


RESEARCH

Open Access

# DermoNet: densely linked convolutional neural network for efficient skin lesion segmentation



Saleh Baghersalimi<sup>1</sup> , Behzad Bozorgtabar<sup>1</sup>, Philippe Schmid-Saugeon<sup>2</sup>, Hazım Kemal Ekenel<sup>3</sup> and Jean-Philippe Thiran<sup>1</sup>

## Abstract

Recent state-of-the-art methods for skin lesion segmentation are based on convolutional neural networks (CNNs). Even though these CNN-based segmentation approaches are accurate, they are computationally expensive. In this paper, we address this problem and propose an efficient fully convolutional neural network, named DermoNet. In DermoNet, due to our densely connected convolutional blocks and skip connections, network layers can reuse information from their preceding layers and ensure high accuracy in later network layers. By doing so, we take advantage of the capability of high-level feature representations learned at intermediate layers with varying scales and resolutions for lesion segmentation. Quantitative evaluation is conducted on three well-established public benchmark datasets: the ISBI 2016, ISBI 2017, and the PH2 datasets. The experimental results show that our proposed approach outperforms the state-of-the-art algorithms on these three datasets. We also compared the runtime performance of DermoNet with two other related architectures, which are fully convolutional networks and U-Net. The proposed approach is found to be faster and suitable for practical applications.

**Keywords:** Fully convolutional neural networks, Lesion segmentation

## 1 Introduction

Skin lesion segmentation is a key step in computerized analysis of dermoscopic images. Inaccurate segmentation could adversely impact the subsequent steps of an automated computer-aided skin cancer diagnosis system. However, this task is not trivial due to a number of reasons, such as the significant diversity among the lesions; inconsistent pigmentation; presence of various artifacts, e.g., air bubbles and fiducial markers; and low contrast between lesion and the surrounding skin, as can be seen in Fig. 1.

In recent years, we have witnessed major advances of convolutional neural networks (CNNs) in many image processing and computer vision tasks, such as object detection [1], image classification [2], and semantic image

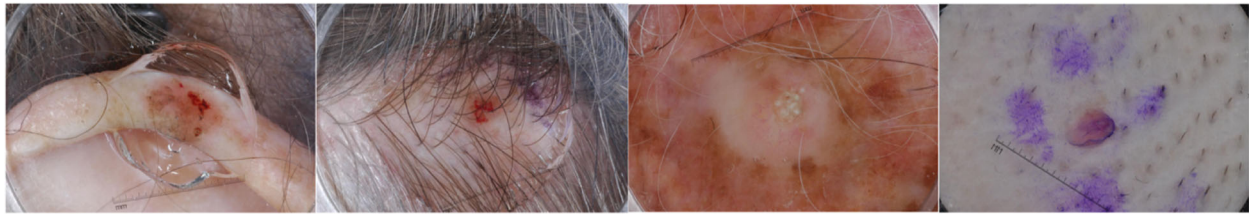
segmentation [3]. A well-known CNN-based segmentation approach, fully convolutional networks (FCNs) [3], tackles per pixel prediction problems by replacing the fully connected layers with convolutions which kernels can cover the entire input image regions. Doing so, FCNs can process any image size and output pixel-wise labeled prediction map. However, the pooling layers in a down-sampling path cause a loss in the image resolution and make the network fragile to handle the lesion boundary details, e.g., fuzzy boundaries. In addition, the fully convolutional layers contain a large number of parameters, which produce a computationally expensive network.

Most of the CNN approaches, such as SegNet [4] and DeconvNet [5], developed for segmentation purposes by using the encoder-decoder structure as the core of their network architecture. Another effective segmentation network is the employment of skip connections for the U-Net [6]. The encoder part is responsible for extracting the coarse features. It is followed by the decoder, which upsamples the features and is trained to recover the input image resolution at the network output. These

\*Correspondence: [saleh.baghersalimi@epfl.ch](mailto:saleh.baghersalimi@epfl.ch)

<sup>1</sup>Electrical Engineering Department, Signal Processing Laboratory (LTS5), École Polytechnique Fédérale de Lausanne (EPFL), Station 11, 1015 Lausanne, Switzerland

Full list of author information is available at the end of the article



**Fig. 1** Sample dermoscopic images from the ISBI 2017 Challenge: Skin Lesion Analysis Toward Melanoma Detection. The presence of artifacts such as hairs on the skin and inconsistent pigmentation making accurate skin lesion segmentation difficult

CNN architectures [4, 5] use a base network adopted from VGG architecture [7], which is already pre-trained based on millions of images. Having said that, they utilize the deconvolution or unpooling layers to recover fine-grained information from the downsampling layers.

Inspired by the residual networks (ResNets) [2], recently, a CNN architecture called DenseNet was introduced in [8]. The core components of the DenseNet are the dense blocks, where each block performs iterative summation of features from the previous network layers. This characteristic enables DenseNet to be more efficient, since it needs fewer parameters. Moreover, each layer can easily access their preceding layers; therefore, it reuses features of all layers with varying scales.

Even though deep convolutional neural networks have been a significant success for the image pixel-wise segmentation, their inefficiency in terms of computational time limits their capability for real-time and practical applications. The motivation for this work is to propose an efficient network architecture for skin lesion segmentation, while achieving the state-of-the-art results. Our contributions can be summarized as follows.

1. Our main aim is to perform an efficient segmentation under limited computational resources, while achieving the state-of-the-art results on skin benchmark datasets.
2. We transform the DenseNets into a fully convolutional network. In particular, our architecture is built from multiple dense blocks in the encoder part, and we add a decoder part to recover the full input image resolution. This helps the multi-scale feature maps from different layers to be penalized by a loss function.
3. The multiple skip connections are arranged between encoder and decoder. In particular, we link the output of each dense block with its corresponding decoder at each feature resolution. Doing so will enable the network to process high-resolution features from early layers as well as high-semantic features of deeper layers.
4. Since we only upsample the feature maps produced by the preceding dense block, the proposed network uses fewer parameters. This enables the network

achieve the best accuracy within a limited computational budget. We have conducted extensive experiments on ISBI 2016, ISBI 2017, and PH2 datasets, and we have shown that the proposed approach is superior to the state-of-the-art skin lesion segmentation methods.

