A POST-PROCESSING METHOD BASED ON FULLY CONNECTED CRFS FOR CHRONIC WOUND IMAGES SEGMENTATION AND IDENTIFICATION

Junnan Zhang and Hanyi Nie

Computer College, NUDT, Changsha, China

ABSTRACT

Chronic wound have a long recovery time, occur extensively, and are difficult to treat. They cause not only great suffering to many patients but also bring enormous work burden to hospitals and doctors. Therefore, an automated chronic wound detection method can efficiently assist doctors in diagnosis, or help patients with initial diagnosis, reduce the workload of doctors and the treatment costs of patients. In recent years, due to the rise of big data, machine learning methods have been applied to Image Identification, and the accuracy of the result has surpassed that of traditional methods. With the fully convolutional neural network proposed, image segmentation and target detection have also achieved excellent results. However, the accuracy of chronic wound image segmentation and identification is low due to the limitation of the deep convolution neural network. To solve the above problem, we propose a post-processing method based on fully connected CRFs with multi-layer score maps. The experiment results show that our method can be used to improve the accuracy of chronic wound image segmentation and identification.

KEYWORDS

Fully Connected CRFs, Chronic Wound Segmentation, Post-processing Method

1. INTRODUCTION

In recent years, with the rise of big data, the field of artificial intelligence has been aroused a broad concern. AlexNet [1], the champion of ImageNet competition in 2012, uses convolutional neural networks for image classification and recognition. Its accuracy exceeds the traditional method significantly, which makes people pay attention to the application of convolutional neural networks in the field of image classification again. Afterward, people continued to innovate (VGG[2], GoogleNet[3], Residual Net[4], DenseNet[5], CapsuleNet[6] and other deep neural networks), and further improve the accuracy of image classification. However, in the field of medical image classification and segmentation, the accuracy of using deep convolution neural networks is relatively low. Therefore, the application of convolutional neural networks in this area is not effective.

Chronic wound including Diabetic foot ulcers, venous leg ulcers, and acne, has long recovery time and need different methods of treatment at various period. The current treatment of chronic wounds usually takes up a significant amount of medical resources and is not easy to treat [7-9]. Long-term hospitalization is a burden for both hospitals and patients. On the one hand, the resources of the hospital are occupied for an extended period, and it is impossible to provide medical services to other patients in urgent need. On the other hand, the long-term hospitalization costs are too high for most of the patients to afford. It is also very common in some remote areas,

patients are far away and inconvenient to see a doctor. These conditions have brought great suffering to the patients.

Hence, it is of great importance to do a post processing method based on fully-connected CRFs for chronic wound images segmentation and identification. This method helps improve the accuracy of segmentation and classification and the results obtained can assist doctors in diagnosis and treatment.

In this paper, we first introduce fully connected CRFs(conditional random fields). Then we propose a post processing method based on fully-connected CRFs for chronic wound images segmentation and identification. Finally we make compare experiment and the results prove the validity of this method.

2. RELATED WORKS

The basic CRF model contains a unary energy function at an independent pixel or image block and a paired energy function in a neighboring pixel block or image block[11-14]. The adjacency CRF structure generated based on that model limits its association between distant pixel points or image blocks within an image, and thus causes an excessively smooth object boundary. To improve the accuracy of segmentation and identification, the researchers extended the basic CRF framework by adding hierarchical connections and higher-order energy functions in the image region[15-18]. However, this method is bound to be limited by the accuracy of unsupervised image segmentation. Although some progress has been made, this limitation affects the ability of region-based CRF method to accurately assign labels to complex image boundary pixels[19]. Although some progress has been made, this limitation affects the ability of region-based methods to accurately assign labels to complex image boundary pixels.
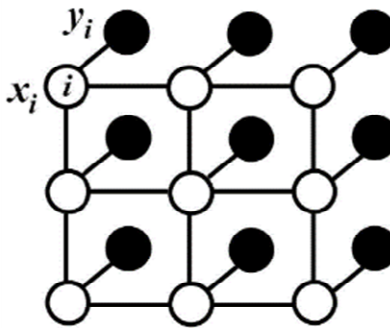

Figure 1. Fully connected CRF model on pixel image[20]

Figure 1. shows a simple fully connected CRF model on pixel image, y_i represents the label of its corresponding pixel point x_i, and all pixels x_i are connected together.

