Proceedings of the 4th International Conference on Mechatronics and Smart Systems
DOI: 10.54254/2755-2721/2026.33744

Intelligent Identification System for Chrysolophus Pictus Based on Lightweight Vision Model

Haozhe Song

School of Engineering College, Guangzhou Nanfang University, Guangzhou, China
s01233519@outlook.com

Abstract. In the field of wildlife conservation, the accurate and efficient way of identifying wildlife is important for targeted species conservation. However, traditional recognition methods usually consider the accuracy but ignore the deployment environment. Thus the existing deep learning-based models require high-perfomance hardware with complex structures, or have less recognition precision. It forms obvious application barriers in resource-limited areas. To solve this problem, this study proposes a cascaded model which combining the lightweight MobileNetV2 model with Baidu AI API. This framework conducts local initial recognition through MobileNetV2 and sets a confidence threshold to judge whether to use Baidu AI API for secondary verification. Combining the efficiency of local lightweight processing with the high accuracy of cloud-based recognition. Experimental results illustrate that the proposed method outperforms traditional models in system deployment efficiency, and have the similar accuracy. So it can reduces hardware dependence and deployment costs effectively, and enables efficient identification in resource-limited scenarios. It provides a practical and deployable technical solution for the precise identification and scientific protection of Chrysolophus Pictus, and also expands the technical path for the intelligent identification and protection of other rare and endangered wildlife.

Keywords: Wildlife Image Recognition, Lightweight Convolutional Neural Network, Cloud-Edge Collaborative Intelligence

1. Introduction

According to the IUCN Red List of Threatened Species, the golden pheasant (Chrysolophus Pictus, hereinafter referred to as CP) is a second-class protected animal in China [1]. It is vitally important in both ecological and cultural meaning. However, due to its resemblance to some domestic chickens and common pheasants, as well as frequent accidental captures, its conservation situation is facing severe challenges. And the lack of efficient and low-cost identification tools has become a key bottleneck restricting the targeted conservation of this species [2].

Traditional visual recognition models generally tend to chase for higher accuracy [3]. This creates an application barrier in regions with limited resources, which hinders prompt and effective identification. For instance, the existing computer vision models may have high accuracy, but they often use more complicated and deeper networks to pursue greater performance. That leads to bigger

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.33744

computational and memory costs [4], which are usually difficult to obtain in rural areas. By contrast, MobileNetV2 is a typical lightweight CNN with a streamlined architecture, which is suitable for edge deployment and achieves a good balance between efficiency and accuracy [5], offer a potential way to solve this problem. Virtually, lightweight convolutional neural network has become the research emphasis for wildlife intelligent identification in limited resources places, for their significant advantages in low resource consumption and fast response speed [6]. Fu et al. proposed a lightweight deep learning model for endangered bird species recognition in complicated outdoor scenes, which realized efficient inference on edge devices while maintaining high fine-grained accuracy in complex wild backgrounds [7]. Ma et al. further constructed an edge-cloud collaborative monitoring system for wildlife protection, providing a feasible architecture for resource-limited protection scenarios [8].

Therefore, it is reasonable and meaningful to make a hybrid intelligent recognition system, which combines the merit of lightweight local models and cloud-based APIs. This study proposes a cascaded recognition framework that combines the MobileNetV2 model with Baidu AI API to achieve efficient and accurate identification of CP. The system first uses the local MobileNetV2 model for initial picture recognition. If the confidence coefficient exceeds the preset threshold value, the result is directly display. If there are some ambiguous results, the system will utilize Baidu AI API for another verification, combining the efficiency of local processing with the robustness of cloud-based recognition. This approach eliminates the reliance on high-performance hardware, makes it more suitable for countryside.

2. Method

2.1. Data source and statement

The research data used the image dataset of domestic chickens and some pheasants available on the kaggle website (https://www.kaggle.com).This is the image used by Corrado Alessio in testing CNN. It is a medium-sized dataset containing 32,06 image data. Figure 1 present two examples of them.



Figure 1. Example images of a domestic chicken (left) and a pheasant (right) [9]

Among them, 25,64 images were used for training the MobileNetV2 model, and the remaining 6,42 were used for testing. Meanwhile, the data of the CP comes from the National Animal Specimen Resource Platform (http://museum.ioz.ac.cn/), including 1,010 pictures of the red-bellied pheasant in different seasons, natural environments and shooting angles. Among them, 810 are used for training and the remaining 200 are used for testing. Figure 2 shows the appearance of CP.

