ASED http://e-ASED.org
Animal Systematics, Evolution and Diversity

Anim. Syst. Evol. Divers. Vol. 40, No. 4: 266-275, October 2024
https://doi.org/10.5635/ASED.2024.40.4.008

Application of Deep Learning-Based Object Detection Models to Classify Images of *Cacatua* Parrot Species

Jung-Il Kim#, Jong-Won Baek#, Chang-Bae Kim*

Department of Biotechnology, Sangmyung University, Seoul 03016, Korea

ABSTRACT

Parrots, especially the *Cacatua* species, are a particular focus for trade because of their mimicry, plumage, and intelligence. Indeed, *Cacatua* species are imported most into Korea. To manage trade in wildlife, it is essential to identify the traded species. This is conventionally achieved by morphological identification by experts, but the increasing volume of trade is overwhelming them. Identification of parrots, particularly *Cacatua* species, is difficult due to their similar features, leading to frequent misidentification. There is thus a need for tools to assist experts in accurately identifying *Cacatua* species *in situ*. Deep learning-based object detection models, such as the You Only Look Once (YOLO) series, have been successfully employed to classify wildlife and can help experts by reducing their workloads. Among these models, YOLO versions 5 and 8 have been widely applied for wildlife classification. The later model normally performs better, but selecting and designing a suitable model remains crucial for custom datasets, such as wildlife. Here, YOLO versions 5 and 8 were employed to classify 13 *Cacatua* species in the image data. Images of these species were collected from eBird, iNaturalist, and Google. The dataset was divided, with 80% used for training and validation and 20% for evaluating model performance. Model performance was measured by mean average precision, with YOLOv5 achieving 0.889 and YOLOv8 achieving 0.919. YOLOv8 was thus better than YOLOv5 at detecting and classifying *Cacatua* species in the examined images. The model developed here could significantly support the management of the global trade in *Cacatua* species.

Keywords: *Cacatua* parrots, CITES, conservation, deep learning, image classification, object detection

INTRODUCTION

Although trade in wildlife is known to directly or indirectly influence the decline in biodiversity, such trade continues to increase worldwide (Scheffers et al., 2019; Hughes, 2021). Various international conventions and conservation bodies, such as the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES), protect wildlife from extinction by managing its global trade. Among avian orders, parrots (Psittaciformes) are the most traded wildlife group worldwide, as live companion pets (Bush et al., 2014; Furnell, 2019; Scheffers et al., 2019). Its features, including mimicking different sounds such as the human voice (Bradbury and Balsby, 2016), colorful and complex plumage (Berg and Bennett, 2010), and high intelligence (Cussen, 2017), make parrots particularly sought after as pets, which has in turn increased the trade in the animals (Aloysius et al., 2020; Chan et al., 2021). Among the parrots, the genus *Cacatua*, which consists of 13 species, was one of the most traded parrot genera globally between 1975 and 2016 (Chan et al., 2021). Indeed, according to the National Institute of Biological Resources (NIBR), parrots are the animal most commonly imported into Korea (Seomun and Kim, 2016). From 2009 to 2014, 50 parrot species were imported into Korea, with *Cacatua* species being the largest group, numbering eight imported species (Seomun and Kim, 2016). Among these eight species, two, *C. moluccensis* and *C. tenuirostris*, were included in CITES Appendix I, indicating the prohibition of their trade, while the other six were included in CITES Appendix II, indicating close monitoring of their trade through permits. In Asia, such as in the Philippines and Hong Kong, the illegal trade in *Cacatua* species is known to be ongoing (Andersson

#These authors contributed equally to this work.
***To whom correspondence should be addressed**
Tel: 82-2-2287-5288, Fax: 82-2-2287-0070
E-mail: evodevo@smu.ac.kr

eISSN 2234-8190

et al., 2021; Brandis et al., 2023).

Identification of traded species is the crucial first step to managing wildlife trade and combating illegal activities (Scheffers et al., 2019). Although the identification of traded wildlife has been pursued by using various methods, the conventional approach usually based on morphological identification by experts remains main procedure (Mahendiran et al., 2018; Trail, 2021). However, continuous rise in the volume of trade in wildlife has overwhelmed experts (Esipova et al., 2021; Hughes, 2021). In particular, identification of parrots including *Cacatua* species can be extremely challenging and these are often incorrectly identified when traded (Parr and Juniper, 2010). *Cacatua* species are mostly characterized by white bodies, mainly white or yellow crest, and white or little yellow ear-coverts (Forshaw, 2010; Parr and Juniper, 2010; Del Hoyo, 2020). The identification of *Cacatua* species *in situ* is challenging due to their similar morphological features between examined species (Parr and Juniper, 2010). Thus, tools in classifying globally traded wildlife, especially parrots, are needed to support experts for accurate and rapid wildlife identification.

Various studies have shown that deep learning-based image classification methods are valuable tools for supporting the identification of wildlife by morphology-based approach (Kim et al., 2022; Baek et al., 2023; Cardoso et al., 2023; Kulkarni and Di Minin, 2023). In addition, the vast number of images collected in recent citizen science programs has made deep learning a more powerful tool for wildlife conservation (McClure et al., 2020; Chowdhury et al., 2023). Citizen scientists can upload images to biodiversity-associated platforms, such as eBird (https://ebird.org) and iNaturalist (https://www.inaturalist.org), by using their mobile devices. These databases have the advantage of bringing together images of various individuals from particular species in a broad range of habitats (McClure et al., 2020). Training deep learning models with these image data can be a valuable way of enhancing model performance (Jang and Lee, 2021). Deep learning using the vast amount of data obtained from citizen science programs could help alleviate the burden placed on experts, allowing them to concentrate on more complex identification of traded wildlife (Spiesman et al., 2021).

