Fine-grained detection of caged-hen head states using adaptive Brightness Adjustment in combination with Convolutional Neural Networks

Jia Chen, Qi'an Ding, Wen Yao, Mingxia Shen, Longshen Liu[*]

(*College of Artificial Intelligence, Nanjing Agricultural University, Nanjing 210095, China*)

Abstract: Timely identification and tracking of abnormal hens in stacked cages are of great significance for precision treatment and the elimination of sick individuals. The head features of the caged-hens are used to overcome observation difficulties caused by the cage and feathers blocking, but it is still hard to identify similar head states. To solve this problem, the fine-grained detection of caged-hens head states was developed using adaptive Brightness Adjustment in combination with Convolutional Neural Networks (FBA-CNN). Grid Region-based CNN (R-CNN), a convolution neural network (CNN), was optimized with the Squeeze-and-Excitation (SE) and Depthwise Over-parameterized Convolutional (DO-Conv) to detect layer heads from cages and to accurately cut them as single-head images. The brightness of each single-head image was adjusted adaptively and classified through the deep convolution neural network based on SE-Resnet50. Finally, we returned to the original image to realize multi-target detection with coordinate mapping. The results showed that the AP@0.5 of layer head detection using the optimized Grid R-CNN was 0.947, the accuracy of classification with SE-Resnet50 was 0.749, the F1 score was 0.637, and the mAP@0.5 of FBA-CNN was 0.846. In summary, this automated method can accurately identify different layer head states in layer cages to provide a basis for follow-up studies of abnormal behavior including dyspnea and cachexia.

Keywords: Grid R-CNN, squeeze-and-excitation, Depthwise Over-parameterized Convolutional, adaptive brightness adjustment, fine-grained detection

DOI: 10.25165/j.ijabe.20231603.7507

1 Introduction

Early warning of poultry diseases is very important for poultry health and production. Automated disease detection systems are of great significance to the poultry industry while preventing huge losses and keeping a good poultry welfare status[1,2].

To monitor the layers without human help, automation, and intelligent technology are applied to collect the phenotypic characteristics. To date, a variety of sensor technologies have been applied to monitor the body condition of chickens. Temperature measurements with non-contact methods[3,4] and wearable sensors[5,6] can help estimate chicken temperatures and determine abnormal individuals. Using sound detection, abnormal sounds[7-10], such as coughs, can be detected in chickens. Using machine vision, the shape and color of droppings[11] which have been associated with broiler digestive diseases can be detected as well. Image and video processing of poultry is an important part of abnormal poultry detection. Zhuang et al.[12] extracted the posture features of healthy and sick chickens to establish feature vectors. The posture of broilers was analyzed through machine learning algorithms to detect sick individuals. Okinda et al.[13] collected depth images of chickens

in top view and obtained characteristic parameters, such as eccentricity and walking speed. Six models, including Artificial Neural Network (ANN), were used to classify sick and healthy chickens with an accuracy of 97.8%. Zhuang and Zhang et al.[14] proposed a deep convolution neural network (CNN) model based on image information to detect broilers and identify their health status. Li et al.[15] extracted $a*$ from $L*a*b*$ space to detect chicken combs from images and used a support vector machine (SVM) to identify abnormal chickens from the comb color. Bi et al.[16] extracted the texture features of chicken heads and the geometric features of chicken eye pupils to form the feature vector for sick chicken recognition, using which an SVM can be used to identify sick chickens. Chen et al.[17] used Faster R-CNN to locate the heads and bodies of chickens, and fused semantic regional features to classify sick chickens.

For caged hens, given the limited range of layer activities and limited shielding of heads, head detection approaches offer an important reference to evaluate chicken health status. Chicken heads are defined as small targets in the image of the chicken group and in order to increase the precision of small target detection and decrease the rate of false or missed detection, different attention modules have been applied to improve the accuracy of the algorithm for small target detection[18-20]. Feature fusion for small targets can increase the information dimension[21-23]. These methods can detect small targets with significant phenotypic differences, but it remains difficult to identify the detail of chicken heads, including the state of the beak and the shape or color of combs, which has always been linked with *Leucocytozoonosis caulleryi*[24], infectious bronchitis[25] with uneven illumination and high-density breeding.

To solve these problems, the fine-grained detection of caged-hens head states was proposed using adaptive Brightness Adjustment in combination with Convolutional Neural Networks

(FBA-CNN). The accuracy of Grid R-CNN[26] was improved using Squeeze-and-Excitation[27] and Depthwise Over-parameterized Convolutional[28] to detect caged-hens head as 'head'. Each Region of Interest (RoI) of the head was cut as a small image and the adaptive brightness adjustment method was combined with brightness evaluation to keep these cut images within a similar brightness. The optimized images were classified by SE-Resnet50 and other deep convolution neural network backbones to identify the head state with similar features. Later, these small targets with coordinate and classification information were mapped to the original image in order to perform multi-detection of caged-hen heads in the image.

2   Materials and methods

2.1   Animals and experimental site

The experiment was conducted at Nanjing Lukou Poultry Development Co., Ltd., China, from October to November 2020 using 54-week-old laying hens, which were a mix of Highland gray and Highland brown. With the fixed light outside the cage, these high-density breeding caged hens were under uneven illumination as Figure 1. From both the actual photo in Figure 1a and the render in Figure 1b, it's shown that the caged hens in the same cage may be under different illumination because of their position and occlusion relations.

a. Actual photo of caged-hen breeding

b. Render of caged-hen breeding

Note: a and b show that the caged hens in the same cage may be under different illumination because of their position and occlusion relations.

Figure 1   High-density breeding caged hens were under uneven illumination

With the help of a vet, these layers were divided into five cages with different numbers of healthy and unhealthy individuals to simulate actual cage poultry breeding; for example, cage 1 had four healthy layers and four abnormal layers. Figure 2 shows the extracted head features of laying hens in the experiment.

