基于深度学习和水平集的彩色图像分割方法

杨　宇[1]，崔　陶[2,3]

(1. 四川大学数学学院，成都 610064；2. 四川大学华西第二医院，成都 610041；

3. 出生缺陷与相关妇儿疾病教育部重点实验室，成都 610041)

摘　要：本文结合基于全卷积神经网络的语义分割与基于水平集的图像分割方法，使用 Deep-Lab V2 与 Distance Regularized Level Set Evolution(DRLSE)模型对一般彩色图像进行分割. 通过在 DRLSE 模型中加入一个新形状能量项，该方法提高了零水平集的演化速度. 数值模拟结果验证了方法的有效性.

关键词：深度学习；水平集方法；图像分割；距离正则化水平集演化；DeepLab

中图分类号：O29　　　**文献标识码**：A　　　DOI：10.19907/j.0490-6756.2021.041004

A color image segmentation method based on deep learning and level set

YANG Yu[1]，CUI Tao[2,3]

(1. School of Mathematics，Sichuan University，Chengdu 610064，China；

2. West China Second University Hospital，Sichuan University，Chengdu 610041，China；

3. Key Laboratory of Birth Defects and Related Diseases of Women and Children (Sichuan University)，

Ministry of Education，Chengdu 610041，China)

Abstract：In this study，we combine the semantic segmentation technology based on full convolution neural network and the image segmentation technology based on level set method and uses Deeplab V2 and Distance Regularized Level Set Evolution (DRLSE) model to realize general color image segmentation. To improve the evolution speed of the zero level set segmentation，a new shape energy term is added to the DRLSE model. Numerical simulations verify the efficiency of our method.

Keywords：Deep Learning；Level set method；Image segmentation；DRLSE；DeepLab

1　Introduction

Since the 1970s，the problem of image segmentation has attracted the attention of many scholars. The level set method has been developed in recent 20 years，and has achieved good results in the field of image segmentation. The idea is to represent the two-dimensional plane contour as the zero level set of the three-dimensional implicit function，which is called the Level Set Function (LSF)，and evolves the LSF according to the Partial Differential Equation (PDE). In recent years，many segmentation models based on level set have been presented.

Recently，deep learning and Convolutional Neural Networks (CNN) have made a lot of

breakthroughs in image segmentation. In these methods, CNN is trained by lots of data to effectively extract object features and convert the image to be segmented into simple feature images for pixel level recognition and classification, so as to realize the corresponding mapping relationship between the image to be segmented and the desired segmentation result. This method has strong learning ability, which can deal with complex natural image segmentation problems and multi-objective segmentation problems. This method has robustness and can realize image semantic understanding and segmentation. But the quantity and quality of data required by such methods are positively related to the difficulty of the segmentation task, which costs much more time compared with the usual algorithms, and it usually needs tens of thousands of data to segment the image well.

1. 1   Overview on level set method

The Active Contour Model (ACM)[1] proposed by Kass, etc. creatively in the study of image segmentation methods lays the foundation for the level set method. The level set method was first proposed by Osher et al.[2] in 1988 to study the outer contour of fire. The idea is to express the two-dimensional closed curve as the zero level set of three-dimensional implicit function, and the evolution of the zero level set is used to approximate the edge contour of the target. The concept of reinitialization is given by Chop[3]. Reinitialization means that only updating the values in the vicinity of the zero level set during the evolution of the zero level set, and it will reduce the amount of calculation. However, if the operation is repeated many times, the amount of calculation will increase. The classical segmentation model Mumford Shah (M-S) model was proposed by Mumford et al.[4], which used the minimization of the energy functional to drive the evolution of zero level set. Then Chan et al.[5] proposed the C-V model to simplify the energy functional of M-S model. The model can be used to segment the image with blurred contour, since it does not rely on

gradient correlation information, and has certain robustness for noise. All the models mentioned above need to be reinitialized in the process of numerical computations, so there are two difficulties. One is whether the regularized LSF after reinitialization can ensure that the zero level set approaches the target edge, and the other is how to select the appropriate time for reinitialization in the actual calculation process. Li et al.[6] proposed a method to solve the above two problems, namely DRLSE method. By introducing the regularized term, LSF is guaranteed to be a signed distance function and the reinitialization step is avoided.

Generally speaking, active contour models can be divided into two categories: edge based segmentation model and region based segmentation model. Let's take a look at several edge based segmentation models. Li et al.[7] put forward the double well potential function and updated the DRLSE model. However, the model still has some shortcomings such as unidirectional evolution and reduction of convergence rate. Wang et al.[8] proposed an Adaptive Level Set Evolution (ALSE) model. Compared with DRLSE model, this model reduces the dependence on the initial contour set and accelerates the evolution of zero level set, but there are still some shortcomings. Zhang et al.[9] proposed a level set evolution model driven by the optimal area energy term. This model has the characteristics of high efficiency and stability. It is robust for noise, but it is not adaptive and sensitive to initialization. Here are some region based segmentation models. Chan and Vese[5] put forward a Piecewise Constant (PC) model, which can segment images with blurred edges, but it is not appropriate for images with uneven intensity. Li et al.[10] proposed a Region Scalable Fitting (RSF) model, which can segment non-uniform images, but has a large amount of computation and is sensitive to the initial contour. Zhang et al.[11] proposed Local Image Fitting (LIF) model. Compared with RSF model, LIF converges more fast, but the ro-

bustness of initialization needs to be improved. Wang et al.[12] proposed a Local Gaussian Distribution Fitting (LGDF) model, which uses LGDF energy to calculate the variance statistics of images, and can segment regions with similar strength but different variances. However, LGDF model takes much time to segment. Ding et al.[13] added a logarithmic energy term to the RSF model and proposed Laplacian optimization of the region-scalable fitting (RSFLog) model. Compared with RSF model, RSFLog segmentation model has higher segmentation speed, accuracy and robustness. Fast segmentation and good robustness will be achieved by introducing the local pre-fitting (LPF) energy into variational level set method in Ref. [14]. Xia et al.[15] improved CV model, proposed a new Smooth convergence function based on the region confidence.

