

Article

Skin Lesion Segmentation from Dermoscopic Images Using Convolutional Neural Network

Kashan Zafar ¹, Syed Omer Gilani ^{1,*}, Asim Waris ¹, Ali Ahmed ¹, Mohsin Jamil ²,
Muhammad Nasir Khan ³ and Amer Sohail Kashif ¹

¹ Department of Biomedical Engineering and Sciences, School of Mechanical and Manufacturing Engineering, National University of Sciences and Technology, Islamabad 44000, Pakistan; kashan293@gmail.com (K.Z.); asim.waris@smme.nust.edu.pk (A.W.); talpur54@outlook.com (A.A.); amer.kashif@smme.nust.edu.pk (A.S.K.)

² Department of Electrical and Computer Engineering, Memorial University of Newfoundland, Newfoundland, St. John's, NL A1C 5S7 P.O. Box 4200, Canada; mjamil@mun.ca

³ Department of Electrical Engineering, University of Lahore, Lahore 54590, Pakistan; muhammad.nasir@ee.uol.edu.pk

* Correspondence: omer@smme.nust.edu.pk; Tel.: +92-51-90856052

Received: 28 January 2020; Accepted: 9 March 2020; Published: 13 March 2020



Abstract: Clinical treatment of skin lesion is primarily dependent on timely detection and delimitation of lesion boundaries for accurate cancerous region localization. Prevalence of skin cancer is on the higher side, especially that of melanoma, which is aggressive in nature due to its high metastasis rate. Therefore, timely diagnosis is critical for its treatment before the onset of malignancy. To address this problem, medical imaging is used for the analysis and segmentation of lesion boundaries from dermoscopic images. Various methods have been used, ranging from visual inspection to the textural analysis of the images. However, accuracy of these methods is low for proper clinical treatment because of the sensitivity involved in surgical procedures or drug application. This presents an opportunity to develop an automated model with good accuracy so that it may be used in a clinical setting. This paper proposes an automated method for segmenting lesion boundaries that combines two architectures, the U-Net and the ResNet, collectively called Res-Unet. Moreover, we also used image inpainting for hair removal, which improved the segmentation results significantly. We trained our model on the ISIC 2017 dataset and validated it on the ISIC 2017 test set as well as the PH² dataset. Our proposed model attained a Jaccard Index of 0.772 on the ISIC 2017 test set and 0.854 on the PH² dataset, which are comparable results to the current available state-of-the-art techniques.

Keywords: melanoma; dermoscopic images; convolutional neural networks; U-Net; ResNet; image inpainting; Jaccard Index; ROC curve

1. Introduction

Computer-aided technologies for the diagnostic analysis of medical images have received significant attention from the research community. These are efficiently designed and modified for the purposes of inter-alia segmentation and classification of the region of interest (ROI) [1], which in this instance involves cancerous regions. Needless to mention, the effective treatment of cancer is dependent on early detection and delimitation of lesion boundaries, particularly during its nascent stages because cancer generally has the characteristic tendency of delayed clinical onset [2]. Every year, nearly 17 million people are affected by cancer and about 9.6 million people die due to delayed diagnosis and treatment [3]. This makes cancer the leading causes of death worldwide [4]. In the case of skin cancer, it is one of the most prevalent types of the disease in both adults and children [5] and occurs or

originates in the epidermal tissue. Various computer-aided techniques have been proposed for cancer boundary detection from dermoscopic images [3].

Among the different types of skin cancers, melanoma is not only the most dangerous and aggressive in nature due to its high metastasis rate, but also has the greatest prevalence [4]. Melanoma is a malignant type of skin cancer that develops through the irregular growth of pigmented skin cells called melanocytes [5]. It can develop anywhere on the epidermal layer of the skin and presumably may also affect the chest and back and propagate from the primary site of the cancer [6]. Its incidence rate has risen up to 4–6% annually and has the highest mortality rate among all types of skin cancers [4]. Early diagnosis is crucial as it increases the five-year survival rate up to 98% [7,8].

From the above details pertaining to the incidence and mortality rate associated with melanoma, timely diagnosis becomes all the more necessary for providing effective treatment to the affected. Insofar as the detection and segmentation of lesion boundaries, there are two streams of methodologies: first, traditional methods that usually resort to visual inspection by the clinician, and second, semi-automated and automated methods, which mostly involve point-based pixel intensity operations [9,10], pixel clustering methods [11–14], level set methods [15], deformable models [16], deep-learning based methods [17–19], et cetera.

Be that as it may, most of the methods being used today are not semi-automated because the accuracy associated therewith is generally prone to errors due to the following reasons: the inherent limitations of the methods [20], and changing character of dermoscopic images induced due to the fluorescence and brightness inhomogeneities [10]. For this very reason, the world has shifted toward more sophisticated methods, inter-alia, the convolutional neural networks (CNNs) [21].

In this paper, we intend to exploit the properties and model architectures based on CNN for skin lesion boundary delimitation and segmentation. In addition, we propose our own novelty within the already available techniques which greatly increases the segmentation accuracy, which is image inpainting. Image inpainting, together with other image preprocessing techniques such as morphological operations, is used to remove the hair structures contained within the dermoscopic images that otherwise handicap the architecture because of complexities present in the images.

This research examines the accuracy of the proposed technique together with the adoption of the proposed preprocessing method. We also benchmark our proposed scheme with other available methods by way of results through network accuracy, Jaccard Index, Dice score, and other performance metrics that aid us in comparison.

1.1. Literature Review

This section delineates and chalks-out the relevant work done on the issue of segmentation of skin lesions. It is done with an added emphasis and focus on the recent studies that have incorporated deep-learning methods for the aforementioned purpose of lesion segmentation.

At the outset, it is contended that accurate segmentation and delimitation of skin lesion boundaries can aid and assist the clinician in the detection and diagnosis process, and may later also help toward classification of the lesion type. There has been a gamut of studies done for the purposes of segmentation and classification of skin lesions, and for a general survey of these, the reader can refer the following two papers authored by Oliveira et al. [3], and Rafael et al. [22].

