Detection of abnormal chicken droppings based on improved Faster R-CNN

Min Zhou[1*], Junhui Zhu[1], Zhihang Cui[1], Hongying Wang[1], Xianqiu Sun[2]
(1. *College of Engineering, China Agricultural University, Beijing 100083, China*;
2. *Shandong Minhe Pastoral Co., Yantai 265600, Shandong, China*)

Abstract: The characteristics of chicken droppings are closely related to the health of chickens. Veterinarians often judge the health of a chicken by looking at whether the chicken poop is normal. At present, the inspection of abnormal chicken droppings in chicken coops relies on manual observation, which is inefficient, accurate varies from person to person, labor-intensive, and has the risk of cross-infection. To achieve efficient, accurate, and intelligent identification of abnormal chicken droppings, an abnormal chicken droppings detection method based on improved Faster Region-based Convolutional Neural Network (Faster R-CNN) was proposed in this study. In the feature extraction network stage, deformable convolution was used and combined with Path Augmentation-Feature Pyramid Network (PA-FPN) to improve the extraction ability of features at different scales. In the Region Proposal Network (RPN) stage, the *K*-means++ algorithm was used to cluster the dataset and obtain the Anchor-ratio which is more suitable for the chicken poop object, and the FocalLoss classification loss function was used to improve the classification ability of difficult samples. In the regional convolutional network stage, the region of interest calibration algorithm was used instead to obtain more accurate localization information. The experimental results show that the improved Faster R-CNN structure can reach an accuracy of 98.8% for abnormal chicken poop detection, and the average accuracy mean value was improved by 27.8%. The results can provide a key core technology support for establishing an efficient abnormal chicken droppings online detection system.

Keywords: abnormal chicken droppings, Faster R-CNN, detection, non-destructive monitoring, PA-FPN
DOI: 10.25165/j.ijabe.20231601.7732

Citation: Zhou M, Zhu J H, Cui Z H, Wang H Y, Sun X Q. Detection of abnormal chicken droppings based on improved Faster R-CNN. Int J Agric & Biol Eng, 2023; 16(1): 243–249.

1 Introduction

The characteristics of chicken droppings are closely related to the health of chickens, and there are significant differences in shape and color between normal and abnormal chicken droppings[1-4]. For large-scale chicken breeding enterprises, the observation of chicken droppings properties is an important means to determine the presence of sick chickens. However, manual inspection is time-consuming and labor-intensive and cannot detect sick chickens in a timely manner, which can have a significant impact on chicken health and enterprise economic efficiency. Therefore, it is important to realize automatic and efficient detection of sick chickens to reduce labor intensity and improve the level of farm automation.

With the development of machine vision technology and its advantages of non-destructive and high accuracy[5], it has started to be widely used in chicken behavior identification[6-8], chicken welfare status monitoring[9,10], and sick chicken identification. Li et al.[11] performed multiple video acquisitions and extracted the chicken crown region, and established a classifier model based on the crown color using the support vector machine (SVM) algorithm, which can determine the sick chickens and locate their specific locations. Bi et al.[12] constructed a sick chicken recognition model based on chicken head features by extracting the crown and pendant regions and combining the geometric features of chicken eyes, and the correct rate of sick chicken recognition could reach 92.5%. Zhuang et al.[13] selected SVM classifier to achieve sick chicken recognition by comparing and analyzing the posture features and feature vectors of different features for broiler contour and bone structure. The traditional machine learning-based recognition methods mentioned above have high accuracy but require manual feature image extraction, which has a large limitation and poor model generalization ability.

Due to the advantages of Convolutional Neural Networks (CNN) with high detection accuracy, strong generalization ability, and automatic extraction of image features, deep learning techniques have also been applied in the field of livestock and poultry breeding disease detection[14-16]. Quach et al.[17] compared two deep networks, VGGNet and ResNet, for four common diseases in chickens, and achieved an accuracy of 74.1%. Zhuang et al.[18] achieved the automatic detection of diseased broiler chickens in a flock by fitting a chicken skeleton model based on a single-stage detection (SSD) network and fusing feature pyramid techniques. However, all of the above research methods are susceptible to the influence of the house environment and individual chicken overlap, as well chicken droppings are rich in chicken disease information and will not be affected by these[19]. Normal chicken droppings are usually in shades of brown and pretty solid with a kind of small covered white on top, abnormal droppings tend to be blood red or green in color and have an irregular shape. Therefore, abnormal

Received date: 2022-06-08 **Accepted date:** 2022-10-17
Biographies: **Junhui Zhu**, Master candidate, research interest: deep learning, livestock health, Email: jhz0203@163.com; **Zhihang Cui**, Master candidate, research interest: livestock health, Email: jaiteczh@163.com; **Hongying Wang**, PhD, Professor, research interest: livestock health, Email: hongyingw@cau.edu.cn; **Xianqiu Sun**, Livestock Pastor, research interest: feeding management of broiler chickens, Email: sunxianqiu@126.com.
***Corresponding author: Min Zhou,** PhD, Associate Professor, research interest: livestock health, machine vision. China Agricultural University, Beijing 100083, China. Tel: +86-18519651989, Email: zhoumin2016@ cau.edu.cn.

chicken droppings detection provides a new idea for sick chicken identification.

