Multimedia Tools and Applications
https://doi.org/10.1007/s11042-018-6784-9

Hatching egg classification based on CNN with channel weighting and joint supervision

Lei Geng [1,2] • Huasong Liu [1,2] • Zhitao Xiao [1,2] • Tingyu Yan [1,2] • Fang Zhang [1,2] • Yuelong Li [3]

Received: 28 June 2018 / Revised: 20 September 2018 / Accepted: 15 October 2018
Published online: 23 October 2018

Abstract

Convolutional neural networks (CNNs) show state-of-the-art performance in tackling a variety of visual tasks. It is expected that a CNN can be applied to the 9-day hatching eggs classification. These hatching eggs are divided into fertile eggs and dead eggs. Because of the inter-class similarity and intra-class difference issues in 9-day hatching eggs datasets, the CNN classification method combining channel weighting (squeeze-and-excitation module) and joint supervision is proposed to improve the classification accuracy. We use the center loss and softmax loss together as a joint supervision signal. With such joint supervision, the CNN can obtain the deep features with inter-class dispersion and intra-class compactness, which enhances the discriminative and generalization powers. Simultaneously, channel weighting is adopted in feature extraction, which is added in each convolutional layer to make better use of the channel features. The experimental results demonstrate that the proposed method successfully solves the classification problem of hatching eggs. The accuracy of our method is 98.8%.

Keywords CNN · Channel weighting · Joint supervision · Hatching eggs · Center loss

Preliminary results of this work were presented at the 3rd International Symposium on Artificial Intelligence and Robotics (ISAIR) 2018, Nanjing, China.

✉ Zhitao Xiao
xiaozhitao@tjpu.edu.cn

[1] School of Electronics and Information Engineering, Tianjin Polytechnic University, NO. 399 Binshui West Street Xiqing District, Tianjin 300387, China

[2] Tianjin Key Laboratory of Optoelectronic Detection Technology and Systems, NO. 399 Binshui West Street Xiqing District, Tianjin 300387, China

[3] School of Computing Engineering, Tianjin Polytechnic University, NO. 399 Binshui West Street Xiqing District, Tianjin 300387, China

1 Introduction

Vaccinations are important means of the prevention and control of avian influenza at present. The vaccine is made by using the avian influenza virus, which is inoculated and cultured in living egg embryos before being inactivated. In the process of virus collection, if mixed with the necrosis of dead embryos, the same batch of collected viruses will be contaminated. Hence, fertility detection and classification of egg embryos before virus collection are crucial. Recently the method for fertility detection and classification of hatching eggs became artificial. However, this method is inefficient and vulnerable to subjective factors. Moreover, it costs a great amount of material resources and manpower. Fertility detection of hatching eggs can usually be divided into four periods: 5-day, 9-day, 14-day and 16-day. Hatching eggs have different features in different hatching periods. In particular, the 9-day dead hatching eggs are very similar to fertile ones in their morphology, which causes difficulties in classification. Fertile eggs at a period later than 9 days are injected with the avian influenza virus to continue culturing, and the dead eggs are recycled to make raw materials for feed. Therefore, it is very significant to improve the state of the art in 9-day hatching eggs classification problems.

Currently, the detection methods for hatching eggs include hyperspectral imaging [9, 28], multi-information fusion [26, 27], machine vision [18, 22], etc. A near infrared hyperspectral imaging system [9, 28] is developed to detect early embryos. The support vector machine (SVM) method [28] is used to establish classification of fertile and infertile egg models, which is based on the image, spectrum and image-spectrum fusion information. The two types of spectral transmission characteristics extracted from the original and Gabor-filtered images are used for K-means clustering in [9]. Xu et al. [26] establish a BP neural network to perform egg embryo condition identification based on the information fusion of image features, which contain areas with egg embryo blood vessels and blackspots extracted from the RGB space, and the standard of each component in the Lab color space, temperature and transmittance. Zhang et al. [27] propose a method that fuses the computer vision technique and the impact excitation technique in which the computer vision model adopts the Learning Vector Quantization (LVQ) artificial neural network. The weighted fuzzy C-means clustering algorithm [18] is utilized to find the threshold to segment the blood vessels of the hatching eggs. Fertility is detected by assessing the quality of the blood vessels. A non-destructive inspection system [22] is designed. It detects and eliminates speckle noise using the minimum univalve segment assimilating nucleus principle, and the nearest neighbor method is applied to classify hatching eggs. Among these methods, machine vision technology shows the best performance. However, the preprocessing is very complicated and the extracted features are not discriminative. Furthermore, the accuracy is not high enough.

With the popularity of CNNs, CNNs have been successfully used to solve many problems (e.g., wound intensity correction and segmentation [14] and, low illumination underwater light field image reconstruction [13]). The CNN method in [13] is far superior to the traditional method [17]. Meanwhile, many general methods [2, 6, 21] have been proposed to improve the performance of CNNs. It is expected that these latest methods can be applied to the classification task of hatching eggs to improve the classification accuracy. The "squeeze-and-excitation" (SE) module [6] adaptively recalibrates channel-wise feature responses by explicitly modeling the interdependencies between channels. Wen et al. [21] proposes a new supervision signal called the center loss. With the joint supervision of the softmax loss and center loss, we could train a robust CNN to obtain the deep features with two key learning objectives, the inter-class dispersion and intra-class compactness, as much as possible. A novel deep convolutional neural network architecture [2] is proposed, where Inception modules have been replaced with depth-wise separable convolutions. Moreover, some emerging methods are expected to solve the hatching egg classification task. These include reinforcement learning [11], brain intelligence [10] and deep adversarial metric learning [23].

