Skin prick test wheal detection in 3D images via convolutional neural networks

Conference Paper · October 2021

DOI: 10.1109/CI-IBBI54220.2021.9626125

3 authors, including:

Juan C. Pena
Universidad Tecnológica de Bolívar
1 PUBLICATION **0** CITATIONS
SEE PROFILE

Andrés G Marrugo
Universidad Tecnológica de Bolívar
79 PUBLICATIONS **335** CITATIONS
SEE PROFILE

Skin prick test wheal detection in 3D images via convolutional neural networks

Juan C. Peña, Jose A. Pacheco, Andrés G. Marrugo
Facultad de Ingeniería, Universidad Tecnológica de Bolívar, Cartagena, Colombia.
juancpb0727@gmail.com
josepachecoserna@gmail.com
agmarrugo@utb.edu.co

***Abstract*—The skin prick test (SPT) is performed to diagnose different types of allergies. This medical procedure requires measuring the size of the skin wheals that appear when the test is performed. However, the manual measurement method is cumbersome and suffers from intra- and inter-observer errors. Thus, multiple approaches have been developed to improve the reproducibility of the test. This work aims to improve part of the automated reading of the SPT to improve the reliability of the wheal detection procedure through the use of convolutional neural networks (CNN). Our proposal starts from the 3D images of the SPT from the arm of patients. They are processed for global surface removal, and then a CNN is trained to produce an output mask that detects the wheals. Finally, the contour of each wheal and its largest diameter is obtained. Encouraging results with mean difference 0.966 *mm* and mean coefficient of variation 7.29% show that the proposed method provides reliable automated skin wheal detection.**

***Index Terms*—SPT, skin prick test, wheal, 3D image, convolutional neural network.**

I. Introduction

The skin prick test is one of the most performed tests on patients to identify allergic reactions to various substances. The key aspect of the test is the size of the wheal that appears after the skin is pricked. The physician measures the diameter of the wheal to determine the degree of sensitization to an allergen [1]. The reading of the test has been performed using different devices and techniques, making evaluation and comparison between different studies challenging [2]. Therefore, the Global Allergy and Asthma European Network determined a standard guideline for performing the test, including aspects such as the distance between two pricks ($\geq$ 2 cm) and the minimum diameter of the wheal to be considered a positive reaction (3 mm) [3]. The wheals are often measured by considering them as ellipses and measured along their longest diameter [4].

Conventionally, the measurement of wheals has been performed directly on the patient's forearm, using a millimeter ruler as shown in Fig. 1. Note that often the wheals have irregular shapes, which impede a reliable manual measurement. As a result, this Method suffers from intra- and inter-observer errors [5], [6].

Other authors have attempted to tackle this problem with conventional 2D digital imaging of the region of interest. However, despite sophisticated image processing software, the results are not entirely reliable [7] or have problems with skin tone [8]. Alternatively, Pineda et al. [9] recently proposed a robust method using a 3D imaging system and measuring the wheal size automatically. This method is based on pyramidal surface decomposition and Principal Component Analysis (PCA) to detect the wheals and perform parametric fitting to calculate their diameter. Nevertheless, the automated detection stage requires lots of parameter tuning that may not work well in all circumstances.

We propose to replace the wheal detection stage with a deep-learning-based approach. Specifically, we use a U-Net architecture for the automated segmentation and diameter estimation of the SPT wheals [10]. The U-Net architecture has been successfully used in many medical domains for reliable segmentation. The network used can segment multiple skin wheals in a single pass.

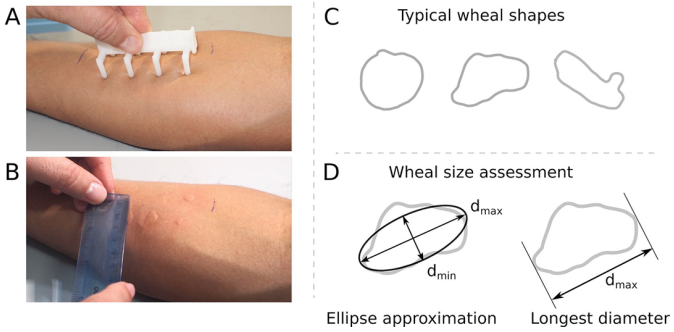

Fig. 1: **(A)** SPT is performed on the patient's forearm. **(B)** After 15 minutes, the diameter of the wheals are measured with a ruler. **(C)** Common wheal shapes. **(D)** The wheals are approximated as ellipses for measuring their diameter.

