This is a pre print version of the following article:

Skin Lesion Segmentation Ensemble with Diverse Training Strategies / Canalini, Laura; Pollastri, Federico; Bolelli, Federico; Cancilla, Michele; Allegretti, Stefano; Grana, Costantino. - 11678:(2019), pp. 89-101. (Intervento presentato al convegno International Conference on Computer Analysis of Images and Patterns tenutosi a Salerno, Italy nel Sep 3-5) [10.1007/978-3-030-29888-3_8].

Springer, Cham
Terms of use:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

12/10/2024 11:23

(Article begins on next page)

Skin Lesion Segmentation Ensemble with Diverse Training Strategies

Laura Canalini, Federico Pollastri, Federico Bolelli, Michele Cancilla, Stefano Allegretti, and Costantino Grana

Dipartimento di Ingegneria "Enzo Ferrari"
Università degli Studi di Modena e Reggio Emilia
Via Vivarelli 10, Modena MO 41125, Italy
{name.surname}@unimore.it

Abstract. This paper presents a novel strategy to perform skin lesion segmentation from dermoscopic images. We design an effective segmentation pipeline, and explore several pre-training methods to initialize the features extractor, highlighting how different procedures lead the Convolutional Neural Network (CNN) to focus on different features. An encoder-decoder segmentation CNN is employed to take advantage of each pre-trained features extractor. Experimental results reveal how multiple initialization strategies can be exploited, by means of an ensemble method, to obtain state-of-the-art skin lesion segmentation accuracy.

Keywords: Deep Learning · Convolutional Neural Networks · Transfer Learning · Skin Lesion Segmentation

1 Introduction

Since 2016, the International Skin Imaging Collaboration (ISIC) has been gathering public datasets and hosting multiple challenges and workshops, stressing the relevance of skin lesion analysis [9]. Many efforts have been given in order to aid professional dermatologists in the detection of malignant melanoma which, being the most dangerous type of skin cancer, holds a substantial death rate.

Skin lesion segmentation is a fundamental step in the automated melanoma detection process, defined as the recognition of the set of pixels that constitute the skin lesion within the image. This task can be especially troublesome due to the vast variety of skin characteristics among different people, and the subjectivity of the definition of skin lesion borders. State-of-the-art approaches on this field have proved once again the effectiveness of deep learning algorithms, as a matter of fact, Convolutional Neural Networks (CNNs) are currently the cornerstone of medical images analysis. Segmentation CNNs are able to extract features through a contracting path and exploit them to generate a segmentation mask across the expanding path; the size of feature maps decreases progressively in the former, whereas in the latter it increases back to the input resolution thanks to up-sampling operators and fractionally-strided convolutions, thus producing an encoder-decoder architecture. U-Net [29] is a noteworthy example of such kind

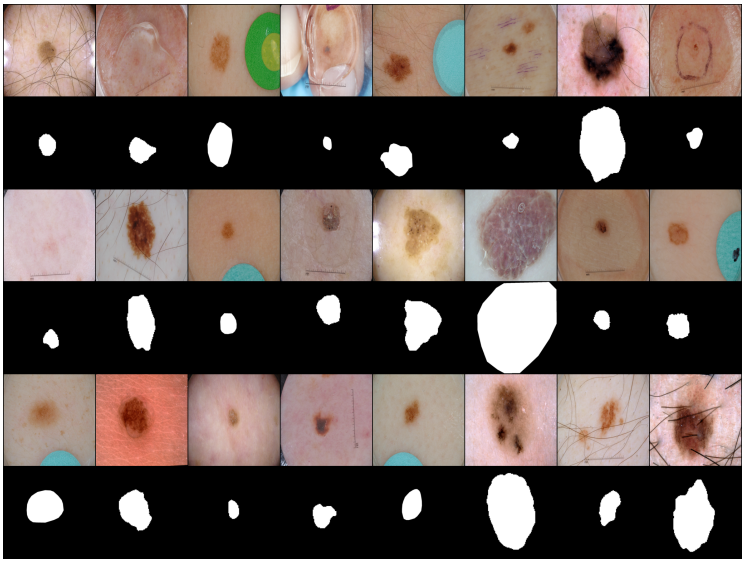


Fig. 1. Samples from the ISIC dataset: dermoscopic skin images coupled with their ground truth segmentation masks.

of neural networks, it is characterized by an equal number of layers in the two distinct paths and by skip connections, which have the purpose of concatenating features between the two distinct sections.

