HHS Public Access

Author manuscript

Annu Int Conf IEEE Eng Med Biol Soc. Author manuscript; available in PMC 2021 September 17.

Published in final edited form as:

Annu Int Conf IEEE Eng Med Biol Soc. 2019 July ; 2019: 1002–1005. doi:10.1109/EMBC.2019.8856665.

Diabetic Wound Segmentation using Convolutional Neural Networks

Can Cui,

Department of Electrical and Computer Engineering, University of Michigan, Ann Arbor, MI, USA

Karl Thurnhofer-Hemsi,

Department of Computer Languages and Computer Science, University of Málaga, Bulevar Louis Pasteur, 35, 29071 Málaga, Spain

Reza Soroushmehr,

Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA

Abinash Mishra,

Kingsbrook Jewish Medical Center, New York, USA

Jonathan Gryak,

Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA

Enrique Domínguez,

Department of Computer Languages and Computer Science, University of Málaga, Bulevar Louis Pasteur, 35, 29071 Málaga, Spain

Kayvan Najarian,

Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA

Ezequiel López-Rubio

Department of Computer Languages and Computer Science, University of Málaga, Bulevar Louis Pasteur, 35, 29071 Málaga, Spain

Abstract

Image segmentation is a common goal in many medical applications, as its use can improve diagnostic capability and outcome prediction. In order to assess the wound healing rate in diabetic foot ulcers, some parameters from the wound area are measured. However, heterogeneity of diabetic skin lesions and the noise present in images captured by digital cameras make wound extraction a difficult task. In this work, a Deep Learning based method for accurate segmentation of wound regions is proposed. In the proposed method, input images are first processed to remove artifacts and then fed into a Convolutional Neural Network (CNN), producing a probability map. Finally, the probability maps are processed to extract the wound region. We also address the

problem of removing some false positives. Experiments show that our method can achieve high performance in terms of segmentation accuracy and Dice index.

I. Introduction

Diabetes-related foot ulcers (DFU) remain a significant source of patient morbidity, poor health-related quality of life, and increased mortality. It is well known that DFU surface area changes [1] and wound tissue structure [2] are highly predictive of subsequent wound closure [3]. However, manual calculation of the surface area is time consuming. Automated techniques can help physicians extracting some characteristics from the wound area. One approach is to use a convolutional neural network (CNN). For instance, CNN is used in [4] for segmentation of psoriasis lesions affecting human skin biopsy.

Images of diabetic skin lesions are more challenging due to color and texture heterogeneity, as well as illumination variations, reflection and presence of noise. Subudhi *et al.* [5] proposed a segmentation method based on k-means clustering to segment diabetic wounds. An automatic tissue identification in wound images was proposed by Veredas *et al.* [6], where color and texture patterns are fed into a set of multilayer perceptrons. Another framework for automatic wound segmentation and healing prediction is proposed in [7] where a kind of auto-encoder architecture called ConvNet is used for segmentation.

In this paper, we propose a CNN-based method for extraction of diabetic lesion regions from skin images. We include the well-known U-Net deep network for image segmentation [8]. We also adapt the CNN architecture of previous works based on melanoma detection [9], [10] to segment diabetic wound lesions. The novelty is that we design a simplified model to process a wide range of diabetic skin wounds, incorporating an adaptive thresholding method to remove artifacts. We also perform a pre-processing step to homogenize the input, hence the images are generated by standard cameras; and include some post-processing operations to remove noisy artifacts and to select the correct wound regions based on mean probabilities of the connected components. To have a better evaluation of the performance of our model, we also design an SVM binary classification model for image segmentation, which has been successfully used in classification and regression problems, and compare the outputs of the CNN models with the SVM model.

The rest of paper is organized as follows. In Section II, the proposed model is presented and both pre-processing and post-processing steps are explained in detail. Comparative results are reported in Section III, where the dataset and the parameter selection procedure are also described. Finally, the conclusions and future works are presented in Section IV.

II. Proposed Model

In this section the wound segmentation model is explained in detail. The schematic diagram of the proposed method is depicted in Fig. 1.

