

A graph-based superpixel segmentation method for measuring pressure ulcers

(Accepted – journal – Latim America)



A graph-based superpixel segmentation method for measuring pressure ulcers

CITATION DETAILS:

Felipe Moreira de Assunção, Rodolfo Herman Lara e Silva, Alexei Manso Correa Machado, Paulo Sérgio Silva Rodrigues, Zenilton K. G. Patrocínio Jr., Silvio Jamil Ferzoli Guimarães (2023). A graph-based superpixel segmentation method for measuring pressure ulcers. Technical Report ImScience/PUC Minas #01/2023.

A graph-based superpixel segmentation method for measuring pressure ulcers

Felipe Moreira de Assunção, Rodolfo Herman Lara e Silva, Alexei Manso Correa Machado, Paulo Sérgio Silva Rodrigues, Zenilton K. G. Patrocínio Jr., Silvio Jamil Ferzoli Guimarães

June 8, 2023

Abstract Monitoring wound healing is a necessary procedure to help health services control pressure ulcers. The correct diagnosis depends on clinical observations by doctors and nurses during patient visits. The evaluation of the wound area represents one of the most important data. Usually, health professionals assess ulcers through visual inspection, using rulers and decals. These ones, in direct contact with these lesions, may cause discomfort and inducing other infections, and consequently, worsen the patient's clinical condition. Understanding and knowing these injuries allows for better preventive and therapeutic actions. In this paper, we aim to present an automatic and effective method for ulcer delineation according to the following pipeline: (i) graph-based superpixel segmentation; (ii) superpixel feature extraction; (iii) superpixel classification; (iv) ulcer segmentation; and (v) feature description. The main idea is to automatically compute pressure ulcer measurements for identifying the lesion area, allowing the follow-up of the scar evolution. Our graph-based superpixel segmentation method outperformed five other state-of-the-art approaches, as well as deep learning models, reaching 92.6% sensitivity, 98.6% specificity, 97.6% precision, 96.6% accuracy, and 90.4% intersection over the union.

Keywords Superpixel · graphs · pressure ulcers · health

1 INTRODUCTION

One of the complications that arise from prolonged hospitalization is skin lesions that can progress to pressure ulcers [26, 32, 37]. Pressure ulcers are chronic injuries resulting from excessive compression of soft tissue against bony

prominences and intricate surfaces or medical devices. Patients with ulcers have extended lengths of stay, favoring new illnesses that generate physical, emotional, and financial wear for the public/private health systems [36]. Moreover, chronic injuries are among the leading causes of increased morbidity and mortality in patients with chronic-degenerative diseases and the elderly [35]. The skin, when injured, soon starts the healing process through a dynamic, continuous, complex, and interdependent process, composed of a series of overlapping phases [17]. The use of rulers and decals in direct contact with ulcer lesions during a visual inspection by health professionals can cause discomfort to the patient, as well as increase the risk of infection and worsen the patient's clinical condition. This strategy is often imprecise, subjective, and without a well-defined standard among health professionals [30, 31]. It also lacks a standardized method of assessment, leading to inconsistencies in diagnosis and treatment. Furthermore, the reliance on visual inspection alone may not provide a comprehensive understanding of the ulcer, such as its depth and size, which can impede accurate diagnosis and treatment planning. Lesion healing can be better observed using monitoring scales through the development of computational methods that make the clinical assessment through digital images, as they provide more accurate and reliable information [31].

This paper presents an effective method for ulcer delineation that uses the Dynamic and Iterative Spanning Forest (DISF) method as its central core. We hypothesize that more effective superpixel delineation method may highlight the area of the lesion making it possible the better understanding the behavior of the lesion. For computing the pressure ulcer measurements, we have used a pipeline composed of: (i) image acquisition; (ii) pre-processing and superpixel segmentation; (iii) feature extraction; (iv) data preparation; (v) classification; and (vi) post-processing, as shown in Fig. 1. An additional contribution of this work

Silvio Jamil F. Guimarães
ImScience/PUC-Minas – Belo Horizonte 31980-110, Brazil
E-mail: sjamil@pucminas.br

is a comparison of the effectiveness of five other superpixel methods in the pre-processing step of the pipeline.

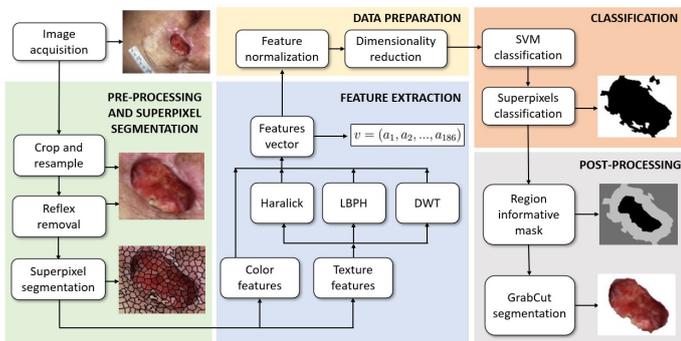


Fig. 1: The proposed methodology for measuring pressure ulcers is based on superpixel computation. (Similar to Silva and Machado [31]).

This work is organized as follows. Section 2 briefly describes some related works. Section 3 describes materials and methods for automatic pressure ulcer segmentation. Section 4 presents and discusses the experimental results, while Section 5 draws conclusions and future works.

2 RELATED WORKS

Veredas et al. [34] proposed a method based on k -means clustering for image segmentation with a comparison of three different machine learning approaches to classify each segmented region as the appropriate tissue type. The results reached a reliability index of around 95%. Deng et al. [9] presented a new pipeline for the automatic extraction of flaky corneal ulcer areas. First, a semi-automatic method identifies the cornea for each image. Then, a four-step approach segments the ulcer region within the cornea as follows: (i) identify/adjust the color information of reflective areas; (ii) use Simple Linear Iterative Clustering (SLIC) to segment each image into 1000 superpixels; (iii) employ Support Vector Machine (SVM) to classify superpixels into two classes; and (iv) *smooth* the ulcer segmentation results through erosion and dilation. They used 150 clinical images. Results demonstrated that the method significantly surpassed two classic segmentation algorithms: active contour and Otsu threshold.