Among deep learning-based image classification models, object detection models were developed based on convolutional neural networks, which are basic and standard in the field of image classification. These models can detect and classify objects in images based on the features extracted from the given dataset. In this process, the 'objects' refers to the target species in images, 'detecting objects' means recognizing the presence of objects and determining the location of these within the images, while 'classifying objects' means the model assigning the detected objects to the most morphologically similar species in the dataset. This ability has led to the widespread and successful utilization of object detection models for classifying wildlife (Kim et al., 2022; Baek et al., 2023; Cardoso et al., 2023; Roy et al., 2023). The You Only Look Once (YOLO) series models currently dominate among object detection models. Among these models, YOLO version 5 (YOLOv5) has been found to outperform most object detection models, in terms of both accuracy and speed (Jocher et al., 2022), and has been widely applied for the classification of various organisms (Ahmad et al., 2022; Tannous et al., 2023; Vo et al., 2023). Meanwhile, YOLO version 8 (YOLOv8) was also recently developed (Jocher et al., 2023) and revealed to be a state-of-the-art object detection algorithm (Wang et al., 2023). For standard datasets, such as the COCO dataset consisting of various objects, the more recent model commonly outperforms the previous one (Jocher et al., 2023). However, for custom datasets, especially for wildlife, this is not always the case (Dang et al., 2023; Sheng et al., 2023). There is thus a need to choose the best model for a particular custom dataset, or to design a novel model suitable for current data.

Here, we developed deep learning-based models for classifying *Cacatua* species using images collected from eBird, iNaturalist, and Google. The YOLOv5 and YOLOv8 models, both widely used and relatively advanced, were employed for classifying 13 *Cacatua* species. Model performance was evaluated using the metrics of precision, recall, and mean Average Precision (mAP). The classification results are also presented through a confusion matrix. The outcomes of this study should aid the classification of such species for conservation purposes.

MATERIALS AND METHODS

Fig. 1 presents a flowchart of the entire methodology of this study. Images of 13 *Cacatua* species were collected, and the species in the images were labeled. The labeled images were split into the training, validation, and test sets. Models were trained using the training set and validation set. During the training process, object detection and classification were learned simultaneously. Finally, the performances of the models were evaluated using the test set.

Data collection

Images of 13 *Cacatua* species were collected from both the eBird (https://ebird.org) and the iNaturalist (https://www.inaturalist.org). To minimize imbalance in the data, additional images of species for which fewer than 50 images were retrieved from these two databases were collected from Google (https://www.google.com). The number of images from each database is detailed in Appendix 1. In the eBird database,

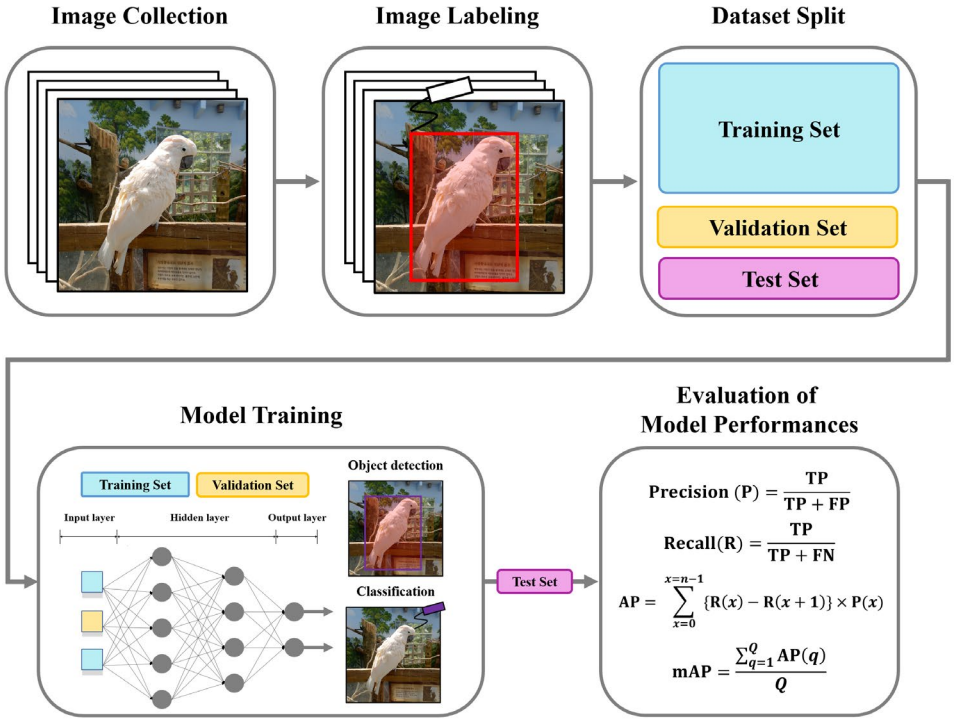

Fig. 1. The flowchart of the overall methodology for classifying 13 *Cacatua* species. TP, true-positive; FP, false-positive; FN, false-negative; AP, Average Precision; mAP, mean Average Precision.

