2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

# Lesion Border Detection of Skin Cancer Images Using Deep Fully Convolutional Neural Network with Customized Weights

R. Kaur, H. GholamHosseini*, *Senior Member, IEEE,* and R. Sinha, *Member, IEEE*

*Abstract*— **Deep learning techniques have been widely employed in semantic segmentation problems, especially in medical image analysis, for understanding image patterns. Skin cancer is a life-threatening problem, whereas timely detection can prevent and reduce the mortality rate. The aim is to segment the lesion area from the skin cancer image to help experts in the process of deeply understanding tissues and cancer cells' formation. Thus, we proposed an improved fully convolutional neural network (FCNN) architecture for lesion segmentation in dermoscopic skin cancer images. The FCNN network consists of multiple feature extraction layers forming a deep framework to obtain a larger vision for generating pixel labels. The novelty of the network lies in the way layers are stacked and the generation of customized weights in each convolutional layer to produce a full resolution feature map. The proposed model was compared with the top four winners of the International Skin Imaging Collaboration (ISIC) challenge using evaluation metrics such as accuracy, Jaccard index, and dice co-efficient. It outperformed the given state-of-the-art methods with higher values of the accuracy and Jaccard index.**

## I. Introduction

Skin cancer is the excessive growth of abnormal cells in the outermost skin layer. Cancerous cells overgrow and damage surrounding tissues to form a malignant tumor known as melanoma. Melanoma is the most life-threatening form of skin cancer that spreads rapidly into the body, causing death if left untreated. It is estimated that new melanoma cases will increase by 5.8%, and the ratio of deaths due to melanoma will increase by 4.8% worldwide by 2021 [1]. In 2020, melanoma affected 2 million people in the USA [2], 15,229 in Australia in 2019 [3], and 26,158 in New Zealand in 2018 [4]. Melanoma is the fourth most common type of skin cancer in New Zealand. According to the American cancer society [5], the annual cost of skin cancer treatment in the U.S. is approximately $8.1 billion; $4.8 billion for nonmelanoma, and $3.3 billion for melanoma. As per the New Zealand census, $28.6 million costs are involved in cancer treatment, thus making it the most expensive procedure for the health sector.

The primary cause of melanoma is direct exposure to UV radiation, and the ageing population is more prone to developing cancer than other age groups. The risk factors that contribute to melanoma formation include fair complexion, overexposure to sun rays, sunburn, genetic history, and weak immune system [6, 7]. Early detection and treatment help increase the survival rate; therefore, it is vital to detect this disease timely. However, there are few drawbacks in traditional clinical procedures, such as tedious, time-consuming, expensive, accuracy variations, and limited availability of resources. Therefore, we aim to propose an automatic segmentation technique using deep learning to extract infected areas from the skin photographs and assists clinicians to speed up the process of understanding the nature of lesion patterns. It is crucial to analyze the shape, texture, and size of the lesion since benign moles closely resemble melanoma moles, thus, creating confusion.

Our choice of methodology is deep learning networks because of recent advances in artificial intelligence in solving complex and challenging problems. In the literature, it has been seen that CNNs are successfully applied to classification problems, where their use to solve semantic segmentation problems has emerged recently. For example, the authors in [8] designed a deep convolutional neural network (DCNN) of 50 layers by adding residual blocks presenting a two-phase framework of segmentation followed by classification. They proved that deeper networks are efficient in producing more detailed feature maps for recognition. They were ranked as the winner in ISBI 2016 challenge as compared to VGG16 and other participated teams. Bi et al. gave a multi-block fully convolutional neural network (FCN) to learn local coarse maps in the initial stages and finely detailed features in the later stages. They claimed that their method produced improved detection for the $PH^2$ dataset [9] but did not give higher performance than the ISIC challenge participants. Yuan et al. [10] proposed an end-to-end trainable DCNN containing 19 layers incorporating Jaccard distance as a loss function. They proved by comparing a different combination of hyperparameters that selecting suitable parameters is important to gain high performance. Another work by Al-Masni et al. and Goyal et al. [11, 12] proposed a full resolution convolutional network (FrCN) that calculates each pixel's resolution to produce a fully segmented map. Fengyine et al.[13] designed a network by including different main, spatial, and channel-wise attention branches and fusing their outcomes to generate high-resolution feature blocks. Their network is efficient in producing accurate lesion boundaries. Many other researchers proposed networks by customizing, extending, and applying transfer learning. The recent survey can be seen in [14] for comparative analysis.