In order to solve the above problem, a fully connected CRF model based on the energy function of all pixel pairs on the image is proposed[21]. Although the fully-connected CRF model has been used for image semantic segmentation[22-25].But the computation complexity of the fully connected model limits its usage on hundreds of image regions that are segmented on the image. Therefore, the accuracy of segmentation is still limited by the accuracy of the unsupervised image segmentation algorithm that generates these regions. In order to solve this problem, the researchers connected all the pixel pairs in the image to achieve more fine segmentation and recognition, at the cost of an explosive increase in the amount of parameters. To overcome this problem, the pairwise energy function is defined as a linear combination of Gaussian kernels in

arbitrary feature space[21] and this CRF distribution is approximated by the mean field. The approximation is iteratively optimized through a series of message passing steps, each of which updates a single variable by fusing the information of all other variables. The update of the mean field can be achieved by Gaussian filtering in the feature space. The computational complexity of message passing can be reduced from quadratic to linear by employing efficient high-dimensional filtering approximations.

The fully-connected CRF model is defined as follows:

Suppose random field X, whose elements are $\{X_1, X_2, ..., X_n\}$, where the value range of X_i is the set of labels L:$\{L_1, L_2, ..., L_k\}$. Suppose another random domain I, whose elements are $\{I_1,...,I_n\}$. I is the information of the pixel on the image with the input size N, and X is the label of each pixel. I_j is the color vector of pixel j, and X_j is the label assigned to pixel j. Conditional random field (I, X) is determined by a Gibbs distribution:

$$P(X|I) = \frac{1}{Z(I)} \exp\left(-\sum_{c \in C_G} \emptyset_c(X_c|I)\right)$$

G=(v, e) in the above equation is an image to be labeled, and each pixel point pair c is an element in the set of pixel pair C_G on this image G, which contains an energy function φ_C The Gibbs energy function of the label $x \in L^N$ is:

$$E(x|I) = \sum_{c \in C_G} \emptyset_c(x_c|I)$$

The maximum posterior estimate of the conditional random field label is:

$$x^* = argmax_{x \in L^N} P(x|I)$$

For the convenience of representation, the latter part will use $\Psi_c(x_c)$ to represent $\varphi_c(x_c|I)$.

In the fully connected pixel-pair CRF model, G is a complete graph of the label X, C_G is the set of all unary and binary pixel groups, so the corresponding Gibbs energy function is

$$E(X) = \sum_i \Psi_u(x_i) + \sum_{i<j} \Psi_p(x_i, x_j)$$

The values of i and j in the above formula range from 1 to N. This unary energy function $\Psi_u(x_i)$ is calculated separately for each pixel by a classifier that generates a label assignment distribution function based on image features. The binary energy function is calculated as follows:

$$\psi_p(x_i, x_j) = \mu(x_i, x_j) \sum_{m=1}^K \omega^{(m)} k^{(m)}(f_i, f_j)$$

The $k^{(m)}$ in the formula is a Gaussian kernel function. Then we get:

$$k^{(m)}(f_i, f_j) = \exp\left(-\frac{1}{2}(f_i - f_j)^T \Lambda^{(m)}(f_i - f_j)\right)$$

The f_i and f_j in the formula are the feature vectors of the pixels i and j in the arbitrary feature space, $\omega^{(m)}$ is the linear combination weight, and μ is a label-compatible function. Label compatible function $\mu(x_i, x_j)=1$ if xi≠xj, otherwise its value is 0. Each Gaussian kernel $k^{(m)}$ is characterized by a symmetric positive definite matrix $\Lambda^{(m)}$. For image segmentation and identification problems, the energy functions of two kernels can be used. I_i and I_j in the following formula represent color vectors, and p_i and p_j represent the position of pixel points.

$$k(f_i, f_j) = \omega^{(1)} \exp\left(-\frac{|p_i - p_j|^2}{2\theta_\alpha^2} - \frac{|I_i - I_j|^2}{2\theta_\beta^2}\right) + \omega^{(2)}\exp(-\frac{|p_i - p_j|^2}{2\theta_\gamma^2})$$

The first half of the formula is the appearance kernel function. The neighboring pixels with the same color are more likely to be the same class of pixel. The degree of distance and color is controlled by the parameters θ_α and θ_β. The second part is the smooth kernel function, which is used to remove isolated small areas[11].