A. Pre-processing

The photographs of the wounds are heterogeneous in terms of illumination, dimension, artifacts and other characteristics. Illumination is one of the main issues that must be considered in pre-processing. Therefore, we first reduce the noise effect of white spots caused by light reflection by the application of a modified version of a mean filter. A threshold, σ , is used to select those pixels with the highest brightness. Next, the kernel $K = [1, 1, 1; 1, 0, 1; 1, 1, 1]$ is generated and convoluted over all the selected pixels: In this way, the white spots are substituted by the mean value from all neighbor pixels, without taking into account the high value of the processed white pixels.

Moreover, color constancy is used to normalize image color to eliminate the influence of different lighting situation [11]. Shades of gray [12] can estimate the color of light source for a color image $\mathbf{I}$,

$$e_c = \frac{1}{k} \left(\frac{\int (\mathbf{I}_c(x))^p d\mathbf{x}}{\int d\mathbf{x}} \right)^{1/p} \quad (1)$$

where $e_c, c \in \{R, G, B\}$, is the component of the color of light source in the c -th channel, $\mathbf{I}_c$ is the component of image $\mathbf{I}$ in the c -th channel, $\mathbf{x}$ is the position of a pixel, and k is the normalization constant to ensure $\mathbf{e} = [e_R, e_G, e_B]$ to be a unit vector. Moreover, p is a norm degree. Then, we can get the coefficient d_c following the von Kries diagonal model [13] to transform the image I as follows:

$$d_c = \frac{1}{\sqrt{3}e_c}, c \in \{R, G, B\} \quad (2)$$

$$\begin{pmatrix} \mathbf{I}'_R \\ \mathbf{I}'_G \\ \mathbf{I}'_B \end{pmatrix} = \begin{pmatrix} d_R & 0 & 0 \\ 0 & d_G & 0 \\ 0 & 0 & d_B \end{pmatrix} \begin{pmatrix} \mathbf{I}_R \\ \mathbf{I}_G \\ \mathbf{I}_B \end{pmatrix} \quad (3)$$

where $[\mathbf{I}_R, \mathbf{I}_G, \mathbf{I}_B]^T$ denotes the intensity of an image acquired in the unknown light source, and $[\mathbf{I}'_R, \mathbf{I}'_G, \mathbf{I}'_B]^T$ denotes the intensity of the transformed image.

B. Convolutional Neural Network

The CNN learns a function f that computes a probability map $\mathbf{Y}$ from an input, such as RGB image $\mathbf{X} \in [0, 255]^{h \times w \times 3}$,

$$f(\mathbf{X}) = \mathbf{Y}, \mathbf{Y} \in [0, 1]^{h \times w} \quad (4)$$

where h and w represent the height and the width of the image, respectively. We also assumed that the range of pixel values is between 0 and 255. For the proposed model described in this section, we use two different CNNs to compute the segmented image. On one side, we use a patch-wise based network [9] and extract a set of local and global patches from the input image. For each patch centered on a determined pixel, f computes

the probability to be wound or non-wound region for this pixel. The local patches would focus on the texture around the central pixel, helping to determine the wound's border. The larger patch will extract more global features from the image. The structure of the network is composed of two paths that are fed by the local and global patches. Each path $b \in \{1, 2\}$ consists of two sets of layers $g_i^{(b)}$ performing a convolution and ReLU operations:

$$g_i^{(b)} = \text{ReLU}(W_i * g_{i-1}^{(b)} + b_i), i \in \{1, 2\} \quad (5)$$

where W_i and b_i are the weights and biases of the i -th layer.

The output of the two paths of layers are connected by a dense layer and then followed by another one with two classes. Finally, a softmax layer is appended to the end to produce a probability map. Specific details of the implementation of this neural network can be found in [9].

On the other hand, the U-Net [8] is a type of fully convolutional network that consists of a contracting path and an expanding path. The contracting path is based on a combination of convolutional and ReLU layers, arriving to multi-channel feature maps, that are reconstructed to the desired number of classes by the expanding path. A combination of 23 convolutional layers and a strong use of data augmentation achieve a very efficient deep network with few data samples.

C. Post-processing

The output of this CNN model is a probability map $\mathbf{Y} \in [0, 1]^{h \times w}$, which shows how likely each pixel belongs to the wound area. The higher the intensity value is, the more likely this pixel belongs to the wound area. The goal of the post-processing is to obtain a binary image, $\tilde{\mathbf{Y}} \in \{0, 1\}^{h \times w}$, that represents a mask segmentation of the source image to remove false positives as many as possible.