1.2　Overview on image segmentation based on deep learning

In recent years, deep learning technology and neural network have achieved many important research results in the field of image processing. CNN[16] is one of the branches of artificial neural network. After a certain degree of data training, it can effectively extract image information and realize recognition and classification at pixel level, which is often used in image processing related issues. In 2015, Long et al.[17] proposed Full Convolutional Network (FCN) and realized semantic segmentation of images. The input image of FCN is not limited by the size of the input image. The convolution operation is used to predict the image category at the pixel level. The size of output result is consistent with the size of the input image. It is the first time to use FCN to solve the problem of image segmentation, but there are also some problems such as the segmentation effect is not accurate enough and the amount of calculation is too large.

Chen et al.[18] proposed the network model DeepLab V1 in 2015. DeepLab V1 innovatively applied hole convolution to vgg-16 network, which expanded the receptive field and improved

the segmentation accuracy of the model. In 2016, Chen et al.[19] proposed DeepLab V2 to extract global semantic information by atrous spatial pyramid pooling (ASPP) to enhance the ability of multi-scale object recognition. Although the accuracy has been improved, the amount of calculation has become very large because of the addition of ASPP module. Then, Chen et al.[20] proposed DeepLab V3, adding convolution layer and global average pooling in ASPP module. By combining serial connection hole convolution with ASPP module, more abundant global semantic information can be obtained.

In Section 2, we review the level set method and DRLSE model. The method based on deep learning and DRLSE model is presented in Section 3. Numerical experiments are shown in Section 4. The conclusion is presented in Section 5 with some remarks.

2　Basic knowledge

2.1　Level Set Method

We will briefly introduce level set method in this section. Suppose the curve C is expressed as

$$C(t) = \{(x,y), \varphi(x,y,t) = c\} \tag{1}$$

The evolution of level set function is governed by

$$\frac{\partial \varphi}{\partial t} = F(p,t) |\nabla \varphi| \tag{2}$$

where $F(p,t)$ is the speed function that controls the motion of the contour. The vertical view of level set function is shown in Fig. 1. We define a measure of distance and use $d_c(x,y)$ represents the shortest distance from a point (x, y) to the initial curve C, then LSF can be initialized as a general Signed Distance Function (SDF),

$$\varphi(x,y) = \begin{cases} -d_C(x,y), & (x,y) \text{ is inside } C, \\ 0, & (x,y) \text{ is on the edge of } C, \\ d_C(x,y), & (x,y) \text{ is outside } C \end{cases} \tag{3}$$

where (x, y) represents a pixel in the image. During the process of iterative evolution, some numerical values may have very serious deviation in the calculation process. The reason is that the level set function will gradually become irregular

near the zero level set in the process of evolution, which will lead to the gradual accumulation of errors. In order to prevent such errors, reinitialization is introduced. To a certain extent, the level set maintains stability and regularity in the evolution process, and reduces the risk of error occurrence. However, there are still some defects in reinitialization, for example, it greatly increases the computational complexity in the evolution process of LSF, it is difficult to choose the appropriate time to reinitialize, and it can not guarantee that the regular LSF after reinitialization can drive the zero level set function to the edge of the target.
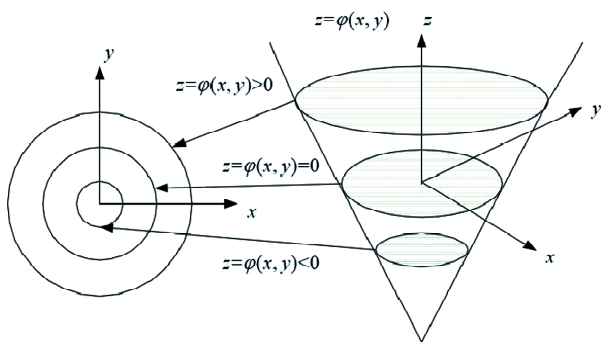

Fig. 1   Vertical view of level set function

2.2   The DRLSE model

In order to solve the problem of reinitialization in the iterative process, Li, etc. [6] proposed the DRLSE model which was is used. Let $\varphi: \Omega \rightarrow \mathbf{R}$ be a level set function on the domain Ω. The energy functional form of DRLSE model is as follows

$$E(\varphi) = \mu R(\varphi) + \lambda L(\varphi) + \alpha A(\varphi) \tag{4}$$

where $R(\varphi)$ represents the distance regular term, which is used to ensure the smoothness of LSF in the evolution process and avoid reinitializing LSF during the evolution process; $L(\varphi)$ represents the length term, which is used to drive the zero level set to evolve towards the edge of the target; $A(\varphi)$ represents the area term. By minimizing the area term, the level set can be contracted ($\alpha > 0$) or expanded ($\alpha < 0$) to guide the level set to evolve towards the target step by step. While μ, λ, α correspond to the distance regular term, the coefficient

of length term and area term respectively. A potential function $P(s)$ is proposed, which is defined as follows

$$P(s) = \frac{1}{2}(s-1)^2 \tag{5}$$

which has a minimum value at s = 1. Thus, the distance regular term $R(\varphi)$ can be defined as follows

$$R(\varphi) = \iint_{\Omega} P(|\nabla \varphi|) \mathrm{d}x \mathrm{d}y \tag{6}$$

which can guarantee that the minimum value of distance regular term $R(\varphi)$ can be obtained when $s=1$. At the same time, it also ensures that SDF has the characteristics of "$|\nabla \varphi| = 1$" when $P(s)$ has a minimum value at $s= 1$. The edge indicator function g is introduced by

$$g = \frac{1}{1 + |\nabla G_\sigma * I|} \tag{7}$$

where C_σ is a Gaussian filter function with a standard deviation of σ. The convolution in (7) is used to smooth the image to reduce the noise. In general, the minimum value of g is obtained when LSF evolves to the boundary of the target. The length term $L(\varphi)$ and area term $A(\varphi)$ can be defined by g as follows,

$$L(\varphi) = \iint_{\Omega} g\delta(\varphi)|\nabla \varphi| \mathrm{d}x \mathrm{d}y \tag{8}$$

$$A(\varphi) = \iint_{\Omega} gH(-\varphi) \mathrm{d}x \mathrm{d}y \tag{9}$$

where δ is dirac function, H is Heaviside function.