In this experiment, three main head features were detected: normal (i.e., healthy), abnormal (i.e., unhealthy), and layer heads with beaks open. Several phenotypic abnormalities with actual reference values were selected as the research points, including

Figure 2   Extracted head features of laying hens

white combs, shrinking combs, and withered eyes because these phenotypic characteristics can offer important references[24,25] to evaluate chicken health status as listed in Table 1.

Table 1   Possible associated problems with different phenotypic characteristics

Focus	Phenotypic characteristics	Possible associated problems
Beak	Dyspnea	Infectious bronchitis and other respiratory diseases
	Long time open-beak	Heat stress
Comb	White combs	*Leucocytozoonosis caulleryi*, hepatic rupture
	Shrinking combs	Dysplasia
Eyes	Long time close-eyes	Mental sluggishness, symptoms of some diseases

2.2   Instrument

Images of the layers were obtained using a 2048×1536 Hikvision (DS-2CD3135F-L) camera fixed on a bracket with adjustable height; the distance between the camera and the cage was less than 45 cm. The field of view of the camera covers one single cage. Each cage of layers was photographed without interference for no less than 60 min intervals. The experimental setup is shown in Figure 3.
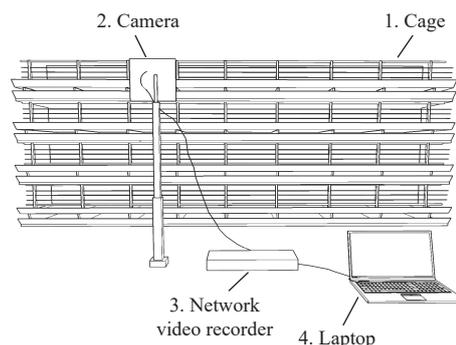
Figure 3   Schematic of the experimental setup

2.3   Data annotation

The image frames for each cage were extracted from the video every 15s. A total of 1350 images were obtained for the cages, with 1300 images available after data filtering. The layer heads were annotated as 'head,' 'om,' and 'unhealth' by Labelme, an open-source graphical image annotation tool inspired by MIT and CSAIL. The corresponding COCO dataset was made by labelme2coco, a Python package. To expand the data, the mixup[29] operation was used for data enhancement to create virtual samples in the training process as a new example could be formed by a weighted linear interpolation of two randomly-sampled examples.

2.4　Layer head detection algorithm combined with attention mechanism

To detect layer heads as a small target, Grid R-CNN was specifically optimized with SE and DO-Conv to improve the accuracy of the whole model. DO-Conv was used to replace traditional convolution to increase the expression ability of the network; the SE module was applied in the backbone to learn the weight coefficients of different channels.

2.4.1　Grid R-CNN

For Grid R-CNN[26], the feature maps obtained by a CNN backbone helped to extract features for each Region of Interest (RoI) individually based on region proposals. These RoI features are used to perform classification and localization for the corresponding proposals. Subsequently, a grid-guided mechanism is used for localization instead of offset regression, and a fully convolutional network is used to predict the locations of predefined grid points, and then, used them for the accurate object bounding box. A fully convolutional network is adopted in the grid prediction branch and outputs a fine spatial layout (probability heatmap). From the spatial layout, the grid points of the bounding box aligned with the object could be located. With the grid points, the accurate object bounding box could be determined with a feature map level information fusion approach.

2.4.2　Depthwise Over-parameterized Convolutional

Depthwise Over-parameterized Convolutional (DO-Conv)[28] can enhance the convolution layer by additional depthwise convolution, in which each input channel uses a different 2D kernel for convolution, and the composition of these two convolutions constitutes over-parameterization. In the depthwise convolution, the convolution kernel is split into three channels and convolutes for each channel without changing the depth of the input feature. In this way, the output feature with the same number of channels as the input feature could be obtained.

DO-Conv adds learnable parameters, and the generated linear operation can be represented by a single convolution layer, which means that the multi-layer composite linear operation used by over-parameterization can be folded into a compact single-layer form after the training stage, and then only a single-layer is used at inference time to reduce the computation to be exactly equivalent to a conventional layer. There are two kinds of calculation in DO-Conv: feature composition and kernel composition. The DO-Conv based on kernel composition was used here, its main operational process can be summarized by Equations (1) and (2).

$$O = W'P \tag{1}$$

$$W' = D^{\mathrm{T}}W \tag{2}$$

where, W' is the new weight, which should be used for traditional convolution with input features P to obtain results with the same receptive field as traditional convolution; D^{T} is the transpose of D, which is the kernel weights; O is the output.

Compared with traditional convolution, DO-Conv increases weight parameters without reducing calculation speed to increase efficiency.

2.4.3　Squeeze-and-Excitation

The Squeeze-and-Excitation (SE)[27] module is a combination of squeeze and excitation. The feature map obtained by convolution should be squeezed to obtain the global features. Subsequently, the global features are subject to excitation to learn the relationship between different channels and obtain the weights of these channels. Finally, the original feature map should be multiplied to obtain the final features. Thus, it can learn to use global information to selectively improve informative features and suppress less useful ones.

The feature maps are first passed through a squeeze operation, which produces a channel descriptor by aggregating feature maps across their spatial dimensions. The above descriptor can produce an embedding of the global distribution with channel-wise feature responses, which allows all its layers to use the information from the global receptive field of the network. The aggregation is followed by an excitation operation, which takes the form of a simple self-gating mechanism that takes the embedding as input and produces a weight for each channel. Finally, these obtained weights are applied to the feature maps to generate the output of the SE block for the next step in subsequent layers of the network.