Table 1. The dataset containing 13 *Cacatua* species examined in this study

No.	Species	Total	Training set	Validation set	Test set
1	*Cacatua alba*	114	72	18	24
2	*Cacatua citrinocristata*	107	68	17	22
3	*Cacatua ducorpsii*	95	60	15	20
4	*Cacatua galerita*	5,912	3,783	945	1,184
5	*Cacatua goffiniana*	323	206	51	66
6	*Cacatua haematuropygia*	84	53	13	18
7	*Cacatua leadbeateri*	795	508	127	160
8	*Cacatua moluccensis*	72	46	11	15
9	*Cacatua ophthalmica*	89	56	14	19
10	*Cacatua pastinator*	254	162	40	52
11	*Cacatua sanguinea*	1,643	1,051	262	330
12	*Cacatua sulphurea*	64	40	10	14
13	*Cacatua tenuirostris*	3,896	2,493	623	780

images were collected by the image downloader provided by Google. Meanwhile, images of research grade were collected from the iNaturalist using Inat_images (Huerta-Ramos and Luštrik, 2021). Moreover, to obtain images from Google, searches were performed using scientific and common names as keywords, and then images of the species were collected using AutoCrawler (Kim, 2004). *Cacatua* species in the images were identified carefully using morphological features specific to each species, with reference to the taxonomic literatures (Forshaw, 2010; Parr and Juniper, 2010; Del Hoyo, 2020). Images for which accurate identification of the species could not be achieved based on morphological features were removed. Image data collection and selection following the process conducted in this study has been successfully applied to construct the dataset for training and evaluating deep learning models (Kim et al., 2022; Baek et al., 2023). Images with

more than 320 × 320 pixels and a resolution of 72 dpi were used in this study. The whole body of the *Cacatua* species was set as objects in the images because the morphological features that can be applied to distinguish between *Cacatua* species (i.e., crest, eye ring, plumage, and base of the tail) are located throughout the body. The ground truth bounding box was labeled using LabelImg (Tzutalin, 2015). The dataset was split randomly into a training set of 8,598 images (64%), a validation set of 2,146 images (16%), and a test set of 2,704 images (20%) (Table 1).

Model training

In this study, the YOLOv5 model (Jocher et al., 2022) and the YOLOv8 model (Jocher et al., 2023) were employed to classify 13 *Cacatua* species. These models are divided into five network structures: n, s, m, l, and x. Among these structures, this study used YOLOv5x and YOLOv8x, which showed the highest classification performance (Jocher et al., 2023). The examined models were trained with 300 maximum epochs at a batch size of 16 and an input image size of 640 × 640. The epoch represents the training period, and the batch size represents the number of training images in each iteration that constitute one epoch. Early stopping functions and data augmentation were performed to prevent overfitting. The model training stopped early at the epoch when mAP did not increase for 10 epochs after setting patience to 10. Two data augmentation methods, namely, Albumentation (Buslaev et al., 2020) and Mosaic Augmentation (Austrheim et al., 2014), were applied in the training process. The Albumentation method included horizontal flip, translation, zoom-in, and augmentation of HSV (Hue, Saturation, and Value) (Buslaev et al., 2020). The experimental platform of these models is based on the Rocky Linux 8 operating system, which uses two Intel Xeon Gold 6326 CPUs, eight 64 GB REG.ECC DDR4 SDRAMs, and Nvidia RTX A5000 Graphics with 24G memory. The experimental program is based on Python 3.11.3, CUDA 12.2, cuDNN 8.9.3, and Pytorch 2.0.1. All experimental programs were executed in a virtual environment using the Anaconda prompt. The required packages for the experimental program were imported from PyPI, SciPy, and Pytorch to ensure the proper functioning and execution of the experiments.

Evaluation of model performance

To evaluate the performance of models, the following metrics were used: precision, recall, and mAP. These were evaluated after the completion of training using the test set. Precision refers to the proportion of true results among predicted results by the model, while recall means the proportion of results correctly predicted by the model among the total true results. mAP is an index reflecting both precision and recall. The precision and recall were calculated using formulas (1) and (2), respectively. The true-positive (TP) and false-positive (FP) were defined by using Intersection over Union (IoU). Model predictions were considered TP or FP when the IoU value was greater or less than the threshold, respectively. TP was situation in which the prediction of detecting objects and classification by the model examined was the same as that of the true label. In contrast, FP occurred when object detection and/or classification predictions of the model differed from the true label. False-negative (FN) results implied that the model did not predict any result despite the presentation of a true label. In this study, the IoU threshold value was determined to be 0.5, the standard value to evaluate the performance of object detection models (Jocher et al., 2022, 2023). The Average Precision (AP) value was calculated using formula (3), with n representing the number of ground truth objects. It balances both precision and recall and is based on calculation of the area under a precision-recall curve to optimize detection and classification models. mAP was calculated using formula (4), with Q representing the number of queries of the dataset and AP (q) representing the AP of a given query q. The inference time was calculated as the average time required to process each image. Finally, the classification results of the models were presented through a confusion matrix.

$$\text{Precision} = \frac{\text{True positive}}{\text{True positive} + \text{False positive}} \quad (1)$$

$$\text{Recall} = \frac{\text{True positive}}{\text{True positive} + \text{False negative}} \quad (2)$$

$$\text{Average Precision (AP)} = \sum_{x=0}^{x=n-1} \{\text{Recall}(x) - \text{Recall}(x+1)\} \times \text{Precision}(x) \quad (3)$$

$$\text{mean Average Precision (mAP)} = \frac{\sum_{q=1}^{Q} \text{Ap}(q)}{\text{Q}} \quad (4)$$

RESULTS

This study compared the performance of YOLOv5 and YOLOv8 models in detecting and classifying 13 *Cacatua* species in 13,448 examined images. Table 2 shows the precision, recall, and mAP values of YOLOv5 and YOLOv8 models used in the study. The AP values were 0.878 for YOLOv5 and 0.919 for YOLOv8. For the YOLOv5 model, the precision values for 13 *Cacatua* species ranged from 0.688 for *C. sulphurea* to 0.975 for *C. galerita*. Similarly, for YOLOv8, the lowest precision value was 0.743 for *C. sulphurea*, while the highest was 0.987 for *C. galerita*. The precision differed most markedly between the models for *C. alba*: 0.700 for YOLOv5

