



Skin lesion segmentation by using object detection networks, DeepLab3+, and active contours

Fatemeh BAGHERI¹, Mohammad Jafar TAROKH^{2,*} , Majid ZIARATBAN³ 

¹Department of Computer Engineering, Faculty of Engineering, Golestan University, Gorgan, Iran

²Department of Industrial Engineering, K. N. Toosi University of Technology, Tehran, Iran

³Department of Electrical Engineering, Faculty of Engineering, Golestan University, Gorgan, Iran

Received: 30.07.2020

Accepted/Published Online: 27.09.2022

Final Version: 28.11.2022

Abstract: Developing an automatic system for detection, segmentation, and classification of skin lesions is very useful to aid well-timed diagnosis of skin diseases. Lesion segmentation is a crucial task for automated diagnosis of skin cancers, as it affects significantly the accuracy of the subsequent steps. Varieties in sizes and locations of lesions, and the lesions with low-contrast boundaries make this task very challenging. In this paper, a three-stage CNN-based method is presented for accurate segmentation of lesions from dermoscopic images. At the first step, normalization, approximate locations and sizes of lesions are estimated. Due to the importance of the normalization stage, three CNN-based networks (Mask R-CNN, RetinaNet, and YOLOv3) are used for the lesion detection. A convolutional network is presented and used to combine the results of the object detection networks with a novel approach. The output of the first stage is a normalized cropped image containing the detected lesion in the center. At the second stage, segmentation, a CNN in a DeepLab3+ structure, is used to extract the lesion from the normalized image. Finally, an active contour method is used as the postprocessing to enhance the boundary of the segmented lesion. The proposed method is evaluated on well-known datasets. Experiments show that the proposed method outperforms all the previous state-of-the-art methods.

Key words: Skin lesion segmentation, Mask R-CNN, RetinaNet, Yolo, DeepLab, Active contour

1. Introduction

Skin cancer is the most common cancer worldwide. One out of every three people with cancer has skin cancer. In 2020 according to the estimations by the American Cancer Society, in the United State, About 100,350 new melanomas will be diagnosed. For this reason, diagnosing and treatment of it as early as possible is crucial to prevent the death of patients. Common and routine diagnostic methods are expensive because of the need for qualified specialists as well as advanced equipment. Recent advances in computer solutions, which have improved accuracy and efficiency, have made the fast and accurate diagnosis of this disease promising. In general, automatic analysis of skin images involves three stages of lesion segmentation, feature extraction, and lesion classification. Lesion segmentation is a very important step, as it significantly affects the accuracy of the lesion classification. Due to the large variety of lesions in sizes, locations, tissues, and colors, segmentation of lesions from skin images is difficult. Other problems include presence of dark hair and blood vessels as well as the light reflections, which is also found in many images on the lesion [1].

In recent decades, several methods have been proposed for automatic skin lesion segmentation. These approaches can be generally divided into two categories of unsupervised and supervised methods. Unsupervised

*Correspondence: mjtarokh@kntu.ac.ir

skin lesion segmentation methods do not require training data. However, unsupervised approaches are not necessarily the appropriate and powerful solutions for all problems. These methods do not present reasonable accuracy in challenging cases [2].

Recent studies employing convolutional neural networks (CNNs) in image segmentation have outperformed other classical methods. This technique is used in various medical applications, such as the detection of organs from CT images, classification of skin cancers, segmentation of medical images to diagnose brain tumors on MRI imaging, segmentation of MRI images of the heart, and segmentation of skin lesions.

1.1. Related works

Yuan et al. presented a full automatic skin lesion segmentation method based on a 19-layer CNN. They used the Jaccard index as a cost function in their method, and adjusted the imbalance between the pixel tissue lesions. They tested their method on datasets of the IEEE International Symposium on Biomedical Imaging (ISBI) 2016 and PH2. However, they had problems with some challenging cases such as the images with low-contrast lesions [4]. Yu et al. proposed a two-step method using deep residual networks (FCRN) for segmentation and classification of skin lesions. They tested their method on the ISBI 2016 dataset and achieved an accuracy of 94.9% [2].

Yuan et al., in 2017, developed a method using deep convolutional-deconvolutional neural networks (CDNN) for skin lesion segmentation. Their approach was ranked first in the ISBI 2017 Challenge with 76.5% of the Jaccard index [5].

In 2018, Li and Shen proposed a deep learning approach to distinguish melanoma from other skin diseases. They first segmented the skin lesions. At the segmentation stage, they presented a framework based on multiscale fully convolutional residual networks and then evaluated their method on the ISBI 2017 dataset. The Jaccard index value for their method on this dataset was 75.3% [6].

In 2018, Al-Masni et al. introduced a full-resolution convolutional network (FrCN) for lesion image segmentation. In their approach, to have better pixel-level lesions segmentation, the network learned from the full features of each pixel of the input images. They did not use any preprocessing or postprocessing operations. To evaluate their method, they performed experiments on the ISBI 2017 and PH2 datasets and compared the results with other similar methods. Their method achieved 77.11% and 84.79% Jaccard index for the mentioned datasets, respectively [7].

Baghersalimi et al. developed a full convolutional neural network called DermoNet to segment skin lesions. At DermoNet, because of the existence of blocks with dense connections and also skip connections between the different layers, the network layers can reuse the previous layers' information and ensure high accuracy in the subsequent layers of the network. In doing so, high-level feature representations learned in the middle layers of varying scales and resolutions were used to segment the lesions. Quantitative evaluation of this method was performed on the three ISBI 2016, ISBI 2017, and PH2 datasets. The Jaccard index obtained on these datasets were 82.5%, 78.3%, and 85.3%, respectively [8].

In 2018, Qian et al. proposed a two-stage method for lesion segmentation with optimized training method and ensemble postprocess. Their method is a two-stage process including detection and segmentation. They used Mask R-CNN to detect the location of the lesion and cropped the lesion from images. Following the detection, segmentation part segments the cropped image and predicts the region of lesion. In the segmentation part, they designed an encode-decode architecture of network. They evaluated their method on a part of the training set of ISBI 2018 dataset and achieved 81.6% in Jaccard index [9].

A method based on attention-guided networks was proposed in [10] in 2021. In the last step of the encoding part, densely connected convolutions were used. They also proposed an attention-guided filter module, channel spatial fast attention-guided filter. Another attention-based network was introduced in [11] called FAC-Net. This network was based on combination of the feedback fusion block (FFB) and the attention mechanism block (AMB) to obtain better feature mapping.

A network with feature adaptation transformer was presented in [12] for skin lesion segmentation. The encoder of the transformer utilized a sequence-to-sequence prediction method and increased the receptive field based on the self-attention mechanism.

In [13], an encoder-decoder based network named Ms Red was proposed in which a multiscale residual encoding module (MsR-EFM) and a multiscale residual decoding module (MsR-DFM) were used as the encoder and decoder, respectively. In addition, a novel multiresolution, multichannel feature fusion module (M^2F^2) was proposed to improve the segmentation performance.

For many applications, both local information on the lesion and general conceptual information on it are required to enhance the classification accuracy. It is not possible to combine these two types of information in public architectures of deep learning. Numerous researchers have used multistream architectures to solve it [14]. Shen et al. used three convolutional networks where each of them receives lesion information from different sights as input. They constructed the final feature vector by bringing together the features obtained as the output of these three networks [15]. A similar approach was developed in 2016 by Kawahara and Hamarneh. They used a multistream CNN to classify skin lesions. The streams worked on different versions of the image resolution [16]. These studies suggest that combining multiple CNNs with different details can improve the final performance. In this paper, a novel combination of object detection networks (ODNs) is proposed to improve the performance of the proposed method.

