Semantic Segmentation of Diabetic Foot Ulcer Images: Dealing with Small Dataset in DL Approaches

Niri Rania[1(✉)], Hassan Douzi[1], Lucas Yves[2], and Treuillet Sylvie[2]

[1] IRF-SIC Laboratory, Ibn Zohr University, Agadir, Morocco
rania.niri@edu.uiz.ac.ma
[2] PRISME Laboratory, Orléans University, Orléans, France

Abstract. Foot ulceration is the most common complication of diabetes and represents a major health problem all over the world. If these ulcers are not adequately treated in an early stage, they may lead to lower limb amputation. Considering the low-cost and prevalence of smartphones with a high-resolution camera, Diabetic Foot Ulcer (DFU) healing assessment by image analysis became an attractive option to help clinicians for a more accurate and objective management of the ulcer. In this work, we performed DFU segmentation using Deep Learning methods for semantic segmentation. Our aim was to find an accurate fully convolutional neural network suitable to our small database. Three different fully convolutional networks have been tested to perform the ulcer area segmentation. The U-Net network obtained a Dice Similarity Coefficient of 97.25% and an intersection over union index of 94.86%. These preliminary results demonstrate the power of fully convolutional neural networks in diabetic foot ulcer segmentation using a limited number of training samples.

Keywords: Diabetic Foot Ulcer (DFU) · Medical images segmentation · Deep learning · Fully convolutional networks · U-Net · Data augmentation

1 Introduction

Type 2 diabetes mellitus can seriously damage several body's organs over time. Diabetic foot ulceration is the most serious diabetes-related complication [1], and represents a major health problem. In a diabetic patient, this pathology is due to neuropathy, infection or peripheral arterial disease of the lower limb resulting in the formation of skin ulcers and may lead to subsequent lower limb amputation. It is estimated that a lower limb or part of a lower limb is lost somewhere in the world every 30 s as a consequence of diabetes [2]. Diabetic foot ulcers (DFU) have a negative impact on patient's quality of life and result in an important social and economic burden on the public health system. The early prevention and detection of diabetic foot ulcer and the use of an adequate treatment are the key to the management of diabetic foot and to prevent foot amputations.

The management of diabetic foot ulcers requires a prolonged assessment process and patients need frequent clinical evaluation to check regularly their ulcers. Therefore, ulcers monitoring is primordial to improve the healing rate and speed and to select an

A. El Moataz et al. (Eds.): ICISP 2020, LNCS 12119, pp. 162–169, 2020.
https://doi.org/10.1007/978-3-030-51935-3_17

efficient treatment. In current practice, DFU assessment is typically based on visual examination. Clinicians evaluate the wound healing status by manual measurements of the ulcer, including length, width, surface area and volume. There are limited manual methods for the assessment of diabetic foot lesions. Ulcer perimeters (length and width) are measured by using a simple ruler, surface area is approximated by tracing the wound outline on a transparent sheet and volume is obtained by filling the ulcer with a physiological liquid [3]. Moreover, these methods are uncomfortable, sometimes painful and are in direct contact with the ulcer bed which can carry high risk of infection and contamination [4]. In another hand the current traditional methods of wound healing assessment are often rely on the subjective diagnosis of the expert and are time consuming. Therefore, an automatic assessment tool is needed to assist clinicians for a more accurate management and optimal diabetic foot ulcer care.

Nowadays, the use of imaging technology for automatic DFU/wound assessment and measurements has increased considerably to become a common practice. Following this trend, many research works have started to perform wound assessment in medical environment using imaging devices [5]. The major advantages of using digital cameras or smartphones is that photography does not require contact with the skin and it also can help not only to measure the ulcer area but also to analysis the different types of tissue inside the wound bed. DFU assessment will be more objective, accurate and less time consuming for health professionals and patients. Therefore, these methods require the use of image processing and computer vision techniques for image analysis.

Several approaches based on Machine Learning (ML) techniques have been experimented on wound and DFU segmentation [6–9]. Mainly, these works performed the segmentation task using traditional ML algorithms such as SVM classifiers after manual feature extraction. The majority of these methods require the extraction of texture and color descriptors from images such as HOG, SIFT, LBP etc. These descriptors may be affected by light conditions, image resolutions and also skin shape and shades. Thus, the traditional ML methods are not robust due to their reliance on the handcrafted extracted features.