2.5　Layer head classification based on local information

2.5.1　Adaptive brightness adjustment method

The uneven illumination of the cage results in poor contrast of some layer heads in images, making it difficult to identify important characteristics, such as color and shape. To address this problem, based on the layer head detection method above, an adaptive brightness adjustment method was proposed. The comprehensive brightness evaluation method with the root mean square (RMS) based on gray scale and RGB space was used to evaluate the brightness of the local area with each layer head as the main target using Equations (3)-(5).

$$V_{b1} = \frac{\sum_{i=1}^{w \cdot h} V_{\mathrm{gr}_i}^2}{w \cdot h} \tag{3}$$

$$V_{b2} = \sqrt{0.241 \cdot \frac{\sum_{i=1}^{w \cdot h} V_{r_i}^2}{w \cdot h} + 0.691 \cdot \frac{\sum_{i=1}^{w \cdot h} V_{g_i}^2}{w \cdot h} + 0.068 \cdot \frac{\sum_{i=1}^{w \cdot h} V_{b_i}^2}{w \cdot h}} \tag{4}$$

$$V_b = \frac{V_{b1} + V_{b2}}{2} \tag{5}$$

where, V_{b1} and V_{b2} are the results of the two brightness evaluations; w and h are the width and height of the picture, respectively; V_{gr} is the gray value of point i; V_r, V_{g_i}, and V_{b_i} are the value of points i in RGB channels.

After evaluation, the brightness could be adjusted based on the reference value using Equation (6).

$$V = \alpha * \frac{V_{\mathrm{ref}}}{V_b} * V_{\mathrm{origin}} \tag{6}$$

where, V_{ref} is the reference value of the layer head's local area with good illumination; V_{origin} is the original value of each point; the weight α is obtained from the experiment. Through this method, the detected layer head target could be standardized with illumination, and its phenotypic characteristics could be better presented.

2.5.2　Layer head classification based on CNN

On the basis of the above adaptive brightness adjustment, the layer head can be classified through the CNN, so that healthy and unhealthy layer heads could be classified based on phenotypic characteristics, such as color and shape. The classification model includes input, backbone, neck, classification head, and output. The SE-Resnet50 backbone network was used.

SE-Resent50 references VGG19 and adds a residual learning shortcut connection to learn the residual terms. In residual learning, if the learned residual value is 0, it is equivalent to making an identity mapping to avoid network performance degradation. In this experiment, it was embodied as two kinds of bottleneck (btnk); the SE block described in Section 2.2.3 was added in the btnk, and

could automatically remove low-weight noise points, retain high-weight noise points, and reduce parameter calculations.

2.6 Multi-classification caged-hen heads detection with FBA-CNN

The Overall process of FBA-CNN included different parts, as shown in Figure 4. Grid-CNN, which was improved by SE and DO-Conv to detect small targets accurately and the detect coordinate could help to determine the local region with one single head in it.

With the adaptive brightness adjustment method, each region could keep with similar lighting conditions and the details such as beak and comb could be identified the classification model combined with CNN and attention mechanism was applied to classify the heads into different types according to these details. In classification, different kinds of layer heads were annotated manually as "head" for normal layer heads, "om" for open mouths, and "unhealth" for unhealthy layer heads.
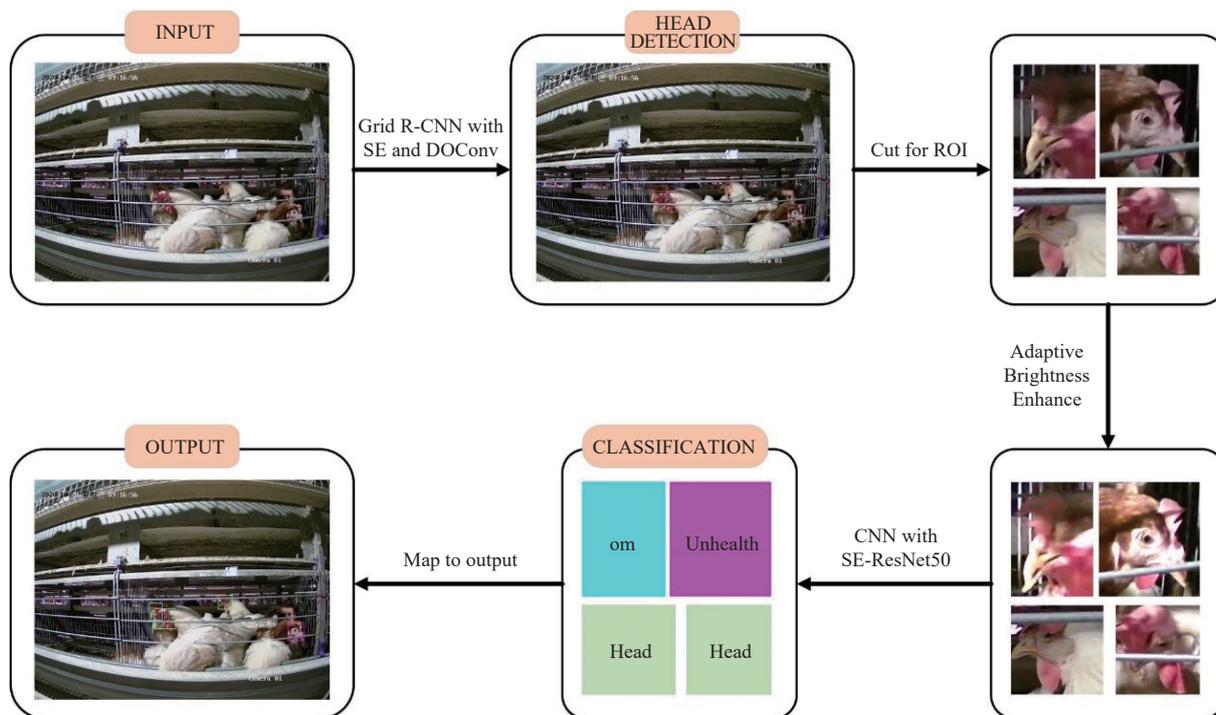
Note: FBA-CNN: Fine-grained detection of caged-hens head states was developed using adaptive Brightness Adjustment in combination with Convolutional Neural Networks. R-CNN: Region-based Convolutional Neural Network; SE: Squeeze-and-Excitation; DOConv: Depthwise Over-parameterized Convolutional; ROI: Research of Interest. "head" for normal layer heads, "om" for the open beak, and "unhealth" for unhealth layer heads.