We hereinafter present a review of the literature vis-à-vis two aspects (i.e., preprocessing and segmentation techniques, respectively). Both aspects have a direct effect on the outcome of the results (the prediction) and therefore, both are catered into the broader scheme of methodology presented in this paper. Additionally, since dermoscopic images have varying complexities and contain different textural, intensity, and feature inhomogeneity, it becomes necessary to apply prior preprocessing techniques so that inhomogeneous sections can be smoothed out.

1.1.1. Preprocessing Techniques

Researchers encounter complications while segmenting skin lesions due to low brightness and the noise present in the images. These artifacts affect the accuracy of segmentation. For better results, Celebi et al. [23] proposed a technique that enhances image contrast by searching for idyllic weights for converting RGB images into grayscale by maximizing Otsu's histogram bimodality measure. Optimization resulted in a better adaptive ability to distinguish between tumor and skin and allowed for accurate resolution of the regions, whereas Beuren et al. [24] described the morphological operation that can be applied on the image for contrast enhancement. The lesion is highlighted through the color morphological filter and simply segmented through binarization. Lee et al. [25] proposed a method to remove hair artifacts from dermoscopic images. An algorithm based on morphological operations was designed to remove hair like artifacts from skin images. Removing hair, characterized as noise, from skin images has a noteworthy effect on segmentation results. A median filter was found to be effective on noisy images. A nonlinear filter was applied to images to smooth them [26]. Celebi et al. [27] established a concept where the size of the filter to be applied must be proportional to the size of the image for effective smoothing.

Image inpainting is a preprocessing technique used for both removing parts from an image and for restoration purposes, so that the missing and damaged information in images is restored. It is of vital importance in the field of medical imaging and through its application, unnecessary structures or artifacts from the images (i.e., hair artifacts in skin lesions images) can be removed [28–30].

1.1.2. Segmentation Techniques

Most image segmentation tasks use traditional machine learning processes for feature extraction. The literature explains some of the important techniques used for accurate segmentation. Jaisakthi et al. [31] summarizes a semi-supervised method for segmenting skin lesions. Grab-cut techniques and K-means clustering are employed conjunctively for segmentation. After the former segments the melanoma through graph cuts, the latter fine-tunes the boundaries of the lesion. Preprocessing techniques such as image normalization and noise removal measures are used on the input images before feeding them to the pixel classifier. Mohanad Aljanabi et al. [32] proposed an artificial bee colony (ABC) method to segment skin lesions. Utilizing fewer parameters, the model is a swarm-based scheme involving preprocessing of the digital images, followed by determining the optimum threshold value of the melanoma through which the lesion is segmented, as done by Otsu thresholding. High specificity and Jaccard Index are achieved by this algorithm.

Pennisi et al. [33] introduced a technique that segments images using the Delaunay triangulation method (DTM). The approach involves parallel segmentation techniques that generate two varying images that are then merged to obtain the final lesion mask. Artifacts are removed from the images after which one process filters out the skin from the images to provide a binary mask of the lesion, and similarly, the other technique utilizes Delaunay triangulation to produce the mask. Both of these are combined to obtain the extracted lesion. The DTM technique is automated and does not require a training process, which is why it is faster than other methods. M Emre Celebi et al. [34] provides a brief overview of the border detection techniques (i.e., edge based, region based, histogram thresholding, active contours and clustering, etc.) and especially pays attention to evaluation aspects and computational issues. Lei Bi et al. [35] suggested a new automated method that performed image segmentation using image-wise supervised learning (ISL) and multiscale super pixel based cellular automata (MSCA). The authors used probabilistic mapping for automatic seed selection that removes user-defined seed selection; afterward, the MSCA model was employed for segmenting skin lesions. Ashnil Kumar et al. [36] introduced a fully convolutional network (FCN) based method for segmenting dermoscopic images. Image features were learned from embedded multi-stages of the FCN and achieved an improved segmentation accuracy (than previous works) of skin lesion without employing any preprocessing part (i.e., hair removal, contrast improvement, etc.). Yading Yuan et al. [37] proposed a convolution deconvolutional neural network (CDNN) to automate the process of the segmentation of

skin lesions. This paper focused on training strategies that makes the model more efficient, as opposed to the use of various pre- and post-processing techniques. The model generates probability maps where the elements correspond to the probability of pixels belonging to the melanoma. Berseth et al. [38] developed a U-Net architecture for segmenting skin lesions based on the probability map of the image dimension where the ten-fold cross validation technique was used for training the model. Mishra [17] presented a deep learning technique for extracting the lesion region from dermoscopic images.

This paper combines Otsu's thresholding and CNN for better results. U-Net based architecture was used to extract more complex features. Chengyao Qian et al. [39] proposed an encoder decoder architecture for segmentation inspired by DeepLab [40] and ResNet 101 was adapted for feature extraction. Frederico Guth et al. [41] introduced a U-Net 34 architecture that merged insights from U-Net and ResNet. The optimized learning rate was used for fine tuning the network and the slanted triangular learning rate strategy (STLR) was employed.

2. Materials and Methods

2.1. Dataset Modalities

We trained and tested our CNN model on dermoscopic skin images acquired from two publicly accessible datasets (i.e., PH² [42] and ISIC 2017 [43]), the latter provided by the "International Skin Imaging Collaboration" (ISIC). Example images from both datasets are shown in Figure 1.

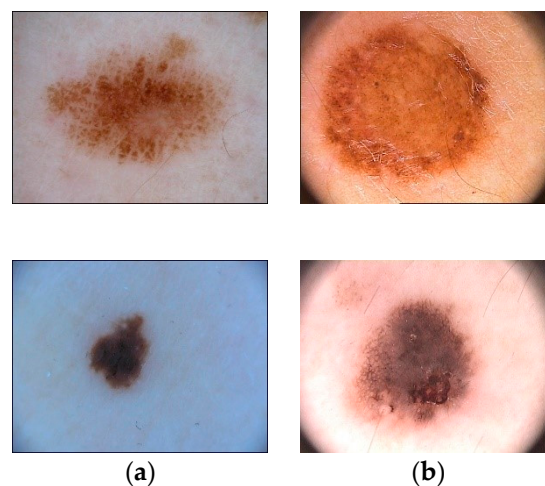


Figure 1. Examples of the ISIC-17 dataset (a) and PH² dataset (b).

