IMAGE PROCESSING, PATTERN RECOGNITION

Skin lesion segmentation method for dermoscopic images with convolutional neural networks and semantic segmentation

Dang N.H. Thanh [1], Nguyen Hoang Hai [2], Le Minh Hieu [3], Prayag Tiwari [4], V.B. Surya Prasath [5,6,7,8]
[1] Department of Information Technology, School of Business Information Technology,
University of Economics Ho Chi Minh City, Vietnam,
[2] Faculty of Computer Science, Vietnam-Korea University of Information and Communication Technology
– The University of Danang, Vietnam,
[3] Department of Economics, University of Economics, University of Danang, Vietnam,
[4] Department of Information Engineering, University of Padua, Italy,
[5] Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH USA,
[6] Department of Pediatrics, University of Cincinnati, OH USA,
[7] Department of Biomedical Informatics, College of Medicine, University of Cincinnati, OH USA,
[8] Department of Electrical Engineering and Computer Science, University of Cincinnati, OH USA

Abstract

Melanoma skin cancer is one of the most dangerous forms of skin cancer because it grows fast and causes most of the skin cancer deaths. Hence, early detection is a very important task to treat melanoma. In this article, we propose a skin lesion segmentation method for dermoscopic images based on the U-Net architecture with VGG-16 encoder and the semantic segmentation. Base on the segmented skin lesion, diagnostic imaging systems can evaluate skin lesion features to classify them. The proposed method requires fewer resources for training, and it is suitable for computing systems without powerful GPUs, but the training accuracy is still high enough (above 95 %). In the experiments, we train the model on the ISIC dataset – a common dermoscopic image dataset. To assess the performance of the proposed skin lesion segmentation method, we evaluate the Sorensen-Dice and the Jaccard scores and compare to other deep learning-based skin lesion segmentation methods. Experimental results showed that skin lesion segmentation quality of the proposed method are better than ones of the compared methods.

Keywords: image segmentation, medical image segmentation, semantic segmentation, melanoma, skin cancer, skin lesion, deep learning, cancer.

Introduction

Melanoma skin cancer is one of the most dangerous forms of skin cancer. It grows fast and causes most of the skin cancer deaths. For cancer in general and skin cancer in particular, early detection is a very important task, because doctors can help to stop the metastatic – one of the most popular causes of cancer death. One important method for diagnosing melanoma is the ABCD rule [1, 2]. To improve the diagnostic quality by ABCD rule, it is necessary to segment skin lesions from dermoscopic images. Based on the segmented region, features of skin lesions will be extracted to evaluate the lesion.

The skin lesion segmentation problem plays an important role in medical image processing. Several methods were studied, including learning-based [3, 4] and non-learning-based approaches such as thresholding and level set methods [2, 5]. In this paper, we mainly focus on learning-based methods that become a hot research trend.

In recent years, deep learning is an efficient approach to solve image processing problem, including image segmentation. In that, the artificial neural networks (ANNs) [6] and convolutional neural networks (CNNs) [3] became the most powerful tools in image processing, pattern recognition, computer vision, and other fields of science, engineering, and technology [3]. CNNs are applied to solve many medical image segmentation problems such as segmentation of tumors, human organs, brain, and bone.

For skin lesion segmentation, there are several methods based on CNNs such as the method with fully convolutional-deconvolutional networks [3], the method using deep fully convolutional networks with Jaccard distance [4], and the method based on multistage fully convolutional networks [7]. All these methods are developed

based on the fully convolutional networks (FCNs). Many works denoted that training on FCNs is complicated, and FCNs are not sensitive enough for segmenting small details and low-intensity regions as in the case of skin lesion [8]. Moreover, FCNs typically require a large amount of training data.

Some other models based on CNN such as the high-resolution CNN [9] and combined deep convolution networks and unsupervised learning [10] are also proposed for segmenting skin lesions. However, the accuracy of the methods for skin lesion segmentation, especially for low-density regions of skin lesions is not high. The skin lesion segmentation methods based on dense deconvolution networks [11, 12] were proposed. Although these methods are good enough for skin lesion segmentation, they cannot reliably segment low-intensity regions. Some other skin lesion segmentation methods based on CNN were proposed [13, 14, 15]. However, the methods cannot work on colorful images directly. Therefore, we need to convert dermoscopic images to grayscale images, process on separate channels, or normalize colors. Some methods are only applied after skin lesion images were preprocessed, such as remove hair, extract regions of interest (ROI), remove shadow and shading effects, etc. One of the effective CNN-based architectures for medical image segmentation is U-Net [16, 8]. It is effective to segments the image by pixels instead of a whole. Therefore, it will improve accuracy for medical image segmentation. Moreover, U-Net-based methods usually require less memory for training with a small training set. Rashika et al. proposed a skin lesion segmentation method based on U-Net [16]. However, the method is not effective to segment small details and narrow bands. In other words, the method of Rashika is not effective for skin lesions with low-intensity regions.