For broiler digestive tract diseases, Wang et al.[19] compared Faster Region-based Convolutional Neural Network (Faster R-CNN) with the YOLO-V3 network and adjusted the training parameters to achieve the identification of abnormal chicken droppings with four different traits, but there was a more obvious phenomenon of missed detection. Mbelwa et al.[20] addressed the above shortcomings and used two common chicken disease droppings as the research object, using a lightweight network architecture and combined with the model was not tested and validated. The current studies on the detection of abnormal chicken droppings are all about single dropping detection under experimental conditions, but not under actual farming environment and the comprehensive performance of the detection model is not good, as the images of chicken poop of different ages vary greatly and have different shapes under actual environment, which requires stronger feature extraction and generalization ability of the detection network.

In order to realize the detection of abnormal chicken droppings in an actual farming environment, this study proposed an improvement based on the Faster R-CNN network. The three stages of the original network are improved to construct an abnormal chicken droppings detection model. The improved network in this paper has higher detection accuracy and practical detection results.

2 Materials and methods

2.1 Dataset construction

2.1.1 Data collection

In this study, chicken droppings image data collection was carried out in a broiler house with six rows of eight layers in a stacked cage mode. A conveyor belt is installed between the upper and lower cages to clean the chicken droppings regularly, the image captured at the end of the manure cleaning belt is shown in Figure 1. A camera was fixed 40 cm above the end of the conveyor belt for chicken droppings image acquisition, the image acquisition system is shown in Figure 2.

Figure 1 A sample of the image captured at the end of the manure cleaning belt

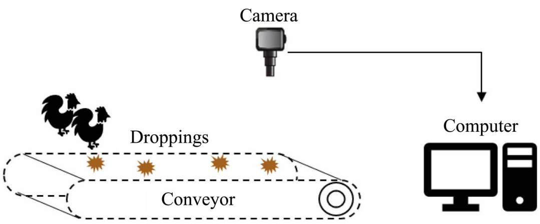


Figure 2 Image acquisition system used in this study

There are obvious differences in chicken droppings images at different day ages (Figure 3). In this study, chicken droppings images were tracked and collected within one growth cycle (43 d for broiler) and used frame extraction to intercept the video stream to obtain clear chicken droppings images. Through screening, a total of 2098 chicken droppings images were obtained at different day ages with a resolution of 3000×4000 pixels.

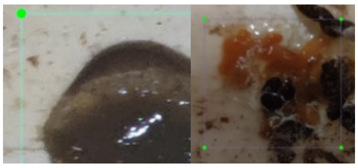
a. 3 days old b. 15 days old c. 30 days old

Figure 3 Images of chicken droppings at different ages

2.1.2 Data enhancement and annotation

In order to fully adjust the sample proportion of the data set, prevent overfitting of the model training, and improve the generalization ability of the model detection, the data set is expanded and adjusted by data augmentation. The enhanced dataset was further filtered to obtain a total of 5314 chicken droppings images. During the dataset annotation process, combining the characteristics of the constructed dataset and professional veterinary advice, the abnormal chicken droppings were divided into three categories: abnormal shape category, abnormal water content, and abnormal color category, as shown in Figure 4. Among them, the abnormal shape category means that the chicken droppings are unshaped and mixed with feed; the abnormal water category includes abnormal water content, and abnormal shape and water content, which means that the chicken droppings are watery and the water content is obviously abnormal, mostly because the chickens have diarrhea or drink too much water; the abnormal color category includes abnormal color and abnormal shape and color, which means that there are abnormal color chicken droppings such as blood droppings.

a. Abnormal shape b. Abnormal water, abnormal shape, and water



c. Abnormal color, abnormal shape, and color

Figure 4 Different abnormal chicken droppings traits

In order to adjust the sample ratio and reduce the training load of the network, a total of 1000 chicken droppings images were selected for the detection model and the image size was adjusted to 1333×800 pixels. The detailed configuration of the dataset is listed in Table 1.