In conclusion, $L(\varphi)$ can calculate the curve integral of φ along the zero level set, and the minimum value is obtained when the zero level set is located at the edge of the target to be segmented; as for $A(\varphi)$, it can accelerate the evolution of LSF. When the initial contour $\varphi_{initial}$ is outside the target contour, a positive value of α will cause the $\varphi_{initial}$ shrinks rapidly inward; when $\varphi_{initial}$ is within the target contour, a negative α value can make the $\varphi_{initial}$ expands rapidly. Li et al. [7] also proposed the improvement of the potential function $P(s)$. They used double well potential function $P_2(s)$ which is defined as follows,

$$P_2(s) = \begin{cases} \dfrac{1}{2}(s-1)^2, & s \geqslant 1, \\ \dfrac{1}{4\pi^2}(1-\cos(2\pi s)), & s < 1 \end{cases} \qquad (10)$$

This potentialfunction can ensure that the SDF characteristic is maintained near the zero level set, i.e. $|\nabla\varphi| = 1$. In the region far away from the zero level set, LSF is a constant and remains unchanged, i.e. $|\nabla\varphi| = 0$. In order to have the above properties, the potential function must have two minimum points $s=0$ and $s=1$ (mark $P_2(s)$ as $P(s)$ in the following). By (6)(8) and (9), we have the energy functional as follows,

$$E(\varphi) = \mu\iint\limits_{\Omega} P(|\nabla\varphi|)\mathrm{d}x\mathrm{d}y +$$

$$\lambda\iint\limits_{\Omega} g\delta(\varphi)|\nabla\varphi|\mathrm{d}x\mathrm{d}y + \alpha\iint\limits_{\Omega} gH(-\varphi)\mathrm{d}x\mathrm{d}y$$

$$(11)$$

The corresponding gradient flow of the energy $E(\varphi)$ is

$$\frac{\partial\varphi}{\partial t} = \mu\mathrm{div}(d_p(|\nabla\varphi|)\nabla\varphi) +$$

$$\lambda\delta(\varphi)\mathrm{div}\left(g\frac{\nabla\varphi}{|\nabla\varphi|}\right) + \alpha g\delta(\varphi) \qquad (12)$$

where $d_p(s) = \dfrac{P'(s)}{s}$. The corresponding φ in the energy functional (12) with the minimum value is the final contour we need.

2.3　Semantic segmentation based on deep learning

For the current semantic segmentation with deep learning, it originated in 2015. Long et al.[17] proposed the FCN. FCN is an evolution of convolutional neural network, which is often used in the prediction and classification of image processing, because the output results of CNN are mostly probability. FCN can classify each pixel of the input image, so as to realize the pixel level image segmentation, namely semantic segmentation. Different from CNN using full connection layer after convolution layer, FCN adopts deconvolution layer. The purpose of FCN is to upsampling the feature image generated by the convolution layer at the end, so that the convoluted image can be restored to the size of the previous input image. Thus, the output result not only retains the space information contained in the input image, but also can predict and classify each pixel of the original input image.

The neural network used in this paper is DeepLab V2 developed from FCN. DeepLab V2 has three characteristics: one is the use of hole convolution, which changes the operation mode of traditional convolution operation, expands the receptive field and maintains the resolution on the basis of unchanged input data; the second is to propose Atrous Spatial Pyramid Pooling (ASPP) for multi-scale processing of input information; the third is to combine fully connected Conditional Random Field (CRF) to process the input information The convoluted image is restored accurately. As shown in Fig. 2, the pixel level image obtained by convolution of the input original image has a good contour after CRF recovery.
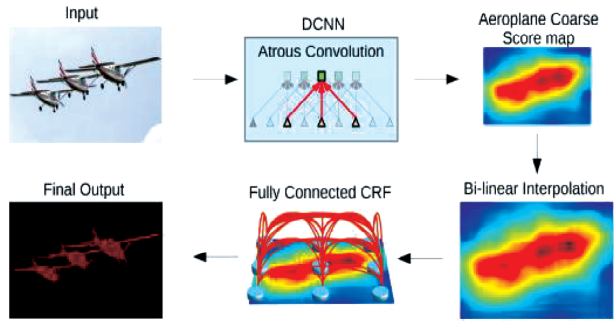

Fig. 2　Functions of DeepLab

3　Preprocessing based on deep learning and improved DRLSE model

Compared with the traditional DRLSE model segmentation method, this paper has two main improvements: (i) The neural network based semantic segmentation is used for image preprocessing. (ii) The image information obtained by semantic segmentation is transformed into an item of energy functional, which is added to the iterative process of DRLSE model.

3.1　Preprocess

The DRLSE model is usually segmented by the gray image of the input image. However, if the gray image of the original image is directly segmented, the result is often very inaccurate, or

even very bad, as shown in Fig. 3 below.

(a) originalimage (b) grayscale image (c) segmentation result
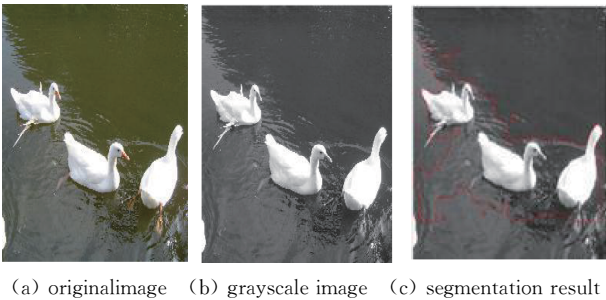
Fig. 3 Segmentation of grayscale image

Our goal is to segment the three ducks in (a) of Fig. 3. It can be seen that (a) is a picture with less color types, and the color (white) of the target to be segmented is clearly different from the background color (green). Therefore, in the gray image of (a), that is, the gray level of the target in (b) is also significantly different from that of the background. However, after segmentation by DRLSE model, the effect is shown in (c), which can only roughly describe the area where the target is located, and can not accurately drive to the edge of the target. Therefore, the Otsu threshold algorithm, which is often used for image preprocessing in DRLSE model, will be introduced below.

Otsu threshold algorithm is an adaptive threshold algorithm proposed by Japanese scholar Otsu[21]. It divides the original image into two parts, the target and the background, according to the gray characteristics of the image itself. The measure standard of Otsu threshold algorithm is the maximum inter class variance, that is, if the variance between the target and the background is larger, the difference between the target and the background is greater, and the probability that the target and the background are wrongly divided into the same category is smaller. In Fig. 4, (a) is transformed into (b) after Otsu threshold algorithm, and (b) is segmented by level set method of DRLSE model to obtain (c). Compared with (d), (c) has a great improvement on the approximation of the target contour, and objectively it has been very close to the target edge, which is a good segmentation. However, when there are

many kinds of colors between the target and the background in the input image or the difference is not obvious enough, the preprocessing effect through Otsu threshold algorithm is not evident, as shown in Fig. 5.

