Journal of Diabetes Science and Technology
2016, Vol. 10(2) 421–428

DOI: 10.1177/1932296815599004
dst.sagepub.com
SAGE

An Automatic Assessment System of Diabetic Foot Ulcers Based on Wound Area Determination, Color Segmentation, and Healing Score Evaluation

Lei Wang, MS[1], Peder C. Pedersen, PhD[1], Diane M. Strong, PhD[2], Bengisu Tulu, PhD[2], Emmanuel Agu, PhD[3], Ron Ignotz, PhD[4], and Qian He, MS[3]

Abstract

Background: For individuals with type 2 diabetes, foot ulcers represent a significant health issue. The aim of this study is to design and evaluate a wound assessment system to help wound clinics assess patients with foot ulcers in a way that complements their current visual examination and manual measurements of their foot ulcers.

Methods: The physical components of the system consist of an image capture box, a smartphone for wound image capture and a laptop for analyzing the wound image. The wound image assessment algorithms calculate the overall wound area, color segmented wound areas, and a healing score, to provide a quantitative assessment of the wound healing status both for a single wound image and comparisons of subsequent images to an initial wound image.

Results: The system was evaluated by assessing foot ulcers for 12 patients in the Wound Clinic at University of Massachusetts Medical School. As performance measures, the Matthews correlation coefficient (MCC) value for the wound area determination algorithm tested on 32 foot ulcer images was .68. The clinical validity of our healing score algorithm relative to the experienced clinicians was measured by Krippendorff's alpha coefficient (KAC) and ranged from .42 to .81.

Conclusion: Our system provides a promising real-time method for wound assessment based on image analysis. Clinical comparisons indicate that the optimized mean-shift-based algorithm is well suited for wound area determination. Clinical evaluation of our healing score algorithm shows its potential to provide clinicians with a quantitative method for evaluating wound healing status.

Keywords

diabetic foot ulcer, tele-medicine, wound image analysis, clinical validation of wound assessment, healing score

Clinicians generally evaluate wounds manually using standardized scales and indices based on visual examination.[1] The main shortcomings of this approach are (1) visual inspection is somewhat subjective and potentially inaccurate for wound area determination and tissue classification, (2) wound examination observations are not always recorded in a consistent format, and (3) manual assessment of a patient's wound and subsequent recording of findings constitute extra clinical workload.[2,3] Quantitative assessment of diabetic foot ulcers at regular time intervals combined with a physician's experienced assessments is likely to provide a more reliable basis for selecting an appropriate treatment regimen.

In previous works, computer-aided methods have been developed[1-5] to determine wound size and perform tissue classification, based on computer vision techniques and with the goal of standardizing wound assessment. Published results indicate that these image-processing-based methods permit the definition of "standard" wound healing and minimize inter- and intraobserver variations. These methods, however suffer from drawbacks including (1) high cost and large physical dimension of the wound imaging equipment, (2) computationally demanding algorithms that do not produce

[1]Electrical and Computer Engineering Department, Worcester Polytechnic Institute, Worcester, MA, USA
[2]School of Business, Worcester Polytechnic Institute, Worcester, MA, USA
[3]Computer Science Department, Worcester Polytechnic Institute, Worcester, MA, USA
[4]Plastic Surgery Department, UMASS Medical School, Worcester, MA, USA

Corresponding Author:
Lei Wang, Electrical and Computer Engineering Department, Worcester Polytechnic Institute, 100 Institute Rd, Worcester, MA 01609 USA.
Email: lwang1@wpi.edu

