Area Determination of Diabetic Foot Ulcer Images Using a Cascaded Two-Stage SVM-Based Classification

Lei Wang*, *Student Member, IEEE*, Peder C. Pedersen, *Senior Member, IEEE*, Emmanuel Agu, Diane M. Strong, and Bengisu Tulu, *Member, IEEE*

***Abstract*—The standard chronic wound assessment method based on visual examination is potentially inaccurate and also represents a significant clinical workload. Hence, computer-based systems providing quantitative wound assessment may be valuable for accurately monitoring wound healing status, with the wound area the best suited for automated analysis. Here, we present a novel approach, using support vector machines (SVM) to determine the wound boundaries on foot ulcer images captured with an image capture box, which provides controlled lighting and range. After superpixel segmentation, a cascaded two-stage classifier operates as follows: in the first stage, a set of *k* binary SVM classifiers are trained and applied to different subsets of the entire training images dataset, and incorrectly classified instances are collected. In the second stage, another binary SVM classifier is trained on the incorrectly classified set. We extracted various color and texture descriptors from superpixels that are used as input for each stage in the classifier training. Specifically, color and bag-of-word representations of local dense scale invariant feature transformation features are descriptors for ruling out irrelevant regions, and color and wavelet-based features are descriptors for distinguishing healthy tissue from wound regions. Finally, the detected wound boundary is refined by applying the conditional random field method. We have implemented the wound classification on a Nexus 5 smartphone platform, except for training which was done offline. Results are compared with other classifiers and show that our approach provides high global performance rates (average sensitivity = 73.3%, specificity = 94.6%) and is sufficiently efficient for a smartphone-based image analysis.**

***Index Terms*—Cascaded classifier, diabetic ulcer, support vector machines (SVM), wound image assessment.**

Manuscript received August 2, 2016; revised October 27, 2016; accepted November 14, 2016. Date of publication November 18, 2016; date of current version August 18, 2017. This work was supported by the National Science Foundation under Grant IIS-1065298. *Asterisk indicates corresponding author.*

*L. Wang is with the Department of Electrical and Computer Engineering, Worcester Polytechnic Institute, Worcester, MA 01609 USA (e-mail: wanglei19841124@gmail.com).

P. C. Pedersen is with the Department of Electrical and Computer Engineering, Worcester Polytechnic Institute.

E. Agu is with the Department of Computer Science, Worcester Polytechnic Institute.

D. M. Strong and B. Tulu are with the Foisie School of Business, Worcester Polytechnic Institute.

Digital Object Identifier 10.1109/TBME.2016.2632522

I. Introduction

IN THE United States, foot ulcers related to type-2 diabetes constitute a significant health issue affecting 5–6 million people [1]. Moreover, diabetes-related wounds resulted in approximately 73 000 lower limb amputations in 2010 [2]. Alexiadou *et al.* [3] Indicate that up to 85% of all diabetic foot ulcer-related amputations are preventable with appropriate clinical interventions and effective patient self-management. In spite of this optimistic claim, the incident rate of foot ulcers is increasing, and effective foot ulcer care imposes a substantial work load for clinicians. Financially, chronic ulcer care adds around U.S.$9–U.S.$13 billion to the annual costs of managing diabetes itself, according to a recent study [4].

Doctors base their wound assessment primarily on visual examination and manual measurements performed either directly on the wounds or on high-resolution wound images. However, such an ad hoc assessment approach does not establish a comprehensive clinical benchmark. A reliable wound assessment can only be achieved by regularly performing accurate measurements of the wound area, analyzing its colors, and the relative sizes of different wound tissues, including proliferation, infected area, slough, or necrosis. Due to the lack of consistency, even with the assistance of tools such as pressure ulcer scale for healing [5], an objective assessment of wound healing rate cannot be ensured. Hence, PC-/laptop-based evaluation of foot ulcers using computer vision and image processing techniques represents an improved approach to accurate chronic wound assessment. Automatic detection of foot ulcer size and tissue composition is especially useful for both clinicians and diabetic patients to monitor the wound healing status and for more effective wound care.

Based on published results of other wound analysis systems [6]–[8], current ulcer image analysis systems mainly focus on addressing the following tasks: 1) high-quality wound image capture, 2) wound boundary determination and wound area measurement, 3) wound tissue classification within the detected wound area, and 4) wound healing rate assessment. This paper focuses on task 2, wound boundary determination. Generally speaking, there are two broad types of methods: 1) "direct" wound boundary determination methods that utilize image analysis techniques and 2) "indirect" wound boundary determination methods, which utilize object recognition techniques that

are popular in computer vision, wherein the wound area is the object to be recognized.

Prior work on "direct" wound boundary determination: In earlier investigations, active contour models and level set-based methods were commonly applied to directly measure the wound areas semiautomatically [9]–[11]. However, these approaches all suffer from problems, such as dependence on the initial contour setting, sensitivity to illumination conditions, and instability when dealing with the poorly defined wound boundaries or false edges. In our previous work in foot ulcer segmentation [12], [13], the mean shift segmentation method and a combined region and boundary algorithm were utilized to find the foot outline, and then three different algorithmic paths were selected based on the wound location, either 1) on the sole, 2) on the edge of the foot, or 3) on the toe, to determine the foot ulcer area. However, these methods suffered from performance limitations when dealing with foot ulcer images with complicated skin textures, such as surgical sutures or other interfering structures very near the wound area.

Prior work on "indirect" wound boundary determination: In recently published works [6], [14], the bottom-up object recognition scheme based on machine learning was applied to wound tissue classification and indirect wound area determination (by grouping all the regions classified as one type of wound tissue). Generally, these approaches consist of following three key steps: *Step 1*: image segmentation; *Step 2:* feature extraction within each segment; and *Step 3:* classifier training on a large number of features from either the wound or nonwound segments. Support vector machine (SVM) [14]–[16] and artificial neural network (ANN) [6] were applied to train the wound tissue classifier. For example, a cascaded classifier based on ANN and Bayesian committee machine was described in [6]. According to published results, the bottom-up scheme provides promising wound tissue classification accuracy on wound images containing mainly wound and only a small amount of surrounding healthy skin tissue. Here, only color and texture features were used to describe each superpixel, which may be sufficient for the classification between wound and healthy skin regions, but when more complicated regions are involved, this feature set no longer suffices. Therefore, when applying SVM and ANN methods to a foot ulcer image where the actual ulcer occupies only a relatively small part of the overall image and where the image may also contain an arbitrary background surrounding the foot, the boundary determination accuracy is not adequate. To address such scenarios, other descriptors, such as bag-of-word (BoW) representation of local scale invariant feature transformation (SIFT) features [17], commonly used in general object recognition tasks need to be investigated.

Wound analysis software systems in current use: There are several software tools for wound analysis and monitoring presently available, such as PictZar Digital Planimetry Software [18], WITA [19], MOWA [20]. However, none of these software packages have yet incorporated automatic wound detection methods to minimize the clinician's initial involvement. A different approach to wound monitoring software and hardware was proposed in [21]. They developed a new "wound mapping" device, which is based on electrical impedance spectroscopy and involves a multifrequency characterization of the electrical properties of wound tissue under an electrode array. Using this approach clinically, however, requires major changes to the daily clinical routine in wound care. Another evidence-based wound assessment tool, called "Silhouette TM," was developed by Aranz Medical [22]. The system includes a 3-D camera with laser guidance that captures a 3-D image of the wound. However, the wound boundary must still be manually delineated, after which the system measures the area, depth, and volume of the wound and its healing progress.

