Blood vessels segmentation of hatching eggs based on FCN with dilated convolution and channel weighting

Lei Geng [1, 2, a], Ling Qiu [1, 2, b], Jun Wu [1, 2, c, *], Zhitao Xiao [1, 2, d]

[1] Tianjin Key Laboratory of Optoelectronic Detection Technology and Systems, Tianjin, 300387, China;

[2] School of Electronics and Information Engineering, Tianjin Polytechnic University, Tianjin, 300387, China.

[a] genglei@tjpu.edu.cn, [b] 2521936474@qq.com, [c] zhenkongwujun@163.com, [d] xiaozhitao@tjpu.edu.cn

Abstract

The fertility detection of the egg embryos is a key step in vaccine preparation. Segmentation of blood vessels of the hatching eggs images is important in the process of quantitative analysis, which is instructive in the fertility detection of the egg embryos. In this paper, a method for the segmentation of blood vessels of the hatching eggs images based on FCN with dilated convolution and channel weighting is proposed. Firstly, CLAHE and Gamma correction are used the green channel of the image to enhance the contrast. Then, in order to adapt to network training, the enhanced images are divided into patches to expand the data. Finally, the dilated convolution instead of the standard convolution to increase the receptive field and the module of the channel weighting is embedded in FCN to supervise the experiment. The results show that the MAE of the proposed method can reach 0.028 and IOU can reach 0.889. The proposed method can obviously improve the segmentation performance.

Keywords

Semantic segmentation, FCN, dilated convolution, channel weighting, hatching eggs.

1. Introduction

Vaccination is the specific implementation of immunization against infectious diseases and an effective means to control and eliminate infectious diseases. Currently, the preparation of influenza vaccines is usually used in the form of incubating viruses in egg embryos. Due to the individual differences of egg embryos and the various situations during inoculation, many egg embryos will die [1]. Therefore, the fertility detection of egg embryos is of great significance.

The fertility detection of hatching embryos is crucial in the preparation of vaccines, but there are many difficulties. Firstly, the diversity of samples seriously affects the judgment of fertility. For the aborted and dead embryos, the characteristics of the embryos are unstable and the blood vessels distribution is irregular. Some weak embryos still have the characteristics of normal embryos, the difference is not significant, resulting in a very difficult to analyze characteristics of egg embryos. Secondly, the methods that use of optical [2-3] and electrical potential [4-7] are susceptible to eggshell thickness and the external environment and they cannot accurately locate its characteristic parameters. Finally, the instability of egg embryos imaging can also seriously affect its characteristic. At present, there are many schemes judge the fertility by extracting the blood vessel of the embryo images [8-9]. In many egg embryos images, there are some tiny blood vessels that are important features of egg embryos. It is difficult to divide these tiny blood vessels with traditional methods. Some small cracks on the eggshell and some stomata will also affect the extraction of features. In addition, although the detection of the fertility of egg embryos by the traditional methods has been greatly improved, a large number of image processing techniques are used and the processing is too complicated [10-11]. It is difficult to meet the requirements of engineering.

In recent years, the deep learning has made a great breakthrough, combining the shallow features to form abstract deep features and discovering the distributed features of the data. Compared with traditional methods, deep learning enables the computer to learn from observational data and solve the problem by itself according to the results of the study. As the branch of the deep learning, fully convolutional neural network (FCN) [12] is proposed based on semantic segmentation. The architecture of the U-net [13] is a semantic segmentation network based on FCN, which is suitable for medical image segmentation. The encoder-decoder structure is adopted in the network. On the one hand, the space dimension of the pooling layer is gradually reduced by encoder; on the other hand, the details and spatial dimensions of the object are gradually restored by the decoder. In addition, there is a skip connection between the encoder and the decoder to help the decoder fix the details of the target better. In this paper, the experiment is tested on the images of hatching eggs. Firstly, contrast limited adaptive histogram equalization (CLAHE) and gamma correction are used the green channel of the image to enhance the contrast. Secondly, the enhanced images are divided into patches to expand the data. Thirdly, on the basis of the structure of the U-net, the dilated convolution instead of the standard convolution to increase the receptive field and the channel weighting is also embedded in FCN for training. Finally, make prediction on the test set with the trained model and verify the feasibility.

2. Algorithm Description

2.1 Dilated Convolution

The operation of the pooling can increase the receptive field and improve the performance of the network. However, the pooling operation will reduce the resolution and some information will be lost in the process of upsampling. The dilated convolution [14] can exponentially increase the receptive field without diminishing the spatial dimension. Dilated convolution introduces the dilation rate based on the standard convolution, which can obtain more dependencies of the pixels in a wider receptive field and save more computation cost than using large convolution kernel. The dilated convolution is shown in Fig. 1.