In this paper, we propose a new method combining channel weighting (SE module) and joint supervision to improve the accuracy of hatching egg classification in CNNs. SE module improves the representational power of a network by explicitly modelling the interdependencies between the channels of its convolutional features. With the joint supervision, the softmax loss forces the deep features of different classes to stay apart and the center loss efficiently pulls the deep features of the same class to their centers. More importantly, the SE module also has an advantageous effect on the optimization of the center loss. Compared with other similar methods [5, 12, 14, 24, 25], the discriminative power of the deeply learned features can be highly enhanced. Our main contributions are summarized as follows: (i) We propose a new method to obtain more discriminative features. With joint supervision and channel weighting, the highly discriminative features can be obtained for robust hatching egg classification, as supported by our experimental results. (ii) We show that the proposed method is very easy to implement in the CNNs. Our CNN models are trainable and can be directly optimized by the standard SGD.

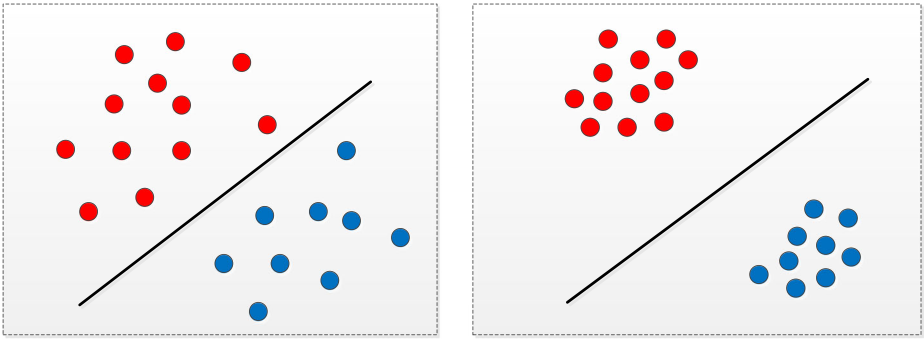
2 Method

2.1 Joint supervision

In the task of CNN classification, the deep features need to be separated so that new data can still be correctly classified without label prediction. The classical softmax loss function has been proved to have good feature separation ability, as shown in Fig. 1a. However, the distance between the intra-classes in many datasets is even larger than the distance between the inter-classes from the perspective of clustering, which is similar to the 9-day hatching eggs dataset. As shown in Fig. 1, in practice, it is expected that features need to be separable, discriminative and generalized. Therefore, some improvement strategies for the loss function are gradually carried out. For example, contrastive loss [19] and triplet loss [16] use the supervised information from the inter-class and intra-class features to train the model. As alternative approaches, they reduce the intra-class features' variations. Contrastive loss can easily make the data between positive and negative samples unbalanced, and triplet loss certainly solves this imbalance problem compared to the previous methods, which require a more complicated sampling strategy. In response to these problems, Wen et al. [6] proposes the center loss function, which can effectively improve the discriminative power of deep features in CNNs. The center loss (E_{center}) function is presented as follows:

$$E_{center} = \frac{1}{2} \sum_{i=1}^{m} \left\| x_i - c_{y_i} \right\|_2^2 \tag{1}$$

In Eq. 1, $x_i \in R^d$ denotes the i-th deep feature, belonging to the y_i-th class. d is the feature dimension. c_{y_i} represents the y_i-th category center of the depth feature. The size of the mini-batch is m.



(a) Separable Features (b) Discriminative Features

Fig. 1 Deeply learned features

Discriminative features can be well-classified by the nearest neighbor (NN) [3] or k-nearest neighbor (k-NN) [4] algorithms, which do not necessarily depend on the label prediction. However, the softmax loss only encourages the separability of features. The softmax loss ($E_{softmax}$) function is presented as follows:

$$E_{Softmax} = -\sum_{i=1}^{m} \log \frac{e^{W_{y_i}^T x_i + b_{y_i}}}{\sum_{j-1}^{n} e^{W_j^T x_i + b_j}} \tag{2}$$

In Eq. 2, $W_j \in R^d$ denotes the j-th column of the weights $W \in R^{d*n}$ in the last fully connected layer and $b \in R^n$ is the bias term. The number of classes is n.

The softmax loss enlarges the inter-class features differences and the center loss efficiently minimizes the intra-class distances of the deep features. Hence, the discriminative power of the deep features can be highly enhanced with the joint supervision. The joint supervision is presented as follows:

$$E = E_{Softmax} + E_{center} = -\sum_{i=1}^{m} \log \frac{e^{W_{y_i}^T x_i + b_{y_i}}}{\sum_{j-1}^{n} e^{W_j^T x_i + b_j}} + \frac{\lambda}{2} \sum_{i=1}^{m} \left\| x_i - c_{y_i} \right\|_2^2 \tag{3}$$

The CNNs are trained under the joint supervision of the softmax loss and center loss, along with a hyper parameter λ to balance the two supervision signals. Joint supervision can be easily implemented in CNNs and optimized by stochastic gradient descent (SGD). If λ is set to 0, it is equivalent to using only the softmax loss function. In Sec. 3.4, we will explain how λ influences the distribution through experiments. Compared with contrastive loss and triplet loss, the center loss and softmax loss do not require construction of complex training samples. The training is more convenient and efficient.