More recently, deep learning models such as Convolutional Neural Networks (CNN) and Fully Connected Deep Networks (FCDN) have been applied to the problem of ulcer assessment. Lu et al. [22] proposed a CNN model and a color correction method to segment non-homogeneous medical images. When applied to 300 lesion images, their approach reached an average accuracy of 83%. Liu et al. [21]

introduced a lightweight neural network to perform wound segmentation. Training and test procedures used a more extensive set of 950 images. After a post-processing step, the best model achieved 98% accuracy, 91% sensitivity, and 93% precision. A minimal dataset of 22 high-resolution images of pressure ulcers was used by Zahia et al. [38] with a CNN model to classify tissue types, reaching a global average of 92% accuracy. The efficiency of DNNs in the segmentation task was improved by Li et al. [18] through background removal techniques, reaching 95% precision and an intersection over union (IoU) of 86%. Blanco et al. [6] proposed a new method for wound segmentation that uses an annotated set of dermatological ulcers to train deep learning models for identifying superpixels that represent ulcerated skin. Goyal et al. [13] presented a set of techniques to recognize the presence of infection and ischemia in diabetic foot ulcers using computerized methods. Their method uses a new feature descriptor with an ensemble of CNNs to recognize and identify the region of interest in foot images and focus on finding the salient features in this area. Silva and Machado [31] used an SVM classifier combined with a modified version of the GrabCut segmentation method for measuring the area affected by pressure ulcers. The primary motivation for their work was related to the difficulty of manually monitoring cases of pressure ulcers. Generally, lesions are evaluated by measuring the affected area with materials that include adhesive labels and rulers in direct contact with the lesion, leading to high inaccuracy due to the degree of subjectivity of the process. Three region segmentation methods were evaluated using a superpixel strategy, and the results were used to extract color and texture descriptors. After, the GrabCut method was applied to delineate the ulcer-affected region from the rest of the image. The experimental evaluation using the MEDETEC dataset (with 105 images of pressure ulcers) showed that the association of SVMs with superpixel segmentation surpassed current methods based on deep learning.

Reis et al. [28] proposed a deep CNN named InSiNet to detect benign and malignant lesions. Scebba et al. [29] presented the Detect-and-Segment (DS), a deep-learning approach to producing wound segmentation maps with high generalization capabilities. For this approach, dedicated deep neural networks detected the wound position, isolated the wound from the perturbing background, and computed a wound segmentation map. Eldem et al. [10] discussed and analyzed pressure wound segmentation using different encoder-decoder-based segmentation models.

Most new superpixel segmentation methods adopt a three-step pipeline: (i) initial seed sampling; (ii) superpixel delineation; and (iii) seed recomputation. But an inconvenience of that approach is the need to limit the size of the initial seed set to the desired number of superpixels that could hinder finding *relevant seeds* that are able to generate an

accurate object edge delineation. Dynamic Iterative Spanning Forest (DISF) [4] is a method for computing superpixel with a dynamic arc-weight estimation. DISF method demonstrates to better preserve relevant object edges, especially for lower numbers of superpixels, compared to the state-of-the-art approaches.

3 MATERIALS AND METHODS

3.1 Image acquisition

We have used images from MEDETEC dataset [23], a public repository of open wounds such as venous leg ulcers, arterial leg ulcers, and pressure ulcers, among others. The ulcer dataset is composed of 105 images of pressure ulcers with a resolution of 410×560 pixels.

3.2 Pre-processing and superpixel segmentation

For the pre-processing step, specular light reflections [16] were removed from the original image. These are bright white-colored regions characterized by high intensity, saturation, and contrast. The resulting images were resized to 240×300 pixels. Finally, the removal and reconstruction of the light reflection regions were performed using an inpainting method [33]. The method takes the binary image as parameters, which informs the regions to be reconstructed, and the size of the radius of neighbors to be considered for the reconstruction, a radius of size $r = 3$ was used. The algorithm fills the light reflection regions with neighboring pixels, recovering the colors hidden by the reflection. An expert manually annotated the pressure ulcers into ulcer and non-ulcer regions. In the following, a superpixel segmentation is done.

Superpixels are groups of connected pixels that share similar characteristics according to a predicate [4]. The superpixel methods used in this work are described as follows:

1. Entropy Rate Superpixels (ERS) method [20] is a graph-based algorithm formulated as a graph maximization problem. ERS uses the entropy rate of a random walk on a graph as a criterion to generate high-quality superpixels.
2. Simple Linear Iterative Clustering (SLIC) method [1] generates superpixels by clustering pixels based on their color similarity and proximity in the image plane;
3. Superpixels Extracted via Energy-Driven Sampling (SEEDS) method [5] is based on a hill-climbing optimization with efficient exchanges of pixels between superpixels. The energy function that is maximized is based on enforcing homogeneity of the color distribution within superpixels;

4. Linear Spectral Clustering (LSC) method [19] produces compact and uniform superpixels with high memory efficiency, and it is able to preserve the global properties of images with low computational costs;
5. Simple Non-Iterative Clustering (SNIC) method [2] is an improved version of the SLIC algorithm. SNIC is non-iterative, enforces connectivity from the start, requires less memory since the distance map is not considered, and is faster;
6. Dynamic and Iterative Spanning Forest (DISF) method [4] finds more relevant seeds, reconstructs relevant edges along with iterations, and guarantees the desired number of superpixels. DISF assures optimal spanning forests for path costs based on dynamic arc-weight estimation, being faster as the desired number of superpixels grows.

Figure 2 presents superpixel segmentation examples for DISF, ERS, SNIC, LSC, SLIC, and SEEDS.