Unfortunately, in order to be trained, deep learning algorithms require huge amounts of data, which are often hard to obtain and particularly expensive to annotate, especially in medical fields. The need for a large annotated medical imaging dataset can be mitigated by pre-training neural networks with an already existing collection of natural images, like ImageNet [10]. Indeed, the first convolutional layers of every CNN learn to recognize simple elements like lines and colors, making them useful across different tasks and datasets [23]. Learning features from wider datasets and re-using them in different tasks can be seen as a form of transfer learning. The choice of the dataset employed to learn low-level features can introduce biases towards certain features and characteristics, and this should be taken into account especially when dealing with medical imaging. For example, CNNs trained using ImageNet are strongly biased in recognizing textures rather than shapes [13].

In this work, we address the problem of skin lesion image segmentation taking into account the trade-off between training a model from scratch and employing images of a very different nature through transfer learning. In particular, we design a CNN-based ensemble model that, exploiting different pre-training strategies, obtains the state-of-the-art performance on skin lesion segmentation when trained on the ISIC skin lesions dataset and on synthetically generated samples.

The rest of the paper is organized as follows. In Section 2 relevant proposals on image segmentation are summed up. Section 3 describes the pre-training strategies adopted to push the effectiveness of the proposed ensemble strategy, which is then described in Section 4. In Section 5 the performance of our proposal are compared with state-of-the-art model, highlighting its validity. Finally, Section 6 draws some conclusions.

2 Related Work

In the last decades, a lot of effort has been devoted to solving the problem of automated skin lesion segmentation, promoting the development of different approaches: histogram thresholding [26, 39], edge-based [8, 30], region-merging [12, 31], and supervised learning [11]. Nevertheless, CNNs are now one of the most powerful tools in machine learning and computer vision. They are able to learn features directly from input image data, with no need of hand-crafted features. In a typical classification task a convolutional network produces a single label as output. However, in other visual tasks such as segmentation [17, 36], a pixel-wise output information is required, having a critical point in the mapping of low resolution features into input resolution.

In 2015, Ronneberger *et al.*proposed the so called U-Net [29], which consists of a contracting encoder network that extracts high resolution features from the input image and followed by an expanding decoder model to produce a full-resolution segmentation. This model employs skip connections between the downsampling and upsampling path, applying a concatenation operator in order to provide coarse contextual information when upsampling.

In DeepLabv1 [6], atrous convolutions were introduced. Instead of repeatedly reducing the dimension of feature maps, each convolutional filter of the network is upsampled in size and filled with zeroes in between filter values. This allows to avoid the use of pooling operations, which are responsible for the spatial resolution reduction, and thus cause a lower output segmentation quality. DeepLabv3+ [7] is a state-of-the-art segmentation CNN. It employs atrous convolutions, residual blocks [16], and a decoder module to refine the segmentation results.

In 2017, Mask R-CNN [15] was able to surpass other instance segmentation methods. This approach combines object detection, which aims to classify and localize individual objects using a bounding box, and semantic segmentation, which has the goal of classifying each pixel given a fixed set of classes. This strategy extends Faster R-CNN [14], adding a Fully Convolutional Network (FCN) to predict segmentation masks for each Region of Interest (RoI) in a pixel-to-pixel fashion.

3 Learning to Extract Features

In order to design an ensemble of CNNs that can effectively segment skin lesion images, we select DeepLabv3+ [7] with a ResNet-101 [16] backbone as the

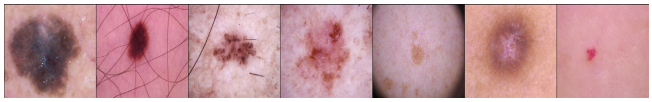


Fig. 2. One sample of each class from the third ISIC challenge task. From left to right: Melanoma, Melanocytic nevus, Basal cell carcinoma, Actinic keratosis, Benign keratosis, Dermatofibroma, and Vascular lesion.