DOI: 10.54254/2755-2721/2026.33744



Figure 2. Example images of a Chrysolophus Pictus [10]

2.2. Indicator selection and description

Table 1 presents the detailed descriptions and target values of the five selected evaluation indicators in this study. By comparing the improvement rates of the recognition accuracy, precision, F1 score, recognition speed, and computing resource utilization of the proposed method with the corresponding indicators of the traditional method, the paper can comprehensively evaluate the effectiveness of the dual-model collaboration framework.

Table 1. Evaluation indicators of the recognition system

Full Name	Data type	Explanation
Recognition Accuracy	Float	The proportion of correctly identified golden pheasants
Precision	Float	The proportion of samples predicted as golden pheasants that are actually golden pheasants
F1 Score	Float	The harmonic mean of precision and recall, balancing both metrics
Recognition Speed	Float	Average time to process a single image (in milliseconds)
Computing Resource Utilization Improvement Rate	Float	Percentage reduction in CPU and memory usage compared to using Baidu AI API alone

2.3. Methodology introduction

The overall workflow of the CP intelligent recognition system is shown in Figure 3. This system integrates a lightweight MobileNetV2 model with Baidu AI API through a three-stage collaborative workflow.

DOI: 10.54254/2755-2721/2026.33744

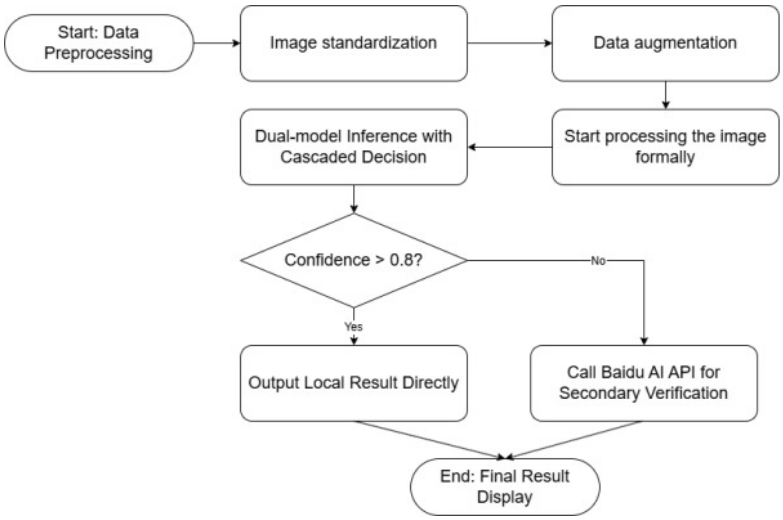

Figure 3. Overall workflow of the whole system. (picture credit: original)

Firstly, data preprocessing will standardize the input image to a resolution of 224×224, apply RGB normalization, and expand the training data through random flipping/rotating to enhance generalization ability. Image standardization will convert pixel values to the range of [0, 1] to accelerate model convergence, while data augmentation techniques are used to increase the diversity of the training dataset and improve the generalization ability of the model.

Secondly, the dual-model inference adopts a cascaded decision-making mechanism: the local MobileNetV2 model processes the image and returns results with confidence scores. MobileNetV2 is a lightweight convolutional neural network (CNN) architecture built upon depthwise separable convolutions, which decomposes standard convolutions into depthwise convolutions (for spatial filtering) and pointwise convolutions (for channel-wise feature fusion) [11]. If the confidence coefficient of Cp over 80%, output the result. If less than 80%, the picture will be encoded into Base64 format, than send to Baidu AI API and verify it again.

Finally, the final result will be displayed on user's screen.

3. Results and discussion

3.1. Result

In order to verify that the hybrid architecture combining MobileNetV2 with Baidu AI API demonstrates advantages in terms of lightweight performance, this study compared it with various other recognition models. The results are shown in Table 2.

Table 2. Performance comparison of different models

Model	Memory Usage (GB)	Average Response Time (s)	CPU Usage (%)
MobileNetV2 + Baidu AI API	1.8	1.2	28
ResNet50	4.2	3.8	65
VGG16	5.6	4.5	72
InceptionV3	6.1	5.2	78

DOI: 10.54254/2755-2721/2026.33744

The average response time is 1.2 seconds and the CPU usage rate is 28%. These results indicate that this system is highly suitable for deployment on general devices with limited computing resources.