To improve the accuracy of segmenting low-intensity regions of skin lesions, we combine U-Net with VGG-16 architecture. VGG-16 is one of the most effective CNN architecture for semantic segmentation and it can work well on low-intensity pixels [17]. In the proposed CNN architecture, VGG-16 will be used for replacing the encoder. Therefore, the proposed skin lesion segmentation method will work more effectively, even with low-intensity regions. In this paper, we propose a method to segment skin lesion with a CNN architecture based on VGG-16 encoder [18, 19] for U-Net and semantic segmentation method. The proposed method does not require a large training data because it utilizes the advantages of U-Net. Therefore, it suits to computing systems without powerful GPUs. Moreover, the method works directly on colorful images and does not require any preprocessing tasks. It can segment low-intensity regions because of a combination of VGG-16 encoder. Further, the proposed method utilizes the advantages of the semantic segmentation method [20]. Semantic segmentation plays a vital role in the field of computer vision. It is one of the high-level tasks that paves the way towards complete scene understanding.

The rest of the article is organized as follows. Section II presents the proposed skin lesion segmentation method for dermoscopic images with convolutional neural networks and semantic segmentation. Section III presents experimental results and the comparison. Finally, Section IV concludes the article.

The proposed skin lesion segmentation method

Convolutional neural network

In the field of deep learning, the convolutional neural networks are powerful and effective tools to process many problems in science, engineering, technology, business, management, and medicine. CNNs are a subclass of deep neural networks that are widely used for analyzing visual imagery. CNNs use a variation of multilayer perceptrons designed to process. They are stimulated by biological processes of connectivity of the neurons of humans and animals.

For the skin lesion segmentation problem, some approaches of deep learning were used: method using convolutional-deconvolutional networks, the method using fully connected convolutional networks, the method based on deep residual networks, etc. In the proposed method, we combine with the ImageNet network [19] to design the CNN architecture to process the skin lesion segmentation problem.

Before proposing the CNN architecture for the skin lesion segmentation problem, we explain some terminologies related to the proposed CNN architecture:

Image input layer [19]: inputs images to a network and applies data normalization.

Convolution 2D layer [21]: applies sliding-convolutional filters to the input image. This layer convolves the input image by moving the filter along the input image by directions (vertical or horizon) and computes the dot product of the weights and the input image. Finally, a bias term is added. In the convolution layer, there are various components including filters and stride, dilated convolutions, feature maps, zero paddings, output size, number of neurons, learnable parameters, and number of layers. The learnable parameters will be updated during network training.

Batch normalization layer [22]: normalizes each input channel x_i of the input image x across a mini-batch. The first layer normalizes the activations $\hat{x}_i$ of each channel by subtracting the mini-batch mean μ_B and dividing by the mini-batch standard deviation σ_B^2. Then, the layer shifts the input by a learnable offset β and scales it by a learnable scale factor γ:

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}}, \; y_i = \gamma \hat{x}_i + \beta.$$

The learnable parameters β and γ will be updated during network training.

ReLU layer [23]: performs a threshold operation to each element of the input image x, where any values less than zero is set to zero:

$$f(x_i) = \begin{cases} x_i, & x_i \geq 0 \\ 0, & x_i < 0 \end{cases}.$$

The ReLU layer does not change the input image size.

Max pooling 2D layer [24]: performs down-sampling by dividing the input image into rectangular pooling regions and computing the maximum of each region.

Max unpooling 2D layer [24]: unspools the output of the max pooling layer.

Softmax layer [25]: apply the softmax function to the input image x. The softmax function has the following form:

$$y_r(x_i) = \frac{e^{a_r(x_i)}}{\sum\limits_{j=1}^{k} e^{a_j(x_i)}}, r = 1, ..., k, \quad \sum\limits_{j=1}^{k} y_j = 1 \quad ,$$

where $0 \leq y_r \leq 1$ and $a = (a_1, ..., a_k)$ is a *k*-dimensional vector of arbitrary real values.

Classification output layer [25]: computes the cross-entropy loss for multi-class classification problems with mutually exclusive classes. The classification output layer usually follows the softmax layer.