This paper describes a smartphone-based wound image assessment system (bottom-up object recognition scheme based) for assessing diabetic foot ulcers. For the wound image capture, we utilized the image capture box described in our previous work [12], but with improved lighting and with verification that our system introduces a minimal geometric distortion. The image capture box plays an important role to ensure the accurate wound area determination by providing consistent illumination, foot position, and range controlling. For wound area determination, we utilized an image processing technique in the form of a cascaded two-stage approach based on machine learning implemented with a SVM to determine the wound area. SVM-based methods have proven effective when applied to similar clinical problems, such as the MRI brain image segmentation and X-ray bone image analysis [14], [23]. Our proposed training methodology includes the following steps: 1) a simple linear iterative clustering (SLIC) method is used to segment the image into a number of superpixels; 2) significant color and texture features are extracted from these superpixels. Different descriptors are extracted for each of the two-stage SVM classifiers (color feature and BoW representation of dense scale invariant feature transformation (DSIFT) feature for the first stage, color feature and wavelet-based texture feature for the second stage); 3) a first-stage training on a set of k SVM binary classifiers is implemented using different subsets of the entire training images dataset; and 4) a second-stage training on a single SVM binary classifier is implemented using the incorrectly classified test instances from the first stage. The first two steps (segmentation and feature extraction) for the wound boundary determination method are similar to the steps 1 and 2 in the training methodology above (the wound boundary determination process can be viewed as the testing or classification part in the standard machine-learning framework). Subsequently, we apply the two classifiers in a cascaded fashion to find all wound superpixels, based on which we can also indirectly delineate the wound boundary. Finally, we refine the delineated wound boundary using a conditional random field (CRF) technique. The wound boundary determination is completely automatic requiring no human intervention and can handle wounds appearing near the edge of the foot, which is a significant improvement compared to our previous work [12], [13]. The entire system was implemented on Nexus 5 smartphone and tested on 15 subjects at the Wound Clinic and Limb Preservation Center in the University of Massachusetts (UMass) Medical School.

This paper is organized as follows: Section II provides an overview of the foot ulcer image assessment system. Section III introduces the image capture box and the evaluation of the de-

gree of geometric distortion it introduces. In Sections IV–VI, the major three key modules (superpixel segmentation, feature descriptor extraction, and wound classifier training) of our foot ulcer area determination are discussed in detail, with associated experimental results. Finally, Section VII gives an overall conclusion and assessment of the proposed system.

II. Methodology Overview

We captured 100-ft ulcer images over a period of two years at the Wound Clinic at the UMASS Medical School by tracking 15 patients, based on an approved IRB protocol. In most cases, the foot ulcers were located on the sole of the foot. To facilitate the image capture process, we used an improved version of our image capture box as described in Section III. The image capture box maintains consistent illumination, has an optical path of fixed length, and is convenient for elderly patients with limited mobility. We utilized a Nexus 5 smartphone whose F2.0 autofocus lens provides a well-focused high-resolution image over a specified distance range. Hundred photographs were selected, which were considered to be an appropriate sized dataset for wound classifier training. Since the objective is an efficient smartphone-based implementation, we down sampled the original image (with pixel dimensions 3264×2448) by a specific factor in both the horizontal and vertical directions. The rationale for this down-sampling procedure is to speed up the following wound area determination, without losing essential information. The down-sampling factor is empirically set as 4 which will provide a final resolution as 816×612. Based on the observation, this factor provides the optimal balance between processing time efficiency and wound recognition accuracy. For implementation, we apply the API from android bitmap operation library provided by Google.

Furthermore, we framed the wound images to remove irrelevant background information (thanks to the image capture box, we can safely rule out the irrelevant background by simply framing the original image with a rectangle with a fixed size, center, and orientation). The final pixel dimensions of each foot image are 560×320, which has been shown empirically to provide a good balance between the wound resolution and the processing efficiency.

Fig. 1 depicts the overall process for the wound boundary determination on diabetic foot ulcer images. Since our wound boundary determination system is based on the bottom-up object recognition scheme, the first critical task is to segment the wound images accurately into a number of superpixels. In published wound analysis works [6], [14], the mean-shift-based segmentation algorithm [24] has been widely applied because of its outstanding region uniformity preservation. However, in terms of forming superpixels, the mean shift procedure only provides a preliminary result with many oversegmented regions. Hence, a region merge procedure is required to group a number of these oversegmented regions together to form a superpixel representation that better preserves the original boundary information. Three superpixel segmentation algorithms, the graph-based, the mean-shift–based, and the J-SEG-based algorithms, have been described and compared on wound images in [14]. The best

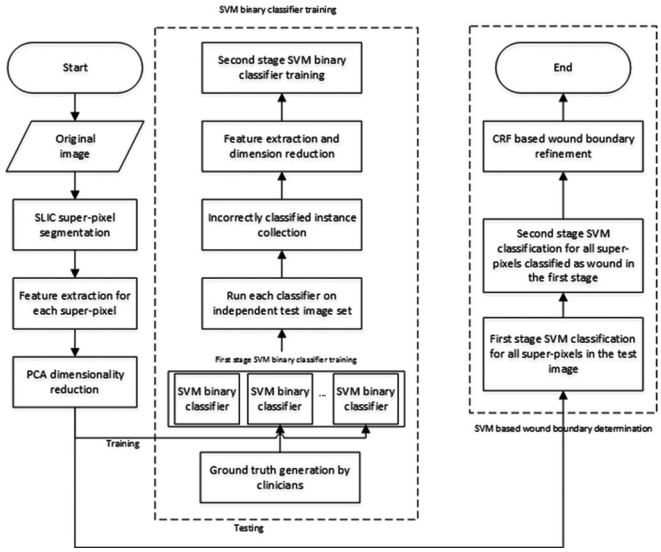

Fig. 1. Complete flowchart for the wound boundary determination system (the details are introduced in the Section II).

results were obtained with the J-SEG algorithm [25]. Both the mean-shift and J-SEG-based algorithms are controlled by several parameters that need to be tuned to adjust the number of output regions. Furthermore, even with the same parameter settings, our experimental results have demonstrated that both the mean shift and the J-SEG methods result in a large variation in the resulting number of superpixels for visually similar wound images. Since our system is intended for the smartphone-based platform, the number of regions will impact the computation time. Hence, we prefer an algorithm which allows direct control of the number of output regions. For this purpose, we chose the SLIC [26] for superpixel segmentation. The SLIC algorithm has been compared to other state-of-art superpixel methods in [27] and has been shown to correctly identify image boundaries, while at the same time providing speed and memory efficiency, which are desirable characteristics for a smartphone-based implementation.

For each superpixel in a given sample image, a feature descriptor is generated based on local color and texture information. Experimental findings in [6] and [14] show that the color-based features are important parameters for determining wound boundaries. Meanwhile, texture is also relevant given that these neighborhood-based features are appropriately selected to describe the difference between the wound and nonwound regions. In [14], several color and texture descriptors, previously applied in dermatological imaging systems, have been explored for wound tissue classification. The experimental results show that the mean color descriptor (MCD), the dominant color descriptor (DCD) in different color spaces (RGB, normalized-RGB, CIE Lab), and the statistics of the color histogram serve as most useful color features for wound tissue classification. With respect to texture, the gray level cooccurrence matrix (GLCM) and wavelet transformation-based features have been shown to be the most effective features. In our paper, we further explore the feasibility of applying another popular visual descriptor used

Fig. 2. Image capture box illustration: (a) Actual product of the image capture box; and (b) wound image captured using the warm LED light.

in other object recognition tasks, such as human recognition and outlining, scene detection, and classification: the BoW representation of the DSIFT descriptors [17]. Based on the experimental results, we extract different feature descriptors for the two-stage classification. The mean-color information (in CIE Lab and normalized RGB space), color histogram statistics, GLCM texture, and BoW representation of the DSIFT features are used for the first stage. The dominant color information and wavelet transformation-based texture are used for the second stage. In each stage, the dimension of feature descriptors can be reduced by applying principal component analysis (PCA) techniques.