In contrast, deep learning (DL) methods do not require manual feature extraction, as this process is done automatically during the training phase. After the successful impact of these methods on various fields in science during the past few years, nowadays, DL approaches are widely used for many different medical image analysis tasks [10]. Recent approaches applied to wound/DFU use deep neural networks to segment or extract feature descriptors from wound images. Wang et al. [11] proposed a wound segmentation technique in an end-to-end style including wound segmentation and healing progress measurement. The research work is based on an encoder-decoder architecture to perform wound segmentation and was trained on a dataset containing 650 images. Goyal et al. [12] implemented a novel fast DL architecture for DFU classification called DFUNet which classified foot lesions into two classes (normal skin and abnormal one). In this work, they used a database of 397 images. In another work [13], Goyal et al. proposed a two-tier transfer learning method using fully convolutional networks (FCNs) [14] on a larger database of 705 images.

The methods based on Deep learning require massive amount of labelled training data. In biomedical field, images collection is not easily accessible and image annotation is a critical task. Since that the training set should be annotated only by medical experts, which is very time consuming, the training set is small. Our approach provides the use of DL methods for diabetic foot ulcer segmentation using a small database.

This research work is part of the STANDUP project http://www.standupproject.eu/ which aim to prevent diabetic foot ulceration at an early stage and also to help clinicians to monitor the ulcer healing over time using a portable and cost-effective system. This system is composed of a smartphone and a small thermal camera called FlirOne in order to combine color and thermal information for an accurate assessment. Temperature indicators from the thermal image can help to detect tissue infection and inflammation. The final system can be used as a low-cost device to help and assist health care professionals during the examination of DFU by automatically calculating the ulcer area after segmentation and to carry out wound tissue analysis based on color and temperature. Also, it could be used as self-management tool to engage patients with foot ulceration in their own wound care routine. For an optimal DFU diagnosis, this application can be used as telemedicine tool to transmit information about the healing status between clinicians and patients located far from health centers to avoid supplementary costs due to transportation, ambulances, etc.

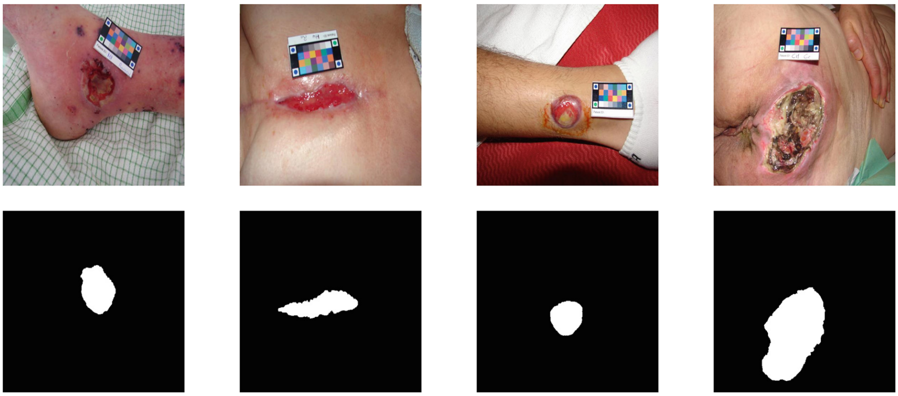

Fig. 1. Sample images from the CW training set illustration of high-resolution CW and the corresponding ground truth masks.

2 Methodology

2.1 DFU Database

In the present work, chronic wounds (CW) images from ESCALE database [15] were used as training set. This database contains 92 high resolution chronic wounds images acquired in several medical centers in France using different commercial digital cameras. A patch was included in the field of view as a scale factor for dimensional measurements and also to estimate the white balance for color calibration [16]. Considering the purpose of this work which is the segmentation of diabetic foot skin lesion, the method was trained on different chronic wounds images and tested only on DFU images. Test data set was obtained from diabetic patients during clinical exams of DFU at CHRO “Centre Hospitalier Regional d’Orleans”, located in Orléans, France. The images were acquired free-handedly using a smartphone camera and without any strict protocol. The whole database has been annotated by experts into two classes i.e. ulcer area and healthy skin. The training set images are in RGB while their ground truth masks are binary (see Fig. 1).