The joint supervision can make deep features more discriminative. Most of the available CNNs only use the softmax loss as a supervision signal, and the resulting deep features would contain large intra-class variations. However, if the CNN is only supervised by the center loss, then the deep features and centers will be degraded to zeros (the center loss is very small). To achieve discriminative feature learning, it is necessary to combine them to supervise the CNN jointly.

2.2 Channel weighting

The performance of a network can be improved by adjusting the network structure or changing the features extraction method. The output of the convolutional layer in the traditional deep networks do not consider the dependence on channels. Channel weighting, that is, the SE module, establishes nonlinear modeling between channels, recalibrates the channel-wise feature responses [1] adaptively, which improves the network's feature learning ability, and can learn to use global information to selectively emphasize informative features and suppress less useful ones through feature recalibration.

As seen in Fig. 2, Original convolution extracts informative features by fusing spatial and channel-wise information together within local receptive fields in Fig. 2a. Channel weighting is based on original convolution, and some other operations are added for feature recalibration in Fig. 2b. The features are recalibrated mainly through the following three steps. (i) Squeeze: Global average pooling is used to compress the features after convolution. This operation is aimed at obtaining the global information of each channel. (ii) Excitation: This is similar to the threshold mechanism in a recurrent neural network. The threshold mechanism sets up two fully connected layers, the first one reduces the input feature dimension to the original 1/R and is then activated by the ReLu [15] activation function, and the second one changes the input feature back to the original dimension. The weight is normalized from 0 to 1 after sigmoid activation. (iii) Reweight: The normalized weight is weighted by the scale operation on the features of each channel.

The SE module is very simple to construct and easy to apply to ordinary deep convolutional networks [8] without introducing new layers or functions. The SE module can be used directly with existing state-of-the-art architectures whose convolutional layers can be strengthened by direct replacement with their SE counterparts. The SE module is computationally lightweight and imposes only slight increases in model complexity and computational burden. Through the modeling of the dependencies between the channels, the SE can adjust every channel feature in

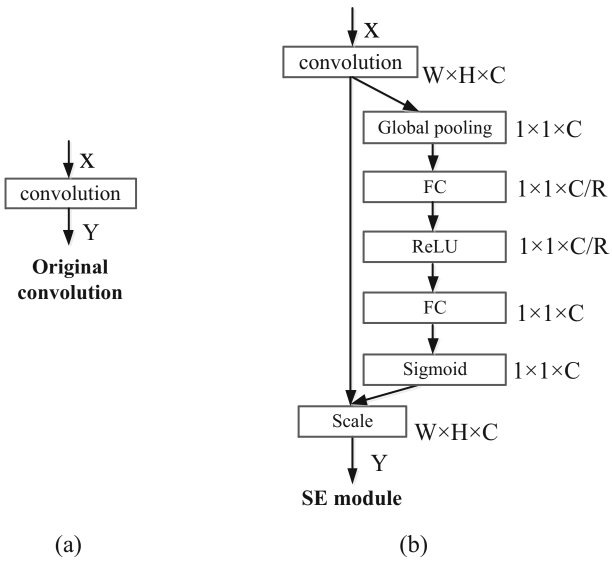


Fig. 2 **a** Original convolution, **b** SE module

a learning manner and show good performance both in terms of model and computational complexity. More importantly, the SE module also has an advantageous effect on the optimization of the center loss.

2.3 Network design

According to the size and characteristics of the dataset, this paper proposes SJ-CNN (the SE module and joint supervision based on a convolution neural network). The basic network structure consists of three convolutional layers and two max pooling layers. The input image is 227×227. In the first convolutional layer, an 11×11 convolution kernel with a stride of 4 is used. The pooling window is 3 with a stride of 2, and the down-sampling of the feature map followed by a max pooling layer. Then, the second convolutional layer includes a 5×5 convolution kernel with a stride of 1 and a padding operation that preserves the size of the input feature map. The next layer is a max pooling layer that uses a pooling window of 3, with a stride of 2. Behind this pooling layer, the network is divided into two branches. One connects to a fully connected layer whose output is 256-dimensional, and the other branch connects to the third convolutional layer, which uses a 3×3 convolution kernel with a stride of 4 and a padding operation. This branch is also followed by a fully connection layer whose output is 256-dimensional after the convolutional layer. Then, the output of the two branches is weighted by the sum and sent to the last fully connected layer. The operation of the branch weighting increases the diversity of feature learning. Meanwhile, it can effectively reduce the occurrence of gradient vanishing. The basic network structural parameters are shown in Table 1.

More importantly, the center loss and SE module are added as supervision signals based on the basic network structure. The SE module is added in each convolutional layer to provide the channel-wise supervision signal. It allows the network to perform feature recalibration, by which it can learn more useful features and enhance the generalization power. Meanwhile, the center loss and softmax loss are used together as the joint supervision signal by fusing the inter-class differences and intra-class compactness, through which the network can learn more discriminative features. The SJ-CNN network structure is shown in Fig. 3 (the operator symbols represent sum operations and the arrows represent the connections between layers).