Table 1 Detailed configuration of the data set

Label	Training set	Validation set	Test set	Total number
Normal sample	412	42	115	569
Abormal sample	308	38	85	431
Total number	720	80	200	1000

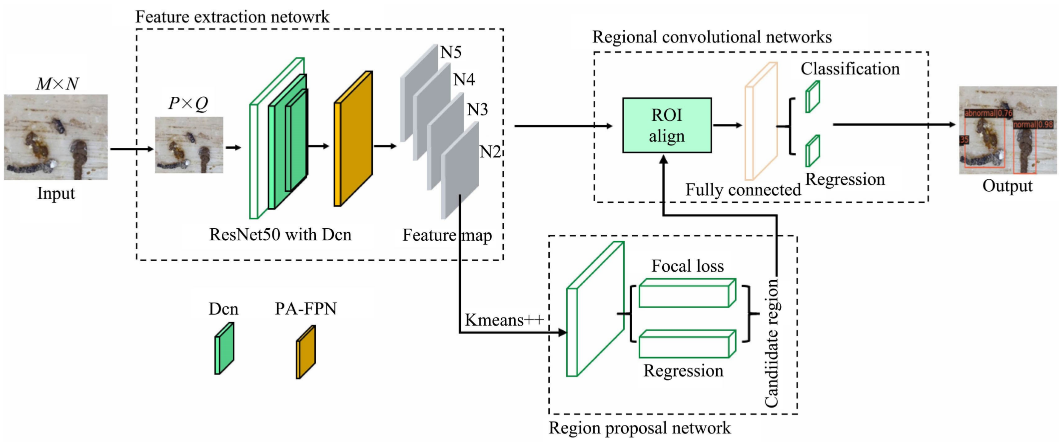
Note: Normal samples are images without abnormal droppings and abnormal samples are images containing abnormal droppings. Abnormal shape category, abnormal water content, and abnormal color category.

2.2 Abnormal chicken droppings detection model

2.2.1 Improved Faster R-CNN structure

In view of the characteristics of chicken droppings detection, the basic Faster R-CNN network has the following shortcomings: 1) The network only processes the last layer of the shared feature map of the feature extraction network, but the deep feature map has a large receptive field, and it is difficult to retain the information of small shallow targets, which causes the network to miss the detection of abnormal chicken droppings in small areas; 2) In the training process of Region Proposal Network (RPN) network, the default value of anchor point ratio for generating candidate frames is (1:1, 1:2, 2:1), which is difficult to adapt to the different shapes of chicken droppings images and affects the detection accuracy.

To address the above shortcomings, this study carried out three optimizations based on the existing Faster R-CNN network: 1) Incorporated the deformable convolution and Path Augmentation-Feature Pyramid Network (PA-FPN) structures; 2) Optimized Region Proposal Network; 3) Changed to the Region of Interest (ROI) Align algorithm the structure of the Faster R-CNN model built. The structure of the improved Faster R-CNN model is shown in Figure 5.



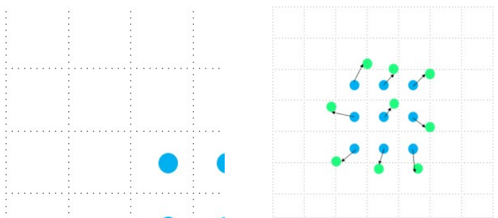
Note: $M×N$ is the initial image size; $P×Q$ is the adjusted input size; Dcn: Deformable convolution; PA-FPN: Path Augmentation-Feature Pyramid Network; ROI: Region of Interest. The white batches (N2-N5) are the shared feature maps obtained by convolutional layers.

Figure 5 Structure of the improved Faster R-CNN model

2.2.2 Fusion of Deformable Convolution and PA-FPN for Feature Extraction Network

This study used ResNet50 as the basic backbone feature extraction network and introduced deformable convolutional layers. The standard convolutional operation is difficult to adapt to the geometric deformation of the target, and the phenomenon of missed detection is easy to occur. Deformable convolution (Dcn) is a sampling position offset added to the traditional convolution operation, which improves the problem of restricted perceptual field and poor adaptation caused by the fixed size and position of the convolution kernel in the traditional convolution operation[21]. With the same convolutional kernel size, it has a larger perceptual field and is more adaptable to irregular target contours, as shown in Figure 6, so it is more suitable for chicken droppings images. In this study, the 3×3 standard convolutional layers in Stages 3, 4, and 5 of the ResNet50 network were replaced with deformable convolutions, and the adjusted network structure is shown in Figure 7.

