

Received May 27, 2019, accepted June 12, 2019, date of publication June 19, 2019, date of current version July 2, 2019. *Digital Object Identifier 10.1109/ACCESS.2019.2923962*

Chronic Wound Healing Assessment System Based on Different Features Modalities and Non-Negative Matrix Factorization (NMF) Feature Reduction

ASHRAF KHALIL¹, MOHAMMED ELMOG[Y](https://orcid.org/0000-0002-2504-6051)^{®2}, (Member, IEEE), MOHAMMED GHAZ[AL](https://orcid.org/0000-0001-7264-1323)³, (Senior Member, IEEE), CONNOR BURNS²,

AND AYMAN EL-BAZ¹⁹², (Senior Member, IEEE)
¹Computer Science and Information Technology, Abu Dhabi University, Abu Dhabi 59911, UAE

²Department of Bioengineering, University of Louisville, Louisville, KY 40208, USA

³Department of Electrical and Computer Engineering, Abu Dhabi University, Abu Dhabi 59911, UAE

Corresponding author: Ayman El-Baz (aselba01@louisville.edu)

ABSTRACT Identification and treatment of chronic wounds (CWs) are considered economic and social challenges, especially with respect to bedridden and elderly persons. CWs do not follow a predictive course of healing within a particular period. Their treatment and management costs are very high. Also, CWs decrease quality of life for patients, which cause severe pain and discomfort. In this paper, we proposed a comprehensive wound healing assessment framework based on current appearance, texture, and prior visual appearance analysis to handle different types of CWs depending on extracting various tissue types. The framework provided an accurate evaluation tool for the CW healing process depending on extracting and fusing significant features from the CWs RGB images. Non-negative matrix factorization (NMF) was used to retrieve the most significant features to reduce computation time. The gradient boosted trees (GBT) classifier was used to classify different tissue types. Finally, the healing assessment of the CWs depended on calculating the improvement in the area of the necrotic eschar, slough, granulation, and healing epithelial tissues. The framework was trained and tested using 377 RGB images from Medetec wound database and national pressure ulcer advisory panel website. The proposed system achieved an average accuracy of 96% for tissue classification, which helps in obtaining an accurate CW healing assessment. This result can be considered as a promising result when compared to the other state-of-the-art techniques.

INDEX TERMS Wounds and injuries, wound healing, texture analysis, color analysis, tissue segmentation.

I. INTRODUCTION

Awound can be defined as deterioration and injury in the regular anatomical structure and function of the patient's skin. They emerge from various pathological operations that can start either externally or internally in any human organ. Wounds are divided into two main categories depending on healing time, which are acute and chronic wounds [1]. Acute wounds (AWs) emerge as a loss of skin tissue during a surgery or accident, which naturally are repaired in a short time with a

The associate editor coordinating the review of this manuscript and approving it for publication was Wenming Cao.

systematic healing process. However, chronic wounds (CWs) do not follow a systematic healing process and take longer to heal than AWs. The most common types of CWs are pressure ulcers (PUs), venous and arterial ulcers (VAUs), diabetic foot ulcers (DFUs), and burns. In 2010, CWs affected approximately 6.5 million patients in the USA, which cost \$25 billion for the treatment and management annually [2]–[4]. In this paper, we will concentrate on CWs diagnosis and healing assessment.

The first stage in assessing and handling CWs is to determine and recognize the wound's area with its various tissue types. During the wound healing process, the variety

and color of wound tissues are the prominent changes in the CW area. Therefore, describing various wound tissues and tracking the percentage of each tissue are a trusted clinical technique for estimating the wound healing process. In most clinical practices, wound assessment is primarily based on visual inspection by a dermatologist, which is subjective, time-consuming, and potentially error-prone. Therefore, a computer-aided diagnosis (CAD) system is in demand to process and analyze color images of CW. Such a CAD system can be utilized to segment and classify different wound tissues and present an objective, accurate assessment for various types of CWs [5].