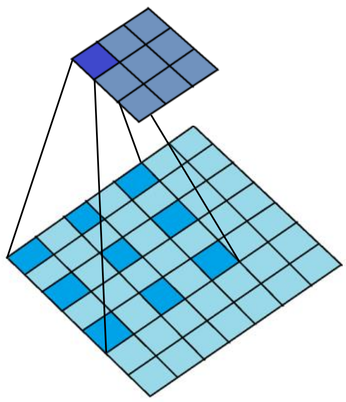

Fig. 1 The schematic of dilated convolution

2.2 Channel Weighting

The block of the squeeze-and-excitation [15] uses the global information of the feature map of the convolution layer to nonlinearly model the dependencies of each channel to enhance the learning ability of network. The module based on the feature dimension explicitly models the interdependence between feature channels. The importance of each feature channel is automatically obtained through learning and the network selectively enhances the useful features according to the degree of importance so that the useful features are fully utilized. The implementation of the SE module mainly includes: (1) The global average pooling is used to compress the characteristics of each channel output by the convolution layer to obtain the information of the global receptive field of the input characteristic channel. (2) A mechanism with a sigmoid function is used to generate a learnable weight for each of the characteristic channel to explicitly model the correlation between the channels. In order to limit the complexity of the model and enhance the generalization ability, the mechanism uses two fully connected layers (FC). The first FC reduces the dimension to 1/8 and follows a rectified linear unit (ReLU).The second FC recovers its dimension. This approach introduces both nonlinearity to better fit of complex correlations between dimensions and also reduces the amount of computation. (3) The weight of the output of sigmoid function is a measure of the importance of each characteristic channel. The weight is weighted to the previous channel characteristics by multiplication and the original feature is recalibrated on the channel dimension. The structure of the SE module is shown in Fig. 2.

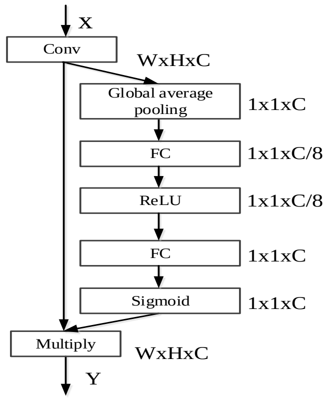

Fig. 2 The structure of the SE module

2.3 FCN with Dilated Convolution and Channel Weighting

As is shown in Fig. 3, the first half of the network consists of the repeated application of two groups of the dilated convolution, the kernel size is 3*3 and dilation rate is 2. Every dilated convolution follows by a ReLU and every group dilated convolution follows by a 2*2 max pooling operation with stride 2. In order to reduce the loss of feature information, double the number of channels after each pooling. The latter half of the network consists of the repeated application of two groups of the transposed convolution and the kernel size is 3*3. The features of the shallow layers of the network contain many details and the deep features are more abstract. The characteristics of the shallow layer were weighted by SE module and the channel characteristics were recalibrated by learning. In order to produce more accurate results, skip connection is introduced between the results of the channel weighting and high level features. Each connection follows by two dilated convolution, half the number of channels after each upsampling. The kernel size of the dilated convolution is 3*3 and dilation rate is 2. Every dilated convolution follows by a ReLU. At the final layer a 1*1 convolution is used to map each 32 component feature vector to 2 classes.

3. Experimental Results and Analysis

3.1 Image Preprocess

3.1.1 Image Enhancement

The algorithm of the image enhancement aims to improve image quality and make it clearer in content. In order to accelerate the speed of network training, the green channel with high contrast of the image is selected for processing in this paper. Due to the uneven illumination and the movement of in the photographing process, the image quality is not good. In order to make the network converge as soon as possible, the green channel image is normalized first. And then, the normalized images are processed by CLAHE, which improves the contrast and clarity of blood vessels in hatching eggs images. Finally, the results of CLAHE were corrected by Gamma and the dynamic range of the image was enhanced to achieve contrast stretch. The results of the image enhancement are shown in Fig. 4.