3. METHOD

The image segmentation and identification network based on the deep neural network outputs a score map and then get the pixel point semantic segmentation label. The score map is generally smooth, so applying short-range CRF for post-processing will have the opposite effect. In order to overcome the above problems, the researchers used the CRF model of [10] as a post-processing method. The score map has a score for each pixel point label classification, that is, the probability that each pixel point may be assigned a label value. This probability is calculated separately for each pixel, so it can be input as a unary energy function into the energy function of the fully-connected CRF:

$$\Psi_u(x_i) = -logP(x_i)$$

In the above formula, $P(x_i)$ is the pixel point assignment label probability output by the deep neural network based image segmentation and identification network. The binary energy function is still related to the color and position of the pixel, and the weight of the pixel color and position is modified by changing the relevant parameters. The fully-connected CRF uses the deep neural network based image segmentation and identification score map of the network output as input, and outputs the fine segmentation result after several fully-connected CRF iterations. Compared with the output of image segmentation and identification method based on the deep neural network, the details of the output of the post processing method based on fully-connected CRFs are clearer.

However, the single score map contains insufficient information on chronic wound images, and the results by the post processing method are poor. To solve the problem, we convert multiple scale feature maps into score maps, and superimpose multiple score maps as a multi-scale score map for the input of fully-connected CRF post-processing methods. The formula is as follows:

$$P(x_i) = \sum_{j=1}^{M} \varphi^{(j)}P_j(x_i)$$

In the above formula, $P_j(x_i)$ is the assigned label probability of pixel i on the jth score graph, and $\lambda^{(j)}$ is the weight of this probability value. The binary energy function remains unchanged, and the parameter weights are modified according to the characteristics of the chronic wound image.

In the optimization of the parameters, we use the method of from coarse to fine. The score map is superimposed using the last three layers. The sum of the three weights is 1. The initial value ratio is (2:2:6), and then the parameters are optimized. The update step is 0.05 until the optimal value is selected. According to the paper[10], the weight parameters ω_2 and θ_γ of the binary energy function are fixed to the default value of 3. The parameter ω_1 has a value range of [5, 10], and the update step is 1 each time. The parameter θ_α has a value range of [50, 100] and the update step size is 10. The parameter θ_β has a value range of [3, 10] and the update step size is 1.

4. EXPERIMENT

To verify the effectiveness of post-processing methods for chronic wound image segmentation and recognition based on fully connected CRF, we designed two comparison experiments: (1) Compare the post-processing results obtained by the post-processing method based on fully connected CRF with the single-layer score map and the original segmentation results of the deep neural network. (2) Compare the post-processing results obtained by the post-processing method based on fully connected CRF with the single-layer score graph and the post-processing results obtained by the post-processing method based on fully connected CRF with the multi-layer score graph. We use the original results of paper[26], and choose the best chronic wound image segmentation and identification network:MobileNet-0.75-skip-fcn16 as the chronic wound image segmentation and identification network to verify the effectiveness of our method.

EXPERIMENT ENVIRONMENT

We use an NVIDIA Geforce 1080Ti GPU to speed up parameter learning and evaluate the learned model on a computer with Intel Core i7-8700K CPU @ 3.70GHz and 32GB RAM. The program runs on a 64-bit windows10 home operating system with CUDA 9.0 and Tensorflow 1.7.0-GPU installed.

DATA SET

We use the data set that is built by the article[26]. The dataset is collected partly from cooperated medical institutions and partly from Medetec Wound Database. We made the size of the images a uniform resolution (512 by 512 pixels).

GROUND TRUTH.

The chronic wound image is manually selected for the chronic wound area and the background area. The red area on the ground truth is the chronic wound area, and the black area is the background.

The evaluation standard is as follows:
TP: the ground-truth is positive and the prediction is positive.
FN: the ground-truth is positive but the prediction is negative.
FP: the ground-truth is negative but the prediction is positive.
TN: the ground-truth is negative and the prediction is negative.
We use accuracy, mean intersection-over-union (mIoU), and dice similarity coefficient (DSC) to compare the result. They are computed as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}$$