Implementing a non-invasive, automatic wound diagnosis and monitoring system has great significance [1]. CAD systems can provide a cost-effective and quantitive solution for precise observation for wound healing state [6]. Furthermore, accurate diagnosis and monitoring of the CWs are crucial for providing effective treatment. The CAD systems can be used to monitor the wound healing by segmenting various tissues and tracking the changes in tissues present in the wound, or its surrounding areas, over time. From this perspective, CW assessment could be based on the analysis of significant regions within different tissue types, each having homogeneous color and texture features. The main tissues of the CWs can be classified into four different types, which are necrotic eschar (black color tissue), slough (yellow/white color tissue), granulation (red color tissue), and healing epithelial tissue (pink color skin) [1]. Fig. [1](#page-1-0) shows the various tissue types in some examples of different CWs.

FIGURE 1. The various tissue types in different CWs: (a) diabetic foot ulcer, (b) venous leg ulcer, (c) pressure ulcer, and (d) burn.

Generally, the wound healing process can be recognized as the tissue color gradually progresses from black to yellow

to red as the wound heals. At first, the wound mostly appears covered with necrotic eschar and slough tissues. Then, the granulation tissue begins to grow slowly from the base of the wound and replaces the necrotic eschar and slough tissues. Also, the granulation tissue begins to fill the wound cavity. Once the wound is filled with granulation tissue, healing tissue begins to grow from the wound edges to color the whole area of the wound. Therefore, monitoring and tracking the changes in each tissue type present a reliable method to assess the wound healing process [4], [7].

The rest of this paper is organized into five sections. Section 2 introduces the current related work in CW diagnosis and wound healing assessment. It also discusses the limitations of the current work and how we overcame these limitations in our proposed framework. In Section 3, the proposed wound healing framework is discussed. The main building blocks of the proposed system are elucidated in more detail. Section 4 describes the used dataset and experimental results. The discussion is presented in Section 5. Finally, Section 6 concludes the paper and discusses the future research directions of the authors.

II. RELATED WORK

The CW diagnosis and monitoring are considered an active research area in the medical image analysis field. There are many research groups that are working in different wound types. For example, Veredas *et al.* [8] introduced a wound area detection system based on statistical color models. They created color histogram models based on the k-means clustering approach for four various wound tissue types, which are granulation, slough, necrosis, and periphery tissues. From a Bayesian perspective, they utilized back-projections of color pixels on the generated histogram models to calculate an estimation of the posterior probability of a pixel to belong to the four tissue classes. The resulting probabilistic model had been complemented with topological models of tissue distribution. Then, Gaussian and morphological filters are used as noise reduction techniques. Finally, they applied the region growing technique with thresholding approaches to get the final region of interest (ROI) of the wound and healing areas. In this paper, the authors only presented a tissue classification system for PU images based on identifying three wound tissues in addition to the periphery tissues.

Hani *et al.* [4] developed a system that detects the beginning of a PU healing by using the hemoglobin content as a marker. They applied the principal component analysis (PCA) technique to whiten the data. They used independent component analysis (ICA) technique to extract gray level hemoglobin from PU images. Then, they implemented a k-means clustering technique to segment detected regions of granulation tissue. In this paper, the authors did provide an effective measure for the healing process for PUs. They detected and segmented the granulation tissue and ignored the other tissue types.

Song and Sacan [3] proposed an automatic image segmentation and wound region identification system for DFU images. They tested four various segmentation techniques, which are k-means, edge detection, thresholding, and region growing. They proposed a parameter optimization procedure to fine-tune each tested segmentation technique automatically. Their proposed system was based on two different artificial neural networks (ANNs) for identification, which are multi-layer perceptron (MLP) and radial basis function (RBF) networks. In this paper, the authors only identified the area of the DFU without further analysis.

Loizou *et al.* [7] presented a system to assess the progress of wound healing for DFUs based on texture analysis. They used the snake approach to segment wound areas. Then, fifteen various texture and four different geometrical features are extracted from the segmented images. The first drawback of this system was that it needed a more robust initialization method for the segmentation technique. Second, the ground truth (GT) of the DFU images was not generated or approved by an expert.

Wang *et al.* [6] proposed a wound analysis system that runs on Android smartphones. They applied the accelerated meanshift algorithm to segment wound images. The healing status is assessed based on the red-yellow-black color evaluation model. They detected the foot outline by finding the largest connected component in the segmented image. In this paper, the authors did not compute a score for the healing status of the wound.

