



OPEN ACCESS

EDITED BY

Mohanned Alhussien,
Technical University of
Munich, Germany

REVIEWED BY

João Souza-Junior,
Federal University of Ceara, Brazil
Ramana Pidaparti,
University of Georgia, United States

*CORRESPONDENCE

Fuyang Tian
✉ fytian@sdau.edu.cn
Jingjing Chen
✉ joyjchan@gmail.com

†These authors have contributed
equally to this work and share first
authorship

SPECIALTY SECTION

This article was submitted to
Animal Behavior and Welfare,
a section of the journal
Frontiers in Veterinary Science

RECEIVED 08 October 2022

ACCEPTED 13 December 2022

PUBLISHED 06 January 2023

CITATION

Yu Z, Liu L, Jiao H, Chen J, Chen Z,
Song Z, Lin H and Tian F (2023)
Leveraging SOLOv2 model to detect
heat stress of poultry in complex
environments.
Front. Vet. Sci. 9:1062559.
doi: 10.3389/fvets.2022.1062559

COPYRIGHT

© 2023 Yu, Liu, Jiao, Chen, Chen,
Song, Lin and Tian. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Leveraging SOLOv2 model to detect heat stress of poultry in complex environments

Zhenwei Yu^{1,2†}, Li Liu^{3†}, Hongchao Jiao^{2,4}, Jingjing Chen^{5*},
Zheqi Chen³, Zhanhua Song¹, Hai Lin^{2,4} and Fuyang Tian^{1*}

¹College of Mechanical and Electronic Engineering, Shandong Agricultural University, Taian, China, ²Shandong Provincial Key Laboratory of Animal Biotechnology and Disease Control and Prevention, Shandong Agricultural University, Taian, China, ³School of Economics and Social Welfare, Zhejiang Shuren University, Hangzhou, China, ⁴Department of Animal Science and Technology, Shandong Agricultural University, Taian, China, ⁵Department of Digital Urban Governance, Zhejiang University City College, Hangzhou, China

Heat stress is one of the most important environmental stressors facing poultry production. The presence of heat stress will reduce the antioxidant capacity and immunity of poultry, thereby seriously affecting the health and performance of poultry. The paper proposes an improved FPN-DenseNet-SOLO model for poultry heat stress state detection. The model uses Efficient Channel Attention (ECA) and DropBlock regularization to optimize the DenseNet-169 network to enhance the extraction of poultry heat stress features and suppress the extraction of invalid background features. The model takes the SOLOv2 model as the main frame, and uses the optimized DenseNet-169 as the backbone network to integrate the Feature Pyramid Network to detect and segment instances on the semantic branch and mask branch. In the validation phase, the performance of FPN-DenseNet-SOLO was tested with a test set consisting of 12,740 images of poultry heat stress and normal state, and it was compared with commonly used object detection models (Mask R CNN, Faster RCNN and SOLOv2 model). The results showed that when the DenseNet-169 network lacked the ECA module and the DropBlock regularization module, the original model recognition accuracy was 0.884; when the ECA module was introduced, the model's recognition accuracy improved to 0.919. Not only that, the recall, AP0.5, AP0.75 and mean average precision of the FPN-DenseNet-SOLO model on the test set were all higher than other networks. The recall is 0.954, which is 15, 8.8, and 4.2% higher than the recall of Mask R CNN, Faster R CNN and SOLOv2, respectively. Therefore, the study can achieve accurate segmentation of poultry under normal and heat stress conditions, and provide technical support for the precise breeding of poultry.

KEYWORDS

heat stress, image segmentation, poultry, SOLOv2, biotechnology

1. Introduction

With the continuous growth of the world's population, people have higher requirements for agricultural production. At present, under the extensive and intensive production system, there are more than 3.535 billion animals raised under this system in the world, and the annual output of milk and meat reached 798 million tons and 3,029 million tons respectively (1, 2). According to existing data, it is predicted that global meat consumption will increase by 70% until 2050 (3). Intensified farming has obviously become the main production method to ensure sufficient supply of meat, eggs, and milk from all over the world. But from another perspective, there are many stressors in intensified farming, such as restraint, warmth, density, and immunity (4, 5). Compared with natural grazing and free-range breeding, heat stress (HS) is an important aspect of all stressors, which causes greater harm to poultry production activities.

HS refers to a series of non-specific reactions produced by the body under high temperature stimulation that exceeds the upper limit of the isothermal zone (6). Constant temperature animals have their own isothermal zone, within this range, the animal body can rely on its own physical regulation function to maintain the balance of body temperature, so as to obtain the best body metabolism and physiological functions (7, 8). However, when the temperature exceeds the upper limit temperature of the isothermal zone, the animal's heat dissipation capacity is hindered, what will break the body temperature balance and steady state maintained by the body's own physical regulation, leading to the heat accumulation in the animal's body, resulting in increased body temperature and increased metabolism, which will damage the health and productivity of the animal. And if the animal continues to be in a state of heat stress, symptoms of organ failure will appear, which can lead to animal death if it is not found and treated in a timely and effective manner (9).

Compared with other animals, poultry is more sensitive to high temperature environment. Due to the high body temperature of the poultry itself, rapid metabolism and lack of sweat glands, especially in summer (10, 11), when the poultry is in a high temperature and high humidity environment, the heat loss through evaporation and heat dissipation is reduced, which intensifies the harm of HS. Therefore, Under the general trend of intensive and large-scale development of poultry breeding, how to reduce the impact of HS has become one of the key issues in the process of poultry breeding.

HS brings significant negative impact on the immune system of poultry. For example, it can affect the immunity organs and cytokines of the poultry and do harm to humoral immunity mediated by B cells and T cells. On the other hand, Heat stress can transform the kind of specific immunity of poultry from cell-mediated immunity to humoral immunity, which will increase the susceptibility of poultry to pathogens. Heat

stress can also weaken the immune function of poultry against emerging pathogens, so that the number of antibodies after vaccination cannot meet expectations. Therefore, this study designed a poultry heat stress detection model based on machine vision. The model can accurately segment poultry under heat stress from healthy ones, and establish a classifier of heat state. If the system based on this model applied in an actual farm, it can quickly provide heat stress detection results for the breeders, and then provide technical support for the accurate breeding and welfare breeding of poultry.

Traditionally, the temperature-humidity index (THI) is used to determine whether poultry is under HS (12). In general, a THI value of 21 is considered as the threshold for chicken heat stress. However, in the actual breeding process, the temperature and humidity in the poultry house are quite different in space. As an indirect indicator, THI cannot directly reflect the heat stress state of poultry located in different positions of the poultry house. With the rapid development of computer vision and bioacoustic technology, poultry behavior monitoring methods based on images and sounds have been widely studied and applied, thus breaking through the limitations of traditional monitoring methods (13–15). Aydin et al. (16) used 3D images captured by the Kinect depth camera to identify the lying and standing states of broilers, and to detect the number and duration of lying states. According to the negative correlation between lying time and health status of broilers, an indirect method for evaluating the health status of broiler legs was proposed. Pu et al. (17) proposed an automatic convolutional neural network-based method to recognize the chicken behavior within a poultry farm using a Kinect sensor. Compared with the actual results, this test result achieved 99.17% accuracy. In addition, there are also studies on the monitoring of heat stress in poultry. Du et al. (18) concentrated on building a machine learning-based hen vocalization identification algorithm to estimate their thermal comfort status. Using the support vector machine (SVM) algorithm to build a classifier for heat stress state. Data from several studies suggest that the classification performance of the optimal SVM model was 95.1 4.3% (the sensitivity parameter) and 97.6 1.9% (the precision parameter). The typical symptoms of heat stress in chickens are reduced activity and increased drinking time. Therefore, Lin et al. (19) analyzed the changes in exercise and drinking time of chickens under different THI values, and proposed to use time-lapse images and deep learning algorithms to monitor drinking time and coordinates of chickens. Further analysis showed that the detection accuracy rate of this method for chickens was 98.16%, and the tracking accuracy rate for chickens was 98.94%. In conclusion, the analysis of poultry sound and the monitoring of drinking time can accurately determine the heat stress state of poultry. However, these research methods have many shortcomings. For example, it is impossible to accurately locate the position of poultry under heat stress, and long-term

monitoring will affect the normal physiological activities of poultry, which is not suitable for actual production.