978-1-6654-0855-4/21/$31.00 ©2021 IEEE

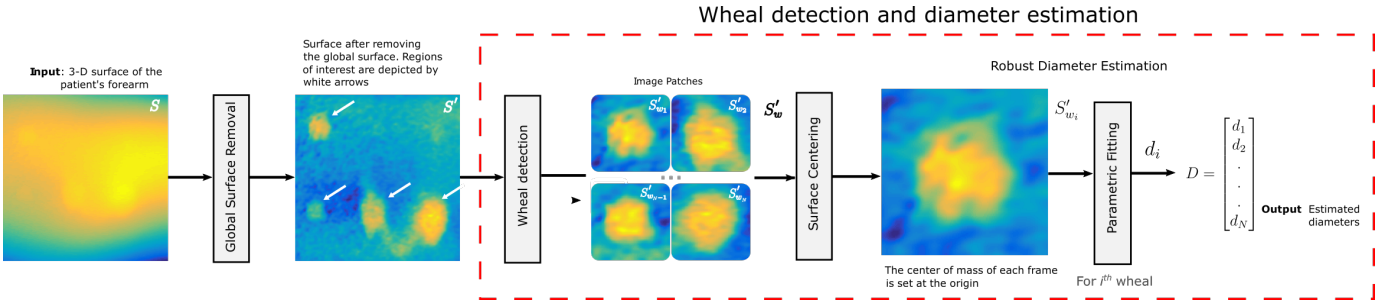

Fig. 2: The stages of the classical method replaced by our proposal are enclosed in the red box.

II. Method

We cast the problem of the detection of wheals obtained from a 3D reconstruction of the SPT as a binary classification problem. The method is based on training a CNN for wheal detection and then determining their longest diameter by a parametric approximation due to the irregular shape of the wheals.

The method proposed by Pineda et al. [9] consisted of multiple stages. First, the wheals are detected via multi-scale filtering with Laplacian of Gaussians, second decomposing the image into patches, and third determining their centers of mass using PCA. Here, we propose to simplify this multi-stage approach using a U-Net style CNN that directly outputs the detected wheals.

Fig. 2 shows a block diagram of the conventional method. The dashed-line block shows the stage we propose to replace with a CNN. Specifically, the input is the 3D representation of the arm, which goes through a global surface removal stage. S' is the output of this stage, where white arrows indicate the wheals. Moreover, the preprocessed surface is the input to our proposed network.

We use these images because they are less likely to have noise or information unrelated to the wheals. These are converted to intensity images with an accompanying mask of the detected wheals. This procedure was carried out with the conventional method.

A. Image acquisition

The images used to train and validate the neural network are the same from Pineda et al. [9], in which the fringe projection system shown in Fig. 3 is used, and the 3D images of the forearm are like the one shown in Fig. 4.

B. Pre-processing

For the CNN input, images are constructed from 3D depth information, the data is handled as one channel depth image. All images were zero-padded to have all training data of the same size. Also, all depth images were normalized from range [0, 255] to range [0, 1].

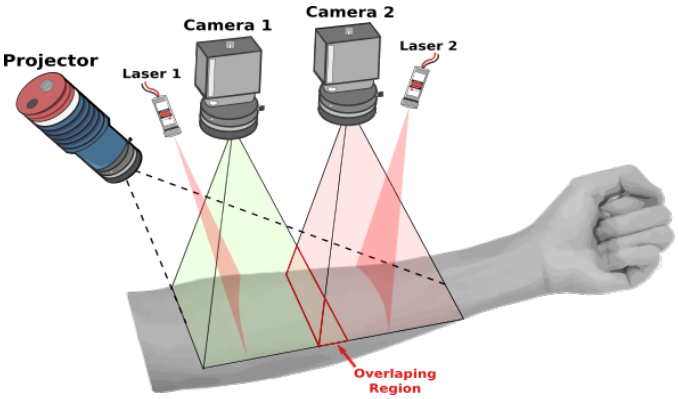

Fig. 3: Fringe projection system, consisting of 2 cameras, 2 lasers and 1 projector, capable of reconstructing an area of 150 mm x 250 mm [9].