Figure 4   FBA-CNN flowchart used in this study

3　Results

3.1　Experimental setting

3.1.1　Sample setting

The 54-week-old laying hens were divided into five cages. After the screening, 1300 images were obtained and included activities of the layers at different times and in different states. As listed in Table 2, different kinds of layer heads were annotated manually as "head" for normal layer heads, "om" for the open beak, and "unhealth" for unhealthy layer heads, which corresponded to 3696, 174, and 2710, labels respectively.

3.1.2　Evaluation indices

The evaluation indices used in the experiment were divided into three parts. For the evaluation of the layer head detection algorithm, Average Precision with IoU 0.5 (AP@0.5), Average Precision with IoU 0.50 to 0.95 (AP@0.5:0.95), and Average Recall (AR) were used. To evaluate the layer head classification algorithm, four different indices (Accuracy, Precision, Recall, and F1-score) were used. As the combination of the layer head detection algorithm and classification algorithm, multi-target detection was evaluated by the AP of each target and mean Average Precision (mAP). These indices were defined as follows:

Based on True Positive (TP), true Negative (TN), False Positive (FP), and False Negative (FN), Accuracy (Acc), Precision,

Table 2　Layer head annotations

Image ID	Number of labels		
	Head	Om	Unhealth
10	3	1	4
12	2	2	2
16	3	0	3
54	3	1	2
56	1	1	4
62	2	1	3
182	5	0	2
186	4	0	2
216	4	0	3
443	3	0	4
501	4	0	2
538	3	0	3

Note: "Head" for normal layer heads, "om" for the open beak, and "unhealth" for unhealthy layer heads.

and Recall can be calculated using Equations (7)-(9)[30].

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

$$Precision = \frac{TP}{TP + FP} \tag{8}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{9}$$

where, Acc is the proportion of all predictions that are correct; Precision is the proportion of all positive predictions that are correct; Recall is the proportion of all real positive observations that are correct; TP is the number of positive classes predicted as positive classes; TN is the number of negative classes predicted as negative classes; FP is the number of negative classes predicted as positive classes; FN is the number of positive classes predicted as negative classes.

On this basis, the F1-score was used for further evaluation. It considered accuracy and recall of the classification model at the same time, which can be regarded as a weighted average of accuracy and recall. The equation of F1-score is as follows:

$$\text{F1-score} = 2\frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{10}$$

Intersection over Union (IoU), which is the intersection and union ratio of the ground truth and the detection result, is expressed as follows:

$$\text{IoU} = \frac{GT \cap DT}{GT \cup DT} \tag{11}$$

where, GT is the ground truth, and DT is the detection truth. According to IoU, the accuracy of the detection target in the image can be calculated and the AP can be calculated using Equation (12):

$$\text{AP}_C^{\text{IoU}} = \frac{\sum \text{Precision}_C^{\text{IoU}}}{N(\text{TotalImages})_C} \tag{12}$$

where, $N(\text{TotalImages})_c$ is the total number of images with category C and $\text{Precision}_C^{\text{IoU}}$ is the precision of category C when the IoU is set. Since layer head detection is detection with one single category, the equations can be written as Equations (13) and (14).

$$\text{AP@0.5} = \frac{\sum \text{Precision}_{\text{head}}^{0.5}}{N(\text{TotalImages})_{\text{head}}} \tag{13}$$

$$\text{AP@0.5:0.95} = \frac{\sum_{\text{IoU}=0.5}^{0.95} \text{AP}^{\text{IoU}}}{10} \tag{14}$$

Since one single category had to be detected in this experiment, AR could be calculated using Equation (15) in which $N(\text{TotalPositive})_{\text{head}}$ is the total number of layer head positive samples in these images.

$$\text{AR} = \frac{\sum_{\text{IoU}=0.5}^{0.95} \text{TP}_{\text{head}}^{\text{IoU}}}{10 \cdot N(\text{TotalPositive})_{\text{head}}} \tag{15}$$

In the multi-target detection experiment, in addition to the AP calculation with IoU=0.5 for different categories, the concept of mAP@0.5 was introduced for comprehensive evaluation as,

$$\text{mAP@0.5} = \frac{\sum \text{AP}_C^{0.5}}{N(\text{class})} \tag{16}$$

where, $\text{AP}_C^{0.5}$ is the AP of category C with IoU=0.5, and $N(\text{class})$ is the total of all categories.

3.1.3 Parameter settings

The model used in this study depended on two kinds of software environment. The software environment information for layer head detection and classification is listed in Table 3.

In this layer head detection experiment, three algorithms (Grid R-CNN, Faster R-CNN, and Cascade R-CNN) were compared with different optimizations. In the layer head classification experiment,

the deep convolution neural networks with three different backbones (e.g., Resnet18, Resnet50, and SE-Resnet50) were compared. The parameter settings for the two experiments are listed in Table 4. To ensure an effective comparison, all hyper-parameters were consistent among the models.

Table 3　Software environment information for different models used in this study

Model kind	Software environment information	
Detection model	GPU	GeForce RTX 3090
	CUDA	11.0
	Python	3.7.10
	PyTorch	1.7.0
	OpenCV	4.5.2
	MMCV	1.3.5
Classification model	GPU	GeForce GTX 1660 Ti
	CUDA	10.1
	Python	3.7.9
	PyTorch	1.3.0
	OpenCV	4.4.0
	MMCV	1.3.10

Table 4　Hyper-parameter settings for the different models used in this study

Model kind	Hyper-parameter setting	
Detection model	Max Epoch	100
	Initial learning rate	0.02
	Step to reduce lr	10,19
	Weight decay	0.9
	momentum	0.0001
	trainset:valset:testset	7:2:1
Classification model	Max Epoch	200
	Initial learning rate	0.1
	Step to reduce lr	60,120,180
	Weight decay	0.9
	momentum	0.0001
	trainset:valset:testset	7:2:1

3.2　Evaluation of layer head detection

Comparative experiments were carried out with the parameter settings in Table 4 in order to verify the effect of the optimization operators. Under the same conditions, the effects of the original algorithm, the improved algorithm with the SE backbone, and the improved algorithm with the two optimizations in the backbone were compared. Figure 5 shows the Precision-Recall (PR) curve comparison for Grid R-CNN with the different backbones.

