# Semi-Automatic Skin Lesion Segmentation via Fully Convolutional Networks

*Lei Bi*[1], *Jinman Kim*[1], *Euijoon Ahn*[1], *Dagan Feng*[1,4] and *Michael Fulham*[1,2,3]

[1]School of Information Technologies, University of Sydney, Australia
[2]Department of PET and Nuclear Medicine, Royal Prince Alfred Hospital, Australia
[3]Sydney Medical School, University of Sydney, Australia
[4]Med-X Research Institute, Shanghai Jiao Tong University, China

## ABSTRACT

Segmentation of skin lesions is considered as an important step in computer aided diagnosis (CAD) for melanoma diagnosis. There have many attempts to segment skin lesions in a semi- or fully-automated manner. Existing methods, however, have problems with over- or under-segmentation and do not perform well with challenging skin lesions such as when a lesion is partially connected to the background or when image contrast is low. To overcome these limitations, we propose a new semi-automated skin lesion segmentation method that incorporates fully convolutional networks (FCNs) with multi-scale integration. We leverage the use of FCNs to derive high-level semantic information with simple user interaction e.g., a single click to accurately segment skin lesions of various complexity. Our experiments with 379 skin lesion images show that our proposed method achieves better segmentation results when compared to the state-of-the-art skin lesion segmentation methods for challenging skin lesions.

***Index Terms—*** Segmentation, Fully Convolutional Networks (FCN), Skin lesion, Melanoma

## 1. INTRODUCTION

Malignant melanoma has one of the most rapidly increasing incidences in the world and causes considerable morbidity and mortality [1]. Early diagnosis is particularly important since melanoma can be cured with excision if detected early [2]. Even for experienced dermatologists, however, diagnosis by human vision can be subjective, inaccurate and non-reproducible [1]. This is attributed to the complexity of lesion segmentation due to variations in size and shape, fuzzy lesion boundaries and different skin colors. Motivated by these difficulties, there has been great interest in developing computer-aided diagnosis (CAD) systems that can assist dermatologists' clinical evaluation [1, 3-5].

Lesion segmentation is a fundamental requirement for a melanoma CAD. A number of segmentation methods have been recently proposed to segment skin lesions. These methods include multi-scale superpixel based cellular automata (MSCA) [6], saliency based skin lesion segmentation method (SSLS) [7], deformable model (DM) [1], texture distinctive lesion segmentation (TDLS) [3], adaptive thresholding (AT) [4], level set proposed by Chan et al (C-LS) [8]. However, all these previous semi- and fully-automatic methods rely on low-level image features such as color and texture features. Low-level features do not understand image-wide variations and therefore cannot provide accurate segmentation when lesions have complex shapes and are low-contrast to the background.

Recently, fully convolutional networks (FCNs) have achieved great success in object detection and segmentation related problems [9]. Such success is primarily attributed to the ability of FCNs to improve feature representation with a high-level of semantics [10]. In addition, different to the existing methods that usually depend on many hand-engineered and pre-processing steps, FCN can be trained with efficient inference of learning in an end-to-end way, which takes an image as input and outputs the object detection or segmentation results directly. However, there is a scarcity of medical image training data together with annotations due to the large cost and complicated acquisition procedures [11]. Consequently, without sufficient training data to cover all the skin lesion variations, it cannot provide accurate segmentation for challenging and rare skin lesions (not covered in the training data).

Motivated by the work proposed by Xu et al [12], we propose a new skin lesion segmentation method using semi-automated FCNs and multi-scale integration. Our semi-automated method takes advantage of using a priori knowledge to solve the segmentation problems for challenging and rare skin lesions. In addition, the proposed multi-scale integration is capable of providing precise boundaries of skin lesions of various sizes and contrasts. The novelty of our algorithm when compared to previous studies is as follows: (1) we leverage FCNs to extract high-level semantic features to accurately segment the skin lesions with high efficiency; and (2) the proposed method is able to segment challenging and rare skin lesions with minimum amount of user interactions (single click).