2.2 Training Convolutional Neural Networks

Our objective is to classify DFU images at a pixel level classification into ulcer and non-ulcer classes. We performed the segmentation by using fully convolutional networks. These models are the state of the art of semantic segmentation which allows a pixelwise prediction [14]. The advantage of using FCNs is that the entire image can be directly used as an input to the network. Therefore, a class label will be assigned for each pixel of the image which leads to an efficient segmentation of the ulcer area. In our case, the ulcer area will be represented by white pixels, healthy skin and background by black pixels. Searching for a suitable model to our small database, we focused on U-Net proposed by Ronneberger et al. [17]. This neural network model was specially created to perform semantic segmentation of small datasets. It was designed for biomedical images, and its specific symmetric architecture produces precise segmentation using few images for training. Figure 2 shows the architecture of the U-Net.

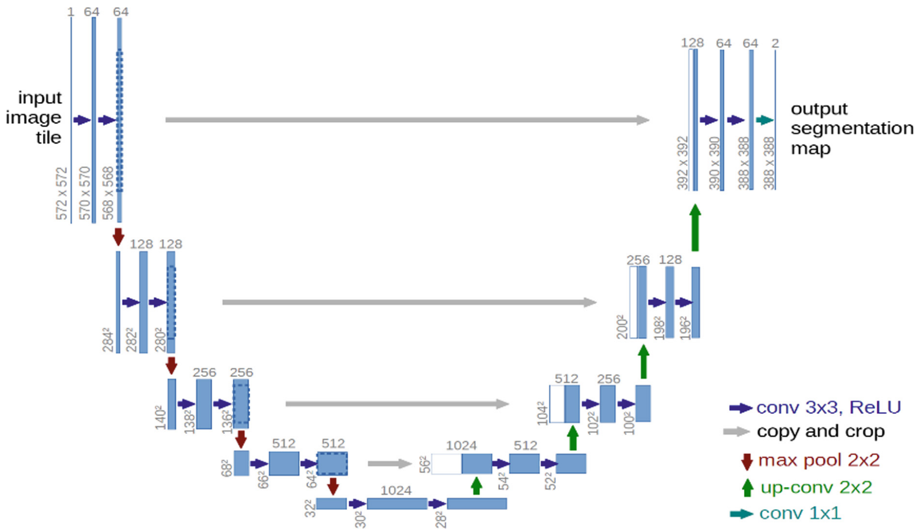

Fig. 2. U-Net architecture based on the paper by Olaf Ronneberger et al.

Since the images were taken in different hospital sites and with different cameras, they had different resolutions. Therefore, all the images were resized to a resolution of 512×512 to perform the U-net training. After that, the database was divided into two sets, resulting in a training set consisting of 92 CW images and validation set consisting of 22 DFU images. To perform our training, we used the Keras framework with TensorFlow backend [18]. The parameters of the kernels were optimized by using the adaptive moment estimation (Adam) with a learning rate set at 0.0001.

Two methods have been chosen to evaluate diabetic foot ulcer segmentation using the proposed approach based on U-Net, which are V-Net [19] and SegNet [20]. As an improvement of U-Net architecture, Milletari et al. proposed an architecture called V-Net to segment prostate MRI volumes. The tested model is the same architecture of the original V-Net but with modified convolution kernel to perform 2D image segmentation. SegNet network is a parallel encoder-decoder based architecture, for every encoder layer there is a corresponding decoder. Unlike U-Net, this network does not contain any skip connections. SegNet uses a pre-trained VGG16 as encoder. We used different image sizes to fit the original input size of each architecture.

2.3 Data Augmentation

The number of samples in our database is limited and not sufficient to feed a deep learning neural network. Therefore, due to the small size of our database, data augmentation is required to increase the size of training set and to avoid overfitting [21]. Different techniques have been used to randomly deform the input image and correspondent ground truth segmentation map. Each image was horizontally flipped, rotated, translated and zoomed (see Fig. 3). After data augmentation, the database size has been increased by four and contains now 368 images instead of 92 images.

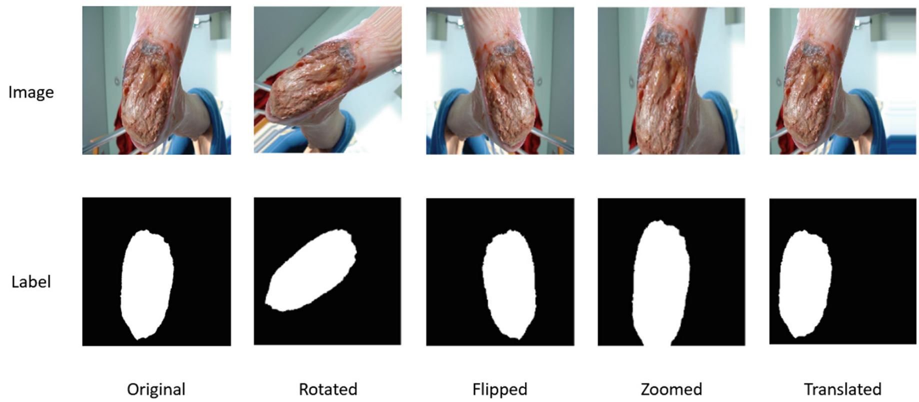

Fig. 3. An example of data augmentation results.