In order to make up for the insufficiency of existing research, for the images of the complex breeding background in actual production. Based on the existing research, this study improves the DenseNet-169 network and fuses it with the Feature Pyramid Network (FPN) as the backbone network of the SOLOv2 network. A novel classification model (FPN-DenseNet-SOLO) for identifying heat stress status in poultry is proposed. The main contributions of this study are as follows:

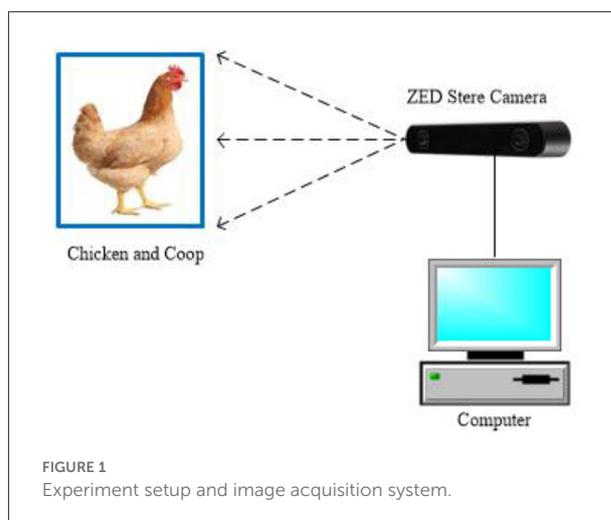
- (1) For complex background images, a deep learning-based recognition and segmentation model of poultry heat stress state is proposed. The model can complete the accurate segmentation of poultry in normal state and heat stress state, and provides a basis for analyzing the classification model of poultry breeding state.
- (2) Optimize the DenseNet-169 network. By introducing Efficient Channel Attention and DropBlock regularization, the extraction of poultry heat stress features is strengthened, and the extraction of invalid background features is suppressed, thereby improving the recognition accuracy and the generalization ability of the network.
- (3) Taking the SOLOv2 model as the main framework, the optimized DenseNet-169 is used as the backbone network to fuse the FPN, and instances are detected and segmented on the semantic branch and the mask branch. In this way, the model can solve the problem of gradient disappearance, have strong anti-fitting ability, improve the accuracy of the model, and provide technical support for the monitoring of poultry heat stress state in actual production.

The rest of the paper is organized as follows: Section 2 introduces and summarizes the related work. Section 3 describes the detailed information of the poultry heat stress data set, PoultryHS. Section 4 describes the proposed novel detection model for poultry heat stress in detail. Section 5 provides the evaluation and analyses of the experiment performance. Finally, Section 6 summarizes this paper.

2. Materials and methods

2.1. Experiment design and image acquisition

Experiments were conducted between September 10–13, 2021 at Animal Experiment Station of College of Animal Science and Technology, Shandong Agricultural University, Tai'an, Shandong, China. Twenty 13-week-old Hy-Line brown laying hens were randomly selected from the farm and reared. These laying hens are housed in environmentally controlled rooms. Each laying hen is kept in a separate cage, which



is connected to each other. Video recording was performed using three ZED 2 cameras to completely record the behavior of the laying hens in the ring control room. It is worth noting that human interference should be avoided during video acquisition. The ZED 2 camera was pre-installed in an environmental control room, the ZED camera was positioned 2.0 m above the ground and precisely at the center of the operation corridor, as shown in Figure 1. The ZED camera was connected to an Intel core i7-11800H CPU, 4.6 GHz, 16 GB physical memory, Microsoft Windows 10 PC installed with the ZED for Windows Software Development Kit (SDK) via a USB port. Additionally, the surveillance camera was used to acquire videos (for observation, labeling, and verification) at 30 fps MOV format. The temperature of the environmental control room was set to 27°C and the relative humidity to 60% at the beginning of the experiment. Raised in this environment for 48 h to ensure that the research subjects fully adapt to the new environment. After 48 h, the data collection of laying hens in normal state was carried out. After data acquisition was completed, the temperature of the environmental control room was set to 36°C and the relative humidity was 75%. When the experimental environment reached the set parameters, the image data of laying hens under heat stress were collected. According to the experimental study by Aydin and Mortensen, normal feeding was performed during the experiment, with free access to food and water, and received 16 h of light per day, the light periods having an approximate light intensity of 30 lx.

2.2. Definition of heat stress behavior of poultry in images

The breeding environment has an important impact on the growth and production of animals. In the face of different stresses, animals usually exhibit different behavioral

characteristics (20, 21). Under high temperature conditions, poultry change their behavioral and physiological balance to thermoregulate and thus lower their body temperatures. And under the influence of heat stress conditions, poultry usually show behaviors such as wing opening, nervous system disorders, short breathing, panting, wing drooping, increased water intake, and reduced food intake (6). As shown in Figure 2A, it can be noticed that the poultry that developed heat stress in this experiment presented behaviors such as wing droop, wing spread, and open mouth panting. These behaviors can be clearly observed in the images compared to the poultry in the normal state (Figure 2B). Behaviors such as wing droop, wing spread, and open mouth panting can increase heat dissipation and thus ensure their own heat stability. This can be seen as a direct behavior of heat stress. At the same time, chickens will drink more water and peck less, which are indirect behavioral effects of heat stress on chickens.

2.3. Image pre-processing

A total of 12,741 clear original images were collected by the image acquisition device, and a data set of the poultry heat stress (PoultryHS) was constructed for the first time. The size of these acquired raw images is $2,976 \times 2,976$ pixels. First, resized the original image to 512×512 pixels to facilitate model training. Then, Labelme software was used to label the images of laying hens to generate a mask map. To draw the bounding boxes, we followed the guidelines of the reference Pascal VOC 2010 Challenge (22). Two classification labels are defined: (1) Normal: laying hens under normal conditions; (2) Heat: laying hens under heat stress. As described above, when the poultry were in a suitable environment (temperature of 27°C and relative humidity of 60%), the collected behaviors belonged to the "Normal" label; when the poultry were in a hot environment (temperature of 36°C and relative humidity of 75%), they showed heat stress behavioral characteristics belonged to the "Heat" label. The accuracy of the study was assessed against manually labeled images. The dataset is randomly shuffled and divided into training, test, and validation sets in a 7:2:1 ratio.