We compared our model in the task of Lesion Segmentation, part 1 of the 2017 ISBI Skin Lesion Analysis Toward Melanoma Detection challenge. We evaluated our model on the ISIC-17 test data consisting of 600 images to compare its performance with state-of-the-art pipelines. Additionally, our model was also tested on the PH² dataset with its 200 dermoscopic images including 40 melanoma, 80 common nevi, and 80 atypical nevi images.

2.2. Proposed Methodology

In this section, we introduce our devised methodology, which was trained and tested on the datasets (details presented later), and the subsequent results are reported and discussed. At the outset, it is pertinent to mention that we proposed a method that out-performed other similar available methods, both in terms of model accuracy and in pixel-by-pixel similarity measure, also called the intersection over union overlap (sometimes also referred to as the Jaccard Index). We herein proceed to describe, point by point, the various subsections of the proposed method.

2.2.1. Image Preprocessing

Images are preprocessed using resizing, scaling, hair removal and data centering techniques before being given as input to the CNN model. For noise removal, morphological operations are applied. We obtained promising results by applying preprocessing practices, which are as follows.

Image Resizing: It is good practice to resize images before they are fed into the neural network. It allows the model to convolve faster, thereby saving computational power and dealing with memory constraints. Dermoscopic images vary in size and to overcome such individual differences, the images and their corresponding ground truths are down sampled to 256×256 resolution. All the RGB images are in the JPEG file format while the respective labels are in the PNG format.

Image Normalization and Standardization: Images are normalized before training to remove poor contrast issues. Normalization changes the range of pixel values, rescaling the image between 0 and 1 so that the input data is centered around zero in all dimensions. Normalization is obtained by subtracting the image from its mean value, which is then divided by the standard deviation of the image.

Hair Removal: Dermoscopic images contain hair-like artifacts that cause issues while segmenting lesion regions. A series of morphological operations are applied to the image to remove these hair-like structures. The inpainting algorithm [44] is then applied to replace the pixel values with the neighboring pixels, explained as follows:

- Conversion of RGB (Figure 2a) to grayscale image (Figure 2b);
- Black top-hat filter [45,46] is applied to the grayscale image;
- Inpainting algorithm is implemented on the generated binary mask; and
- Inpainting of the hair occupied regions with neighboring pixels.

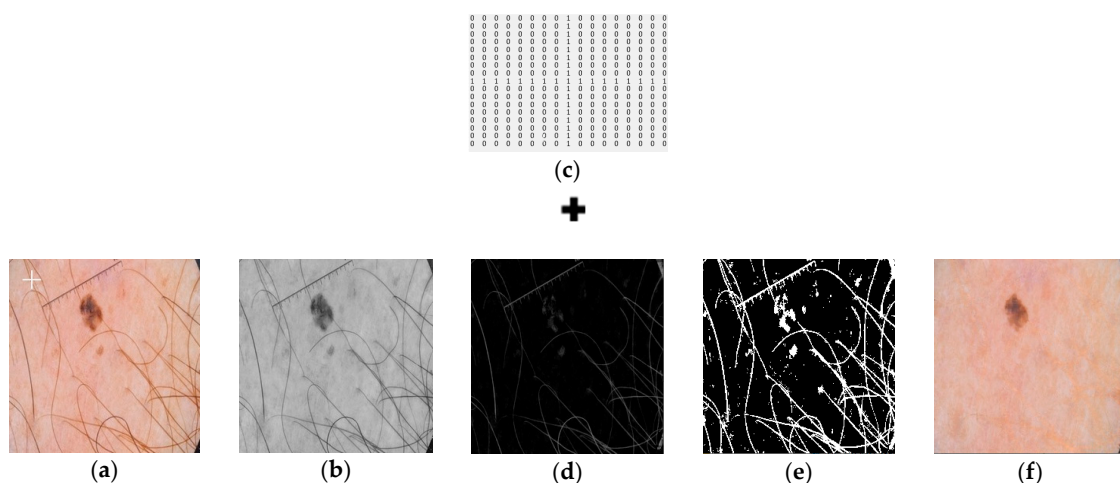


Figure 2. Representation of a single image as it is passed through the hair removal algorithm. Left to right: (a) Input image; (b) Grayscale image; (c) Cross shaped structuring element employed during morphological operations; (d) Image obtained after applying black top-hat filter; (e) Image after thresholding; (f). Final image obtained as output.

A 17×17 cross shaped structuring element is defined, as shown in Figure 2c. Black top-hat (or black hat filter) filtering is obtained by subtracting closing of image from original image. If A is the original input image and B is the closing of the input image, then black top-hat filter is defined by Equation (1):

$$\text{Black Hat}(A) = A_{\text{BH}} = (A \cdot B) - A \quad (1)$$

Closing morphological operation is the erosion of the dilation of set A and B . Closing fills small holes in the region while keeping the initial region sizes intact. It preserves the background pixels that are like the structuring element, while eliminating all other regions of the background.

The image obtained after applying the closing operation on a grayscale image is subtracted from the image itself to obtain hair like structures. Binary mask of the hair elements is obtained by applying a threshold value of “10” on the image obtained from the black top-hat filter. Images obtained from the black top-hat filter and after thresholding, respectively, are highlighted in Figure 2d,e.

The image based on the fast marching method was employed [47]. The inpainting algorithm replaces the hair structures with the bordering pixels of the image to restore the original image. This technique is commonly used in recovering old or noisy images. The image to be inpainted and the mask obtained after thresholding was used to inpaint those hairy regions that were extracted with the neighboring pixels and output was achieved (Figure 2f).

2.2.2. Model Architecture

Deep learning architectures are currently being used to solve visual recognition and object detection problems. CNN models have shown good impact over semi-automated methods for semantic segmentation. The U-Net architecture, which is based on an encoder–decoder approach, has revealed significant results in medical image segmentation. The output of these networks are binary segmentation masks.

In general, CNN models are the combination of layers (i.e., convolutional, max pooling, batch normalization, and activation layer). CNN architectures have been widely used in computer assisted medical diagnostics.

