Segmentation of Right Ventricular MR Image Based on Deep Neural Network: Dilated DenseNet of Two Level Losses

Zhao,Xingrong[1], Liu,Zheng[1,2], Zhao,Heming[1], Chen,Xueqin[1*]
[1]School of Electronic and Information Engineering, Soochow University, China
[2]Suzhou Industrial Park Institute of Services Outsourcing, China
Email:1339974037@qq.com，chenxueqin@suda.edu.cn
*Corresponding author

Abstract—**Right ventricular function assessment plays a very important role in humans' health assessment. It's hard to segment right ventricle accurately and quickly because it's an irregular crescent shape and it changes a lot from one to others, especially diseased ones. In different kinds of deep convolutional neural network(DCNN) algorithms，Dilated-DenseNet is widely used because of good features processing method, but it still has the problem of overfitting in image segmentation. So in this paper, a improved Dilated-DenseNet method is proposed to segment right ventricular MR image. Comparing to some original algorithms, the result is much more precise and stable. The dice metric of train set is 0.91 and test set is 0.90,which shows that overfitting problem has been alleviatd.**

Keywords—right ventricular segmentation; convolutional neural network; Dilated-DenseNet;dice metric

I. INTRODUCTION

Cardiovascular disease is the main threat to human health and the Incidence and mortality is increasing year by year. Magnetic Resonance Image of heart is now the most popular technology to judge the condition of human's Cardiovascular system because it make no damage to people. The right ventricular is an important part to help physicians make right judgment. But the segmentation of right ventricular is a big challenge for physicians. So how to segment right ventricular precisely and quickly is the key to solve problems.

In recent years, people have proposed a variety of ventricular segmentation algorithms mainly divided into two categories: traditional algorithms with building mathematical models and deep learning algorithms[1], in particular convolutional neural network. In 2008, Somkantha proposed a novel edge following technique to segment left ventricular and the result is more accurate than the classical contour models[2]. In 2012, Elbaz use a new Active Shape Model to segment right ventricular. It's more accurate than traditional ASM[3]. In 2013 Punithakumar segment right ventricular via point correspondence and the dice metric in test set is 0.882[4]. In 2014, Zhang applied multi-atlas selection to right ventricular segmentation. In their experiments, the system performs bad in apical slices of the heart, which decreases overall performs and the dice metric is about 0.71[5]. In 2016, Bai put forward a semi-supervised learning algorithm based on neural network to solve the problem of the shortage of MR images' labels. The dice metric is 0.89 on their trained system[6]. In 2017, Molaei proposed a deep neural networks to segment left ventricle and receive a good result[7].

Most papers' experiments in recent years are based on the data set released by Project RVSC. It includes sixteen patients' MRI image of heart and the labels labeled by professional experts. The rest of thirty-two patients have only released the MRI image without labels to the public. So, the first challenge for us in right ventricular segmentation is that we don't have enough training set to train a robust system. The second challenge is that it's difficult to segment the apical slice images precisely because the shape of right ventricular have changed a lot and the region is too small to segment rightly.Most traditional algorithms with building mathematical models can't segment apical images and many deep learning algorithms based on neural networks will go overfitting easily without enough training sets. Aiming at the problems above, segmentation of right ventricular MR image based on dilated densenet of two level losses is proposed. This network can perform robustly and find the inner relationship between training set and labels within a relatively small data set. The experiment results show that the proposed right ventricular segmentation method is very effective for apical slices segmentation.

II. PROPOSED METHOD

The right ventricular segmentation system mainly includes three modules: (1) the data preprocessing module; (2) the dilated densenet training module; (3) the dilated densenet testing module. The system structure is shown in Fig.3.