## 2. METHODS

### 2.1. Semi-Automated Fully Convolutional Networks (FCN)

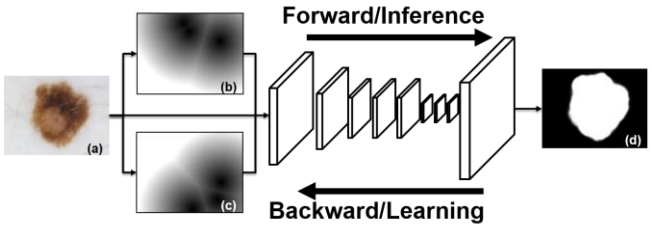


Figure 1. Semi-automated FCN at different stages. (a) input skin lesion image; (b) and (c) fore- and background channels derived from user clicks; (d) FCN produced probabilistic map (segmentation result) of (a).

The FCN architecture has downsampling and upsampling parts [9]. The downsampling part has convolutional and max-pooling layers, which were widely used in convolutional neural networks (CNN) for image classification task [13]. The upsampling part has convolutional and deconvolutional layers, which upsample the feature maps and output the score masks. By combining the downsampling and upsampling parts, the FCN architecture was able to extract the high-level semantic information while predicting the pixel-wise score mask.

Different to the existing FCN based methods, we create two extra channels that encode the probability of the foreground (skin lesion) and the background calculated from the user interactions. Basically, the two channels indicate the probability of a pixel belonging to the fore- and background. The calculation of these two channels was achieved by applying Euclidean distance transform to measure the shortest distance from a pixel to the user clicked points.

For each training image, we randomly sampled a couple of foreground and background clicks from the ground truth. There were 9 combinations of fore- and background clicks, which consist of {1, 2, 3} foreground and {0, 1, 2, 3} background clicks. We also determined that the number of background clicks should be always less/equal than the number of foreground clicks because foreground clicks are more useful for detecting challenging and rare skin lesions.

*2.1.1. Training*

There is a scarcity of medical image training data together with annotations. Compared with the limited data in the medical domain, there are much more data that are available in the field of general images. Existing literature show that fine-tuning deep convolutional networks (DCNs) can alleviate the problem of insufficient training data, where the lower layers of the fine-tuned network are more general filters (trained on general images) while those in the higher layers are more specific to the target problem [11, 14]. Therefore, we used the off-the-shelf MatConvNet [15] version of a FCN trained on the PASCAL VOC 2011 dataset. Data augmentation techniques including random crops and flips were used to provide extra material for training. To achieve more precise details of pixel predication, we fine-tuned a stride-8 FCN architecture (FCN-8s) on the skin images. It took about 3 days to fine-tune 200 epochs with a batch size of 20 on a Titan X GPU. The FCN converges in approximately the 150$^{th}$ epoch and longer training time gives finer delineation results.

*2.1.2. Testing*

For testing images, we used the same technique to generate the random clicks of the fore- and the background channels. The trained model will also work if no click is provided, where in this case, both the fore- and the background channel will have the largest distance. Furthermore, the trained model can also work in an iterative semi-automated manner. The user can iteratively select a new pixel (seed) and then this new seed will be transformed into a new fore- or background channel and enter into the trained model for a new segmentation. Leveraging the high-end GPU, our semi-automated method is highly efficient, where it takes around 200ms to segment one image for a single iteration.

### 2.2. Multi-scale Integration and Post-refinement

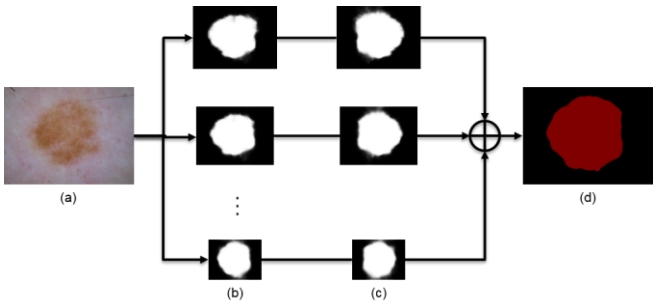


Figure 2. Multi-scale integration, where (a) is the input image, (b-c) are the semi-automatic FCN generated probabilistic maps on different scales (b) and flips (c), and (d) final integrated segmentation result.