3. Methodology

Driven by artificial intelligence technology, convolutional neural networks (CNNs) have also achieved rapid development in computer vision. Image recognition and classification technology based on CNNs can realize automatic recognition and classification of target objects; by training a large number of data sets, computers can process, analyze and recognize images efficiently and accurately, which greatly improves production efficiency (23, 24). In the intensive and large-scale animal husbandry mode, there are often some abnormal behavior

monitoring problems of livestock and poultry, such as stress response, disease monitoring, behavior recognition, etc. Using CNNs to accurately monitor abnormal behavior of poultry can effectively solve such problems. It is of great significance to improve the economic benefits and management level of the poultry breeding industry.

In order to achieve accurate and rapid detection of the heat stress state of laying hens, this paper proposed to introduce Efficient Channel Attention (ECA) after each convolutional layer on the basis of DenseNet-169 network, and randomly hidden some feature blocks of the research object through DropBlock regularization. In this way, not only the generalization ability of the model can be improved, but also the adaptability of the model to identify poultry in different poses can be enhanced. The improved DenseNet-169 network is used as the SOLOv2 backbone network to fuse the Feature Pyramid Network (FPN) to detect and segment instances on the semantic branch and mask branch. This research will provide technical support for the fine breeding of poultry and animal welfare.

3.1. SOLOv2

SOLOv2 is mainly composed of five parts: fully convolutional networks (FCN), feature pyramid network, mask kernel branch, mask feature branch and semantic branch (Category branch). For the input image, the SOLOv2 network divides it into $S \times S$ grids, and performs feature extraction through the fully convolutional network and the feature pyramid network to determine whether the center of the instance falls into a certain grid. The grids that meet the conditions will enter the semantic branch and the mask branch, and use the corresponding instances to judge the semantics and the size and position of the mask, respectively (25).

The output of the network is divided into two branches: the classification branch and the semantic branch. The structure of the classification branch is $S \times S \times C$, where C is the number of categories. While, the semantic branch structure is $H \times W \times S^2$, where H and W are the high and wide resolution of the semantic output, generally $1/4$ of the original image, and S^2 is the maximum number of instances predicted. One-to-one correspondence with the grid of classification branches in a top-to-bottom, left-to-right manner. According to the principle, the network predicts category and semantic segmentation based on the location at the same time, and combines the results to achieve instance segmentation. Semantic prediction is divided into convolution kernel branch and feature branch. Matrix non-maximum suppression is proposed in SOLOv2. The idea is derived from softening non-maximum suppression, but the computational efficiency is close to that of fast non-maximum suppression. SOLOv2 has a simple structure, and the one-step instance segmentation method adopted has high efficiency and accuracy.



FIGURE 2
Poultry in different states in the dataset. (A) Poultry in heat stress. (B) Poultry in normal state.

3.2. DenseNet

The SOLOv2 network has a complex structure and generates a high number of parameters. In order to reduce the number of parameters and improve the model performance, the DenseNet-169 network is selected as the backbone network in this study. The robust structure of dense block in DenseNet network makes the number of output feature maps of each convolutional layer small. Compared with traditional convolutional networks, it requires fewer parameters to make the feature transfer more efficient, make the network easier to train, make the number of parameters smaller, and make the performance improvement higher. At the same time, the model uses the feature maps of all previous layers as input and its own feature maps as input to all subsequent layers, ensuring maximum information transfer between layers in the network. In this way, all layers can be connected and the gradient disappearance can be effectively mitigated. The DenseNet network combines the advantages of Resnet and Inception, but simply improves the performance of the DenseNet network by combining them, but directly from the perspective of optimal features. By using the two structures of feature reuse and bypass setting, the amount of parameters of the network can be reduced, and the problem of gradient disappearance can be alleviated (26, 27).

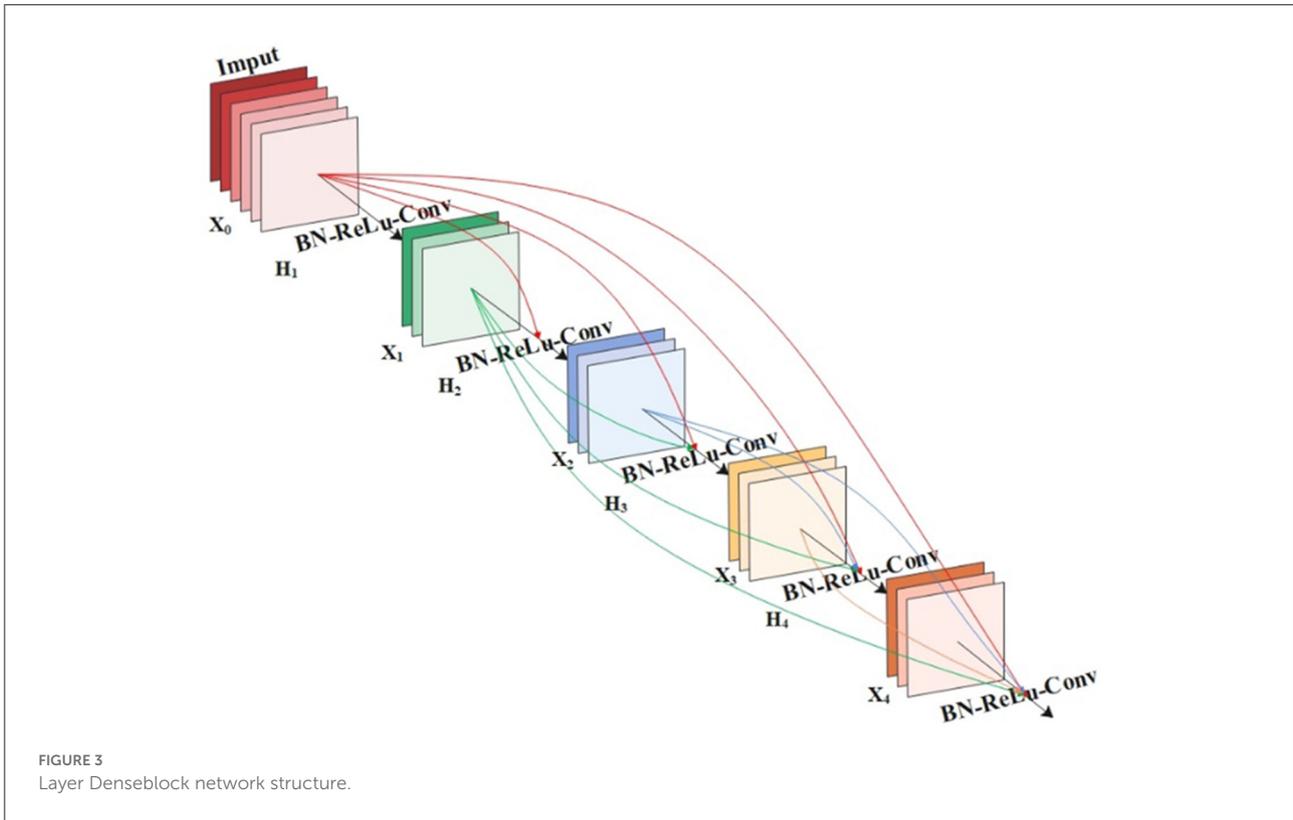
DenseNet is a densely connected convolutional neural network, mainly composed of DenseBlock and Transition Layer. The architecture contains convolutional layers, pooling layers, and densely connected modules with a growth rate of 4, where the growth rate is to keep the channel feature dimension moderate. As shown in Figure 3, DenseNet contains multiple

DenseBlock modules, DenseBlock consists of BN + ReLU + Conv (1×1) + BN + ReLU + Conv (3×3), and the layers between DenseBlocks are called Transition Layers, which consist of BN + Conv (1×1) + Average Pooling (2×2). Since the dimension of the output feature map is too large, an 1×1 convolution is added to the Transition Layers module for dimensionality reduction, thereby improving computational efficiency.