The rest of this paper is organized as follows: Section 2 presents the related work. Section 3 describes the proposed network architecture in detail. Section 4 conveys and discusses the experimental results. Finally, section 5 concludes the paper.

## 2 Related work

Recently, deep learning has ushered in a new era of computer vision and image analysis. It is even more remarkable that the trained models on big dataset seem to transfer to many other problems such as detection technology [1, 9, 10] and semantic segmentation [3]. In particular, recent works on applying CNNs to image segmentation demonstrate superior performance over classical methods in terms of accuracy. In particular, convolutional neural networks can be adapted to FCNs [3] and perform semantic segmentation by replacing the fully connected layer of a classification network with a convolutional layer. However, due to the resolution loss in the down-sampling steps, the predicted lesion segmentation lacks lesion boundary details. Recently, several alternatives have been presented in the literature to address this shortcoming in FCNs. SegNet [4] and DeconvNet [5] are two examples of these approaches built upon auto-encoder network. In encoder, they both use the convolutional network from VGG16 for image classification. DeconvNet keeps two fully connected layers from VGG16, but SegNet discards them to decrease the number of parameters. Different from FCN in which the segmentation mask is recovered with only one deconvolution layer, the decoder network is composed of multiple deconvolution and unpooling layers both in SegNet and DeconvNet, which identify pixel-wise class labels and predict segmentation masks.

U-Nets [6] have shown to yield very good results in different segmentation benchmarks. In the U-Net architecture, there are skip connections from encoder layers to their corresponding decoder layers. These skip

connections help the decoder layers to recover the image details from the encoder part. As a result, a faster convergence and a more efficient optimization process are obtained during the training. Farabet et al. [11] proposed a segmentation method, where the raw input image is decomposed through a multi-scale convolutional network and produces a set of coarse-to-fine feature maps. Bozorgtabar et al. [12] proposed a skin segmentation method, which integrates fine and coarse prediction scores of the intermediate network layers. Simon et al. [13] used DenseNets to deal with the problem of semantic segmentation, where they achieved state-of-the-art results on urban scene benchmark datasets such as CamVid [14].

In addition, post-processing techniques such as conditional random fields (CRF) have been a popular choice to enforce consistency in the structure of the segmentation outputs [15]. Zheng et al. [16] proposed an interpretation of dense CRFs as recurrent neural networks (RNN). In their segmentation method, CRF-based probabilistic graphical modeling is integrated with deep learning techniques.

Our proposed DermoNet is based on fully convolutional neural network. Unlike the FCN, in the DermoNet architecture, the outputs of the encoders are linked into the corresponding decoder to recover lost spatial information. The main difference between DermoNet and U-Net is that the encoder in DermoNet consists of four dense blocks with each block having four layers, whereas the encoder of U-Net is a path followed by the typical architecture of a convolutional neural network as can be seen in Fig. 2.

### 3 Method

In this section, we propose a CNN-based architecture to perform lesion segmentation. Our network, DermoNet, consists of an encoder and a decoder; the encoder starts with a block, which performs the convolution on an input image with a kernel size of  $7 \times 7$  and a stride of 2, and

followed by the max pooling with stride of 2. In DermoNet, the output feature dimension of each layer within a dense block has  $k$  feature maps, where they are concatenated to the input. This procedure is repeated four times for each dense block; the output of the dense block is the concatenation of the outputs of the previous layers as in Eq. 1.

$$x_l = F_l ([x_{l-1}, x_{l-2}, \dots, x_0]) \tag{1}$$

where  $x_l$  denotes the output feature of the  $l$ th layer.  $F(\cdot)$  is a nonlinear function defined as a convolution followed by a rectifier non-linearity (ReLU), and  $[\dots]$  denotes the concatenation operator. By using dense blocks, we enable the network to process high-resolution features from early layers as well as high-semantic features of deeper layers.

Similar to the encoder, the decoder consists of four blocks, with each block having three layers. Each decoder block is composed of a convolutional layer with a kernel of size  $1 \times 1$ , a full-convolution layer with a kernel of size  $3 \times 3$  followed by an upsampling by a factor 2 and a convolutional layer with a kernel of size  $1 \times 1$ . The network ends with three last convolutional layers and two bilinear upsampling steps by a factor of 2 in order to generate a segmented image with the same size as the input. Table 1 presents the architectural details of the proposed DermoNet.

Figure 3 illustrates an overview of the proposed architecture; the encoder could be found on the right side of the figure while the decoder is shown on the left side.

Since FCNs perform the image pixel-wise classification, cross-entropy loss is usually used for the segmentation task. However, a skin lesion usually occupies a small portion of a skin image. Consequently, the segmentation network trained with cross-entropy loss function tends to be biased toward the background image rather than lesion itself. Different variants of the cross-entropy loss have