1.2. Contributions

- Adding a normalization stage before segmentation to reduce the complexity of the segmentation stage.
- Adding a postprocessing stage after segmentation for fine-tuning the result of the segmentation stage.
- Use of three CNN-based ODNs to increase the accuracy of the normalization stage.
- A novel combination approach to combine the results of object detectors.
- Converting the results of object detectors to image matrices and concatenating these results with the input dermoscopic image.
- Introducing a deep convolutional network to combine the results of object detectors with considering the input dermoscopic image and the ground-truth mask, simultaneously.
- Data augmentation of training images as well as using different modes of test images to cover varieties of lesions.
- Employing CNN-based segmentation stage followed by the active contour method to consider both overall characteristics and image-specific features of lesions in skin images.

2. Proposed method

In recent years, most semantic segmentation methods have been proposed based on convolutional neural networks and in a single stage. One of the main problems of these methods is that their performances are sensitive to the varieties in size and location of objects in images. In lesion segmentation applications, existence of very large and very small lesions reduces the accuracy of single-stage segmentation methods. Moreover,

different lesion locations increase the network complexity and reduce the segmentation accuracy. Thus, it is better to perform a presegmentation step in order to normalize the size and location of lesions in images. Doing that, at the segmentation stage, the network faces lower complexity and functions better. In the proposed method, a normalization stage is considered before the segmentation stage to estimate the size and location of a lesion in an image.

2.1. Image normalization stage

The most important part of the proposed method is the lesion bounding box detection in the normalization stage. It is because the results of this stage substantially affect the result of the segmentation stage. We use object detection networks (ODNs) to estimate the bounding box of lesions.

Numerous CNN-based methods have been developed for object detection applications, such as R-CNN [17], fast R-CNN [18], faster R-CNN [19], mask R-CNN [20], single shot multibox detector (SSD) [21], you only look once (Yolo) [22], and Retinanet [23].

Due to the importance of the normalization stage, a method is proposed to combine the results of ODNs. Several ODNs have been investigated in our experiments and finally, three networks producing the highest lesion detection accuracies on the augmented validation set of the ISBI 2017 dataset are selected as follows:

- Mask R-CNN with ResneXt 101 as a backbone network,
- RetinaNet with Resnet 101 as a backbone network,
- Yolov3 with Darknet as a backbone network.

2.2. Mask R-CNN

In 2018, He et al. presented an object segmentation framework called mask R-CNN. They added a branch to the faster R-CNN structure to predict the object mask, so that, in addition to the classification and bounding box regression branches presented in the faster R-CNN, an object mask is also predicted for each region of interest (ROI) [20]. The mask branch is a full convolutional network which applies to any ROI and predicts the segmentation mask.

Incorporation of this new branch added a small computational overhead, but it can produce pixel-to-pixel alignment between input and output of network and ultimately produce good results [20]. As the term of L_{mask} has been added to the loss function, the object detection accuracy of the mask R-CNN is expected to be higher than that of the faster R-CNN.

2.3. RetinaNet

In 2018, Lin et al. presented a single-step method to detect objects in images. They tried to solve the extreme foreground-background class imbalance in the training phase. Thus, they changed the standard cross entropy loss and assigned more suitable weights to make it more effective. They added the factor $(1 - p_t)^\gamma$ to the standard cross entropy loss function and named it focal loss. They defined focal loss (FL) function as follows:

$$FL(p_t) = -(1 - p_t)^\gamma \log(p_t) \quad (1)$$

$$p_t = \begin{cases} p & \text{if } y = 1 \\ 1 - p & \text{otherwise} \end{cases}, \quad (2)$$

where $\gamma \geq 0$ is the focusing parameter. This parameter adjusted the rate at which easy samples were down-weighted. When focusing parameter was set to zero, the focal loss was equivalent to the cross entropy. Moreover, as it was increased, the effect of the modulating factor was increased. In their experiments, the value of γ was set to 2 to have the best performance [23]. We also set γ to 2 due to reasonable results in our experiments.

2.4. Yolo

Yolo is an object detection algorithm based on a convolutional neural network which consists of only one network. Unlike R-CNN methods, it does not use the region proposal. This algorithm divides the image into several subregions. It then, for each region, predicts the boxes and their probabilities of belonging to the classes. One weakness of the initial version of Yolo was that it did not detect very small objects [22]. In order to improve the initial version, a second version, Yolo v2, was released [24]. To capture low-level features, Yolo v2 concatenates feature maps from previous layers. It is still used in most state-of-the-art methods. Yolo v3 uses Darknet as its backbone, which originally has 53 layers trained on Imagenet dataset. Specifically, 53 more layers have been added to its structure for the task of detection [25].

2.5. Combination of the results of ODNs

Three ODNs (mask R-CNN, Retinanet, and Yolov3) are trained with an augmented training set. In addition to estimation of the bounding box, the ODNs also provide the score of detection at their output. The detection score shows the degree of confidence of the ODN in its detection. In this paper, a novel approach is proposed to combine the results of ODNs. The result of each ODN is converted to an image, in which an elliptic-shaped foreground bounded with the estimated bounding box. The values of the foreground pixels are set to the detection score. The images created from the results of the three ODNs for a sample image in Figure 1a are illustrated in Figure 1b. The images constructed from the result of mask R-CNN, RetinaNet, and Yolov3 are denoted by I_{Mask} , I_{Retina} , and I_{Yolo} , respectively. These images are concatenated with the input RGB image and a 6-channel image is made as follows:

$$I_{Comb} = Concat(I_{Red}, I_{Green}, I_{Blue}, I_{Mask}, I_{Retina}, I_{Yolo}). \quad (3)$$

A convolutional network (we call it *CombNet*) that receives I_{Comb} as an input is implemented. The details of the CombNet are given in Figure 2. The goal of using this network is the estimation of the bounding box of the lesion based on the results of ODNs as well as the input image. To implement this idea, the CombNet is trained to create a binary image similar to the ground truth image at its output. The bounding box of the largest connected component of the binary image at the output of the CombNet is considered the final bounding box. To create training data for the CombNet, the images of the training set of the ISIC2017 dataset are augmented and 5000 new images are created and applied to three ODNs (mask R-CNN, RetinaNet, and Yolov3). A total of 5000 6-channel images, I_{Comb} , are made by concatenating the RGB channels and the images constructed from the output of the ODNs according to Eq. (3). The procedures of constructing the required train sets for training the networks in each stage are illustrated in Figure 3.

Figure 4 displays the procedures of the proposed method. As shown in the figure, the input RGB image is normalized using the bounding box obtained based on the output of the CombNet. To normalize a test image, a subimage limited to the obtained bounding box should be cropped from the test image and resized to a predefined image size.

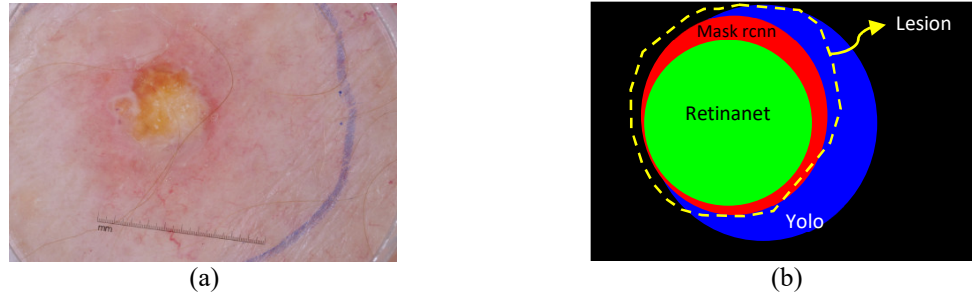


Figure 1. (a) A sample image. (b) The ovals created from the results of the ODNs. Dashed line marks the boundary of the lesion in the ground-truth.