Figure 5 shows the PR curve comparison for Grid R-CNN with the different backbones. Grid R-CNN optimized by both SE and DO-Conv achieved the best performance for IoU in the range of 0.55 to 0.70 and showed greatly improved results compared with the original algorithm. The larger the IoU, the greater the advantages. The PR curves of Grid R-CNN with SE show that the SE had a good optimization effect here. The further addition of DO-Conv, an improvement of traditional convolution, also improved the detection accuracy to a certain extent.

Table 5 lists a comparison of the algorithms with the different optimizations. The AP of layer head detection was improved with DO-Conv and SE, and Grid R-CNN with SE and DO-Conv showed the best performance. The AP@0.5, AP@0.5:0.95, and AR of this algorithm were 0.947, 0.639, and 0.639, respectively. Compared

with the original Grid R-CNN algorithm, the indices were improved by 0.017, 0.055, and 0.049, respectively. The two optimizations improved Cascade R-CNN in the same way; however, for Fast R-

CNN, the SE improved the original algorithm more than the application of both optimizations. This may have been due to the different network structures, but this remains to be explored.
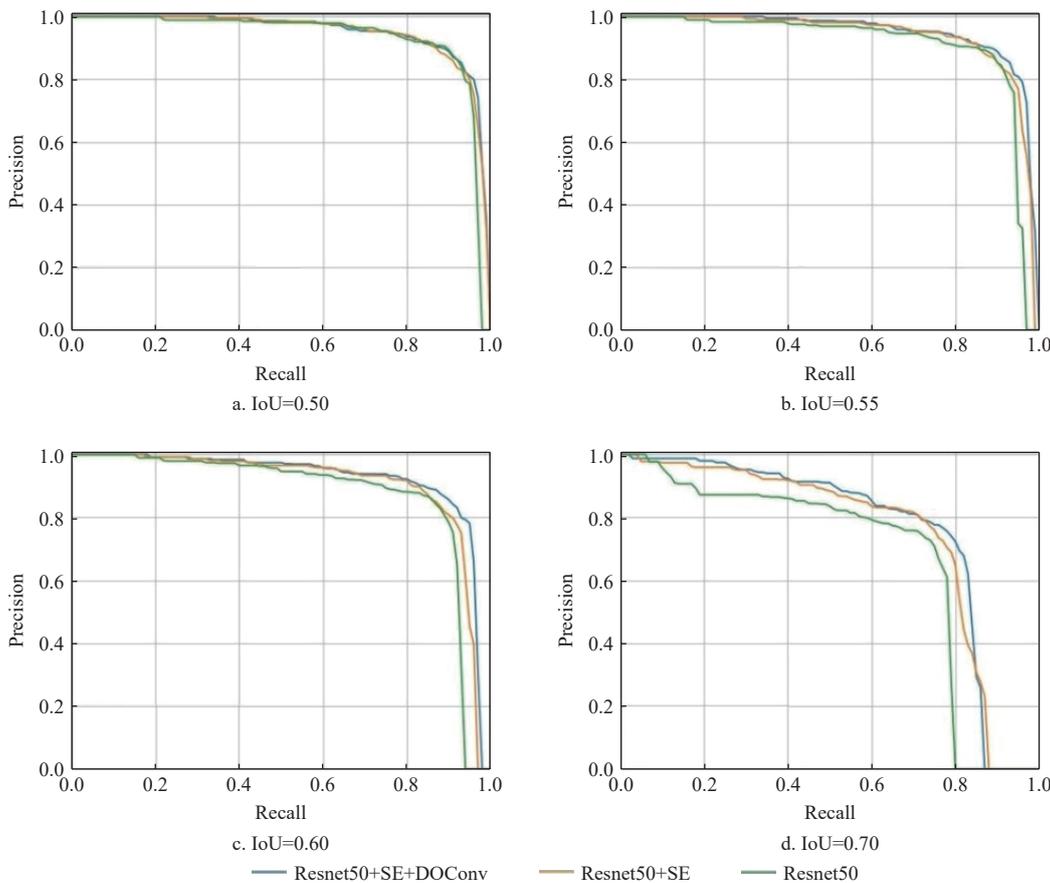
Note: IoU: Intersection over Union

Figure 5   Precision-recall (PR) curve comparison for the Grid R-CNN convolution neural network (CNN) with different backbones for Resnet50, Resnet50+ SE, and Resnet50+SE+DO-Conv

Table 5   Performance comparison of algorithms with different optimizations to detect heads

Algorithm	Backbone	AP@0.5	AP@0.5:0.95	AR
Grid R-CNN	Resnet50+SE+DOConv	0.947	0.639	0.639
	Resnet50+SE	0.944	0.634	0.640
	Resnet50	0.930	0.584	0.590
Faster R-CNN	Resnet50+SE+DOConv	0.910	0.572	0.574
	Resnet50+SE	0.924	0.621	0.623
	Resnet50	0.886	0.596	0.597
Cascade R-CNN	Resnet50+SE+DOConv	0.928	0.598	0.600
	Resnet50+SE	0.919	0.607	0.608
	Resnet50	0.902	0.599	0.600

Note: R-CNN: Region-based Convolutional Neural Network.

3.3 Evaluation of layer head classification

After the layer head detection, an adaptive brightness adjustment method was performed on the small image with the layer head as the main target. Images under different lighting conditions were adjusted via normalization, as shown in Figure 6.