Skin lesion dataset

For the dataset of dermoscopic skin lesions images, we use the International Skin Imaging Collaboration (ISIC) 2017: *https://www.isic-archive.com*. The total size of the dataset is 5.4GB and it includes about 2000 dermoscopic images with given ground truth segmented by dermoscopic experts and super-pixel masks. The ID of images has the form: ISIC_00xxxxx. The ground truth for segmentation has the same ID as the corresponding image ISIC_00xxxxx_segmentation. Moreover, it also provides a validation set with 150 images and a test set with 600 images. All dermoscopic images of the dataset are stored in RGB-colour and the JPEG format. Ground truth and super-pixel masks are stored in the PNG format.
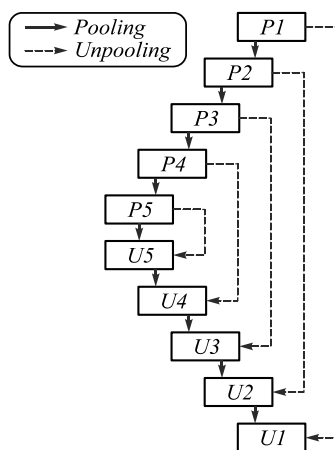

Fig. 1. The general proposed CNN architecture

CNN architecture for skin lesion segmentation

The proposed CNN architecture for skin lesion segmentation includes five max pooling layers (and five max unpooling layers). General CNN architecture is presented in Figure 1.

In fig. 1, the blocks P1-P5 use max pooling layers and the blocks U1-U5 use max unpooling layers.

The order of layers for block P1 is an image input layer, convolution 2D layer, batch normalization layer, ReLU layer, and max pooling 2D layer. For block P2 – convolution 2D layer, batch normalization layer, ReLU layer, convolution 2D layer, batch normalization layer, ReLU layer, max pooling 2D layer. For blocks P3, P4 and P5 – convolution 2D layer, batch normalization layer, ReLU layer, convolution 2D layer, batch normalization layer, ReLU layer, convolution 2D layer, batch normalization layer, ReLU layer, max pooling 2D layer.

The order of layers for blocks U5, U4, U3 – max unpooling 2D layer, convolution 2D layers, batch normalization layer, ReLU layer, convolution 2D layers, batch normalization layer, ReLU layer, convolution 2D layers, batch normalization layer, ReLU layer. For block U2 – max unpooling 2D layer, convolution 2D layer, batch normalization layer, ReLU layer, convolution 2D layer, batch normalization layer, ReLU layer. For block U1 – max unpooling 2D layer, convolution 2D layer, batch normalization layer, ReLU layer, softmax layer, and classification output layer (pixel classification layer).

As we mentioned above, the proposed CNN architecture used VGG-16 encoder with the U-Net architecture. The algorithm for skin lesion segmentation method for dermoscopic images with semantic segmentation and CNN is presented in Algorithm 1.

Algorithm 1. The skin lesion segmentation method for dermoscopic images with semantic segmentation and convolutional neural network.

> **Input**: The input dermoscopic images of skin lesion *v*.
> **Output**: The segmented skin lesion image *u*.
>
> **Function** *u=SemanticSegCNN(v)*
>   *Step 1*: Pretrain the proposed model on a training set.
>   *Step 2*: Implement the semantic segmentation method.
>   *Step 3*: Improve segmentation quality:
>   – Apply Gaussian filter.
>   – Fill the holes and filter out small segments.
>   – Compute the scores to assess the segmentation quality of the proposed method.
> **End**.

To suit to weak computing systems, in the first step, we only need to implement the training process on a small number of patterns of a dataset.

We implement the steps outlined in Algorithm 1 for improving the deep learning model for accurate delineation of skin lesions in dermoscopic images. After pretraining data, we implement the semantic segmentation method proposed by Brostow et al. [20]. The detail of the semantic segmentation method in the second step can be found in the work [20].

We note that to implement the semantic segmentation method in MATLAB, we use *semanticseg* function. To apply Gaussian filter after segmentation, we use *imgauss-filt* function, and to fill the holes and filter out small segments, we use *imfill* function with default settings.

Experimental results and discussions

We implement the training process and the proposed skin lesion segmentation method on MATLAB 2018b. The configuration of the computing system is Windows 10 Pro with Intel Core i5, 1.6 GHz, 4GB 2295 MHz DDR3 RAM without GPU. If the training process is implemented on computing systems with powerful GPU, performance will be better. Our proposed method suits for both training on CPU and GPU.

Image segmentation quality assessment metrics

For the segmentation task, we use the Sorensen-Dice and the Jaccard metrics to assess quality [26, 27].

Let consider that X is segmented regions that we need to assess the quality, Y is the corresponding ground truth.

Sorensen-Dice similarity [28] is computed as follows:

$$dice(X,Y) = \frac{2\,|\,X \cap Y\,|}{|\,X\,| + |\,Y\,|},$$

where $|\cdot|$ denotes the set cardinality (the number of elements of a set). The value of the Sorensen-Dice similarity metric is between 0 and 1 (or 0% to 100%). The higher the Sorensen-Dice value, the better the segmentation result.