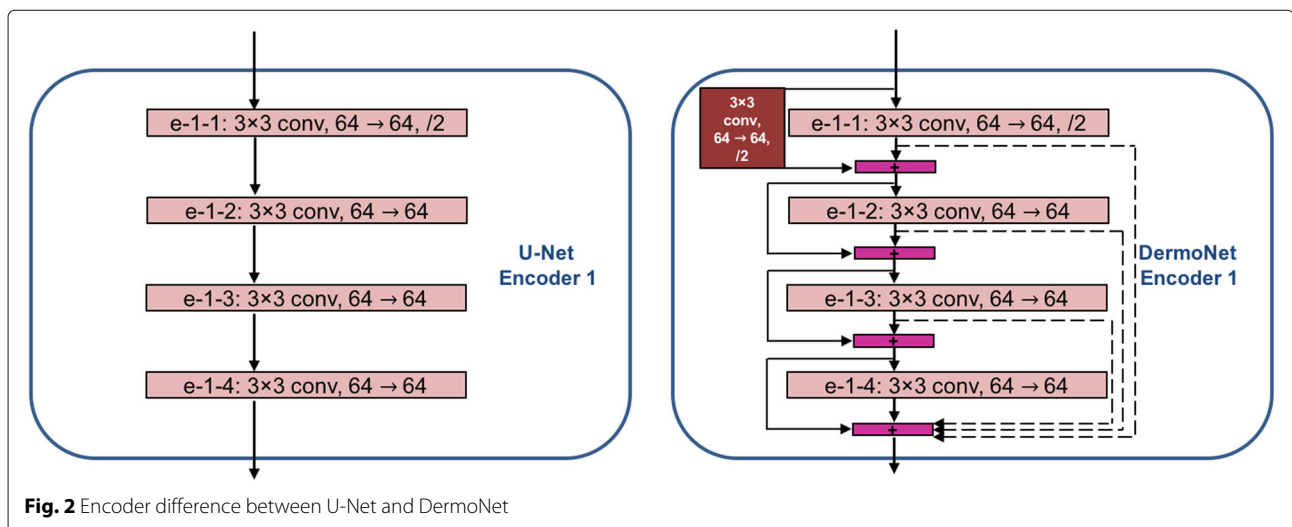


Fig. 2 Encoder difference between U-Net and DermoNet

**Table 1** Architectural details of the proposed DermoNet

Input image, 384×512×3					
Encoder			Decoder		
Input	Filter	Output	Input	Filter	Output
i-1	7×7, /2	192×256×64	d-4-1	1×1	6×8×128
i-2	3×3, /2	96×128×64	d-4-2	3×3, *2	12×16×128
e-1-1	3×3, /2	48×64×64	d-4-3	1×1	12×16×256
e-1-2	3×3	48×64×64	d-3-1	1×1	12×16×64
e-1-3	3×3	48×64×64	d-3-2	3×3, *2	24×32×64
e-1-4	3×3	48×64×64	d-3-3	1×1	24×32×128
e-2-1	3×3, /2	24×32×128	d-2-1	1×1	24×32×32
e-2-2	3×3	24×32×128	d-2-2	3×3, *2	48×64×32
e-2-3	3×3	24×32×128	d-2-3	1×1	48×64×64
e-2-4	3×3	24×32×128	d-1-1	1×1	48×64×16
e-3-1	3×3, /2	12×16×256	d-1-2	3×3, *2	96×128×16
e-3-2	3×3	12×16×256	d-1-3	1×1	96×128×64
e-3-3	3×3	12×16×256	o-1	3×3, *2	192×256×32
e-3-4	3×3	12×16×256	o-2	3×3	192×256×32
e-4-1	3×3, /2	6×8×512	o-3	2×2, *2	384×512×1
e-4-2	3×3	6×8×512			
e-4-3	3×3	6×8×512			
e-4-4	3×3	6×8×512			

Output image, 384×512×1

Here, /2 and \*2 denote downsampling operator using strided convolution and upsampling using a factor of 2, respectively.

been devised to address this problem, which focus on the class balancing [17]. However, this class balancing strategy brings additional computation cost during the training procedure. In this paper, we use a loss function based on Jaccard distance ( $L_J$ ) [18], which is complementary to the Jaccard index:

$$L_J = 1 - \frac{\sum_{i,j}(t_{ij}p_{ij})}{\sum_{i,j}t_{ij}^2 + \sum_{i,j}p_{ij}^2 - \sum_{i,j}(t_{ij}p_{ij})} \quad (2)$$

where  $t_{ij}$  and  $p_{ij}$  denote the target and prediction output at image pixel  $(i, j)$ , respectively. The Jaccard index measures the intersection over the union of the labeled segments for each class and reports the average. It takes into account both the false alarms and the missed values for each class. Our experimental results disclose that this loss function is more robust compared to the classical cross-entropy loss function. In addition, it is well suited to the imbalanced classes of the foreground and background, respectively.

#### 4 Results and discussion

The output of DermoNet model is binarized to a lesion and compared with the ground truth provided by clinicians. As the evaluation metrics, *Jaccard coefficient* (JC) and *Dice similarity coefficient* (DSC) are used,

which measure the spatial overlap between the obtained segmentation mask and the ground truth, respectively. They are defined as follows:

JC =  $\frac{TP}{TP+FN+FP}$  DSC =  $\frac{2 \times TP}{2 \times TP+FN+FP}$  where TP, FP, and FN denote the number of true positives, false positives, and false negatives, respectively.