3.3 Feature extraction

After segmentation into superpixels, color and texture features were extracted for the RGB , $L^*a^*b^*$, Luv , and normalized RGB color spaces. For color descriptors, the values of mean, variance, asymmetry, frequency and intensity of the two highest peaks in the histograms were extracted for each component of the color space after the application of a moving average filter. The centroids of each superpixel in each color space were used, totaling 96 color features. Regarding the texture ones, the features were extracted based on co-occurrence matrices [3], Local Binary Pattern Histograms (LBPH) [25], and Haar Wavelet Transform [8]. We used the Haralick descriptors of contrast, correlation, energy, entropy, and homogeneity, determined at the angles of 0° , 45° , 90° and 135° at a distance of one pixel. We extract the features for each of the components R , G , B , a^* , b^* , u and v . Five more descriptors were extracted and computed from the combination of RG , RB , GB , a^* , b^* , and uv . A total of 60 texture features were extracted, considering Haralick and its variations. For the LBPH, the mean, variance, entropy, and energy were computed for the histogram location of each superpixel in each of the color space components, totaling 12 features. Finally, the Discrete Wavelet Transform (DWT) was used to extract the energy and entropy descriptors. For this method, three sub-images contain the details in horizontal, vertical, and diagonal orientation in the third wavelet decomposition level for each RGB component, totaling nine sub-images. Each sub-image calculated energy and entropy descriptors for each superpixel, resulting in 18 features. In the end, a total of 186 descriptors were extracted from each superpixel.

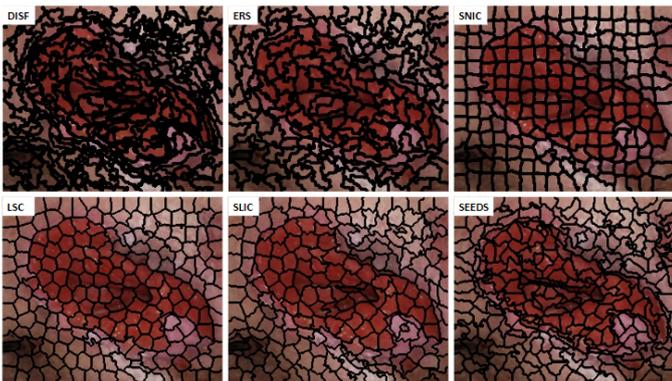


Fig. 2: Superpixels segmentation examples for DISF, ERS, SNIC, LSC, SLIC and SEEDS

3.4 Data preparation

The Wrapper algorithm [14], in conjunction with a decision tree [27], was used to classify the features for selecting a subset of relevant features. However, the Hill-Climbing searching algorithm [24], optimized by cross-validation [15], could select the most relevant attributes for the classification task. The feature was normalized in the range from 0 to 1.

3.5 Classification

For classification purposes, the 105 images were randomly separated into training and test sets. The training and validation sets were composed of 84 images, equivalent to 80% of the dataset, while the test set consisted of 21 images. An SVM with an RBF kernel was used since it has shown to be effective in medical image classification tasks [34,35]. Parameters C and γ were defined using the grid search and the cross-validation technique to find the best parameters. The parameter C is the regularization parameter that controls the trade-off between the margin and the misclassification of training examples. A small value of C creates a wider margin with more misclassified training examples, while a large value of C creates a narrow margin with fewer misclassified examples. The parameter γ is specific to the RBF kernel and controls the shape of the decision boundary. A small value of γ creates a smoother decision boundary, while a large value creates a more complex decision boundary. For the grid search, we adjust the space of search in $2^{-10} \rightarrow 2^{10}$ with the steps of 0.1. During the search, we adopted a 10-fold cross-validation.

3.6 Post-processing

The output classification was transformed into a mask to be used in the interactive segmentation made by the GrabCut method. Instead of scribbles drawn by users, an automatic

mask was computed according to the following steps: (i) applying the morphological dilation filter with a 3×3 cross-type structuring element to the SVM classification result to fill small discontinuities found between any two misclassified superpixels; (ii) using the Canny algorithm [12] to detect contours, where the region of the image that is covered by the most prominent closed contour receives the label of probable ulcer while the label of non-ulcer is assigned to the region around it; (iii) applying a skeletonization process to the closed contour [39], until 65% of its area is eroded (the branches of the images are removed by erosion); and (iv) the remaining region is labeled as an ulcer. Thus, the main idea is to determine the region inside the ulcer to be used as a mask. This mask is used for the GrabCut method to perform the segmentation, pointing out the regions of ulcers.

3.7 Evaluation measures

For measuring the quality of the ulcer identification, we have used the following measures: (i) Sensibility (S); (ii) Specificity (E); (iii) Precision (P); (iv) Accuracy (A); and (v) Intersection over Union (I). The running time (TE) was also computed.

4 EXPERIMENTAL RESULTS

4.1 Configuration settings

The main application was developed in Java 11, the classification in Weka 3.8 [11], data analysis in Python, while the segmentation methods varied between Java using OpenCV library [7] (SLIC, LSC, SEEDS, GrabCut), C (DISF), and Matlab (SNIC and ERS). The experiments were performed on a desktop with the following settings: Intel Core i7-8700 CPU at 3.20 GHz with 16 GB and Nvidia GeForce RTX 2060 GPU running Windows 10 (64 bits).

In the pre-processing step, we removed and reconstructed specular light reflections as shown in Figure 3 according to the following steps. First, the original image is converted

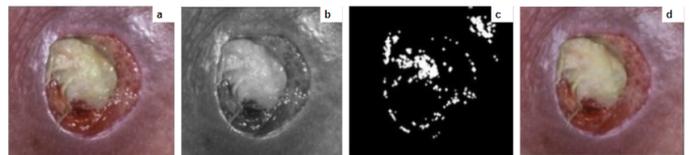


Fig. 3: Process of removal and reconstruction of light reflection regions. (a) Original image; (b) Grayscale image; (c) Binary image with regions of light reflection in white color; (d) Result of the image after applying the inpainting method and the average filter.

into a grayscale image, which is binarized for showing the light reflection in white color. An inpainting is applied to the binary image for reconstructing the original image without the light reflections. For smoothing the image, an average filter is then applied. As previously related, the set of 105 images was randomly divided into training and test sets. The set used for training (and cross-validation) was composed of 84 images.

The number of iterations was set to 400 for LSC, SLIC, and SEEDS since no further changes were observed in the formation of superpixels. The parameter related to the compactness factor is associated with the different superpixel shapes that affect the smoothness of the edges between tissue types in the ulcer bed and between the skin and ulcer regions. Considering the edges between different tissue regions, for each superpixel segmentation method, the compactness factor was set between 0.05 and 0.065 for the LSC, with the best value being 0.06. For SLIC, the parameter varied from 10 to 25; the best value was 20. For SEEDS, we varied the parameter from 1 to 4; the best value was 3. For SNIC, the parameter was in the range of [18, 21], and the best value was 20. The compactness values were based on previously published suggestions, as described in the original articles of the respective methods.