For this purpose, a CNN architecture was trained on an ISIC 2017 dataset. The network architecture (as shown in Figure 3) takes insight from both U-Net and ResNet. The contracting path (convolutional side) is based on the ResNet architecture, and the expansive path (deconvolutional side) is based on the U-Net pipeline. Overall, the network performs in an encoder–decoder fashion and is composed of 50 layers (ResNet-50). Input images of resolution 256×256 are fed into the model. The convolutional network architecture is shown in Table 1.

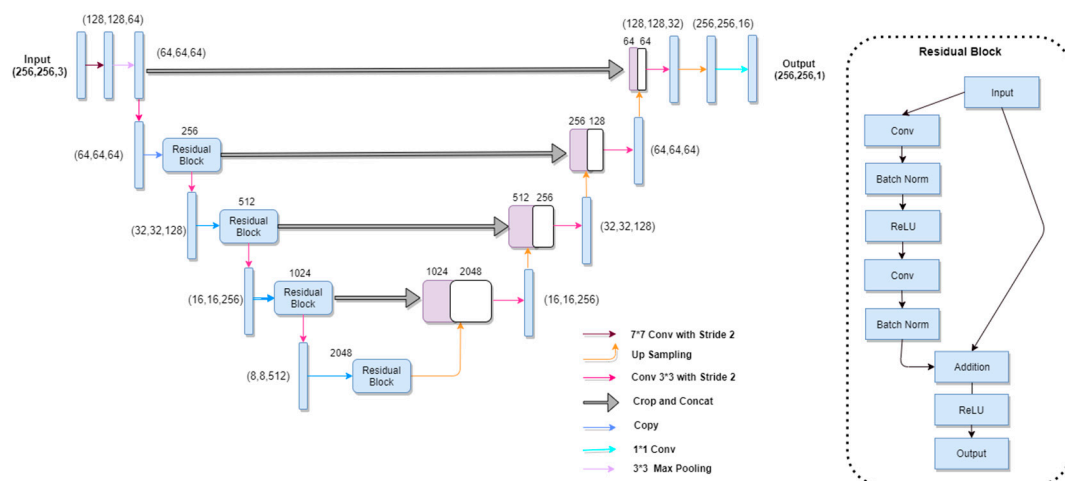


Figure 3. Schematic diagram representing UResNet-50. The encoder shown on the left is the ResNet-50, while the U-Net decoder is shown on the right. Given in the parenthesis is the channel dimensions of the incoming feature maps to each block. Arrows are defined in the legend.

On the contracting side, after the first convolutional layer, a max pooling layer is defined with a kernel of 3×3 and a stride of 2 that halves the input dimension. Repetitive blocks are introduced with three convolutional layer per block; the 1×1 convolutional layer is defined before and after each 3×3 convolutional layer. It reduces the number of channels in the input before the 3×3 convolutional layer and again, the 1×1 is defined to restore dimensions. This is called the “Bottleneck” design, which reduces the training time of the network.

Table 1. Convolutional network architecture based on ResNet-50.

Layer Name	Output Size	Kernel Size & No. of Filters
<i>Conv 1</i>	128×128	$7 \times 7, 64, \text{Stride } 2$
<i>Max Pooling</i>	64×64	$3 \times 3, \text{Stride } 2$
<i>Conv 2</i>	64×64	$\begin{bmatrix} 1 \times 1 \times 64 \\ 3 \times 3 \times 64 \\ 1 \times 1 \times 256 \end{bmatrix} \times 3$
<i>Conv 3</i>	32×32	$\begin{bmatrix} 1 \times 1 \times 128 \\ 3 \times 3 \times 128 \\ 1 \times 1 \times 512 \end{bmatrix} \times 4$
<i>Conv 4</i>	16×16	$\begin{bmatrix} 1 \times 1 \times 256 \\ 3 \times 3 \times 256 \\ 1 \times 1 \times 1024 \end{bmatrix} \times 6$
<i>Conv 5</i>	8×8	$\begin{bmatrix} 1 \times 1 \times 512 \\ 3 \times 3 \times 512 \\ 1 \times 1 \times 2048 \end{bmatrix} \times 3$

After 5 units of downsampling, the dimension ranges to 8×8 and 2048 filters. In contrast, the deconvolutional side or expansive path (as shown in Table 2) consists of 10 layers that perform deconvolution.

Table 2. The deconvolution architecture based on U-Net.

Layer Name	Kernel	Output Size & No. of Filters
U1	2×2	$16 \times 16 \times 2048$
D1	3×3	$16 \times 16 \times 256$
D2	3×3	$16 \times 16 \times 256$
U2	2×2	$32 \times 32 \times 256$
D3	3×3	$32 \times 32 \times 128$
D4	3×3	$32 \times 32 \times 128$
U3	2×2	$64 \times 64 \times 128$
D5	3×3	$64 \times 64 \times 64$
D6	3×3	$64 \times 64 \times 64$
U4	2×2	$128 \times 128 \times 64$
D7	3×3	$128 \times 128 \times 32$
D8	3×3	$128 \times 128 \times 32$
U5	2×2	$256 \times 256 \times 32$
D9	3×3	$256 \times 256 \times 16$
D10	3×3	$256 \times 256 \times 16$
Output	1×1	$256 \times 256 \times 1$

2.2.3. Network Training

We trained our model for 100 epochs and applied data augmentation during runtime, which enhances the performance as more data increases the predictability of the model so that it can classify better, thereby producing a significant effect on the segmentation results. We rotated images in three dimensions, which increased the dataset thricefold.

Early stopping is defined and the learning rate is reduced if the model loss does not decrease for 10 epochs. Our model stopped after approximately 70 epochs. Transfer learning was employed for training the model on our dataset, utilizing pre-trained weights obtained through training on the ImageNet dataset. Table 3 shows the hyperparameters used to train our model.

Table 3. Hyperparameters maintained during training.

Name	Value
Input Size	$256 \times 256 \times 3$
Batch Size	16
Learning Rate	1×10^{-3}
Optimizer	Adam
Epoch	100
Loss Function	Binary Crossentropy

3. Results

Model Evaluation

Our model was evaluated on images obtained from the International Skin Imaging Collaboration ISIC 2017. We trained our CNN model on the training group of ISIC 2017, which consisted of 2000 skin lesion images. During this process, a total training accuracy of 0.995 was obtained for 70 epochs. The variations of accuracy between the training and validation group during training is highlighted in Figure 4.