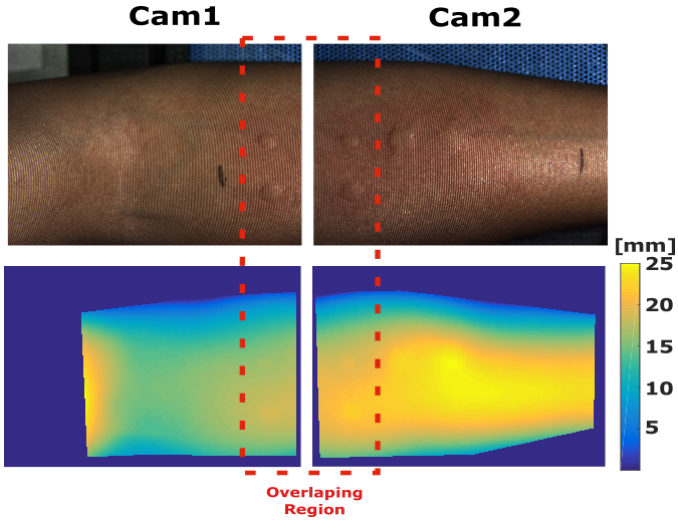

Fig. 4: Reconstruction 3D of the arm. The area marked by the red box indicates the region where both observation systems overlap [9].

C. Data augmentation

We performed data augmentation to improve the robustness of the model. It was carried out through the Keras image data preprocessing API [11]. We obtained an increased number of images performing random

operations such as horizontal and vertical flip, rotation (range from 0° to 180°), cropping (range from 0% to 20%) and zoom (range from 0% to 20%).

In total, 14 images were acquired for data training before data augmentation, and a total of 47 wheals were identified. After data augmentation, a total of 84 training images were obtained. The number of wheals after data augmentation is variable due to random geometric transformations; however, it is well above 200.

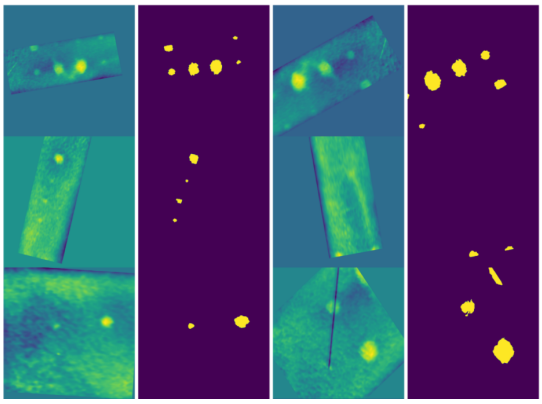

Fig. 5: Transformed images and masks for data augmentation.

Fig. 5 shows an example of the transformed images for data augmentation, where the first and third columns are entry images with random transformation arguments and the second and fourth columns are the corresponding masks with the same transformation arguments. The first row has two transformed images from the same original image, with different transformation arguments.

D. CNN architecture

The implemented network has a U-Net type architecture due to the efficiency obtained in other medical segmentation tasks [10].

Fig. 6 shows a diagram of the used CNN. The network has 18 convolutional layers distributed in 9 main blocks. The decoder is consists of 4 blocks, each one conformed by two repeats of 3×3 convolution layers, followed by batch normalization and ReLU. A 2×2 max-pooling with stride 2 follows each decoding block for downsampling, where the number of filters is doubled. The skip connection consists of two repetitions of 3×3 convolutional layer, followed by batch normalization and ReLU. The encoder consists of 4 block, each one conformed by a 2×2 transposed convolution layer, followed by batch normalization, ReLU, a concatenation with corresponding feature map from decoding blocks, and two repetitions of 3×3 convolutional layers, followed by batch normalization and ReLU. In each upsampling, the number of filters is halved. The number of filters start at 64, in the bridge block are 1024 and finish at 64.