In the DenseNet network, there is an inseparable relationship between any two layers, that is, the input of any layer in the network is the superposition of the output results of all previous layers. The result obtained by this layer will also be transmitted to the next layer as input with the previous output, which is transmitted down in turn. Use the feature outputs of all previous layers as the input of the current layer, that is, X_0, X_1, \dots, X_{l-1} are the feature maps from the first layer to the $l-1$ layer, pass through the l -th layer through the cascade connection, and finally use the composite function H_l (*) to obtain the output X_l .

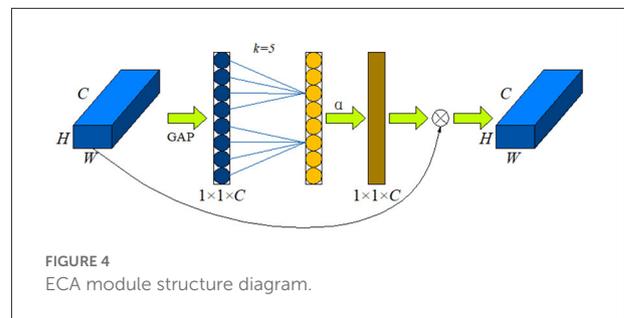
$$X_l = H_l([X_0, X_1, \dots, X_{l-1}]) \quad (1)$$

Where H_i is BN + ReLU + 1×1 Conv + BN + ReLU + 3×3 Conv; BN means Batch Normalization; ReLU is the activation function; Conv is the convolution; $[X_0, X_1, \dots, X_{i-1}]$ means splicing all the feature layers before the i -th layer.



3.3. Efficient Channel Attention (ECA-Net)

Chicken cages, troughs, water troughs and other poultry co-exist within the feeding area are the backgrounds in the actual production environment, which are complex and diverse. Based on the backbone network, the efficiency of extracting behavioral features from images is improved by adding an efficient attention mechanism to increase the weights of important features, thus strengthening the poultry heat stress behavioral features and suppressing the invalid background features. This study adopts ECA-Net, a lightweight attention module that can improve the performance of deep convolutional neural networks. By using the efficient attention module to combine the depth and spatial information of the feature map, the extraction of important features can be effectively suppressed while the extraction of non-important features can be effectively suppressed, so that the accuracy of target recognition in complex environments can be effectively improved (28). Figure 4 is the structure diagram of ECA-Net. C is the number of channels, H is the height of the input data, W is the width of the input data, k is the local interaction size of one-dimensional convolution, GAP is the global average pooling, σ is the sigmoid activation function. The same below. Firstly, global average pooling was performed on the input feature map, and a single value was used to represent the feature layer of each channel; Second,



used a one-dimensional convolution of size k to generate weights for each channel, obtained the interdependence between each channel, and added a sigmoid activation function for normalization; Finally, the weights generated by each channel were multiplied onto the input feature map to enhance the extraction of important features.

ECA-Net uses a one-dimensional convolution of size k to interact across channels to replace the fully connected layer, which can effectively reduce the computational complexity and complexity of the fully connected layer, and then generate weights for each channel.

$$\omega = \delta (CID_k(y)) \tag{2}$$

where ω is the channel weight, δ is the sigmoid activation function, and CID is one-dimensional convolution. The more channels of the input feature map, the greater the k value that requires local interaction, so the k value is proportional to the number of channels C . In this paper, the k value is determined adaptively by a function related to the channel dimension.

$$C = 2^{(\gamma \cdot k - b)} \tag{3}$$

Therefore:

$$k = \left\lceil \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rceil_{\text{odd}} \tag{4}$$

Where $\lceil t \rceil_{\text{odd}}$ is the odd number closest to t , γ , and b is set to 2 and 1, respectively.

3.4. DropBlock regularization

Changes in the surrounding environment may lead to a reduction in the recognition accuracy and the over-fitting of the DenseNet-121 network. In this paper, the DropBlock regularization model was adopted to avoid over-fitting by randomly hiding some feature maps, so as to extract higher robust features.

DropBlock is an improved version of Dropout, which can remove semantic information more effectively than Dropout (29, 30). DropBlock works on the entire space block. Dropout regularization generally works by randomly hiding neurons in fully connected layers (31), however, it is not very effective when used in convolutional layers. The reason is that with the deepening of feature extraction, the feature map gradually becomes smaller, and the receptive field gradually becomes larger. Each feature on the feature map corresponds to a receptive field range, and the corresponding semantic information can be learned through adjacent position elements, and then loses its effect. On the other hand, DropBlock hides the feature map by setting the whole block element, blocks the learning semantic information of adjacent positions, and normalizes the feature map that is not hidden, so as to achieve the regularization effect of the convolution layer.

s and γ are two important parameters in DropBlock: s represents the size of the hidden block. Generally, the network takes 3, 5, and 7, and the effect is best when $s = 7$; γ represents the number of hidden activation units. The relationship between them is shown in Eq. (5).

$$\gamma = \frac{1 - \rho}{s^2} \cdot \frac{f_s^2}{(f_s - s + 1)} \tag{5}$$

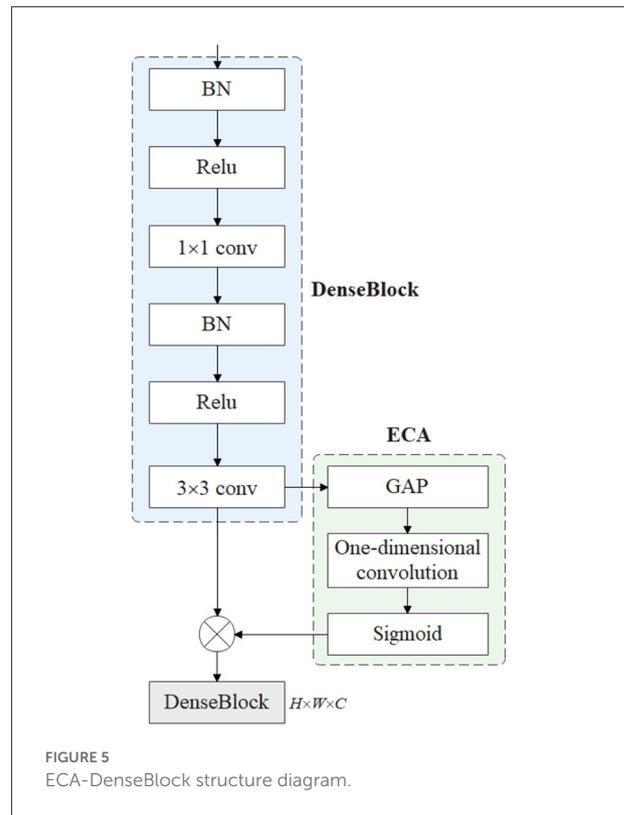


FIGURE 5
ECA-DenseBlock structure diagram.

Where ρ represents the activity threshold probability of the activation unit, and f_s represents the size of the feature map there.