Jaccard similarity [28, 27] related to the Sorensen-Dice similarity:

$$jaccard(X,Y) = \frac{dice(X,Y)}{2 - dice(X,Y)}.$$

The range of the Jaccard value is in [0, 1]. The higher the Jaccard value, the better the segmentation result.

Accuracy (%) [26] measures how well a binary segmentation method correctly identifies or excludes a condition:

$$accuracy = \frac{TP + TN}{FP + FN + TP + TN},$$

where *TP*, *TN*, *FP*, *FN* denote true positive, true negative, false positive, and false negative, respectively. They are basic statistical indices that can be evaluated based on predicted segmentation results and the ground truth. Note that the ground truth of the ISIC dataset is handled by experienced dermatologists.

Sensitivity (%) [26] is evaluated as the proportion of real positives that are correctly identified:

$$sensitivity = \frac{TP}{TP + FN},$$

Specificity (%) [26] is computed as the proportion of real negatives that are correctly identified:

$$specificity = \frac{TN}{TN + FP},$$

We also note that Dice similarity and Jaccard similarity can be evaluated via *TP*, *TN*, *FP*, *FN*:

$$dice = \frac{2TP}{2TP + FP + FN},$$

To evaluate Dice and Jaccard scores, we need the ground truth. All ground truths are given in the ISIC dataset. These ground truths were segmented by experienced dermatologists.

Image Datasets

We use the dermoscopic images of skin lesions of the ISIC dataset for the challenge of 2017. All images are of high definition. To suit processing on our computing system, we resize all images to standard size 256×256 pixels and store them in PNG format.

We select 20 images to present the visual results of the proposed skin lesion segmentation method. Moreover, the acquired results for a test set are used for comparison.

Fig. 2 shows all selected images (20 images) for testing. All images used for the test are colorful images. Our proposed method works directly on colorful images without converting to grayscale or extracting separate channels.

All selected images used for the segmentation task are original without any preprocessing tasks. We only resize dermoscopic images to standard size 256×256 pixels. As can be seen in Fig. 2, these images include shading effects, hairs and color intensity of skin lesion is different together varying from low-intensity to high-intensity.

For the training process, we use the proposed CNN architecture with the stochastic gradient descent with momentum [29]. Otherwise, we use the configuration: initial learning rate is 0.001, max epochs – 200, mini-batch size – 32. The number of iterations is the same as max epochs.

The training accuracy is presented in Figure 3. As can be seen, with iteration under 90, accuracy increases very fast. After that, it increases slower. With the above setting, the accuracy of our training result is 95.79 % after 200 iterations. This is a very good result, especially for the small size of the training data.

The segmented results by the proposed method are presented in Fig. 4 – 5. Fig. 4 is for black-white segmentation. The white region denotes the skin lesion. Figure 5 presents segmented results on the dermoscopic images of skin lesions. We must notice that the red border region is segmented by our proposed method, the green border region – by the given ground truth (is segmented by experienced dermatologists). As can be seen, the difference is very small.

The proposed skin lesion segmentation method has some advantages: although we only train on small data, the proposed method still works effectively; it can work directly on colorful images and no need to convert to grayscale or process on separate channels; no need to remove hairs; no need to extract the regions of interest

(ROI). Otherwise, the proposed method can segment low-intensity regions of skin lesions well. The other considered methods usually use image enhancement algorithms before processing.
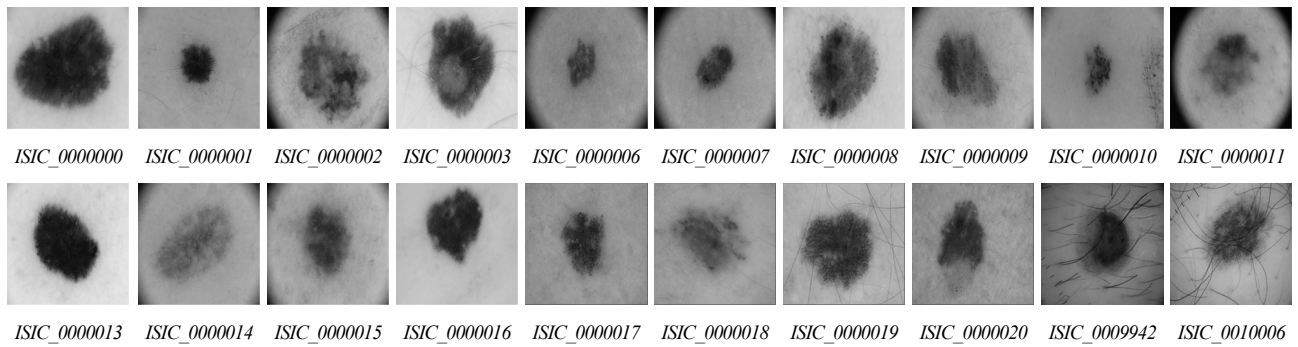