$$\text{mIoU} = \frac{TP}{TP + FP + FN}$$

$$DSC = \frac{2 \times TP}{2 \times TP + FP + FN}$$

5. RESULTS

The figure below shows a comparison of post-processing results obtained by the post-processing method based on fully connected CRF with the single-layer score map and the original segmentation results of the deep neural network. The first column(A) on the left are the original images, the second column(B) are the ground truth, the third column(C) are the Mobilenet origin results, and the fourth column(D) images are images obtained by merging the first and the third columns of images, which is convenient for viewing the segmentation effect. The fifth column(E) are the results obtained by using the post-processing method with single-layer score map and the sixth column(F) are the images obtained by merging the first and the fourth column images. Tt can be found that the image processed by the post-processing method with single-layer score map has a certain improvement in the edge detail compared with the original segmentation result. However, it could cause anti-effect such as first and third lines of the images.
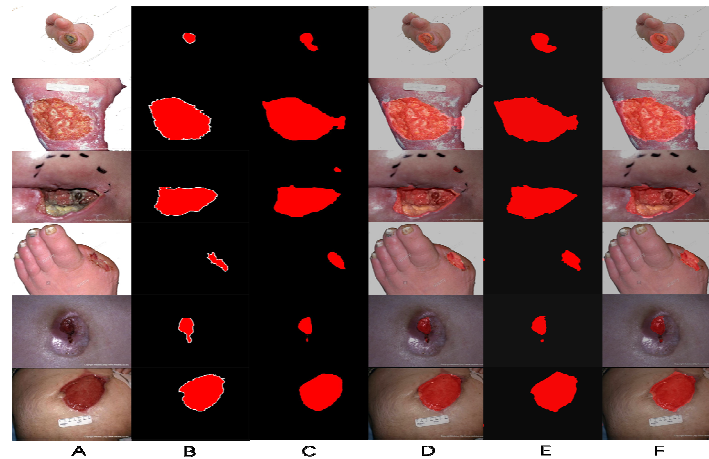

Figure 2. Comparison of the results of neural networks and post-processing method with single-layer score map
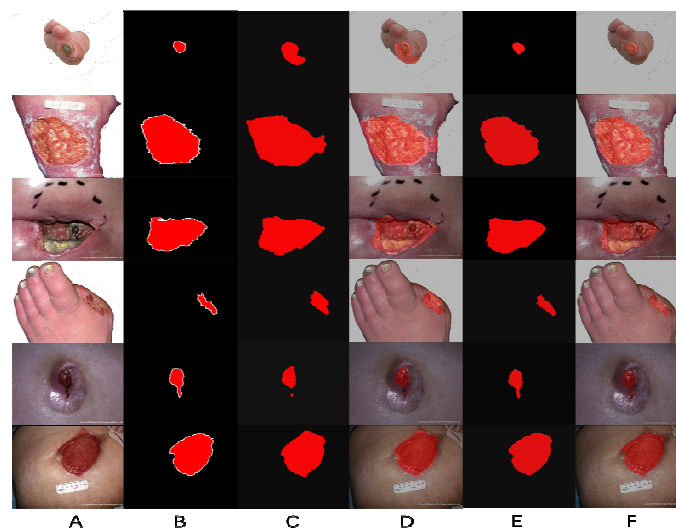

Figure 3. Comparison of the results of post-processing method with single-layer score map and post-processing method with multi-layer score map

The first column(A) on the left are the original images, the second column(B) are the ground truth, the third column(C)(also E in figure 2) are the results obtained by using the post-processing method with single-layer score map, and the fourth column(D)(also F in figure 2) images are images obtained by merging the first and the third columns of images, which is convenient for viewing the segmentation effect. The fifth column(E) are the results obtained by using the post-processing method with multi-layer score map and the sixth column(F) are the images obtained by merging the first and the fourth column images.

Compared with the two results in the above figure, we can find that it is better to use the post-processing method with multi-score map to process the original segmentation image. It can also correct some misclassifications and make the contours of chronic wound areas more detailed. The table below compares the accuracy of the two methods.

Table 1. Accuracy comparison

Method	Accuracy (%)	mIoU (%)	DSC (%)
Origin Results	98.26	85.76	92.33
Single-layer score map	98.23	85.80	92.35
Multi-layer score map	**98.36**	**86.08**	**92.52**

The comparison results of the show that the post-processing method with multi-layer score map is better than the post-processing method with single-layer score map and achieves the highest of the three indexes. That proves the validity of our method.

6. CONCLUSION

In this paper, we first introduce the CRF algorithm and its improved version of the fully connected CRF algorithm. Then we innovatively combined multi-scale score map with the fully connected CRF algorithm and propose a chronic wound image segmentation and identification post-processing method based on fully connected CRF. This post-processing method overcomes the insufficient information using the single score map as the input of the fully connected CRF, and thus can get better results. The post-processing method is divided into two steps: Firstly, we calculate the score maps of the corresponding last three layers of the feature maps, then combine the three-layer score maps as the input of the fully connected CRF, and refine the parameters to obtain the best results. The results of the experiments have proved that the post-processing method based on fully connected CRF with multi-layer score map can optimize the origin segmentation results obtained by deep convolution neural networks.