Table 2. Precision, recall, and mean Average Precision (mAP) values for the examined models in 13 *Cacatua* species

Species	Precision		Recall		mAP	
	YOLOv5	YOLOv8	YOLOv5	YOLOv8	YOLOv5	YOLOv8
Cacatua alba	0.700	0.886	0.833	0.780	0.819	0.885
Cacatua citrinocristata	0.956	0.959	0.835	0.893	0.929	0.973
Cacatua ducorpsii	0.901	0.885	0.818	0.773	0.880	0.893
Cacatua galerita	0.975	0.987	0.992	0.983	0.991	0.991
Cacatua goffiniana	0.917	0.907	0.899	0.899	0.949	0.932
Cacatua haematuropygia	0.887	0.961	0.944	0.889	0.981	0.954
Cacatua leadbeateri	0.950	0.957	0.957	0.953	0.970	0.980
Cacatua moluccensis	0.854	0.922	0.730	0.737	0.838	0.867
Cacatua ophthalmica	0.818	0.900	0.565	0.739	0.695	0.894
Cacatua pastinator	0.867	0.917	0.870	0.783	0.925	0.925
Cacatua sanguinea	0.930	0.938	0.947	0.936	0.959	0.968
Cacatua sulphurea	0.688	0.743	0.390	0.588	0.622	0.693
Cacatua tenuirostris	0.970	0.986	0.982	0.977	0.993	0.994
Average	0.878	0.919	0.828	0.841	0.889	0.919

and 0.886 for YOLOv8. The average recall values were 0.828 for YOLOv5 and 0.841 for YOLOv8. In the two models, the recall value of *C. sulphurea* was the lowest with value of 0.390 for YOLOv5 and 0.588 for YOLOv8. In addition, the recall value of *C. galerita* was the highest at 0.992 for YOLOv5 and 0.983 for YOLOv8. The recall value for the *C. sulphurea*, which showed the lowest value in both models, differed most markedly between the two models. The mAP value was 0.889 for the YOLOv5 model, but 0.919 for YOLOv8. For the YOLOv5 model, the AP values ranged from 0.622 for *C. sulphurea* to 0.993 for *C. tenuirostris*. While, the lowest AP was 0.693 for *C. sulphurea*, while the highest was 0.994 for *C. tenuirostris* in the YOLOv8 model. The inference time of YOLOv5 and YOLOv8 models were presented as 8.7 ms and 11.0 ms, respectively.

The classification results of the two models are presented as a confusion matrix (Fig. 2). The average rates at which the models correctly classified the 13 *Cacatua* species were 79.0% for YOLOv5 and 85.0% for YOLOv8. For YOLOv5, the lowest correct classification rate was 41.2% for *C. sulphurea*, while the highest was 98.9% for *C. galerita* (Fig. 2A). This low rate for *C. sulphurea* was due to its misclassification as *C. galerita* (41.2%), *C. sanguinea* (5.9%), and high background FN (11.8%). The misclassification of *C. sulphurea* as *C. galerita* was the most common type of classification error among the results of the YOLOv5 model. This was followed by the misclassifications of *C. ophthalmica* as *C. galerita* and *C. moluccensis* as *C. alba*, at rates of 26.1% and 25.0%, respectively. Meanwhile, for the YOLOv8 model, *C. galerita* was classified most accurately, at a rate of 99.1%, while *C. sulphurea* was classified most inaccurately, at 47.1% (Fig. 2B). *Cacatua sulphurea*, which showed the lowest correct classification rate, was misclassified as *C. galerita* (35.3%), *C. citrinocristata* (5.9%), *C. sanguinea* (5.9%) and background FN (5.9%). Similar to the classification results of YOLOv5, for the YOLOv8 model the most common misclassification was *C. sulphurea* as *C. galerita*. Although the misclassification rates of *C. ophthalmica* as *C. galerita* and *C. moluccensis* as *C. alba* were higher than 20.0% for the YOLOv5 model, these rates were only 4.3% and 6.2%, respectively, for YOLOv8.

DISCUSSION

Over the past few years, object detection models, particularly YOLOv5 and YOLOv8, have been successfully applied to the classification of wildlife (Shetty and Ashwath, 2023; Vo et al., 2023). Despite the classification of traded wildlife is fundamental for conservation, few studies have applied deep learning models to classify traded wildlife (Kim et al., 2022; Baek et al., 2023). Although the genus *Cacatua* in particular has been commonly traded legally and illegally, to the best of our knowledge no studies have employed a deep learning model to classify its species. This study thus applied object detection models to classify 13 *Cacatua* species for the first time, which should aid the management and control of their global trade.