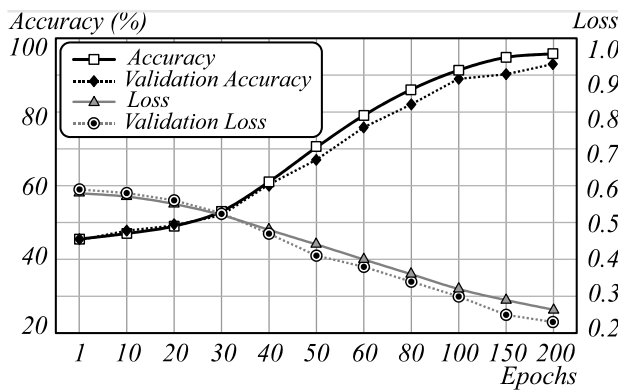*Fig. 2. The dermoscopic images of skin lesions for the test*

Fig. 3. The change of the training accuracy and the training loss by epochs. The final accuracy is 95.79%, the final loss is 0.26. Training with proposed CNN architecture

Tab. 1 shows Accuracy, Sorensen-Dice, Jaccard, Sensitivity, and Specificity scores for a set of 20 selected images from the ISIC dataset segmented by the proposed skin lesion segmentation method. Almost of Sorensen-Dice scores are higher than 0.9 and most of Jaccard scores are higher than 0.8. This is an impressive result. The average Dice score is 0.92 and the average Jaccard score is 0.86.

Tab. 2 presents a comparison of Accuracy, Dice, Jaccard, Sensitivity, and Specificity scores of the proposed method with other skin lesion segmentation methods based on deep learning. As can be seen that Accuracy, Dice, and Jaccard scores of the segmented result by the proposed method are the highest. Hence, our proposed method can compete with other state-of-the-art methods for skin lesion segmentation.
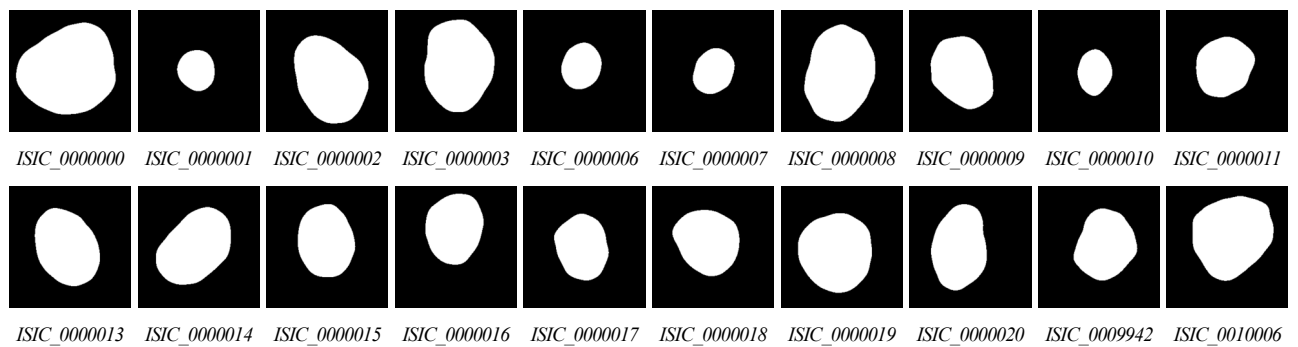

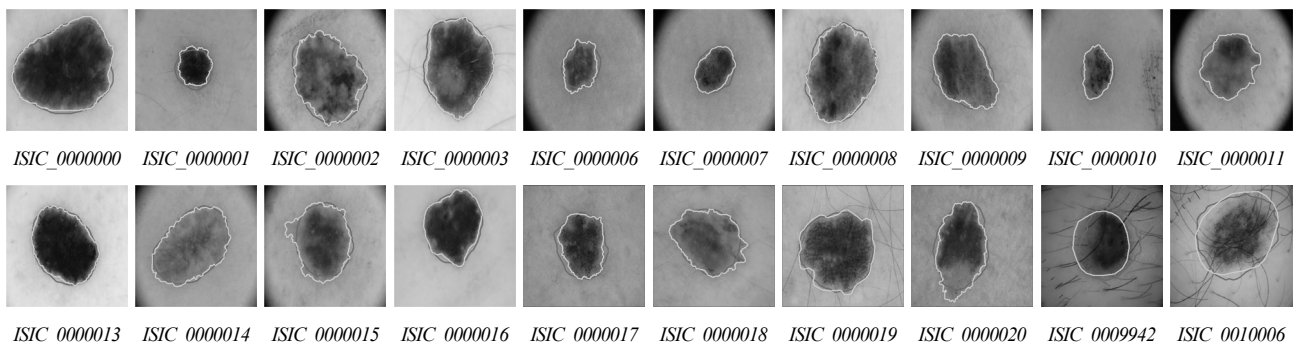*Fig. 4. The segmented results of skin lesions in black-white mode*