#### 4.1 Datasets

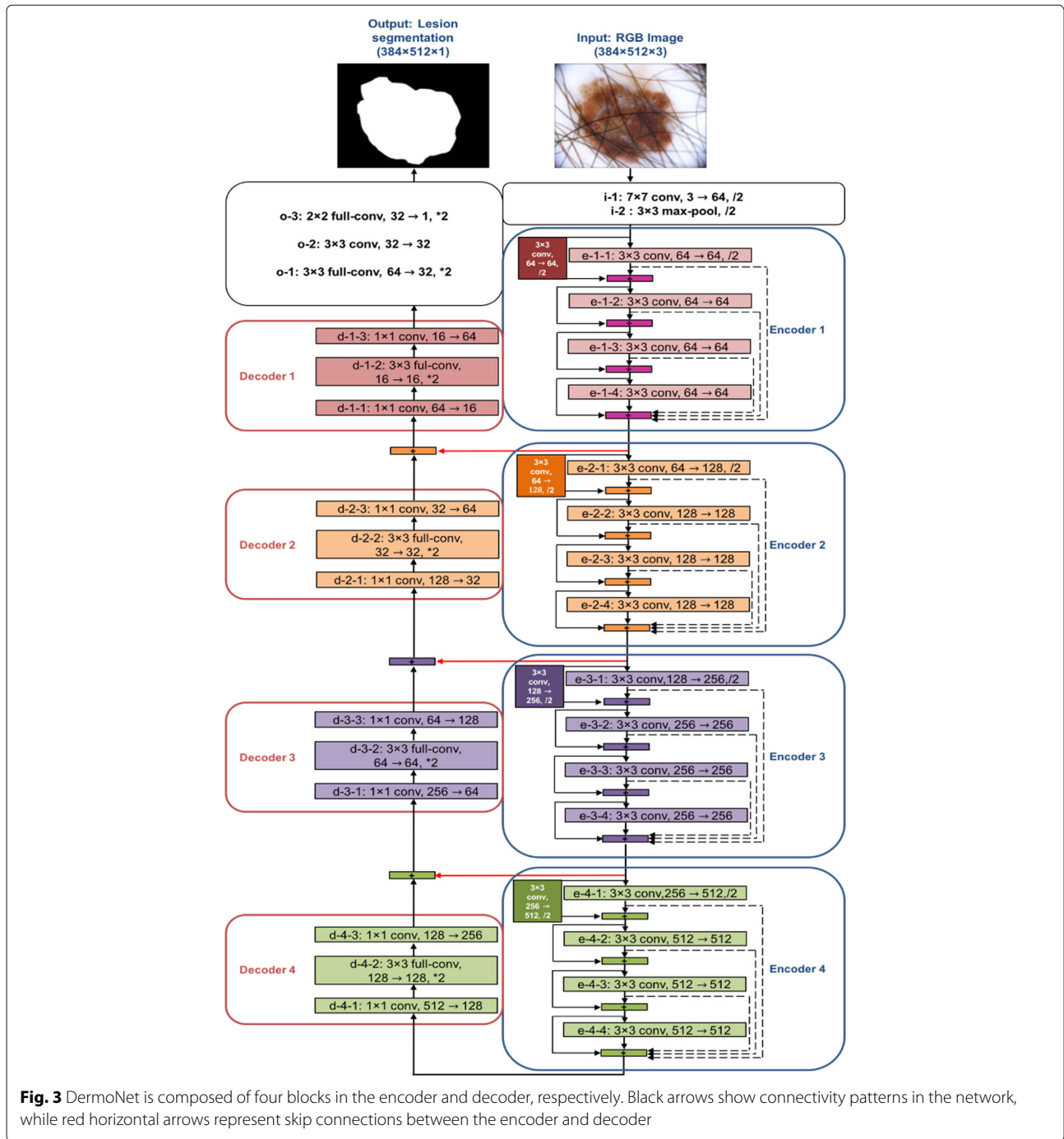
For the experiments, we have used the following three datasets :

**ISBI 2017:** This dataset [19] contains 2000 training dermoscopic images, while there are 600 test images with the ground truths provided by experts. The images sizes vary from 771 × 750 to 6748 × 4499.

**ISBI 2016:** This dataset [20] contains dermoscopic images, where the image sizes vary from 1022 × 767 to 4288 × 2848 pixels. There are 900 training images and 379 test images.

**PH2:** This dataset has been acquired at Dermatology Service of Hospital Pedro Hispano, Matosinhos, Portugal [21] with Tuebinger Mole Analyzer system. This dataset contains 200 dermoscopic test images with a resolution of 768 × 560 pixels.

Table 2 gives a summary about all three datasets.



**Table 2** Datasets summary

Dataset	No. of training set	No. of test set	Image sizes
ISBI 2017	2000	600	771 × 750 → 6748 × 4499
ISBI 2016	900	379	1022 × 767 → 4288 × 2848
PH2	-	200	768 × 560

### 4.2 Implementation details

We have trained our network using the resized RGB images of size  $384 \times 512$  pixels. For the augmentation, we flipped the training images horizontally and vertically and did shrinking via cropping. Then, we normalized each image such that the pixel values would be between 0 and 1. The initial weights of our network are sampled from *Xavier* initialization. Adam optimizer is used as the optimizer for the DermoNet. The base learning rate for the

network is set to  $10^{-4}$ . The maximum number of iteration is 5540. The whole architecture is implemented on the TensorFlow [22]. We used Nvidia Tesla K40 GPU with 12 GB GDDR5 memory for the training. We apply a threshold value of 0.5 to final pixel-wised score to generate lesion mask.

### 4.3 Runtime

To verify the effectiveness of the DermoNet in terms of test execution time, we compare it with two related architectures, namely FCN and U-Net. Table 3 presents the segmentation execution times per image using a system with Intel Core i7-5820K CPU. Due to the densely connected convolutional blocks and having less parameters, the proposed network is found to be faster.

### 4.4 Results on ISBI 2016 dataset

For the experiments on the ISBI 2016 dataset, for training the models, we used either only the training dataset provided by the ISBI 2016 challenge or the augmented version of it, in which we include 6500 dermoscopic images from DermoSafe [23] to the ISBI 2016 training dataset, in order to introduce a wider variety of images. These trained models are then evaluated on the ISBI 2016 test dataset. Obtained results on the ISBI 2016 challenge dataset are given in Table 4. In this challenge, the participants are ranked only based on the JC. In addition, we also report the DSC results. The proposed DermoNet improved the segmentation performance both in terms of Jaccard coefficient and Dice similarity coefficient. As can be seen from the table, in terms of JC, 9.9% and 2.2% absolute performance increase improvement has been achieved with respect to FCN and U-Net, respectively. In terms of DSC, the obtained absolute increase in performance with respect to FCN and U-Net is 7.8% and 1.8%, respectively. As can be seen from the last two rows of the table, DermoNet's performance improves with the use of the additional data provided by DermoSafe. However, even without using the additional DermoSafe data, it still outperforms the state-of-the-art methods. Figure 4 shows several examples of automatic segmentation results on the ISBI 2016 test set with different cases, such as hairy skin, irregular shape, and low contrast. We observe that the proposed DermoNet is able to separate the skin lesions from these artifacts and is robust to different image acquisition conditions.