Subsequently, a cascaded two-stage SVM-based wound boundary determination method is proposed based on the extracted feature descriptors. Before training the algorithm, three experienced clinicians assisted us in generating the ground truth labels of the wound boundaries for all training images. After the ground truth had been determined, k-fold cross validation was performed to train the set of k SVM binary classifiers in the first stage. After applying these k classifiers to different testing sets, we collected all the incorrectly classified superpixels to form a new training dataset for the next stage. In the second stage, we extracted new feature descriptors for these superpixels and trained a new binary SVM classifier. When analyzing a new wound image, we first applied the first-stage SVM classifier to all superpixels. Then, for all the superpixels classified as "wound," we applied the second-stage SVM classifier to remove false positives. Finally, a CRF-based algorithm was applied to refine the determined wound boundary. In Sections IV to VI, each module of the flowchart will be discussed in detail.

III. Improved Image Capture Box

As described in our previous work [12], to capture an image of their wound, the patient places his/her foot with the wound over the opening of the image capture box and a smartphone is attached next to the foot. By using a pair of straight-angled front surface mirrors, the entire foot region can be imaged. In the improved version of the image capture box, a more complicated layout is employed consisting of three LED lights in the box to minimize the shadows. The actual product and the sample image taken with this box are shown in Fig. 2. To fix the smartphone at the right place, we install four elastic metal clips which allow the user to easily slide the phone in and stop at the place where

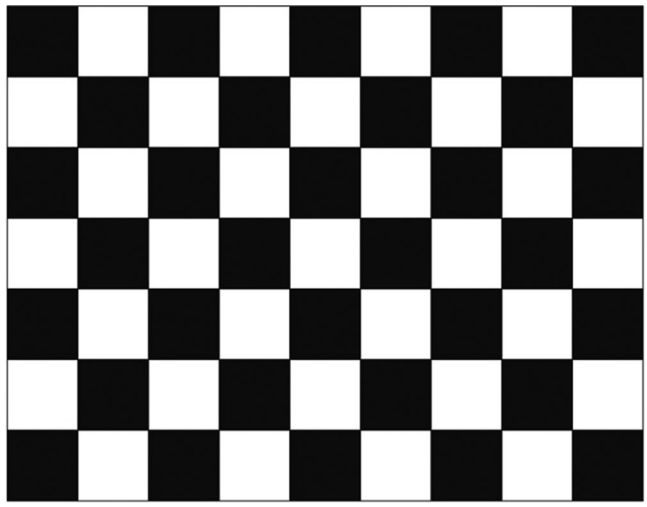

Fig. 3. Grid of black and white squares.

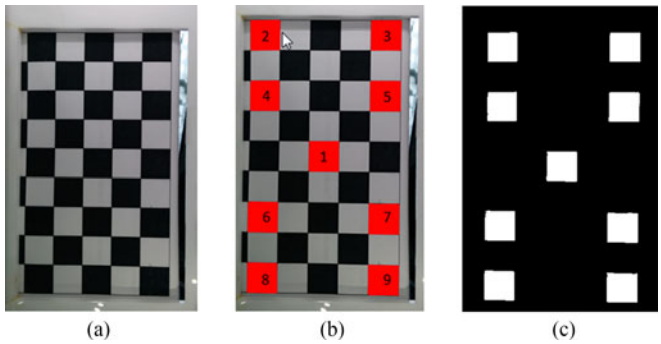

Fig. 4. Geometric distortion estimation for images captured with the image capture box: (a) Original image of a black-white grid sheet; (b) nine selected rectangle areas marked in red for geometric distortion detection; and (c) region area determination results for all rectangle areas.

the camera exactly looks through the hole. When using it, the patient will sit in a chair and place his foot on the opening comfortably. One major concern is the possibility of microbial contamination of the image capture box by the users or the environment. So far, we addressed this problem by wiping the surface of the box with an antimicrobial wipe after each use. A better solution may be a disposable contamination barrier, which will cover the slanted surface of the box except the openings. This will avoid the patient's foot directly touching the surface of the image capture box. We can see from the image in (b) that the plantar surface of the foot is uniformly illuminated.

Since we used a pair of straight-angled mirrors to image the entire foot region, we need to verify that no significant geometric distortion is introduced across the entire image plane. For this test, we prepared a grid of black and white squares on the white paper, as shown in Fig. 3, where the dimension of each square was 15 mm. The grid paper was placed over the opening of the image capture box, and an image was captured, as shown in Fig. 4(a). Next we selected nine black squares at different locations on the image plane, marked these squares

TABLE I
GEOMETRIC DISTORTION ESTIMATION RESULTS

Patch ID	Physical dimension	1	2	3	4	5	6	7	8	9
Area in pixels	N/A	2622	2590	2648	2638	2676	2639	2692	2622	2650
Area in mm^2	225	219	216	220	220	222	220	224	219	220
Error relative to Patch 1 (in %)	+2.7	0	−1.4	+0.5	+0.5	+1.4	+0.5	+2.3	0	+0.5

in red using Photoshop and numbered them 1–9, as shown in Fig. 4(b). After calculating the area of each square, we can determine the geometric distortion by comparing the areas of squares at the edge of the image field with the area of the square in the middle. Because we use the image capture box which provides a fixed object-to-image distance, we can simply convert the area size in terms of pixels to physical units by multiplying a constant scale factor to the former one.

To calculate the square areas, we generated a binary image with the nine selected rectangles highlighted in white, as shown in Fig. 4(c). Then, we applied the mean-shift-based segmentation algorithm to segment the binary image into ten different regions where nine of them represent the nine red rectangles in Fig. 4(b) and the remaining one represents the entire black background region. In this case, the area size for each rectangle can easily be calculated by counting the number of pixels in each region, as shown in Fig. 4(c). The results are displayed in Table I, which shows that the largest deviation from each rectangle area to the middle rectangle is less than 3% of the mean area size. Even when considering the inaccuracy that can possibly be introduced by the square patch delineation using Photoshop, we can conclude that the geometric distortion is small enough to be considered negligible.

IV. SUPERPIXEL SEGMENTATION

A good superpixel segmentation algorithm should have the following three properties: 1) good image boundary adherence (i.e., the edge information in the original image will be preserved in the segmentation result), 2) computationally efficient and with modest memory requirements, and 3) easy to use and control [27]. The most widely used superpixel segmentation algorithms can be divided into two categories [27]: 1) graph-based algorithms, which view each pixel as a node in a graph and use the weight between any two nodes to measure their similarity [28], and 2) gradient ascent-based algorithms, which start from an initial node and iteratively refine the clusters until the convergence criterion is met to form superpixels (mean shift algorithm is the best known in this category [24]). Previous works prove that no existing segmentation method can provide ideal performance under all circumstances [27]. When dealing with a high-resolution image, graph-based algorithms can become very computational demanding if the number of nodes grows rapidly. As mentioned before, in wound tissue classification studies, researchers have mainly used the mean-shift-based algorithm. However, these mode-seeking approaches suffer from several drawbacks: 1) they usually stop at an oversegmented state and require further region merge procedures; 2) several parameters are required to control the performance, and poor parameter tuning can result in efficiency and performance degradation; and 3) the number of superpixels generated cannot be directly controlled. In other bottom-up object recognition applications, the JSEG algorithm, consisting of three steps (color space quantization, J-image calculation, and region growing-based region merge), has also been widely used [14], [25]. However, the number of superpixels generated can also not be controlled. To overcome these shortcomings, in this paper, we utilize the SLIC method [26], [27].