3.5. FPN-DenseNet-SOLO model

The DenseNet-169 network is mainly composed of DenseBlock and Transition Layer. The number and width of output feature maps for each convolutional layer in DenseBlock are lower than the input feature maps. This form of connection makes the transfer of features and gradients more efficient. Figure 5 is a structural diagram of ECA-DenseBlock. DenseBlocks are connected to each other through operations such as BN-ReLU-conv. The ECA-Net module is added to each DenseBlock module to calibrate the weight of each channel on the feature map, increase the weight of important features, and suppress useless features, thereby strengthening the poultry features in heat stress state, while suppressing complex poultry house background features. The ECA-DenseBlock network can effectively speed up the training speed, improve the accuracy of the model, and effectively improve the performance of the network. Since each feature layer contains the output information of the previous feature layer, the information flow is enhanced, so only a few feature maps are needed to complete the network training.

Figure 6 shows the improved DenseNet-169 network structure. First, the input color image goes through a convolutional layer with a 7×7 convolution kernel, which adjusts the number of channels of the image and extracts effective information. Then connect a DropBlock regularization module and an ECA-DenseBlock module in turn. The DropBlock module hides the feature map by setting the whole block element, blocks the learning semantic information of adjacent positions, and normalizes the feature map that is not hidden, so as to realize the regularization of the convolution layer and improve the generalization ability of the model. ECA-DenseBlock is the core part of the improved model. The improved network contains a total of 4 ECA-DenseBlock modules (the number of dense connections is 6, 12, 32, 32). A Transition Layer is connected after each ECA-DenseBlock. Among them, 1×1 conv and average pooling are used to adjust the number of channels to prevent the feature dimension from growing too fast. After the features are extracted by the dense connection structure with the addition of attention mechanism, DropBlock regularization is added to prevent overfitting. Using global average pooling and sigmoid activation function, the generated weights of each channel are weighted onto the feature map by multiplication, and finally the processed feature map is output.

The original SOLOv2 network mainly consists of 5 parts: fully convolutional networks (FCN), feature pyramid network (FPN), mask kernel branch, mask feature branch and category branch. Before the model predicts instance information, it must perform feature extraction through a convolutional network. Compared with FCN, Densenet can fuse the features of each layer of the neural network through dense connections, and the features are reused, which can reduce the amount of calculation and the number of parameters. Its jump structure enables the input feature maps of each layer to be directly connected to the final loss function, accepting the supervision of the final loss function, solving the problem of gradient disappearance, making the network easy to train, and obtaining strong anti-fitting ability.

In Figure 7, H, W, and We are the height, width and number of channels of the output feature map I and mask feature map F of the feature pyramid, respectively. S represents the height or width of the feature map after alignment; C is the total number of semantic categories, which is equivalent to the number of channels of the semantic branch output feature map; D is the convolution kernel weight, which is equivalent to the number of channels of the mask kernel G. The input image is divided into $S \times S$ grids, the total category is C, and the input space is $H \times W \times E$ before being sent to the semantic branch through the feature pyramid, and aligned to $S \times S \times E$. After semantic branch processing (multiple 3×3 convolutions), it is expanded into an output space of $S \times S \times C$. It means that $S \times S \times C$ -dimensional outputs are finally generated, and the semantic category probability is predicted for each grid. Predict the size

and position of the mask in the mask branch, and take the output feature map I of the feature pyramid as the input. Finally, the mask kernel $G \in \mathbb{R}^{S \times S \times D}$ and mask feature map $F \in \mathbb{R}^{H \times W \times E}$ are output after two branches of calculation and learning.

4. Experimental results and analysis

4.1. The setup of experiments

The software environment experimental platform is Ubuntu 18.04 LTS 64-bit system. Python 3.6 is selected for programming language and Pytorch 1.1 is picked up as the deep learning framework equipped with Intel[®] Xeon(R)CPU E5-2683 V3 processor and NVIDIA GeForce GTX 1080Ti GPU, respectively.

The parameters of the CNNs were set as follows: Adam optimizer was used to optimize the model, and the cross entropy was used as the loss function. The batch size was 64, the number of iterations was 100, the initial learning rate was 0.01, the learning rate was reduced to 1/10 of the initial value if the loss value of the validation set does not decrease after 10 epochs. To minimize randomness during training, set random seeds to 100.

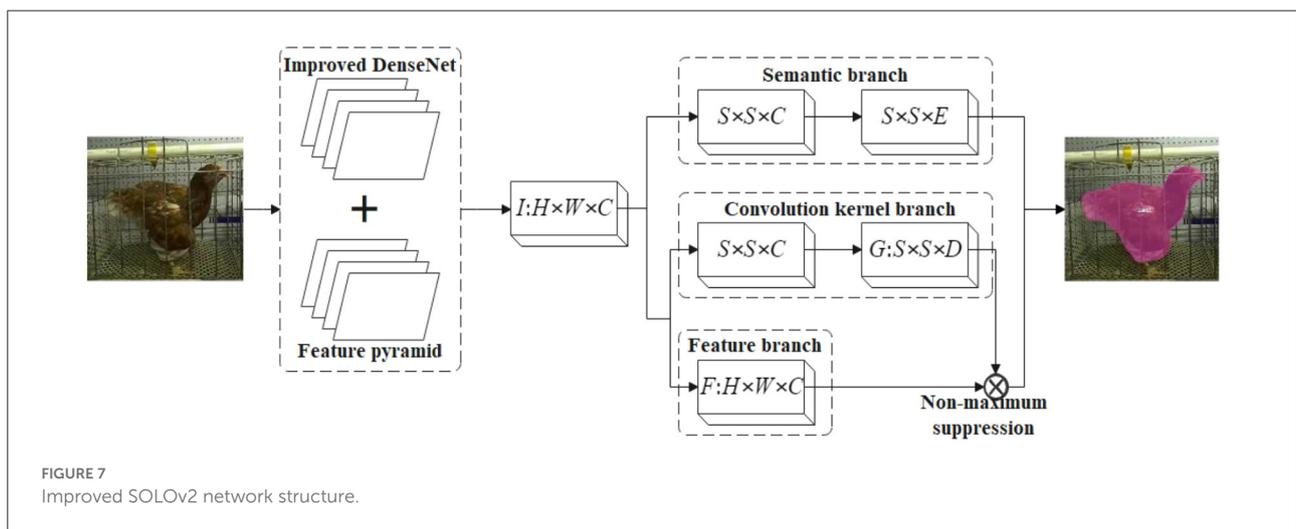
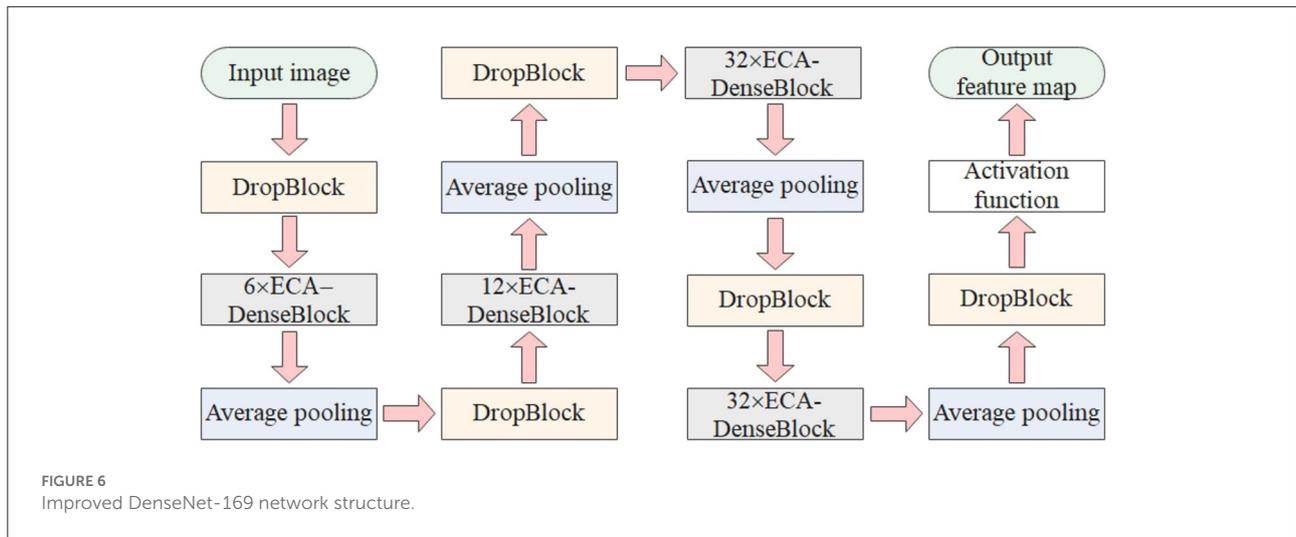
4.2. Loss function