### 4.5 Results on PH2 dataset

In these experiments, we have used the trained models obtained in Section 4.4 and evaluated them on the

**Table 3** Comparison of average runtime (s) per image

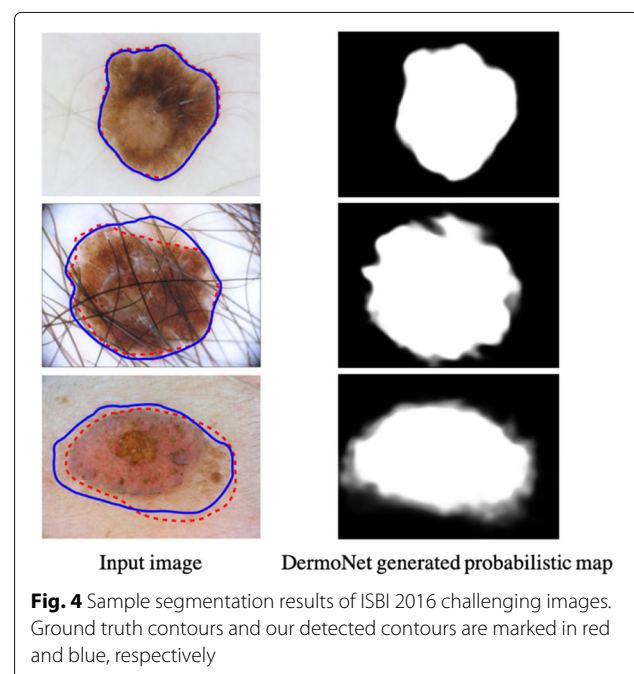
Method	FCN	U-Net	DermoNet
Run time (s)	0.145	0.092	0.081

**Table 4** Performance comparison between the proposed segmentation and other state-of-the-art methods on ISBI 2016 challenge test set

Method	JC (%)	DSC (%)
HED [24]	79.3	88.4
SegNet [4]	70.0	82.3
UiT-Seg [25]	80.6	NA
IHPC-CS [20]	79.9	NA
CNN-CRF [20]	79.7	NA
FCN - (JC) trained with DermoSafe data	72.6	81.6
U-Net - (JC) trained with DermoSafe data	80.3	87.6
DermoNet (JC) trained without DermoSafe data	81.1	88.2
DermoNet (JC) trained with DermoSafe data	82.5	89.4

Here, JC denotes Jaccard coefficient

200 skin images from the PH2 dataset. We have also compared the performance of the proposed lesion segmentation method with superpixel-based saliency detection approaches [26–28] on the PH2 dataset. Attained results are given in Table 5. From the experimental results, it can be observed that DermoNet which is trained using DermoSafe data has outperformed the other skin lesion segmentation methods. Due to dense connectivity in DermoNet, each layer is connected with all subsequent layers and allows later layers to bypass features and to maintain the high accuracy of the final pixel classification layer in a deeper architecture with fewer parameters. As a result, this brings additional performance gains.



**Table 5** Performance comparison between the proposed segmentation and other state-of-the-art methods on PH2 dataset

Method	JC (%)	DSC (%)
Adaptive thresholding [29]	72	80
Yen thresholding [30]	45	55
Level set active contours [31]	76	83
Statistical region growing [32]	43	61
Bootstrap segmentation [26]	60	75
Contextual hypergraph [27]	63	77
Multi-scale segmentation [28]	76	86
FCN - (JC) trained with DermoSafe data	75.8	84.7
U-Net - (JC) trained with DermoSafe data	84.8	91.1
DermoNet (JC) trained without DermoSafe data	84.3	90.2
DermoNet (JC) trained with DermoSafe data	85.3	91.5

Here, JC denotes Jaccard coefficient

#### 4.6 Results on ISBI 2017 dataset

For the experiments on the ISBI 2017 dataset, for training the models, we used either only the training dataset provided by the ISBI 2017 challenge or the augmented version of it, in which we include 6500 dermoscopic images from DermoSafe [23] to the ISBI 2017 training dataset. These trained models are then evaluated on the ISBI 2017 test dataset. Table 6 compares the performance of DermoNet with the state-of-the-art algorithms on ISBI 2017 dataset. Many teams evaluated their segmentation algorithms during the ISBI 2017 challenge. Among them, the top two teams used different variations of a fully convolutional network in their segmentation methods. For example, Yuan et al. [18] proposed a method based on deep fully convolutional-deconvolutional neural networks (CDNN) to segment skin lesions in dermoscopic image. NLP LOGIX [33] used a U-Net architecture followed by a CRF as post-processing in their segmentation method. Here, we observed that the proposed DermoNet outperforms the other teams' approaches.

**Table 6** Performance comparison between the proposed segmentation and other state-of-the-art methods on ISBI 2017 challenge test set

Method	JC (%)
FCN-ensemble [18]	76.5
Modified U-Net [33]	76.2
FCN - (JC) trained with DermoSafe data	66.49
U-Net - (JC) trained with DermoSafe data	75.50
DermoNet (JC) trained without DermoSafe data	77.5
DermoNet (JC) trained with DermoSafe data	78.3

Here, JC denotes Jaccard coefficient

#### 4.7 Effect of loss function

As described in Section 3, due to imbalanced classes, cross-entropy loss function would not be suitable for the skin lesion segmentation task. Therefore, we used Jaccard distance instead, which enabled the DermoNet's training to focus more on lesion pixels over background. To also empirically analyze the effect of the loss function, we compare the performance of DermoNet using Jaccard distance or cross-entropy on ISBI 2016, 2017 and PH2 dataset. As can be seen from Table 7, using Jaccard distance as the loss function improves the performance significantly compared to using cross-entropy as the loss function.

#### 4.8 Qualitative comparison

In this section, we provide qualitative comparison between DermoNet, FCN, and U-Net. Figure 5 shows some tricky cases from ISBI 2017 challenge dataset. In this figure, from left to right, we have the original image, ground truth, the output of DermoNet, U-Net, and FCN, respectively. As can be observed, DermoNet provides better results compared to FCN and U-Net and is able to separate the skin lesion from artifacts such as ink markings and air bubbles.