Fig. 5. The segmented skin lesions of dermoscopic images: the red (dark) border – our segmented result, the green (light) border – the ground truth

Table 1. Accuracy, Dice, Jaccard, Sensitivity, and Specificity scores of segmented results
by the proposed method on 20 selected dermoscopic images of skin lesions

IDs	Accuracy	Dice	Jaccard	Sensitivity	Specificity
ISIC_0000000	0.965	0.962	0.928	0.966	0.964
ISIC_0000001	0.978	0.854	0.745	0.787	0.995
ISIC_0000002	0.961	0.939	0.885	0.906	0.988
ISIC_0000003	0.967	0.951	0.907	0.943	0.980
ISIC_0000006	0.979	0.888	0.799	0.852	0.993
ISIC_0000007	0.983	0.909	0.833	0.852	0.998
ISIC_0000008	0.972	0.959	0.922	0.938	0.991
ISIC_0000009	0.967	0.928	0.866	0.921	0.981
ISIC_0000010	0.982	0.889	0.800	0.878	0.991
ISIC_0000011	0.972	0.922	0.855	0.908	0.986
ISIC_0000013	0.981	0.962	0.926	0.938	0.996
ISIC_0000014	0.976	0.956	0.916	0.951	0.985
ISIC_0000015	0.959	0.911	0.837	0.942	0.964
ISIC_0000016	0.976	0.944	0.894	0.922	0.991
ISIC_0000017	0.965	0.895	0.810	0.819	0.998
ISIC_0000018	0.972	0.940	0.886	0.955	0.978
ISIC_0000019	0.973	0.957	0.918	0.974	0.973
ISIC_0000020	0.955	0.907	0.829	0.887	0.977
ISIC_0009942	0.929	0.829	0.708	0.778	0.972
ISIC_0010006	0.952	0.931	0.872	0.931	0.964
AVERAGE	0.968	0.922	0.857	0.902	0.983

Table 2. Accuracy, Dice, Jaccard, Sensitivity, and Specificity scores of the proposed method
and the other skin lesion segmentation methods based on deep learning for the test set

Method	Accuracy	Dice	Jaccard	Sensitivity	Specificity
Yuan [3]	0.934	0.849	0.765	0.825	0.975
Berseth [30]	0.932	0.847	0.762	0.820	0.978
Bi et al. [6]	0.934	0.844	0.760	0.802	**0.985**
Rashika et al. [16]	0.928	0.868	0.842	**0.930**	0.842
Li et al. [15]	0.950	0.839	0.753	0.855	0.974
Proposed	**0.967**	**0.915**	**0.846**	0.904	0.980

For execution speed, it takes around 8 hours to complete the training process. This is a normal result to compare to other training methods. The training task is always heavy and takes a lot of time. With given pretrained data, our proposed method only takes less than 1 second to complete the segmentation task.

Conclusions

In this paper, we proposed a CNN architecture for skin lesion segmentation for dermoscopic images based on convolutional neural networks and a skin lesion segmentation method based on that proposed CNN architecture and semantic segmentation. Our proposed method can work effectively even with a small size of training data. It gives a very good and impressive result, but no require any preprocessing tasks, such as hair removal, ROI extraction, image enhancement, etc. Our proposed method is good enough to compare to other state-of-the-art methods for skin lesion segmentation.

In future work, we can apply some preprocessing tasks such as image inpainting algorithms [31, 32] to remove hair before applying the segmentation task. This is necessary to increase the accuracy of both training tasks and the segmentation task.

References

[1] Kunz M, Stolz W. ABCD rule. Source: ⟨https://dermoscopedia.org/ABCD_rule⟩.

[2] Thanh DNH, Erkan U, Prasath VBS, Kumar V, Hien NN. A skin lesion segmentation method for dermoscopic images based on adaptive thresholding with normalization of color models. IEEE 2019 6th International Conference on Electrical and Electronics Engineering 2019: 116-120.

[3] Yuan Y, Lo YC. Improving dermoscopic image segmentation with enhanced convolutional-deconvolutional networks. IEEE J Biomed Health Inform 2017; 23(2): 519-526.