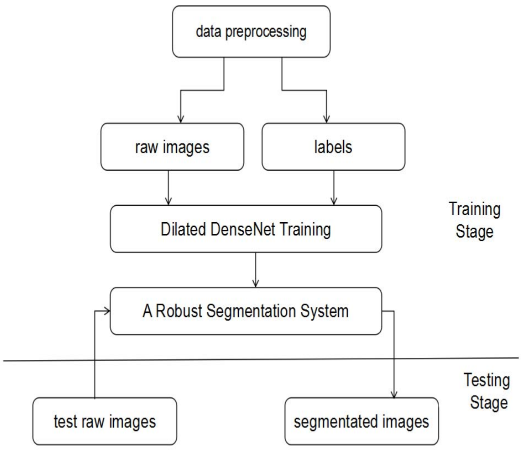


Figure.1 Right ventricular segmentation system

The first step is to do data preprocessing including image augmentation images' reshape. The training set will be shaped as a four-dimensional tensor including batch_size, image_height, image_width and channels, because a platform called Keras is used in this experiment. For training convenience, the batch_size is set for 8 in this paper. It means the system will be trained from batch to batch and each batch has 8 images.

In the training stage, each patch of 8 images will be put into the network to generate an image which classify pixels into two classes:right ventricular pixel and background pixel. Then it should be compared with the label given to get an error. Error Back Propagation is used to minimize the generated error to train the deep neural network and help the network learn the non-linear relationship between raw images and labels[8]. After 3000 iterations of all training batches, a robust right ventricular segmentation system was born.

In the testing stage, test raw images just need to be put into the already trained system. Then the system will give the segmented images of right ventricular automatically.

A. Data Preprocessing

The data set have 243 physician-segmented images and 3697 additional unlabeled images. Since supervise deep neural network is used in this paper, we can only use the 243 physician-segmented images.218 images are used as training set and the rest 25 images are used as test set to value the system's performs. It's necessary to do image enhancement. The standard procedure is to apply affine transformations to the data: random rotations, translations, zooms, shears and elastic deformations[9]. The result is shown in Fig.1, one MR image can be rotated randomly with elastic deformations.

The goal of data augmentation is to prevent overfitting, which is a normal problem existed in supervised deep neural network.

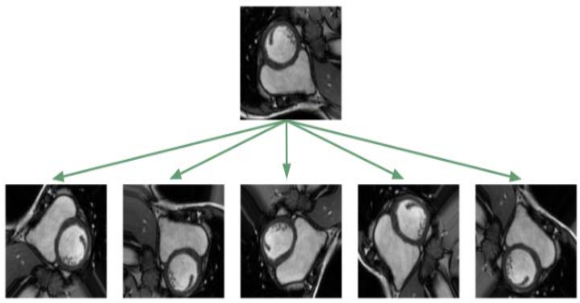
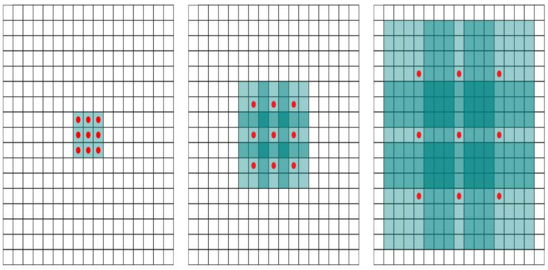


Figure.2 image augment

B. Dilated densenet

Dilated densenet is the combination of dilated convolution and densenet. In my experiment, densenet without dilated convolution usually classify some pixels far from right ventricular incorrectly which obviously are background pixels. After dilated convolution introduced, the probability of this phenomenon decreased greatly.

Dilated convolution has a dilation rate parameter, which is different from ordinary convolution. As is shown in Figure.2, the red dots represent convolution kernel and the green background represent receptive field. It's obviously that dilated convolution increases the receptive field of the convolution kernel while keeping the number of parameters unchanged. At the same time, it can ensure that the size of the feature map of the output remains unchanged[10]. So dilated convolution can help the model classify the pixels which is far from right ventricular as background pixels correctly because of bigger receptive field.



(a) (b) (c)

Figure.3 (a)kernel is 3*3;dilation rate is 1;receptive field is 3*3.(b)kernel is 3*3;dilation rate is 2;receptive field is 5*5.(c)kernel is 3*3;dilation rate is 4;receptive field is 7*7.