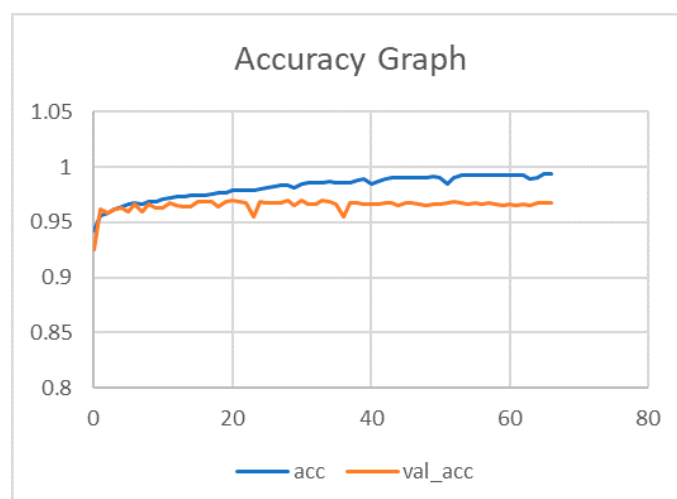


Figure 4. Training and validation accuracy of the proposed convolutional neural network model for 70 epochs.

The model was tested on the validation and test set taken from the ISIC 2017 dataset. Furthermore, the model was also tested on the PH² dataset comprising of 200 dermoscopic images. The ground truths were also available in order to check the performance of the proposed CNN model. All images went through the preprocessing step before being fed into the CNN architecture as described earlier. Parameters of convolutional layers were set during the training process. During the evaluation process, the model parameters were not changed in order to assess our model's performance on the pre-set parameters. The results of multiple subjects are shown in Figure 5.

The receiver operative characteristics (ROC) curve was used to evaluate the performance binary classifiers. ROC is a plot between the true positive rate (Sensitivity) as a function of false positive rate (Specificity) at different thresholds. This study emphasizes segmenting the lesion region, with 1 representing the lesion region and 0 representing the black region of the image. The ROC curve is the best evaluation technique that defines separability between classes. Each datapoint in a curve shows the values at a specific threshold. Figure 6 shows the ROC curve of the model on the ISIC test set.

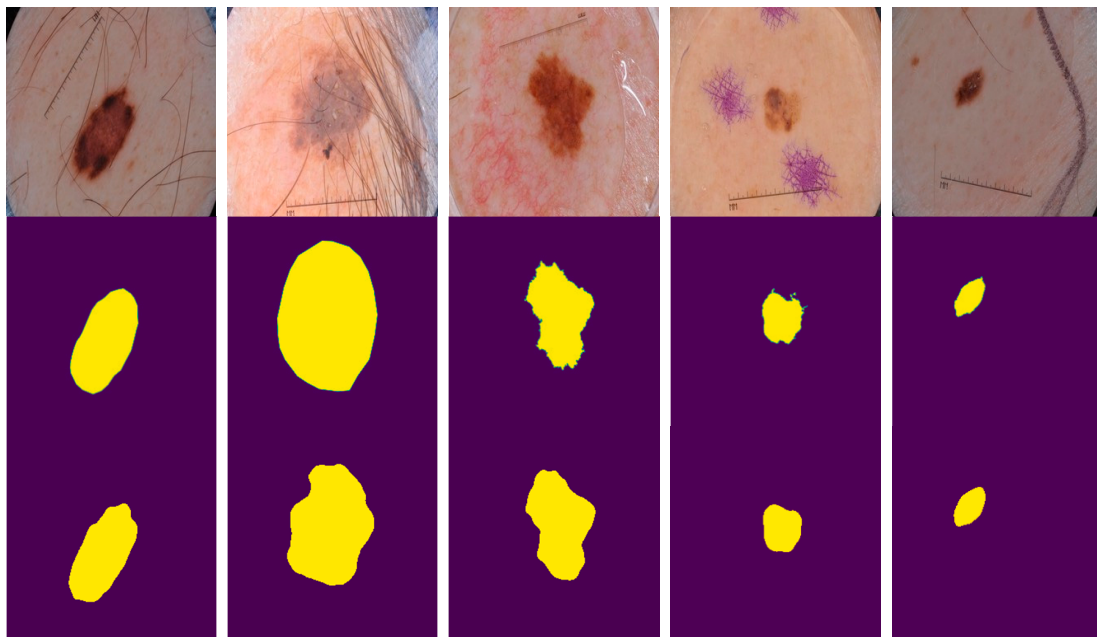


Figure 5. Example results of multiple patients. The first row contains the original images of five patients from the test set. The second row contains corresponding ground truths as provided. The third row contains predicted masks from the proposed method.

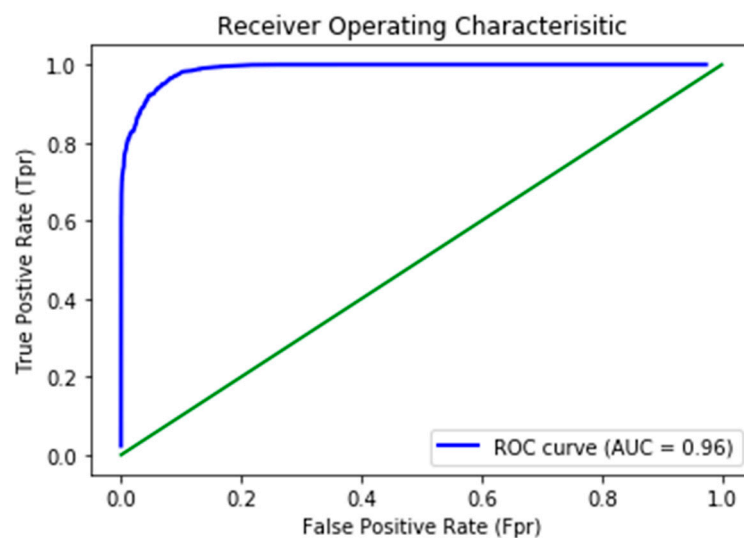


Figure 6. Receiver operative characteristics (ROC) curve generated on the ISIC-17 test set.