According to the present results in terms of mAP, the YOLOv8 model showed higher values, reflecting better precision and recall, than the YOLOv5 model (Table 2). The higher mAP values indicated improved performance of the model in detecting and classifying the 13 *Cacatua* species. In addition, the accurate classification rates of YOLOv8 were higher

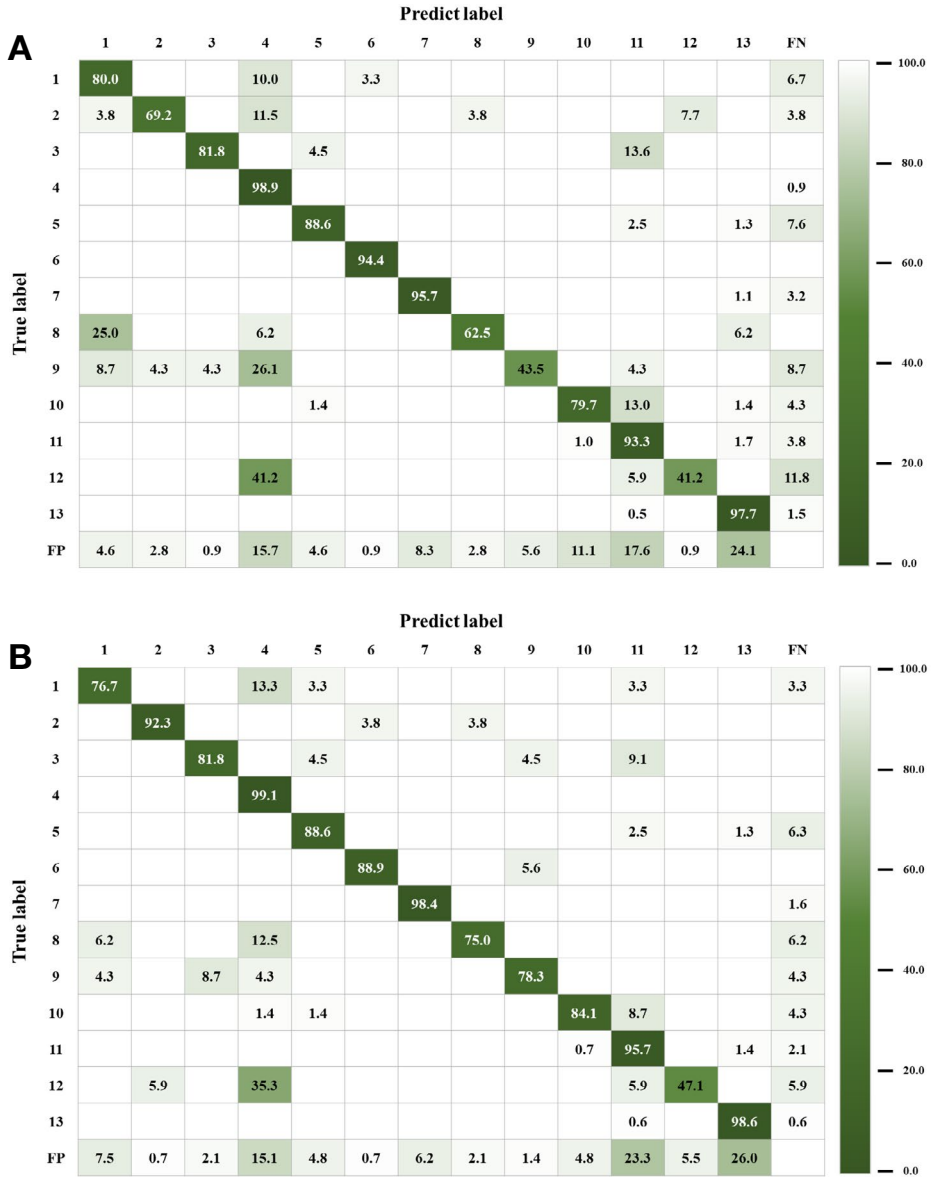

Fig. 2. Confusion matrix of the YOLOv5 and YOLOv8 models to classify 13 *Cacatua* species. A, The confusion matrix of the YOLOv5 model; B, The confusion matrix of the YOLOv8 model. Numbers 1 to 13 indicate the 13 *Cacatua* species classified in this study (see Table 1). FP indicates the background false-positive, while FN points the background false-negative.

than those YOLOv5 due to the lower classification errors in YOLOv8 compared to YOLOv5. In the YOLOv5 model, the misclassification rates of *C. ophthalmica* as *C. galerita* and of *C. moluccensis* as *C. alba* were 26.1% and 25.0%, respectively (Fig. 2A). The YOLOv8 model corrected these rates of misclassification of *C. ophthalmica* as *C. galerita* and *C. moluccensis* as *C. alba* (at 4.3% and 6.2%, respectively), compared with YOLOv5 (Fig. 2B). The improved mAP and classification results of the YOLOv8 model might be due to the difference in model architecture from YOLOv5. The YOLOv5 model utilizes the anchor box method which of initially presents the object predicted by the model in the form of boxes (Jocher et al., 2022). The size of the anchor boxes was determined when the object detection model was developed (Zhong et al., 2020). Although this method has the advantages of high detection accuracy, rapid detection, and low requirements for computational resources (Yan et al., 2021), the model performance can be inferior if the size of anchor boxes is not appropriate for the custom dataset (Zhong et al., 2020). To overcome this challenge, the YOLOv8 model utilizes the anchor-free method, Fully Convolved One-Stage (Jocher et al., 2023). This method initially presents the objects predicted

by the models as a novel center-ness branch and then provides boxes of suitable size for the predicted objects (Tian et al., 2019). This enables more specific learning of the model in the process of detecting objects for the provided dataset (Tian et al., 2019). Indeed, previous studies reported improved performance of YOLOv8 compared to YOLOv5 in their custom datasets due to difference in model architecture (Sary et al., 2023; Shetty and Ashwath, 2023; Paramita et al., 2024).