Figure 6 shows cases where the ground truth is wrongly labeled, and it leads to a very low Jaccard coefficients (JC) even though the output of the segmentation is correct. In this figure, from left to right, we have the original image, ground truth, and DermoNet, U-Net, and FCN output.

Finally, Fig. 7 shows some of the challenging cases among all the ISBI 2017 testing images where all three models (DermoNet, U-Net, and FCN) performed poorly. In these cases, the contrast between lesion and skin is very low.

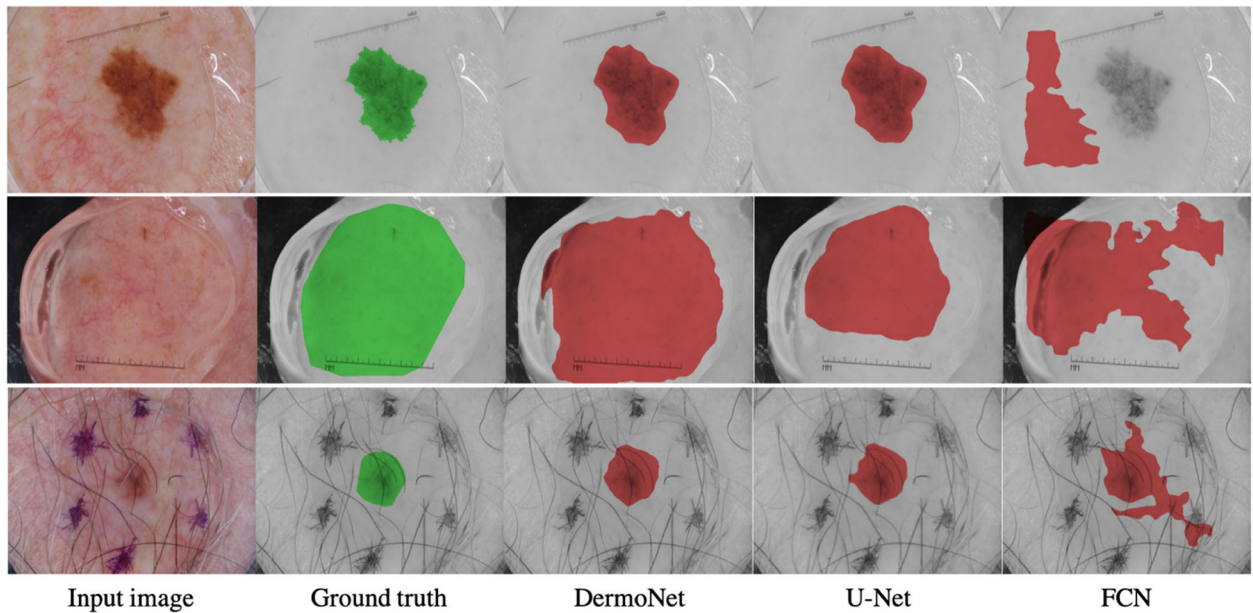
### 5 Conclusion and future work

In this paper, we have presented a new fully convolutional neural network architecture for automatic skin lesion segmentation. The idea behind DermoNet is sharing features across all encoder blocks and taking benefit of reusing features, while remaining densely connected

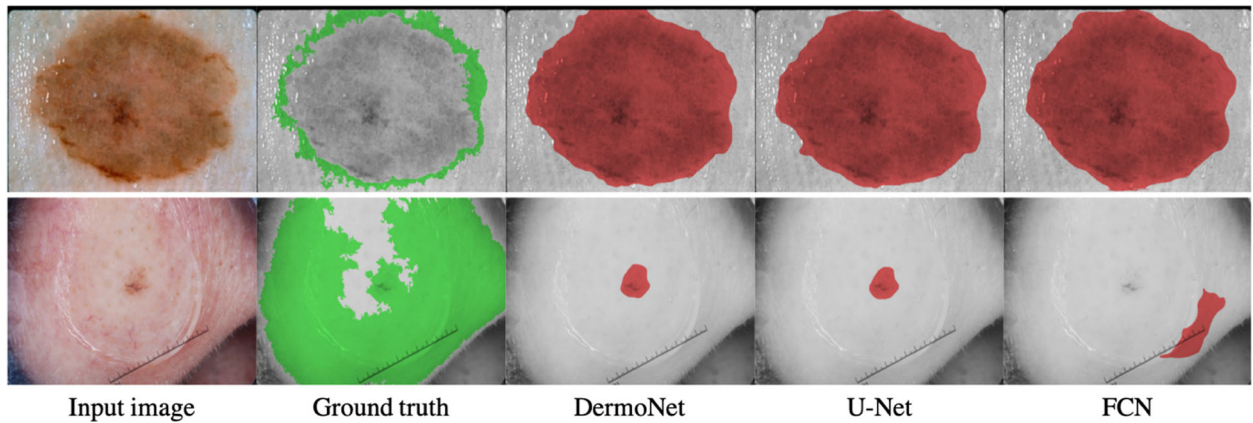
**Table 7** Performance comparison of the proposed segmentation on ISBI 2016 and 2017 and PH2 dataset when using Jaccard coefficient or cross-entropy loss for training

Dataset	Method	JC (%)
ISBI 2016	DermoNet (CE)	79.16
	DermoNet (JC)	82.5
PH2	DermoNet (CE)	72.14
	DermoNet (JC)	85.3
ISBI 2017	DermoNet (CE)	75.4
	DermoNet (JC)	78.3