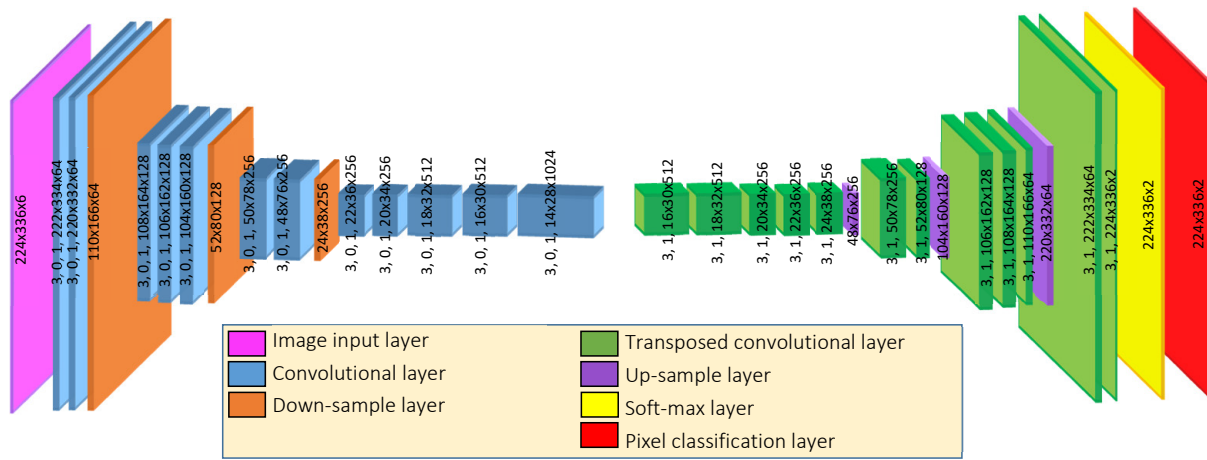


Figure 2. The proposed convolutional network, CombNet, to combine the results of the ODNs.

In some cases, in the lesion detection stage, the bounding box may be estimated smaller than the correct one. In these cases, by cropping the bounding box from the image, some parts of the lesion will be missed. Let H_{BB} and W_{BB} be the height and width of the estimated bounding box. To reduce this type of errors, a margin of size $0.3 W_{BB}$ pixels from left and right sides, and a margin of size $0.3 H_{BB}$ pixels from top and bottom sides of the detected lesion are considered. Hence, a box of size $(1.6H_{BB}) \times (1.6W_{BB})$ is cropped from the input image so that the center of the detected lesion is approximately on the center of the cropped box. The cropped box is resized to a normal image of size 224×336 pixels. The final bounding box estimated by the CombNet and the normalized cropped image for the instance in Figure 1 are depicted in Figure 5. In the segmentation stage, the lesion is segmented from the normalized image. Then the normalized segmented image (which is a 224×336 pixels binary image) is first resized back to the size of the cropped box and placed in the same position as the cropped box had been extracted from the original image. Therefore, as shown in Figure 4, at the output of the "Inverse of normalization" block, we have a segmented binary image with the size of the original image. To fine-tune the boundary of the segmented lesion, the binary image is applied to the postprocessing stage.

2.6. Segmentation stage

The network in the segmentation stage receives the normalized image at its input and produces a binary image containing the segmented lesion. Several methods and networks have been used to image semantic segmentation

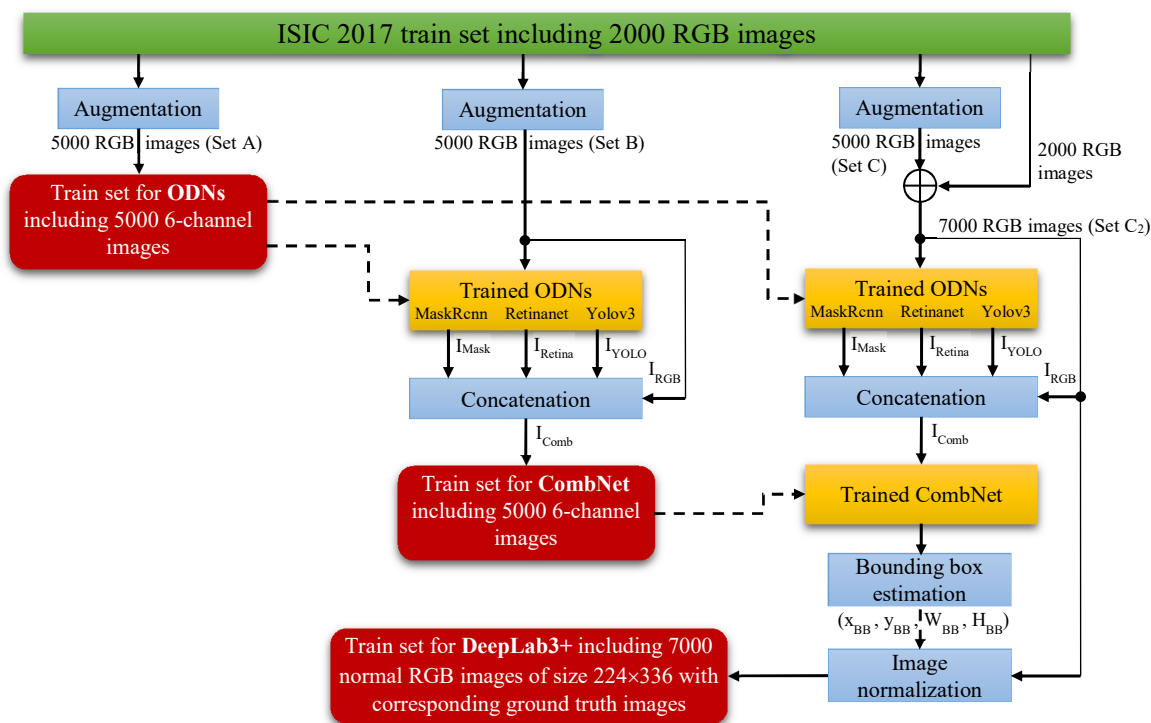


Figure 3. The procedures of constructing the required train sets for training the networks in each stage.

in various applications. DeepLab architecture is one of the most recent structures and has performed well in many applications [26]. In general, the Deeplab architecture is based on a combination of two most common architectures of spatial pyramid pooling and encoder-decoder network [27]. Different DeepLab structures have been improved over time. DeepLab v1 [26], DeepLab v2 [28], DeepLab v3 [29], and DeepLab v3+ [27] are the various structures of DeepLab. DeepLab v1 uses atrous convolution to control the resolution at which feature responses are computed [26]. DeepLab v2 applies atrous spatial pyramid pooling (ASPP) to segment objects on multiple scales [28]. To capture more information, DeepLab v3 augments the ASPP module via image-level feature. To facilitate the training, it also includes batch normalization parameters. DeepLab v3+ extends DeepLab v3 to include an effective decoder module to refine the segmentation results [27].

In the segmentation stage of our proposed method, DeepLab3+ structure is used with VGG19 pretrained network as a backbone network.

2.7. Postprocessing

To modify the boundary of the lesion segmented by the DeepLab, we use active contour method as post-processing. Active contour-based methods have been widely used in image segmentation as well as in lesion segmentation [30–33]. These methods can be categorized into two main groups: edge-based methods [34] and region-based methods [35]. Since many images of the ISBI dataset do not include certain edges, a region-based active contour method is used as developed by Chan and Vese [36]. They proposed their method to improve the segmentation performance in some cases such as images containing objects with low-contrast boundaries, high noise, smooth contours, or multiobjects in different shapes. The main concept of the region-based active

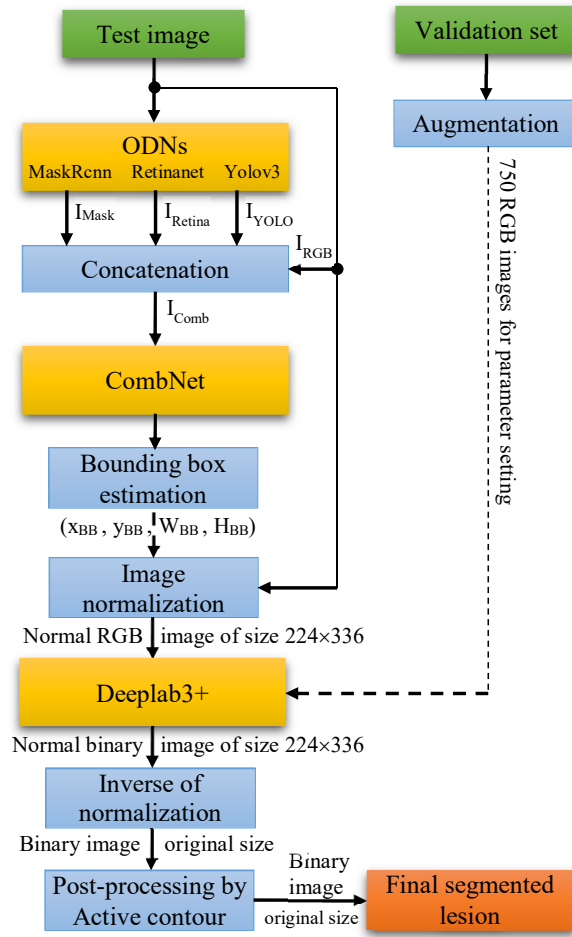


Figure 4. Procedures of the proposed method.