Meanwhile, in both models, the AP value of *C. sulphurea* was lower than those of the other species. This may have been because relatively few images of this species were used for the training of both models (Spiesman et al., 2021; Kim et al., 2022; Baek et al., 2023). In addition, this species was misclassified as *C. galerita* at a rate higher than any other misclassification for both models, at 41.2% for YOLOv5 and 35.3% for YOLOv8 (Fig. 2). The morphological similarities between *C. sulphurea* and *C. galerita* may explain the lower accurate classification values for *C. sulphurea*. Indeed, it has been reported that it is difficult to distinguish *C. sulphurea* and *C. galerita*, due to its similar morphological features of an erectile yellow crest and predominantly white plumage (Forshaw, 2010; Parr and Juniper, 2010; Del Hoyo, 2020). Insufficient images can also lead to high misclassification between morphologically similar species (Kim et al., 2022; Baek et al., 2023). Therefore, to increase the AP and classification accuracy of *C. sulphurea*, future studies should collect comprehensive images of this species including morphological features that distinguish these from other species in images.

This study showed that insufficient image data contributed to decreased the model performance and increased the classification errors between morphologically similar species. It is similar to previous studies (Kim et al., 2022; Alzubaidi et al., 2023; Baek et al., 2023; Bjerge et al., 2023). This is an issue because the collection of data on rare species, such as endangered ones, is more challenging. Moreover, these particular rare species further exacerbate the issue of data imbalance, which is recognized as a significant factor contributing to decreased model performance (Chan et al., 2023). To further improve model performance, various data augmentation methods will be applied in future studies. In this study, the Albumentation method was used, which included horizontal flip, translation, zoom-in, and HSV augmentation, along with the Mosaic Augmentation method. Additional data augmentation methods, such as vertical flip, RGB shift, and rotation, will be employed in future studies to enhance model performance. Moreover, the generative adversarial network model has been recently successfully applied to generate image data of rare species (Zhang et al., 2023). Therefore, this model should be applied to enhance the dataset of *Cacatua* species, especially *C. sulphurea*, in future work. In addition, the YOLOv9 model was newly developed (Wang et al., 2024) very recently, so this new model should be used to classify *Cacatua* species. Finally, the models developed here will be provided as a mobile application to aid the classification of traded *Cacatua* species *in situ*.

In conclusion, this study applied the YOLOv5 and YOLOv8 models to classify 13 *Cacatua* species in images. According to the model performance as assessed by mAP, the YOLOv8 model outperformed the YOLOv5 model in classifying *Cacatua* species. This is the first study to apply object detection models to classify species of this genus, and the ongoing development of these models may help classify these species rapidly and accurately *in situ*. However, it should be emphasized that recent taxonomic research results should be applied (Spiesman et al., 2021). Indeed, *C. citrinocristata* was considered a subspecies of *C. sulphurea* before the taxonomic study of subspecies of *C. sulphurea* began by Collar and Marsden (2014). However, this was recently reconsidered, and it is now managed as a species in various international conventions and conservation bodies by following the 2023 International Ornithological Congress (IOC) World Bird List v. 13.2 (https://www.worldbirdnames.org). Moreover, to establish an accurate image database based on the morphological features of the examined species, the knowledge of taxonomic experts is needed. In this respect, cooperation with taxonomic experts is essential for the development of deep learning models for the more accurate classification of wildlife.

ORCID

Jung-Il Kim: https://orcid.org/0000-0001-7851-2493
Jong-Won Baek: https://orcid.org/0000-0003-2398-2134
Chang-Bae Kim: http://orcid.org/0000-0002-1040-7600

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

ACKNOWLEDGMENTS

This work was supported by a grant from the National Institute of Biological Resources (NIBR), funded by the Ministry of Environment (MOE) of the Republic of Korea (NIBRE 202411). We used the images from the Macaulay Library at the Cornell Lab of Ornithology during model training, and we thank the thousands of eBird participants and organizations for their contributions.

REFERENCES

Ahmad I, Yang Y, Yue Y, Ye C, Hassan M, Cheng X, Wu Y, Zhang Y, 2022. Deep learning based detector YOLOv5 for identifying insect pests. Applied Sciences, 12:10167. https://doi.org/10.3390/app121910167

Aloysius SLM, Yong DL, Lee JG, Jain A, 2020. Flying into extinction: understanding the role of Singapore's international parrot trade in growing domestic demand. Bird Conservation International, 30:139-155. https://doi.org/10.1017/S0959270919000182

Alzubaidi L, Bai J, Al-Sabaawi A, Santamaría J, Albahri AS, Al-dabbagh BSN, Fadhel MA, Manoufali M, Zhang J, Al-Timemy AH, Duan Y, Abdullah A, Farhan L, Lu Y, Gupta A, Albu F, Abbosh A, Gu Y, 2023. A survey on deep learning tools dealing with data scarcity: definitions, challenges, solutions, tips, and applications. Journal of Big Data, 10:46. https://doi.org/10.1186/s40537-023-00727-2

Andersson AA, Gibson L, Baker DM, Cybulski JD, Wang S, Leung B, Chu LM, Dingle C, 2021. Stable isotope analysis as a tool to detect illegal trade in critically endangered cockatoos. Animal Conservation, 24:1021-1031. https://doi.org/10.1111/acv.12705

Austrheim G, Speed JDM, Martinsen V, Mulder J, Mysterud A, 2014. Experimental effects of herbivore density on aboveground plant biomass in an alpine grassland ecosystem. Arctic, Antarctic, and Alpine Research, 46:535-541. https://doi.org/10.1657/1938-4246-46.3.535

Baek JW, Kim JI, Kim CB, 2023. Deep learning-based image classification of turtles imported into Korea. Scientific Reports, 13:21677. https://doi.org/10.1038/s41598-023-49022-3

Berg ML, Bennett ATD, 2010. The evolution of plumage colouration in parrots: a review. Emu-Austral Ornithology, 110:10-20. https://doi.org/10.1071/MU09076