The literature study showed that deep learning has gained satisfactory results for the segmentation of skin lesions. However, there is room for developing a more efficient network for generating accurate lesion boundaries. Because

Ranpreet Kaur, is a PhD candidate at Auckland University of Technology, New Zealand (phone: 00642041794908; e-mail: ranpreet.kaur@aut.nz).

*Hamid GholamHosseini., Associate Professor in the Electrical & Electronics Engineering Department, Auckland University of Technology, New Zealand (e-mail: hamid.gholamhosseini@aut.ac.nz).

Roopak Sinha, Head of Computer Science Department, Auckland University of Technology, New Zealand. (email: roopak.sinha@aut.ac.nz).

manually marking the lesion boundaries and generating ground truths is a costly and labor-intensive process. Moreover, annotations marked by a dermatologist are not always accurate that can further affect the performance of the software diagnostic tool. Our few contributions are (a) designing a new FCNN framework by replacing traditional fully convolutional layers with $1 \times 1$ convolutional layers focusing on calculating each pixel's probability. (b) fusing elements of shallow layers that have local information to enhance segmentation results. (c) customized weights are initialized in each convolution layer using a leakyReLU function.

The remainder of this paper is organized into sections as section II will explain the datasets. In section III, we will elaborate on the methodology of the proposed work. In section IV, the experimental analysis will be presented, and finally, we will conclude the paper in section V.

## II. MATERIALS

The proposed study employed three benchmark skin cancer datasets. The first two datasets were from *The International skin Imaging challenge (ISIC)*, namely ISIC 2016 [15], ISIC 2017 [16]. These sets contain dermoscopic skin cancer images of types benign and malignant. The ISIC 2016 contains 900 training, 379 test samples, and 2000 training and 366 test images in ISIC 2017. The corresponding ground truth images are available, which are annotated by the experts for performing a comparison between segmented and ground images. The second dataset used for testing purposes only is the $PH^2$ [9] obtained from the *Dermatology Service of Hospital Pedro Hispano (Matosinhos, Portugal)*. The dataset has 160 nevus images and 40 dermoscopic melanoma images. The images are resized to $192 \times 256$ dimensions using the nearest neighbor interpolation method to reduce the computation time and to increase the network's performance.

## III. METHODOLOGY

A dense convolutional neural network is developed with 16 convolutional layers organized into 5 blocks followed by batch normalization, leakyReLU, and pooling layers as shown in Fig. 1. The blocks of the network are cascaded with transposed layers, crop layer, and double $1 \times 1$ convolution layer. The network consists of multiple convolutional layers, calculates the input data features by sliding kernels over an entire image. In the convolutional layer, few parameters are artificially set, such as kernel size, the number of kernels, stride, and padding. The number of kernels is increasing in each convolutional layer with kernel size $3 \times 3$. The padding defines the style of handling border samples if the input image does not fit properly, and stride is the step size of the moving kernel over an image. Another parameter set in the convolutional layer is a customized function proposed by He et al. [17] to calculate the weights using:

$$\sigma^2 = \frac{2}{(1+s^2)m},\ where\ m = kernelsize(i) * kernelsize(j) * numchannels \quad (1)$$

The weights are initialized by setting the mean value of normalized distribution to zero and variance is computed by using (1). Here, *s* defines the scale of the leakyReLU layer that is set as '*0.3*' after fine-tuning, and '*m*' is calculated by multiplying the depth and width of a kernel with the number of channels. Next, a leakyReLU activation function is used to improves the network's performance without adding any extra computational cost or creating any overfitting problem. leakyReLU function passes the input samples *'x'* to the next layer if it is positive, and multiply the sample with any *scalar* value if it is negative using the following equation:

$$ReLU(x) = \begin{cases} scalar * x, & x < 0 \\ x, & x \geq 0 \end{cases} \quad (2)$$

A batch normalization layer is used between convolution and leakyReLU to normalize each input data according to mini-batch size. A mini-batch size is a variation of the gradient descent algorithm that splits the training dataset into small batches to calculate network error and update network coefficients. Batch size is the number of samples processed before the network is updated.