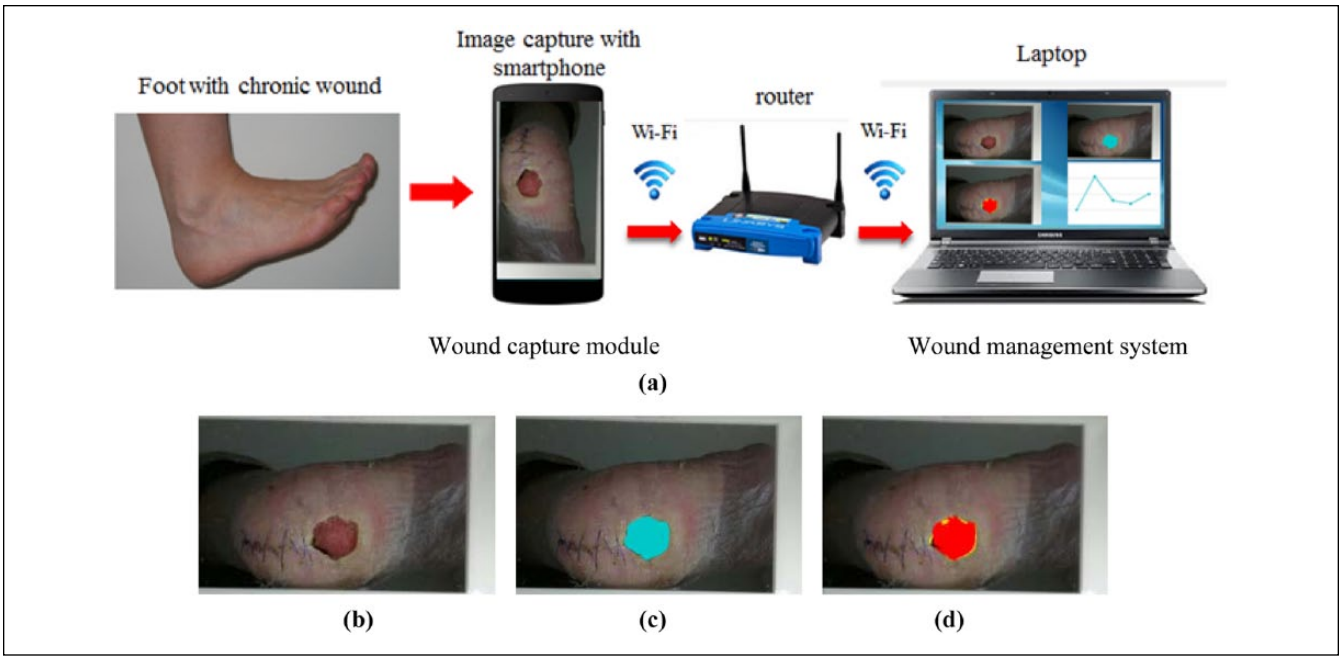

Figure 1. Collaborative system based on a smartphone and a laptop. (a) System overview, composed of a smartphone and a laptop (laptop displays the original wound image, wound image with wound boundary outlined, the image with wound color segmentation and calculates the wound area and healing score). (b)-(d) An example output of our wound analysis for one visit by one patient: (b) the original wound image, (c) the image with the wound area detected, and (d) the image with the color segmentation in the wound area.

results in real time, and (3) lack of wound healing status measures for assessing time sequences of wound images.

This study presents a new real-time wound assessment system for wound clinics, implemented collaboratively on a smartphone and a laptop. The major objectives have been to facilitate foot ulcer image capture and to formulate accurate, yet computationally efficient algorithms for wound area measurement, wound tissue analysis and a score for healing status assessment.

Methods

System Overview

Our collaborative system consists of a smartphone and a laptop, communicating over Wi-Fi via a router in a peer-to-peer mode as shown in Figure 1a. Image capture is performed by the high resolution camera on the smartphone (Nexus 5 smartphone, Android platform) and an image capture box to simplify the photographing of foot ulcers on the sole of the foot, especially for patients with limited mobility, as illustrated in Figure 2a.

Next, the captured image is transmitted via the Wi-Fi network to a laptop (Dell Inspiron 15R) for wound assessment, the results of which are displayed by the wound management system on the laptop. As shown in Figure 1a, clinicians can view the wound healing trend by comparing the current image and assessment results with the corresponding images and results from a patient's past visits to the wound clinic.

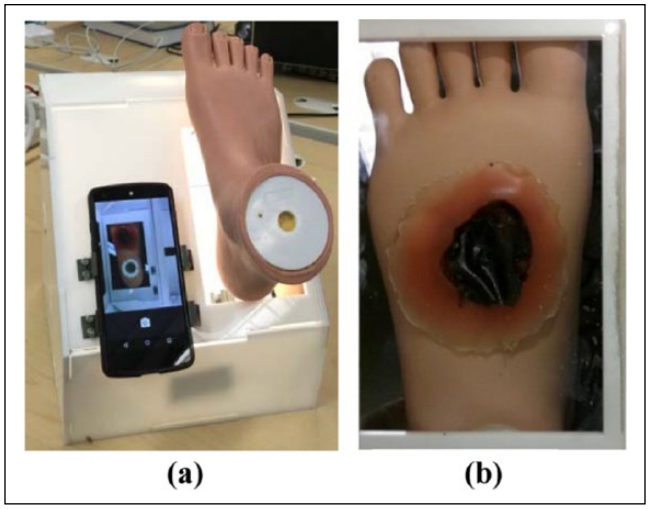

Figure 2. Image capture box: (a) the image capture box with smartphone and foot model; (b) wound image captured using the warm LED light.