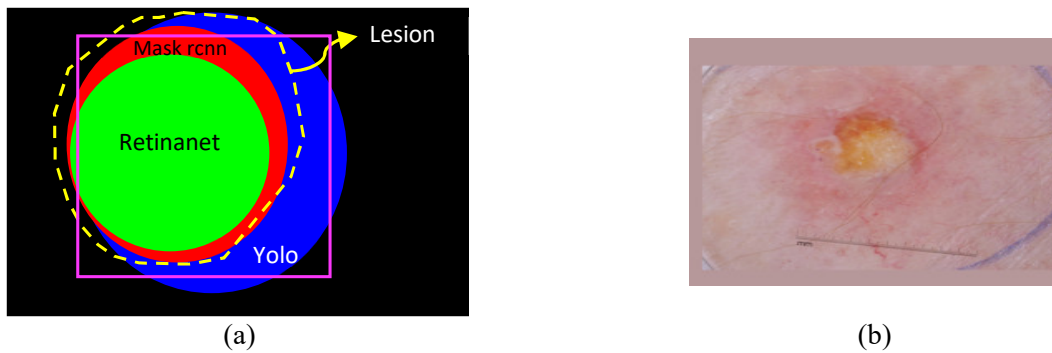


Figure 5. The pink rectangle in (a) is the final bounding box estimated by the CombNet, and the normalized cropped image (b) for the image in Figure 1.

contour method is the concept of energy, which is the sum of inside and outside energies [36]:

$$F_1(C) + F_2(C) = \int_{inside(C)} |v_0(x, y) - C_1|^2 dx dy + \int_{outside(C)} |v_0(x, y) - C_2|^2 dx dy. \quad (4)$$

The first step of the region-based active contour method is to create a mask or contour and consider it an initial object (*inside* (C)). Outside of this mask is considered the background (*outside* (C)). The curve C is represented via a Lipschitz function by:

$$\begin{cases} C = \partial\omega = \{(x, y) \in \Omega : \phi(x, y) = 0\}, \\ \text{inside } (C) = \omega = \{(x, y) \in \Omega : \phi(x, y) > 0\} \\ \text{outside } (C) = \Omega \setminus \bar{\omega} = \{(x, y) \in \Omega : \phi(x, y) < 0\} \end{cases}, \quad (5)$$

where Ω is an image and C is a curve in the image. ω expresses the object area. Therefore, $C \subset \Omega$, $\omega \subset \Omega$, $C = \partial\omega$. *inside* (C) denotes the region ω , and *outside* (C) denotes the region $\Omega \setminus \bar{\omega}$. Energy is calculated by $F(C_1, C_2, C) = F(C_1, C_2, \phi)$ in equation 6 to develop the curve in new area and position.

$$\begin{aligned} F(C_1, C_2, \phi) = & \mu \int_{\Omega} |\nabla H(\phi)| dx dy + v \int_{\Omega} \nabla H(\phi) dx dy \\ & + \lambda_1 \int_{\Omega} \nabla |u_0(x, y) - C_1|^2 H(\phi) dx dy \\ & + \lambda_2 \int_{\Omega} \nabla |u_0(x, y) - C_2|^2 (1 - H(\phi)) dx dy \end{aligned} \quad (6)$$

Euclidean distance was used to measure the color distance between two colors. Like in [36], we set $\lambda_1 = \lambda_2 = 1$ and $v = 0$. The length parameter μ has a scaling role. For small values of μ , smaller objects will be detected; for larger values of μ , only larger objects are detected, or objects formed by grouping [36]. Since lesions in dermoscopic images are not very small, in our experiments, the value of μ was set to 0.1×255^2 . Equations (7) and (8) show the calculation of C_1 and C_2 .

$$C_1(\phi) = \frac{\int_{\Omega} v_0(x, y) H(\phi(x, y)) dx dy}{\int_{\Omega} H(\phi(x, y)) dx dy} \quad \text{if } \int_{\Omega} H(\phi(x, y)) dx dy > 0, \quad (7)$$

$$C_2(\phi) = \frac{\int_{\Omega} v_0(x, y) (1 - H(\phi(x, y))) dx dy}{\int_{\Omega} (1 - H(\phi(x, y))) dx dy}, \quad \text{if } \int_{\Omega} 1 - (H(\phi(x, y))) dx dy > 0, \quad (8)$$

where $v_0(x, y)$ is a pixel value of the image and $H(\phi(x, y))$ is a heaviside function. From (7) and (8), C_1 and C_2 are weighted average values of v_0 inside C and outside C , respectively. More details about the heaviside function and other parameters of the region-based active contour can be found in [36]. We use the active contour method only for fine-tuning of the boundary of the segmented lesion. On the other hand, in some cases such as images in Figures 6a–6c, a lesion area contains some high-contrast regions. Large number of iterations causes the active contour algorithm to incorrectly converge to these high-contrast regions. Hence, the number of iterations in the active contour method is limited and set proportional to the size of the segmented lesion as follows:

$$\text{Iteration_number}_{Active \text{ contour}} = 0.03(H_i + W_i) \quad (9)$$

where H_i and W_i are the height and width of the lesion segmented by the DeepLab network. To improve the performance of both normalization and segmentation stages, it is better to obtain and use the outputs of some additional modes of an input image. In the normalization stage, totally 4 modes of a test image are

considered as follows:

- The input image
- Horizontal and vertical flipped versions of the input image. The estimated bounding boxes are flipped horizontally and vertically, respectively.
- The normalized image rotated by 180 degrees. The estimated bounding box is rotated by 180 degrees.

The final bounding box is computed by averaging coordinates of corners of four bounding boxes estimated according to the above modes. Moreover, in the segmentation stage, 8 different modes of a normalized image are considered as follows:

- The normalized image
- Horizontal and vertical flipped versions of the normalized image. The obtained binary images are flipped horizontally and vertically, respectively.
- The normalized image rotated by ± 90 , ± 45 , and $+180$ degrees. The obtained binary images are rotated by ∓ 90 , ∓ 45 , and 180 degrees, respectively.

The final output of the segmentation stage is a binary image and is achieved by combining these eight outputs. In the segmented image, a pixel is considered a lesion pixel, if the corresponding pixel has nonzero values at least in 3 out of 8 binary images of the above modes.

3. Experiments

3.1. Datasets

The proposed lesion segmentation method is evaluated on well-known datasets of skin lesions. PH2 dataset was collected by the dermatology service of Hospital Pedro Hispano and the research group of the Universidade do Porto [37]. It contains 200 dermoscopy images of size 768×560 pixels.

The ISBI 2016 and ISBI 2017 skin lesion challenge datasets were prepared by the International Skin Imaging Collaboration (ISIC). ISBI 2016 contains 900 and 379 images for training and testing, respectively. Image sizes of this dataset vary from 566×679 pixels to 2848×4228 pixels [38].