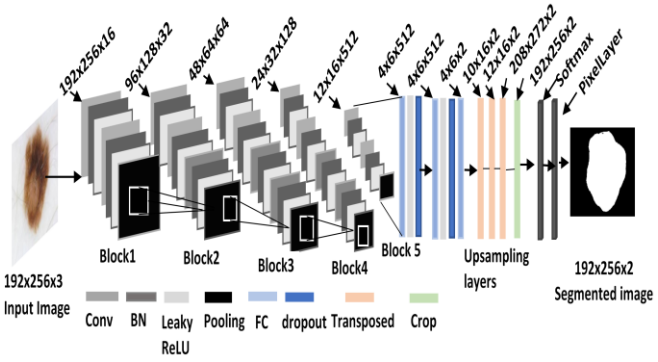


Figure 1. The proposed architecture of the CNN

The network is designed to obtain deep features; however, spatial resolution is lost as the network goes deeper and deeper due to the successive use of pooling layers reducing feature map size. Therefore, two up-sampling layers are embedded at the end to transform the feature map back to its original dimensions. Moreover, the shallow layers contain more local information about the object domain. Thus, the output of pooling layers is fused elementwise to enhance the results. At the end of the network, a double $1 \times 1$ convolutional layer is used to convert features into dense feature maps instead of fully connected layers. Then Softmax function [18] is used to transform the feature map coefficients between 0 and 1 using (3) so that input vector elements can be interpreted as pixels' probabilities.

$$softmax(\vec{x})_i = \frac{e^{x_i}}{\sum_{n=1}^{K} e^{x_j}} \quad (3)$$

$\vec{x}_i$ is an input vector, $e^{x_i}$ is a standard exponential function for the input vector $x_i$, $K$ is the number of classes predicted by the network, and $e^{x_j}$ is the exponential function for the output vector. The exponential function used in the Softmax layer normalizes the output of the neural network into the probability distribution over predicted outcomes. Finally, the pixel classification layer applied cross-entropy loss function to identify loss between the predicted $P$ and target $T$ samples using:

$$Loss = \frac{1}{N} \sum_{i=1}^{K} \sum_{n=1}^{N} w_i T_{ni} \log(P_{ni}) \quad (4)$$

Here, $N$ is the number of observations, $K$ is the number of classes, and $w$ is a vector of weights determined by the network for each class. The entropy loss is mostly used to evaluate the performance of medical image segmentation compared to the ground truth images.

## IV. RESULTS

This section illustrates the experimental results produced by the network trained with ISIC 2016, and ISIC 2017 training samples, and tested on ISIC 2016-2017 and $PH^2$ test datasets. Fig. 2 shows the accuracy pattern versus the number of iterations. It presents how the network offered higher accuracy with the increase in the iterations on ISIC 2016 test set. Similarly, it presents the performance of the network for ISIC 2017 and $PH^2$ test data offering rising accuracy over the iterations. Moreover, the network objectively evaluated based on performance metrics such as accuracy (ACC), Jaccard index (JAC), and Sorensen Dice coefficient (DICE) and compared with the other related studies in Table I. ACC defines the number of correctly identified pixels over the total number of pixels, JAC index is the overlapping ratio between correctly classified pixels divided by a total number of ground truth pixels and predicted pixels in that class. The third parameter DICE indicates the matching score between the predicted boundary of each class and the truly segmented boundary in the ground truth.