Densenet is proposed by Gao Huang and Zhuang Liu in 2016. The structure of densenet can be summarized as the formula(1) where $[x_0, x_1, ..., x_{l-1}]$ refers to the concatenation of the feature-maps produced in layers $0,...l-1$. $H_l(\cdot)$ is defined as a composite function of three consecutive operations: batch normalization(BN), followed by a rectified linear unit(ReLU) and a 3*3 convolution(Conv)[11].

$$x_l = H_l([x_0, x_1, ..., x_{l-1}]) \quad (1)$$

C. Proposed Two level losses

Two level losses are proposed in this paper and their repressions are shown as follows:

$$L1 = -y_T \log(y_P) - (1 - y_T)\log(1 - y_P) \quad (2)$$

where y_T is the standard segmentation of right ventricular given by medical specialists, y_P is the result of prediction given by the already trained model. This formula called log likelihood loss function is used because pixels are divided as two classes: right ventricular pixel and background pixel.

$$L2 = 1 - \frac{2 * X \cap Y}{X + Y} \quad (3)$$

Where X is the region of predict result given by training model, Y is the target region of right ventricular given by professional doctors.

Two level losses strategy trains the neural network alternately with pixel loss and shape loss and initialize the other network alternately. This strategy can effectively prevent the neural network form reaching local optima. Two levels of losses thought is learn from Generative Adversarial Net(GANs). GANs has two networks includes a generative models and a discriminative model. Generative model generate data approaching the true data and discriminative model judge the authenticity of the data from generative model. The two losses of GANs are confrontational. In this paper, the two losses start from different angles: pixel level and shape level. But they have the same purpose: segment the right ventricular. Then the model can easily cross the local optima to reach the global optima and prevent overfitting.

III. EXPERIMENT AND ANALYSIS

We processed 243 MR images from 16patients, contained in the database provided by Project RVSC. Each patient's data consisted of 11-18 slices, result in 243 images. In the experiment, training set of 243 images has separated to 218 images for training and 25 images for testing.

As we can see form figure 4, the left column is the raw images, the middle column is the prediction of right ventricular segmentation given by the trained model in this experiment and the right column is the segmentation of right ventricular given by professional physicians.

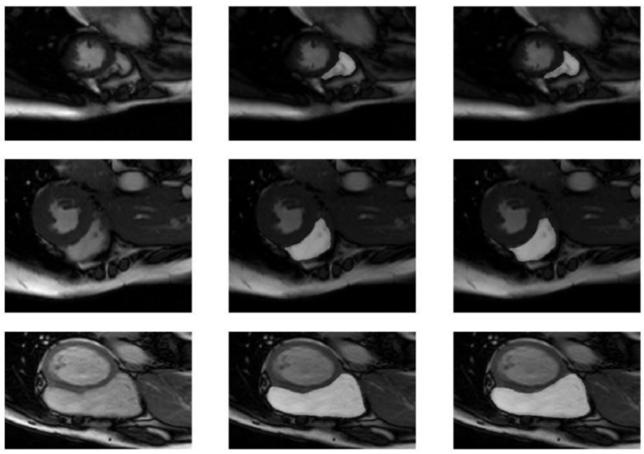


Figure.4 Result of right ventricular segmentation

The dice metrics of the three segmentation test examples in Figure 4 are 0.900, 0.941 and 0.973. Most segmentation methods proposed previously perform are bad in apical images. But in my experiment, the model is so robust that can segment apical images precisely. That's the reason why dice metric can be improved in this paper.

Table I. DICE METRIC OF DIFFERENT METHODS

	#labelled	#unlabelled	Right ventricular
Multi-atlas	80	-	0.840
Semi-supervised learning	80	240	0.888
Proposed	**243**	**-**	**0.900**

It can be seen in Table 1, compared with different methods proposed by other papers, dilated densenet of two losses performs much better in test set. The dice metric improved obviously when using the method proposed in this paper.