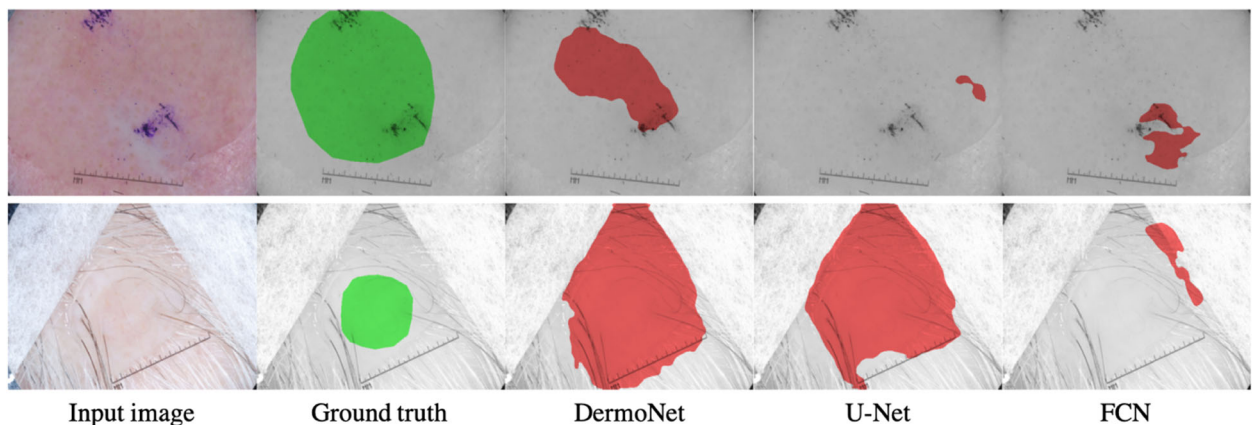
Here, JC denotes Jaccard coefficient and CE denotes cross-entropy, respectively



**Fig. 5** Cases where our proposed DermoNet outperforms FCN and U-Net



**Fig. 6** Cases where all segmentation methods failed (mis-segmentation)



**Fig. 7** Cases where the results of all the networks are suboptimal



to provide the network with more flexibility in learning new features. The proposed network has fewer parameters compared to existing baseline segmentation methods that have an order of magnitude larger memory requirement. Moreover, it improves state-of-the-art performance on challenging skin datasets, without using neither additional post-processing nor pre-training. We have achieved an average Jaccard coefficient of 82.5% on the ISBI 2016 Skin Lesion Challenge dataset, 85.3% on the PH2 dataset, and 78.3% on ISBI 2017 Skin Lesion Challenge dataset. In our future work, we plan to apply the proposed segmentation with some modifications in the network architecture on standard semantic segmentation benchmarks, e.g., MSCOCO, to show the generalization capability of the proposed framework.

#### Abbreviations

CDNN: Convolutional-deconvolutional neural network; CNN: Convolutional neural networks; CRF: Conditional random fields; DSC: Dice similarity coefficient; FCN: Fully convolutional networks; FN: False negative; FP: False positive; ISBI: International symposium on biomedical imaging; JC: Jaccard coefficient; ReLU: Rectifier non-linearity; ResNet: Residual network; RNN: Recurrent neural networks; TP: True positive

#### Acknowledgments

The authors would like to thank Mr. Philippe Held, CEO and Founder of DermoSafe, for his support and for giving us access to the DermoSafe's database of images of pigmented skin lesions, which helped us to achieve the mentioned results.

#### Funding

This work was supported by the Swiss Commission for Technology and Innovation CTI fund no. 25515.2 PFLS-LS for the project entitled "DermoBrain: advanced computer vision algorithms and features for the early diagnosis of skin cancer."

#### Availability of data and materials

The ISBI 2016 [20] datasets analyzed during the current study are available in <https://challenge.kitware.com/#phase/566744dcccad3a56fac786787>. The ISBI 2017 [19] datasets analyzed during the current study are available in <https://challenge.kitware.com/#challenge/583f126bcad3a51cc66c8d9a>. The PH2 [21] datasets analyzed during the current study are available in <https://www.dropbox.com/s/k88qukc20jnbuo/PH2Dataset.rar>. The datasets of DermoSafe [23] that are analyzed during the current study are not publicly available due to the protection of patient privacy.

#### Authors' contributions

SBS and BB conceived and designed the methods. SBS performed the experiments. PSS, HKE, and JPT supervised the project.

#### Competing interests

The authors declare that they have no competing interests.

#### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

#### Author details

<sup>1</sup>Electrical Engineering Department, Signal Processing Laboratory (LTS5), École Polytechnique Fédérale de Lausanne (EPFL), Station 11, 1015 Lausanne, Switzerland. <sup>2</sup>DermoSafe SA, EPFL Innovation Park, Bâtiment D, 1015 Lausanne, Switzerland. <sup>3</sup>Department of Computer Engineering, 34469 Maslak, Istanbul, Turkey.