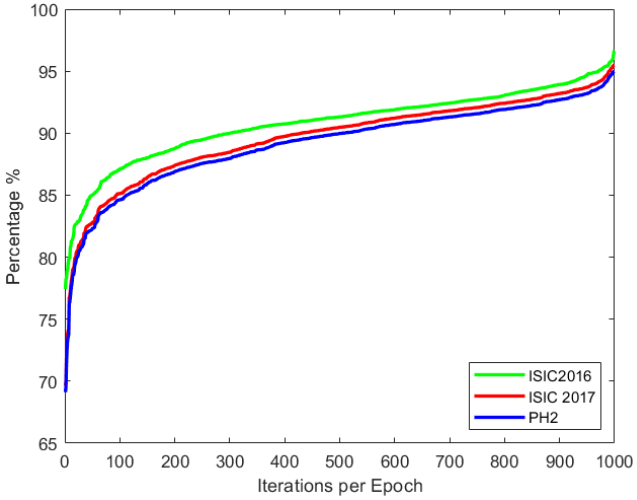


Figure 2. Performance on the ISIC 2016, 2017 & $PH^2$ test sets.

These metrics are calculated using elements of the confusion matrix i.e., true positives, true negatives, false positives, and false negatives (TP, TN, FP, and FN, respectively) using the following formulas:

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (5)$$

$$\text{Jaccard Index} = \frac{\text{TP}}{\text{TP+FP+FN}} \quad (6)$$

$$\text{DICE} = \frac{2*\text{TP}}{2*\text{TP+FP+FN}} \quad (7)$$

The value of these parameters is expected high for obtaining better segmentation results. Table I shows the quantitative results of the proposed model on the adopted datasets. The highest ACC and JAC index score were recorded for the network as compared to the teams participated in ISIC 2016 and 2017 challenge. The primary parameter used by the ISIC challenge to rank their winners is the JAC index. The proposed network outperformed the other state-of-the-art methods with an ACC of 96.3% and JAC index of 84.5% on ISIC 2016 and 95.9% ACC and 78.9% JAC on ISIC 2017. Additionally, the performance analysis of the trained model on the $PH^2$ test dataset showed higher values of ACC 94.9% and JAC index 90.9% as compared to other related studies. It is worth mentioning here that we obtained lesser values of the DICE score that defines the overlapping ratio of boundary lines between segmented and target output. This is because any pre-processing technique has not been applied to enhance noisy images. In future work, the emphasis will be on improving noisy images to enhance segmentation results.

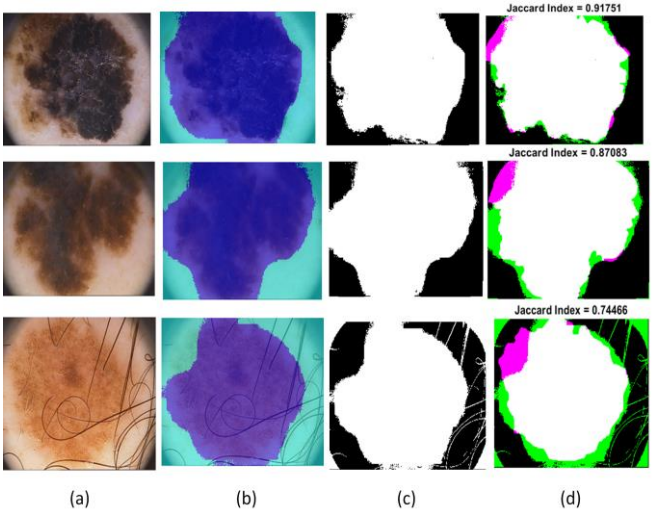


Figure 3. Segmentation results on few images.

