66980–66989, 2019, doi: 10.1109/ACCESS.2019.2918274

[16]W. Xiang, Z. Song, G. Zhang, and X. Wu, "Birds Detection in Natural Scenes Based on Improved Faster RCNN," *Appl. Sci.*, vol. 12, no. 12, p. 6094, 2022. https://doi.org/10.3390/app12126094

[17]Z. Qiu, X. Zhu, C. Liao, D. Shi, Y. Kuang, Y. Li, and Y. Zhang, "Detection of bird species related to transmission line faults based on lightweight convolutional neural network," *IET Gener. Transm. Distrib.*, vol. 16, no. 5, pp. 869–881, 2022. https://doi.org/10.1049/gtd2.12333

[18]A. M. Hilal, H. Alsolai, F. N. Al-Wesabi, M. K. Nour, A. Motwakel, A. Kumar, I. Yaseen, and A. S. Zamani, "Fuzzy Cognitive Maps with Bird Swarm Intelligence Optimization-Based Remote Sensing Image Classification," *Comput. Intell. Neurosci.*, vol. 2022, p. 4063354, 2022. https://doi.org/10.1155/2022/4063354

[19]K. Wang, F. Yang, Z. Chen, Y. Chen, and Y. Zhang, "A Fine-Grained Bird Classification Method Based on Attention and Decoupled Knowledge Distillation," *Animals*, vol. 13, no. 2, p. 264, 2023. https://doi.org/10.3390/ani13020264

[20]C. -W. Lin, S. Hong, M. Lin, X. Huang, and J. Liu, "Bird posture recognition based on target keypoints estimation in dual-task convolutional neural networks," *Ecol. Indic.*, vol. 135, p. 108506, 2022. https://doi.org/10.1016/j.ecolind.2021.108506

[21]Yoon, H. . (2023). A Quantitative Evaluation for Usability under Software Quality Models. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3), 24–29. https://doi.org/10.17762/ijritcc.v11i3.6194

[22]Anna, G., Jansen, M., Anna, J., Wagner, A., & Fischer, A. Machine Learning Applications for Quality Assurance in Engineering Education. Kuwait Journal of Machine Learning, 1(1). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/109