The adaptive brightness adjustment method successfully improved the contrast of insufficiently bright images with a single layer head as the main target, clearly revealing phenotypic characteristics, such as the color and shape of the layer head. Moreover, other detailed actions (e.g., open beak) could be observed as well.
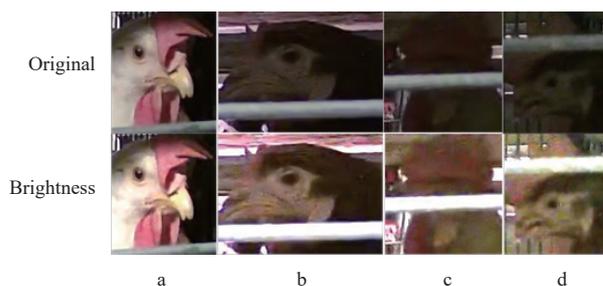
a. Layer head with suitable brightness b-d: Layer heads under different lighting conditions, obscured by the cage, or obscured by other layers

Figure 6   Results of adaptive brightness adjustment

Deep CNNs with different backbones were constructed to classify the processed layer head images, and the performances of three backbones (SE-Resnet50, Resnet50, and Resnet18) were compared (Table 6).

Table 6 lists that the number of backbone layers could influence the performance of the algorithm. Comparing the two backbones with the same number of backbone layers, compared with Resnet50, the Accuracy, Precision, Recall, and F1-score of SE-ResNet50 were improved by 0.022, 0.064, 0.012, and 0.023, respectively. With adaptive brightness adjustment, the classification method could receive a small increase in accuracy and precision, and the recall of SE-ResNet50 and Resnet50 has been improved

greatly. In summary, SE-ResNet50 with adaptive brightness adjustment could help to classify normal layer heads and abnormal layer heads accurately.

Table 6   Performance comparison of convolution neural networks (CNNs) with different backbones to classify heads

Operation	Backbone	Accuracy	Precision	Recall	F1-score
With adaptive brightness adjustment	SE-ResNet50	0.749	0.752	0.599	0.637
	ResNet50	0.727	0.688	0.587	0.614
	ResNet18	0.710	0.506	0.484	0.471
Without adaptive brightness adjustment	SE-ResNet50	0.746	0.517	0.509	0.498
	ResNet50	0.726	0.491	0.496	0.488
	ResNet18	0.714	0.491	0.489	0.479

3.4   Evaluation of fine-grained detection

The performances of layer head detection and layer head classification were both verified and the performance of multi-target detection was compared using the different algorithms.

Figure 7 shows the multi-target detection results of FBA-CNN and the other three algorithms. In Test 1, FBA-CNN detected most of the targets when other algorithms failed to detect some of them. Some unmarked layer heads were also detected using FBA-CNN. In Test 2, owing to the high similarity between the targets, Grid R-CNN and Faster R-CNN both had repeated recognition of the same target, while Cascade R-CNN had false detection. Compared with these algorithms, FBA-CNN detected most of the layer heads and classified them correctly, reflecting good accuracy. Table 7 lists the performance comparison of the proposed method of this study with the three other algorithms for multi-target detection for layer heads.
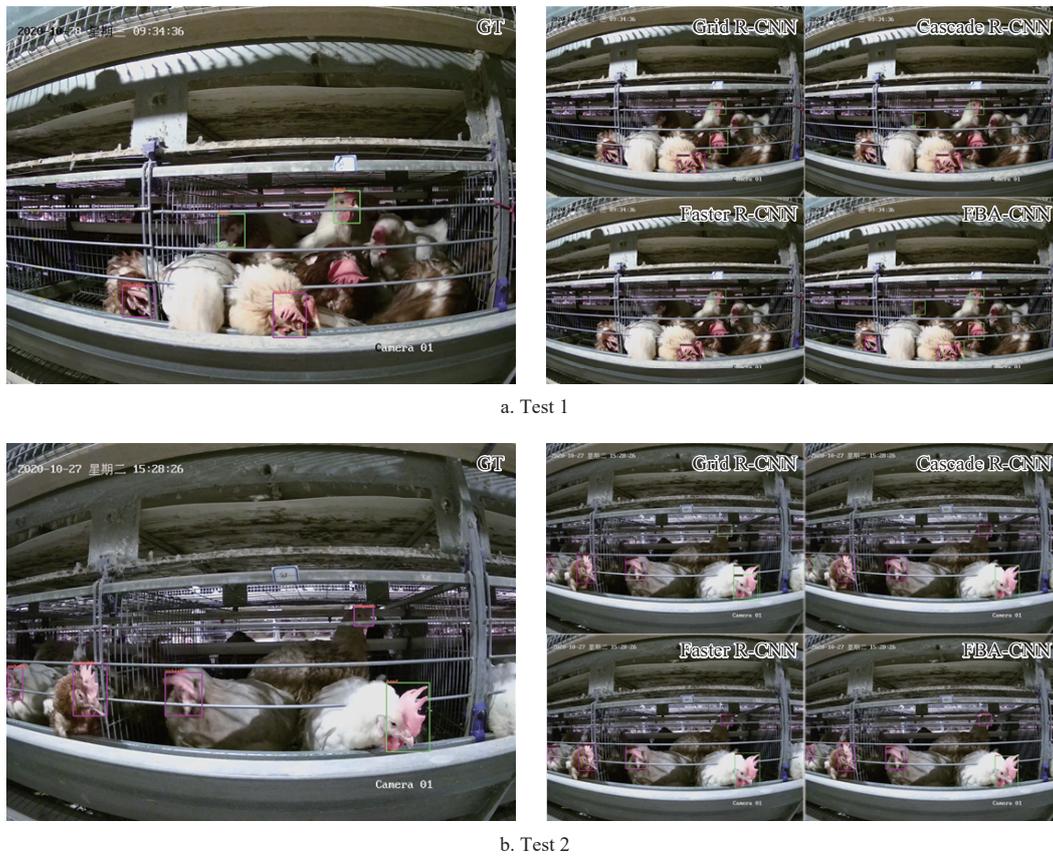
a. Test 1

b. Test 2

Note: Test 1 and Test 2, where the image to the top represents the Ground-Truth (GT) image and the image to the bottom is the comparison of different algorithms. The box in the images with different color is the detection box for different classification.