Wound Image Capture

Because most foot ulcers occur on the sole of the foot, we utilize an image capture box[6] to facilitate wound image capture as illustrated in Figure 2. The image capture box was designed to be compact, rugged and inexpensive, made from quarter-inch white acrylic. The interior is illuminated with warm white LED light, which provides consistent lighting

for imaging. The slanted surface of the box can be flipped over to accommodate patients with a wound on either the right or left foot. When the patient visits the wound clinic, the nurse helps the patient to place his or her foot on the opening of the slanted surface, at which point the nurse (and the patient) can view the foot ulcer on the smartphone screen. After confirming that the entire wound area is visible, the nurse initiates image capture using a voice command. To protect patients from possible microbial contamination, the foot contact area of the box is cleaned after each use.

Wound Image Management

The foot ulcer image captured on the smartphone is transmitted with the patient's ID (no personal identification information is transmitted for confidentiality) to the laptop over the Wi-Fi connection. The patient ID is used as the unique identifier to add new wound data into the database along with date of image acquisition. On the laptop, the acquired wound image, along with its assessment results (wound area; red, yellow, and black area components; healing score) are saved into a wound image database established on a secure local database server on the laptop. Clinicians with access to the database can view a patient's wound images and healing trend remotely.

Wound Assessment Algorithms

During wound image processing, the wound area is determined by an augmented mean-shift-based algorithm,[7,8] consisting of three modules: (1) mean-shift-based image segmentation, which groups all image pixels into a number of homogeneous regions; (2) a fast method for detecting the largest connected component, to recognize the foot outline; and (3) final wound boundary determination achieved by analyzing the internal and external boundaries of the foot outline. We augmented the algorithm from our previous work,[6] which only detected wounds inside the foot outline, to also incorporate a combined region and boundary algorithm.[9] Based on the wound location, one of two different algorithmic paths is selected, as described in the next paragraph. Following the determination of the wound area, the wound tissue classification is performed based on a red-yellow-black color model.

For wounds enclosed within the foot outline, the largest connected internal region within the foot outline is considered the wound area. For wounds near the boundaries of the foot outline, we designed a turning points detection algorithm to better determine the wound boundary along the foot outline. The input to this algorithm is a group of corner points (prominent structural elements in an image[10]) on the external foot outline. Among these corner points, we identified three turning points, characterized by the greatest direction change, and determine the wound boundary by connecting them into a closed contour, either along the foot outline or by an approximation arc. Next, the image morphology technique called closing is applied to refine the determined wound boundary (Figure 1c). Closing is a common image processing method to fill the small holes and smooth the target boundary.[11]

Within the defined wound area, we identified the wound tissues based on color segmentation, using the well-known red-yellow-black (RYB) wound evaluation model.[12] The color segmentation was implemented using a concept similar to Bag-of-Words,[13] which groups CIE Lab color features for all wound pixels into 3 clusters (read, yellow, and black) by applying the clustering K-Means algorithm.[14] In the RYB model (Figure 1d), red wound tissue indicates healing; yellow tissue indicates infection or slough that is not ready to heal; black tissue indicates necrosis.

Wound Healing Score

To create a measure of wound healing status, we translate the raw wound assessment results into a numerical value called healing score (S_n) using equations 1 to 3. Such a single-valued healing score will provide patients and caregivers with a simple measure of the wound healing or wound deterioration relative to the status at the initial visit. This score can range from 0 to 10. The larger the score is, the better the healing status is. The fundamental principle underlying the healing score design is the RYB evaluation model. The calculation of the healing score is described in the 3 steps below.

Step 1: For each patient, a reference score of 5 is assigned to the wound image collected at the first visit to the wound clinic.

Step 2: At each subsequent visit, the weighed area of the wound is calculated by applying equation 1, where WA_n represents the weighted area of the wound at the nth visit. RA_n, YA_n, and BA_n represent the red, yellow, and black tissue areas, respectively. $[W_R, W_Y, W_B]$ is the vector of weights for red, yellow, and black tissue areas, respectively. From clinical observations, changes in yellow and black tissue areas influence the healing status more than do changes in the red tissue area, which can be expressed as $W_R < W_Y < W_B$. In our case, we empirically determined an appropriate weight vector to be [1, 1.5, 2.5].

$$WA_n = W_R RA_n + W_Y YA_n + W_B BA_n \quad (1)$$

Step 3: Compute a relative healing score using equation 2 to compare WA_n with WA_0, the weighted area for the first wound image of the current patient. The parameter G is an empirically determined gain value, ranging from 0 to 1, to control the dynamic range of the healing score such that our assessment results match clinicians' judgments.

$$S_n = 1 - \frac{WA_n - WA_0}{WA_0} * G = (1+G) - G\frac{WA_n}{WA_0} \quad (2)$$

Figure 3. Software interface screenshot for presenting wound images to clinicians. Clinicians click the "Next Image" button to view the next image for current patient, click the "Next Patient" button to score the images for the next patient, and click the "Next Phase" button to score all images again with different given information.