Table II. DICE METRIC of DIFFERENT LOSSES

	L1	L2	**Two level losses**
Training set	0.91	0.88	**0.91**
Test set	0.87	0.81	**0.90**

As can be clearly seen from table 2, L1 is better than L2 because L1 is the pixel level loss which is much finer than the shape level loss. But they have the same problem called overfitting that test set is far less effective than training set. Using two level losses strategy to train the model, the phenomenon of overfitting has been greatly alleviated. Though the training performance decreased than L1 loss, two level losses can achieve a much better result in test set, which is close to the training performance.

IV. CONCLUSION

Right ventricular segmentation based on dilated densenet using two level losses is proposed in this paper. This method can effectively segment the region of right ventricular. The experimental results show that the right ventricular segmentation system has certain superiority in the aspect of overfitting after the model alternately trained by two losses, and the dice metric improved in test set and the time of segmentation is greatly reduced. Future work includes experimenting with conditional random field(CRF)[12] to learn about the location structure information so that the systerm can decrease the rate of misjudgment pixels, which obviously don't belong to right ventricular. Jointing this technology may improve the accuracy of segmentation result. As well as performing high level data augmentation techniques[13] to strengthen the model is a key factor.

ACKNOWLEDGMENT

This work is supported by the national natural science foundation of China under grant No.61340004 and by Jiangsu higher vocational college teachers professional lead a dvanced research project under grant No.2017GRGDYX058.

REFERENCES

[1] Litjens G, Kooi T, Bejnordi B E, et al. A survey on deep learning in medical image analysis[J]. Medical Image Analysis, 2017, 42(9):60-88.

[2] Somkantha K, Theera-Umpon N, Auephanwiriyakul S. Left ventricular segmentation of cardiac magnetic resonance images using a novel edge following technique[C]// Cybernetics and Intelligent Systems, 2008 IEEE Conference on. IEEE, 2008:169-174.

[3] ElBaz MS, Fahmy AS. Active shape model with inter-profile modeling paradigm for cardiac right ventricle segmentation[J]. MICCAI,2012:691-698.

[4] Punithakumar K, Noga M, Boulanger P. Cardiac right ventricular segmentation via point correspondence[C]// Engineering in Medicine & Biology Society. Conf Proc IEEE Eng Med Biol Soc, 2013:4010.

[5] Zhang Y, Song Q, Chen C, et al. Accurate segmentation for right ventricles based on affinity propagation clustering and multi-Atlas selection[C]// IEEE International Conference on Medical Imaging Physics and Engineering. IEEE, 2014:171-175.

[6] Bai W, Oktay O, Sinclair M, et al. Semi-supervised Learning for Network-Based Cardiac MR Image Segmentation[C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2017:253-260.

[7] Molaei S, Shiri M, Horan K, et al. Deep Convolutional Neural Networks for left ventricle segmentation[J]. Conf Proc IEEE Eng Med Biol Soc, 2017, 2017:668-671.

[8] Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors[M]// Neurocomputing: foundations of research. MIT Press, 1988:533-536.

[9] Simard P Y, Steinkraus D, Platt J C. Best practices for convolutional neural networks applied to visual document analysis[C]// International Conference on Document Analysis and Recognition. IEEE Computer Society, 2003:958.

[10] F.Yu,V.Koltun,"Multi-scale context aggregation by dilated convolutions", arXiv:1511.07122, 2015.

[11] Huang G, Liu Z, Maaten L V D, et al. Densely Connected Convolutional Networks[C]// CVPR. 2017.

[12] Lafferty J D, Mccallum A, Pereira F C N. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data[C]// Eighteenth International Conference on Machine Learning. Morgan Kaufmann Publishers Inc. 2001:282-289.

[13] Wang S H, Lv Y D, Sui Y, et al. Alcoholism Detection by Data Augmentation and Convolutional Neural Network with Stochastic Pooling[J]. Journal of Medical Systems, 2018, 42(1):2.