ISBI 2017 dataset is the latest version of ISBI datasets that includes the segmentation ground truth for all the training, test, and validation sets, and is available online at [39]. It contains 2750 dermoscopy images ranging in size from 540×722 to 4448×6748 pixels of which 2000 are intended for training, along with 150 and 600 images for validation and testing, respectively.

DermQuest is a dataset of nondermoscopic skin images. It consisted of 137 images of size 1043×1640 pixels [40].

3.2. Data augmentation

As can be seen in Figure 3, three different train sets were constructed by augmentation of the official train set of the ISBI 2017 dataset. The first, second, and third sets consisted of 5000, 5000, and 7000 images, respectively. Combinations of some geometric transformations were applied on the images of the official train set of the ISBI 2017 and three new sets (Set A, Set B, and Set C) each one including 5000 transformed RGB images were produced. Applied geometric transformations were rotation, random horizontal and vertical flipping, image resizing and cropping, and brightness changing. The rotation angle, resizing factor, and brightness changing factor in the augmentation procedure were randomly selected from $[0 \ 180]$, $[0.8 \ 1.2]$, and $[0.7 \ 1.2]$, respectively.

Set A was used for training ODNs. Three trained ODNs were applied on each image of the Set B. As shown in Figure 3, by concatenation I_{Mask} , I_{Retina} , I_{YOLO} , and the input RGB image of ODNs, a 6-channel image I_{Comb} was made. By performing this procedure for all images of the Set B, a set of 5000 6-channel images was constructed which was used for training the CombNet. By adding 2000 original training images to the images of the Set C, Set C2 including 7000 images was constructed. By applying the procedure illustrated in right side of Figure 3, a set of 7000 normal RGB images was made which was used for training of the DeepLab3+ network.

In addition, 150 images of the official validation set of the ISBI 2017 were augmented and a total of 750 images were obtained which were used as the validation set to prevent overfitting in training the DeepLab3+ network.

3.3. Evaluation metrics

For evaluating semantic segmentation methods, the following metrics have been used in the literature. Sensitivity (SEN) represents the rate of pixels of correctly detected skin lesion. Specificity (SPE) is the rate of pixels of nonskin lesions classified correctly [41]. The Jaccard index (JAC) is an intersection over union (IOU) of the result of the segmented lesions with the ground truth masks. Index of Dice (DIC) measures the similarity of the segmented lesions through ground truth. Accuracy (ACC) shows the overall performance of the segmentation [42]. All these criteria have been computed from confusion matrix elements as follows:

$$SEN = \frac{TP}{TP + FN}, \quad (10)$$

$$SPE = \frac{TN}{TN + FP}, \quad (11)$$

$$JAC = \frac{TP}{TP + FN + FP}, \quad (12)$$

$$DIC = \frac{2 \cdot TP}{(2 \cdot TP) + FP + FN}, \quad (13)$$

$$ACC = \frac{TP + TN}{TP + FP + TN + FN}. \quad (14)$$

3.4. Results

In skin lesion segmentation applications, the most important metric is the Jaccard index, as in the ISBI challenges, different methods have been ranked based on their Jaccard values. The Jaccard values obtained by the proposed method over the ISBI 2017 dataset are reported in Table 1. In the first, second, and third rows of this table, only one ODN, respectively Mask RCNN, RetinaNet, and Yolov3 was used for lesion detection. In the last 3 rows, combinations of ODNs were used. In the fourth row, the results of ODNs were combined by using an averaging approach, in which the height, width, and coordinates of the top-left corner of the final bounding box were calculated by averaging those of the bounding boxes estimated by ODNs. Table 2 compares the proposed method and other methods over the ISBI 2017 dataset. Moreover, a comparison among different methods in terms of the Jaccard index over PH2, ISBI 2016, and ISBI 2017 datasets are given in Table 3. To

evaluate our method on a nondermoscopic skin images, the results on the DermQuest dataset were obtained and reported in Table 4. Results reported in Tables 2-4 show that our method outperformed other state-of-the-art methods over different well-known datasets. All experiments were performed by using a 6GB NVIDIA GeForce RTX2060 graphic card. The values of the learning rate and mini-batch size of all networks were set to 0.001 and 10 images, respectively.

Table 1. Comparison among various modes of the proposed methods over the ISBI 2017 dataset.

Lesion detection				Segmentation	Postprocessing	
Mask R-CNN	RetinaNet	Yolov3	Combination approach	DeepLab3+	Active contour	<i>JAC</i>
×				×		78.82
	×			×		78.67
		×		×		78.84
×	×	×	Averaging	×		79.08
×	×	×	CombNet	×		79.63
×	×	×	CombNet	×	×	80.02

Table 2. Comparison among various methods based on different metrics over the ISBI 2017 dataset

	<i>SEN</i>	<i>SPE</i>	<i>ACC</i>	<i>DIC</i>	<i>JAC</i>
CDNN [5]	82.50	97.50	93.40	84.90	76.50
Li et al. [6]	82.00	97.80	93.20	84.70	76.20
ResNet [43]	82.20	98.50	93.40	84.40	76.00
U-Net [44]	-	-	-	77.00	62.00
FrCN [7]	85.40	96.69	94.03	87.08	77.11
DermoNet [8]	-	-	-	-	78.30
DSNet [45]	87.5	95.5	-	-	77.5
Tschandl et al. [46]	-	-	-	-	77.0
FAC-Net [10]	74.53	99.35	95.94	86.97	78.85
CSAG and DCCNet [11]	81.06	97.43	93.63	84.91	74.27
FAT Net [12]	83.92	97.25	93.26	85.0	76.53
Ms RED [13]	-	-	94.10	86.46	78.55
Kaur et al. [47]	-	-	94.3	65.9	77.8
Proposed method	88.56	96.25	94.37	87.62	80.02

4. Discussion

In the normalization stage, the most important part is the detection of the skin lesions. In our method, three state-of-the-art ODNs have been used. These ODNs have three different structures and three different backbone networks. In other words, each ODN detected lesions by completely different approaches. Proper combination of different methods can improve the performance. We proposed a novel combination method to improve the detection accuracy and consequently the final segmentation performance. As shown in Table 1, combining the

Table 3. Comparison among various methods in terms of Jaccard index over different datasets

	PH2	ISBI 2016	ISBI 2017
FrCN [7]	84.79	-	77.11
DermoNet [8]	84.3	81.1	78.30
DSNet [45]	87.00	-	77.5
SDRR [48]	76.00	67.00	-
MSCA [49]	72.33	66.19	47.93
SSLS [50]	68.16	57.20	44.77
FCN [51]	82.15	81.37	73.12
MFCN [52]	83.99	84.64	-
DCL-PSI [53]	85.90	85.92	77.73
CSAG and DCCNet [11]	-	86.23	74.27
FAT Net [12]	89.62	85.30	76.53
Ms RED [13]	90.14	87.03	78.55
Kaur et al. [47]	88.8	87.1	77.8
Proposed method	89.96	87.72	80.02

Table 4. Comparison among various methods based on different metrics over the DermQuest dataset