[4] Yuan Y, Chao M, Lo Y-C. Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance. IEEE Trans Med Imaging 2017; 36(9): 1876-1886.

[5] Thanh DNH, Hien NN, Prasath VBS, Thanh LT, Hai NH. Automatic initial boundary generation methods based on edge detectors for the level set function of the chan-vese segmentation model and applications in biomedical image processing. In Book: Satapathy SC, Bhateja V, Nguyen BL, Nguyen NG, Le D-N, eds. Frontiers in intelligent computing: theory and applications. Singapore: Springer: 2020: 171-181.

[6] Bi L, Kim J, Ahn E, Feng D. Automatic skin lesion analysis using large-scale dermoscopy images and deep residual networks. arXiv preprint 2017. Source: ⟨https://arxiv.org/abs/1703.04197⟩.

[7] Bi L, Kim J, Ahn E, Kumar A, Fulham M, Feng D. Dermoscopic image segmentation via multistage fully convolutional networks. IEEE Trans Biomed Eng 2017; 64(9): 2065-2074.

[8] Chen Y, Cao Z, Cao C, Yang J, Zhang J. A modified U-Net for Brain MR image segmentation. In Book: Sun X, Pan Z, Bertino E, eds. Cloud computing and security. Cham: Springer; 2018: 233-242.

[9] Xie F, Yang J, Liu J, Jiang Z, Zheng Y, Wang Y. Skin lesion segmentation using high-resolution convolutional neural network. Comput Methods Programs Biomed 2020; 186: 105241.

[10] Bozorgtabar B, Sedai S, Roy PK, Garnavi R. Skin lesion segmentation using deep convolution networks guided by local unsupervised learning. IBM J Res Dev 2017; 61(4-5): 6.

[11] Li H, He X, Zhou F, Yu Z, Ni D, Chen S, Wang T, Lei B. Dense deconvolutional network for skin lesion segmentation. IEEE J Biomed Health Inform 2019; 23(2): 527-537.

[12] He X, Yu Z, Wang T, Lei B, Shi Y. Dense deconvolution net: Multi path fusion and dense deconvolution for high resolution skin lesion segmentation. Technol Health Care 2018; 26: 307-316.

[13] Goyal M, Oakley A, Bansal P, Dancey D, Yap MH. Skin lesion segmentation in dermoscopic images with ensemble deep learning methods. IEEE Access 2020; 8: 4171-4181.

[14] Zafar K, Gilani SO, Waris A, Ahmed A, Jamil M, Khan MN, Kashif AS. Skin lesion segmentation from dermoscopic images using convolutional neural network. Sensors 2020; 20(6): 1601.

[15] Li Y, Shen L. Skin lesion analysis towards melanoma detection using deep learning network. Sensors 2018; 18(2): 556.

[16] Mishra R, Daescu O. Deep learning for skin lesion segmentation. IEEE International Conference on Bioinformatics and Biomedicine (BIBM) 2017: 1189-1194.

[17] Yao Y, Luo Z, Li S, Fang T, Quan L. MVSNet: Depth inference for unstructured multi-view stereo. In Book: Ferrari V, Hebert M, Sminchisescu C, Weiss Y, eds.Computer Vision – ECCV 2018. Cham: Springer; 2018: 785-801.

[18] Karen S, Zisserman A. Very deep convolutional networks for large-scale image recognition. International Conference on Learning Representations (ICLR-2015) 2015.

[19] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM 2017; 60(6): 84-90.

[20] Brostow GJ, Julien F, Roberto C. Semantic object classes in video: a high-definition ground truth database. Patt Recogn Lett 2008; 30(2): 88-97.

[21] Le Cun Y, Boser B, Denker J, Henderson D, Howard RE, Hubbard W, Jackel L. Handwritten digit recognition with a back-propagation network. In Book: Touretzky DS, ed. Advances in neural information processing systems 2. San Francisco: Morgan Kaufmann; 1990: 396-404.

[22] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. Proceedings of the 32nd International Conference on International Conference on Machine Learning (ICML'15) 2015: 448-456.

[23] Vinod N, Hinton GE. Rectified linear units improve restricted boltzmann machines. Proceedings of the 27th international conference on machine learning (ICML-10) 2010: 807-814.

[24] Nagi J, Ducatelle F, Di-Caro GA, Ciresan D, Meier U, Giusti A, Nagi F, Schmidhuber J, Gambardella LM. Max-pooling convolutional neural networks for vision-based hand gesture recognition. IEEE International Conference on Signal and Image Processing Applications (ICSIPA2011) 2011: 342-347.

[25] Bishop CM. Pattern recognition and machine learning, New York: Springer, 2006.