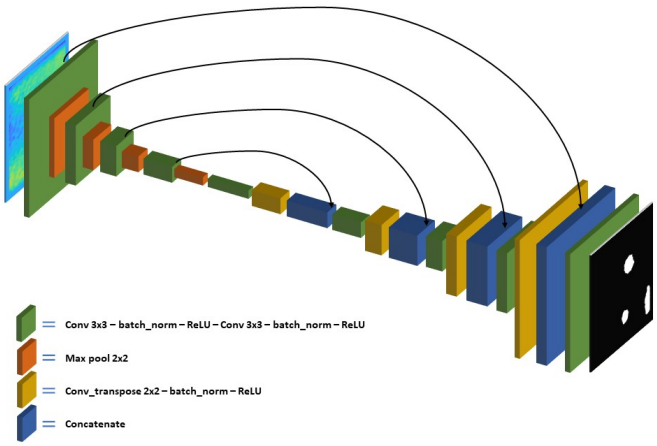

Fig. 6: Visual representation of the CNN architecture.

III. Results and Discussion

The CNN was trained during 150 epochs, with an Adam loss function. The training dataset contains 14 images, and the testing dataset contains 4 images.

The results obtained with CNN used in the predictions, evidence a high performance even when the entry data quantity is not as large as desired.

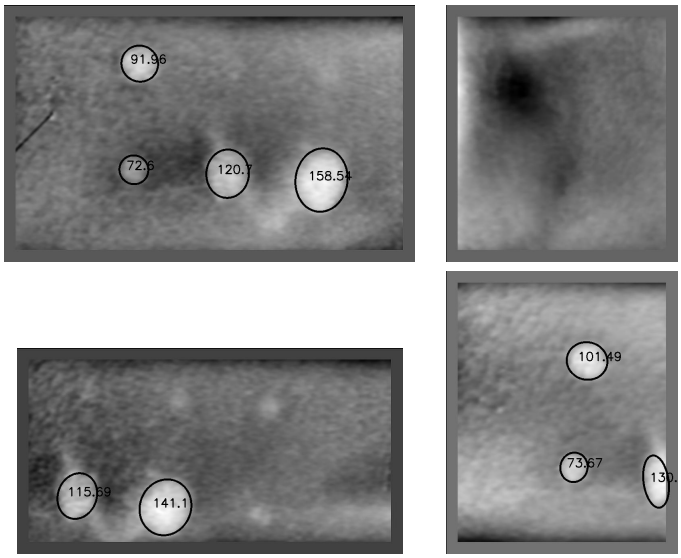

Fig. 7: CNN output (in pixels) from the untransformed images with aproximated ellipses drawn.

Output images are shown in Fig. 7. These results show good predictions on the images. First, the image without wheals shows how the CNN can differentiate pixels that are not part of a wheal. Note that large wheals are detected sufficiently well, whereas not all smaller wheals are detected. Moreover, the wheals are measured in pixels, so it is necessary to obtain their value in millimeters.

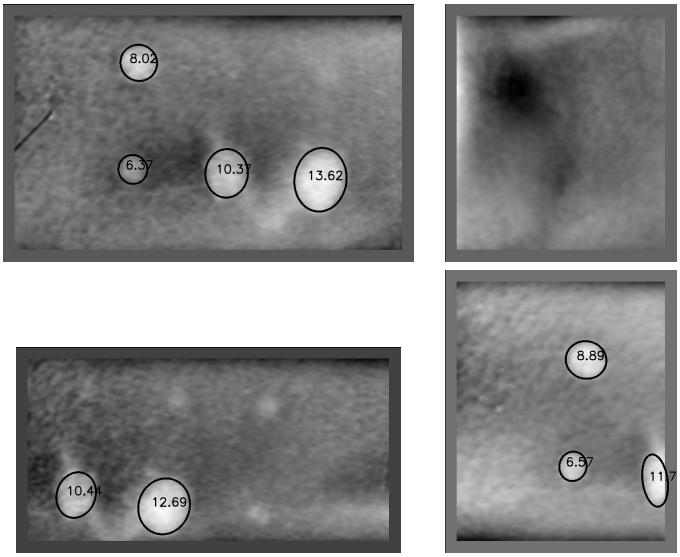

Fig. 8: CNN output (in millimeters) from the untransformed images with aproximated ellipses drawn.

In Fig. 8, we show the diameter of the wheals in millimeters. For this process, we obtain the coordinates (X,Y) of the extreme points of the ellipse and compute its Euclidean distance (in millimeters) in the reconstruction 3D. Images show that most wheals are detected and measured reliably well despite minor problems in the segmentation for very small wheals.