We find that the gain values of 0.5 and 0.4 provide similarly good results. Choosing G = 0.4, we verified that S_n ranges from 0 to 1.4 if we assume that WA_n is bound by $0 < WA_n < 3.5WA_0$.

To normalize S_n into the range [0, 10], we multiply the expression in equation 2 by 10/1.4. This results in the final formulation of the healing score, as given in equation 3. It is easily verified that the healing score increases from 0 (wound condition is seriously degraded) to 10 (wound is completely healed) as the weighted wound area decreases from its upper bound $(3.5WA_0)$ to 0.

$$S_n = 10 - \frac{2.857WA_n}{WA_0} \quad (3)$$

The healing score is a simple, but useful mathematical construct, which is applicable to other types of chronic wounds, such as venous ulcers, possibly requiring a parameter adjustment.

Clinical Assessment of Wound Area and Healing Score

To establish a clinical basis (ground truth) with which to compare our wound area, three experienced wound clinicians outlined the wound area of the wound images in our database independently, using a tablet computer and an electronic pen. Their delineations for a given wound were combined into one ground truth using a majority vote scheme at the pixel level.[1] To assess the accuracy of the wound area determined by the mean shift algorithm relative to the ground truth, we apply the Matthews correlation coefficient (MCC),[15] which is commonly used for the evaluation of binary classification methods.

To provide clinical validation for our healing score, we asked the same three clinicians to independently score the foot ulcer healing status for each wound image over the range from 0 to 10. A computer-based application was designed to present wound images to clinicians. Only the first image is shown initially, and each click of the "Next Image" button brings up a new image for scoring, while retaining the previous images, as shown in Figure 3.

Furthermore, to ascertain whether the quantitative wound data, in addition to the wound image itself, can improve the clinicians' assessment of the wound, we ask each of the clinicians to score each wound image twice. In the first round, only wound images are presented, so that clinicians' scores are based solely on their visual observations. In the second round, the total wound area and the areas of the red, yellow, and black tissues within the wound boundary are also presented. These two sets of scores from the clinicians are compared to the scores, generated by the healing score algorithm, by calculating the Krippendorff's alpha coefficient (KAC).[16] KAC is a statistical measure of the agreement of ratings