The selection of the loss function is one of the important links in the training process of the deep learning network model. Choosing the correct loss function can provide a better convergence direction and obtain better training results. The current general loss function is the L_2 loss function, and the square of the Euclidean norm of the difference between the minimized prediction result and the true value is used as the convergence direction. The function definition is shown in formula (6):

$$l_2(\hat{y} - y) = \left\| \hat{y} - y \right\|_2^2 \quad (6)$$

From the formula, it can be found that when the difference between the predicted value and the real value is large, the L_2 loss function can rapidly decrease the gradient and improve the convergence speed of the model. However, when the predicted value is close to the real value, the convergence speed of the L_2 loss function will be greatly reduced, and the gradient descent will be slow. Therefore, this paper adopts the BerHu loss function, which can effectively combine the L_1 loss function with the L_2 loss function to obtain better convergence. Its definition is shown in formula (7):

$$B(x) = \begin{cases} |x| & |x| \leq c \\ \frac{x^2 + c^2}{2c} & |x| > c \end{cases} \quad (7)$$



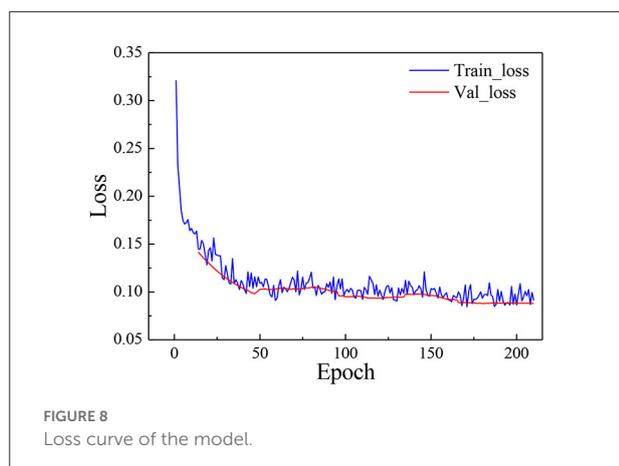
The BerHu function takes c as the limit, and works with the L_2 loss function when it is greater than c to ensure that the gradient decreases rapidly; When it is less than c , it works with the L_1 loss function to ensure that when the predicted value is close to the real value, the gradient of the model can also maintain a certain speed of decline. This paper set $c = k \max_i \left(\left| \hat{y}_i - y \right| \right)$, and set k to different values (0.1, 0.5, and 0.2) for testing. The test results showed that when $k = 0.2$, the model obtained the best output results.

4.3. Model training

In the experiment, the poultry heat stress data set (PoultryHS) was used for training. And the training set, validation set and test set were obtained by dividing them

according to the ratio of 7:2:1. All three used the same classification, which was convenient to objectively measure the model recognition ability. The loss curves of the model on the training set and validation set were shown in Figure 8.

The loss curve can intuitively reflect the dynamic process of training, and can also reflect the convergence of the network through the change of the loss value. Figure 1 showed the variation trend of the loss function with the number of iterations throughout the training process. After the model completed an epoch training, the parameters of the model were adjusted using the validation set. As the epoch increased, the Loss curve of the model gradually decreased and tended to stabilize. Specifically, during the first 9 epochs, the loss value dropped rapidly. Then between 9 and 40 epochs, the rate of decline in the loss value flattened out. Between 40 and 165 epochs, there was a small fluctuation in the loss value. After 165 epochs, the fluctuation of the loss value was further reduced, and the overall value had



become stable, which meant that the network convergence effect was good. At this point the training loss was 0.0871 and the validation loss was 0.0881.

4.4. Detection effect

Since the farm environment in which the dataset was collected was a non-open-air poultry house, external weather factors had little influence on the test results. And the test did not consider the situation of lights out at night, so there is no night-time detection in the recognition results. As can be seen from Figure 9, the FPN-DenseNet-SOLO model had a good detection effect, not only could accurately segment hens from complex environments, but also could accurately identify the posture of hens. According to the different postures and standing angles of hens, the algorithm could show high accuracy and recognition rate, and could also effectively recognize hens with less obvious characteristics.

4.5. Comparison of identification results with classical CNN models

To further analyze the performance of the FPN-DenseNet-SOLO model, the classic CNNs models (Mask R CNN, Faster R CNN, SOLOv2) are compared with the improved models under the same experimental conditions. Use the same training set, validation set and test set to train and test different CNNs models. Table 1 showed the comparison of the classification results of different CNNs models on the test set for poultry heat stress state. In the table, $AP_{0.5}$ and $AP_{0.75}$ represented the AP indicators when the IoU threshold was 0.5 and 0.75, respectively; mAP represented the average of AP corresponding to the increase of the IoU threshold from 0.5 to 0.95 in steps of 0.05.

Each model trained to convergence was applied to the test set for instance detection and segmentation of poultry heat stress, and the final performance evaluation results were obtained as shown in Table 1. From the table, it could be found that the recall, $AP_{0.5}$, $AP_{0.75}$ and mAP of FPN-DenseNet-SOLO model on the test set were higher than other networks. The recall of FPN-DenseNet-SOLO was 0.954, which was 15%, 8.8% and 4.2% higher than the recall of Mask R CNN, Faster R CNN and SOLOv2, respectively. Compared with the original SOLOv2 network, the model proposed in this paper outperformed in various average precision ($AP_{0.5}$, $AP_{0.75}$, and mAP) metrics, 0.978, 0.934, and 0.909, respectively. Especially when the IoU threshold was 0.75, the improvement effect was very obvious, reaching 8.7%. In addition to evaluating the average accuracy of FPN-DenseNet-SOLO, it is also necessary to consider the time spent by the algorithm in the actual segmentation, that is, to reduce the segmentation time while ensuring the accuracy. On the same Graphics Processing Unit (GPU), Mask R CNN, Faster R CNN, SOLOv2, and FPN-DenseNet-SOLO models had an average segmentation time of 0.215, 0.179, 0.078, and 0.092 s to recognize an image, respectively. SOLOv2 and improved SOLOv2 were significantly faster in processing than the other two models, and SOLOv2 model had the shortest split time. FPN-DenseNet-SOLO was only 0.014s slower than SOLOv2, the gap was almost negligible. Considering both average precision and Testing time, the performance of FPN-DenseNet-SOLO is better than SOLOv2. The above results show that the improved method of SOLOv2 network is feasible, and the proposed model can effectively detect the heat stress state of poultry, which highlights the superiority of the model in the identification of heat stress in poultry.