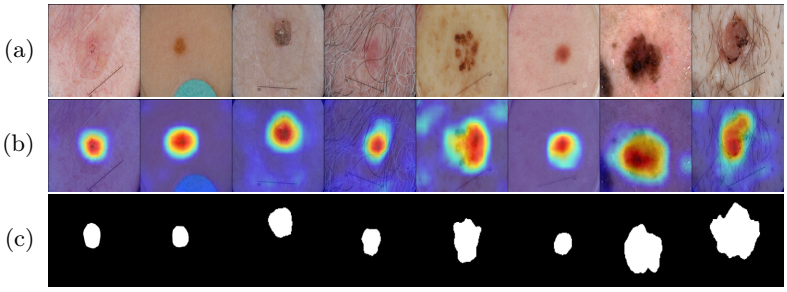


Fig. 3. Samples from the 2017 ISIC validation set (a), with corresponding Grad-Cam heatmaps (b) and segmentation masks (c).

baseline architecture, and explore several pre-training strategies. This Section discloses three different pre-training methods employed to obtain three different initializations of ResNet-101, the features extractor for our segmentation neural network.

3.1 ISIC Task 3

The International Skin Imaging Collaboration, in 2018, hosted a challenge split in three tasks: lesion segmentation, lesion attribute detection and disease classification. A public dataset of 10 015 dermoscopic images was divided by experts into 7 different classes, illustrated in Fig. 2. In the interests of obtaining a features extractor effective on dermoscopic images, we train ResNet-101 to correctly classify skin lesion images in the 7 original classes, taking advantage of the 10 015 annotated samples from the ISIC dataset [34]. The network is trained for 15 epochs, using a weighted Cross-Entropy Loss, an initial learning rate of 0.0001, and the Adam optimizer [20]. Fig 3 provides a useful visualization of which sections of the input image mostly affect the hidden representation that the network obtains after the last convolutional layer, right before employing the fully-connected layers designed for classification. The heatmaps obtained by applying the Grad-Cam technique [32], suggest that our features extractor implicitly learns to coarsely detect the skin lesion.

3.2 Autoencoder

Data annotation for the image segmentation task is extremely expensive, it requires an expert dermatologist to perform an extremely time consuming and