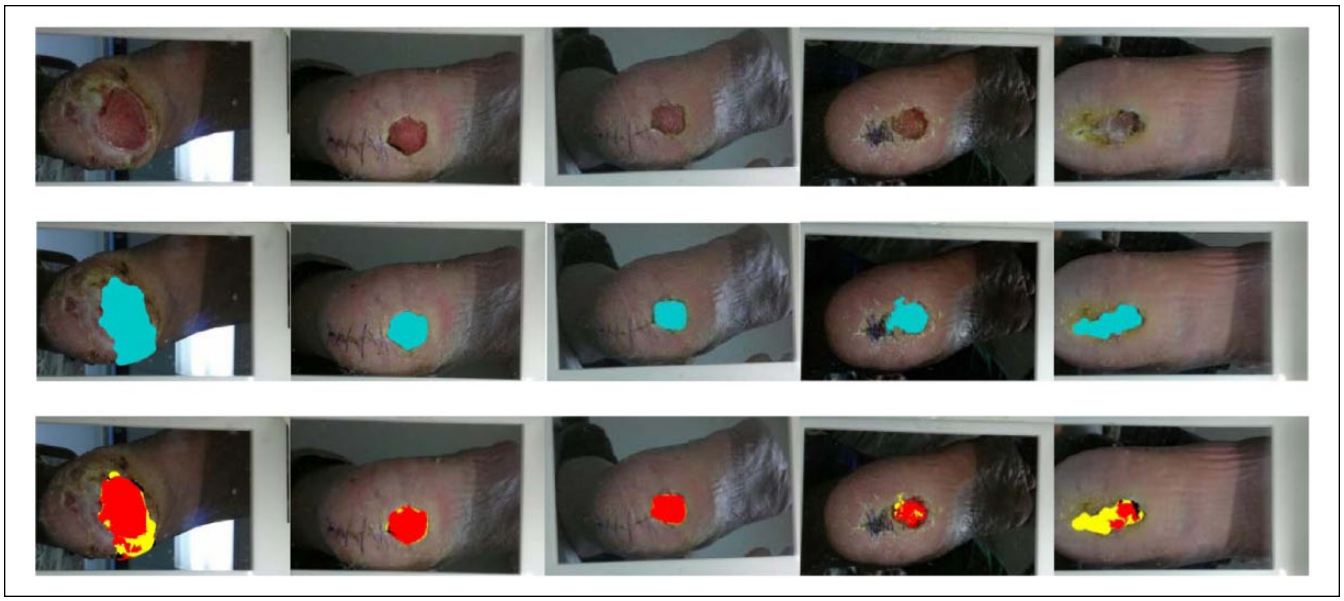

Figure 4. Wound area and tissue classification results for patient 1. Row 1: original foot ulcer images; row 2: wound boundary determination results; row 3: tissue classification results.

given by two or more clinicians. The value of this coefficient ranges in $[-\infty, 1]$, where value 1 indicates perfect agreement and value 0 indicates the absence of agreement. A value less than 0 implies that the disagreements are systematic and exceed what can be expected by chance.[16] The detailed clinical validation results are presented in the "Results" section.

Results

To evaluate our wound assessment method, we have involved 12 patients over a period of one year where each patient was seen over a period ranging from 1 month to 5 months in the Wound Clinic at UMass Medical School, based on an approved institutional review board (IRB) protocol. Among the 12 patients, 9 of them were monitored over at least 2 consecutive visits (2 visits for 3 patients, 3 visits for 4 patients, 4 visits for 1 patient, and 6 visits for 1 patient). In total, 32 foot ulcer images were collected (1 patient, visiting only once, had foot ulcers on both feet) and 28 images were used for the clinical validation of the healing score algorithm.

Clinical Validation Results of the Wound Boundary Determination Algorithm

In Figures 4 and 5, we present the wound area and tissue classification results for two patients, resulting in two time sequences of foot ulcer images. The average MCC value is .68, based on comparison to the wound area delineation by clinicians for the 32 images, which is better than what was obtained with the method used in our previous work (which yielded an average MCC value as .403)[6] and better than the wound recognition method proposed in Wannous et al[1] (with an average MCC value of .45). Since the image capture box is used for acquiring the images, the distance between the foot and the imaging plane is constant. This enables the pixel dimensions from the algorithm measurements to be converted into actual area units (square millimeters in our case) by simply multiplying the pixel dimensions by a constant. The corresponding actual wound area, areas of different wound tissues and the healing scores for the two patients are shown in Tables 1 and 2.

Clinical Validation Results of the Healing Score Algorithm

We utilize KAC to compare the consistency of healing score among the three clinicians (also referred to as "raters"), both for the case where the clinicians are presented with only the wound image and for the case where wound assessment data are also available (Figure 3 illustrates the latter case). The calculated coefficients are referred to as the *interrater consistency coefficients*. The results are shown in the diagonally symmetrical Table 3 where the top number in each cell is the consistency coefficient for wound image only while the bottom number is the coefficient for wound image plus quantitative wound data. We can see that clinicians 1 and 2 have similar assessment about the wound healing status irrespective of whether the quantitative data is presented (KAC > .8 in cell (1, 2)). Clinician 3's assessment differs somewhat from that of the other 2 clinicians (KAC < .5 in cell (1, 3) and the top number in cell (2, 3)). This lack of agreement is also evident in Figure 6, where the red curve (scores from clinician 3) deviates from the green and black curves (scores