In this study, the PA-FPN structure was used on the basis of FPN, which enhances the network's ability to express shallow feature information by adding bottom-up path aggregation routes. The PA-FPN network structure[25] is shown in Figure 8, where M2, M3, M4, and M5 are the feature maps obtained at different stages of ResNet50. The dotted arrow indicates the path of shallow and deep feature aggregation. The PA-FPN structure enhances the network's ability to represent shallow feature maps and improves the network's adaptability to anomalous chicken droppings at different scales.



a. Standard convolution b. Deformable convolution

Figure 6 Standard convolution and deformable convolution

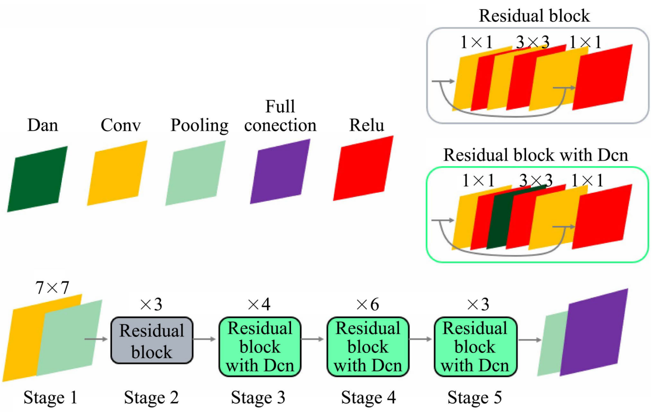


Figure 7 Structure of ResNet50 with Dcn network

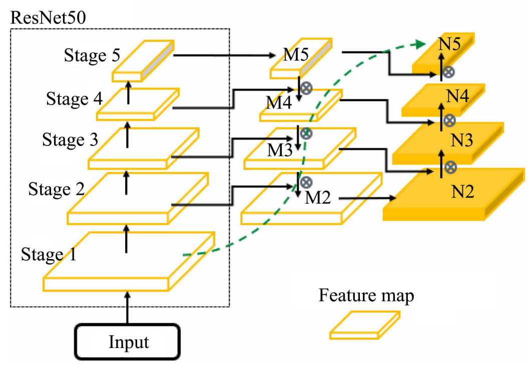


Figure 8 Structure of PA-FPN

2.2.3 RPN optimization based on K-means++ clustering algorithm

The K-means++ clustering algorithm was used to optimize the value of the anchor ratio for a priori anchor box generation to be more suitable for abnormal chicken droppings image detection. In this study, the aspect ratio data of the real target frame were obtained, and three clustering centers (0.16:0.21, 0.3:0.2, 0.12:0.11) were got as the new anchor ratio by this algorithm, and the clustering results are shown in Figure 9. The dots in the figure are the data distribution of the GT boxes aspect ratio of the dataset, the green inverted triangle is the original anchor point ratio, and the red marker is the anchor point ratio after clustering. After the clustering update, the a priori box aspect ratio matches better with the chicken droppings features, which enhances the network's ability to represent irregular chicken droppings images.

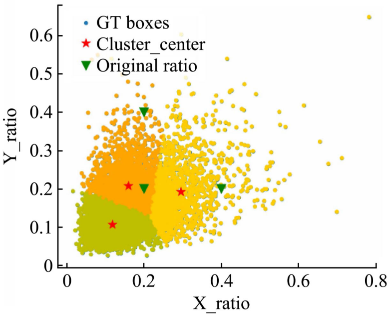


Figure 9 *K*-means++ clustering results

2.2.4 RPN classification loss function optimization

The stacking of chicken droppings and the presence of residual droppings in the scavenging zone at the later stage make it difficult to distinguish the foreground from the background, and the ratio of positive and negative samples appears to be unbalanced. A large number of positive samples containing foreground targets are treated as negative samples, which reduces the convergence speed and detection accuracy of the network. To solve this problem, this study adopted an improved Focal Loss function, as in Equation (1), and introduced equilibrium factors α and γ on the basis of the original cross-entropy loss function as follows:

$$L_{fl}=\begin{cases}-\alpha(1-y')^{\gamma}\log y', & y=1\\ -(1-\alpha)y'^{\gamma}\log(1-y'), & y=0\end{cases} \tag{1}$$

where, y' is the probability of classifying the current sample into a certain class of targets. y=1 indicates a positive sample (candidate frames containing foreground targets to be detected) and y=0 indicates a negative sample (candidate frames containing background regions).

2.2.5 Regional convolutional network optimization

In this study, the Roi Align algorithm was used instead of the original Roi Pooling operation. The improvement of the algorithm enhanced the detection ability of the model for smaller abnormal chicken droppings and improved the localization accuracy of the target detection frame.

3 Results and discussion

3.1 Hyperparameter setting

The Intel(R) Core(TM) i9-10900K(3.6GHz,16GB) processor, NVIDIA GeForce RTX 3070(8GB) display adapter, and 1T SATA mechanical hard drive were used as the hardware platform. The specific training parameters are listed in Table 2.