3 Results

3.1 Evaluation Metrics

In order to evaluate the segmentation results quantitatively, we choose *Intersection Over Union index (IoU)*, and *Dice Similarity Coefficient (DSC)* [22]. They are the most used metrics in the field of medical image segmentation. The DSC coefficient assesses the similarity between the predicted segmented ulcer region bounded by the three deep learning networks and manually labeled ground truth given by experts. These two-evaluation metrics are defined respectively by (Eq. 1) and (Eq. 2):

$$IoU = \frac{TP}{TP + FP + FN} \tag{1}$$

$$DSC = \frac{2TP}{2TP + FP + FN} \tag{2}$$

where TP stands for the true positive samples, FP for the false positive samples and FN for the false negative ones.

3.2 Experimental Results

As mentioned before, the purpose of our research work was to segment DFU images into ulcer and non-ulcer classes. Due to computational resources limitations and memory constraint, the U-net model was trained for just 5 epochs. The network did surprisingly well, after 5 epochs the calculated accuracy was about 95%. Overall, the qualitative results shown in Fig. 4 reveal the power of the network to generate a mask similar to the ground truth. As we can clearly notice that after overlaying the predicted mask on the original image of the ulcer, the network segment correctly the ulcer area with a high precision.

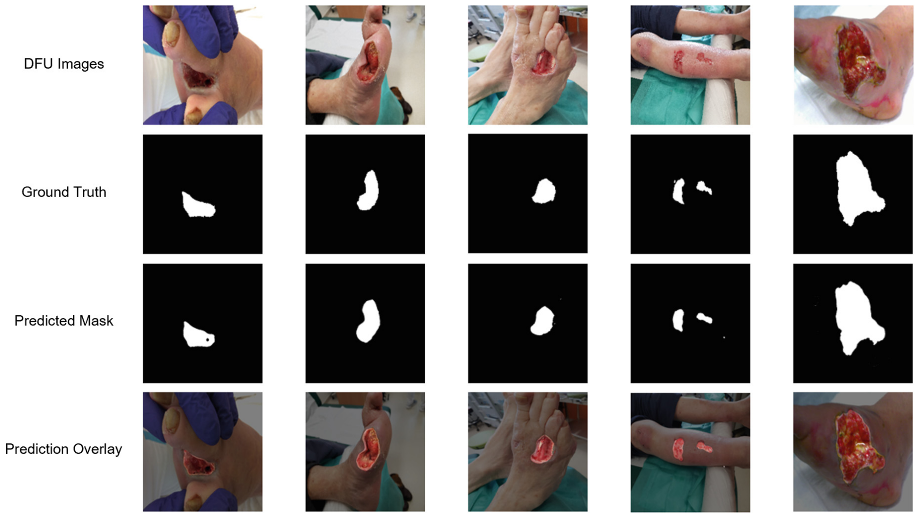

Fig. 4. An illustration of the predicted segmentation map of the proposed network on some images from our DFU validation set.

The quantitative evaluation of the segmentation results was performed on the 22 images of our DFU validation database. Table 1 compares the results of these three methods using test accuracy, mean IoU score and mean DSC coefficient. These results show that U-Net and V-Net methods give both good results compared to the SegNet network. We clearly notice that the best segmentation results are given by the U-Net architecture.

Table 1. DFU Segmentation results. Test Accuracy, Intersection over Union (IoU) and Dice coefficient (DSC) in %.

Method	Accuracy	IoU	DSC
U-Net	**94.96**	**94.86**	**97.25**
V-Net	92.52	92.17	95.74
SegNet	88.94	87.87	93.45

4 Conclusion

To summarize, in this preliminary work our objective was to find an accurate deep learning architecture for diabetic foot ulcer segmentation using a limited number of samples, as data collection and expert labeling are very expensive in the field of medical imaging. The current training database contains only 92 images of different chronic wounds images including foot pathologies. Different methods were tested and evaluated using 22 DFU images. The results show that the U-Net network produces the best DFU segmentation results with a DSC coefficient of 97% after only five epochs of training. After the ulcer segmentation, the next step will be to perform DFU tissue classification using deep learning to help clinicians for an objective analysis of the different tissues withing the ulcer bed. As future work, we aim to develop a mobile system for diabetic foot ulcer evaluation and assessment using images acquired through a smartphone and an add-on thermal camera for a more accurate diagnosis of DFU over time. Moreover, we attempt to perform a 3D reconstruction model of the ulcer for a more precise and robust surface measurements of the wound area.