3 Experiments

3.1 Comparative analysis of network structure

The dataset has a total of 22,000 images, of which 20,000 are used as the training set, 1000 are used as the verification set and the remaining 1000 are used as the testing set. The dataset is composed of fertile egg images and dead egg images classified by an experienced egg

Table 1 The basic network structural parameters

Network structure	Parameters							
SJ-CNN	Conv1 11×11,32	Pool1 Max, 3×3	Conv2 5×5,64	Pool2 Max, 3×3	Conv3 3×3, 128	Fc3–1 256	Fc3–2 256	Fc4 2

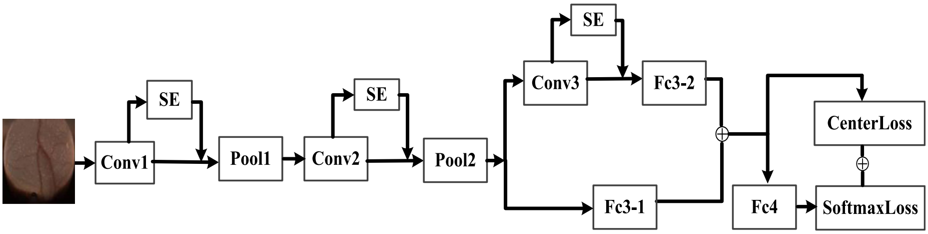


Fig. 3 The SJ-CNN network structure

classification worker. The images of hatching eggs are acquired using a HIKVISION MV-CE013-50GC color industrial camera with an 8-mm HIKVISION MVL-HF0828M-6MP lens. A brightness-adjustable ring light is used to provide the lighting conditions needed for hatching egg image acquisition. The size of the original images is 1280 * 960. The hatching egg images are acquired by an image acquisition and preservation software system, which is based on MFC (Microsoft Foundation Classes). All the experiments below are based on this dataset.

To remove the redundant information, traditional image processing algorithms are used to extract the region of interest from the original images. The hatching egg pre-processed images are shown in Fig. 4.

In Sec. 2.3, SJ-CNN is applied for the 9-day hatching egg dataset based on the basic network structure with an SE module and joint supervision. This network has a strong learning ability. To verify the feasibility of the scheme, different network structures are compared. The batch size is 64. Once the training set trains one epoch, validation is performed to conduct an

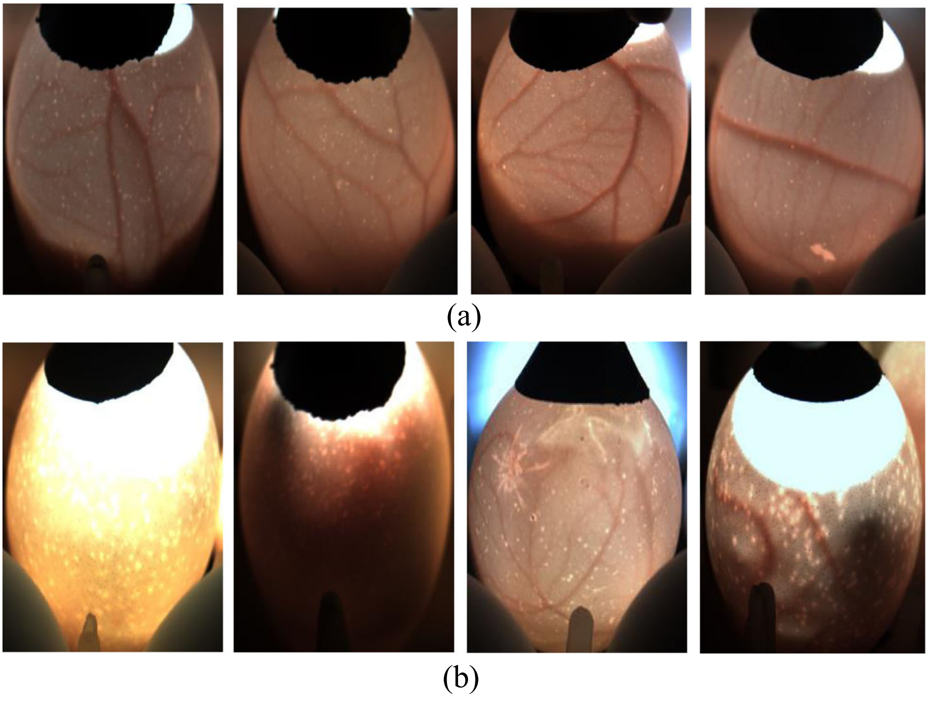


Fig. 4 The 9-day hatching eggs morphology **a** fertile eggs, **b** dead eggs

overall evaluation. This paper compares the first 1000 epochs' accuracy and loss values, as shown in Fig. 5.