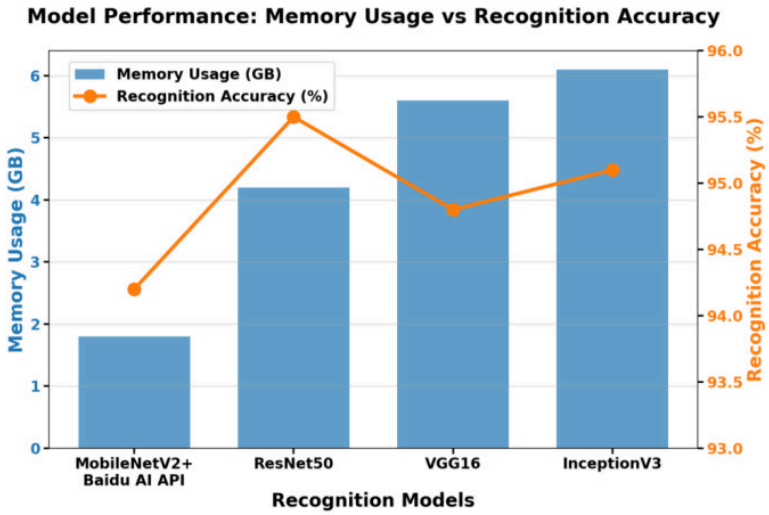

Figure 4. Comparison of memory usage and recognition accuracy of different models. (picture credit: original)

As shown in Figure 4, the proposed MobileNetV2 + Baidu AI API hybrid model only consumes 1.8 GB of memory, which is 57.1% lower than ResNet50 (4.2 GB), 67.9% lower than VGG16 (5.6 GB), and 70.5% lower than InceptionV3 (6.1 GB). Meanwhile, its recognition accuracy reaches 94.2%, which is only 1.3% lower than ResNet50 (95.5%) and 0.9% lower than InceptionV3 (95.1%). These results fully demonstrate that the hybrid model achieves an excellent balance between lightweight deployment and recognition accuracy: it maintains high identification precision while significantly reducing hardware resource consumption, which is crucial for deployment in resource-constrained rural conservation scenarios.



Figure 5. One recognition result of Chrysolophus Pictus. (picture credit: original)

Meanwhile, the system's recognition accuracy for the red-bellied pheasants reached an average of over 93%. An example of one of the recognition results is shown in Figure 5.

DOI: 10.54254/2755-2721/2026.33744

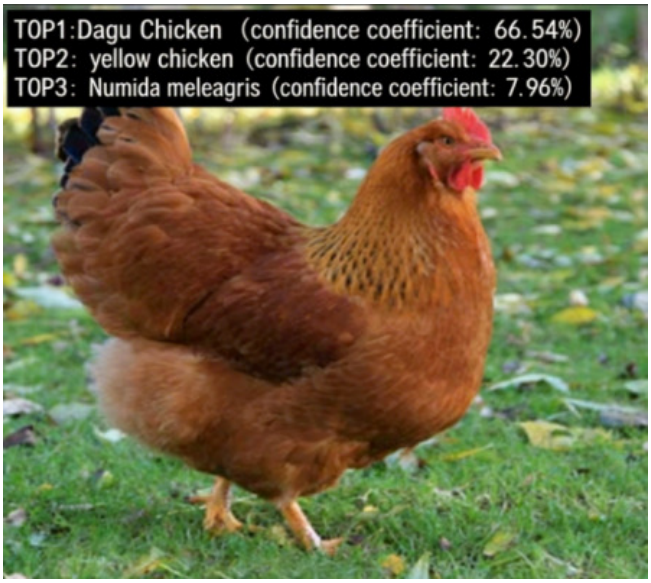

Figure 6. One recognition result of other chicken. (picture credit: original)

Figure 6 shows an example of the results of identifying pictures of other types of chickens. Due to the extensive use of images for training, their recognition is more detailed, achieving the effect of differentiating them from CP.