Acknowledgment. This research work is supported by internal funding sources of PRISME laboratory, Orleans, France and from the European Union's Horizon 2020 under the Marie Sklodowska-Curie grant agreement No 777661.

References

1. American diabetes association: diagnosis and classification of diabetes mellitus. Diabetes Care Suppl 1, 81–90 (2014)
2. International Diabetes Federation: IDF Diabetes Atlas, 8th edn. International Diabetes Federation, Brussels (2017)
3. Humbert, P., Meaume, S., Gharbi, T.: Wound healing assessment. Phlebolymphology **47**, 312–319 (2014)

4. Keast, D.H., et al.: Measure: a proposed assessment framework for developing best practice recommendations for wound assessment. Wound Repair Regenerat. **12**(1), 1–17 (2004)
5. Mohafez, H., Ahmad, S.A., Roohi, S.A., Hadizadeh, M.: Wound healing assessment using digital photography: a review. JBEIM, 3(5) (2016). ISSN: 2055-1266
6. Kolesnik, M., Fexa, A.: Segmentation of wounds in the combined color-texture feature space. In: Medical Imaging 2004: Image Processing, vol. 5370, pp. 549–556. International Society for Optics and Photonics (2004). https://doi.org/10.1117/12.535041
7. Kolesnik, M., Fexa, A.: How robust is the SVM wound segmentation. In: Proceedings of the 7th Nordic Signal Processing Symposium, pp. 50–53 (2006)
8. Wannous, H., Lucas, Y., Treuillet, S.: Efficient SVM classifier based on color and texture region features for wound tissue images. In: Medical Imaging 2008: Computer-Aided Diagnosis, vol. 6915, p. 69152T. International Society for Optics and Photonics (2008). https://doi.org/10.1117/12.770339
9. Wang, L., et al.: Area determination of diabetic foot ulcer images using a cascaded two-stage SVM-based classification. IEEE Trans. Biomed. Eng. **64**(9), 2098–2109 (2017)
10. Litjens, G., et al.: A survey on deep learning in medical image analysis. Med. Image Anal. **42**, 60–88 (2017)
11. Wang, C., et al.: A unified framework for automatic wound segmentation and analysis with deep convolutional neural networks. In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (2015)
12. Goyal, M., Reeves, N.D., Davison, A.K., Rajbhandari, S., Spragg, J., Yap, M.H.: DFUNet: convolutional neural networks for diabetic foot ulcer classification. In: IEEE Transactions on Emerging Topics in Computational Intelligence, pp. 1–12. IEEE (2018)
13. Goyal, M., Yap, M.H., Reeves, N.D., Rajbhandari, S., Spragg, J.: Fully convolutional networks for diabetic foot ulcer segmentation. In: 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (2017)
14. Shelhamer, E., Long, J., Darrell, T.: Fully convolutional networks for semantic segmentation. In: IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 4 (2017)
15. Lucas, Y., Treuillet, S., Albouy, B., Wannous, H., Pichaud, J.C.: 3D and color wound assessment using a simple digital camera. In: 9th Meeting of the European Pressure Ulcer Advisory Panel, Berlin (2006)
16. Wannous, H., et al.: Improving color correction across camera and illumination changes by contextual sample selection. J. Electron. Imag. **21**(2), 023015 (2012)
17. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
18. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. CoRR, vol. abs/1412.6980 (2014)
19. Milletari, F., Navab, N., Ahmadi, S.A.: V-net: fully convolutional neural networks for volumetric medical image segmentation. In: International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT) (2016)
20. Badrinarayanan, V., Kendall, A., Cipolla, R.: Segnet: a deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans. Pattern Anal. **39**(12), 2481–2495 (2017)
21. Perez, L., Wang, J.: The Effectiveness of Data Augmentation in Image Classification using Deep Learning (2017). arXiv preprint arXiv:1712.04621
22. Lee, R.: Dice: measures of the amount of ecologic association between species. Ecology **26** (3), 297–302 (1945)