As seen from the Fig. 5, the accuracy rate of SJ-CNN increases the fastest and the loss value also drops the fastest in the epoch interval from 0 to 100, followed by the basic network combined with joint supervision and the basic network. The structure using the basic network combined with SE exhibits the slowest. Within the 100 to 200 epoch intervals, SJ-CNN is still the fastest changing structure. However, the convergence speed of the basic network combined with SE gradually exceeds the basic network and begins to show the advantages of its feature selection. Within the 200 to 300 epoch intervals, the other three networks except for SJ-CNN tend to be stable and basically convergent after approximately 300 epochs. However, SJ-CNN has not converged after approximately 300 epochs, as the loss value is still decreasing and the accuracy rate is still rising. It does not converge until 500 epochs. The value of the loss after

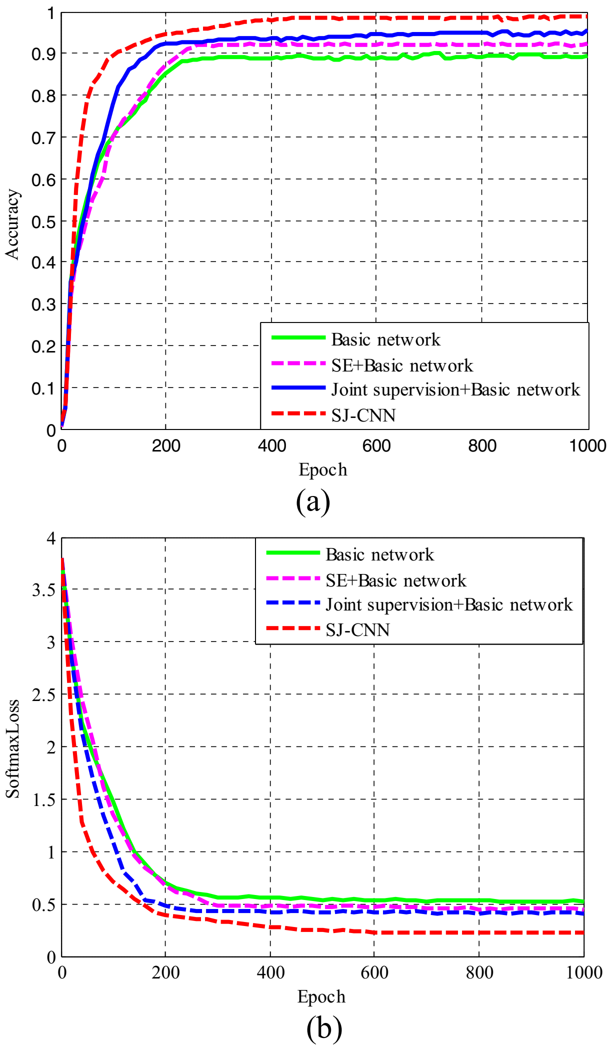


Fig. 5 Experimental analysis graphs of network structure **a** accuracy comparison, **b** loss comparison

convergence is much smaller than those of the other networks, and the accuracy is also higher than those of the other networks. The advantages of channel weighting and joint supervision are highlighted.

On the basis of the basic network, the experiments verify and analyze the necessities of adding the channel weighting module, joint supervision, or combination of the two. First, the basic network and the first two networks are trained and tested to provide a contrast. Finally, the SJ-CNN proposed in this paper is trained and tested. Table 2 shows the accuracy of each network model on the testing set. The accuracy of the basic network is 90.1%. With the addition of the SE module, the accuracy is improved to 92.3%, which shows that the modeling of the channel correlation plays a certain role in improving the accuracy. In the case of joint supervision alone, the accuracy rate increases to 94.6%. Joint supervision has its advantages, learns more discriminative features and improves the accuracy. The proposed SJ-CNN combines the advantages of channel weighting and joint supervision. It not only utilizes the SE module to inhibit the useless features through the strategy of feature recalibration but also utilizes joint supervision to solve the intra-class differences of the dataset, with an accuracy rate of 98.8%. the experiments show that the combination of channel weighting and joint supervision can make deep networks more selective for feature learning. Both of them optimize the network concurrently, which greatly improves the network learning ability and enables learning of more discriminative features so that the accuracy rate will be greatly improved.

3.2 SE module hyperparameter R analysis

The SE module realizes the calibration of the importance of each channel by learning the dependency relationship of each feature channel. To limit the complexity of the model and enhance the generalization ability, two fully connected layers are used in the threshold mechanism in Fig. 2. The first fully connected layer reduces the dimension to 1/R. After the activation of the ReLu function, the dimension is returned to the original one through the second fully connected layer. This section analyzes the selection of the hyperparameter R and compares the accuracy of the training model on the testing set when R is 4, 8, 16, or 32. As shown in Table 3, it can be seen that the training model has the highest accuracy when R is 16. Therefore, 16 is chosen as the multiple of dimension reduction and dimension-enhancement of the fully connected layer.

3.3 Impact of the SE module on center loss

To verify the influence of the SE module on the center loss, SJ-CNN is compared with the basic network combined with joint supervision in Fig. 5.