(a) original image   (b) Otsu

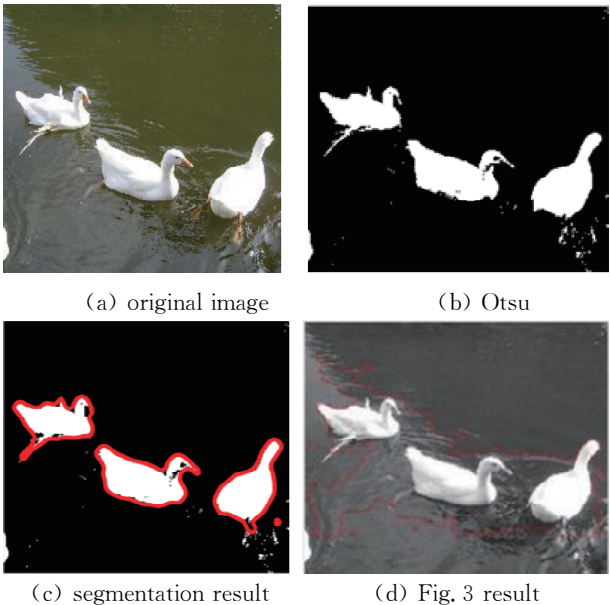(c) segmentation result   (d) Fig. 3 result

Fig. 4 Segmentation of Otsu

In Fig. 5, our target is the bird in (a). After Otsu threshold algorithm, (a) becomes (b). From (b), we can directly see that the image becomes very bad after preprocessing. Therefore, (d) as the result of segmentation is intuitively imprecise and deviates from the target seriously, which is a bad segmentation. Therefore, the DRLSE model pretreated by Otsu threshold algorithm is not suitable for image segmentation of color rich images, which is one of the reasons why DRLSE is commonly used in medical CT image segmentation.

In order to improve this dilemma, we use the new semantic segmentation based on neural network to preprocess the general color image in this paper. In theory, semantic segmentation is the classification of image pixels. Pixels belonging to the same category will be classified into one category. Therefore, the difference between the target and the background after the semantic segmentation preprocessing will be a very significant in color labeling, which is more conducive to the

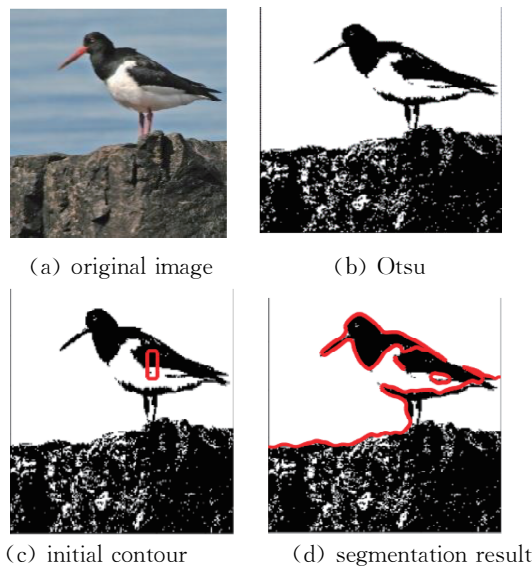DRLSE model for target contour approximation after the gray image.

（a) original image　　　　　（b) Otsu

（c) initial contour　　　（d) segmentation result

Fig. 5　Segmentation of Otsu

（a) original image　（b) DeepLab　（c) grayscale image
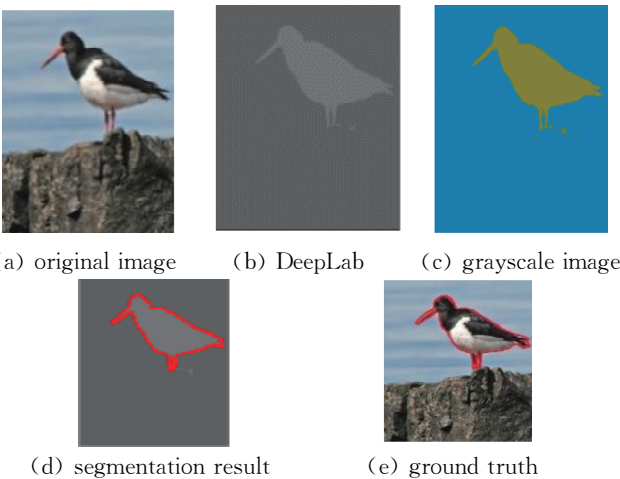
（d) segmentation result　　（e) ground truth

Fig. 6　Segmentation of DeepLabu

In Fig. 6，(a) is obtained after semantic segmentation based on neural network. It is observed that (b) is better than the result using Otsu threshold algorithm for the case withhuge difference between target and background. The figure (c) is the gray image transformed from (b)，and (d) is the segmentation result of DRLSE model，and (e) is the ground truth of the original image. It can be seen that the zero level set which finally stops evolving is very close to the contour of the target.

In Section 4，more examples are presented to prove the superiority of semantic segmentation in

image preprocessing compared with Otsu algorithm.

3. 2　Improvement of DRLSE model

For an image I，when using DRLSE model to segment image I，it is usually necessary to manually select an initial contour $\varphi_{initial}$. Fig. 8 shows the contour of edge indicator function g detected φ_g and the actual contour $\varphi_{groundtruth}$ (hereinafter referred to as φ_{GT})，you can find the φ_g and φ_{GT}，in general，has a similar outline. Generally speaking，in the evolution process of the level set，when the initial contour $\varphi_{initial}$ is selected，the contour represented by zero level set converges to φ_{final} with the indication of edge indicator function g and iterative update of zero level set. The process is shown in Fig. 8，which shows the expansion of the initial contour $\varphi_{initial}$ selected inside the target in the process of level set iteration，and finally gradually expand to the edge of the target to complete the segmentation.
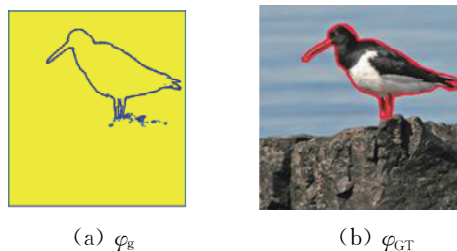

（a) φ_g　　　　　　　　（b) φ_{GT}

Fig. 7　φ_g and φ_{GT}

Recall the semantic segmentation mentioned in Section 3. 1，after preprocessing，the color difference between the target and the background is very significant，which is shown in Fig. 9. In this case，we can obtain a contour which is similar to the final contour directly through the color difference between the target and the background. Because DeepLab is used for preprocessing，this contour is defined as a prior contour φ_{DPL}. Fig. 10 is a comparison of the prior contour φ_{DPL} with the contour φ_g detected by the edge indicator function and the actual contour φ_{GT} of the target. Compared with φ_{DPL}，we can find that φ_g contains some redundant image information，which is often caused by the error in the preprocessing with DeepLab. However，with the itera-

tive update of the zero level set, the redundant information will be gradually filtered out. The outline of φ_g and φ_{DPL} is similar, but φ_{DPL} contains much less redundant information than φ_g.