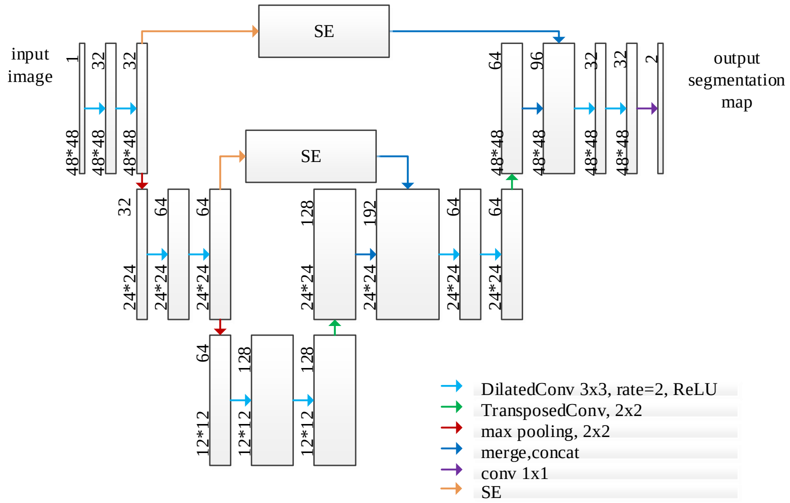

Fig. 3 The network architecture of the proposed method

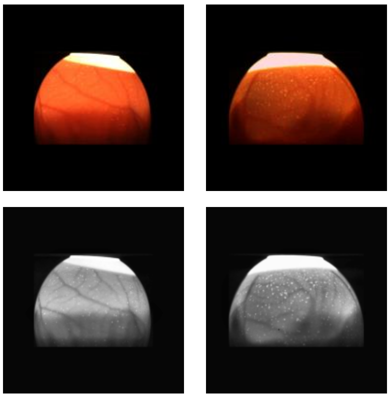

Fig. 4 The preprocess results (First row: images of original. Last row: images of preprocess.)

3.1.2 Data Augmentation

The scale of data has a great impact on the performance of network. Fewer images, time-consuming annotation and the width of blood vessel in images varies from one pixel to a dozen pixels. Therefore, the training of the network is based on the image patches in this paper. The 100 images of the training set are taken as samples. 48*48 image patches are extracted from each image after enhancement, and 1800 image patches are extracted for each image. Accordingly, the same patches extraction operation is performed in the groundtruth. The patches are shown in Fig. 5.

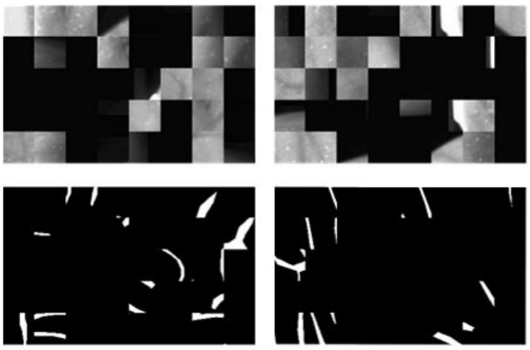

Fig. 5 The patches of the images (First row: the patches of preprocess images. Last row: the patches of groundtruth.)

3.2 Results of Segmentation

3.2.1 Evaluation Parameters

To evaluate the performance of the proposed algorithm, the intersection-over-union (IOU) is used for comparative analysis. As given in formula (1), IOU refers to the ratio between intersection of two sets and union. And the two sets are the GroundTruth and predicted segmentation respectively.

$$IOU = \frac{predicted\ segmentation \cap GroundTruth}{predicted\ segmentation \cup GroundTruth} \tag{1}$$

It can be seen that the larger the value of IOU and the better the segmentation method.

IOU only evaluates the accuracy of segmentation objective, but it cannot illustrate the suppression of background noises. Therefore, the MAE is used to evaluate the accuracy of the whole image. MAE calculates the average difference between the predicted segmentation and the groundtruth in pixel level, and it is defined by formula (2) as follows:

$$MAE = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} \left| S(\mathrm{i},\mathrm{j}) - GT(\mathrm{i},\mathrm{j}) \right| \tag{2}$$

where M and N denote the height and width of the image respectively. $S(\mathrm{i},\mathrm{j})$ is the pixel value of the predicted segmentation. $GT(\mathrm{i},\mathrm{j})$ is the pixel value of the groundtruth. Obviously, the smaller the MAE value is, the closer the predicted segmentation to the groundtruth.