DISF has the parameter α that controls the trade-off between the importance of individual pixels and the spatial coherence of the superpixels. By adjusting the value of α , DISF can balance the trade-off between these two objectives and generate superpixels that are both visually meaningful and spatially coherent. The parameter related to the number of superpixels controls the size of the superpixels generation and is related to the image resolution. Thus, appropriately adjusting this parameter made a significant impact on the SVM classifier performance. Moreover, seeds are designed to compete with each other and conquer the most closely connected nodes, mathematically defining each superpixel as an optimal path tree. For the DISF method, we set the number of seeds to be proportional to the size of 5% of the total image size. So, for an image of 240×300 , the number of seeds is set up to 3,600. For the ERS method, we set the balancing term λ' to 0.5, the kernel bandwidth σ to 5.0, and we adopted an 8-connectivity graph, as suggested by the authors. The parameters setup for the superpixel segmentation step are synthesized in Table 1. The number of generated superpixels is also presented.

After the superpixel generation, each one is described by color and texture features. For the classification, the SVM classifies the region as ulcer or non-ulcer. A trained nurse determined the ground truth from manual segmentation. For the superpixel labeling, only the superpixels containing pixels with a single tissue type were incorporated into the training set. Figure 4 illustrates the superpixel labeling process using DISF, in which the ulcer underwent manual

Table 1: Parameters setup for the superpixel segmentation for DISF, ERS, SNIC, LSC, SLIC, and SEEDS. The number of generated superpixels is also shown.

	DISF	ERS	SNIC	LSC	SLIC	SEEDS
Iterations	–	–	–	400	400	400
Superpixels Size	200	200	200	18	19	300
# Seeds	3600	–	–	–	–	–
Compactness	–	–	20	0.06	20	3
λ'	–	0.5	–	–	–	–
σ	–	5	–	–	–	–
Connectiveness	–	1	–	–	–	–
# Generated Superpixels	200	200	208	208	≈ 208	~ 237

segmentation in (b) and superpixel segmentation in (c). In the end, the superpixels were then mapped to the manual labels to create a superpixel grid in (d), with superpixels in white representing the ulcer, the ones in light gray corresponding to healthy skin, and the ones in dark gray containing labeled pixels from both classes and were therefore not included in the SVM training.

In data preparation, the number of training instances and selected attributes for DISF, ERS, SNIC, LSC, SLIC, and SEEDS are summarized in Table 2, while Table 3 shows the parameters for the training using SVM for each method.

Table 2: Number of training instances and selected attributes for DISF, ERS, SNIC, LSC, SLIC, and SEEDS.

	DISF	ERS	SNIC	LSC	SLIC	SEEDS
# Training Instances	7964	7540	7960	7508	7726	8336
# Selected Attributes	10	11	18	14	10	10

Table 3: Parameters for training step with SVM for DISF, ERS, SNIC, LSC, SLIC, and SEEDS.

	DISF	ERS	SNIC	LSC	SLIC	SEEDS
% of Folds	10	10	10	10	10	10
C	7943.28	39.81	1.59	7.94	15848.93	1.58

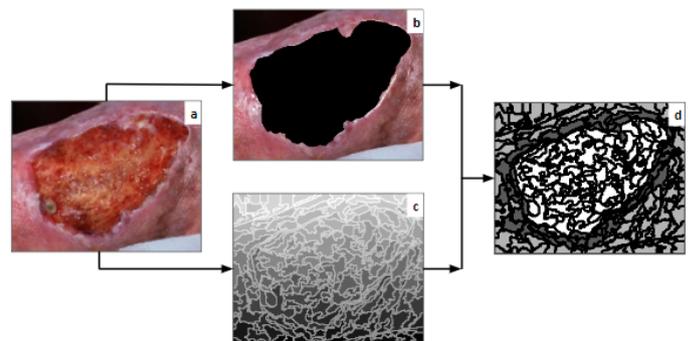


Fig. 4: Superpixel labeling. (a) original image; (b) manual segmentation; (c) superpixels; (d) superpixel labeling: light gray color represents regions of healthy skin, white the regions of ulcer, and dark gray the excluded superpixels composed of both tissues.

4.2 Quantitative Analysis

An SVM classifier was designed for each of the described superpixel methods. Table 4 presents the results during the training step for SVM classifier using DISF, ERS, SNIC, LSC, SLIC, and SEEDS. The performance achieved for each method during the test step, with and without post-processing by the GrabCut, is shown at Table 5.

We compare results in pairs of superpixels to assess possible discrepancies between the metrics results. A two-tailed paired t-test (Student’s t-test) was used in this comparative approach, considering the values obtained from the SVM + Grabcut technique. The DISF confidence intervals for a level of 95% are $S = [88.7, 96.5]$, $E = [97.6, 99.6]$, $P = [96.3, 98.8]$, $A = [95.5, 97.7]$, $I = [86.8, 94.0]$. Among the results, a significant difference was found between the results of Sensitivity (S), Accuracy (A), and Intersection over Union (I) of some pairs of superpixels. In the other metrics, no significant differences were observed considering the value of $\alpha = 0.05$. When evaluating all possible combinations and statistical tests, the classifier trained with the superpixels formed by DISF showed better performance in all metrics used, except processing time (11.0s), slightly above SLIC (9.6s), ERS (10.5s) and SNIC (10.7s), but with a small difference. It is important to note that the execution time of a method in a given programming language can vary depending on several factors, including the nature of the method, the efficiency of the implementation, and the processing power of the machine on which the code is executed. In general, the C language tends to perform faster than most other languages, but other languages may have advantages in specific situations.

4.3 Qualitative Analysis

Figures 5 and 6 show segmentation results with their respective masks and the overlap between the manual segmentation and those generated by GrabCut, for DISF, ERS, SNIC, LSC, SLIC, and SEED methods.