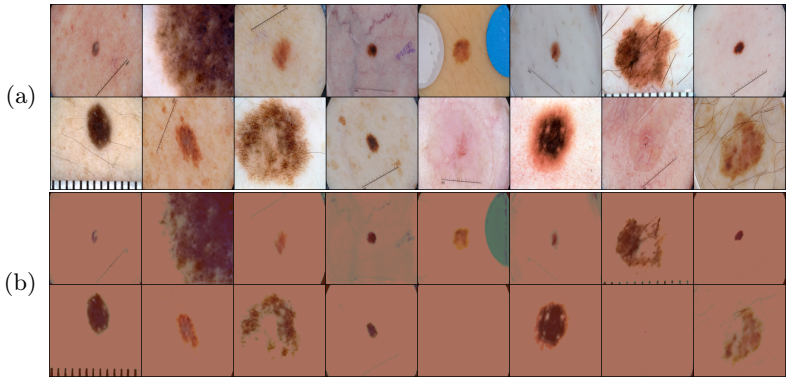


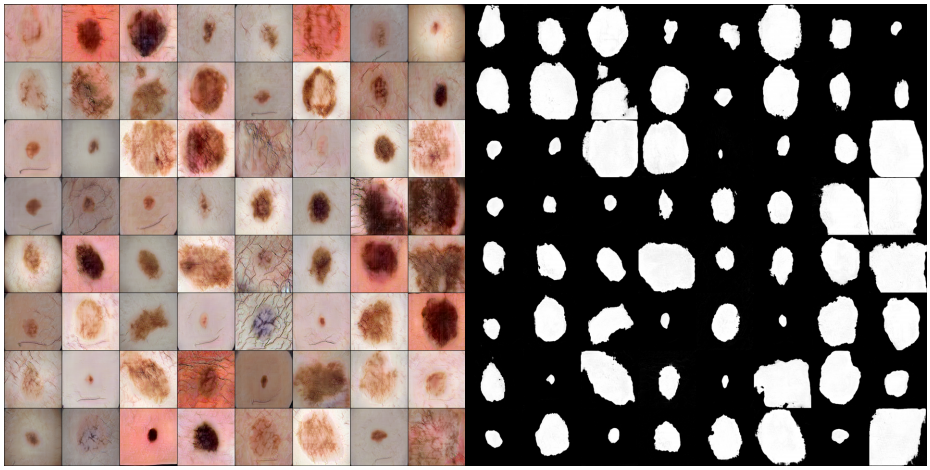
Fig. 4. Samples from the 2017 ISIC validation set (a) and the corresponding autoencoder output (b).

uninspiring job. Therefore, we try to exploit many dermoscopic images that were never given manual annotation through a form of unsupervised learning. Taking advantage of the encoder-decoder architecture of DeepLabv3+, an autoencoder is built in order to obtain a version of ResNet able to map skin lesion images to a meaningful hidden representation. The model of our autoencoder is identical to DeepLabv3+ [7], with 3 output classes (RGB channels). It is trained using, again, the 10 015 images from the ISIC classification task for 10 epochs, employing the mean squared error loss. The learning rate is initially set to 0.0001 and then influenced by the Adam optimizer [20]. Fig. 4 shows that the autoencoder struggles to recreate the original background skin color and texture, but promptly generates a coherent lesion in the correct position, which is the most important element for our final goal.

3.3 GAN

Generative Adversarial Networks (GANs) are often used to create unlabeled examples, which cannot be directly employed for the training of a supervised algorithm [38]. Following the approach introduced in [27, 28], we improve the role of GANs in the training process by designing an architecture able to generate both the skin lesion image and its segmentation mask, making it extremely easy to exploit new synthetic images as additional training data. We modify the GAN proposed by Karras *et al.* [19] in order to feed it 4-channels images: the first three channels are the R, G and B components and the fourth one is the binary segmentation mask. Instead of generating a fixed number of samples, the generator is required to provide new couples image-mask for each training batch. An example of the generated images is shown in Fig. 5.

The GAN is trained to generate 256×256 images and segmentation masks. Most images present realistic details, like a well delivered presence of hair, black corners representing real camera characteristics, and pen marks. Moreover, the



(a) (b)

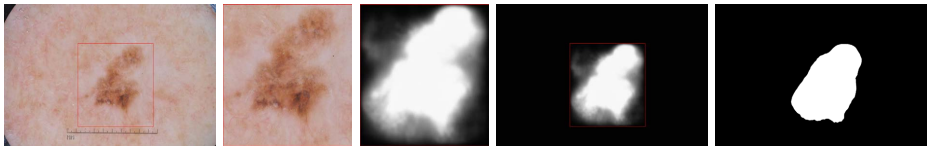
Fig. 5. GAN-generated skin lesion samples (a) and their segmentation masks (b).

generative model produces segmentation masks that look very coherent with the respective generated images.

4 Learning to Segment Skin Lesions

To take advantage of the pre-trained backbones, DeepLabv3+ is fine-tuned following the training protocol outlined in [7]. Input images are resized to 513×513 pixels, the learning rate is initially set to 0.007 and then multiplied by $(1 - \frac{iter}{max_iter})^{0.9}$, following the poly policy [25]. The output stride is fixed to 16, the network is trained for 60 epochs, and the early stopping strategy is employed. During the whole fine-tuning process, the learning rate is multiplied by a factor of 10 for layers outside of ResNet-101, in order to take full advantage of the pre-training process applied to the features extractor. Data augmentation is performed by randomly rotating, flipping, shifting, shearing and scaling input images, and by changing the color contrasts.