Bjerge K, Alison J, Dyrmann M, Frigaard CE, Mann HMR, Høye TT, 2023. Accurate detection and identification of insects from camera trap images with deep learning. PLOS Sustainability and Transformation, 2:e0000051. https://doi.org/10.1371/journal.pstr.0000051

Bradbury JW, Balsby TJS, 2016. The functions of vocal learning in parrots. Behavioral Ecology and Sociobiology, 70:293-312. https://doi.org/10.1007/s00265-016-2068-4

Brandis KJ, Meagher P, Schoppe S, Zawada K, Widmann I, Widmann P, Dolorosa RG, Francis R, 2023. Determining the provenance of traded wildlife in the Philippines. Animals, 13:2165. https://doi.org/10.3390/ani13132165

Bush ER, Baker SE, Macdonald DW, 2014. Global trade in exotic pets 2006-2012. Conservation Biology, 28:663-676. https://doi.org/10.1111/cobi.12240

Buslaev A, Iglovikov VI, Khvedchenya E, Parinov A, Druzhinin M, Kalinin AA, 2020. Albumentations: fast and flexible image augmentations. Information, 11:125. https://doi.org/10.3390/info11020125

Cardoso AS, Bryukhova S, Renna F, Reino L, Xu C, Xiao Z, Correia R, Di Minin E, Ribeiro J, Vaz AS, 2023. Detecting wildlife trafficking in images from online platforms: a test case using deep learning with pangolin images. Biological Conservation, 279:109905. https://doi.org/10.1016/j.biocon.2023.109905

Chan DTC, Poon ESK, Wong ATC, Sin SYW, 2021. Global trade in parrots: influential factors of trade and implications for conservation. Global Ecology and Conservation, 30:e01784. https://doi.org/10.1016/j.gecco.2021.e01784

Chan WH, Fung BSB, Tsang DHK, Lo IMC, 2023. A freshwater algae classification system based on machine learning with StyleGAN2-ADA augmentation for limited and imbalanced datasets. Water Research, 243:120409. https://doi.org/10.1016/j.watres.2023.120409

Chowdhury S, Fuller RA, Ahmed S, Alam S, Callaghan CT, Das P, Correia RA, Di Marco M, Di Minin E, Jarić I, Labi MM, Ladle RJ, Rokonuzzaman M, Roll U, Sbragaglia V, Siddika A, Bonn A, 2023. Using social media records to inform conservation planning. Conservation Biology, 38:e14161.

Collar NJ, Marsden SJ, 2014. The subspecies of Yellow-crested Cockatoo *Cacatua sulphurea*. Forktail, 30:23-27.

Cussen VA, 2017. Psittacine cognition: individual differences and sources of variation. Behavioural Processes, 134:103-109. https://doi.org/10.1016/j.beproc.2016.11.008

Dang F, Chen D, Lu Y, Li Z, 2023. YOLOWeeds: a novel benchmark of YOLO object detectors for multi-class weed detection in cotton production systems. Computers and Electronics in Agriculture, 205:107655. https://doi.org/10.1016/j.compag.2023.107655

Del Hoyo J, 2020. All the birds of the world. Lynx Edicions, Barcelona, pp. 364-366.

Esipova O, Love E, Noakes A, Schatz A, Swartz K, Vallianos C, 2021. Wildlife trafficking detection tools: best practices and application to the illegal rhino horn trade. Conservation International, Arlington, VA, USA, pp. 30-33.

Forshaw JM, 2010. Parrots of the world. Princeton University Press, Princeton, NJ, pp. 1-336.

Furnell S, 2019. Strengthening CITES processes for reviewing trade in captive-bred specimens and preventing mis-declaration and laundering: a review of trade in Southeast Asian parrot species [Internet]. BirdLife International, Cambridge, Accessed 12 Mar 2024, <https://coilink.org/20.500.12592/4njf3d>.

Huerta-Ramos G, Luštrik R, 2021. Inat_Images: v.1.1 [Internet]. Zenodo, Accessed 12 Mar 2024, <https://doi.org/10.5281/zenodo.4733367>.

Hughes AC, 2021. Wildlife trade. Current Biology, 31:R1218-R1224. https://doi.org/10.1016/j.cub.2021.08.056

Jang W, Lee EC, 2021. Multi-class parrot image classification including subspecies with similar appearance. Biology, 10:1140. https://doi.org/10.3390/biology10111140

Jocher G, Chaurasia A, Qiu J, 2023. Ultralytics YOLOv8 [Internet]. Github, San Francisco, CA, Accessed 16 Mar 2024, <https://github.com/ultralytics/ultralytics>.

Jocher G, Chaurasia A, Stoken A, Borovec J, Kwon Y, Michael K, Fang J, Wong C, Yifu Z, Montes D, 2022. ultralytics/yolov5:

v6. 2-yolov5 classification models, apple m1, reproducibility, clearml and deci. ai integrations [Internet]. Github, San Francisco, CA, Accessed 16 Mar 2024, <https://github.com/ultralytics/yolov5>.

Kim JI, Baek JW, Kim CB, 2022. Image classification of amazon parrots by deep learning: a potentially useful tool for wildlife conservation. Biology, 11:1303. https://doi.org/10.3390/biology11091303

Kim YG, 2004. AutoCrawler [Internet]. Github, San Francisco, CA, Accessed 2 Mar 2024, <https://github.com/YoongiKim/AutoCrawler>.