3.2.2 Comparison of Different Methods

Due to the lack of public dataset, the methods of the reference [8] and reference [9] are experimented with 9-day embryo images of acquisition.The comparison results of IOU and MAE are shown in the Table 1 and Table 2 respectively. It can be seen that the proposed algorithm obtains the biggest IOU and the smallest MAE, which proves that the predicted segmentation of the proposed algorithm can achieve the best performance. The main reason is that the dilated convolution can obtain more dependencies of the pixels in a wider receptive field in the proposed algorithm and the module of the channel weighting can nonlinearly model the dependencies of each channel to enhance the learning ability of network.

Table 1. The comparison of the IOU

method	IOU
Shan et al. [8]	0.865
Xu et al. [9]	0.877
Our method	**0.889**

Table 2. The comparison of the MAE

method	MAE
Shan et al. [8]	0.214
Xu et al. [9]	0.115
Our method	**0.028**

Fig. 6 shows visual comparisons of different segmentation results. It can be seen that the method of the Shan is susceptible to image noise, which leads to the serious breakage of blood vessels and is not conducive to the fertility judgment of embryos. The method of the Xu mainly extracts the main blood vessels in the image and ignores the effects of small blood vessels. The proposed method is based on supervised learning and does not require high quality images. The dilated convolution is used to consider the spatial relationship of pixels and the feature extraction process takes into account the dependence of the characteristic channels to enhance the learning ability of the network. It can be seen that the proposed method is better than traditional methods, which is more conducive to the fertility judgment subsequently.