1. First, we use the following equation to select those connected components whose corresponding pixels in $\mathbf{Y}$ are greater than a pre-defined threshold, α .

$$S_t = \underset{n=1, \dots, N}{\operatorname{argmax}} \left\{ |\mathbf{c}^{(n)}| < |\mathbf{c}^{(S_t-1)}| \right\}, t = 1, \dots, N' \quad (6)$$

where $\{\mathbf{c}^{(S_t)}, t = 1, \dots, N'\}$ is the set of selected components and $|\cdot|$ stands for the cardinality of a set. We also assume $\{\mathbf{c}^{(n)}, n = 1, \dots, N\}$ is the set of all connected components, and $|\mathbf{c}^{(S_0)}| = h \cdot w$.

2. Secondly, instead of simply keeping the largest connected component as the wound area, the mean value of the probability intensity for each selected component is computed and the connected component with the largest mean value is selected.

$$n_M = \underset{t=1, \dots, N'}{\operatorname{argmax}} \left(\frac{1}{|\mathbf{c}^{(S_t)}|} \sum_{(i,j) \in c^{(S_t)}} \mathbf{Y}(i,j) \right) \quad (7)$$

3. Thirdly, to reduce the false positive selection, two following rules are added.

- The center of the selected component is close to the border of the image.
- The area of the selected component is too small compared with largest connected component:

$$|\mathbf{c}^{(n_M)}| < \tau \cdot \max_{t=1, \dots, N'} (|\mathbf{c}^{(S_t)}|) \quad (8)$$

where $\tau \in (0, 1)$ is a fixed threshold to remove small components, even if the mean probability of the component is very high.

If the selected connected component does not satisfy the rules of step 3, it is discarded. Then, steps 2 and 3 are repeated with the rest of the $N' - 1$ regions until the two rules are satisfied or no components are remaining, and in this case, the component with the highest mean probability is chosen. Finally, closing, dilation, hole filling and erosion morphology operations are applied to refine the selected wound area to get the final result of segmentation.

III. Experimental Results

A. Dataset

The dataset used for the experiments is composed of high resolution camera photos of wound images, provided by New York University (NYU) [7]. These images were previously cropped by the GrabCut tool [14] to remove some background. In total, the dataset contains 445 images. Among these images, 392 images are randomly selected to do the 4-fold cross validation for hyper-parameter selection and model training. Then, the trained model is tested on unseen test set with 53 images. All images have fixed size with 240×320 pixels. Since our method is patch-wise based, 4500 random pairs of local and global patches, respectively with resolutions of 31×31 and 201×201, are evenly selected from wound area, wound border and the background area on each padded image [10]. Patches are fed into the CNN to train the model.

B. Parameter selection

A 4-fold cross-validation is applied to tune the hyper-parameters (i.e. initial learning rate, $init_lr = \{0.001, 0.005, 0.0001, 0.0005, 0.0001\}$, and momentum, $\mu = \{0.5, 0.8, 0.9, 0.95, 0.99\}$) of our CNN.

In our simulations, $init_lr = 0.0001$ and $\mu = 0.99$ give the best performance. The standard deviation of the pixel accuracy rate of 4 test sets is smaller than 0.7%. For the U-Net, we use the default parameters. In pre-processing, $p = 6$ is selected for shades of gray as is

recommended in the work of Barata *et al.* [11]. After getting the probability map generated by the CNNs, the binary threshold $\alpha = 0.5$ is selected. We also selected at most four connected components (i.e. $N' = 4$) and set τ to be 0.125.

C. Evaluation

From a quantitative point of view, we have chosen well-known measures in order to evaluate the performance of the image segmentation. Specifically, we chose the Dice index and Intersection on Union (*IoU*), which are used to evaluating the overlapping extent of the ground truth and the predicted segmentation. These measures provide values in the interval $[0, 1]$, where higher values indicate better performance and represent the percentage of coincidence with respect to the ground truth (GT). Precision, sensitivity, specificity, pixel accuracy and Matthews Correlation Coefficient (*MCC*) are also used in this work. $MCC \in [-1, 1]$ is more informative in evaluating binary classification problems, because it takes into account the balance ratios of the TP, TN, FP and FN. 1 represents a perfect prediction and -1 indicates complete disagreement.