As seen in Fig. 6, the two initial values of the loss are relatively large, and the loss values are similar in the 0 and 300 epoch intervals. Meanwhile, the loss value of SJ-CNN is slightly

Table 2 Comparison of the accuracy rates of the network structures

Network structure	Accuracy rate
Basic network	90.1%
SE+ Basic network	92.8%
Joint supervision + Basic network	94.6%
SJ-CNN	98.8%

Table 3 Comparison of the accuracy for different hyperparameters R

R	Accuracy
4	90.5%
8	92.0%
16	92.8%
32	90.0%

lower than that of the basic network combined with joint supervision. The center loss value of SJ-CNN near 300 epochs converges faster than that of the joint supervision-based network and the loss value is lower. The feature selection of the SE module highlights its advantages. At approximately 400 epochs, both networks tend to be stable and the loss is essentially convergent. The experiments show that the network obtained after adding the SE module exhibits an advantageous role in optimizing the center loss.

3.4 Adjustment of parameter λ

According to Eq. 2, comparative analysis of the magnification parameters of the center loss is carried out in this paper, as shown in Fig. 6. Within the unequal spacing, we take multiple sets of data from 0 to 0.025; then different networks are trained, and the convergence of the model and the accuracy are tested.

It is very clear that simply using the softmax loss (in this case, λ is 0) is not a good choice since it leads to poor verification performance. Properly choosing the value of λ can improve the verification accuracy of the deeply learned features.

As seen from Fig. 7, when 0 is taken, only the softmax loss is adopted, and the accuracy is the lowest. This is because the deep learning features will include large intra-class changes, which is detrimental to feature learning. As λ increases, the accuracy rate gradually increases, which shows the effectiveness of the center loss. When λ is 0.08, the accuracy reaches its maximum. Then, the accuracy decreases as λ increases, thus indicating that the ratio of the

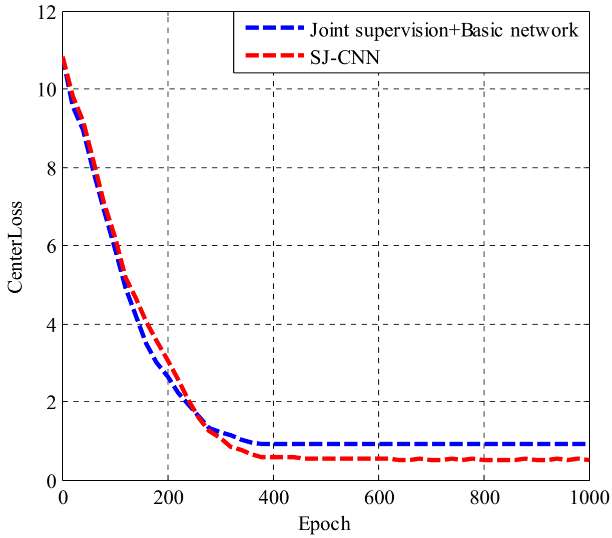


Fig. 6 The influence of SE module on center loss

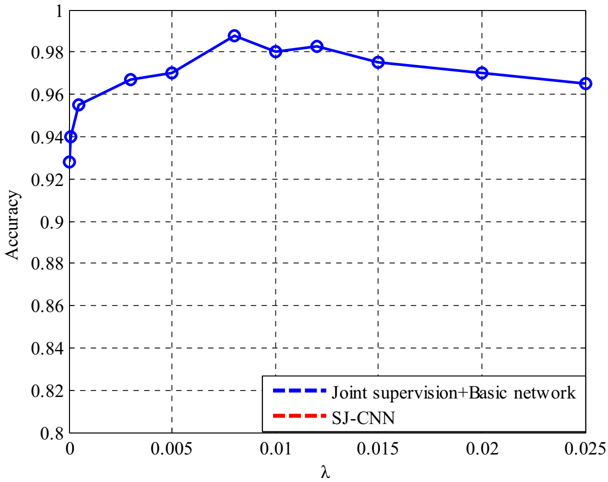


Fig. 7 λ value analysis chart

center loss should not be too large or it will adversely affect the learning ability. Therefore, the correct choice of λ for improving the learning ability of the model is significant. This paper sets λ to 0.08, the value at which the training model is the best.

3.5 Experimental evaluation

To verify the feasibility of the deep learning algorithm proposed in this paper, its experimental results are compared with those of traditional algorithms. The experimental comparison is based on a uniform dataset, which is used in SJ-CNN. The experimental results are shown in Table 4.

As seen from the table, the accuracy of SJ-CNN is higher than those of the other two methods. In the two traditional methods, the extraction of some small blood vessels is still not refined enough, which decrease the final classification accuracy. SJ-CNN extracts more discriminative and generalized features, and the accuracy rate reaches 98.8%.

Furthermore, it is important to compare SJ-CNN with state-of-the-art image classification methods, such as DenseNet [7], SENet [6] and residual attention networks [20]. Table 5 shows the results.

The above results suggest that the SJ-CNN model exhibits high efficiency and good performance. Although these image classification methods have achieved great success in image classification challenges, they are not appropriate for the hatching egg dataset. The image classification challenge dataset is very large and has many classes, but the hatching egg dataset is small and only has blood vessel features. The SJ-CNN model is more appropriate for extracting blood vessel features.