（a）initial contour　　　（b）100 iterations

（c）200 iterations　　　（d）segmentation result
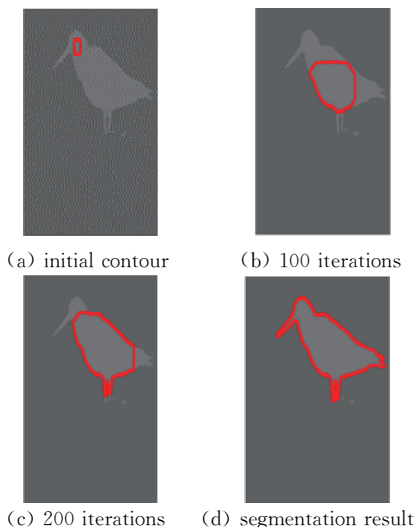
Fig. 8　The iterations

Therefore, it is unreasonable to direct the initial contour $\varphi_{initial}$ to φ_{DPL}. However, we can use φ_{DPL} to guide $\varphi_{initial}$ to evolve rapidly at the beginning of the iterative evolution process from $\varphi_{initial}$ to φ_{GT}, so that it can quickly form a contour similar to φ_{GT}, and then through the previous iterative update, it can better approximate the truth contour φ_{GT}. Therefore, according to the previous analysis, our model is more effective in two cases. One is to segment the image with simple contour, in which φ_{GT} and φ_{DPL} have similar shape; the other is φ_g contains redundant information while φ_{DPL} con tains less or no redundant information.

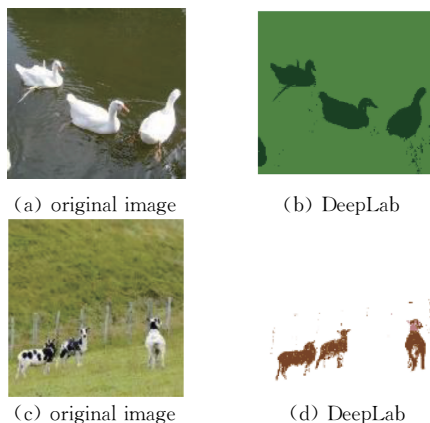（a）original image　　　　　（b）DeepLab

（c）original image　　　　　（d）DeepLab

Fig. 9　Function of Deeplab

Recall the energy functional of DRLSE model mentioned in (11), we will add a new energy term φ_{DPL} to formulation (11) in order to make it have a great contribution to the evolution in the early stage of the evolution of zero level set, that is, it can quickly approach the general outline. According to Ref. [22], we can define it as:

$$E_{DPL}(\varphi, \varphi_{DPL}) = \omega \iint_{\Omega} (\varphi - \varphi_{DPL})^2 \mathrm{d}x\mathrm{d}y \qquad (13)$$

which represents the shape term that drives $\varphi_{initial}$ towards the shape φ_{DPL}, where ω decreases gradually to 0 as iterations increases. This definition is consistent with the purpose of φ_{DPL} to help zero level set approach the contour in the early stage, but does not affect the evolution of zero level set in the later stage. Therefore we have

$$E(\varphi) = \mu \iint_{\Omega} P(|\nabla\varphi|) \mathrm{d}x\mathrm{d}y +$$
$$\lambda \iint_{\Omega} g\delta(\varphi)|\nabla\varphi| \mathrm{d}x\mathrm{d}y + \alpha \iint_{\Omega} gH(-\varphi) \mathrm{d}x\mathrm{d}y +$$
$$\omega \iint_{\Omega} (\varphi - \varphi_{DPL})^2 \mathrm{d}x\mathrm{d}y \qquad (14)$$

（a）original image　　　（b）DeepLab　　　（c）φ_g
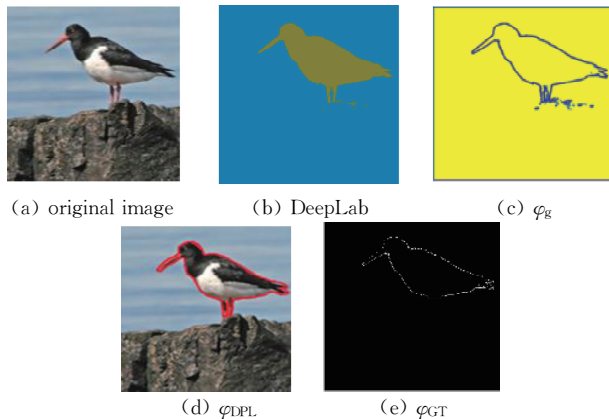
（d）φ_{DPL}　　　　　（e）φ_{GT}