Wheal	Reference Measurement	CNN	Difference	CV (%)
1	7.554	8.020	0.466	4.230
2	5.661	6.370	0.709	8.340
3	8.173	10.370	2.197	16.760
4	11.003	13.620	2.617	15.030
5	10.147	10.440	0.293	2.020
6	12.007	12.690	0.683	3.910
7	4.253	-	-	-
8	4.647	-	-	-
9	5.341	-	-	-
10	8.726	8.890	0.164	1.320
11	-	11.700	-	-
12	5.975	6.570	0.595	6.710
Mean			0.966	7.29

TABLE I: Measurements record (in millimeters) with difference and coefficient of variation (CV) as performance indicators. The symbol (-) represents wheals not detected by the network, wheals not accepted by the physician and indicators not calculated.

In table I, the reference measurements obtained by averaging three measurements from the physician on a computer screen and the proposed CNN are shown, with their difference and coefficient of variation (CV). Note that overall performance is satisfactory with mean difference 0.966 *mm* and mean CV 7.29%. However, there are three cases (2, 3 and 4) that have CVs of 8%, 16% and 15%. In these cases, the presence of pseudopods or extensions of the wheals makes it difficult to agree with the real measurement, and in those large wheals, there is no doubt about their positivity.

IV. Conclusions

We proposed a wheal detection and diameter estimation algorithm based on a U-Net convolutional neural network. Despite the small amount of data, the network successfully detected the wheals in the test dataset. We believe that the preprocessing stage for removing the global shape of the arm gives a reliable 3D surface which greatly facilitates the detection. Further validation studies are required to thoroughly test the robustness of the proposed approach.

Acknowledgement

This work has been partly funded by Universidad Tecnologica de Bolivar (UTB) and MinCiencias (contract 935-2019). J. A. Pacheco thanks the CEIBA Foundation, and J. Peña thanks the Ser Pilo Paga program for an undergraduate scholarship. We also thank Jesús Pineda and Jhacson Meza for their advice and for guiding us through the data. Finally, we thank all the people who encouraged us to carry out this project.

References

[1] H. Pijnenborg, L. Nilsson, and S. Dreborg, "Estimation of skin prick test reactions with a scanning program," *Allergy*, vol. 51, no. 11, pp. 782–788, 1996.

[2] B. Buyuktiryaki, U. Sahiner, E. Karabulut, O. Cavkaytar, A. Tuncer, and B. Sekerel, "Optimizing the use of a skin prick test device on children," *International archives of allergy and immunology*, vol. 162, pp. 65–70, 06 2013.

[3] L. Heinzerling, A. Mari, K. Bergmann, M. Bresciani, G. Burbach, U. Darsow, S. Durham, W. Fokkens, M. Gjomarkaj, T. Haahtela, A. Todo Bom, S. Wöhrl, H. Maibach, and R. Lockey, "The skin prick test – european standards," *Clinical and translational allergy*, vol. 3, p. 3, 02 2013.

[4] G. Konstantinou, P. J. Bousquet, T. Zuberbier, and N. Papadopoulos, "The longest wheal diameter is the optimal measurement for the evaluation of skin prick tests," *International archives of allergy and immunology*, vol. 151, pp. 343–5, 10 2009.

[5] S. Wöhrl, K. Vigl, M. Binder, G. Stingl, and M. Prinz, "Automated measurement of skin prick tests: an advance towards exact calculation of wheal size.," *Experimental Dermatology*, vol. 15, pp. 119–124, Feb. 2006.

[6] X. Justo, I. Díaz, J. Gil, and G. Gastaminza, "Prick test: evolution towards automated reading," *Allergy*, vol. 71, no. 8, pp. 1095–1102, 2016.

[7] W. McCann and D. Ownby, "The reproducibility of allergy skin test scoring and interpretation by board-certified/board-eligible allergists," *Annals of allergy, asthma & immunology*, vol. 89, pp. 368–71, 11 2002.

[8] O. Bulan, "Improved wheal detection from skin prick test images," *Proc. SPIE*, vol. 9024, pp. 138 – 147, 2014.

[9] J. Pineda, R. Vargas, L. A. Romero, J. Marrugo, J. Meneses, and A. G. Marrugo, "Robust automated reading of the skin prick test via 3d imaging and parametric surface fitting," *PLOS ONE*, vol. 14, pp. 1–18, 10 2019.

[10] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241, Springer, 2015.

[11] K. Team, "Keras documentation: Image data preprocessing," 2021.