The ROC curve dictates the model's capability to distinguish between classes accurately. The higher the area under the curve, the higher the network's ability to distinguish two classes more precisely (i.e., either 0 or 1). The AUC of our proposed model was 0.963, illustrating the model's remarkable competence of differentiability.

4. Benchmarks

4.1. Comparison with Different Frameworks

The model was tested both with and without preprocessing on the ISIC-17 dataset to ascertain the efficiency of the hair removing algorithm. A Jaccard of 0.763 (as shown in Table 4) was achieved when the inpainting algorithm was not employed to remove the hair structures from the images, which improved considerably to 0.772 with the implementation of the preprocessing technique.

Table 4. Model performance on the ISIC-2017 test set.

Methods	Jaccard Index
No Preprocessing	Preprocessed
0.763	0.772

For evaluation, we compared our results with the existing deep learning frameworks (enlisted in Table 5) that had been tested on the ISIC-17 dataset. FCN-8s [48] achieved a JI (0.696) and DC (0.783), respectively. Although our proposed method was the deepest among the below listed frameworks, we improved the results by balanced data augmentation and reduced overfitting. Simple U-Net obtained a JI of 0.651 and a DC of 0.768.

Table 5. Comparison with different frameworks on the ISIC-2017 test set.

Methods	Jaccard Index	Dice Coefficient
FCN-8s [47]	0.696	0.783
U-Net [48]	0.651	0.768
II-FCN [49]	0.699	0.794
Auto-ED [50]	0.738	0.824
FCRN [18]	0.753	0.839
Res-UNet (Proposed)	0.772	0.858

Our proposed method is a combination of the ResNet50 based encoder and U-Net based decoder, which achieved a Jaccard index of 0.772 and Dice coefficient of 0.858.

4.2. Comparison with Top 5 Challenge Participants of Leaderboard

The intent was that this research would segment the lesion regions with higher accuracy when compared to other methods. Three different group of images were used to validate our network: (1) the ISIC 2017 test group; (2) ISIC 2017 validation group; and the (3) PH² dataset. The test group consisted of 600 dermoscopic images and the validation group was composed of 150 images. The PH² dataset is a renowned dataset and was used for further evaluation of our network and benchmarking our results with existing methods and participants in the challenge. Table 6 depicts our results in terms of the Jaccard Index as per the challenge's demand, in comparison with the top five participants from the ISIC-17 Challenge. The top ranked participant Yading Yuan et al. [21] obtained a Jaccard index of 0.765.

Table 6. Comparison of results with the challenge participants.

Authors	Model	Jaccard Index
Yading Yuan et al. [21]	CDNN	0.765
Matt Berseth et al. [51]	U-Net	0.762
Popleyi et al. [52]	FCN	0.760
Euijoon Ahn et al. [35]	ResNet	0.758
Afonso Menegola et al. [53]	VGG16	0.754
Proposed Method	Res-UNet	0.772

Different methods has been employed for segmentation purpose. The second ranked participant Matt Berseth et al. [51] obtained a JI (0.762) by employing a U-Net framework. Our technique achieved a Jaccard index of 0.772 with the proposed technique stated earlier. Based on our results, our model performed better than the existing techniques used in the associative field of study.

4.3. Evaluation of Model on the PH² Dataset

To evaluate the robustness of our proposed model, we further tested the architecture on the PH² dataset and compared our segmentations with the existing state-of-the-art techniques. The results are

listed below in Table 7. Our method achieved promising results. FCN-16s achieved a JI of 0.802 and DC of 0.881, respectively. Another framework, Mask-RCNN attained a JI of 0.839 and a DC of 0.907 on the PH² dataset.

Table 7. Comparison with different frameworks on the PH² Dataset

Methods	Jaccard Index	Dice Coefficient
FCN-16s	0.802	0.881
DeepLabV3+	0.814	0.890
Mask-RCNN	0.830	0.904
Multi-Stage FCN	-	0.906
SSLS	-	0.916
Res-UNet (Proposed)	0.854	0.924

5. Conclusions

Skin lesion segmentation is a vital step in developing a computer aided diagnosis system for skin cancer. In this paper, we successfully developed a skin lesion segmentation algorithm using CNN with an advanced hair-removal algorithm that effectively removed hair structures from the dermoscopic images, improving the accuracy considerably. We tested our model architecture on the ISIC-2017 dataset and PH² dataset, and the Jaccard index obtained thereof was 0.772 and 0.854, respectively. Our proposed method achieved promising results compared with the state-of-the-art techniques in terms of the Jaccard index. Furthermore, our CNN model was tested on a PH² dataset along with the ISIC-17 test set and produced better segmentation and performed better than the existing methods in the literature. Empirical results show that the combination of the U-Net and ResNet shows impressive results.

The limited training data used requires extensive augmentation to prevent the model from overfitting. A large dataset is therefore needed for better accuracy and generalization of the model. Furthermore, for it to achieve state-of-the-art results, the model was made to be complex and efficient, which takes more time to train as opposed to the conventional U-Net.

Our future work includes using a larger dataset to reduce overfitting problems and hyper tuning the parameters for more effective training. Additionally, a conditional random field (CRF) application can also be applied to refine the model output.

Author Contributions: Conceptualization, S.O.G. and A.W.; Data curation, K.Z.; Formal analysis, K.Z.; Methodology, K.Z.; Project administration, A.W., S.O.G.; Validation, K.Z., A.A. and M.J.; Visualization, S.O.G.; Writing—original draft, K.Z.; Writing—review & editing, S.O.G., A.W., A.A., M.J.; M.N.K. and A.S.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pham, D.L.; Xu, C.; Prince, J.L. Current methods in medical image segmentation. *Annu. Rev. Biomed. Eng.* **2000**, *2*, 315–337. [[CrossRef](#)] [[PubMed](#)]
2. Macià, F.; Pumarega, J.; Gallén, M.; Porta, M. Time from (clinical or certainty) diagnosis to treatment onset in cancer patients: The choice of diagnostic date strongly influences differences in therapeutic delay by tumor site and stage. *J. Clin. Epidemiol.* **2013**, *66*, 928–939. [[CrossRef](#)] [[PubMed](#)]
3. Oliveira, R.B.; Filho, M.E.; Ma, Z.; Papa, J.P.; Pereira, A.S.; Tavares, J.M.R.S. Computational methods for the image segmentation of pigmented skin lesions: A review. *Comput. Methods Programs Biomed.* **2016**, *131*, 127–141. [[CrossRef](#)] [[PubMed](#)]
4. Matthews, N.H.; Li, W.-Q.; Qureshi, A.A.; Weinstock, M.A.; Cho, E. Epidemiology of melanoma. In *Cutaneous Melanoma: Etiology and Therapy*; Codon Publications: Brisbane, Australia, 2017; pp. 3–22.