Fig. 10　Comparison

Fig. 11 shows the difference between the traditional DRLSE model and the improved DRLSE model in image segmentation after semantic segmentation preprocessing. It can be seen that the contour evolution of the improved model is better than that of the traditional model in redundant information processing. It can be seen from (i) and (j) in Fig. 11 that the contour of the improved model converges to the final contour after 400 iterations, while the contour of the traditional model is still evolving and converges only after 500 iterations.
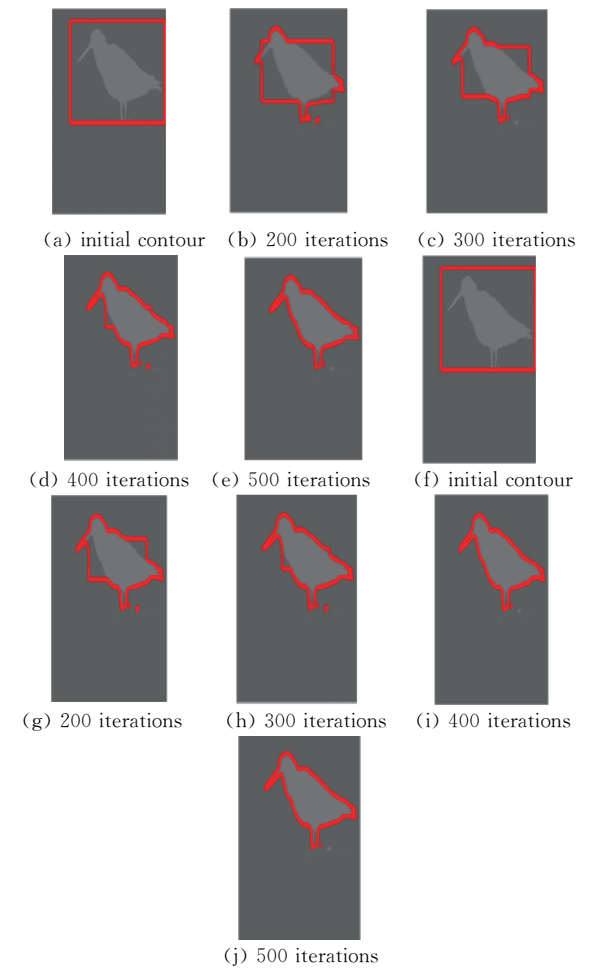
(a) initial contour　(b) 200 iterations　(c) 300 iterations

(d) 400 iterations　(e) 500 iterations　(f) initial contour

(g) 200 iterations　(h) 300 iterations　(i) 400 iterations

(j) 500 iterations

Fig. 11　(a)～(e) traditional DRLSE model，(f)～(j) improved DRLSE model

4　Numerical example

4.1　Data set

In this example，the DeepLab V2 model is trained using Pascal voc2012 data set，a total of 17 125 color images are trained. Pascal voc2012 data set mainly provides label data for supervised learning in visual tasks. In semantic segmentation，it is often the preferred data set. It has 20 categories，including some common animals，such as birds，cats，horses，transportation vehicles，furniture supplies and people.

4.2　Parameter setting

The parameters set in this section include DeepLab V2 and two parts of DRLSE model，as shown in the Tab. 1 and Tab. 2. The software for training and using DeepLab V2 is Python 3.7，and the software using DRLSE model is Matlab 2016a.

In this paper，the parameter ω proposed in Section 3.2 is set as $\omega = \frac{1}{8} e^{-\frac{9}{2} \left[\frac{\lfloor t \rfloor}{100} + 1 \right]}$ in consideration of the decreasing property of ω with iteration times t，where $\lfloor t \rfloor$ is the downward rounding function.

Tab. 1　Parameter setting of DeepLab V2

Parameter	Momentum	BatchSize	LearningRate
Value	0.9	4	9×10^{-4}
	WeightDecay	Classes	Steps
	0.0005	21	40001

Tab. 2　Parameter setting of DRLSE

Parameter	μ	λ	α
Value	0.2	5	± 3

4.3　Measurement criteria

In this paper，we use the Dice coefficient to measure image segmentation，which is defined as follows，

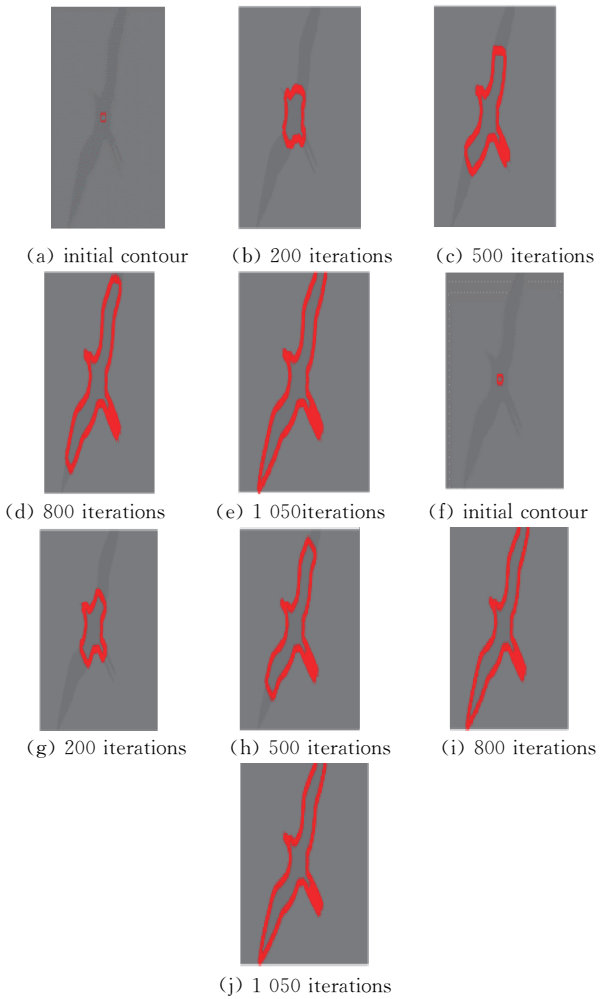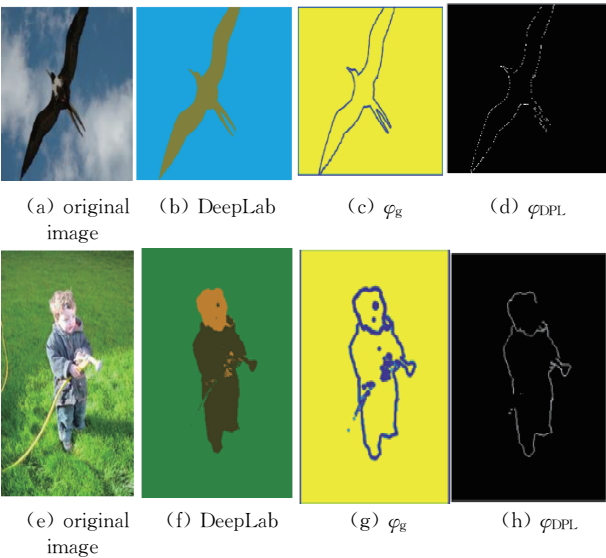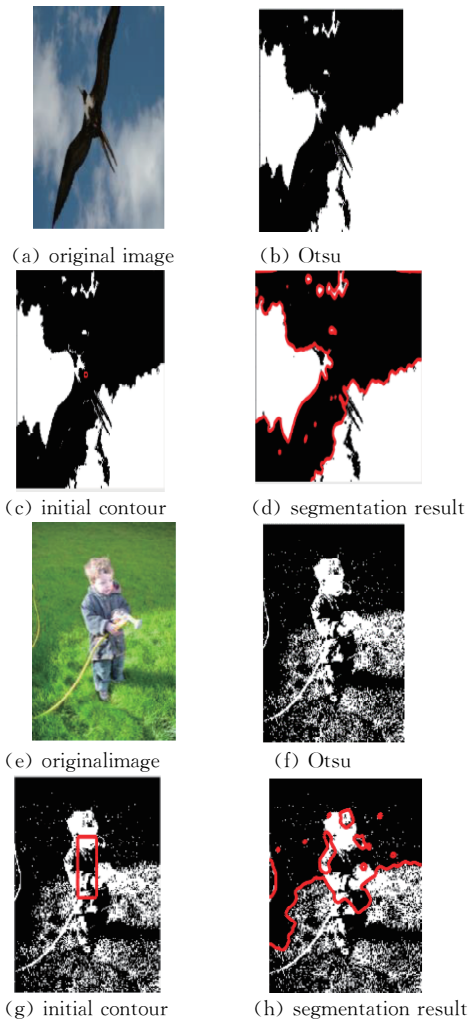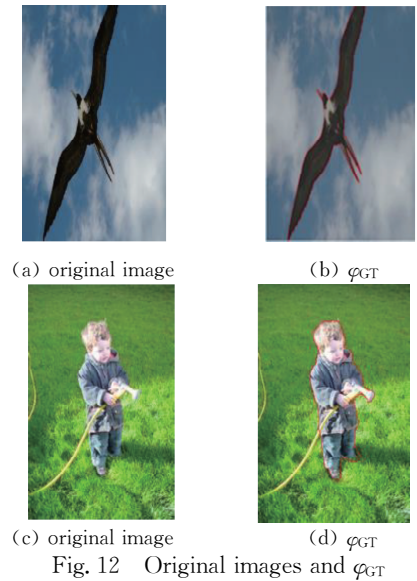$$DC(A,B) = \frac{2|A \cap B|}{|A| + |B|} \tag{15}$$