Before the superpixel segmentation, the original image needs to be transformed from RGB space to CIE Lab space in order to obtain better perceptual accuracy. Similar to the mean shift algorithm, the SLIC clustering takes spatial continuity into account by expanding the original 3-D color range space to a 5-D space, which includes two spatial components. By default, we only need to specify one parameter k, the desired number of approximately equal-sized superpixels. Unlike the random initialization required for the traditional K-means clustering method, the SLIC algorithm places k initial cluster centers uniformly on the image grid with S pixels separating each pair. It is easy to verify that $S = \sqrt{N/k}$, where N denotes the number of pixels (i.e., the product of image height and width). In this case, we can expect the segmented superpixels to be of approximately equal size. In the assignment step, each pixel is assigned to the closest cluster center, which is similar to the traditional K-mean process. However, instead of computing the distance between the target pixels and all cluster centers, SLIC only computes distances from each cluster centers to pixels within a $2S \times 2S$ neighborhood region. This approach narrows down the search region, effectively reducing the number of distance calculations and accelerating the algorithm. Meanwhile, it is also pointed out in [27] that this speeding up is allowed only when a special distance measure D is introduced, as formulated in (1). The novel distance measure D combines distances from color space (d_c) and spatial space (d_s) to effectively reduce the variance of boundary adherence due to changes in the superpixel's size [27]:

$$\begin{aligned} d_c &= \sqrt{(L_i - L_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2} \\ d_s &= \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \\ D &= \sqrt{d_c^2 + \left(\frac{d_s}{S}\right)^2 m^2} \end{aligned} \tag{1}$$

where $(L_i, a_i, b_i, x_i, y_i)$ and $(L_j, a_j, b_j, x_j, y_j)$ represent two points in the 5-D color and spatial spaces. The variables d_c and d_s represent distances in color subspace and the distance

spatial subspace, respectively. As mentioned earlier, S denotes the approximate superpixel size. The parameter m is used to weight the relative importance between color similarity and spatial proximity. It is recommended in [27] that m should be in the range [1], [40] if the algorithm is performed in CIE Lab color space. In our case, we empirically set m to 16 and S to 30. Then, the following process is similar to K-mean algorithm: 1) replace the cluster center with the mean of all pixels in that cluster; 2) calculate the cluster center change and decide whether it converges; and 3) reallocate isolated pixels to nearby cluster centers to enforce connectivity.

It is demonstrated in [6] that the "ground truth" is not always a practical basis for evaluating the performance of superpixel segmentation algorithms. The reasons are: 1) to generate the ground truth, experienced wound clinicians are required to manually segment each wound image into 100 of regions; 2) this process is not only time consuming but also lacks consistency in terms of both the segmentation granularity and boundary determination for each superpixel; and 3) for segmentation algorithms with high level of complexity, it makes more sense to apply goodness evaluation methods, which assess the performance on a number of test images by using some desirable properties of segmented images as goodness measures. Hence, we use the evaluation method used in [6] to assess the superpixel segmentation result of the SLIC algorithm, and compare it to the mean-shift-based algorithm used in [29] and the JSEG algorithm used in [14]. The evaluation is based on two measures: region uniformity measurement U_α and region contrast measurement C_α of the entire segmented image, as defined in (2) and (3). In both these equations, f_i is the color vector in CIE Lab space for pixel i in region R_j, $\bar{f}_j$ is the average value of all f_i in R_j, A_j is the area of the region R_j, and $f_{\min}$ and $f_{\max}$ are the maximum and minimum values in this region. It is obvious that the larger the value of U_α, the better the uniformity is achieved in the current region, where α is the significant area where the measure is evaluated. In (3), c_j is the contrast measure in region R_j and v_j is the weight assigned to region R_j. Based on [6], v_j is modeled as the simple linear contribution of region R_j to the total area of image. The parameter c_{ij} represents the contrast between regions R_j and R_i. Finally, p_{ij} is the adjacency parameter specified as the ratio of the common perimeter of R_j to R_i. More details about the definition and derivation of these formulas can be found in [30]:

$$U_\alpha = 1 - \frac{2\sum_{R_j\in\alpha}\sum_{i\in R_j}\left\|f_i - \overline{f_j}\right\|}{\sum_{R_j\in\alpha} A_j \left\|f_{\max} - f_{\min}\right\|} \tag{2}$$

$$C_\alpha = \frac{\sum_{R_j\in\alpha} v_j c_j}{\sum_{R_j\in\alpha} v_j}$$

$$c_j = \sum_{Adj R_i} p_{ij} c_{ij}$$

$$c_{ij} = \frac{\left\|\overline{f_i} - \overline{f_j}\right\|}{\left\|\overline{f_i}\right\| + \left\|\overline{f_j}\right\|}. \tag{3}$$

We applied the superpixel segmentation evaluation scheme proposed in [6], specifically, we calculated the mean and standard deviation of the U_α and C_α measures for three superpixel segmentation algorithms (with appropriate parameter tuning) when applied to 100 sample images (significant area α in (2) and (3) is set as the entire image area). The results are shown in Table II. In Table III, the same indicators are calculated with the significant area α containing the entire wound area manually delineated by three experienced clinicians (the details about these ground truth generation process will be discussed in detail in Section V). The implementation details of the mean shift + region merge algorithm and JSEG method are beyond the scope of this paper and can be found in [23], [25], and [28]. In Table IV, the efficiencies (running time) of these three segmentation approaches are compared, using the Nexus 5 smartphone platform with a Quad core CPU, 2.3-GHz Krait 400, 2-GB RAM. We can see from Tables II and III that the mean value of region uniformity is generally large and that the mean value of region contrast is generally small. These results were expected since we deliberately tuned the parameters to obtain a relatively large number of superpixels. Besides, the standard deviation is generally small for all cases. Comparing the data in these two tables, it can be concluded that in the wound area, uniformity is smaller and region contrast is higher than for the entire image. This shows that the nature of the wound area is more complicated than that of the background and the healthy skin. Comparing the different columns in both tables, we find that SLIC provides the best performance balance between the uniformity and contrast measures. However, the difference in the performance of

TABLE II
STATISTICS OF REGION UNIFORMITY AND CONTRAST MEASURES ON THE CIE LAB SPACE WITH THE SIGNIFICANT AREA CONTAINING THE ENTIRE IMAGE FOR THREE DIFFERENT SUPERPIXEL SEGMENTATION METHODS

Method	Mean shift		JSEG		SLIC	
	μ	σ	M	σ	μ	σ
U_α	0.867	0.022	0.878	0.051	0.873	0.032
C_α	0.147	0.008	0.166	0.019	0.099	0.027

TABLE III
STATISTICS OF REGION UNIFORMITY AND CONTRAST MEASURES ON THE CIE LAB SPACE WITH THE SIGNIFICANT AREA CONTAINING THE MANUALLY DELINEATED WOUND AREA FOR THREE DIFFERENT SUPERPIXEL SEGMENTATION METHODS

Method	Mean shift		JSEG		SLIC	
	μ	σ	M	σ	μ	σ
U_α	0.692	0.109	0.769	0.162	0.748	0.174
C_α	0.217	0.058	0.233	0.084	0.184	0.061

TABLE IV
COMPUTATION TIME EVALUATION ON NEXUS 5 SMARTPHONE PLATFORM FOR THREE DIFFERENT SUPERPIXEL SEGMENTATION ALGORITHMS

Method	Mean shift	JSEG	SLIC
Computation time	12.7 s	9.7 s	7.2 s

different algorithms is not large. The average region uniformity U_α is larger than 0.86 and the average region contrast C_α is smaller than 0.2 for all three methods. Hence, based on these findings, we conclude that all three superpixel segmentation algorithms can achieve good performance given careful parameter tuning. We selected the SLIC algorithm for several reasons. As mentioned earlier in Section 2.2.3, we only need to tune one parameter for the SLIC algorithm. In contrast, at least two and three parameters need to be tuned in the JSEG shift and mean shift algorithms, respectively [24], [25]. Furthermore, SLIC has the best performance in terms of computation time, as shown in Table IV.