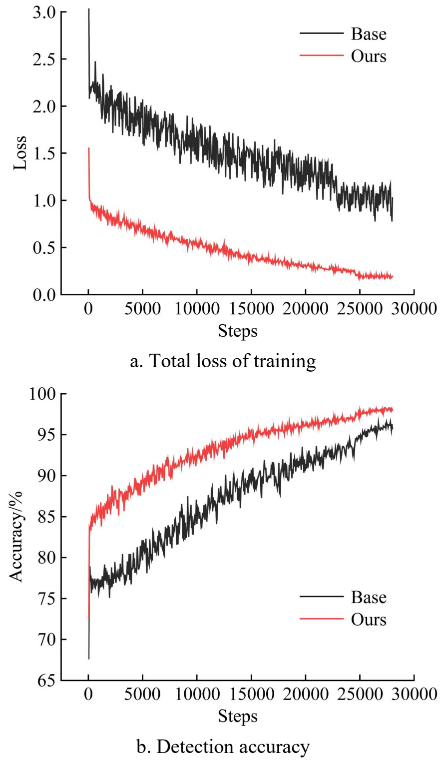
Table 2 Training parameter settings

Training parameter	Setting
Optimizer	SGD
Momentum	0.9
Learning rate	0.0025
Weight decay	0.0001
Batch size	2
Max epoch	40
Warmup	Linear
Warmup ratio	0.001
Warmup iters	500
Weight saving	10

Note: SGD: Stochastic Gradient Descent.

3.2 Comparison of training results

Under the same hardware and hyperparameter settings, the improved model of this study and the original Faster R-CNN were trained separately, and the total loss and accuracy comparison curves are shown in Figure 10. As can be seen from Figure 10a, at the end of the iteration, the total losses of training of this study model and the original Faster R-CNN were stable at about 0.15 and 0.75, respectively; The detection accuracy of this study model was significantly higher than that of the original Faster R-CNN, which can reach 98.8% (Figure 10b). This indicates that the improved model in this study has better convergence performance and detection accuracy.

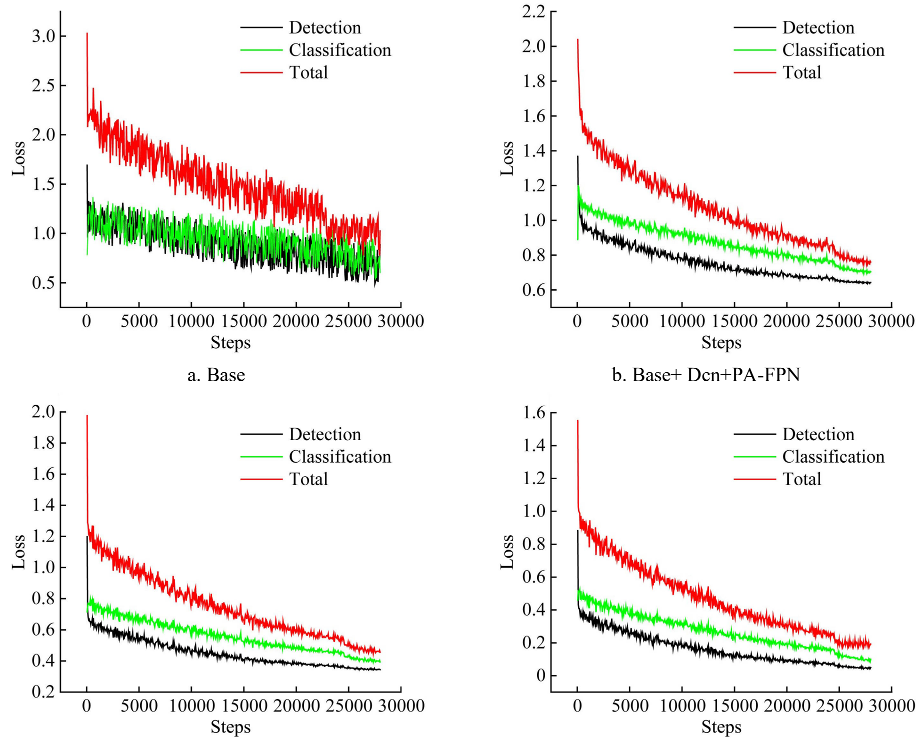


Note: 'Base' means the base model of Faster R-CNN; 'Ours' means the improved Faster R-CNN proposed in this study.

Figure 10 Performances comparison of the two models

To test the improvement effect of the different modules mentioned above on the detection performance, the four network models formed before and after optimizing the different modules were trained separately to visualize the specific training losses and evaluate the performance with accuracy and mean average precision (mAP) as metrics. The training losses of each model are shown in Figure 11, including detection loss, classification loss, and total loss. As can be seen from the figure, compared to the base model of Faster R-CNN (Figure 11a), the model loss curve is smoother and the classification and detection losses are lower after adding the Dcn and PA-FPN structures. After optimizing the RPN network, the model loss is further reduced, and finally, the model in this study (Figure 11d) achieves the best performance.