	<i>SEN</i>	<i>SPE</i>	<i>ACC</i>	<i>DIC</i>	<i>JAC</i>
L-SRM [54]	89.4	92.7	92.3	82.2	70.0
Otsu-R [55]	87.3	85.4	84.9	73.7	64.9
Otsu-RGB [56]	93.6	80.3	80.2	69.3	59.1
Otsu-PCA [57]	79.6	99.6	98.1	83.3	75.2
TDLS [58]	91.2	99.0	98.3	82.8	71.5
Jafari [3]	95.2	99.0	98.7	83.1	81.2
FCN-8s [51]	90.0	99.5	98.9	89.7	82.9
U-net [59]	91.5	99.5	98.7	88.7	81.4
Yuan [4]	91.6	99.6	98.7	89.3	82.0
DFCN-No Interlace [60]	78.0	99.8	98.0	81.62	73.0
DFCN [60]	92.4	99.6	98.9	91.6	85.2
Proposed method	94.87	99.41	99.06	92.31	85.86

results of ODNs using the CombNet improved the Jaccard value by 0.81%, 0.96%, and 0.79% compared to when using only the Mask R-CNN, Retinanet, or Yolov3 as a lesion detector over the ISBI 2017 dataset. To the best of our knowledge, the combination of the object detectors was not used to improve the detection accuracy, especially in the skin lesion detection. Table 1 shows that the averaging combination approach has slightly increased the Jaccard value. However, the final Jaccard index has been significantly increased by using the proposed combination approach. The reason is that the bounding boxes estimated by ODNs were converted to image matrices and concatenated with the input image to be applied at the input of the CombNet. Convolutional layers of the CombNet extracted features from the input image and the results of ODNs, simultaneously. In other words, not only the results of ODNs, but also the features of the lesion in the input image were considered

to estimate the final bounding box. In the training phase of the DeepLab, the network was learned to create the best segmentation masks for a large number of skin images with various specifications at its output. In other words, the global features of lesions in images were considered much more than the image-specific features. In the postprocessing stage, we used the active contour method to employ the image-specific features. Using the Chan-Vese active contour method in the postprocessing stage, 0.39% improvement in Jaccard value has been obtained. This region-based active contour method allowed the boundary of the segmented lesion to be slightly modified. Figure 6 reveals the results of applying our postprocessing to some lesions segmented by the Deeplab3+ segmentation structure. These images are difficult samples from test set of the ISBI 2017 dataset. Very low-contrast borders of lesions, existence of hairs and blood vessels, and nonuniform lesions are of the main challenges in these images. The green, red, and blue colors in this figure show the boundaries of the lesions respectively in the ground-truth, the segmented lesions by the Deeplab, and the final segmented lesion after postprocessing. As shown, after applying the active contour method, the boundary of the lesions have been closer to the correct boundary, because the active contour method considered the image-specific features of lesions and surrounded tissues for each image. In Figures 6c–6h, boundaries of lesions were extended by the active contour. The postprocessing stage limited boundaries of lesions in Figures 6i and 6j. The results in Figure 6 and improving the Jaccard value in Table 1 demonstrated that the active contour method worked well as the postprocessing stage. Figure 7 displays the distributions of various Jaccard values of the test images obtained by the proposed method over the ISBI 2017 dataset. Blue, orange, green, and red colors in this figure are respectively related to the first, second, third, and sixth rows of the Table 1. It can be observed that by using the combination of ODNs as well as fine-tuning the results of the DeepLab by the active contour method the distribution of the Jaccard index has tended to the higher values. Results in Tables 1–3 showed that the proposed method performed well on four different datasets consisting of various dermoscopic and nondermoscopic images. It demonstrated that the proposed method is a robust lesion segmentation method.

5. Conclusion

In this paper, a three-stage method based on CNNs and active contour method was proposed to improve the performance of the skin lesion segmentation. In the first stage, which is the most important stage of the proposed method, the size and location of the lesion were estimated and the cropped image of the detected lesion was normalized. To improve the lesion detection accuracy in the normalization stage, a novel CNN-based method was proposed to combine the bounding boxes estimated by Mask R-CNN, RetinaNet, and Yolov3 object detection networks (ODNs). The experiments showed that the combination of ODNs' results using the proposed CombNet improved the Jaccard value from 78.84% to 79.63% over the ISBI 2017 dataset. In the segmentation stage, a DeepLab3+ structure with the VGG19 pretrained network as the backbone segmented the lesions from the normalized images. To improve the segmentation performance, a postprocessing stage based on a region-based active contours was used. Application of the active contour method to the segmented lesions adjusted the lesions boundaries so that they were closer to the correct boundaries. The proposed method was evaluated on well-known skin lesion datasets. The ISIB2017 dataset is the latest version datasets of ISIC for which the ground truth images of validation and test sets are available. The value of the Jaccard index of the proposed method on this dataset reached 80.02% which was 2.91% and 1.72% greater than the Jaccard values of the state-of-the-art methods proposed in [7] and [8], respectively.

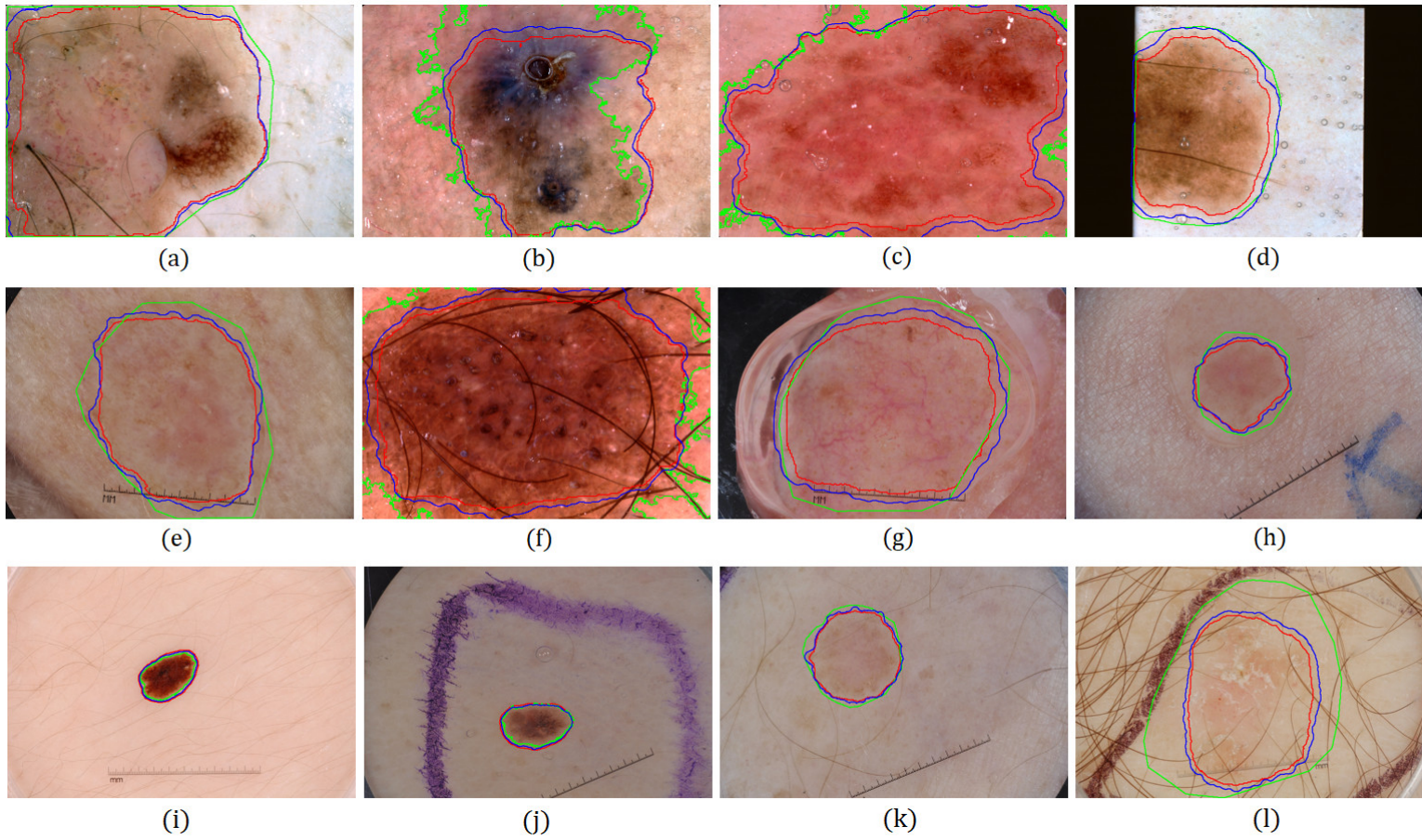


Figure 6. Improvement of the lesion segmentation by applying the active-contour method as postprocessing. The green, red, and blue curves show the boundaries of the lesions in the ground-truth, the segmented lesions by the Deeplab, and the final segmented lesion after applying the active-contour method, respectively.