V. Feature Descriptor Extraction for Superpixels

In terms of selecting appropriate descriptors for wound classification, we concentrate on the color and texture descriptors already applied in the wound tissue classification tasks [6], [14], [31]. Generally speaking, color features are very useful for wound recognition and tissue classification, as demonstrated by the red-yellow-black healing visual assessment scheme [32]. Texture features, even though they have lesser significance, can still be useful in refining the classification results [14]. The superpixel-based local color features used in previous wound tissue classification work include mean color, dominant color, statistics (highest peaks, variance, skewness, energy, and entropy) of color histogram and sampled multidimensional color histogram [31] in various color spaces (RGB, normalized-RGB, HSV, CIE Lab, and CIE Luv). The widely used local texture features have specific measures, determined by the chosen descriptors: 1) energy of each subband of the two-stage wavelet transformation, 2) Gabor transformation-based features on five scales, 3) local binary pattern histograms, 4) normalized texture contrast and anisotropy, and 5) GLCM-based features. Wannous *et al.* [14] compared the extraction efficiency and predictive power of several local color and texture features and concluded that the combined descriptor of mean color, dominant color, and GLCM-based features provided the best performance (92% for specificity and 77% for sensitivity) [14]. Veredas *et al.* [6] applied the combination of mean and variance of local color in RGB, normalized RGB and CIE Luv spaces, histogram statistics, and wavelet-based texture features as region descriptors. It is claimed that a high global wound classification performance (95% for specificity and 79% for sensitivity) was achieved by applying neural networks and Bayesian classifiers [6]. However, both works dealt with wound images that contained predominantly wound tissue, with just a small amount of surrounding healthy skin tissues. In our scenario, the wound boundary determination task is more challenging since the wound image contains large area of healthy skin tissues (sometimes the entire foot) and even some arbitrary background surrounding the foot. After applying a combination of the sets of descriptors proposed in previous works, the wound boundary was not always accurately determined and the sensitivity was not high enough (<65%), which implies that a significant amount of nonwound regions have been incorrectly classified as wound regions. Hence, it became necessary to explore additional features that would improve the accuracy of wound boundary determination.

In other object recognition applications, such as face detection, pedestrian detection, and scene classification, the BoW representation of local features is widely used as a powerful descriptor for classification. BoW treats image features as words, where BoW is a sparse histogram of word counts over the vocabulary of image features. This direction of development is now established as a state-of-the-art approach for visual object category classification [17], [33], [34]. The wound boundary determination task can also be treated as an object recognition task, although the "object," i.e., the wound, is often complex due to lack of characteristic outline and internal structure. Hence, we applied the normalized BoW histogram representation as a supplement to the color and texture features. To construct a less sparse BoW histogram representation (meaning that the number of zeroes in the histogram representation is reduced) and to provide more relevant information to the classifier, we extracted SIFT descriptors for each pixel (or over a denser grid than traditional SIFT) of the image at a fixed scale and orientation using the DSIFT framework [17], [35]. As mentioned in [17], the computation of DSIFT descriptors is usually accompanied by a clustering stage, where the individual SIFT descriptors are reduced to a small vocabulary of visual words, which can then be combined with the BoW model method [17]. The extracted descriptors are clustered to generate *K* (a predefined number) centers (a so-called dictionary of visual words) using a K-means algorithm. Then, each descriptor is classified based on its nearest center as determined by the minimum square distance rule. The descriptors in each superpixel s_i are aggregated into a single normalized histogram $h_i \in R_+^K$ based on the classification results, where *K* is the number of words predefined in the *codebook* (the set of clusters resulted from the K-means algorithm, which in our case was set to 400).

The BoW histogram-based descriptor has two major drawbacks [17]: 1) the resulting classifier is superpixel specific without considering the surrounding regions for each superpixel, and 2) the histograms are still very sparse even when based on DSIFT features, since the superpixels are inherently uniform and a large number of DSIFT features are assigned to the same visual word. To overcome the problems caused by the lack of consideration of the surrounding region of each superpixel and sparse histogram representation, we apply the histograms to superpixel neighborhoods [17]. Let h_i^0 be the histogram associated with superpixel s_i. Then, h_i^N is the histogram obtained by merging the histograms of the superpixel s_i and neighbors which are less than *N* nodes away in the adjacency graph of the current superpixel [17]:

$$h_i^N = \sum_{s_j | D(s_i, s_j) \leq N} h_j^0. \tag{4}$$

In our experiments, we evaluated four combined feature descriptors by applying them to the single-stage binary SVM-based wound classifier, using 100 sample images and employing the leave-one-out cross-validation scheme. More details about the machine-learning techniques we utilized can be found in the Section VI. Descriptors 1 and 2 are formulated based on [6] and

TABLE V
WOUND RECOGNITION PERFORMANCE EVALUATION FOR DIFFERENT COMBINED DESCRIPTORS

Descriptor ID	Classification results (α = entire image)		Classification results (α = wound + healthy skin only)	
	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
1	68.7	86.2	72.2	86.1
2	68.3	86.9	74.5	88.3
3	71.4	87.2	69.4	86.6
4	70.4	87.0	72.3	85.8

1) Descriptor 1: MCD + DCD + GLCM [14], descriptor 2: MCD + color histogram + wavelet [6], descriptor 3: MCD + color histogram + BoW, and descriptor 4: MCD + DCD + color histogram + wavelet + GLCM + BoW; 2) descriptors are evaluated on two different types of image area: the entire image area containing wound bed, healthy skin, and irrelevant background and the area containing only wound bed and small amount of surrounding healthy skin.

TABLE VI
COMPUTATION TIME EVALUATION FOR DIFFERENT COMBINED DESCRIPTOR GENERATION (INCLUDE THE PCA DIMENSION REDUCTION) BASED ON NEXUS 5 SMARTPHONE

Descriptor ID	1	2	3	4
Computation time	4.5	4.2	5.6	8.8

[14] and contain color and texture features, respectively. In descriptor 3, we use the color features introduced in [6] (mean and variance of color in the CIE Lab and normalized RGB spaces, two highest peaks, dominant levels, variance, skewness, energy and entropy of the color histogram in the RGB, and normalized RGB spaces), and we append the color features with the BoW histogram representation of local DSIFT features. Descriptor 4 combines the color and texture features used in descriptors 1–3. After applying the PCA dimension reduction method by maintaining 99% total invariance, the resulting descriptor lengths are 12, 19, 33, 39 for descriptors 1–4, respectively.

As seen in Table V, the general sensitivity and specificity measures when evaluating the entire image show that the descriptor 3 provides the best performance (71.4% for sensitivity and 87.2% for specificity) on the 100-ft wound images. However, when the significant area is confined to the wound itself and small region of surrounding skin, descriptor 2 provides the best performance (74.5% for sensitivity and 88.3% for specificity). These results may indicate that the BoW features are useful for ruling out the irrelevant background information, but is less able to distinguish between the wound and healthy skin comparedto wavelet and GLCM-based texture features. This observation led us to implement the cascaded two-stage classifier design introduced in the next section. Morever, the longest feature descriptor 4, while obviously more computationally demanding as seen from Table VI, provides no major performance improvement. Based on the discussion above, we applied the descriptor 3 and descriptor 1 to the first- and second-stage classification, respectively. The cascade two-stage classification system will be introduced in the next section.