The improvement effect of different optimization methods is listed in Table 3, from which it can be seen that the base feature extraction network, with the addition of Dcn and PA-PFN, significantly improves the mAP value from the original 56.3% to 76.1%, which significantly enhances the feature extraction capability for targets of different scales. The mAP value is further improved by optimizing the RPN network with *K*-means++ algorithm and Focalloss loss function, and finally, the mAP value of the model in this study reaches 84.1%, which is 27.8% higher compared with the original Faster R-CNN, fully illustrating the effectiveness of the proposed improved method.



a. Base b. Base+ Dcn+PA-FPN

c. Base+ Dcn+PA-FPN+ *K*-means++ Focalloss d. Method proposed in this study

Note: 'Base' means the base model of Faster R-CNN.

Figure 11 Loss curves for different models

Table 3 Improvement effects of different optimization methods

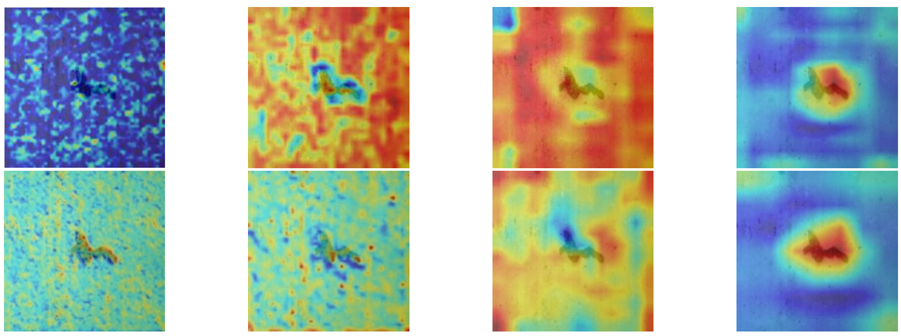
Improvement	ResNet50			
Dcn+PA-FPN	--	√	√	√
K-means++ Focalloss	--	--	√	√
RoiAlign	--	--	--	√
Accuracy/%	93.5	96.2	97.3	98.8
mAP/%	56.3	76.1	80.2	84.1

Note: '√' means the integration of current improvement strategies.

The GRAD-CAM algorithm was used to visualize and analyze the model before and after the improvement to show more clearly the focus areas of the model during training and detection. The heat map visualization of the output of different stages of the model feature extraction network before and after the improvement is shown in Figure 12. The area covered by the final stage of the improved model is significantly better than that of the pre-model, indicating that the network has better extracted the chicken poop features and has better detection performance.

Base

Mothed proposed in this study



a. Stage 1 b. Stage 2 c. Stage 3 d. Stage 4

Figure 12 Heat map of model feature extraction network before and after improvement

3.3 Detection performance comparison and analysis

In order to fully verify the detection performance and generalization ability of the improved model, a test set consisting of different day-old chicken droppings images was selected for detection comparison. In addition to this, the detection was compared with the anchor-free multiple object detectors YOLOv5 model under the same hardware conditions and parameter configuration, the results are shown in Figure 13, where, 10 d means the chicken is 10 days old. As can be seen from the figures, the YOLOv5 model showed significant leakage in the detection performance at different stages of age.