Fig. 3 displays the segmented outputs produced by the proposed model. Fig. 3a presents the original image, Fig. 3b shows the labeled images in which the network marked all nonzero pixels with dark blue color and all background pixels with cyan color. Then in Fig. 3c, the segmented images can be seen, and Fig. 3d shows the overlapping of ground truth and a segmented image representing the JAC score. In this image, green and magenta color displays where the intensities are

TABLE I. QUANTITATIVE ANALYSIS OF THE PROPOSED MODEL ON ISIC 2016, ISIC 2017, AND $PH^2$ TEST DATASET

| ISIC 2016 | | | | ISIC 2017 | | | | $PH^2$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **References** | **ACC** | **JAC** | **DICE** | **References** | **ACC** | **JAC** | **DICE** | **References** | **ACC** | **JAC** | **DICE** |
| Proposed method | **0.963** | **0.845** | 0.623 | Proposed method | **0.959** | **0.789** | 0.607 | Proposed method | **0.948** | **0.909** | 0.743 |
| U. Sanchez [19] | 0.953 | 0.843 | 0.910 | Y. Yuan [20] | 0.934 | 0.784 | **0.849** | Unver [21] | 0.929 | 0.881 | 0.795 |
| L. Yu [8] | 0.949 | 0.829 | 0.897 | M. Berseth [22] | 0.932 | 0.762 | 0.847 | A. Nazi [23] | ----- | 0.870 | 0.879 |
| M. Rahman [24] | 0.952 | 0.822 | 0.895 | L. Bi [25] | 0.934 | 0.760 | 0.844 | L. Bi [26] | 0.942 | 0.839 | **0.906** |
| L. Huang [27] | 0.958 | 0.811 | **0.911** | Menegola [28] | 0.931 | 0.754 | 0.839 | F. Afza [29] | 0.937 | ---- | ----- |

different in the two images. It can be seen that the third sample image is quite noisy, containing hairlines. Thus, the corresponding JAC score is less than the other two images, which confirms that noise artifacts affect the algorithm's performance.

## V. Conclusion

This paper discussed a CNN model for skin lesion segmentation on challenging dermoscopic skin cancer images containing color variations, hairlines, and poor illumination. The network was designed using 16 convolutional layers with additional weight calculation functions in each layer to generate customized weights. Moreover, the up-sampling layers were included to regain the spatial resolution of the output samples. The network's hyperparameters were finely tuned to enhance segmentation results. The network outperformed the other state-of-the-art methods with high segmentation ACC and JAC score. The proposed network can be successfully applied in clinical settings to generate ground truth patterns and understand lesion nature for cancer diagnosis. Future research will evaluate the proposed network using versatile datasets to make it more general for other application domains.

## References


[1] "Skin cancer facts and statistics." https://www.skincancer.org/skin-cancer-information/skin-cancer-facts (accessed 2020).

[2] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2020," *CA: a cancer journal for clinicians*, vol. 70, no. 1, pp. 7-30, 2020.

[3] M. Fransen et al., "Non-melanoma skin cancer in Australia," *Medical Journal of Australia*, vol. 197, no. 10, pp. 565-568, 2012.

[4] "Cancer: Historical summary 1948-2013." Wellington: Ministry of Health. https://www.health.govt.nz/publication/cancer-historical-summary-1948-2013 (accessed 2019).

[5] "Cancer Facts and Figures 2021." https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2021.html (accessed 2020).

[6] F. Bogo, F. Peruch, A. B. Fortina, and E. Peserico, "Where's the lesion?: variability in human and automated segmentation of dermoscopy images of melanocytic skin lesions," *Dermoscopy Image Analysis: CRC Press*, 2015, pp. 82-110.

[7] D. Schadendorf et al., "Melanoma," *Nature reviews Disease primers*, vol. 1, no. 1, pp. 1-20, 2015.

[8] L. Yu, H. Chen, et al., "Automated melanoma recognition in dermoscopy images via very deep residual networks," *IEEE transactions on medical imaging*, vol. 36, no. 4, pp. 994-1004, 2016.

[9] "$PH^2$ Database." https://www.fc.up.pt/addi/ph2%20database.html (accessed 2019).