Figure 7   Multi-target detection using different algorithms

Table 7   Comparison of different algorithms to detect heads with different states

Algorithm	$AP_{0.5}^{head}$	$AP_{0.5}^{om}$	$AP_{0.5}^{unhealthy}$	mAP@0.5
FBA-CNN	0.839	0.829	0.870	0.846
Grid R-CNN	0.665	0.273	0.709	0.549
Faster R-CNN	0.695	0.142	0.720	0.519
Cascade R-CNN	0.668	0.193	0.716	0.526

Owing to the high similarity between "head" and "unhealth" with less light, and fewer "om" samples, it was easy to miss detection and false detection with Grid R-CNN, Faster R-CNN, or Cascade R-CNN. For these three algorithms, AP@0.5 of "head" with these three algorithms was in the range of 0.6 to 0.7, the AP@0.5 of "unhealth" was within the range of 0.7 to 0.8, the AP@0.5 of "om" was ~0.3, and the mAP@0.5 was ~0.5. For FBA-CNN, the AP@0.5 of "head" was 0.839, the AP@0.5of "om" was 0.829, the AP@0.5 of "unhealth" was 0.87, and the mAP@0.5 was 0.846 (representing improvements of 0.297, 0.327, and 0.32, respectively).

4   Discussion

In the cage environment, the layer head, which is easy to detect and informative, has a high reference value to judge layer health status. For example, dyspnea is the symptom of some chicken respiratory diseases, opening beak for a long time could be linked with heat stress, the white combs may be related to *Leucocytozoonosis caulleryi*, and the shrinking combs could

represent dysplasia. The judgment of abnormal chicken requires long-term continuous tracking, but manned patrol may cause chicken stress and it is difficult to continue manual observation around the clock.

In order to continuously monitor the caged hens with the unmanned method to identify the abnormal ones, chicken heads with different detail should be accurately detected and classified. Several object detection algorithms in the breeding field mainly tend to coarse-grained detection, such as the distinguishment between chicken individuals, and the detection of similar chicken heads with slight differences needs to be improved. Moreover, under the Chinese caged-hens breeding condition, with the fixed light source out of the cage, the illumination for the outer chicken head is very different from it for the inner chicken head and it is difficult to identify different types of chicken heads through details such as beak and comb.

It could be seen from Section 3.4 that object detection algorithms like Grid R-CNN could detect 'head' accurately, but the accuracy of 'om' and 'unhealth', including white combs and other abnormalities, is low. We tried a variety of optimization methods to increase the accuracy and found it difficult to improve the fine-grained detection because most of the false detections were inside the cage with less light. To solve this problem, the overall brightness of the image combined with the histogram was adjusted and this method was not suitable for the experimental environment in this study because each chicken head in one image should maintain similar light conditions. Therefore, it was tried to adjust local brightness and apply this method between the detection of all the heads and the classification of different head types.

In FBA-CNN, Gird R-CNN with SE and DO-Conv was used to detect layer heads in the image, which was the sum of 'head', 'om', and 'unhealth'. Each detection area contained a single head and was adjusted by the adaptive brightness method to keep with similar brightness and the details of the chicken head, such as beak and comb, could be highlighted. The experiment in 3.4 mentioned that another reason for the poor performance of object detection algorithms including Gird R-CNN was that there were fewer opening-beak samples, and the CNN model combined with the attention mechanism could better solve this problem to identify different head types. Finally, combined with the coordinate information and classification information by the above model, the original image was mapped to achieve a universal fine-grained detection that could be adaptive under different lighting conditions to accurately detect heads, opening-beak heads, and abnormal heads with phenotypic characteristics including white comb.

Compared with previous studies[15-17], this method decreases the influence of uneven illumination and provides a reliable basis for the fine-grained detection of different types of chicken heads. Each link of this method could be optimized to meet different needs, such as supplementing other abnormal chicken combs like purple combs in the classification link to directly detect abnormal chicken at the risk of Fowl Cholera. With the accurate detection of opening-beak ones, the dynamic abnormal behaviors that could reflect the health status of chickens, such as dyspnea and long-time opening-mouth, could be detected in future research.

5   Conclusions

Fine-grained detection of caged-hens head states using adaptive Brightness Adjustment in combination with Convolutional Neural Networks includes three main parts: head detection, adaptive brightness adjustment, and head classification. The optimized Grid R-CNN algorithm employing both SE and DO-Conv was used to detect caged-hen heads in cages as small targets, which were then optimized through the adaptive brightness adjustment method. The identified heads were classified through a deep neural network with SE-Resnet50 into different types. Finally, combined with the coordinate information and the classification information, the caged-hen heads could be detected in a complex cage environment with the correct classification such as the opening-mouth ones.

The experimental results of this study showed that the AP@0.5 of head detection for the optimized Grid R-CNN was 0.947, the accuracy and F1_score of head classification were 0.749 and 0.637, respectively, and the mAP@0.5 of FBA-CNN was 0.846. In summary, this non-human method could quickly and accurately identify different layer head states in cage poultry breeding to provide a basis for follow-up studies of abnormal behavior including dyspnea and cachexia.

In the subsequent research, the study of the non-human detection method for abnormal caged hens and apply FBA-CNN in dynamic abnormal behavior detection such as dyspnea and mental sluggishness should be further focused on. The problems of cage breeding, including serious occlusion and uneven illumination, still need to be improved and we will add group actions as influence factors to detect abnormal layers more accurately and reliably.

[References]

[1]   Neethirajan S, Tuteja S K, Huang S T, Kelton D. Recent advancement in biosensors technology for animal and livestock health management. Biosensors and Bioelectronics, 2017; 98: 398–407.

[2]   Okinda C, Nyalala I, Korohou T, Okinda C, Wang J T, Achieng, T, et al. A review on computer vision systems in monitoring of poultry: A welfare perspective. Artificial Intelligence in Agriculture, 2020; 4: 184–208.