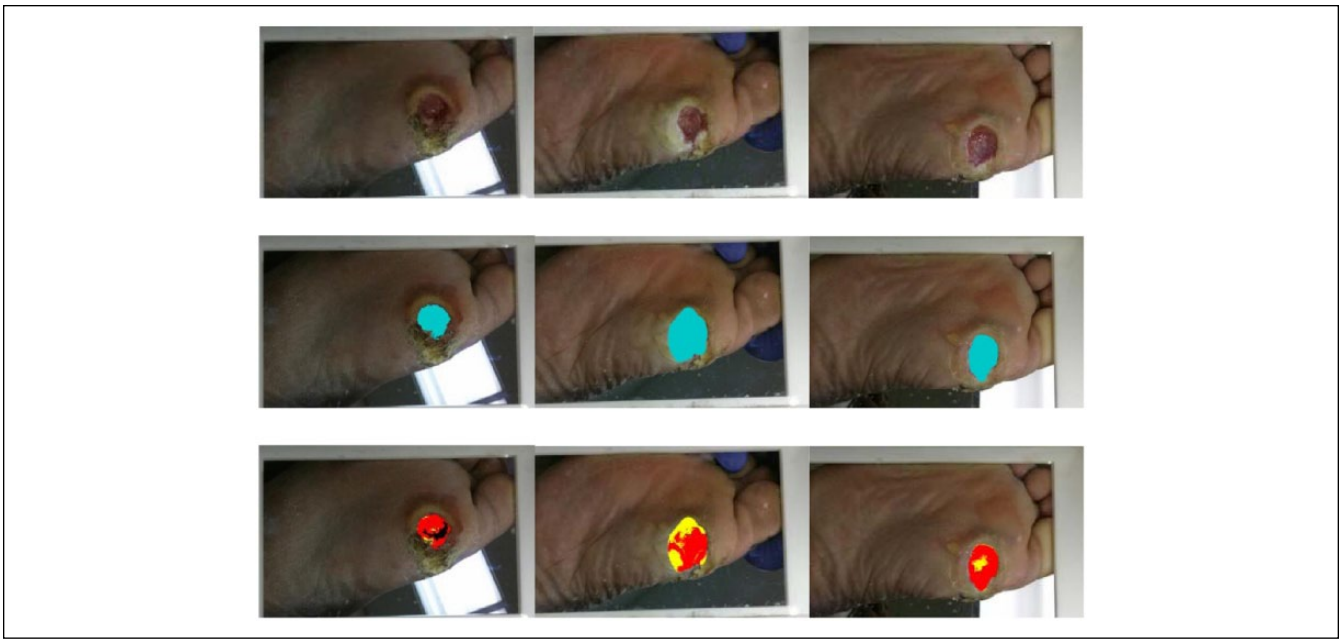

Figure 5. Wound area and tissue classification results for patient 2. Row 1: original foot ulcer images; row 2: wound boundary determination results; row 3: tissue classification results.

Table 1. Wound Assessment Results for Patient 1 (Area Unit: mm^2); for Better Accuracy, 2 Significant Digits for the Healing Scores Are Displayed.

	Image 1	Image 2	Image 3	Image 4	Image 5
Healing score	5 (ref)	7.0	7.4	6.9	7.3
Wound area	928	403	293	279	332
Red area	751	353	283	215	126
Yellow area	158	39	10	43	184
Black area	19	11	0	21	22

Table 2. Wound Assessment Results for Patient 2 (Area Unit: mm^2).

	Image 1	Image 2	Image 3
Healing score	5 (ref)	3.9	5.6
Wound area	249	329	253
Red area	203	247	232
Yellow area	11	82	21
Black area	38	0	0

from clinicians 1 and 2) at several points. Another finding is that clinicians 2 and 3 agree more when the wound quantitative data is also presented (KAC > .6 for the bottom number in cell (2, 3)). Due to our limited number of clinicians and wound samples, our preliminary results indicating that adding wound data can have some influence should be tested with a larger group of clinicians and additional samples.

Table 3. Krippendorff's Alpha Coefficients for the Interrater Consistency Measurements, both for Wound Image Only (Top Values) and for Wound Image Plus Quantitative Wound Data (Bottom Values).

	Clinician 1	Clinician 2	Clinician 3
Clinician 1		.85	.42
		.80	.46
Clinician 2	.85		.46
	.80		.63
Clinician 3	.42	.46	
	.46	.63	

Next, the effect of providing quantitative wound data, in addition to the wound image itself, on the healing scoring of a given clinician (or "rater") is evaluated by determining the KAC between the healing scores with and without the quantitative wound data presented. The evaluation results are reflected in the *intrarater data impact coefficients*. The quantitative wound data consists of healing score, total wound area and area components of red, green and black tissues. The results are given in Table 4 for the three clinicians, showing a modest, but detectable effect (.8 < KAC < .9 for each cell); had there been no effect, KAC would be 1.0. We conclude that adding quantitative data to visual image appears to result in better and/or more consistent assessments, but with our limited set of observations, we cannot generalize as to whether these results would apply in a larger wound sample.