To enhance the fine-tuning process, the official 2017 training set of 2 000 images is enlarged using several dermoscopic images gathered from the public ISIC archive. In order to remove potentially disadvantageous samples, every supplementary image is first fed to a state-of-the art segmentation CNN. Images segmented with an Intersection over Union lower than 0.60 are considered incorrect and thus removed from further analysis. To build the final training set, we select the 1 500 dermoscopic images that offer a ground truth segmentation mask where the skin lesion fills at least the 10% of the image. We thus obtain a 3 500 images training set.



(a) (b) (c) (d) (e)

Fig. 6. Detection-Segmentation Pipeline: sample image from the 2017 ISIC test set with its predicted bounding box overlapped (a), the cropped and resized sample (b), the output of the ensemble method (c) that is rescaled to the original size (d) and the final segmentation mask (e).

4.1 Detection-Segmentation Pipeline

In order to encourage the network to focus on the correct section of input images, we remove redundant background from the input samples. During training, every ground truth segmentation mask is exploited to obtain the bounding box containing the skin lesion, and every dermoscopic image is cropped accordingly (Fig. 6a and Fig. 6b). During inference, input samples are cropped in accordance with bounding boxes obtained through Mask R-CNN [15], pre-trained on the COCO dataset [24] and then fine-tuned on the 2017 ISIC original training set. Mask R-CNN produces bounding boxes with an Intersection over Union of 0.886 on the 2017 public test set. This pipeline, illustrated in Fig. 6, improves the Intersection over Union accuracy of the segmentation framework by 0.045. However, cropping images by means of an object detection network can erase meaningful portions of the segmentation CNN input, propagating the error from the first network to the next one. To tackle this drawback, bounding boxes can be directly supplied to the segmentation CNN, without forcing it to discard the sections of the input image classified as background by the detection tool. We thus feed the segmentation network with a 4-channels input: R, G, B, and the skin lesion bounding box in the form of a binary mask, where white pixels represent the foreground and black pixels represent the background. Improvements delivered by adding an object detection network to the process are shown in Table 1.

Pipeline Strategy	**Validation IoU**
None	0.800
Crop	0.845
4-channels	0.848

Table 1. Detection-Segmentation Pipelines Comparison.

4.2 The Ensemble Method

As observed by Krogh *et al.* [21] and Chandra *et al.* [5], a key point on ensemble learning is that the hypotheses should be as accurate and as diverse as possible. In our work, such properties were prompted by applying different strategies. Even though the performance of the various CNNs are comparable, different networks focus on different features, greatly increasing the effectiveness of the ensemble. Fig. 7 displays how the features extractor pre-trained to build the

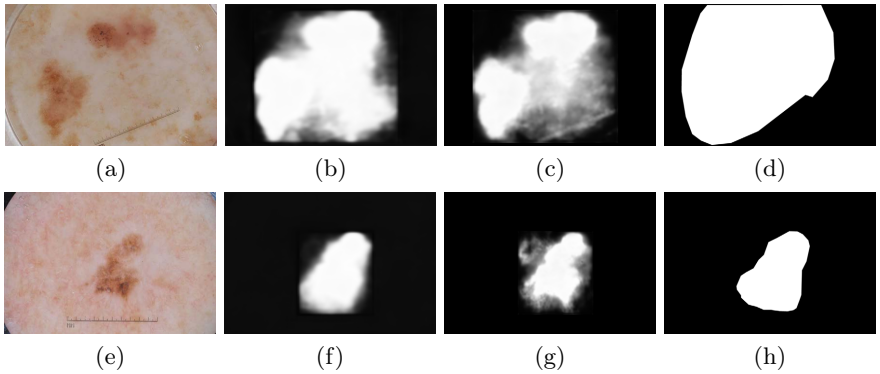


Fig. 7. Examples of model outputs when using different pre-trained feature extractors. (a) and (e) are the original images, (b) and (f) are the outputs when using a features extractor pre-trained with ImageNet, (c) and (g) are the outputs of the CNNs that exploit the features extractor described in Section 3.2. Finally, (d) and (h) represent the ground truth.

autoencoder described in Section 3.2, seems to give much importance to darker shades of colours, whereas the CNN that employs the features extractor formerly used to classify natural images from ImageNet, produces segmentation masks less attentive to this particular feature, focusing on higher level structures instead.