[3]   Zhong Y F, Xiao H S, Liu N, Gao J K, Lin Z Y, Huang Z B. Breeding chicken temperature measurement system based on ZigBee. Science and Technology Innovation Herald, 2015; 12(8): 48. (in Chinese)

[4]   Yang W. Development of wireless wearable sensor equipment for monitoring layers' body temperature and experiment research. Master dissertation. Hangzhou: Zhejiang University, 2017; 72p. (in Chinese)

[5]   McManus C, Tanure C B, Peripolli V, Seixas L, Fischer V, Gabbi A M, et al. Infrared thermography in animal production: An overview. Computers and Electronics in Agriculture, 2016; 123: 10–16.

[6]   Shen M X, Lu P Y, Liu L S, Sun Y W, Xu Y, Qin F L. Body temperature detection method of ross broiler based on infrared rhermography. Transactions of the CSAM, 2019; 50(10): 222–229. (in Chinese)

[7]   Carpentier L, Vranken E, Berckmans D, Paeshuyse J, Norton T. Development of sound-based poultry health monitoring tool for automated sneeze detection. Computers and Electronics in Agriculture, 2019; 162: 573–581.

[8]   Banakar A, Sadeghi M, Shushtari A. An intelligent device for diagnosing avian diseases: Newcastle, infectious bronchitis, avian influenza. Computers and Electronics in Agriculture, 2016; 127: 744–753.

[9]   Mahdavian A, Minaei S, Yang C, Almasganj F, Rahimi S, Marchetto P M. Ability evaluation of a voice activity detection algorithm in bioacoustics: A case study on poultry calls. Computers and Electronics in Agriculture, 2020; 168: 105100.

[10]   Qin F L, Shen M X, Liu L S, Sun Y W, Zheng H H, Lu P Y, et al. Study on recognition algorithm of white feather broiler cough based on audio technology. Journal of Nanjing Agricultural University, 2020; 43(2): 372–378. (in Chinese)

[11]   Wang J T, Shen M X, Liu L S, Xu Y, Okinda C. Recognition and classification of broiler droppings based on deep convolutional neural network. Journal of Sensors, 2019; 2019: 3823515.

[12] Zhuang X L, Bi M N, Guo J L, Wu S Y, Zhang T M. Development of an early warning algorithm to detect sick broilers. Computers and Electronics in Agriculture, 2018; 144: 102–113.

[13] Okinda C, Lu M Z, Liu L S, Nyalala I, Muneri C, Wang J T, et al. A machine vision system for early detection and prediction of sick birds: A broiler chicken model. Biosystems Engineering, 2019; 188: 229–242.

[14] Zhuang X L, Zhang T M. Detection of sick broilers by digital image processing and deep learning. Biosystems Engineering, 2019; 179: 106–116.

[15] Li Y S, Mao W H, Hu X A, Zhang X C. Sick chicken detection with chicken crown color based on machine vision. Robot Technique and Application, 2014; 5: 23–25. (in Chinese)

[16] Bi M N, Zhang T M, Zhuang X L, Zhang X L, Qiao P R. Recognition method of sick yellow feather chicken based on head features. Transactions of the CSAM, 2018; 49(1): 51–57. (in Chinese)

[17] Chen Z B, Hou Y. Research on recognition of fine-grained sick chicken based on DCNN feature fusion. Journal of Lanzhou University of Arts and Science (Natural Sciences), 2020; 34(2): 79–84,104. (in Chinese)

[18] Hou L Y. Research on small object recognition algorithm based on SE-SSD network. Master dissertation. Dalian: Dalian University of Technology, 2020; 62p. (in Chinese)

[19] Zhang H T, Zhang M. SSD target detection algorithm with channel attention mechanism. Computer Engineering, 2020; 46(8): 264–270. (in Chinese)

[20] Wei L, Wang Y, Yao K M. Small target detection based on improved YOLO v4. Software Guide, 2021; 20(7): 54–58. (in Chinese)

[21] Zhu M C, Feng T, Zhang Y. Remote sensing image multi-target detection method based on FD-SSD. Computer Applications and Software, 2019; 36(1): 232–238. (in Chinese)

[22] Zheng P, Bai H Y, Li W, Guo H W. Small target detection algorithm in complex background. Journal of Zhejiang University (Engeering Science), 2020; 54(9): 1777–1784. (in Chinese)

[23] Wang N, Hu J H, Liu R K, Fan L C. Small target detection algorithm based on Bi-SSD. Computer Systems & Applications, 2020; 29(11): 139–144. (in Chinese)

[24] Yin Z Q, Wang K Y, Jia R Y. Pathomorphological observation on experimental *Leucocytozoonsis caulleryi*. Chinese Journal of Veterinary Science, 2002; 22(6): 597–600. (in Chinese)

[25] Cavanagh D. Severe acute respiratory syndrome vaccine development: experiences of vaccination against avian infectious bronchitis coronavirus. Avian Pathology, 2003; 32(6): 567–582.

[26] Lu X, Li B Y, Yue Y X, Li Q Q, Yan J J. Grid R-CNN. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach: IEEE, 2019; pp.7355–7364. doi: 10.1109/CVPR.2019.00754.

[27] Hu J, Shen L, Sun G. Squeeze-and-excitation networks. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City: IEEE, 2018; pp.7132–7141. doi: 10.1109/CVPR.2018.00745.

[28] Cao J M, Li Y Y, Sun M C, Chen Y, Lischinski D, Clhen-Or D, et al. Do-conv: Depthwise over-parameterized convolutional layer. IEEE Transactions on Image Processing, 2022; 31: 3726–3736.

[29] He T, Zhang Z, Zhang H, Zhang Z Y, Xie J Y, Li M. Bags of tricks for image classification with Convolutional Neural Networks. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach: IEEE, 2019; pp.558–567. doi: 10.1109/CVPR.2019.00065.

[30] David M W. Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. Journal of Machine Learning Technologies, 2011; 2(1): 37–63.