[10] Y. Yuan, M. Chao, and Y.-C. Lo, "Automatic skin lesion segmentation using deep fully convolutional networks with Jaccard distance," *IEEE transactions on medical imaging*, vol. 36, no. 9, pp. 1876-1886, 2017.

[11] M. A. Al-Masni, M. A. Al-antari, M.-T. Choi, S.-M. Han, and T.-S. Kim, "Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks," *Computer methods and programs in biomedicine*, vol. 162, pp. 221-231, 2018.

[12] Manu Goyal, Jiahua Ng, Amanda Oakley, and Moi Hoon Yap "Skin lesion boundary segmentation with fully automated deep extreme cut methods", *Proc. SPIE 10953, Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging*, 109530Q (15 March 2019).

[13] F. Xie, J. Yang, J. Liu, et al., "Skin lesion segmentation using high-resolution convolutional neural network," *Computer methods and programs in biomedicine*, vol. 186, pp. 105241, 2020.

[14] R. Kaymak, C. Kaymak, and A. Ucar, "Skin lesion segmentation using fully convolutional networks: A comparative experimental study," *Expert Systems with Applications*, vol. 161, pp. 113742, 2020.

[15] D. Gutman et al., "Skin lesion analysis toward melanoma detection: A challenge at the international symposium on biomedical imaging (ISBI) 2016, hosted by the international skin imaging collaboration (ISIC)," arXiv preprint arXiv:1605.01397, 2016.

[16] N. C. F. Codella 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)," *Int. Symposium on Biomedical Imaging (ISBI 2018)*, Washington, DC, USA, 2018, pp. 168-172, DOI: 10.1109/ISBI.2018.8363547.

[17] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," *in Proc. of the IEEE Int. Conf. on computer vision*, 2015, pp. 1026-1034.

[18] C. M. Bishop, "Pattern recognition and machine learning". *Springer*, 2006.

[19] U. Sanchez, "ISIC 2016 challenge Task 1: Lesion segmentation " 2016.

[20] Y. Yuan, "Automatic skin lesion segmentation with fully convolutional-deconvolutional networks," arXiv preprint arXiv:1703.05165, 2017.

[21] H. M. Ünver and E. Ayan, "Skin lesion segmentation in dermoscopic images with a combination of YOLO and grab cut algorithm," *Diagnostics*, vol. 9, no. 3, pp. 72, 2019.

[22] M. Berseth, "ISIC 2017 – Skin Lesion Analysis Towards Melanoma Detection," *ISIC Challenge, 2017*.

[23] Z. Al Nazi and T. A. Abir, "Automatic skin lesion segmentation and melanoma detection: Transfer learning approach with U-Net and DCNN-SVM," *in Proc. of Int. Conf. on Computational Intelligence*, 2020: Springer, pp. 371-381.

[24] M. U. Rahman, "ISIC 2016 challenge Task 1: Lesion Segmentation," *ISIC 2016*.

[25] L. Bi, J. Kim, E. Ahn, and D. Feng, "Automatic skin lesion analysis using large-scale dermoscopy images and deep residual networks," arXiv preprint arXiv:1703.04197, 2017.

[26] L. Bi, J. Kim, E. Ahn, A. Kumar, M. Fulham, and D. Feng, "Dermoscopic image segmentation via multistage fully convolutional networks," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 9, pp. 2065-2074, 2017.

[27] L. Huang, Y.-g. Zhao, and T.-j. Yang, "Skin lesion segmentation using object scale-oriented fully convolutional neural networks," *Signal, Image, and Video Processing*, vol. 13, no. 3, pp. 431-438, 2019.

[28] A. Menegola, J. Tavares, M. Fornaciali, L. T. Li, S. Avila, and E. Valle, "RECOD titans at ISIC challenge 2017," arXiv preprint arXiv:1703.04819, 2017.

[29] F. Afza, M. A. Khan, M. Sharif, and A. Rehman, "Microscopic skin laceration segmentation and classification: A framework of statistical normal distribution and optimal feature selection," *Microscopy research and technique*, vol. 82, no. 9, pp. 1471-1488, 2019.