4.6. Ablation experiment

According to the comparative analysis of the identification and classification results of different CNNs models on the heat stress state of poultry on the test set, the FPN-DenseNet-SOLO model has the best performance. Therefore, this paper extracts the features of the improved model and analyzes the influence of different backbone networks on the detection results. Ablation experiments were performed on the proposed model, and the results are shown in Table 2. When the ECA-Net module and DropBlock regularization module were not added to the DenseNet-169 network, the original model recognition accuracy was 0.884. The introduction of the ECA-Net module increased the recognition accuracy of the model to 0.905. This shows that the depth and spatial information of the feature map are combined, which can efficiently extract image feature information, such as texture, edge and color, and suppress the extraction of new questions in complex environments. As the convolutional layers deepen, the visual information in the feature map will be further reduced, increasing the amount

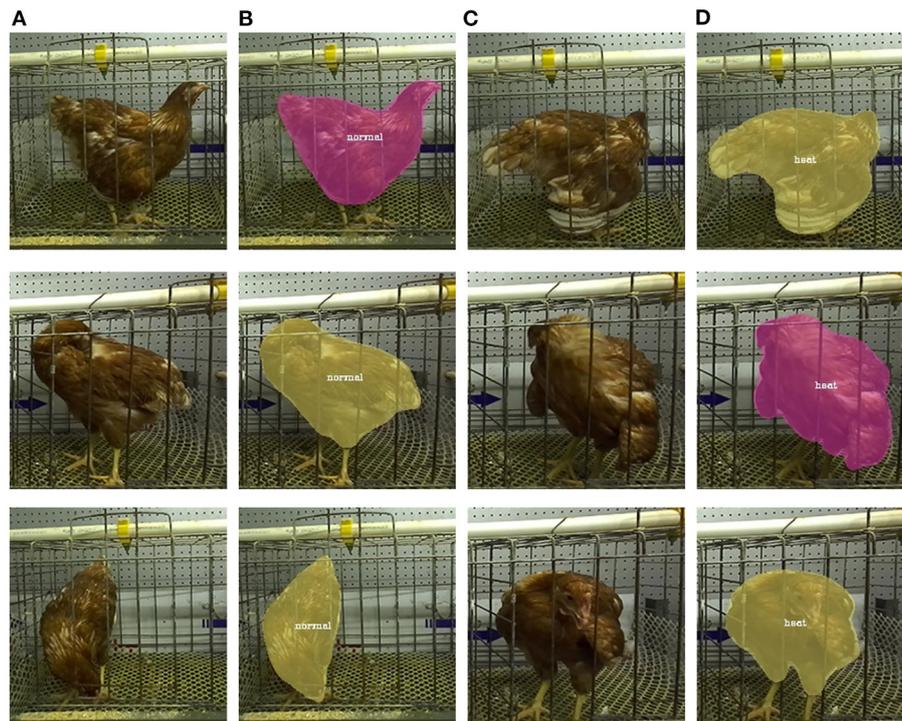


FIGURE 9 Model segmentation effect. (A) Laying hens in normal state; (B) identification results; (C) laying hens in heat stress state; (D) identification results.

TABLE 1 Recognition results of different CNNs models.

Model	Recall	AP _{0.5}	AP _{0.75}	mAP	Testing time (s)
Mask R CNN	0.804	0.828	0.539	0.641	0.215
Faster R CNN	0.866	0.873	0.646	0.713	0.179
SOLOv2	0.912	0.936	0.847	0.792	0.078
FPN-DenseNet- SOLO	0.954	0.978	0.934	0.909	0.092

of abstract information. The addition of the ECA-Net module effectively improves the accuracy of the identification and classification of poultry heat stress in complex poultry house environments. The model recognition accuracy after using the DropBlock regularization module is 91.8%. The function object of this module is the entire space block. By randomly hiding some feature blocks of poultry images, some continuous semantic information on the feature map is deleted. In addition, some independent units can also be randomly discarded, but the deletion of relevant semantic information is not complete, and the corresponding feature information can still be passed to the subsequent network layers. Therefore, the addition of DropBlock regularization can improve the generalization ability of the model. Faced with poultry with different postures and feathers, it can effectively enhance the adaptability of model recognition. After the introduction of ECA-Net and DropBlock

TABLE 2 Results of ablation experiments.

ECA-Net	DropBlock	mAP	Recall
-	-	0.884	0.915
✓	-	0.905	0.932
-	✓	0.897	0.928
✓	✓	0.919	0.954

regularization module improvement, the new model has a recognition accuracy of 0.919 for heat stress in poultry, which is 3.5% higher than the original model. The experimental results show that the addition of ECA-Net and DropBlock regularization module can better extract image features of poultry heat stress.

4.7. Limitations

Compared with the original SOLOv2 model, the detection performance and detection time of FPN-DenseNet-SOLO have been greatly improved. The model inference speed ensures the feasibility of real-time detection, and accurate judgment can be made even when the complex environmental factors of the poultry house are greatly affected.

Nevertheless, as the FPN-DenseNet-SOLO model was applied to edge computing and cloud computing, the size of the model was large, which was not easy to be used in mobile platforms or embedded development platforms. As a result, the application scope of the model was limited and the degree of portability was not high. In future research, we will focus on reducing the size of the model without losing the accuracy of model recognition, so that it can be applied to mobile platforms and embedded development platforms to make the model more widely applicable.

5. Conclusion

To achieve accurate and rapid identification of poultry heat stress state, this paper proposes a poultry heat stress detection algorithm based on FPN-DenseNet-SOLO. The conclusions are as follows.

- (1) In the complex background, the recall, $AP_{0.5}$, $AP_{0.75}$ and mean Average Precision of the FPN-DenseNet-SOLO model on the test set were higher than other networks. The recall of this model was 0.954, which was 15, 8.8, and 4.2% higher than the recall of Mask R CNN, Faster R CNN, and SOLOv2, respectively.
- (2) DenseNet-169 was optimized by introducing ECA and DropBlock regularization, which strengthened the extraction of poultry heat stress features, improved the recognition accuracy and the generalization ability of the network. Taking the SOLOv2 model as the main framework, the optimized DenseNet-169 was used as the backbone network to fuse FPN, and instances were detected and segmented on the semantic branch and mask branch.
- (3) The FPN-DenseNet-SOLO model solved the problem of gradient disappearance, enhanced the anti-fitting ability,

improved the accuracy of the model, and provided technical support for monitoring the heat stress state of poultry in actual production.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The animal study was reviewed and approved by Animal Ethical Review Committee of Shandong Agricultural University.

Author contributions

ZY and FT contributed to the concept and method. LL and JC contributed to the draft and revision of the manuscript. All these authors contributed to the experiments.