ACKNOWLEDGEMENTS

REFERENCES

[1]   A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks." in NIPS, 2012.

[2]   Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv: 1409.1556 (2014).

[3]   Szegedy, Christian, et al. "Going deeper with convolutions." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015.

[4]   He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.

[5]   Huang, Gao, et al. "Densely connected convolutional networks." arXiv preprint arXiv:1608.06993 (2016).

[6]   Sabour, Sara, Nicholas Frosst, and Geoffrey E. Hinton. "Dynamic routing between capsules." Advances in Neural Information Processing Systems. 2017.

[7]   V. Shukla, M. A. Ansari, and S. Gupta, "Wound healing research: a perspective from india," International Journal of Lower Extremity Wounds, vol. 4, no. 1, pp. 7–9, 2005.

[8]   C. K. Sen, G. M. Gordillo, S. Roy, R. Kirsner, L. Lambert, T. K. Hunt, F. Gottrup, G. C. Gurtner, and M. T. Longaker, "Human skin wounds: a major and snowballing threat to public health and the economy," Wound Repair and Regeneration, vol. 17, no. 6, pp. 763–771, 2009.

[9]   J. Posnett and P. Franks, "The burden of chronic wounds in the uk," Diabetic Medicine, vol. 14, no. 5, pp. S7–S85, 2008.

[10]   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 transactions on pattern analysis and machine intelligence, 2018, 40(4): 834-848.

[11]   Shotton J, Winn J, Rother C, et al. Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context[J]. International Journal of Computer Vision, 2009, 81(1): 2-23.

[12]   Triggs B, Verbeek J J. Scene segmentation with crfs learned from partially labeled images[C]//Advances in neural information processing systems. 2008: 1553-1560.

[13]   Gould S, Rodgers J, Cohen D, et al. Multi-class segmentation with relative location prior[J]. International Journal of Computer Vision, 2008, 80(3): 300-316.

[14]   Fulkerson B, Vedaldi A, Soatto S. Class segmentation and object localization with superpixel neighborhoods[C]//Computer Vision, 2009 IEEE 12th International Conference on. IEEE, 2009: 670-677.

[15]   He X, Zemel R S, Carreira-Perpiñán M Á. Multiscale conditional random fields for image labeling[C]//Computer vision and pattern recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE computer society conference on. IEEE, 2004, 2: II-II.

[16]   Kumar S, Hebert M. A hierarchical field framework for unified context-based classification[C]//null. IEEE, 2005: 1284-1291.

[17]   Kohli P, Torr P H S. Robust higher order potentials for enforcing label consistency[J]. International Journal of Computer Vision, 2009, 82(3): 302-324.

[18]   Russell C, Kohli P, Torr P H S. Associative hierarchical crfs for object class image segmentation[C]//Computer Vision, 2009 IEEE 12th International Conference on. IEEE, 2009: 739-746.

[19]   Ladicky L, Russell C, Kohli P, et al. Graph cut based inference with co-occurrence statistics[C]//European Conference on Computer Vision. Springer, Berlin, Heidelberg, 2010: 239-253.

[20]   Zheng S, Jayasumana S, Romera-Paredes B, et al. Conditional Random Fields as Recurrent Neural Networks[J]. 2015:1529-1537.

[21] Krähenbühl P, Koltun V. Efficient inference in fully connected crfs with gaussian edge potentials[C]//Advances in neural information processing systems. 2011: 109-117.

[22] Rabinovich A, Vedaldi A, Galleguillos C, et al. Objects in context[C]//Computer vision, 2007. ICCV 2007. IEEE 11th international conference on. IEEE, 2007: 1-8.

[23] Toyoda T, Hasegawa O. Random field model for integration of local information and global information[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(8): 1483-1489.

[24] Galleguillos C, Rabinovich A, Belongie S. Object categorization using co-occurrence, location and appearance[C]//Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on. IEEE, 2008: 1-8.

[25] Payet N, Todorovic S. ^ 2--Random Forest Random Field[C]//Advances in Neural Information Processing Systems. 2010: 1885-1893.

[26] Liu X, Wang C, Li F, et al. A framework of wound segmentation based on deep convolutional networks[C]//Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2017 10th International Congress on. IEEE, 2017: 1-7.