5. Colditz, G.A. Encyclopedia of Cancer and Society. In *Google Books*; SAGE Publications, Inc.: Thousand Oaks, CA, USA, 2015.
6. The American Cancer Society. Available online: <https://contentsubscription.cancer.org/content/dam/CRC/PDF/Public/8606.00.pdf> (accessed on 9 March 2020).
7. Mohan, S.V.; Chang, A.L.S. Advanced basal cell carcinoma: Epidemiology and therapeutic innovations. *Curr. Dermatol. Rep.* **2014**, *3*, 40–45. [[CrossRef](#)]
8. Guy, G.P.; Machlin, S.R.; Ekwueme, D.U.; Yabroff, K.R. Prevalence and costs of skin cancer treatment in the U.S., 2002–2006 and 2007–2011. *Am. J. Prev. Med.* **2015**, *48*, 183–187. [[CrossRef](#)]
9. Barata, C.; Ruela, M.; Francisco, M.; Mendonca, T.; Marques, J.S. Two systems for the detection of melanomas in Dermoscopy images using texture and color features. *IEEE Syst. J.* **2014**, *8*, 965–979. [[CrossRef](#)]
10. Schaefer, G.; Rajab, M.I.; Celebi, M.E.; Iyatomi, H. Color and contrast enhancement for improved skin lesion segmentation. *Comput. Med. Imaging Graph.* **2011**, *35*, 99–104. [[CrossRef](#)]
11. Gómez, D.D.; Butakoff, C.; Ersbøll, B.K.; Stoecker, W. Independent histogram pursuit for segmentation of skin lesions. *IEEE Trans. Biomed. Eng.* **2008**, *55*, 157–161. [[CrossRef](#)]
12. Maeda, J.; Kawano, A.; Yamauchi, S.; Suzuki, Y.; Marçal, A.R.S.; Mendonça, T. Perceptual image segmentation using fuzzy-based hierarchical algorithm and its application to images Dermoscopy. In Proceedings of the 2008 IEEE Conference on Soft Computing on Industrial Applications, Muroran, Japan, 25–27 June 2008; pp. 66–71.
13. Yüksel, M.E.; Borlu, M. Accurate segmentation of dermoscopic images by image thresholding based on type-2 fuzzy logic. *IEEE Trans. Fuzzy Syst.* **2009**, *17*, 976–982. [[CrossRef](#)]
14. Xie, F.Y.; Qin, S.Y.; Jiang, Z.G.; Meng, R.S. PDE-based unsupervised repair of hair-occluded information in Dermoscopy images of melanoma. *Comput. Med. Imaging Graph.* **2009**, *33*, 275–282. [[CrossRef](#)]
15. Silveira, M.; Nascimento, J.C.; Marques, J.S.; Marcal, A.R.S.; Mendonca, T.; Yamauchi, S.; Maeda, J.; Rozeira, J. Comparison of segmentation methods for melanoma diagnosis in dermoscopy images. *IEEE J. Sel. Top. Signal Process.* **2009**, *3*, 35–45. [[CrossRef](#)]
16. Ma, Z.; Tavares, J.M.R.S. A novel approach to segment skin lesions in dermoscopic images based on a deformable model. *IEEE J. Biomed. Heal. Inform.* **2016**, *20*, 615–623. [[CrossRef](#)] [[PubMed](#)]
17. Mishra, R.; Daescu, O. Deep learning for skin lesion segmentation. In Proceedings of the 2017 IEEE International Conference on Bioinformatics and Biomedicine, BIBM, Kansas City, MO, USA, 13–16 November 2017.
18. Li, Y.; Shen, L. Skin lesion analysis towards melanoma detection using deep learning network. *Sensors* **2018**, *18*, 556. [[CrossRef](#)] [[PubMed](#)]
19. Jafari, M.H.; Karimi, N.; Nasr-Esfahani, E.; Samavi, S.; Soroushmehr, S.M.R.; Ward, K.; Najarian, K. Skin lesion segmentation in clinical images using deep learning. In Proceedings of the International Conference on Pattern Recognition, Cancun, Mexico, 4–8 December 2016.
20. Tamilselvi, P.R. Analysis of image segmentation techniques for medical images. *Int. Conf. Emerg. Res. Comput. Inf. Commun. Appl.* **2014**, *2*, 73–76.
21. Yuan, Y.; Chao, M.; Lo, Y.C. Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance. *IEEE Trans. Med. Imaging* **2017**, *36*, 1876–1886. [[CrossRef](#)] [[PubMed](#)]
22. Korotkov, K.; Garcia, R. Computerized analysis of pigmented skin lesions: A review. *Artif. Intell. Med.* **2012**, *56*, 69–90. [[CrossRef](#)] [[PubMed](#)]
23. Celebi, M.E.; Iyatomi, H.; Schaefer, G. Contrast enhancement in dermoscopy images by maximizing a histogram bimodality measure. In Proceedings of the International Conference on Image Processing ICIP, Cairo, Egypt, 7–10 November 2009; pp. 2601–2604.
24. Beuren, A.T.; Janasieivicz, R.; Pinheiro, G.; Grando, N.; Facon, J. Skin melanoma segmentation by morphological approach. In Proceedings of the ACM International Conference Proceeding Series, Chennai, India, 3–5 August 2012; pp. 972–978.
25. Lee, T.; Ng, V.; Gallagher, R.; Coldman, A.; McLean, D. DullRazor: A software approach to hair removal from images. *Comput. Biol. Med.* **1997**, *27*, 533–543. [[CrossRef](#)]
26. Chanda, B.; Majumder, D.D. *Digital Image Processing and Analysis*, 2nd ed.; PHI Learning Pvt. Ltd.: Delhi, India, 2011.
27. Celebi, M.E.; Kingravi, K.A.; Iyatomi, H.; Aslandogan, Y.A.; Stoecker, W.V.; Moss, R.H.; Malters, J.M.; Grichnik, J.M.; Marghoob, A.A.; Rabinovitz, H.S.; et al. Border detection in Dermoscopy images using statistical region merging. *Skin Res. Technol.* **2008**, *14*, 347–353. [[CrossRef](#)]