Received: 23 November 2018 Accepted: 15 May 2019

Published online: 18 July 2019

#### References

1. J. Redmon, S. K. Divvala, R. B. Girshick, A. Farhadi, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. You only look once: unified, real-time object detection, (Las Vegas, NV, 2016), pp. 779–788
2. K. He, X. Zhang, S. Ren, J. Sun, in *Proceedings of the IEEE conference on CVPR*. Deep residual learning for image recognition, (2016), pp. 770–778
3. J. Long, E. Shelhamer, T. Darrell, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Fully convolutional networks for semantic segmentation, (Boston, MA, 2015), pp. 3431–3440
4. V. Badrinarayanan, A. Kendall, R. Cipolla, SegNet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans Pattern Anal Mach Intell.* **39**(12), 2481–2495 (2017)
5. H. Noh, S. Hong, B. Han, in *Proceedings of the IEEE ICCV*. Learning deconvolution network for semantic segmentation, (2015), pp. 1520–1528
6. O. Ronneberger, P. Fischer, T. Brox, in *Proc. Med. Image Comput. Comput.-Assisted Intervention*. U-net: convolutional networks for biomedical image segmentation (Springer, 2015), pp. 234–241
7. K. Simonyan, A. Zisserman, in *ICLR*. Very deep convolutional networks for large-scale image recognition, (2015)
8. G. Huang, Z. Liu, K. Q. Weinberger, in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Densely connected convolutional networks, (Honolulu, 2017), pp. 2261–2269
9. S. Ren, K. He, R. Girshick, J. Sun, in *Advances in neural information processing systems*. Faster R-CNN: towards real-time object detection with region proposal networks, (2015), pp. 91–99
10. C. Yan, H. Xie, J. Chen, Z. Zha, X. Hao, Y. Zhang, Q. Dai, A fast ughur text detector for complex background images. *IEEE Trans. Multimed.* **20**(12), 3389–3398 (2018)
11. C. Farabet, C. Couprie, L. Najman, Y. LeCun, Learning hierarchical features for scene labeling. *IEEE Trans. Pattern. Anal. Mach. Intell.* **35**(8), 1915–1929 (2013)
12. B. Bozorgtabar, Z. Ge, R. Chakravorty, M. Abedini, S. Demyanov, R. Garnavi, in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. Investigating deep side layers for skin lesion segmentation, (Melbourne, 2017), pp. 256–260
13. S. Jégou, M. Drozdal, D. Vazquez, A. Romero, Y. Bengio, in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. The one hundred layers tiramisu: fully convolutional densenets for semantic segmentation, (Honolulu, 2017), pp. 1175–1183
14. G. J. Brostow, J. Fauqueur, R. Cipolla, Semantic object classes in video: a high-definition ground truth database. *Pattern Recogn. Lett.* **30**(2), 88–97 (2009)
15. L. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A. L. Yuille, DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Trans Pattern Anal Mach Intell.* **40**(4), 834–848 (2016)
16. S. Zheng, S. Jayasumana, B. Romera-Paredes, V. Vineet, Z. Su, D. Du, C. Huang, P. H. S. Torr, in *Proceedings of the IEEE ICCV*. Conditional random fields as recurrent neural networks, (2015), pp. 1529–1537
17. K. Maninis, J. Pont-Tuset, P. A. Arbeláez, L. J. V. Gool, in *International Conference on MICCAI*. Deep retinal image understanding (Springer, 2016), pp. 140–148
18. Y. Yuan, M. Chao, Y. Lo, in *International Skin Imaging Collaboration (ISIC) 2017 Challenge at the International Symposium on Biomedical Imaging (ISBI)*. Automatic skin lesion segmentation with fully convolutional-deconvolutional networks, (2017). <https://arxiv.org/pdf/1703.05165.pdf>
19. N. C. F. Codella, D. Gutman, M. E. Celebi, B. Helba, M. A. Marchetti, S. W. Dusza, A. Kalloo, K. Liopyris, N. K. Mishra, H. Kittler, A. Halpern, in *IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. 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), (Washington, 2017), pp. 168–172
20. D. Gutman, N. C. F. Codella, M. E. Celebi, B. Helba, M. A. Marchetti, N. K. Mishra, A. Halpern, Skin lesion analysis toward melanoma detection: a challenge at the ISBI 2016, hosted by the ISIC (2016). arXiv preprint arXiv:1605.01397
21. T. Mendonça, P. M. Ferreira, J. S. Marques, A. R. S. Marçal, J. Rozeira, in *EMBC, 2013 35th Annual International Conference of the IEEE*. Ph 2-a

- dermoscopic image database for research and benchmarking (IEEE, 2013), pp. 5437–5440
22. M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, TensorFlow: Large-scale machine learning on heterogeneous systems (2015). software available from tensorflow.org. [Online]. Available: <https://www.tensorflow.org/>
  23. DermoSafe. <https://www.dermosafe.com/en/>
  24. S. Xie, Z. Tu, in *IEEE International Conference on Computer Vision (ICCV)*. Holistically-nested edge detection, (Santiago, 2015), pp. 1395–1403
  25. M. Zortea, S. O. Skråvseth, T. R. Schopf, H. M. Kirchesch, F. Godtliessen, Automatic segmentation of dermoscopic images by iterative classification. *J. Biomed. Imaging*. **2011**, 3 (2011)
  26. N. Tong, H. Lu, X. Ruan, M. Yang, in *Proceedings of the IEEE Conference on CVPR*. Salient object detection via bootstrap learning, (2015), pp. 1884–1892
  27. X. Li, Y. Li, C. Shen, A. Dick, A. V. D. Hengel, in *Proceedings of the IEEE ICCV*. Contextual hypergraph modeling for salient object detection, (2013), pp. 3328–3335
  28. B. Bozorgtabar, M. Abedini, R. Garnavi, in *Proc. Int. Workshop Mach. Learn. Med. Imag.* Sparse coding based skin lesion segmentation using dynamic rule-based refinement (Springer, 2016), pp. 254–261
  29. M. Silveira, J. C. Nascimento, J. S. Marques, A. R. S. Marcal, T. Mendonca, S. Yamauchi, J. Maeda, J. Rozeira, Comparison of segmentation methods for melanoma diagnosis in dermoscopy images. *IEEE J. Sel. Top. Sig. Process.* **3**(1), 35–45 (2009)
  30. M. Sezgin, B. Sankur, Survey over image thresholding techniques and quantitative performance evaluation. *J. Electron. Imaging*. **13**(1), 146–168 (2004)
  31. C. Li, C. Kao, J. C. Gore, Z. Ding, Minimization of region-scalable fitting energy for image segmentation. *IEEE Trans. Image Process.* **17**(10), 1940–1949 (2008)
  32. M. Celebi, H. Kingravi, H. Iyatomi, Y. Aslandogan, W. Stoecker, R. Moss, J. Malter, J. Grichnik, A. Marghoob, H. Rabinovitz, S. Menzies, Border detection in dermoscopy images using statistical region merging. *Skin Res. Technol.* **14**(3), 347–353 (2008)
  33. M. Berseth, ISIC 2017-skin lesion analysis towards melanoma detection (2017). <https://arxiv.org/abs/1703.00523>

Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

---