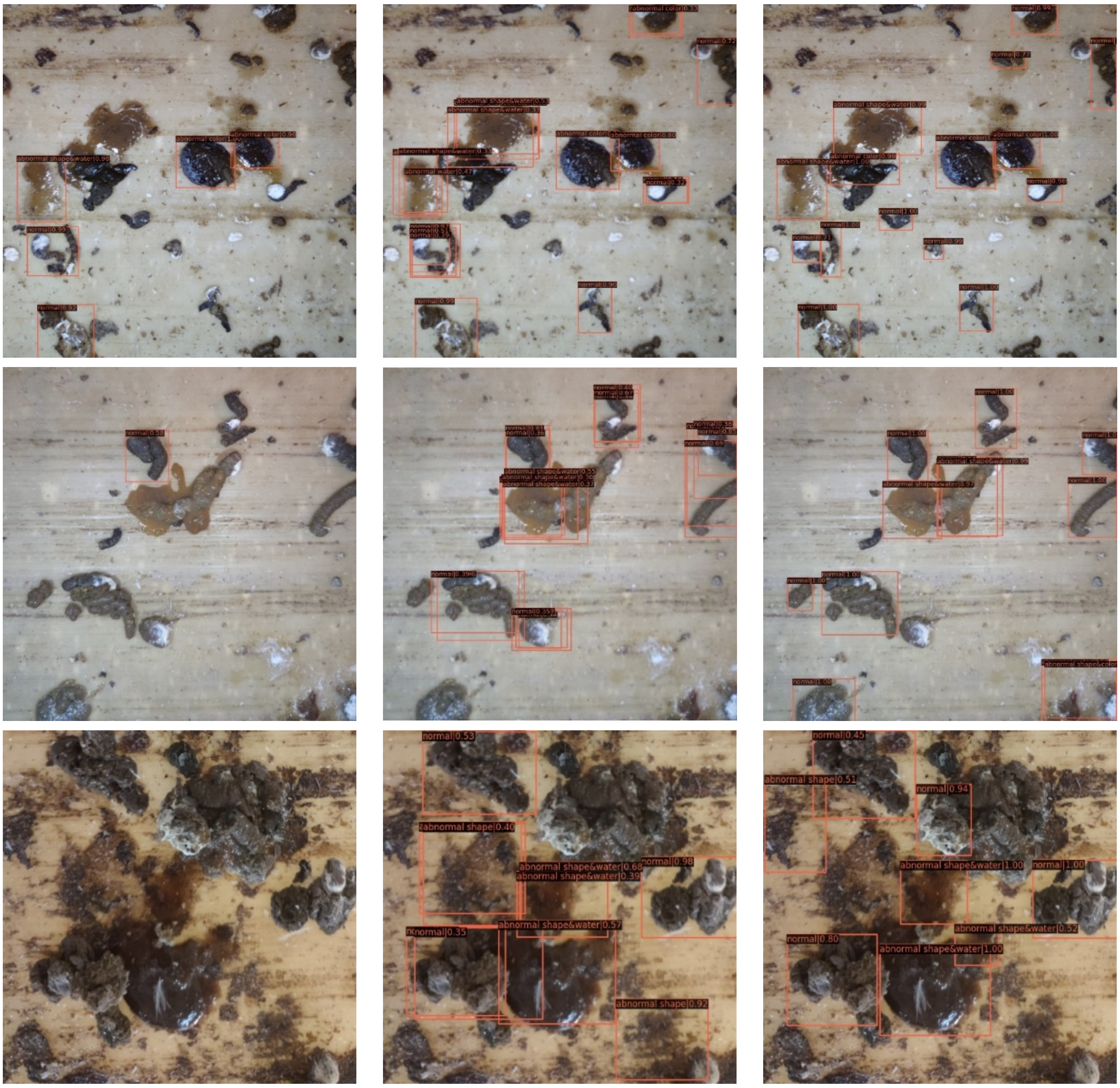
The original Faster R-CNN has the phenomenon of detection frame redundancy, the optimization of the RPN network by using the Focalloss loss function in this study significantly improved the above phenomenon. With the growth of chicken days, the missed detection and detection frame redundancy of the original Faster R-CNN is more obvious, and the detection frame localization is biased. In this study, based on the original feature extraction network, the deformable convolution and PA-PFN structure are fused to fully improve the detection ability of the model for different scale targets. The improved Faster R-CNN still performs well on more complex large-day-old chicken droppings images and improves the detection frame localization accuracy while improving the leakage problem.

In general, by improving the three modules of the original model, this study provides better detection performance for chicken droppings images of variable shapes and different day ages in real environments.

10 days old chicken

20 days old chicken

30 days old chicken



a. YOLOv5 b. Base Faster R-CNN c. Improved Faster R-CNN in this study

Figure 13 Detection results of different models

4 Conclusions

This study proposed an abnormal chicken droppings detection model for practical farming environment based on Faster Region-based Convolutional Neural Network (Faster R-CNN) network, which was adapted and optimized from three aspects respectively. Among them, the feature extraction network was fused with deformable convolution and Path Augmentation-Feature Pyramid Network (PA-FPN) to fully improve the multi-scale target adaptation of the model, the *K*-means++ algorithm and classification Focalloss loss function was used to optimize Region Proposal Network (RPN), and the region of interest calibration algorithm was used to obtain more accurate localization information. The improved Faster R-CNN network achieved an accuracy of 98.8% and the mean average precision (mAP) value is improved by 27.8% compared with the original network. In the task of detecting chicken droppings images of different days of age, the model in this study improved the detection frame localization accuracy while improving the leakage problem, which provides theoretical and technical support to realize efficient online monitoring of abnormal chicken droppings and accurate discrimination of different types of chicken diseases.

Acknowledgements

This study was financially supported by the 'Zhu Cheng Professorial Workstation of China Agricultural University', and the Science and Technology Cooperation Project between China and the United Kingdom under the National Key Research and Development Program (Envirobot an Autonomous Roving Platform for Environment, Health, and Welfare Monitoring of Poultry (Grant No. 2017YFE0122200).

[References]

[1] Sun L Q, Chen S H, Liu T, Liu C H, Liu Y. Pig target tracking algorithm based on multi-channel color feature fusion. Int J Agric & Biol Eng, 2020; 13(3): 180–185.