28. Guillemot, C.; Le Meur, O. Image Inpainting: Overview and recent advances. *IEEE Signal Process. Mag.* **2014**, *31*, 127–144. [[CrossRef](#)]
29. Shen, H.; Li, X.; Cheng, Q.; Zeng, C.; Yang, G.; Li, H.; Zhang, L. Missing information reconstruction of remote sensing data: A technical review. *IEEE Geosci. Remote Sens. Mag.* **2015**, *3*, 61–85. [[CrossRef](#)]
30. Li, X.; Shen, H.; Zhang, L.; Zhang, H.; Yuan, Q.; Yang, G. Recovering quantitative remote sensing products contaminated by thick clouds and shadows using multitemporal dictionary learning. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 7086–7098.
31. Jaisakthi, S.M.; Mirunalini, P.; Aravindan, C. Automated skin lesion segmentation of Dermoscopic images using grabcut and kmeans algorithms. *IET Comput. Vis.* **2018**, *12*, 1088–1095. [[CrossRef](#)]
32. Aljanabi, M.; Özok, Y.E.; Rahebi, J.; Abdullah, A.S. Skin lesion segmentation method for Dermoscopy images using artificial bee colony algorithm. *Symmetry* **2018**, *10*, 347. [[CrossRef](#)]
33. Pennisi, A.; Bloisi, D.D.; Nardi, D.; Giampetruzzi, A.R.; Mondino, C.; Facchiano, A. Skin lesion image segmentation using Delaunay Triangulation for melanoma detection. *Comput. Med. Imaging Graph.* **2016**, *52*, 89–103. [[CrossRef](#)] [[PubMed](#)]
34. Celebi, M.E.; Iyatomi, H.; Schaefer, G.; Stoecker, W.V. Lesion border detection in Dermoscopy images. *Comput. Med. Imaging Graph.* **2009**, *33*, 148–153. [[CrossRef](#)]
35. 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. In Proceedings of the International Symposium on Biomedical Imaging, Prague, Czech Republic, 13–16 April 2016; pp. 1059–1062.
36. Bi, L.; Kim, J.; Ahn, E.; Kumar, A.; Fulham, M.; Feng, D. Dermoscopic image segmentation via multistage fully convolutional networks. *IEEE Trans. Biomed. Eng.* **2017**, *64*, 2065–2074. [[CrossRef](#)]
37. Yuan, Y. Automatic skin lesion segmentation with fully convolutional-deconvolutional networks. *arXiv* **2017**, arXiv:1703.05165.
38. Berseth, M. ISIC 2017-Skin Lesion Analysis Towards Melanoma Detection, International Skin Imaging Collaboration. 2017. Available online: <https://arxiv.org/abs/1703.00523> (accessed on 9 September 2019).
39. Qian, C.; Jiang, H.; Liu, T. *ISIC 2018-Skin Lesion Analysis*. 2018. *ISIC—Skin Image Analysis Workshop and Challenge @ MICCAI 2018 Hosted by the International Skin Imaging Collaboration (ISIC)*; Springer: Berlin, Germany, 2018.
40. Chen, L.-C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. IEEE Transactions on Pattern Analysis and Machine Intelligence DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *40*, 834–848. [[CrossRef](#)]
41. Guth, F.; Decampos, T.E. Skin Lesion Segmentation Using U-Net and Good Training Strategies. Available online: <https://arxiv.org/abs/1811.11314> (accessed on 9 September 2019).
42. Mendonca, T.; Ferreira, P.M.; Marques, J.S.; Marcal, A.R.S.; Rozeira, J. PH²—A dermoscopic image database for research and benchmarking. In Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Osaka, Japan, 3–7 July 2013.
43. Codella, N.C.F.; Gutman, D.; Celebi, M.E.; Helba, B.; Marchetti, M.A.; Dusza, S.W.; Kalloo, A.; Liopyris, K.; Mishra, N.; Kittler, H.; 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 Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2017.
44. Telea, A. An image inpainting technique based on the fast marching method. *J. Graph. Tools* **2004**, *9*, 23–34. [[CrossRef](#)]
45. Thapar, S. Study and implementation of various morphology based image contrast enhancement techniques. *Int. J. Comput. Bus. Res.* **2012**, *128*, 2229–6166.
46. Wang, G.; Wang, Y.; Li, H.; Chen, X.; Lu, H.; Ma, Y.; Peng, C.; Tang, L. Morphological background detection and illumination normalization of text image with poor lighting. *PLoS ONE* **2014**, *9*, e110991. [[CrossRef](#)] [[PubMed](#)]
47. Shelhamer, E.; Long, J.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.

48. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer Assisted Interventions, Munich, Germany, 5–9 October 2015; pp. 234–241.
49. Wen, H. II-FCN for skin lesion analysis towards melanoma detection. *arXiv* **2017**, arXiv:1702.08699.
50. Attia, M.; Hossny, M.; Nahavandi, S.; Yazdabadi, A. Spatially aware melanoma segmentation using hybrid deep learning techniques. *arXiv* **2017**, arXiv:1702.07963.
51. Berseth, M. ISIC 2017-skin lesion analysis towards melanoma detection. *arXiv* **2017**, arXiv:1703.00523.
52. Bi, L.; Kim, J.; Ahn, E.; Feng, D. Automatic Skin Lesion Analysis using Large-scale Der-moscopy Images and Deep Residual Networks. Available online: <https://arxiv.org/ftp/arxiv/papers/1703/1703.04197.pdf> (accessed on 11 December 2019).
53. Menegola, A.; Tavares, J.; Fornaciali, M.; Li, L.T.; Avila, S.; Valle, E. RECOD Titans at ISIC Challenge 2017. Mar. 2017. Available online: <https://arxiv.org/abs/1703.04819> (accessed on 19 November 2017).



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).