Kulkarni R, Di Minin E, 2023. Towards automatic detection of wildlife trade using machine vision models. Biological Conservation, 279:109924. https://doi.org/10.1016/j.biocon.2023.109924

Mahendiran M, Parthiban M, Azeez PA, Nagarajan R, 2018. *In situ* measurements of animal morphological features: a non-invasive method. Methods in Ecology and Evolution, 9:613-623. https://doi.org/10.1111/2041-210X.12898

McClure EC, Sievers M, Brown CJ, Buelow CA, Ditria EM, Hayes MA, Pearson RM, Tulloch VJ, Unsworth RKF, Connolly RM, 2020. Artificial intelligence meets citizen science to supercharge ecological monitoring. Patterns, 1:100109. https://doi.org/10.1016/j.patter.2020.100109

Paramita C, Supriyanto C, Putra KR, 2024. Comparative analysis of YOLOv5 and YOLOv8 cigarette detection in social media content. Scientific Journal of Informatics, 11:341-352. https://doi.org/10.15294/sji.v11i2.2808

Parr M, Juniper T, 2010. Parrots: a guide to parrots of the world. Bloomsbury Publishing, London, pp. 1-436.

Roy AM, Bhaduri J, Kumar T, Raj K, 2023. WilDect-YOLO: an efficient and robust computer vision-based accurate object localization model for automated endangered wildlife detection. Ecological Informatics, 75:101919. https://doi.org/10.1016/j.ecoinf.2022.101919

Sary IP, Andromeda S, Armin EU, 2023. Performance comparison of YOLOv5 and YOLOv8 architectures in human detection using aerial images. Ultima Computing: Jurnal Sistem Komputer, 15:8-13. https://doi.org/10.31937/sk.v15i1.3204

Scheffers BR, Oliveira BF, Lamb I, Edwards DP, 2019. Global wildlife trade across the tree of life. Science, 366:71-76. https://doi.org/10.1126/science.aav5327

Seomun H, Kim HJ, 2016. The guideline for import and export review of CITES species. National Institute of Biological Resources, Incheon, pp. 76-132.

Sheng B, Wu L, Zhang N, 2023. A maturity detection method for *Hemerocallis citrina* Baroni based on lightweight and attention mechanism. Applied Sciences, 13:12043. https://doi.org/10.3390/app132112043

Shetty AD, Ashwath S, 2023. Animal detection and classification in image & video frames using YOLOv5 and YOLOv8. In: 2023 7th International Conference on Electronics, Communication and Aerospace Technology (ICECA). Institute of Electrical and Electronics Engineers, New York, pp. 677-683.

Spiesman BJ, Gratton C, Hatfield RG, Hsu WH, Jepsen S, McCornack B, Patel K, Wang G, 2021. Assessing the potential for deep learning and computer vision to identify bumble bee species from images. Scientific Reports, 11:7580. https://doi.org/10.1038/s41598-021-87210-1

Tannous M, Stefanini C, Romano D, 2023. A deep-learning-based detection approach for the identification of insect species of economic importance. Insects, 14:148. https://doi.org/10.3390/insects14020148

Tian Z, Shen C, Chen H, He T, 2019. FCOS: fully convolutional one-stage object detection. Preprint at arXiv: https://doi.org/10.48550/arXiv.1904.01355

Trail PW, 2021. Morphological analysis: a powerful tool in wildlife forensic biology. Forensic Science International: Animals and Environments, 1:100025. https://doi.org/10.1016/j.fsiae.2021.100025

Tzutalin D, 2015. LabelImg. version 1.8.1 [Internet]. GitHub, San Francisco, CA, Accessed 9 Mar 2024, <https://github.com/tzutalin/labelImg>.

Vo HT, Thien NN, Mui KC, 2023. Bird detection and species classification: using YOLOv5 and deep transfer learning models. International Journal of Advanced Computer Science and Applications, 14:939-947. https://doi.org/10.14569/IJACSA.2023.01407102

Wang CY, Yeh IH, Liao HYM, 2024. YOLOv9: learning what you want to learn using programmable gradient information. Preprint at arXiv: https://doi.org/10.48550/arXiv.2402.13616

Wang G, Chen Y, An P, Hong H, Hu J, Huang T, 2023. UAV-YOLOv8: a small-object-detection model based on improved YOLOv8 for UAV aerial photography scenarios. Sensors, 23:7190. https://doi.org/10.3390/s23167190

Yan B, Fan P, Lei X, Liu Z, Yang F, 2021. A real-time apple targets detection method for picking robot based on improved YOLOv5. Remote Sensing, 13:1619. https://doi.org/10.3390/rs13091619

Zhang Q, Yi X, Guo J, Tang Y, Feng T, Liu R, 2023. A few-shot rare wildlife image classification method based on style migration data augmentation. Ecological Informatics, 77:102237. https://doi.org/10.1016/j.ecoinf.2023.102237

Zhong Y, Wang J, Peng J, Zhang L, 2020. Anchor box optimization for object detection. Preprint at arXiv: https://doi.org/10.48550/arXiv.1812.00469

Received June 12, 2024
Revised September 25, 2024
Accepted September 25, 2024

Appendix 1. The number of images of the 13 *Cacatua* species collected from three databases

No.	Species	eBird	iNaturalist	Google
1	*Cacatua alba*	97	17	–
2	*Cacatua citrinocristata*	26	9	72
3	*Cacatua ducorpsii*	21	12	62
4	*Cacatua galerita*	5,860	52	–
5	*Cacatua goffiniana*	204	119	–
6	*Cacatua haematuropygia*	72	12	–
7	*Cacatua leadbeateri*	517	278	–
8	*Cacatua moluccensis*	66	6	–
9	*Cacatua ophthalmica*	25	8	56
10	*Cacatua pastinator*	237	17	–
11	*Cacatua sanguinea*	1,627	16	–
12	*Cacatua sulphurea*	12	20	32
13	*Cacatua tenuirostris*	2,766	1,130	–