To compare the performance of CNN methods, we use a linear SVM classifier with hinge loss as a baseline, treating the segmentation problem as a pixel-wise binary classification [7]. For each pixel, its intensity value in the RGB channels and those of the surrounding 8 pixels are extracted as features. To balance samples from the wound and remaining areas for training, 1/3 and 1/9 of pixels are randomly selected from the wound and remaining areas, respectively.

D. Results

Fig. 2 presents examples of the qualitative performance of the proposed models. The pre-processing methods reduce misleading factors by removing white spots and adjusting the white balance within the image. The probability map generated by the CNN is successful in eliminating the majority of background regions. Overall, U-Net produces sharper boundaries than those created by the patch-based CNN method. To further refine the segmentation and obtain the predicted mask, we apply post-processing on the raw output of the CNNs. In post-processing, wound selection based on the probability value and the two rules mentioned above work together to reduce the false positive rate, which can improve the performance of wound segmentation. Fig. 2 is an example where the false positive recognition is corrected based on the post-processing.

The CNN models with the selected hyper-parameters were applied on the independent test set. The segmentation results without and with post-processing are shown in Table I. In general, the U-Net model performs better than the patch-based method. Compared to the classical SVM method, the CNN methods achieve higher values for accuracy, sensitivity and specificity, but the most indicative measures are the differences obtained in the Dice index and MCC, which evince a better coincidence with the ground truth.

To further refine the segmentation and obtain the predicted mask, post-processing is applied to the raw output of the CNNs. In our post-processing, instead of simply keeping the connected component with the largest area, wound selection based on the probability value and the two additional rules mentioned above work together to reduce the false positive

rate. The segmentation results with the original post-processing method (keeping the largest component) and our proposed post-processing method are shown separately in Tables I and II. It can be observed that the refined post-processing outperforms the original one. Fig. 3 depicts an inspection of the post-processing procedure to remove false positives. Applying the threshold α to the CNN output, the 4 largest components shown in Fig. 3d and Fig. 3e would be kept. By computing the mean probability values (shown on top of each image) of these components and considering the two additional rules, our methods is able to discard the largest one shown in Fig. 3f, which in this example is a spurious object, and select the correct one shown in Fig. 3e.

IV. Conclusion

In this paper we introduced an automatic segmentation model for diabetic wounds, consisting of pre-processing, CNN segmentation and post-processing. Color constancy and white spot removal were used to reduce the effect of different illuminations. The patch-based CNN model and the U-Net model were trained to create a probability map. A post-processing method based on the probability map was used to refine the segmentation results by using two additional rules and a morphology method. Experiments on the NYU database demonstrated the effectiveness of the U-Net model as well as the patch based model even with a small dataset, but the former one is more computationally efficient and accurate in segmentation. Increasing the number of images in the dataset would most likely improve model performance. The results of the automatic wound segmentation can be used for wound area measurement and feature extraction, which can further aid clinicians in wound diagnosis and prognosis.

ACKNOWLEDGMENT

This work is partially supported by the Ministry of Economy and Competitiveness of Spain under grant TIN2014-53465-R, including funds from ERDF, as well as National Institutes of Health grant 2U24DK076169-11. K.T.H. (FPU15/06512) is funded by a PhD scholarship from the Spanish Ministry of Education, Culture and Sport under the FPU program, and a grant from the University of Málaga.

References

- [1]. Sheehan Peter., "Percent change in wound area of diabetic foot ulcers over a 4-week period is a robust predictor of complete healing in a 12-week prospective trial," *Diabetes care*, vol. 26, no. 6, pp. 1879–1882, 2003. [PubMed: 12766127]
- [2]. Carmen MValenzuela-Silva, "Granulation response and partial wound closure predict healing in clinical trials on advanced diabetes foot ulcers treated with recombinant human epidermal growth factor," *Diabetes Care*, vol. 36, no. 2, pp. 210–215, 2013. [PubMed: 22966096]
- [3]. Kurd Shanu K., "Evaluation of the use of prognostic information for the care of individuals with venous leg ulcers or diabetic neuropathic foot ulcers," *Wound Repair and Regeneration*, vol. 17, no. 3, pp. 318–325, 2009. [PubMed: 19660039]
- [4]. Pal Anabik, Garain Utpal, Chandra Aditi, Chatterjee Raghunath, and Senapati Swapn, "Psoriasis skin biopsy image segmentation using deep convolutional neural network," *Computer Methods and Programs in Biomedicine*, vol. 159, pp. 59 – 69, 2018. [PubMed: 29650319]
- [5]. Subudhi AsitBabu KS and Sabut Sukanta, "Segmentation of diabetic wound by multidimensional clustering for quantitative assessment of healing process," *Current Medical Imaging Reviews*, vol. 14, no. 1, pp. 71–76, 2018.