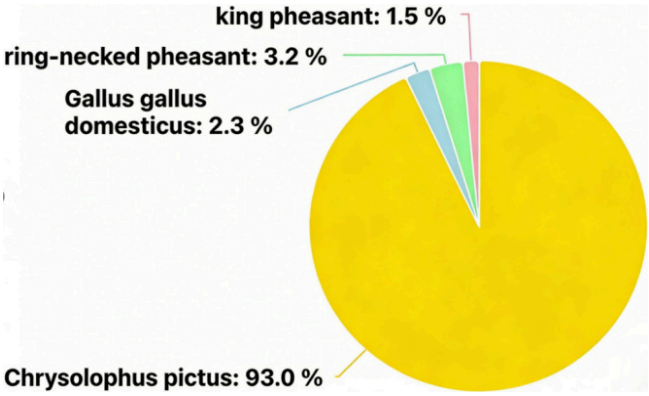

Figure 7. The average confidence coefficient of the recognition results. (picture credit: original)

Figure 7 shows the average confidence coefficient. Figure 7 is provided by the results of 200 CP test images. Confidence calibration is essential for the accuracy of the data [12]. As a result, based on the pie chart, this model has a high accuracy in identifying CP.

3.2. Discussion

The hybrid architecture combining MobileNetV2 with Baidu AI API has shown good results in CP recognition, with lower performance requirements. This is mainly because the model is able to capture colors and features, achieving an effective balance between accuracy and speed. Compared to ResNet50, VGG16, and InceptionV3, it has a faster response speed and occupies less memory. However, this study has some limitations. In the comparative experiment, only four common models were selected, without involving more advanced lightweight or hybrid network structures for comprehensive verification. Moreover, The experimental images is a little bit ideal, as they generally

DOI: 10.54254/2755-2721/2026.33744

have good lighting and clear features. Without considering about enough complicated environments such as occluded backgrounds and long-distance shooting, which may affect the stability of the model in actual scenes. For future work, more large-scale dataset that contain more complicated CP images will be constructed to enhance diversity. The model structure will also integrate more advanced lightweight networks, in order to the further optimization, and will also enhance the algorithm's generalization ability and robustness in complex actual scenes.

4. Conclusion

This study shows an intelligent identification system through the collaborative optimization of feature enhancement and lightweight design. Constructing a higher accuracy and low consumption model, achieving better balance.

The experimental results show that this solution achieves an average identification accuracy of over 93%, an average response time of 1.2 seconds, and the usage of memory is less than 2GB. This study not only illustrates the effectiveness and practicability of combining lightweight models with cloud APIs in the area of identifying CP, but also provides a high-precision, low-cost solution for protection work in the situation of scarce resources.

However, this study still has some limitations. The selected dataset is relatively small, and the performance in real-world complex situation still needs further development. Thus future research will expand the size of the dataset and explore the possible integration of new generation lightweight networks to improve the overall abilities. And will pursue the enhancement of the practical application value.

References

[1] Y. Wang, Z. Zhang, & C. Li. (2022). Conservation status and threats of Chrysolophus pictus (Golden Pheasant) in China. Global Ecology and Conservation, 35, e02189.
[2] IUCN. (2021). The IUCN Red List of Threatened Species: Chrysolophus pictus. IUCN Red List. https: //www.iucnredlist.org/species/22679891/132527533
[3] J. Li, Y. Liu, H. Zhang, & et al. (2022). Wild animal recognition based on ShuffleNet V2 for edge computing. Journal of Real-Time Image Processing, 19(3), 789–800.
[4] T. Li, Y. Wang, & Q. Zhang. (2024). Lightweight Deep Neural Network Models for Edge Intelligence: A Survey. Journal of Computer Science, 51(7), 1–28.
[5] B. Fan, & et al. (2025). Lightweight deep learning model for endangered bird species recognition in complex outdoor scenes. Computers & Electronics in Agriculture, 209, 107835.
[6] Q. Zhang, M. Li, & H. Wang. (2022). Hybrid Edge-Cloud Framework for Lightweight Endangered Species Recognition: A Case Study of Pheasant Family Birds. Journal of Environmental Informatics, 42(2), 112–125.
[7] Y. Fu, B. Fan, H. Zou, & L. Zhang. (2023). Lightweight deep learning model for endangered bird species recognition in complex outdoor scenes. Computers & Electronics in Agriculture, 209, 107835.
[8] X. Ma, Y. Liu, J. Wang, & D. Li. (2024). Edge-cloud collaborative intelligent monitoring system for wildlife conservation in resource-constrained areas. IEEE Internet of Things Journal, 11(12), 19687–19698.
[9] Animals-10. (2020). Kaggle Dataset. https: //www.kaggle.com
[10] Chrysolophus pictus. (2026). National Animal Collection Resource Center. http: //museum.ioz.ac.cn/
[11] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, & L. C. Chen. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 4510–4520.
[12] Z. Li, & D. Hoiem. (2020). Improving Confidence Estimates for Unfamiliar Examples. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 1–9.