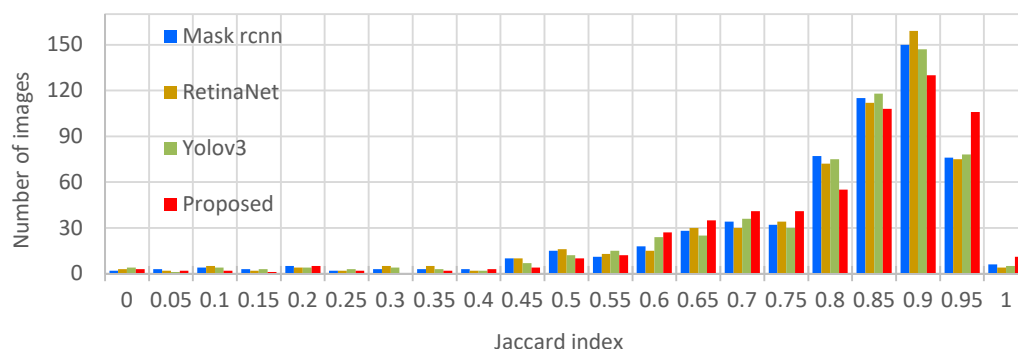


Figure 7. Distribution of various Jaccard values of the test images.

References

- [1] Silveira M, Nascimento J. Comparison of segmentation methods for melanoma diagnosis in dermoscopy images. *IEEE Journal of Selected Topics in Signal Processing* 2009; 3 (1) :35–45. doi: 10.1109/JSTSP.2008.2011119
- [2] Yu L, Chen H, Dou Q, Qin J, Heng P A. Automated melanoma recognition in dermoscopy images via very deep residual networks. *IEEE Transactions on Medical Imaging* 2017; 36 (4) :994–1004. doi: 10.1109/TMI.2016.2642839
- [3] Jafari M H, Nasr-Esfahani E, Karimi N, Soroushmehr S, Samavi S et al. Extraction of skin lesions from non-dermoscopic images using deep learning. *International journal of computer assisted radiology and surgery* 2017; 12 (6) :1021–1030. doi: <https://doi.org/10.1007/s11548-017-1567-8>
- [4] Yuan Y, Chao M, Lo Y C. Automatic skin lesion segmentation using deep fully convolutional networks with Jaccard distance. *IEEE Transactions on Medical Imaging* 2017; 36 (9) :1876–1886. doi: 10.1109/tmi.2017.2695227
- [5] Yuan Y. Automatic skin lesion segmentation with fully convolutional-deconvolutional networks. *arXiv preprint arXiv:1703.05165*. 2017; doi: 10.1109/JBHI.2017.2787487
- [6] Li Y, Shen L. Skin lesion analysis towards melanoma detection using deep learning network. *Sensors* 2018; 18 (2) :1–16. doi: 10.3390/s18020556
- [7] Al-masni MA, Al-antari MA, Choi M, Han S, Kim T. Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks. *Computer Methods and Programs in Biomedicine* 2018; 162 :221–231. doi: 10.1016/j.cmpb.2018.05.027
- [8] Baghersalimi S, Bozorgtabar B, Schmid-saugeon P, Ekenel HK, Thiran J. DermoNet: densely linked convolutional neural network for efficient skin lesion segmentation. *EURASIP Journal on Image and Video Processing* 2019; 71 :1–10. doi: <https://doi.org/10.1186/s13640-019-0467-y>
- [9] Qian C, Liu T, Jiang H, Wang Z, Wang P et al. A Detection and Segmentation Architecture for Skin Lesion Segmentation on Dermoscopy Images. 2018 :2–7. [Online]. Available: <http://arxiv.org/abs/1809.03917>
- [10] Tao S, Jiang Y, Cao S, Wu C, Ma Z. Attention-Guided Network with Densely Connected Convolution for Skin Lesion Segmentation. *Sensors* 2021; 21 (10): 3462. doi: 10.3390/s21103462
- [11] Dong Y, Wang L, Cheng S, Li Y. FAC-Net: Feedback Attention Network Based on Context Encoder Network for Skin Lesion Segmentation. *Sensors* 2021; 21 (15):5172. doi: 10.3390/s21155172
- [12] Wu H, Chen S, Chen G, Wang W, Lei B et al. FAT-Net: Feature adaptive transformers for automated skin lesion segmentation, *Medical Image Analysis* 2022; 76 (9): 102327. doi: 10.1016/j.media.2021.102327
- [13] Dai D, Dong C, Xu S, Yan Q, Li Z et al. Ms RED: A novel multi-scale residual encoding and decoding network for skin lesion segmentation, *Medical Image Analysis* 2022; 75 (6): 102293. doi: 10.1016/j.media.2021.102293

- [14] Litjens G, Kooi T, Bejnordi BE, Arindra A, Setio A et al. A survey on deep learning in medical image analysis. *Medical Image Analysis* 2017; 42 (December 2012) :60–88. doi: 10.1016/j.media.2017.07.005
- [15] Shen W, Yang F, Mu W, Yang C, Yang X et al. Automatic localization of vertebrae based on convolutional neural networks. *Proceedings of the SPIE on Medical Imaging* 2015; 9413 :94132E. doi: 10.1117/12.2081941
- [16] Kawahara J, Hamarneh G. Multi-resolution-tract CNN with hybrid pretrained and skin-lesion trained layers. *Machine learning in medical imaging*. Cham: Springer International Publishing 2016; 10019 :164–171.
- [17] Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 2014:580–587. doi: 10.1109/CVPR.2014.81
- [18] Girshick R. Fast R-CNN. 2015 *IEEE International Conference on Computer Vision (ICCV)* 2015; doi: 10.1109/ICCV.2015.169
- [19] Ren S, He K, Girshick R, Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2017; 39 (6) :1137–1149. doi: 10.1109/TPAMI.2016.2577031
- [20] He K, Gkioxari G, Dollár P, Girshick R. Mask R-CNN. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2020; 42 (2) :386–397. doi: 10.1109/TPAMI.2018.2844175
- [21] Liu W, Anguelov D, Erhan D, Szegedy C, Reed S et al. SSD: Single Shot MultiBox Detector. *Computer Vision – ECCV 2016*. Lecture Notes in Computer Science vol 9905. Springer Cham 2016; doi: 10.1007/978-3-319-46448-0-2
- [22] Redmon J, Divvala S, Girshick R, Farhadi A. You Only Look Once: Unified Real-Time Object Detection. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 2016; doi: 10.1109/CVPR.2016.91
- [23] Lin TY, Goyal P, Girshick R, He K, Dollar P. Focal Loss for Dense Object Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2020; 42 (2) :318–327. doi: 10.1109/TPAMI.2018.2858826
- [24] Redmon J, Farhadi A. YOLO9000: Better, faster, stronger. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 2017; doi: 10.1109/CVPR.2017.690
- [25] Redmon J, Farhadi A. YOLOv3: An Incremental Improvement. *arXiv:1804.02767v1* 2018; [Online]. Available: <http://arxiv.org/abs/1804.02767>
- [26] Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille A L. Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs. *International Conference on Learning Representations* 2015
- [27] Chen L, Zhu Y, Papandreou G, Schroff F, Adam H. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. *Computer Vision – ECCV 2018*. *ECCV 2018*. Lecture Notes in Computer Science 2018; 11211 doi: <https://doi.org/10.1007/978-3-030-01234-2-49>
- [28] Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille AL. “DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2017; 40 (4) :834–848. doi: 10.1109/TPAMI.2017.2699184
- [29] Chen LC, Papandreou G, Schroff F, Adam H. Rethinking Atrous Convolution for Semantic Image Segmentation. In *arXiv preprint* <http://arxiv.org/abs/1706.05587> 2017; [Online]. Available: arxiv.org/abs/1706.05587
- [30] Erkol B, Moss RH, Stanley RJ, Stoecker WV, Hvatum E. Automatic lesion boundary detection in dermoscopy Images Using Gradient Vector Flow Snakes. *Skin Research and Technology* 2005; 11 (1) :17–26. doi: 10.1111/j.1600-0846.2005.00092.x
- [31] Abbas Q, Fondón I, Sarmiento A, Celebi ME. An Improved Segmentation Method for Non-melanoma Skin Lesions Using Active Contour Model. In: *Image Analysis and Recognition*. *ICIAR 2014*. Lecture Notes in Computer Science vol 8815. Springer Cham 2014; 8815 :193–200. doi: 10.1007/978-3-319-11755-3
- [32] Kass M, Witkin A, Terzopoulos D. Snakes: Active Contour Models. *International Journal of Computer Vision* 1998; 1 (4) :321–331. doi: 10.1016/B978-0-12-386454-3.00786-7