- [6]. Veredas F, Mesa H, and Morente L, "Binary tissue classification on wound images with neural networks and bayesian classifiers," *IEEE Trans. on Medical Imaging*, vol. 29, no. 2, pp. 410–427, 2 2010. [PubMed: 19825516]
- [7]. Wang C, Yan X, Smith M, Kochhar K, Rubin M, Warren SM, Wrobel J, and Lee H, "A unified framework for automatic wound segmentation and analysis with deep convolutional neural networks," in *Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 8 2015, pp. 2415–2418.
- [8]. Ronneberger O, Fischer P, and Brox T, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. 2015, vol. 9351 of LNCS, pp. 234–241, Springer.
- [9]. Jafari MH, Karimi N, Nasr-Esfahani E, Samavi S, Soroushmehr SMR, Ward K, and Najarian K, "Skin lesion segmentation in clinical images using deep learning," in *Int. Conf. on Pattern Recognition (ICPR)*, 12 2016, pp. 337–342.
- [10]. Jafari M, Hossein, Nasr-Esfahani Ebrahim, Karimi Nader, Soroushmehr S. M. Reza, Samavi Shadrokh, and Najarian Kayvan, "Extraction of skin lesions from non-dermoscopic images for surgical excision of melanoma," *International Journal of Computer Assisted Radiology and Surgery*, vol. 12, no. 6, pp. 1021–1030, 6 2017. [PubMed: 28342106]
- [11]. Barata C, Celebi ME, and Marques JS, "Improving dermoscopy image classification using color constancy," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 3, pp. 1146–1152, 5 2015. [PubMed: 25073179]
- [12]. van de Weijer J, Gevers T, and Gijsenij A, "Edge-based color constancy," *IEEE Trans. on Image Processing*, vol. 16, no. 9, pp. 2207–2214, 9 2007.
- [13]. von Kries J, "Influence of adaptation on the effects produced by luminous stimuli," *Sources of Color Science*, vol. 3, pp. 109–119, 1970.
- [14]. Rother Carsten, Kolmogorov Vladimir, and Blake Andrew, "Grabcut: Interactive foreground extraction using iterated graph cuts," in *ACM Trans. on graphics (TOG)*. ACM, 2004, vol. 23, pp. 309–314.

Fig. 1.
A schematic diagram of the proposed algorithm.

Fig. 2.

Segmentation results for 3 images of the NYU dataset. From left to right, the raw image, three steps of the segmentation model, and ground truth image are shown for comparison (the 3rd and 4th columns are from the patch-based model, while the 5th and 6th columns are from the U-Net model).

Fig. 3.

Example of the post-processing operation for the wound segmentation refinement. (a) Raw image; (b) probability map generated by the CNN model; (c) the ground-truth mask; (d) the largest connected component; (e) post-processed mask by using the proposed post-processing methods.

TABLE I

Quantitative comparison of wound segmentation models (higher is better).

Method	Precision	Sensitivity	Specificity	Pixel Accuracy	Mean IoU	Dice	MCC
SVM(RGB)	0.564	0.806	0.896	0.877	0.472	0.596	0.594
Patch-based CNN	0.646	0.854	0.932	0.913	0.569	0.700	0.675
U-Net	0.768	0.937	0.960	0.949	0.723	0.823	0.840

TABLE II

Quantitative comparison of wound segmentation models with the proposed post-processing (higher is better).

Method	Precision	Sensitivity	Specificity	Pixel Accuracy	Mean IoU	Dice	MCC
Patch-based CNN	0.722	0.9	0.947	0.934	0.660	0.770	0.753
U-Net	0.830	0.917	0.973	0.966	0.761	0.845	0.839