Figure 5 shows some examples of segmentation with the worst results for each method. One can observe a poor creation of the masks generated from the classification results with the SVM. Among the methods, the SNIC presented the worst segmentation among the examples presented. In

Table 4: Results of accuracy during the training and validation for each SVM classifier using DISF, ERS, SNIC, LSC, SLIC, and SEEDS. The time (in seconds) to build each model is presented.

	DISF	ERS	SNIC	LSC	SLIC	SEEDS
Training Acc.	85.71%	85.85%	88.17%	88.64%	86.67%	88.09%
Validation Acc.	85.86%	86.25%	88.37%	86.85%	86.80%	88.36%
Time (s)	13425.3	3288.11	4197.16	787.02	15744.03	3696.8

Table 5: Average \pm standard deviation of Sensitivity (S), Specificity (E), Precision (P), Accuracy (A), Intersection over Union (I), and Execution Time (T_E) in seconds for DISF, ERS, SNIC, LSC, SLIC, and SEEDS.

	SVM	T_E (s)	SVM+Grab	T_E (s)
DISF	S	91.8\pm8.7	92.6\pm8.6	
	E	85.8 \pm 7.9	98.6\pm2.2	
	P	79.4 \pm 9.4	10.6 \pm 1.9	11.0 \pm 2.1
	A	88.3\pm4.2	96.6\pm2.4	
	I	73.8\pm9.4	90.4\pm8.0	
ERS	S	79.6 \pm 14.4	82.3 \pm 16.2	
	E	88.1 \pm 7.7	98.5 \pm 1.9	
	P	80.3\pm1.0	10.1 \pm 1.9	10.5 \pm 1.9
	A	85.6 \pm 6.3	92.9 \pm 5.4	
	I	66.6 \pm 14.0	80.2 \pm 15.7	
SNIC	S	55.9 \pm 14.1	48.5 \pm 25.8	
	E	88.8\pm7.4	98.6 \pm 2.2	
	P	75.2 \pm 14.5	10.3 \pm 1.6	10.7 \pm 1.7
	A	77.1 \pm 7.2	80.9 \pm 10.3	
	I	47.2 \pm 13.3	47.6 \pm 24.9	
LSC	S	89.8 \pm 8.5	90.5 \pm 10.1	
	E	81.7 \pm 10.0	97.6 \pm 3.8	
	P	75.5 \pm 8.7	15.21 \pm 1.8	15.56 \pm 1.8
	A	85.3 \pm 5.4	95.3 \pm 3.4	
	I	69.2 \pm 8.6	86.9 \pm 10.7	
SLIC	S	88.3 \pm 10.1	92.2 \pm 10.1	
	E	82.8 \pm 9.9	98.3 \pm 2.0	
	P	76.5 \pm 11.2	9.3 \pm 4.6	9.6 \pm 0.5
	A	86.1 \pm 6.4	95.9 \pm 1.8	
	I	69.3 \pm 11.3	89.8 \pm 10.4	
SEEDS	S	86.5 \pm 9.4	86.6 \pm 10.8	
	E	81.1 \pm 10.4	95.2 \pm 3.3	
	P	73.7 \pm 11.2	13.5 \pm 2.2	13.5 \pm 2.2
	A	83.6 \pm 4.0	96.4 \pm 2.1	
	I	66.0 \pm 11.5	89.1 \pm 11.1	

Fig. 5(q), a large region with non-ulcer labels can be seen inside the wound’s lesion. Furthermore, it was observed that the sensitivity, specificity, and precision values were much lower when compared to the other methods. This demonstrates the incredible difficulty of detecting ulcer and non-ulcer regions in these examples. In general, more significant variations in sensitivity were observed, whose values varied from 12% to 66%. This may be related to the selection of color and texture attributes. Perhaps the selected superpixels had a low number of attributes of color and texture characteristics thus hindering the training of the classifier and, therefore, leading to misclassification.

The skeletonization step of the method contributes to generating a mask that inaccurately maps the region that does not present an ulcer, causing this region to be inside the perimeter of the wound. Likewise, we observed the prevalence of the mask that indicates non-ulcer regions within the ulcer area in Figures 5(b), 5(e), 5(h), 5(k), 5(n), and 5(q). This occurrence erroneously indicates that GrabCut should delete these regions. Despite that, some regions labeled as non-ulcer and probably ulcer within the ulcer area were correctly segmented with an ulcer in the GrabCut results in Figure 5(c), 5(f), 5(i), 5(l), 5(o), and 5(r),

with emphasis on the last one, in which part of the region previously considered as an ulcer outside the perimeter of the ulcer in Figure 5(q), is correctly detected as non-ulcer.

In contrast, Figure 6 illustrates the most effective segmentation for each method. Considering the excellent result of the classification generated by the SVM, one can observe that the regions of probable ulcer in the masks of Figs. 6(f), 6(i), 6(l), 6(o), and 6(r) follow the shape of the ulcer and, therefore, present a segmentation closer to manual segmentation performed by the nurse. Figure 7 shows the segmentation results with the trained SVM and the post-processing results using the GrabCut along with the use of DISF method. In the first row, we have the original image. In the second row, one can observe segmentation results with the SVM. It can be noticed that the method had more difficulties in classifying the non-ulcer regions (lower specificity) than the ulcer regions (higher specificity). Segmentation results from post-processing with GrabCut are shown in the third row. From the quantitative point-of-view, it is possible to observe a substantial improvement, showing that the post-processing step could practically altogether remove the false-positive regions and include a large number of false-negative regions, which impact on the effectiveness of the segmentation.

4.4 Discussion

4.4.1 Performance evaluation

Evaluating the performance of an automatic ulcer delineation method against a manual segmentation is a challenging task. To address this, a multi-faceted approach is suggested. Firstly, quantitative metrics such as sensibility (S), specificity (E), precision (P), accuracy (A), and intersection over union (I) can be used to compare the overlap between two segmentations. Secondly, a qualitative visual inspection can be conducted to identify any errors or inconsistencies in the segmentations. Thirdly, consulting a medical expert may provide valuable insights into the accuracy of the algorithm's output. Lastly, statistical tests, such as t-tests may be performed to determine the significance of any differences between the segmentations. A comprehensive evaluation methodology, as presented in this paper, combining these different approaches, can provide a robust assessment of the performance of the automatic ulcer delineation technique.