[26] Csurka G, Larlus D, Perronnin F. What is a good evaluation measure for semantic segmentation. The British Machine Vision Conference 2013.

[27] Thanh DNH, Prasath VBS, Hieu LM, Hien NN. Melanoma skin cancer detection method based on adaptive principal curvature, colour normalisation and feature extraction with the ABCD rule. J Digit Imaging 2020; 33: 574-585.

[28] Abdel AT, Allan H. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. BMC Medical Imaging 2015; 15: 1-29.

[29] Qian N. On the momentum term in gradient descent learning algorithms. Neural Netw 1999; 12(1): 145-151.

[30] Berseth M. ISIC 2017 – Skin lesion analysis towards melanoma detection. arXiv preprint 2017. Source: ⟨https://arxiv.org/abs/1703.00523⟩.

[31] Thanh DNH, Prasath VBS, Son NV, Hieu LM. An adaptive image inpainting method based on the modified mumford-shah model and multiscale parameter estimation. Computer Optics 2019; 43(2): 251-257. DOI: 10.18287/2412-6179-2019-43-2-251-257.

[32] Thanh DNH, Prasath VBS, Dvoenko S, Hieu LM. An adaptive image inpainting method based on euler's elastica with adaptive parameters estimation and the discrete gradient method. Signal Process 2021; 178: 107797.

Authors' information

Dang Ngoc Hoang Thanh (b. 1986) graduated from Belarusian State University in 2008 and MSc in 2009 majoring in Applied Mathematics, graduated Ph.D. of Computer Science (2016) from Tula State University, Russia. He was a lecturer at Hue College of Industry. Currently, he works as an assistant professor at the Information Technology Department, School of Business Information Technology, University of Economics Ho Chi Minh City, Vietnam. He is a member of scientific organization INSTICC (Portugal), ACM (USA), IAENG (Taiwan) and is also a member of international conferences committee, such as IEEE ICCE 2018 (Vietnam), IWBBIO 2018 (Spain), IEEE CoCoNet (Kazakhstan), ICMTEL (U.K.), etc. He has over 70 works on peer-reviewed international journals and conference proceedings, one international patent, and two books. Research interests are image processing, computer vision, machine learning, data mining, and knowledge discovery. E-mail: *thanhdnh@ueh.edu.vn* .

Nguyen Hoang Hai (b. 1986) graduated Bachelor and Master graduation certificate specialized in Computer Science at Belarus State University–Belarus, and he received PhD graduation certificate with the same specialization at the Autonomous University of Barcelona, Spain. Currently, he works as an assistant professor at Computer Science faculty, Vietnam-Korea University of Information and Communication Technology –The University of Danang, Vietnam. His

research area focuses on Internet of Things (IoT), cloud computing and optimization methods for energy saving in interconnection networks, image processing, and machine learning. E-mail: *nhhai@vku.edu.vn* .

Le Minh Hieu (b. 1986) graduated from Belarusian State University in 2009 majoring in Applied Mathematics and graduated MSc (2010), Ph.D. (2018) majoring in Computational Mathematics from Belarusian State University. Currently, he works as a lecturer at the University of Economics, the University of Danang, Vietnam. Research interests are finite difference schemes, nonlinear PDEs, machine learning. E-mail: *hieulm@due.edu.vn* .

Prayag Tiwari (b. 1991) received his Master's Degree majoring in Computer Science from the National University of Science and Technology (MISiS), Russia in 2017, and graduated Ph.D. (2020) majoring in Information Engineering from University of Padua, Italy. Research interests include Machine Learning/Deep Learning, Quantum Theory, Information Retrieval, and IoT. E-mail: *prayag.tiwari@dei.unipd.it* .

Prasath V. B. Surya (b. 1982) graduated Ph.D. from the Indian Institute of Technology Madras in 2010, majoring in Mathematics. Currently, he works as an assistant professor at the Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, and also affiliated with the Biomedical Informatics, Electrical Engineering, and Computer Science departments, University of Cincinnati, USA. He has been a postdoctoral fellow at Mathematics department, University of Coimbra, Portugal (2010-2012). From 2012 to 2017 he was an assistant professor at the Computational Imaging and VisAnalysis (CIVA) Lab of the University of Missouri, USA. He had summer fellowships / visits at Kitware Inc. NY, USA, The Fields Institute, Canada, and Institute for Pure and Applied Mathematics (IPAM), University of California Los Angeles, USA. Research interests are nonlinear PDEs, regularization methods, inverse & ill-posed problems, variational and PDE-based image processing, computer vision with applications in remote sensing, biometrics, and biomedical imaging domains. E-mail: *prasatsa@uc.edu* .