VI. SUPERVISED MACHINE-LEARNING-BASED METHOD FOR WOUND BOUNDARY DETERMINATION

A. Expert Labeling for Supervised Machine Learning

Wannous *et al.* [6], [14] developed graphical user interfaces to deal with the time-consuming task of labeling the whole set of more than 10 000 regions (superpixels) from the segmented wound images. It was reported in [6] that labeling just one image required 6 min. To save the clinicians' time, we proposed a simpler labeling method. For the images used in this study, a group of three expert wound clinicians from the Plastic Surgery Department at the UMASS Medical School delineated the wound boundary for all 100 wound images independently. The setup consisted of a laptop, installed with Photoshop and connected via USB to a Wacom drawing tablet. Using the electronic pen, it took on average only 30 s for each clinician to label one wound image. To generate the ground truth based on a binary label for each superpixel (wound or nonwound), we first assign each pixel a label by applying a majority vote scheme [14]. Then, for each superpixel, another majority vote scheme is applied: if more than half of all pixels in this superpixel are labeled as "wound," we assign it as a wound region; otherwise, we assigned it as nonwound region. To validate our labeling method, we reused the wound boundary delineation based on the label associated with each superpixel in a given image and compared it with the original delineation from the clinicians. The resulting similarity at the pixel level between these two sets of delineations was higher than 96%. Overall, we concluded that the novel method provided labeling results with high accuracy, although the second majority vote may cause some inaccuracy for regions near the wound boundary.

B. Two-Stage Cascaded SVM-Based Machine-Learning Architecture

In [6], Verdedas *et al.* utilized the approach of creating a stable neural network-based classifier from a set of subclassifiers trained by following a cross-validation strategy. The results proved that this approach provides promising wound tissue classification performance when dealing with images containing only wound area and some surrounding healthy skin. We have modified Verdedas's proposal to a two-stage cascade machine-learning architecture based on SVM, as shown in the following steps.

Step 1: Split the entire training image dataset into k subsets of equal size. Since we can control the segmented superpixel number by using the SLIC algorithm and the dimensions of most training images are the same, we have approximately an equal number of superpixels for each training image. Hence, in each subset, we have approximately equal numbers of both images and superpixels.

Step 2: Make the number of wound regions and nonwound regions from k−1 subsets equal. Since we use the image capture box to photograph the foot, sometimes the number of wound regions (i.e., superpixels) is relatively small compared to the number of nonwound regions. This skewed distribution of the number of instances in the different cat-

egories can severely undermine the performance of the trained SVM classifier. Hence, if our training set has m wound superpixels and n nonwound superpixels ($n >> m$) in the subset 1 to $k-1$, we will randomly select only m nonwound regions from n in total. Consequently, our training set has $2m$ training patterns.

Step 3: Follow a classical tenfold cross-validation scheme to train a binary SVM-based classifier on these 2m training patterns. Based on each subset, we further split this subset into ten equal-sized folds. Then, we run the standard RBF-χ^2 kernel-based soft margin SVM (C-SVM) [36] ten times, where on each run, one fold is the validation set while the remaining nine folds are the training set.

Step 4: Use the trained classifier to classify regions in a subset other than the k−1 subsets for training.

Step 5: Collect the incorrectly classified instances from the results in Step 4 into the training set for the second stage.

Step 6: Repeat the first-stage training (Steps 2 to 5) k times and let each subset be the test set exactly once.

Step 7: Train the second-stage SVM binary classifier in the same way as described in Step 3.

Based on the findings in Section IV, we extract wound image descriptors consisting of color features and the BoW histogram representation of DSIFT features for superpixels for the first-stage training set. For the second-stage training set, we extract the descriptors consisting of color features and wavelet-based features. In kernel-based SVM training, the nonlinear separable classes in low-dimensional feature space may become linearly separable when mapping features to higher dimensional space using a suitable kernel function [37]. Hence, the key point of the SVM classifier design remains the choice of the kernel function. Since no universal kernel will fit all applications, we evaluated various classic kernels, including linear, polynomial, perceptron, and RBF-χ^2, and we found that the RBF-χ^2 provided the best performance. After the selection of the RBF-χ^2 kernel, its parameters must be tuned. In our case, we utilized the soft-margin SVM with the RBF kernel. Hence, there are two parameters that need to be tuned: (C, Υ) where C is the regulation parameter and Υ is the kernel parameter [36]. The best parameter combination is often selected by a grid search with exponentially growing sequences of the two parameters [38]. Each point in this grid was used as parameters in the SVM training process and the general testing error (GTE) on the validation sets was evaluated. The optimal parameter combination was the one that provided the lowest GTE. After running the grid search for both stages, we found that the optimal parameters for the first and second stages were ($C = 100$, $\Upsilon = 0.05$) and ($C = 85$, $\Upsilon = 0.12$), respectively.

To evaluate the performance of the two-stage cascaded SVM-based approach more completely, we also compared it to other machine-learning strategies. The strategies we compared against included 1) a single-stage SVM-based classifier with the same configuration as the first-stage classifier in the two-stage approach, and 2) a single-stage classifier based on neural network (ANN) with one hidden layer of 40 neurons [6]. For both single-stage methods, we applied descriptor 2 introduced in Section V (this descriptor is a combination of mean color, color histogram, and wavelet-based texture features). As mentioned in Section IV, for the two-stage SVM-based method, we applied descriptors 3 and 2 for the first and second stages, respectively. Here, we used the specificity and sensitivity as performance measures of the wound boundary determination methods. Despite a slight increase in the computation time (as shown in the bottom row of Table VII), our two-stage SVM-based classifier outperforms both single-stage methods, by being approximately 5% higher in sensitivity and was 6% higher in specificity. Three examples of our wound classification results are shown in Fig. 5. Note that we implemented the SVM algorithm using libSVM library [37].

TABLE VII
WOUND BOUNDARY DETERMINATION PERFORMANCE EVALUATION OF DIFFERENT MACHINE-LEARNING STRATEGIES

Method ID	1	2	3	4
Sensitivity (%)	68.3	66.4	71.4	73.3
Specificity (%)	86.9	83.7	92.8	94.6
Computation time (s)	15.4	16.1	18.8	20.5

Method 1: single-stage SVM-based method, method 2: single-stage ANN-based method, method 3: novel two-stage SVM-based method, method 4: two-stage SVM + CRF refinement. The Computation time measures the average time consumed by the entire wound boundary determination process for 100 images including superpixel segmentation, descriptor generation, and classification, and only method 4 includes the boundary refinement process.

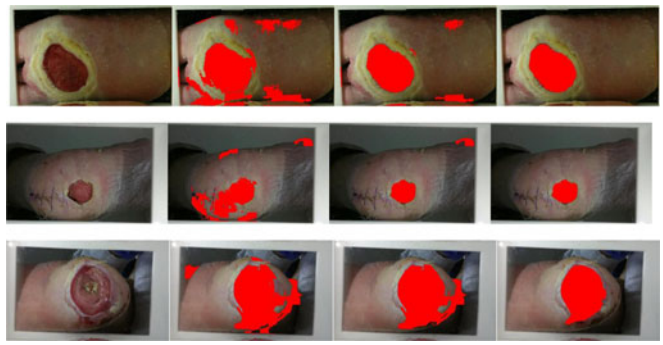

Fig. 5. Examples of wound boundary determination results (the determined wound areas are covered with red color), column 1: the original image, column 2: the boundary determination results by applying our two-stage classifier, column 3: the results after applying the CRF refinement technique, and column 4: the results after the outlier removal or hole filling up.

C. Wound Boundary Refinement

Even though the two-stage SVM-based classifier significantly improves the wound boundary determination performance, some regions may still be misclassified. Most of the erroneously classified regions were located near the wound boundary and were connected to the target wound area. In addition, some misclassified regions may also be located inside the wound bed (holes) and/or in the healthy skin (outliers). To refine the wound boundary, we first generate a binary result image based on the wound boundary determination result, by assigning a value of 255 to the pixels classified as "wound" and 0 to the pixels classified as "nonwound."