For each pixel, its probability of being part of the skin lesion is obtained as the mean value across the selected CNNs. The output is then binarized with a dual-threshold method. A high threshold (0.80) is followed by blob analysis [3, 4] and the biggest object center is assumed to be the tumor center. Afterwards, a lower threshold (0.40) is applied and the final segmentation mask is given by the region which contains the tumor center. Whenever the first high threshold does not yield any object, we only apply the second one and keep its result as segmentation mask.

5 Experimental Results

Experimental results of the proposed networks are summed up in Table 2. The first column shows the detection-segmentation pipeline of each CNN. It is important to notice that, when using the 4-channels strategy described in Section 4.1, the first layer of the features extractor is required to deal with four channels instead of three, and thus to be re-trained from scratch. Since every network described in Section 3 was pre-trained using dermoscopic images, we fine-tune them using the cropping strategy to avoid a random re-initialization of the first layer of filters.

In order to further increase the diversity between predicted masks, and hence the effectiveness of the ensemble architecture, we employ two different loss func-

Pipeline Strategy	Loss	Pre-Training	Validation IoU	Validation TIoU	Test IoU	Test TIoU
Crop	Cross-Entropy	Classes	0.851	0.819	0.841	0.809
	Cross-Entropy	AE	0.856	0.837	0.838	0.810
	Cross-Entropy	GAN	0.854	0.834	0.838	0.814
	Tanimoto	Classes	0.855	0.832	**0.845**	0.818
	Tanimoto	AE	0.857	0.838	0.844	0.816
	Tanimoto	GAN	0.850	0.831	0.841	0.817
4-channels	Cross-Entropy	None	0.847	0.816	0.834	0.805
	Cross-Entropy	ImageNet	0.848	0.823	0.837	0.808
	Tanimoto	None	0.846	0.819	0.831	0.802
	Tanimoto	ImageNet	**0.859**	**0.840**	**0.845**	**0.819**

Table 2. Analysis of the Neural Networks trained for the task. IoU is the Intersection over Union, TIoU is the Threshold Intersection over Union.

tions during training (second column of Table 2): the cross-entropy loss[1] in its binary form and the Tanimoto distance defined as

$$L = 1 - \frac{\sum_{i,j} t_{ij} p_{ij}}{\sum_{i,j} t_{ij}^2 + \sum_{i,j} p_{ij}^2 - \sum_{i,j} t_{ij} p_{ij}} \tag{1}$$

where t_{ij} is the target value of the pixel at coordinates (i, j), and p_{ij} is the real output. Note that t_{ij} is either 0 or 1, while p_{ij} is a real number in range $[0, 1]$.

The third column of Table 2 presents the pre-training procedure applied to each network. Values *Classes*, *AE*, and *GAN* refer to Sections 3.1, 3.2, and 3.3 respectively. The value *None* represents a network trained from scratch, included for comparison. The last four columns of the Table present the accuracy of each network, on both the validation set and the test set from the ISIC 2017 challenge. IoU stands for Intersection over Union, which is the official evaluation metric of the 2017 ISIC challenge. TIoU means Threshold Intersection over Union, an adaptation of Intersection over Union that reflects the number of images in which automated segmentation fails, by giving a score of 0 to images segmented with an Intersection over Union lower than 0.65. Both of the metrics show similar results when testing different initialization methods, but stress the importance of employing a pre-trained feature extractor: CNNs trained from scratch (value *None* in the third column) show the worst results. Table 3 displays the results obtained by merging the output of resulting CNNs through the ensemble strategy described in Section 4.2. The proposed pipeline clearly outperforms state-of-the-art segmentation algorithms when applied on skin lesions.

[1] Cross-Entropy is the standard loss function employed when training DeepLab [6].