where A represents the set corresponding to the final contour φ_{final} of the model evolution of the target，and B represents the set corresponding to the actual real contour φ_{GT} of the target.

4.4　Experimental process

In this experiment，(a) and (c) in Fig. 12 are selected and their sizes are 500×332 and 240×320 pixels，respectively. (a) is the image mentioned in Section3.2 with simple target contour，(c) is the image with more redundant information after DeepLab preprocessing mentioned in Section 3.2. Fig. 12 shows the selected image and the corresponding ground truth φ_{GT}.

The traditional DRLSE model is used for image segmentation. Fig. 13 shows the results preprocessed by Otsu threshold algorithm and the image after DRLSE model segmentation. The final segmentation of the contour φ_{final} and φ_{GT} has a huge gap that can be seen by the naked eye. It can be seen that segmentation by traditional "Otsu threshold algorithm + DRLSE model" will lead to extremely bad results. "DeepLab + DRLSE model" is used in our method. Fig. 14 shows the results of preprocess. It can be seen

that the image preprocessed by DeepLab retains the information of the target contour well. The DRLSE model and the improved DRLSE model are used for image segmentation.

(a) original image　　　　(b) φ_{GT}

(c) original image　　　　(d) φ_{GT}

Fig. 12　Original images and φ_{GT}

(a) original image　　　　(b) Otsu

(c) initial contour　　　　(d) segmentation result

(e) originalimage　　　　(f) Otsu

(g) initial contour　　　　(h) segmentation result

Fig. 13　Segmentation of Otsu

(a) original image　　(b) DeepLab　　(c) φ_g　　(d) φ_{DPL}

(e) original image　　(f) DeepLab　　(g) φ_g　　(h) φ_{DPL}

Fig. 14　Preprocess

(a) initial contour　　(b) 200 iterations　　(c) 500 iterations

(d) 800 iterations　　(e) 1 050iterations　　(f) initial contour

(g) 200 iterations　　(h) 500 iterations　　(i) 800 iterations

(j) 1 050 iterations

Fig. 15　(a)～(e) traditional DRLSE model，(f)～(j) improved DRLSE model

From Fig. 15，it is observed that the improved model converges faster than the traditional DRLSE model under the same iteration times. It can be seen from (i) and (j) in Fig. 15 that the

segmentation of the model proposed in this paper has been completed after 800 iterations，while（e）shows that the traditional model can converge to the final contour after 1 050 iterations. （b）（c）（g）and（h）of Fig. 16 show that the model proposed in this paper has a faster convergence rate at the same iteration number. It can be seen from （d）and（i）in Fig. 16 that both models have rough outlines after 300 iterations，but there is still redundant information waiting to be processed（red part of the head）. It can be seen from （e）and（j）in Fig. 16 that the improved model completes the processing of redundant information and converges to the final contour after 650 iterations，while the traditional model completes segmentation after 750 iterations. Therefore，compared with the traditional DRLSE model，the convergence speed of the proposed model is improved，and the convergence accuracy will be compared in Section 4. 5.
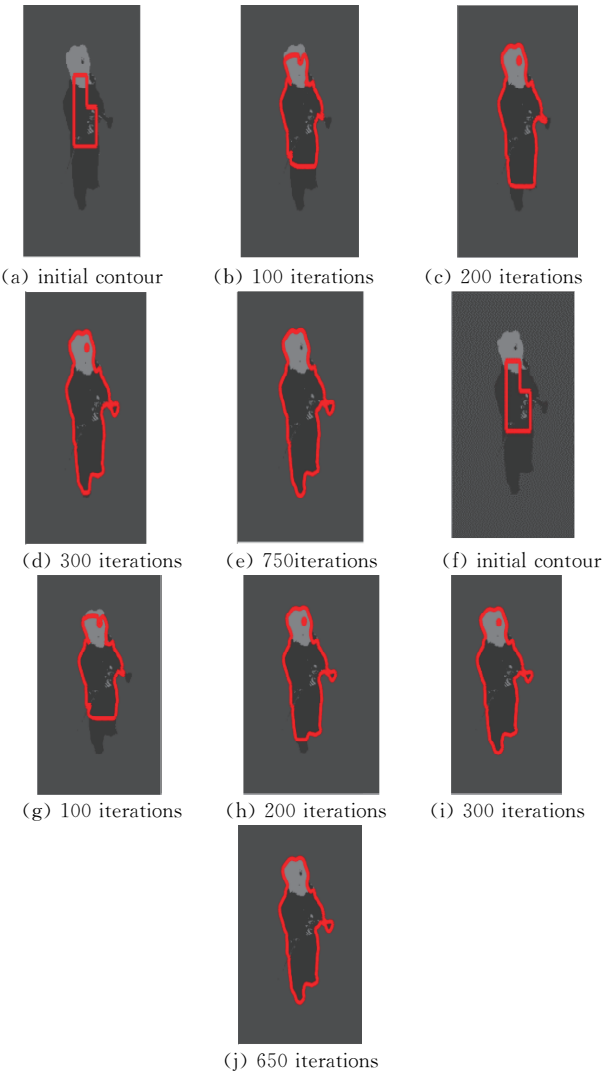
(a) initial contour　　(b) 100 iterations　　(c) 200 iterations