Following that there are three major steps for wound boundary refinement. The first step is to apply the CRF technique to relabel the misclassified regions near the actual wound boundary. Ideally, all correctly classified regions should retain their labels after the CRF refinement, while the labels for misclassified regions should be flipped. Otherwise, extra errors might be introduced by the refinement process itself. The second step is to break narrow bridges (if there are any) and fill the remaining holes (if there are any) inside the determined wound area by applying the closing image morphological operation (we used the circular structure element with radius $= 3$ pixels). For the last step, we try to remove the outlier regions if there are any. This is easy because outliers are always disconnected from the main wound area and small in size. We can simply run a connected region detection method [39] and only keep the largest connected region. Examples of the results of boundary refinement after the two previous steps are shown in Fig. 5, column 4. By comparing the last two images on the third row, significant refinement can be observed on the boundary determined by last two steps when CRF alone fails to provide accurate refinement.

Since the last two steps are straightforward and only involve basic image processing techniques, in the remaining part of this section, we will focus on the first refinement step. CRF and level set are two candidate techniques, which have been used for postboundary refinement in similar applications (e.g., refine the determined boundary of an individual in an image) [40]. As mentioned in Section I, the level set-based algorithm can be unstable especially when the actual wound boundary is poorly defined. Hence, we decided to utilize CRF to recover more precise boundaries by further reducing the misclassification occurring near the edges of the wound. CRF provides a natural way to incorporate such constraints by including them in a pairwise edge potential of the model. The refinement problem can be formulated as the task of finding the most probable labels for all pixels in a given image. It includes the use of both unary pixel properties and pairwise relations between pixel labels. In image processing using CRF, the labeling problem is transformed into an energy minimization problem. The energy is usually defined as the sum of a series of unary terms and pairwise terms, which indicate individual label preferences (unary term) and spatial coherence, respectively (pairwise term).

Let $P(c|G;w)$ be the conditional probability of the set of class label assignments c, given the adjacency graph $G(S,E)$ and scalar weight w. The energy function in [17] and [41] is shown as

$$-\log(P(c|G,w)) = \sum_{s_i \in S} \psi(c_i|s_i) + w \cdot \sum_{(s_i,s_j)\in E} \phi(c_i,c_j|s_i,s_j). \tag{5}$$

Our unary potentials ψ are defined directly by the probability outputs (provided by the libSVM library) for each superpixel as

$$\psi(c_i|s_i) = -\log(P(c_i|s_i)). \tag{6}$$

Regarding the pairwise term in (5), our pairwise edge potential φ is identical to the one introduced in [23], defined as

$$\phi(c_i,c_j|s_i,s_j) = \left(\frac{L(s_i,s_j)}{1+\|s_i-s_j\|}\right)[c_i \neq c_j] \tag{7}$$

where [.] is the zero-one indicator function and $\|s_i - s_j\|$ is the norm of color difference between any two superpixels in the Luv (or CIE Lab) color space where the perceptible color difference can be directly calculated. $L(s_i,s_j)$ is the shared boundary length between superpixels and acts here as a regularizing term which discourages small isolated regions.

In applications of CRF for computer vision tasks, the unary and pairwise potentials are represented by a weighted summation of many simple features, and so the parameters of the model are trained by maximizing their conditional log likelihood. In the formulation in (5), the CRF model simply has one weight w, which represents the tradeoff between spatial regularization and our confidence in the classification. We estimated w by applying cross validation on the training data [17]. Once our model has been trained, we wish to find the most probable labeling c^*, i.e., the labeling that maximizes the conditional probability in (5). The optimal labeling is found by carrying out inference with the multilabel graph optimization library of [42] using α-expansion. Since the CRF is defined on the superpixel graph, inference is very efficient, taking less than half a second per image on a typical modern PC [17]. Based on our experimental results, the CRF refinement consumes 2.2 s on average on the Nexus 5 smartphone, which is still fast enough for processing at interactive rates. Examples of the results of CRF refinement are shown in Fig. 5, column 3. To demonstrate the refinement effectiveness, we reevaluate the specificity and sensitivity measures based on the CRF-refined boundary determination results on the same testing images in Section IV. By comparing columns 2 and 5, a 3.8% improvement on sensitivity and 1.2% improvement on specificity can be observed.

VII. Discussion and Conclusion

A novel smartphone-based system for automatic wound boundary determination on diabetic foot ulcer images has been presented in this paper. We tracked 15 patients in the Wound Clinic at UMASS over a two-year period resulting in 100-ft high-resolution ulcer images that were captured using an image capture box. The images were first segmented into superpixels by using the SLIC algorithm, which our evaluation has shown to outperform other widely used segmentation algorithms. The main contribution of this paper is the development of a cascaded two-stage SVM-based classifier to determine the wound boundaries. The inputs of the classifiers were color and texture descriptors of superpixels. To train the supervised classifiers, we asked three experienced clinicians to delineate the wound boundary on 100 images and designed a novel way to generate the ground truth label for each superpixel. In the first stage, k C-SVM-based classifiers were trained by a k-fold cross-validation strategy on the entire training dataset. In the second stage, only the incorrectly classified instances, when k classifiers were applied in the first stage, were used as the training set to train another C-SVM-based classifier. To optimize performance,

we used different combinations of color and texture descriptors for the two training stages. Finally, the determined boundary was further refined utilizing a number of image processing techniques, including morphological operations, connected region detection, and a CRF-based relabeling method, to either remove the outlier nonwound regions or fill undetected wound regions.

To evaluate our two-stage binary classification system, we compared its performance with two other machine-learning strategies: a single-stage SVM classifier and an ANN-based classifier. Moreover, the classification performance on different color and texture descriptors were also compared to a single-stage SVM machine-learning strategy. During the evaluation, we found that the sensitivity was the performance measurement that varied the most across different machine-learning approaches analyzed in this paper. In contrast, almost every machine-learning strategy used here can achieve promising specificity (higher than 92%) and there was only small variance between methods (<3%). According to our results, the mean color + color histogram + wavelet texture feature (applied in our second-stage training) showed the best ability to distinguish between wound regions and surrounding skin regions. However, the dominant color + color histogram + BoW histogram of DSIFT features (applied in our first-stage training) provided the best sensitivity when classifying irrelevant background and foot tissue. Furthermore, our two-stage SVM-based classifier provided the best sensitivity rate to determine wound boundary from foot ulcer images containing wound regions, as well as the foot, healthy skin, and irrelevant background.

We implemented the entire wound classification system on a Nexus 5 android smartphone and measured the computation time for each module. The average time required for our two-stage classification method was 17.5 s. Compared with the method proposed in [14], the running time of our two-stage approach only increased slightly, i.e., by 2 s. However, the performance of boundary determination has been improved significantly (12% improvement on sensitivity and 8% on specificity). We only implemented the offline classification system on the smartphone, excluding the classifier training process. Moreover, the SVM-based training strategy is good at dealing with descriptors containing a large number of elements on a relatively small training set. Hence, the PCA dimension reduction was not really helpful and we did not use it in the actual smartphone implementation as it would further increase the computation time.

As mentioned earlier, the good performance of wound area determination relies on the image capture box. In another words, our approach can achieve robust accuracy only when the image capture condition, including illumination, foot position, and foot-to-camera distance are consistent. All previous discussion regarding method selection for superpixel segmentation and feature descriptor generation are also based on this well-controlled image database.

To obtain a more precise estimation of the appropriateness of our machine-learning approach and extend its feasibility, a few tasks should be addressed in the future. 1) More clinicians should be recruited to delineate wound boundaries in order to minimize the impact of inter- and intraobserver variability, which was reported as an unavoidable problem to undermine the classification efficiency when using manual delineation as the ground truth; 2) the current wound image database should be extended to further enhance the classifier training. Hence, not only do more type-2 diabetic patients need to be recruited, but we also considered training our supervised wound classifier on other type of wound images (for example, burn wounds), and also wound images taken under varying lighting conditions. In these cases, we believe that our methodology would be easy to apply in other similar scenarios; and 3) cloud computing techniques could be exploited to update the wound classification model using an online machine-learning strategy to add more flexibility to our method, rather than prestore the trained model files on the smartphone.