Method	Test IoU	Test TIoU
Ours (ensemble)	**0.850**	0.827
SegAN [35]	0.785	—
GAN Augmented [27]	0.781	—
DCL-PSI [2]	0.777	—
DeepLabv3+* [7]	0.769	—
(RE)-DS-U-ResnetFCN34 [22]	0.772	—
SegNet* [1]	0.767	—
Challenge winners [37]	0.765	—
Tiramisu* [18]	0.765	—
U-Net* [29]	0.740	—

Table 3. Performance of the proposed method compared to the state-of-the-art. All the models have been trained and tested on the 2017 ISIC dataset. Reimplemented/retrained models are identified by *. IoU is the Intersection over Union, TIoU is the Threshold Intersection over Union.

6 Conclusion

With this paper, we tackled the problem of skin lesion segmentation by a different perspective, introducing a novel ensemble strategy to improve state-of-the-art results. We presented multiple ways to initialize a features extractor without the need to employ biases-inducing datasets. The designed segmentation pipeline takes advantage of the multiple pre-training methods, improving the overall performance. Experimental results on the ISIC 2017 dataset show the effectiveness of our approach on skin lesion segmentation. The source code of the proposed model is available in [33].

Acknowledgments This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825111, DeepHealth Project.

References

1. Badrinarayanan, V., Kendall, A., Cipolla, R.: SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. IEEE transactions on pattern analysis and machine intelligence **39**(12), 2481–2495 (2017)
2. Bi, L., Kim, J., Ahn, E., Kumar, A., Feng, D., Fulham, M.: Step-wise integration of deep class-specific learning for dermoscopic image segmentation. Pattern Recognition **85**, 78–89 (2019)
3. Bolelli, F., Baraldi, L., Cancilla, M., Grana, C.: Connected Components Labeling on DRAGs. In: International Conference on Pattern Recognition (2018)
4. Bolelli, F., Cancilla, M., Grana, C.: Two More Strategies to Speed Up Connected Components Labeling Algorithms. In: International Conference on Image Analysis and Processing. pp. 48–58. Springer (2017)

5. Chandra, A., Yao, X.: Evolving Hybrid Ensembles of Learning Machines for Better Generalisation. Neurocomputing **69**(7-9), 686–700 (2006)
6. Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L.: Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs. arXiv preprint arXiv:1412.7062 (2014)
7. Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H.: Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. In: Proceedings of the European Conference on Computer Vision (ECCV). pp. 801–818 (2018)
8. Chung, D.H., Sapiro, G.: Segmenting Skin Lesions with Partial-Differential-Equations-Based Image Processing Algorithms. IEEE transactions on Medical Imaging **19**(7), 763–767 (2000)
9. Codella, N., Gutman, D., Celebi, 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). In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). pp. 168–172 (2018)
10. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: ImageNet: A Large-Scale Hierarchical Image Database. In: 2009 IEEE conference on computer vision and pattern recognition. pp. 248–255. Ieee (2009)
11. Ganster, H., Pinz, P., Rohrer, R., et al.: Automated Melanoma Recognition. IEEE transactions on medical imaging **20**(3), 233–239 (2001)
12. Gao, J., Zhang, J., Fleming, M.G.: A novel multiresolution color image segmentation technique and its application to dermatoscopic image segmentation. In: Proceedings 2000 International Conference on Image Processing. vol. 3, pp. 408–411. IEEE (2000)
13. Geirhos, R., Rubisch, P., Michaelis, C., Bethge, M., Wichmann, F.A., Brendel, W.: ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness. arXiv preprint arXiv:1811.12231 (2018)
14. Girshick, R.: Fast R-CNN. In: Proceedings of the IEEE international conference on computer vision. pp. 1440–1448 (2015)
15. He, K., Gkioxari, G., Dollár, P., Girshick, R.: Mask R-CNN. In: Proceedings of the IEEE international conference on computer vision. pp. 2961–2969 (2017)
16. He, K., Zhang, X., Ren, S., Sun, J.: Deep Residual Learning for Image Recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
17. Jafari, M.H., Nasr-Esfahani, E., Karimi, N., Soroushmehr, S.M.R., Samavi, S., Najarian, K.: Extraction of skin lesions from non-dermoscopic images for surgical excision of melanoma. International Journal of Computer Assisted Radiology and Surgery **12**(6), 1021–1030 (2017)
18. Jégou, S., Drozdzal, M., Vazquez, D., Romero, A., Bengio, Y.: The One Hundred Layers Tiramisu: Fully Convolutional DenseNets for Semantic Segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 11–19 (2017)
19. Karras, T., Aila, T., Laine, S., Lehtinen, J.: Progressive Growing of GANs for Improved Quality, Stability, and Variation. In: International Conference on Learning Representations (2018)
20. Kingma, D.P., Ba, J.: Adam: A Method for Stochastic Optimization. arXiv preprint arXiv:1412.6980 (2014)
21. Krogh, A., Vedelsby, J.: Neural Network Ensembles, Cross Validation, and Active Learning. In: Advances in neural information processing systems. pp. 231–238 (1995)