4.4.2 Limitations

For segmentation and delineation of pressure ulcers, major limitations are their variation in appearance and shape, as well as the presence of noise and artifacts in the images. In addition, the presence of other types of skin lesions can

make specific detection of pressure ulcers more difficult. Another limitation is the need to segment the wound area into different layers, such as necrotic tissue, granulation, and epithelialization, which can be an additional challenge.

4.4.3 Benchmark

In order to provide a comparative approach, we highlight the related works and the results found, as shown in Table 6.

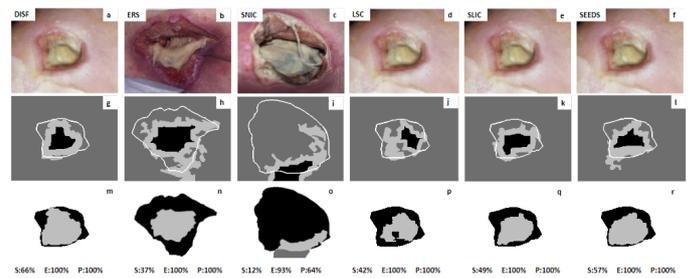


Fig. 5: Examples of ineffective segmentation for DISF, ERS, SNIC, LSC, SLIC, and SEEDS methods. Original images (row 1), masks (row 2) and overlay (row 3) between manual segmentation and DISF methods (column 1), ERS (column 2), SNIC (column 3), LSC (column 4), SLIC (column 5), and SEEDS (column 6). Masks indicate dark gray non-ulcer regions, probably light gray and black ulcers. It also shows the shape of the wound traced in white. Overlapping images indicate regions of false negatives in black, false positives in dark gray, and true positives in light gray. Sensitivity (S), specificity (E), and precision (P) are provided.

Regarding the state-of-the-art, including methods based on deep learning, no method outperformed the proposed method in more than one metric. A lot of papers use a restricted number of metrics to evaluate the proposed method. Regarding sensitivity, the work by Reis et al. [28] presents better results, followed by the work of Blanco et al. [6] and Silva and Machado [31]. In Reis et al. [28], sensitivity is 4.9% higher, however, the proposed method exceeds specificity by 7.4% and accuracy by 2%. In Silva and Machado [31], sensitivity is 1.1% higher, however, the proposed method exceeds specificity by 1.7%, precision by 3.2%, accuracy by 0.6% and intersection over union by 1.4%. Regarding accuracy, Eldem et al [10] presents better results, being 3.1% better than the proposed method, followed by Liu et al. [21], being 1.6% higher than the proposed method. However, the proposed method outperforms by 2%, 4.3% and 5.8% the results presented in [21], in terms of sensitivity, precision, and intersection over union, respectively. In the context of medical image analysis, it is considered important that the results can be explained, that may not always occurs for

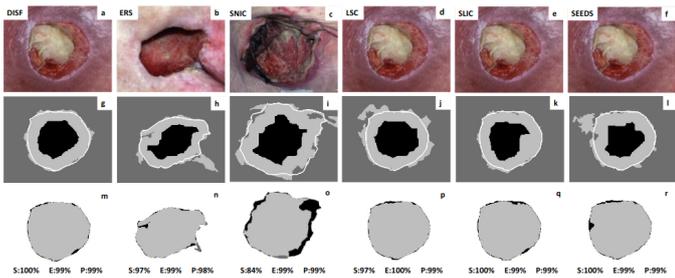


Fig. 6: Examples of effective segmentation. Original images (row 1), masks (row 2), and overlay (row 3) between manual segmentation and DISF (column 1), ERS (column 2), SNIC (column 3), LSC (column 4), SLIC (column 5), and SEEDS (column 6). Masks indicate dark gray non-ulcer regions, probably light gray and black ulcers. It also shows the shape of the wound traced in white. Overlapping images indicate regions of false negatives in black, false positives in dark gray, and true positives in light gray. Sensitivity (S), specificity (E), and precision (P) are provided.

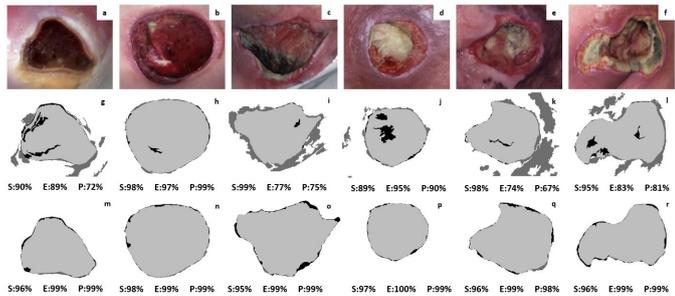


Fig. 7: Example of the effectiveness of GrabCut segmentation using DISF. The original images are shown on the first row, the SVM results on the second row and the GrabCut results on the last row. Images indicate regions of false negatives in black, false positives in dark gray, and true positives in light gray. Sensitivity (S), Specificity (E) and Precision (P) are also provided.

deep learning methods, made clearer, and that the behavior of the methods can be predicted in terms of performance.

It is worth mentioning previous works that use the same methodology and database, the ones developed by Silva and Machado [30, 31]. DISF showed better results for the automatic measurement of pressure ulcers with regard to specificity, precision, and intersection over union when associated with the post-processing technique with GrabCut, according to Table 7.

5 CONCLUSION

This work presented a graph-based superpixel segmentation method for segmenting and identifying pressure ulcer. Based on graph and superpixel representation, it improved

Table 6: Performance metrics reported by methods in the literature for digital image segmentation of chronic lesions. ‡ Number of training + validation (test) images. Studies that used only public datasets are marked with P. *Test based on 10-fold cross-validation. **Test based on 5-fold cross-validation. † The number of test images was not informed. Methods: J-value Segmentation (J-SEG); Support Vector Machines (SVM); Neural Networks (NN); Convolutional Neural Networks (CNN or ConvNet); Fully Convolutional Network (FCN); Conditional Random Field (CRF); Deep Neural Networks (DNN); Convolutional Networks combined (MobileNet-UNet); Simple Linear Iterative Clustering (SLIC). Metrics: Sensitivity (S); Specificity (E); Precision (P); Precision (A), Intersection over union (I).