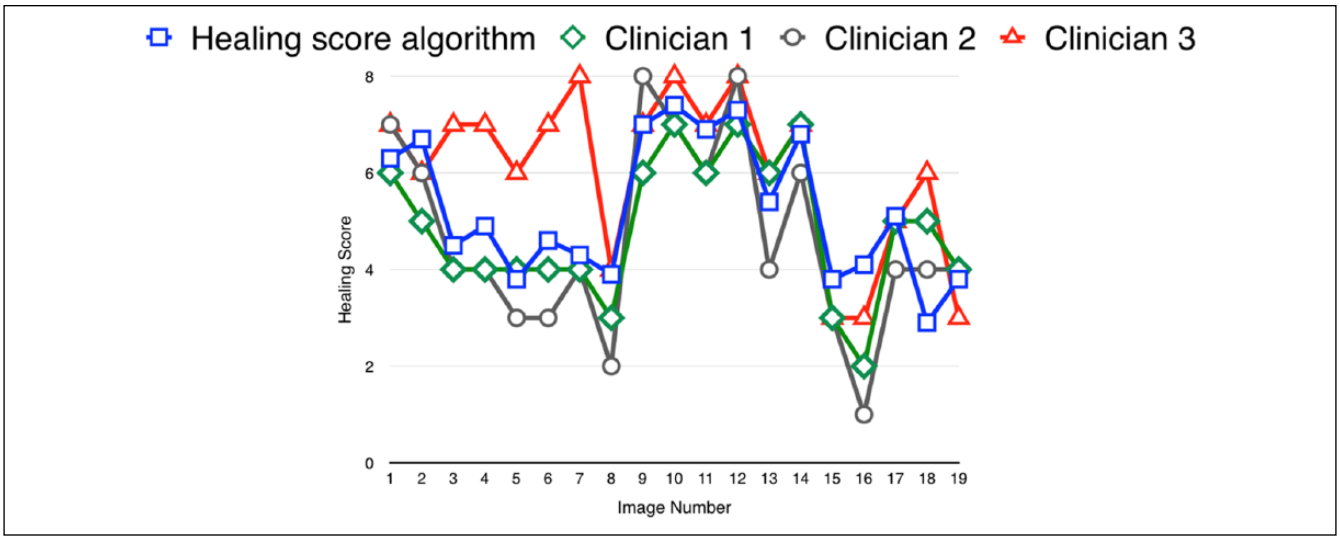

Figure 6. Healing scores by the three raters (green, black, and red for clinicians 1, 2, and 3, respectively), and by the healing score algorithm (blue), for the case where both wound images and wound data are presented.

Table 4. Krippendorff's Alpha Coefficients for the Intrarater Data Impact Measurements.

Clinician number	1	2	3
Intrarater data impact coefficients	.81	.80	.86

Table 5. Krippendorff's Alpha Coefficients for the Clinical Validity Measurements, both for Wound Image Only (Top Values) and for Wound Image Plus Quantitative Wound Data (Bottom Values).

Clinician number	1	2	3
Clinical validity coefficients	.73	.68	.42
	.66	.81	.46

The agreement between the algorithm-based healing score and the clinician-based healing score is measured similarly, using the KAC. The results are given in Table 5, where the measured coefficients are called the *clinical validity coefficients*. As with the interrater consistency coefficients, the results are provided for both the case where the clinicians see only the wound image (top values) and the case where they see both the wound image and the quantitative wound data (bottom data). The values in Table 5 show that our healing score algorithm agrees well with clinician 2 (KAC > .8 especially when quantitative wound data is presented) and has an acceptably good agreement with clinician 1 (KAC > .6). The KAC value for the scoring results from clinician 3 and our algorithm is less than .5, possibly indicating differences in evaluation criteria.

The actual healing scores for 19 wounds (the wound image for each patient's initial visit is the reference image) given by 3 clinicians (for the case where both the wound images and quantitative wound data are presented) and by our algorithm are shown in Figure 6. From this figure, we can see that the scores given by our algorithm are a reliable quantitative indicator of the wound healing trend. Overall our algorithm provides a promising quantitative assessment that approximates well the average score from three clinicians.

Discussion and Conclusion

Reliable foot ulcer measurements that provide good wound healing status assessment are likely to be important for accelerating the healing and for reducing the risk of lower limb amputation for type 2 diabetes patients. The results presented indicate that efficient image processing algorithms and cost-effective image capture devices meet clinical needs. In our prior work, wound assessment was implemented efficiently on a smartphone-alone system,[6] which enables patients to self-manage their foot ulcers and gives the patients' clinicians the possibility for asynchronous evaluation.

As demonstrated in the "Results" section, our wound boundary determination algorithm provides clinically valid results. The average total computing time for the wound boundary determination, color segmentation and healing score calculation algorithm, implemented on a laptop CPU (Intel i5 2.5 GHz), is around 6 seconds for the images in dimension of 816 × 612. It indicates that our wound analysis method is time-efficient enough for a real-time wound assessment implementation.

The use of the image capture box ensures consistent lighting, image distance and quality images, permitting effective wound assessment. The experience from the wound clinic indicates that our image capture box can be used by patients, if necessary with the assistance of a nurse or caregiver.