22. Li, X., Yu, L., Fu, C.W., Heng, P.A.: Deeply Supervised Rotation Equivariant Network for Lesion Segmentation in Dermoscopy Images. In: OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis, pp. 235–243. Springer (2018)
23. Li, Y., Yosinski, J., Clune, J., Lipson, H., Hopcroft, J.E.: Convergent Learning: Do different neural networks learn the same representations? In: FE@ NIPS. pp. 196–212 (2015)
24. Lin, T.Y., Maire, M., Belongie, S., et al.: Microsoft COCO: Common Objects in Context. In: European conference on computer vision. pp. 740–755. Springer (2014)
25. Liu, W., Rabinovich, A., Berg, A.C.: ParseNet: Looking Wider to See Better. arXiv preprint arXiv:1506.04579 (2015)
26. Pellacani, G., Grana, C., Seidenari, S.: Algorithmic reproduction of asymmetry and border cut-off parameters according to the abcd rule for dermoscopy. Journal of the European Academy of Dermatology and Venereology **20**(10), 1214–1219 (2006)
27. Pollastri, F., Bolelli, F., Grana, C.: Improving Skin Lesion Segmentation with Generative Adversarial Networks. In: 31st International Symposium on Computer-Based Medical Systems (2018)
28. Pollastri, F., Bolelli, F., Paredes, R., Grana, C.: Augmenting Data with GANs to Segment Melanoma Skin Lesions. In: Multimedia Tools and Applications Journal. MTAP, Springer (2019)
29. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional Networks for Biomedical Image Segmentation. In: International Conference on Medical image computing and computer-assisted intervention. pp. 234–241. Springer (2015)
30. Rubegni, P., Ferrari, A., Cevenini, G., et al.: Differentiation between pigmented Spitz naevus and melanoma by digital dermoscopy and stepwise logistic discriminant analysis. Melanoma research **11**(1), 37–44 (2001)
31. Schmid, P.: Lesion Detection in Dermatoscopic Images Using Anisotropic Diffusion and Morphological Flooding. In: Proceedings 1999 International Conference on Image Processing. vol. 3, pp. 449–453. IEEE (1999)
32. Selvaraju, R.R., Cogswell, M., Das, A., et al.: Grad-CAM: Visual Explanations From Deep Networks via Gradient-Based Localization. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 618–626 (2017)
33. Source Code of the Proposed Model. `https://github.com/PollastriFederico/skin_lesion_segmentation_ensemble`, accessed on 2019-06-21
34. Tschandl, P., Rosendahl, C., Kittler, H.: The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Scientific data **5** (2018)
35. Xue, Y., Xu, T., Huang, X.: Adversarial learning with multi-scale loss for skin lesion segmentation. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). pp. 859–863. IEEE (2018)
36. 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 **36**(4), 994–1004 (2017)
37. Yuan, Y., Chao, M., Lo, Y.C.: Automatic skin lesion segmentation with fully convolutional-deconvolutional networks. arXiv preprint arXiv:1703.05165 (2017)
38. Zheng, Z., Zheng, L., Yang, Y.: Unlabeled Samples Generated by GAN Improve the Person Re-Identification Baseline in Vitro. arXiv preprint arXiv:1701.07717 (2017)
39. Zortea, M., Flores, E., Scharcanski, J.: A simple weighted thresholding method for the segmentation of pigmented skin lesions in macroscopic images. Pattern Recognition **64**, 92–104 (2017)