[2] Carpentier L, Berckmans D, Youssef A, Berckmans D, van Waterschoot T, Johnston D, et al. Automatic cough detection for bovine respiratory

disease in a calf house. Biosystems Engineering, 2018; 173: 45–46.

[3] Aziz N A, Othman M F B. Binary classification using SVM for sick and healthy chicken based on chicken's excrement image. Pertanika Journal Science and Technology, 2017; 25: 315–324.

[4] Li X H, Nie C S, Liu Y C, Chen Y, Lyu Z, Wang L, et al. The genetic architecture of early body temperature and its correlation with Salmonella pullorum resistance in three chicken breeds. Frontiers in Genetics, 2020; 10: 1287. doi: 10.3389/fgene.2019.01287.

[5] Chen B Q, Wu Z H, Li H Y, Wang J. Research of machine vision technology in agricultural application: today and the future. Science & Technology Review, 2018; 36(11): 54–65. (in Chinese)

[6] Lao F D, Du X D, Teng G H. Automatic Recognition method of laying hen behaviors based on depth image processing. Transactions of the CSAM, 2017; 48(1): 155–162. (in Chinese)

[7] Leroy T, Vranken E, Van Brecht A, Stuelens E, Sonck B, Berckmans D. A computer vision method for on-line behavioral quantification of individually caged poultry. Transactions of the ASABE, 2006; 49(3): 795–802.

[8] Abdanan Mehdizadeh S, Neves D P, Tscharke M, Nääs I A, Banhazi T M. Image analysis method to evaluate beak and head motion of broiler chickens during feeding. Computers and Electronics in Agriculture, 2015; 114: 88–95.

[9] Kashiha M, Pluk A, Bahr C, Vranken E, Berckmans D. Development of an early warning system for a broiler house using computer vision. Biosystems Engineering, 2013; 116(1): 36–45.

[10] Liu Y H, Liu X L, Hou R Y, Huang Y K, Lu H S. Research on chicken thermal comfort discrimination method based on machine vision. Heilongjiang Animal Husbandry and Veterinary Medicine, 2018(19): 11–14. (in Chinese)

[11] Li Y S, Mao W H, Hu S W, Zhang X C. A method for detecting diseased chickens based on machine vision recognition of crown color. Robot Technique and Applications 2014; 5: 23–25. (in Chinese)

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

[13] 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.

[14] Cuan K X, Zhang T M, Huang J D, Fang C, Guan Y. Detection of avian influenza-infected chickens based on a chicken sound convolutional neural network. Computers and Electronics in Agriculture, 2020; 178: 105688. doi: 10.1016/j.compag.2020.105688.

[15] Fang P, Hao H Y, Wang H Y. Behavior recognition model of stacked-cage layers based on knowledge distillation. Transactions of the CSAM, 2021; 52(10): 300–306. (in Chinese)

[16] Liu H W, Chen C H, Tsai Y C, Hsieh K W, Lin H T. Identifying images of dead chickens with a chicken removal system integrated with a deep learning algorithm. Sensors, 2021; 21(11): 3579. doi: 10.3390/s21113579.

[17] Quach L D, Pham-Quoc N, Tran D C, Hassan M F. Identification of chicken diseases using VGGNet and ResNet models. In: INISCOM 2020: Industrial Networks and Intelligent Systems, Springer, 2020; 334: 259–269. doi: 10.1007/978-3-030-63083-6_20.

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

[19] 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. doi: 10.1155/2019/3823515.

[20] Mbelwa H, Mbelwa J, Machuve D. Deep convolutional neural network for chicken diseases detection. International Journal of Advanced Computer Science and Applications, 2021; 12(2): 759–765.

[21] Dai J F, Qi H Z, Xiong Y W, Li Y, Zhang G D, Hu H, et al. Deformable convolutional networks. In: Proceedings of the IEEE international conference on computer vision, IEEE, 2017; pp.764–773.

[22] Liu S, Qi L, Qin H F, Shi J P, Jia J Y. Path aggregation network for instance segmentation. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City: IEEE, 2018; pp.8759–8768. doi: 1109/CVPR.2018.00913.

[23] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, IEEE, 2016; pp.770–778.

[24] Lin T Y, Dollár P, Girshick R, He K M, Hariharan B, Belongie S. Feature pyramid networks for object detection. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2017; pp.936–944.

[25] Chen H Y, Zhao P, Yan H W. Crack detection by multi-scale Faster RCNN with fused attention. Optoelectronic Engineering, 2021; 48(1): 61–71. doi:10.12086/oee.2021.200112.

[26] Lin T Y, Goyal P, Girshick R, He K M, Dollár P. Focal loss for dense object detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020; 42(2): 318–327.