Reference	Approach	Dataset‡	S	E	P	A	I
Veredas et al. 2009	Mean shift, growing region, NN and Bayesian Classifiers	113*	78.7	94.7	91.5	-	-
Wannous et al. 2010	J-SEG and SVM	25†	77.0	92.0	-	-	-
Veredas et al. 2015	K-means and SVM	90(23)	-	-	88.1	-	-
Wang et al. 2015	CNN	500(150)	-	-	-	-	47.3
Wang et al. 2017	SLIC, SVM and CRF	100*	73.3	94.6	-	-	-
Liu et al. 2017	CNN and CRF	760(190)	90.6	-	93.3	98.2	84.6
Liu et al. 2017	CNN	300†P	-	-	-	-	83.0
Dhane et al. 2017	Clustering Spectral Fuzzy	70(70)	87.3	95.7	-	91.5	79.0
Silva et al. 2018	Filtering and Morphologic operations	110(110)P	81.6	-	90.8	81.3	-
Zahia et al. 2018	CNN	18(4)	-	-	-	92.0	-
Li et al. 2018	DNN and Filtering	760(190)	-	-	94.7	-	86.3
Li et al. 2019	DNN and Filtering	760(190)	-	-	95.3	-	86.5
Blanco et al. 2019	QTDU, CNN	217†P*	97.0	97.4	-	-	-
Goyal et al. 2020	CNN and henceforth TML	7136 (2038)**	88.6	92.1	91.8	90.3	-
Silva et al. 2021	SEEDS, SVM and GrabCut	84(21)P	93.7	96.9	94.4	96.0	89.0
Niri et al. 2021	FCN	164(55)	-	-	-	-	92.9
Reis et al. 2022	CNN	9514(501)P	97.5	91.2	-	94.6	-
Scabbia et al. 2022	UNet	84(21)P**	-	-	-	-	83.0
Eldem et al. 2022	MobileNet-UNet	90(15)P**	-	-	-	-	99.7
Proposed	DISF, SVM and GrabCut	84(21)P*	92.6	98.6	97.6	96.6	90.4

Table 7: Comparison of performance among works on MEDTEC dataset. Metrics: Sensitivity (S); Specificity (E); Accuracy (P); Accuracy (A); Intersection over Union (I).

Method	S	E	P	A	I
Filtering [30]	81.6	-	90.8	81.3	-
SEEDS [31]	93.7	96.9	94.4	96.0	89.0
DISF	92.6	98.6	97.6	96.6	90.4

the results of pressure ulcer assessment when compared to the state-of-the-art. The proposed method combines classical, unsupervised filters, clustering, feature extraction, dimensionality reduction, and supervised machine learning algorithms. In particular, the superpixel method was compared to state-of-the-art superpixel methods, such as the ERS, SNIC, LSC, SLIC, and SEEDS methods. The results obtained by DISF were better for the automatic measurement of pressure ulcers concerning specificity, precision, and intersection over union.

Further work may consider new interactive segmentation methods. Also, the possibility of applying the methodology in other more extensive databases and for different types of chronic lesions must be examined. For this purpose, ensembles that combine shallow and deep learning techniques are promising as they may take advantage of different perspectives.