In general, the semi-automated FCN outputs are feasible. However, FCN was designed for object detection on general images that cannot provide accurate boundary definitions of skin lesions of various sizes and contrast. Therefore, we used a multi-scale strategy with flip on each scale to segment the skin lesions (Figure 2). For each input skin lesion image, we downsized for 7 different scales and we averaged the segmentation results to produce the final probability map. Downsizing instead of upsizing was used mainly due to the fact that FCN usually produce over-segmented results with many isolated regions for upsized skin lesion images.

The final probabilistic map was converted into a binary segmentation result via thresholding (50% of the maximum). For segmentation refinement, we followed previous work of Glaister et al [3] and used a morphological dilation process

to smooth the boundary, fill small holes and use connected thresholding to remove small isolated single pixels.

## 3. RESULTS AND DISCUSSION

### 3.1. Materials and Experimental Setup

We evaluated our proposed method on the ISBI 2016 skin lesion challenge dataset [16]. The challenge dataset consists of 900 training and 379 testing dermoscopy images. Manually annotated lesions from dermatologists were used as the ground truth.

We compared our proposed method to the related and the state-of-the-art methods on skin lesion segmentation. For our method, we only used the segmentation results produced via using a single foreground click (random selected, see Section 2.1 for more details), where segmentation results can be further improved with extra clicks. The comparison methods include: (i) MSCA [6] – multi-scale superpixel based cellular automata; (ii) SSLS [7] – saliency based skin lesion segmentation; (iii) FCN [9] – fully convolutional networks (FCN-8s); (iv) RW [17] – seeded random walker; and (v) GC [18, 19] – grow cut. The attributes of each method are provided in Table 1. For both (iv) and (v), we have used extra seeds to make the segmentation results comparable, where the extra seeds were randomly sampled from the ground truth.

**Table 1:** Attributes of our method and the comparison methods.

| | MSCA | SSLS | FCN | GC | RW | Our |
|---|---|---|---|---|---|---|
| **Training Data Required** | × | × | √ | × | × | √ |
| **Semi-automatic** | × | × | × | √ | √ | √ |
| **Deep learning** | × | × | √ | × | × | √ |
| **Year** | 2016 | 2015 | 2015 | 2005 | 2006 | 2016 |

The most common skin lesion segmentation evaluation metrics were used for comparison including: (i) dice similarity coefficient (Dic.) – measures the overlap between the ground truth (denoted as $g$) and the algorithm produced results (denoted as $a$); (ii) Jaccard index (Jac.) – measures the overlap over the union of $g$ and $a$; and (iii-v) pixel-level sensitivity (Sen.), specificity (Spe.) and accuracy (Acc.).

In addition, we calculated the pixel-level receiver operating characteristic (ROC) curve and the precision-recall (PR) curve for additional comparisons. Both ROC and PR curve have been widely used for object detection related problems on general images [20].

Furthermore, we compared our method with the top 5 performing results (out of 28 teams) on the ISBI 2016 skin lesion segmentation challenge.

### 3.2. Results

Table 2, Figure 3 and Figure 4 show that our method performed the best across different measurements. Table 3 shows that our method also achieved competitive results when compared with the top 5 out of 28 teams on the ISBI 2016 skin lesion segmentation challenge.

**Table 2:** Segmentation results of our method compared to the other comparison methods, where **red** and **blue** represent the best and the second best results.

| | Dic. | Jac. | Sen. | Spe. | Acc. |
|---|---|---|---|---|---|
| **MSCA** | 75.88 | 66.19 | 78.30 | 91.31 | 85.68 |
| **SSLS** | 69.97 | 57.20 | 70.04 | **97.31** | 84.67 |
| **GC** | 50.65 | 35.71 | 70.50 | 75.72 | 71.61 |
| **RW** | 57.52 | 42.81 | 72.31 | 77.07 | 74.28 |
| **FCN** | **88.64** | **81.37** | **91.70** | 94.90 | **94.13** |
| **Our** | **90.24** | **83.36** | **91.82** | **95.23** | **95.10** |

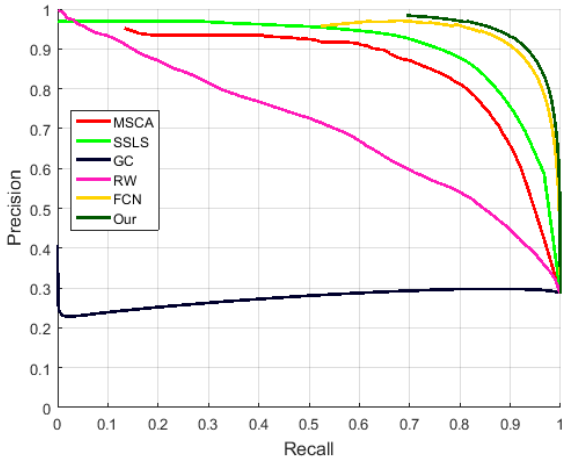

Figure 3. PR cuve of our method and other comparison methods.