- [33] Mete M, Sirakov NM. Lesion detection in dermoscopy images with novel density-based and active contour approaches. *BMC Bioinformatics* 2010; 11 (S23) doi: 10.1186/1471-2105-11-23
- [34] Caselles V, Kimmel R, SAPIRO G. Geodesic Active Contours. *International Journal of Computer Vision* 1997; 22 (1) :61–79. doi: 10.1023/A:1007979827043
- [35] Vese LA, Chan TF. A Multiphase Level Set Framework for Image Segmentation Using the Mumford and Shah Model. *International Journal of Computer Vision* 2002; 50 (3) :271–293. doi: 10.1023/A:1020874308076
- [36] Chan TF, Vese LA. Active contours without edges. *IEEE Transactions on Image Processing* 2001; 10 (2) :266–277. doi: 10.1109/83.902291
- [37] Mendonça T, Ferreira PM, Marques JS, Marcal AR, Rozeira J. PH2 - A dermoscopic image database for research and benchmarking. 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2013 :5437–5440. doi: 10.1109/EMBC.2013.6610779
- [38] Gutman D, Codella NC, Celebi E, Helba B, Marchetti M 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;
- [39] Codella NCF, Gutman D, Celebi ME, Helba B, Marchetti MA et al. Skin Lesion Analysis Toward Melanoma Detection: a Challenge at The 2017 International Symposium on Biomedical Imaging (ISBI) Hosted By The International Skin Imaging Collaboration (ISIC). 2018 IEEE 15th International Symposium on Biomedical Imaging 2018; doi: 10.1109/ISBI.2018.8363547
- [40] DermQuest, The art, science and practice of dermatology, <http://www.dermquest.com>, 2010.
- [41] Al-antari MA, Al-masni MA, Park SU, Park JH, Metwally MK et al. An automatic computer-aided diagnosis system for breast cancer in digital mammograms via deep belief network. *Journal of Medical and Biological Engineering* 2018; 38 :443–456. doi: 10.1007/S40846-017-0321-6
- [42] Powers D. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies* 2011; 2 (1) :37–63.
- [43] Bi L, Kim J, Ahn E, Feng D. Automatic Skin Lesion Analysis using Large-scale Dermoscopy Images and Deep Residual Networks. <http://arxiv.org/abs/1703.04197> 2017 :6–9. [Online]. Available: arxiv.org/abs/1703.04197
- [44] Lin BS, Michael K, Kalra S, Tizhoosh HR. Skin lesion segmentation: UNets versus clustering. 2017 IEEE Symposium Series on Computational Intelligence (SSCI) 2017; doi: 10.1109/SSCI.2017.8280804
- [45] Hasan MK, Dahal L, Samarakoon PN, Tushar FI, Martí R. DSNet: Automatic dermoscopic skin lesion segmentation. *Computers in Biology and Medicine* 2020; 120 (April) :103738. doi: 10.1016/j.compbiomed.2020.103738
- [46] Tschandl P, Sinz C, Kittler H. Domain-specific classification-pretrained fully convolutional network encoders for skin lesion segmentation. *Computers in Biology and Medicine* 2019; 104 :111–116. doi: 10.1016/j.compbiomed.2018.11.010
- [47] Kaur R, GholamHosseini H, Sinha R. Skin lesion segmentation using an improved framework of encoder-decoder based convolutional neural network, *International Journal of Imaging Systems and Technology* 2022: 1-16 doi:10.1002/ima.22699
- [48] Bozorgtabar B, Abedini M, Garnavi R. Sparse Coding Based Skin Lesion Segmentation Using Dynamic Rule-based Refinement Behzad. 7th International Conference on Machine Learning in Medical Imaging In conjunction with MICCAI 2016 2016; doi: 10.1007/978-3-319-47157-0
- [49] Bi L, Kim J, Ahn E, Feng D, Fulham M. Automated skin lesion segmentation via image-wise supervised learning and multi-scale superpixel based cellular automata. 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI) 2016 :1059–1062. doi: 10.1109/ISBI.2016.7493448

- [50] Ahn E, Kim J, Bi L, Kumar A, Li C et al. Saliency-Based Lesion Segmentation Via Background Detection in Dermoscopic Images. *IEEE Journal of Biomedical and Health Informatics* 2017; 21 (6) :1685–1693. doi: 10.1109/JBHI.2017.2653179
- [51] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 2015 :3431–3440.
- [52] Bi L, Kim J, Ahn E, Kumar A, Fulham M et al. Dermoscopic Image Segmentation via Multistage Fully Convolutional Networks. *IEEE Transactions on Biomedical Engineering* 2017; 64 (9) :2065–2074. doi: 10.1109/TBME.2017.2712771
- [53] Bi L, Kim J, Ahn E, Kumar A, Feng D et al. Step-wise integration of deep class-specific learning for dermoscopic image segmentation. *Pattern Recognition* 2019; 85 :78–89. doi: 10.1016/j.patcog.2018.08.001
- [54] Celebi ME, Kingravi HA, Iyatomi H, Aslandogan YA, Stoecker WV et al. Border detection in dermoscopy images using statistical region merging. *Skin Research and Technology* 2008; 14 (3) :347–353. doi: 10.1111/j.1600-0846.2008.00301.x
- [55] Cavalcanti PG, Yari Y, Scharcanski J. Pigmented Skin Lesion Segmentation on Macroscopic Images Review of Recent Pigmented Skin Lesion Segmentation Methods. *2010 25th International Conference of Image and Vision Computing New Zealand 2010*; doi: 10.1109/IVCNZ.2010.6148845
- [56] Cavalcanti PG, Scharcanski J, Lopes CBO. Shading Attenuation in Human Skin Color Images. *Advances in Visual Computing Lecture Notes in Computer Science* 2010; 6453 :190–198. doi: 10.1007/978-3-642-17289-2
- [57] Cavalcanti PG, Scharcanski J. Automated prescreening of pigmented skin lesions using standard cameras. *Computerized Medical Imaging and Graphics* 2011; 35 (6) :481–491. doi: 10.1016/j.compmedimag.2011.02.007
- [58] Glaister J, Member S, Wong A, Clausi D, Member S. Segmentation of Skin Lesions From Digital Images Using Joint Statistical Texture Distinctiveness. *IEEE Transactions on Biomedical Engineering* 2014; 61 (4) :1220–1230. doi: 10.1109/TBME.2013.2297622
- [59] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention* 2015: 234–241. doi: 10.1007/978-3-319-24574-4-28
- [60] Nasr-Esfahani E, Rafei S, Jafari MH, Karimi N, Wrobel JS et al. Dense pooling layers in fully convolutional network for skin lesion segmentation. *Computerized Medical Imaging and Graphics* 2019; 78 :101658. doi: 10.1016/j.compmedimag.2019.101658