We believe that our system can readily be integrated into the work-flow of a wound clinic, thanks to compact design and ease of use. It provides data for objective evaluation by recording, for each patient visit, the original and processed wound images (as shown in Figures 1b-1d), along with wound area and healing score. While actual implementation in a wound clinic may take different formats, operational issues were briefly addressed in the "Wound Image Capture" and "Wound Image Management" sections.

The clinical assessment results of our healing score algorithm shows very good agreement with clinicians' scores and thus indicates its strong potential for automated quantitative wound healing assessment. More validation data is needed to further evaluate our algorithm. Unavoidably, healing score determination by clinicians is influenced by experience and training, similar to what has been observed in wound area determination.[17,18] We expect that a future research effort will include information regarding wound texture[1,14] into our wound boundary determination algorithm. Doing so will likely require switching to an algorithm based on machine learning that may be better suited for integrating clinically determined texture information into wound boundary determination for better area determination.

Abbreviations

CPU, central processing unit; IRB, institutional review board; KAC, Krippendorff's alpha coefficient; LED, light-emitting diode; MCC, Matthews correlation coefficient; RYB, red-yellow-black; UMASS, University of Massachusetts.

Acknowledgments

The authors would like to thank the participating clinicians from UMASS Medical School for their help with this study.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This research was supported by National Science Fundation under Grant IIS-1065298.

References

1. Wannous H, Treuillet S, Lucas Y. Robust tissue classification for reproducible wound assessment in telemedicine environment. *J Electron Imaging*. 2010;19(2):023002.
2. Malian A, Azizi A, Van den Heuvel FA, Zolfaghari M. Development of a robust photogrammetric metrology system for monitoring the healing of bedscores. *Photogrammetric Record*. 2005;20(111):241-273.
3. Veredas F, Mesa H, Morente L. Binary tissue classification on wound images with neural networks and Bayesian classifiers. *IEEE Trans Med Imaging*. 2010;29(2):410-427.
4. Plassman P, Jones TD. MAVIS: A non-invasive instrument to measure area and volume of wounds. *Med Eng Phys*. 1998;20(5):332-338.
5. Plassmann P, Jones TD. Improved active contour model with application to measurement of leg ulcers. *J Electron Imaging*. 2003;12(2):317-326.
6. Wang L, Pedersen PC, Strong DM, Tulu B, Agu E, Ignotz R. Smartphone-based wound assessment system for patients with diabetes. *IEEE Trans Biomed Engineer*. 2015;62(2): 477- 487.
7. Comaniciu D, Meer P. Mean shift: a robust approach toward feature space analysis. *IEEE Trans Pattern Anal Machine Intel*. 2002;24(5):603-619.
8. Christoudias CM, Georgescu B, Meer P. Synergism in low level vision. *IEEE Proc 16th Internat Conf Pattern Recog*. 2002;4:150-155.
9. Chang F, Chen CJ. A component—labeling algorithm using contour tracing technique. *Computer Vision, Graphics, and Image Processing: Image Understanding*. 2004;93(2): 206-220.
10. Harris C, Stephens M. A combined corner and edge detector. *Proc 4th Alvey Vision Conf*. 1998:147-151.
11. Gonzalez RC. *Digital Image Processing*. 3rd ed. Upper Saddle River, NJ: Prentice Hall; 2008.
12. Kransner D. Wound care how to use the red-yellow-black system. *Am J Nursing*. 1995;95(5):44-47.
13. Lazebnik S, Schmid C, Ponce J. Beyond bags of features: spatial pyramid matching for recognizing natural scene categories. *Proc IEEE Conf Computer Vision Pattern Recog*. 2006;2: 2169-2178.
14. Hartigan JA, Wong MA. Algorithm AS 136: a K-mean clustering algorithm. *J Royal Stat Soc C*. 1979;28(1):100-108.
15. Matthews BW. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *J Biochimica et Biophysica Acta (BBA)—Protein Structure*. 1975;405(2): 442-451.
16. Hayes AF, Krippendorff K. Answering the call for a standard reliability measure for coding data. *J Communication Methods Measures*. 2007;1(1):77-89.
17. Wunderlich RP, Peters EJG, Armstrong DG, Lavery LA. Reliability of digital videometry and acetate tracing in measuring the surface area of cutaneous wounds. *J Diabetes Res Clin Pract*. 2000;49(2-3):87-92.
18. Rogers LC, Bevilacqua NJ, Armstrong DG, Andros G. Digital planimetry results in more accurate wound measurements: a comparison to standard ruler measurements. *J Diabetes Sci Technol*. 2010;4(4):800-802.