Table 4 Experimental results of different methods

Methods	Feature extraction methods	Accuracy rate
Xu, Q. L et al. (2014)[22]	SUSAN	97.8%
Shan, B et al.(2010)[18]	WFCM	98.1%
SJ-CNN	SJ-CNN	98.7%

Table 5 Comparisons with the state-of-the-art methods

Methods	Accuracy rate
DenseNet-12 [23]	96.8%
SE-BN-Inception[10]	97.1%
SE-ResNet-50 [10]	97.8%
Residual Attention Networks [30]	98.2%
SJ-CNN	98.7%

SJ-CNN still has a 1.3% recognition error. Most of the images recognition error are attenuated hatching eggs in the dataset. These eggs belong to dead eggs. The attenuated hatching egg are almost similar to fertile eggs in morphology. The deeply learned features are mainly blood vessel features. The attenuated eggs also have blood vessels. Therefore, the next work is to improve the recognition accuracy of attenuated hatching eggs. Figure 8 shows fertile eggs and attenuated eggs as follows.

4 Conclusion

In this work, we propose a hatching egg classification method based on a CNN with channel weighting and joint supervision. To enhance the discriminative power of the deep features, the joint supervision of the softmax loss and center loss is used as the supervision signal to train the model. With the joint supervision, the CNN model can obtain the deep features with two key learning objectives, the inter-class dispersion and the intra-class compactness, as much as possible. To improve the representative power of a network, channel weighting is adopted in each convolutional layer. The channel weighting explicitly models the interdependencies between the channels of the convolutional features. Quantitative experiments on the 9-day hatching egg dataset verify that the proposed method is effective in classifying hatching eggs. However, some dead eggs are attenuated eggs and their morphology is similar to that of fertile eggs. However, there are slight differences in the blood vessel features. These embryos easily lead to recognition error. Therefore, the network needs to be further optimized.

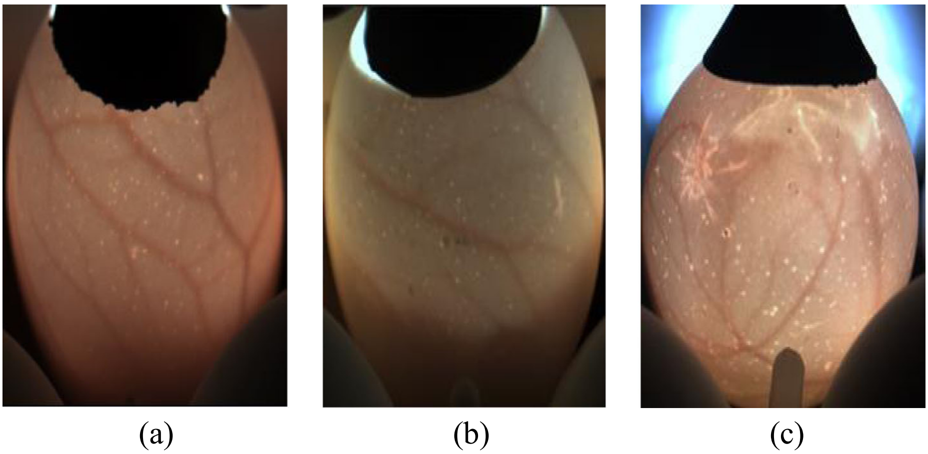


Fig. 8 **a** fertile eggs, **b c** attenuated eggs

Acknowledgements This work was supported by the National Natural Science Foundation of China under grant No. 61771340, Tianjin Science and Technology Major Projects and Engineering under grant No. 17ZXHLSY00040, No. 17ZXSCSY00060 and No. 17ZXSCSY00090, the Plan Program of Tianjin Educational Science and Research under grant No.2017KJ087, and the Tianjin Natural Science Foundation under grant No. 17JCQNJC01400.