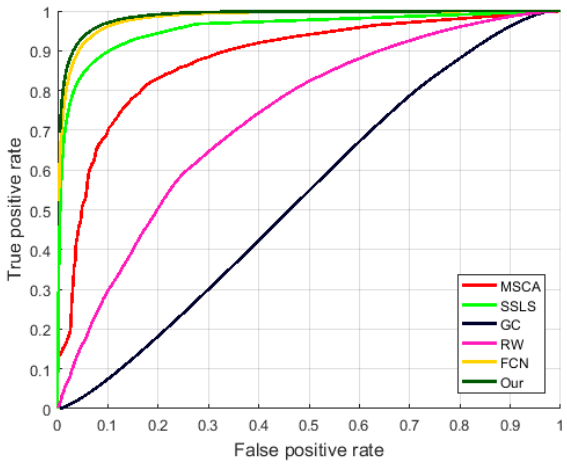

Figure 4. ROC cuve of our method and other comparison methods.

**Table 3:** Segmentation results of our method compared to the top 5 out of 28 teams on the ISBI 2016 skin lesion segmentation challenge.

| | Dic. | Jac. | Sen. | Spe. | Acc. | Rank |
|---|---|---|---|---|---|---|
| **Our** | **90.24** | **83.36** | **91.82** | 95.23 | 95.10 | - |
| **ExB** | **91.00** | **84.30** | 91.00 | 96.50 | **95.30** | 1 |
| **CUMED** | 89.70 | 82.90 | 91.10 | 95.70 | 94.90 | 2 |
| **Rahman** | 89.50 | 82.20 | 88.00 | **96.90** | **95.20** | 3 |
| **SFU** | 88.50 | 81.10 | **91.50** | 95.50 | 94.40 | 4 |
| **TUM** | 88.80 | 81.10 | 83.20 | **98.70** | 94.60 | 5 |

### 3.3. Discussion

In Table 2 and Figures 3 and 4 we show that our method was the best performed for skin lesion segmentation when compared with the existing methods. When compared with RW and GC, both MSCA and SLSS produce better segmentation results, and we suggest that this relates to their more advanced feature representations. These unsupervised methods, however, were not able to separate the skin lesions from artifacts e.g., color band (Figure 5, first two columns). Although FCN achieved competitive results (~2% lower to our method in Jaccard measure), it failed to segment the rare lesions, where similar lesions were not found in the training dataset (Figure 5, last column).

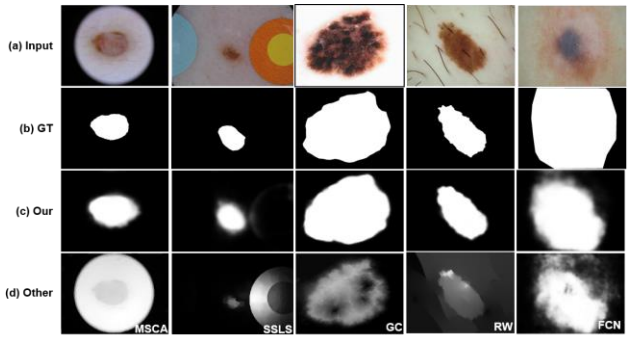


Figure 5. Segmentation results from 5 example studies. (a) input images, (b) ground truth (GT), (c) segmentation results of our method and (d) segmentation results of other methods including MSCA, SSLS, GC, RW and FCN from left to right.