(d) 300 iterations　　(e) 750iterations　　(f) initial contour

(g) 100 iterations　　(h) 200 iterations　　(i) 300 iterations

(j) 650 iterations

Fig. 16　（a）～（e）traditional DRLSE model，（f）～（j）improved DRLSE model

4. 5　Comparison of methods

Tab. 3 shows the DC coefficients of pictures （a）and（c）in Fig. 12 respectively，which are：（i）DC coefficients obtained by using the traditional DRLSE model after preprocessing by Otsu algorithm；（ii）DC coefficients of segmentation results by traditional DRLSE model after being processed by DeepLab V2；（iii）DC coefficients of segmentation results with the improved DRLSE model after being processed by DeepLab V2.

Tab. 3　Comparison

Method	Otsu+DRLSE	DeepLab+ DRLSE	DeepLab+ improved DRLSE
DC(m)	<0. 5	0. 972	0. 968
DC(n)	<0. 5	0. 838	0. 833

There is almost no difference in the segmentation accuracy between the two models. The accuracy of the model proposed in this paper is slightly different in the order of the thousandth，but in terms of convergence speed，the improved model proposed in this paper has more advantages than the traditional model.

5　Conclusions

Compared with the traditional DRLSE model，this paper uses neural network to segment general color images，and merges the preprocessing information into a new energy item，which is added to the DRLSE model to segment the color image and obtain the contour. The segmentation accuracy is well retained，the number of iterations of zero level set is reduced，and the speed of segmentation is accelerated.

In the future，we will generate our method to the problem with several objects for semantic segmentation，and perform semantic segmentation for more kinds of color images. We will optimize the setting of parameter ω in this paper and reduce the artificial adjustment.

References：

[1]　Kass M，Witkin A，Terzopoulos D. Snakes：active contour models ［J］. Int J Comput Vis，1988，

1：321.

［2］　Osher S，Sethian J A. Fronts propagating with curvature-dependent speed：algorithms based on Hamilton-Jacobi formulations ［J］. J Comput Phys，1988，79：12.

［3］　Chopp D L. Computing minimal surfaces via level set curvature flow［J］. J Comput Phys，1991，106：77.

［4］　Mumford D B，Shah J. Optimal approximations by piecewise smooth functions and associated variational problems［J］. Commun Pure Appl Math，1989，42：577.

［5］　Chan T F，Vese L A. Active contours without edges［J］. IEEE T Image Process［J］. 2001，10：266.

［6］　Li C，Xu C，Gui C，et al. Level set evolution without re-initialization：a new variational formulation ［C］// 2005 IEEE computer society conference on computer vision and pattern recognition（CVPR'05）. Piscataway：IEEE，2005.

［7］　Li C，Xu C，Gui C，et al. Distance regularized level set evolution and its application to image［J］. IEEE T Image Process，2010，19：3243.

［8］　Wang Y，He C. An adaptive level set evolution equation for contour extraction［J］. Appl Math Comput，2013，219：11420.

［9］　Zhang X，Weng G. Level set evolution driven by optimized area energy term for image segmentation ［J］. Optik，2018，168：517.

［10］　Li C，Kao C Y，Gore J C，et al. Minimization of region-scalable fitting energy for image segmentation ［J］. IEEE T Image Process，2008，17：1940.

［11］　Zhang K，Song H，Zhang L. Active contours driven by local image fitting energy［J］. Pattern Recognit，2010，43：1199.

［12］　Wang L，He L，Mishra A，et al. Active contours driven by local Gaussian distribution fitting energy ［J］. Signal Process，2009，89：2435.

［13］　Ding K，Xiao L，Weng G. Active contours driven by region-scalable fitting and optimized Laplacian of Gaussian energy for image segmentation ［J］. Signal Process，2017，134：224.

［14］　Ding K，Xiao L，Weng G. Active contours driven by local pre-fitting energy for fast image segmentation［J］. Pattern Recognit Lett，2018，104：29.

［15］　Xia X，Ge L，Meng H Y. Image segmentation based on an improved CV model［J］. J Sichuan Univ：Nat Sci Ed（四川大学学报：自然科学版），2017，54：1185.

［16］　Krizhevsky A，Sutskever I，Hinton G E. Image net classification with deep convolutional neural networks［C］// Bartlett P，Pereira F，Burges C，et al. Advances in neural information processing systems. San Francisco：Margan Kaufmann Publishers，2012.

［17］　Long J，Shelhamer E，Darrell T. Fully convolutional networks for semantic segmentation［C］// Proceedings of the IEEE conference on computer vision and pattern recognition. Piscataway：IEEE，2015.

［18］　Chen L C，Papandreou G，Kokkinos I，et al. Semantic image segmentation with deep convolutional nets and fully connected crfs［EB/OL］. https：//arXiv：1412.7062，2014.

［19］　Chen L C，Papandreou G，Kokkinos I，et al. DeepLab：Semantic image segmentation with deep convolutional nets，atrous convolution，and fully connected crfs ［J］. IEEE T Pattern Anal，2017，40：834848.

［20］　Chen L C，Papandreou G，Schroff F，et al. Rethinking atrous convolution for semantic image segmentation ［EB/OL］. https：// arXiv：1706.05587，2017.

［21］　Otsu N. A threshold selection method from gray-level histograms ［J］. IEEE T Syst Man Cy-B，1979，9：62.

［22］　Ngo T A，Lu Z，Carneiro G. Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance ［J］. Med Image Anal，2017，35：159.