References

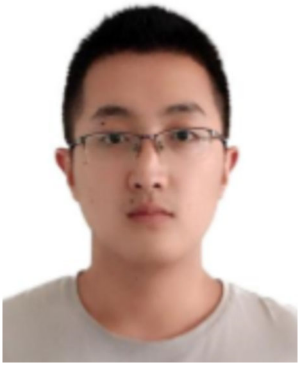
1. Chen L, Zhang H, Xiao J, et al (2016) SCA-CNN: Spatial and Channel-Wise Attention in Convolutional Networks for Image Captioning. pp. 6298–6306
2. Chollet F (2017) Xception: Deep Learning with Depthwise Separable Convolutions. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1800–1807
3. Cover TM, Hart PE (1967) Nearest neighbor pattern classification. Information Theory, IEEE Transactions 13(1):21–27
4. Fukunaga K, Narendra PM (1975) A branch and bound algorithm for computing knearest neighbors. Computers, IEEE Transactions 100(7):750–753
5. Gao L, Guo Z, Zhang H et al (2017) Video captioning with attention-based LSTM and semantic consistency. IEEE Transactions on Multimedia 19(9):2045–2055
6. Hu J, Shen L, Sun G (2017) Squeeze-and-excitation networks. arXiv preprint arXiv:1709.01507
7. Huang G, Liu ZH, et al (2016) Densely connected convolutional networks. arXiv preprint arXiv: 1608.06993
8. Larochelle H, Hinton G (2010) Learning to combine foveal glimpses with a third-order Boltzmann machine. International Conference on Neural Information Processing Systems. Curran Associates Inc., pp. 1243–1251
9. Liu L, Ngadi MO (2013) Detecting fertility and early embryo development of chicken eggs using near-infrared hyperspectral imaging. Food Bioprocess Technol 6(9):2503–2513
10. Lu H, Li Y, Chen M et al (2017) Brain intelligence: go beyond artificial intelligence. Mobile Networks & Applications 2017(7553):1–8
11. Lu H, Li Y, Mu S et al (2017) Motor anomaly detection for unmanned aerial vehicles using reinforcement learning. IEEE Internet Things J 99:1
12. Lu H, Li Y, Uemura T et al (2017) FDCNet: filtering deep convolutional network for marine organism classification. Multimedia Tools & Applications 2017(2):1–14
13. Lu H, Li Y, Uemura T, et al (2018) Low illumination underwater light field images reconstruction using deep convolutional neural networks. Futur Gener Comput Syst
14. Lu H, Li B, Zhu J, et al (2017) Wound intensity correction and segmentation with convolutional neural networks. Concurrency & Computation Practice & Experience, 29(6)
15. Nair V, Hinton GE (2010) Rectified linear units improve restricted Boltzmann machines. International Conference on Machine Learning, Haifa, Israel, 2010: 807–814
16. Schroff F, Kalenichenko D, Philbin J (2015) FaceNet: A Unified Embedding for Face Recognition and Clustering. Proceedings of the 2015 conference on computer vision and pattern recognition, Boston, 815–823
17. Serikawa S, Lu H (2014) Underwater image dehazing using joint trilateral filter. Pergamon Press, Inc 40 (1): 41–50
18. Shan B (2010) Fertility Detection of Middle-stage Hatching Egg in Vaccine Production Using Machine Vision. International Workshop on Education Technology and Computer Science. ETCS pp.95–98
19. Sun Y, Wang X, Tang X (2014) Deep learning face representation by joint identification-verification. Adv Neural Inf Proces Syst 27:1988–1996
20. Wang F, Jiang MQ, et al (2017) Residual attention network for image classification. arXiv preprint arXiv: 1704.06904
21. Wen Y, Zhang K, Li Z, et al (2016) A discriminative feature learning approach for deep face recognition. Computer vision – ECCV 2016. Springer International Publishing, 2016:499–515
22. Xu QL, Cui FY (2014) Non-destructive detection on the fertility of injected SPF eggs in vaccine manufacture. Chinese control and decision conference. pp 1574-1579
23. Xu X, He L, Lu H et al (2018) Deep adversarial metric learning for cross-modal retrieval. World Wide Web-internet & Web Information Systems 2018:1–16

24. Xu X, He L, Shimada A et al (2016) Learning unified binary codes for cross-modal retrieval via latent semantic hashing. Neurocomputing 213:191–203
25. Xu X, Shen F, Yang Y et al (2017) Learning discriminative binary codes for large-scale cross-modal retrieval. IEEE Trans Image Process 26(5):2494–2507
26. Xu Y, Xu A, Xie T et al (2015) Automatic sorting system of egg embryo in biological vaccines production based on multi-information fusion. Transactions of the Chinese Society for Agricultural Machinery 46(2):20–26
27. Zhang W, Tu K, Liu P et al (2012) Early fertility detection of hatching duck egg based on fusion between computer vision and impact excitation. Transactions of the Chinese Society for Agricultural Machinery 43(02):140–145
28. Zhu ZH, Liu T, Xie DJ, Wang QH, Ma MH (2015) Nondestructive detection of infertile hatching eggs based on spectral and imaging information. Int J Agric & Biol Eng 8(4):69–76

Lei Geng Associate professor at the School of Electronics and Information Engineering, Tianjin Polytechnic University. He received his Ph. D. degree from the School of Precision Instrument and Optoelectronics Engineering, Tianjin University in 2012. His research interest covers image processing and pattern recognition, intelligent signal processing technology and system, DSP system research and development.

Huasong Liu Master candidate at the School of Electronics and Information Engineering, Tianjin Polytechnic University. He received his bachelor degree from Tianjin Polytechnic University in 2016. His research interest covers pattern recognition and machine learning.

Zhitao Xiao Professor at the School of Electronics and Information Engineering, Tianjin Polytechnic University. He received his Ph. D. degree from the School of Electronics and Information Engineering, Tianjin University in 2003. His research interest covers intelligent signal processing, image processing and pattern recognition. Corresponding author of this paper.



Tingyu Yan Master candidate at the School of Electronics and Information Engineering, Tianjin Polytechnic University. He received his bachelor degree from Qilu University of Technology in 2015. His research interest covers pattern recognition and machine learning.

Fang Zhang Associate professor at the School of Electronics and Information Engineering, Tianjin Polytechnic University. She received her Ph. D. degree from the School of Precision Instrument and Optoelectronics Engineering, Tianjin University in 2009. Her research interest covers image processing and pattern recognition, and optical interference measurement technique.

Yuelong Li Associate Professor at the School of Computer Science and Software Engineering, Tianjin Polytechnic University. He received his PhD degree in Computer Science from the Peking University in 2012. He is visiting at Professor Edwin Hancock's Lab in the University of York. His research interests focus on face recognition, face alignment, shape modeling, and data embedding. He is a reviewer for many journals in the field, such as IEEE Transactions on Image Processing, IEEE Transactions on Systems, Man and Cybernetics: Systems, IEEE Transactions on Circuits and Systems for Video Technology, Neurocomputing, and so forth.