Acknowledgment

The authors would like to thank all the reviewers for their constructive comments which greatly improve the scientific quality of this paper.

References

[1] K. M. Buckley *et al.*, "Reducing the risks of wound consultation: Adding digital images to verbal reports," *Wound Ostomy Continence Nursing*, vol. 36, no. 2, pp. 163–170, 2009.
[2] NIH's National Diabetes Information Clearing House, National Institutes of Health, Bethesda, MD, USA, 2011.
[3] K. Alexiadou and J. Doupis, "Management of diabetic foot ulcers," *Diabetes Therapy*, vol. 3, no. 1, pp. 1–15, 2012.
[4] J. B. Rice *et al.*, "Burden of diabetic foot ulcers for medicare and private insurers," *Diabetes Care*, vol. 37, no. 3, pp. 651–658, 2014.
[5] C. S. J. Cuddigan and E. Ayello, "Pressure ulcers in America: Prevalence, incidence, and implications for the future. An executive summary of the national pressure ulcer advisory panel monograph," *Adv. Skin Wound Care*, vol. 14, no. 4, pp. 208–215, 2001.
[6] F. Veredas *et al.*, "Binary tissue classification on wound images with neural networks and Bayesian classifiers," *IEEE Trans. Med. Imag.*, vol. 29, no. 2, pp. 410–427, Feb. 2010.
[7] E. S. Papazoglou *et al.*, "Image analysis of chronic wounds for determining the surface area," *Wound Repair Regeneration*, vol. 18, no. 4, pp. 349–358, 2010.
[8] F. J. Veredas *et al.*, "Wound image evaluation with machine learning," *Neurocomputing*, vol. 164, pp. 1–11, 2015.
[9] T. D. Jones and P. Plassmann, "An active contour model for measuring the area of leg ulcers," *IEEE Trans. Med. Imag.*, vol. 19, no. 12, pp. 1202–1210, Dec. 2000.
[10] P. Plassmann and T. D. Jones, "Improved active contour models with application to measurement of leg ulcers," *SPIE Electron. Imag.*, vol. 12, no. 2, pp. 317–326, 2003.
[11] L. Wang *et al.*, "Wound image analysis system for diabetics," *Int. Soc. Opt. Photon.*, vol. 8669, p. 866924, 2013.
[12] L. Wang *et al.*, "Smartphone-based wound assessment system for diabetic patients," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 2, pp. 477–488, Feb. 2015.
[13] L. Wang *et al.*, "An automatic assessment system of diabetic foot ulcers based on wound area determination, color segmentation, and healing score evaluation," *J. Diabetes Sci. Technol.*, vol. 10, no. 2, pp. 421–428, 2015.
[14] H. Wannous *et al.*, "Enhanced assessment of the wound-healing process by accurate multiview tissue classification," *IEEE Trans. Med. Imag.*, vol. 30, no. 2, pp. 315–326, Feb. 2011.
[15] K. Wantanajittikul, "Automatic segmentation and degree identification in burn color images," in *Proc. IEEE Biomed. Eng. Int. Conf.*, 2011, pp. 169–173.
[16] M. Kolesnik and A. Fexa, "Segmentation of wounds in the combined color-texture feature space," *Proc. SPIE*, vol. 5370, pp. 549—556, 2004.
[17] B. Fulkerson *et al.*, "Class segmentation and object localization with superpixel neighborhoods," in *Proc. 2009 IEEE 12th Int. Conf. Comput. Vis.*, 2009.

[18] M. F. Ahmad Fauzi *et al.*, "Computerized segmentation and measurement of chronic wound images," *Comput. Biol. Med.*, vol. 60, pp. 74–85, 2015.
[19] D. Filko *et al.*, "Wita—Application for wound analysis and management," in *Proc. 2010 12th IEEE Int. Conf. e-Health Netw., Appl. Serv.*, 2010, pp. 1–6.
[20] MOWA: Mobile Wound Analyzer, Apr. 1, 2016.
[21] S. A. Weber *et al.*, "Remote wound monitoring of chronic ulcers," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 2, pp. 371–377, Mar. 2010.
[22] Silhouette: The Future of Wound Assessment, Jan. 30, 2016.
[23] R. Mukherjee *et al.*, "Automated tissue classification framework for reproducible chronic wound assessment," *Biomed Res. Int.*, vol. 2014, 2014, Art. no. 851582.
[24] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 603–619, May 2002.
[25] Y. Deng and B. S. Manjunath, "Unsupervised segmentation of color-texture regions in images and video," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 8, pp. 800–810, Aug. 2001.
[26] R. Achanta *et al.*, "SLIC superpixels," EPFL, Lausanne, Switzerland, Tech. Rep. 149300, Jun. 2010.
[27] R. Achanta *et al.*, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal.*, vol. 34, no. 11, pp. 2274–2281, Nov. 2012.
[28] Y. Boykov and G. Funka-Lea, "Graph cuts and efficient N-D image segmentation," *Int. J. Comput. Vis.*, vol. 70, no. 2, pp. 109–131, 2006.
[29] C. M. Christoudias *et al.*, "Synergism in low level vision," *Object Recog. Supported User Interact. Serv. Robots*, vol. 4, no. 2, pp. 150–155, 2002.
[30] M. D. Levine and A. M. Nazif, "Dynamic measurement of computer generated image segmentations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 7, no. 2, pp. 155–164, Feb. 1985.
[31] M. Kolesnik and A. Fexa, "Multi dimensional color histograms for segmentation of wounds in images," in *Proc. Int. Conf. Image Anal. Recog.*, 2005, pp. 1014–1022.
[32] D. Kransner, "Wound care how to use the red-yellow-black system," *Amer. J. Nursing*, vol. 95, no. 5, pp. 44–47, 2011.
[33] T. Leung and J. Malik, "Representing and recognizing the visual appearance of materials using three-dimensional textons," *Int. J. Comput. Vis.*, vol. 43, no. 1, pp. 29–44, 2001.
[34] J. Zhang *et al.*, "Local features and kernels for classification of texture and object categories: A comprehensive study," *Int. J. Comput. Vis.*, vol. 73, no. 2, pp. 213–238, 2007.
[35] S. Lazebnik *et al.*, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, 2006, vol. 2, pp. 2169–2178.
[36] D. Making, "Soft margin SVM," *Advances*, vol. 3, no. 4, pp. 1–4, 2003.
[37] C.-W. Hsu *et al.*, "A practical guide to support vector classification," *BJU Int.*, vol. 101, no. 1, pp. 1396–1400, 2008.
[38] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Min. Knowl. Discovery*, vol. 2, pp. 121–167, 1998.
[39] H. Samet and M. Tamminen, "Efficient component labeling of images of arbitrary dimension represented by linear bintrees," *IEEE Trans. Pattern Mach. Intell.*, vol. 10, no. 4, pp. 579–586, Jul. 1988.
[40] X. He *et al.*, "Multiscale conditional random fields for image labeling," in *Proc. 2004 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, 2004, vol. 2, pp. 695–702.
[41] C. Sutton and A. McCallum, "An introduction to conditional random fields for relational learning," in *Introduction Statistical Relational Learning*. Cambridge, MA, USA: MIT Press, 2007, pp. 93–126.
[42] Y. Boykov *et al.*, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 11, pp. 1222–1239, Nov. 2001.

Authors' photographs and biographies not available at the time of publication.