Funding

This study was partly supported by Postdoctoral research funding (grant number 040-760083) and National Key Research Program of China (2021YFD1300405).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Pulido M, Barrera-González J, Badgery W, Rodrigo-Comino J, Cerdà A. Sustainable grazing. *Curr Opin Environ Sci Health*. (2018) 5:42–6. doi: 10.1016/j.coesh.2018.04.004
2. Wang J, Shen M, Liu L, Xu Y, Okinda C. Recognition and classification of broiler droppings based on deep convolutional neural network. *J Sensors*. (2019) 2019:1–10. doi: 10.1155/2019/3823515

3. Berckmans D. General introduction to precision livestock farming. *Anim Front.* (2017) 7:6–11. doi: 10.2527/af.2017.0102
4. El-Sabroun K, El-Deek A, Ahmad S, Usman M, Dantas MRT, Souza-Junior JBF. Lighting, density, and dietary strategies to improve poultry behavior, health, and production. *J Anim Behav Biometeorol.* (2022) 10:1–17. doi: 10.31893/jabb.22012
5. Hieu TV, Qui NH, Quyen NTK. Mitigating feather pecking behavior in laying poultry production through tryptophan supplementation. *J Anim Behav Biometeorol.* (2022) 10:1–4. doi: 10.31893/jabb.22011
6. Álvarez EAO, Mejia AMG, Sepúlveda CAG, Saraz JAO. Total heat loss in broilers fed with different lipid sources. *J Anim Behav Biometeorol.* (2022) 10:1–7. doi: 10.31893/jabb.22029
7. Benincasa NC, Sakamoto KS, Silva IJO, Lobos CMV. Animal welfare: impacts of pre-slaughter operations on the current poultry industry. *J Anim Behav Biometeorol.* (2020) 8:104–10. doi: 10.31893/jabb.20014
8. Mascarenhas NMH, Furtado DA, de Souza BB, de Oliveira AG, de Costa ANL, Feitosa JV., et al. Thermal environment characterization of laying hen-housing systems. *J Anim Behav Biometeorol.* (2022) 10:1–9. doi: 10.31893/jabb.22008
9. Pinheiro DG, Filho JADBF, Machado NAF. Effect of wetting method on the broiler transport in Brazilian Northeast. *J Anim Behav Biometeorol.* (2020) 8:168–73. doi: 10.31893/jabb.20022
10. Maciel MP, Carvalho CCS, Aiura FS, Aiura ALO, Maranhão CMA, Silva DB, et al. Use of alternative tiles in the free-range broilers rearing: thermal environment, productive performance and physiological responses. *J Anim Behav Biometeorol.* (2021) 9:1–6. doi: 10.31893/jabb.21008
11. Sugiharto S. Alleviation of heat stress in broiler chicken using turmeric (*Curcuma longa*) - a short review. *J Anim Behav Biometeorol.* (2020) 8:215–22. doi: 10.31893/jabb.20028
12. Perera WNU, Dematawewa CMB. Thermal comfort differences in poultry houses and its influence on growth performance of broiler strains. *Acta Horticult.* (2017) 415–20. doi: 10.17660/ActaHortic.2017.1152.56
13. Del Valle JE, Pereira DF, Mollo Neto M, Filho LRG, Salgado DDA. Unrest index for estimating thermal comfort of poultry birds (*Gallus gallus domesticus*) using computer vision techniques *Biosyst Eng.* (2021) 206:123–34. doi: 10.1016/j.biosystemseng.2021.03.018
14. Okinda C, Lu M, Liu L, Nyalala I, Muneri C, Wang J, et al. A machine vision system for early detection and prediction of sick birds: a broiler chicken model. *Biosyst Eng.* (2019) 188:229–42. doi: 10.1016/j.biosystemseng.2019.09.015
15. Zhuang X, Bi M, Guo J, Wu S, Zhang T. Development of an early warning algorithm to detect sick broilers. *Comput Electron Agric.* (2018) 144:102–13. doi: 10.1016/j.compag.2017.11.032
16. Aydin A. Using 3D vision camera system to automatically assess the level of inactivity in broiler chickens. *Comput Electron Agric.* (2017) 135:4–10. doi: 10.1016/j.compag.2017.01.024
17. Pu H, Lian J, Fan M. Automatic recognition of flock behavior of chickens with convolutional neural network and kinect sensor. *Int J Pattern Recogn Artif Intell.* (2018) 32:1850023. doi: 10.1142/S0218001418500234
18. Du X, Carpentier L, Teng G, Liu M, Wang C, Norton T. Assessment of laying hens' thermal comfort using sound technology. *Sensors.* (2020) 20:473. doi: 10.3390/s20020473
19. Lin C-Y, Hsieh K-W, Tsai Y-C, Kuo Y-F. Automatic monitoring of chicken movement and drinking time using convolutional neural networks. *Trans ASABE.* (2020) 63:2029–38. doi: 10.13031/trans.13607
20. Abdel-Hamid SELS, Fattah AFA. How the diurnal hours affects maintenance behavior in Japanese quail concerning rearing modification systems. *J Anim Behav Biometeorol.* (2020) 8:276–81. doi: 10.31893/jabb.20035
21. Mota-Rojas D, Broom DM, Orihuela A, Velarde A, Napolitano F, Alonso-Spilsbury M. Effects of human-animal relationship on animal productivity and welfare. *J Anim Behav Biometeorol.* (2020) 8:196–205. doi: 10.31893/jabb.20026
22. Everingham M, Van Gool L, Williams CKI, Winn J, Zisserman A. The pascal visual object classes (VOC) challenge. *Int J Comput Vis.* (2009) 88:303–38. doi: 10.1007/s11263-009-0275-4
23. Nyalala I, Okinda C, Kunjie C, Korohou T, Nyalala L, Chao Q. Weight and volume estimation of poultry and products based on computer vision systems: a review. *Poult Sci.* (2021) 100:101072. doi: 10.1016/j.psj.2021.101072
24. Yu Z, Liu Y, Yu S, Wang R, Song Z, Yan Y, et al. Automatic detection method of dairy cow feeding behaviour based on YOLO improved model and edge computing. *Sensors.* (2022) 22:3271. doi: 10.3390/s22093271
25. Koh EJY, Amini E, McLachlan GJ, Beaton N. Utilising convolutional neural networks to perform fast automated modal mineralogy analysis for thin-section optical microscopy. *Miner Eng.* (2021) 173:107230. doi: 10.1016/j.mineng.2021.107230
26. Yan L, Han J, Yue Y, Zhang L, Qian Y. Sentiment analysis of short texts based on parallel DenseNet. *Comput Mater Continua.* (2021) 69:51–65. doi: 10.32604/cmc.2021.016920
27. Zeng M, Xiao N. Effective Combination of DenseNet and BiLSTM for keyword spotting. *IEEE Access.* (2019) 7:10767–75. doi: 10.1109/ACCESS.2019.2891838
28. Liu J, Xu X, Shi Y, Deng C, Shi M. RELAXNet: residual efficient learning and attention expected fusion network for real-time semantic segmentation. *Neurocomputing.* (2022) 474:115–27. doi: 10.1016/j.neucom.2021.12.003
29. Yao J, Xing W, Wang D, Xing J, Wang L. Active dropblock: method to enhance deep model accuracy and robustness. *Neurocomputing.* (2021) 454:189–200. doi: 10.1016/j.neucom.2021.04.101
30. You J, Liu W, Lee J. A DNN-based semantic segmentation for detecting weed and crop. *Comput Electron Agric.* (2020) 178:105750. doi: 10.1016/j.compag.2020.105750
31. Poernomo A, Kang DK. Biased dropout and crossmap dropout: learning towards effective dropout regularization in convolutional neural network. *Neural Netw.* (2018) 104:60–7. doi: 10.1016/j.neunet.2018.03.016