Table 3 shows that our method can achieve similar results when compared with the best performing teams in the ISBI 2016 skin lesion segmentation challenge (0.96% lower to the best). In addition, our method further improves segmentation of the challenging skin lesions, as shown in Figure 6. In this example, the top two teams had a DSC value of 60.50% (ExB) and 71.10% (CUMED). In contrast, our proposed method had a DSC of 80.21% with a single foreground click and a DSC of 92.19% with an additional background click. These results reflect the advantages of our method in segmenting rare and challenging skin lesions.

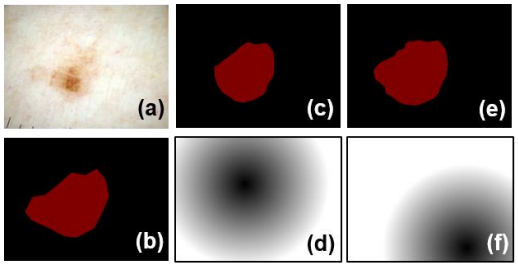


Figure 6. Segmentation results from one example study. (a) input image; (b) ground truth; (c) segmentation results (DSC=80.21%) of our method with a single foreground click (d); and (e) segmentation results (DSC=92.19%) of our method with foreground click (d) and an additional background click (f).

## 4. CONCLUSION AND FUTURE WORK

We propose a new semi-automated skin lesion segmentation method that employs semi-automated FCNs and multi-scale integration. Our experiments show the advantages of using our method for segmenting challenging skin lesions of various complexity. In future work, we will use our method in a larger sample of complex skin lesion images.

## REFERENCES

[1] Z. Ma, et al., "A Novel Approach to Segment Skin Lesions in Dermoscopic Images Based on a Deformable Model," IEEE Journal of Biomedical and Health Informatics, 2015.
[2] M. E. Celebi, et al., "A methodological approach to the classification of dermoscopy images," Computerized Medical Imaging and Graphics, 2007.
[3] J. Glaister, et al., "Segmentation of skin lesions from digital images using joint statistical texture distinctiveness," IEEE Transactions on Biomedical Engineering, 2014.
[4] M. Silveira, et al., "Comparison of segmentation methods for melanoma diagnosis in dermoscopy images," IEEE Journal of Selected Topics in Signal Processing, 2009.
[5] E. Ahn, et al., "Saliency -based Lesion Segmentation via Background Detection in Dermoscopic Images," IEEE Journal of Biomedical and Health Informatics, 2016.
[6] L. Bi, et al., "Automated skin lesion segmentation via image-wise supervised learning and multi-scale superpixel based cellular automata," in IEEE ISBI, 2016.
[7] E. Ahn, et al., "Automated Saliency-based Lesion Segmentation in Dermoscopic Images," in IEEE EMBC, 2015.
[8] T. F. Chan, et al., "Active contours without edges for vector-valued images," Journal of Visual Communication and Image Representation, 2000.
[9] J. Long, et al., "Fully convolutional networks for semantic segmentation," in IEEE CVPR, 2015.
[10] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv:1409.1556, 2014.
[11] H.-C. Shin, et al., "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning," IEEE Transactions on Medical Imaging, 2016.
[12] N. Xu, et al., "Deep Interactive Object Selection," in CVPR, 2016.
[13] K. Chatfield, et al., "Return of the devil in the details: Delving deep into convolutional nets," arXiv preprint arXiv:1405.3531, 2014.
[14] L. Bi, et al., "Automatic detection and classification of regions of FDG uptake in whole-body PET-CT lymphoma studies," Computerized Medical Imaging and Graphics, 2016.
[15] Vedaldi and K. Lenc, "Matconvnet: Convolutional neural networks for matlab," in ACM international conference on Multimedia, 2015.
[16] D. Gutman, et al., "Skin Lesion Analysis toward Melanoma Detection: A Challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC)," arXiv preprint arXiv:1605.01397, 2016.
[17] L. Grady, et al., "Random walks for image segmentation," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006.
[18] V. Vezhnevets and V. Konouchine, "GrowCut: Interactive multi-label ND image segmentation by cellular automata," Proc. of Graphicon, 2005.
[19] L. Bi, et al., "Cellular automata and anisotropic diffusion filter based interactive tumor segmentation for positron emission tomography," in IEEE EMBC, 2013.
[20] X. Li, et al., "Contextual hypergraph modeling for salient object detection," in IEEE ICCV, 2013.