References

1. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süsstrunk, S.: Slic superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **34**(11), 2274–2282 (2012). DOI 10.1109/TPAMI.2012.120
2. Achanta, R., Süsstrunk, S.: Superpixels and polygons using simple non-iterative clustering. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017)
3. Arvis, V., Debain, C., Berducat, M., Benassi, A.: Generalization of the cooccurrence matrix for colour images: Application to colour texture classification. *Image Analysis & Stereology* **23**, 63–72 (2011)
4. Belem, F.C., Guimaraes, S.J.F., Falcao, A.X.: Superpixel segmentation using dynamic and iterative spanning forest. *IEEE Signal Processing Letters* **27**, 1440–1444 (2020). DOI 10.1109/lsp.2020.3015433. URL <http://dx.doi.org/10.1109/LSP.2020.3015433>
5. den Bergh, M.V., Boix, X., Roig, G., Gool, L.V.: Seeds: Superpixels extracted via energy-driven sampling. *arXiv* (2013)
6. Blanco, G., Traina, A.J., Traina Jr., C., Azevedo-Marques, P.M., Jorge, A.E., de Oliveira, D., Bedo, M.V.: A superpixel-driven deep learning approach for the analysis of dermatological wounds. *Computer Methods and Programs in Biomedicine* **183**, 105,079 (2020). DOI 10.1016/j.cmpb.2019.105079
7. Bradski, G.: *The OpenCV Library*. Dr. Dobb's Journal of Software Tools (2000)
8. Daubechies, I.: Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics* **41**(7), 909–996 (1988). DOI <https://doi.org/10.1002/cpa.3160410705>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.3160410705>
9. Deng, L., Huang, H., Yuan, J., Tang, X.: Superpixel based automatic segmentation of corneal ulcers from ocular staining images. In: *2018 IEEE 23rd International Conference on Digital Signal Processing (DSP)*, pp. 1–5 (2018). DOI 10.1109/ICDSP.2018.8631795
10. Eldem, H., Ülker, E., Yaşar Işıklı, O.: Encoder–decoder semantic segmentation models for pressure wound images. *Imaging Science Journal* **70**(2), 75 – 86 (2022). DOI 10.1080/13682199.2022.2163531. URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85146679831&doi=10.1080%2f13682199.2022.2163531&partnerID=40&md5=18716b8f8bea5bdf989d82453afd6ec3>. Cited by: 0
11. Frank, E., Hall, M.A., Holmes, G., Kirkby, R., Pfahringer, B., Witten, I.H.: *Weka: A machine learning workbench for data mining.*, pp. 1305–1314. Springer, Berlin (2005). URL <http://researchcommons.waikato.ac.nz/handle/10289/1497>
12. Gonzalez, R.C., Woods, R.E.: *Digital Image Processing (3rd Edition)*. Prentice-Hall, Inc., USA (2006)
13. Goyal, M., Reeves, N.D., Rajbhandari, S., Ahmad, N., Wang, C., Yap, M.H.: Recognition of ischaemia and infection in diabetic foot ulcers: Dataset and techniques. *Computers in Biology and Medicine* **117**, 103,616 (2020). DOI 10.1016/j.cplbiomed.2020.103616
14. Khalid, S., Khalil, T., Nasreen, S.: A survey of feature selection and feature extraction techniques in machine learning. *2014 Science and Information Conference* pp. 372–378 (2014)
15. Kohavi, R.: A study of cross-validation and bootstrap for accuracy estimation and model selection. pp. 1137–1143. Morgan Kaufmann (1995)
16. Kudva, V., Prasad, K., Guruvare, S.: Detection of specular reflection and segmentation of cervix region in uterine cervix images for cervical cancer screening. *RBM - Revue Europeenne de Technologie Biomedicale* **38**(5), 281–291 (2017). DOI 10.1016/j.irbm.2017.08.003
17. LE, E., JM, B., oldberg M, L, M., L, M., M, S.: Revised national pressure ulcer advisory panel pressure injury staging system: Revised pressure injury staging system. *J Wound Ostomy Continence Nurs* **43**, 585–597 (2016). DOI 10.1097/WON.0000000000000281
18. Li, F., et al.: A composite model of wound segmentation based on traditional methods and deep neural networks. *Comput. Intel. Neurosc.* **2018** (2018)
19. Li, Z., Chen, J.: Superpixel segmentation using linear spectral clustering. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1356–1363 (2015). DOI 10.1109/CVPR.2015.7298741
20. Liu, M.Y., Tuzel, O., Ramalingam, S., Chellappa, R.: Entropy rate superpixel segmentation. In: *CVPR 2011*, pp. 2097–2104 (2011). DOI 10.1109/CVPR.2011.5995323
21. Liu, X., et al.: A framework of wound segmentation based on deep convolutional networks pp. 1–7 (2017)
22. Lu, H., et al.: Wound intensity correction and segmentation with convolutional neural networks. *Concurr. Comp-Pract. E.* **29**(6), e3927–n/a (2017)
23. MEDETEC: Wound database: stock pictures of wounds (2007). URL <http://www.medetec.co.uk/index.html>
24. Nilsson, N.J.: *Artificial Intelligence: A New Synthesis*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (1998)
25. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**(7), 971–987 (2002). DOI 10.1109/TPAMI.2002.1017623
26. Peko, L., Barakat-Johnson, M., Gefen, A.: Protecting prone positioned patients from facial pressure ulcers using prophylactic dressings: A timely biomechanical analysis in the context of the covid-19 pandemic. *International wound journal* **17**(6), 1595–1606 (2020)
27. Quinlan, J.: Decision trees and decision-making. *IEEE Transactions on Systems, Man, and Cybernetics* **20**(2), 339–346 (1990). DOI 10.1109/21.52545
28. Reis, H.C., Turk, V., Khoshelham, K., Kaya, S.: Insinet: a deep convolutional approach to skin cancer detection and segmentation. *Medical & biological engineering & computing* (2022)
29. Scebba, G., Zhang, J., Catanzaro, S., Mihai, C., Distler, O., Berli, M., Karlen, W.: Detect-and-segment: A deep learning approach to automate wound image segmentation. *Informatics in Medicine Unlocked* **29**, 100,884 (2022). DOI <https://doi.org/10.1016/j.imu.2022.100884>. URL <https://www.sciencedirect.com/science/article/pii/S2352914822000375>
30. e Silva, R.H.L., Machado, A.M.C.: A computational method for semi-automatic measurement of pressure ulcers. *Wound Repair Regen* **26**(4), 332–339 (2018). DOI 10.1111/wrr.12650
31. e Silva, R.H.L., Machado, A.M.C.: Automatic measurement of pressure ulcers using support vector machines and grabcut. *Computer Methods and Programs in Biomedicine* **200**, 105,867 (2021). DOI <https://doi.org/10.1016/j.cmpb.2020.105867>
32. Tang, J., Li, B., Gong, J., Li, W., Yang, J.: Challenges in the management of critical ill covid-19 patients with pressure ulcer. *International Wound Journal* **17**(5), 1523–1524 (2020). DOI <https://doi.org/10.1111/iwj.13399>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/iwj.13399>
33. Telea, A.: An image inpainting technique based on the fast marching method. *J. Graph. Tools* **9**(1), 23–34 (2004)

34. Veredas, F.J., Luque-Baena, R.M., Martín-Santos, F.J., Morilla-Herrera, J.C., Morente, L.: Wound image evaluation with machine learning. *Neurocomput.* **164**(C), 112–122 (2015). DOI [10.1016/j.neucom.2014.12.091](https://doi.org/10.1016/j.neucom.2014.12.091). URL <https://doi.org/10.1016/j.neucom.2014.12.091>
35. Wang, L., Pedersen, P.C., Agu, E., Strong, D.M., Tulu, B.: Area determination of diabetic foot ulcer images using a cascaded two-stage svm-based classification. *IEEE Transactions on Biomedical Engineering* **64**(9), 2098–2109 (2017). DOI [10.1109/TBME.2016.2632522](https://doi.org/10.1109/TBME.2016.2632522)
36. Wild, T., Prinz, M., Fortner, N., Krois, W., Sahora, K., Stremitzer, S., Hölzenbein, T.: Digital measurement and analysis of wounds based on color segmentation. *European Surgery* **40**, 5–10 (2014). DOI [10.1007/s10353-008-0378-0](https://doi.org/10.1007/s10353-008-0378-0)
37. Young, S., Narang, J., Kumar, S., Kwizera, E., Malik, P., Billings, S.D., Ko, J.S., Fernandez, A.P.: Large sacral/buttocks ulcerations in the setting of coagulopathy: A case series establishing the skin as a target organ of significant damage and potential morbidity in patients with severe covid-19. *International wound journal* **17**(6), 2033–2037 (2020)
38. Zahia, S., et al.: Tissue classification and segmentation of pressure injuries using convolutional neural networks. *Comput. Meth. Prog. Bio.* **159**, 51–58 (2018)
39. Zhang, T.Y., Suen, C.Y.: A fast parallel algorithm for thinning digital patterns. *Commun. ACM* **27**(3), 236–239 (1984). DOI [10.1145/357994.358023](https://doi.org/10.1145/357994.358023). URL <https://doi.org/10.1145/357994.358023>