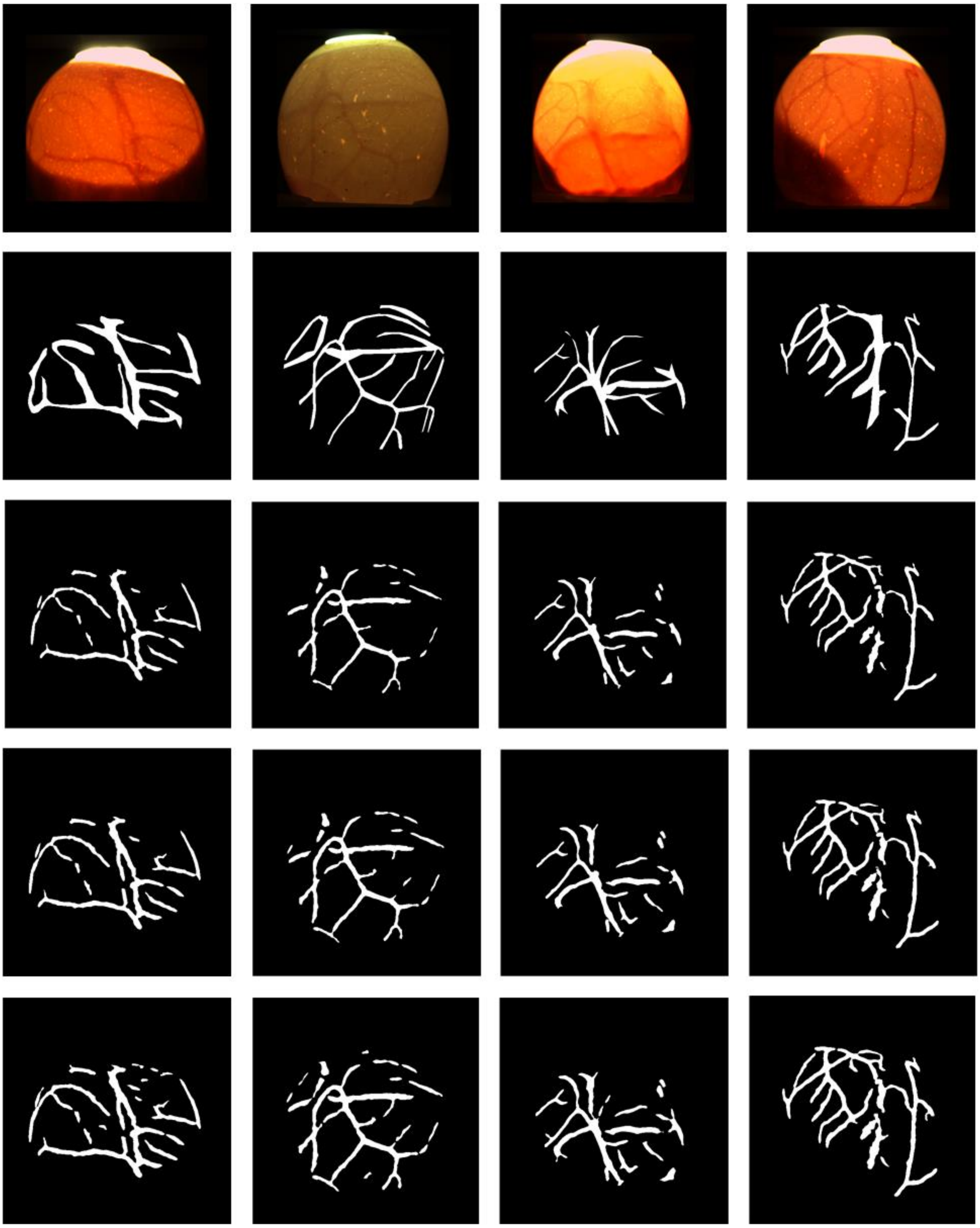

Fig. 6 The segmentation results different methods (The first row to the fifth row are original images, the groundtruth, the results of the method of the Shan, the results of the method of the Xu, the results of the proposed method)

4. Conclusion

This paper proposes a method of blood vessels segmentation of the hatching eggs images based on FCN with dilated convolution and channel weighting. Firstly, some images of the training database are preprocess to enhance contrast. Secondly, the pretreated images are expanded to fit the network training. Thirdly, using the convolution of dilated convolution instead of the standard convolution method. Finally, taking into account the degree of interdependence between the feature channels, we introduce the channel weighting module, which is embedded in the FCN for training. Experimental results show that the network using dilated convolution and channel weighting has better feature resolution and better segmentation performance.

Acknowledgements

This work was supported by Basic Application Research Project of China National Textile and Apparel Council under grant No. J201509, Tianjin Natural Science Foundation under grant No.17JCQNJC01400, Tianjin Science and Technology Major Projects and Engineering under grant 17ZXSCSY00090.

References

[1] Stephenson J. Bird Flu Vaccines [J]. Journal of the American Medical Association, 2007, 297 (11): 1182-1182.

[2] Smith D P, Mauldin J M, Lawrence K C, et al. Detection of fertility and early development of hatching eggs with hyperspectral imaging[C]. XIth European Symposium on the Quality of Eggs and Egg Products, Doorwerth, The Netherlands, 2005: 176-180.

[3] Liu L, Ngadi M O. Detecting Fertility and Early Embryo Development of Chicken Eggs Using Near-Infrared Hyperspectral Imaging [J]. Food and Bioprocess Technology, 2013, 6(9): 2503-2513.

[4] Romanoff A L, Cottrell C L. Fertility Study of Fresh Eggs by Radio Frequency Conductivity and Dielectric Effect [J]. Experimental Biology and Medicine, 1939, 42(1): 298-301.

[5] Romanoff A L, Frank K. High Frequency Conductivity and Dielectric Effect of Fresh Fertile and Infertile Hens'Eggs [J]. Experimental Biology and Medicine, 1941, 47(2): 527-530.

[6] Romanoff Alexis L. Detection of Fertility in Fresh Eggs [J]. World's Poultry Science Journal, 1947, 3(1): 12-14.

[7] Romanoff A L. Evaluation of future perish-ability of intact fresh eggs by radio-frequency conductivity [J]. Journal of Food Science, 1949, 14(4): 310-313.

[8] Shan B. Fertility Detection of Middle-stage Hatching Egg in Vaccine Production Using Machine Vision[C]. Proceedings of the Second International Workshop on Education Technology and Computer Science, Wuhan, China, 2010: 95-98.

[9] Xu Q L, Cui F Y. Non-destructive detection on the fertility of injected SPF eggs in vaccine manufacture[C]. Proceedings of the 26th Chinese Control and Decision Conference, Changsha, PEOPLES R CHINA, 2014: 1574-1579.

[10] Schellpfeffer M A, Kolesari G L. Microbubble Contrast Imaging of the Cardiovascular System of the Chick Embyro [J]. Ultrasound in Medicine and Biology, 2012, 38(3): 504-510.

[11] Coucke P M, Room G M, Decuypere E M, et al. Monitoring Embryo Development in Chicken Eggs using Acoustic Resonance Analysis [J]. Biotechnology Progress, 2010, 13(4): 474-478.

[12] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation [C]. Proceedings of the 2015 Conference on Computer Vision and Pattern Recognition, Boston, MA, 2015: 3431-3440.

[13] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation [J]. 2015, 9351:234-241.

[14] Yu F, Koltun V. Multi-Scale Context Aggregation by Dilated Convolutions[C]. arXiv: Computer Vision and Pattern Recognition, Boston, MA, 2015: 1-13.

[15] Hu J, Shen L, Sun G. Squeeze-and-Excitation Networks [J]. 2017.