Seixas *et al.* [9] implemented a segmentation approach for skin wound images. Their system is based on proposing an approach to find a seed for the region growing segmentation technique. Also, they utilized the energy of each color channel for the RGB images to enhance the range of the possible values for each tissue type. The main concern of this article was to find a good seed for the wound segmentation technique. The authors did not present any results of the evaluation of their proposed segmentation method.

Wang *et al.* [10] proposed a deep learning technique to segment and analyze the area of the wounds automatically. They proposed a deep convolutional neural network (CNN) to segment the wound area. Also, CNN is used to extract the significant features to detect infection via support vector machines (SVM) technique and predicted the healing process via Gaussian process (GP) regression. In this paper, the authors did not compute a score for the healing status of the wound.

Badea *et al.* [11] proposed a classification system that distinguishes burn wounds from healthy skin based on a CNN network. The proposed CNN was based on the MatConNet architecture. The authors did not make any preprocessing procedure for the tested images.

Wang *et al.* [5] implemented a boundary determination system for DFU images based on a simple linear iterative clustering (SLIC) technique to make superpixel segmentation on a smartphone. Different color and texture features were extracted from segmented images to be supplied to the classifier. They proposed a cascaded two-stage SVM classifier to determine the wound boundaries. Finally, they refined the detected wound boundary by using the conditional random field method. The authors intended to recruit more clinicians to delineate CW boundaries to minimize the effect of intra-/inter-observer differences. Also, they planed to improve the CW classification model to add more flexibility to their system.

Goyal *et al.* [12] proposed an automatic segmentation system for DFU images based on a fully CNN network. They implemented a two-tier transfer learning method by training the fully convolutional network (FCN) for the segmentation of a DFU and its surrounding skin. In this paper, the authors did not determine different pathologies for DFU as a multiclass classification problem. Also, they did not develop a user-friendly interface for their system.

Garcia-Zapirain *et al.* [13], [14] proposed a classification framework for PUs images based on 3D CNN networks, which depended on retrieving and representing various features. They represented these features as different modalities of PU RGB images. These modalities were fused and provided to the 3D CNN with multiple paths to detect PU area and its various tissues. The authors did not discuss the healing assessment of CW. Table [1](#page-3-0) lists a summary of some current related work.

In summary, the basic limitations of the current studies in wound analysis can be formulated in the following points. First, the automatic wound segmentation is not a trivial computational task; this technique has many challenges, such as vague boundaries, irregular shape, heterogeneous tissue colors, and various skin colors [8]. Therefore, a segmentation technique is needed to extract each tissue type from the CW images precisely. Second, most of the CW images are captured by using a regular camera with variable lighting conditions, which affects the wound images' quality [7]. Third, the assessment of the wound healing rate is considered a challenging task that depends on the segmentation of various tissue types [5], [9].

To overcome the limitations mentioned above, we proposed a wound healing assessment framework to analyze various CWs types based on color, texture, and prior visual appearance analysis. Eleven features are extracted to summarize the image contents as well as assess the healing process. These features are five color spaces, local entropy, local range, local standard deviation, gray level co-occurrence matrix (GLCM), scale-invariant feature transform (SIFT), and prior visual appearance. Then, non-negative matrix factorization (NMF) is used to retrieve the most significant features. Tissue classification is done be supplying the extracted features to the gradient boosted trees (GBT) classifier. We classified four different types of tissues, which are eschar, slough, granulation, and healing tissues in addition to the other background skin and objects. Finally, we evaluated the tissue healing rate depending on the changes in the area of each tissue type to represent the wound healing process visually.

TABLE 1. A summary of some current related work (AUC: Area under the curve, ACC: Accuracy, DSC: Dice similarity coefficient, Sens.: Sensitivity, Spec.: Specificity, MCC: Matthews correlation coefficient, and PAD: Percentage area distance).

FIGURE 2. The proposed healing assessment framework for CWs.

III. THE PROPOSED CW HEALING ASSESSMENT FRAMEWORK

The ultimate aim of the proposed system is to develop an objective non-invasive system for CWs healing assessment.

It utilizes eleven different features to extract the main characteristics of the wounds. The proposed system classifies different types of CW tissues from RGB images and assesses the healing process in different types of wounds. Fig. [2](#page-3-1)

shows the architecture of the proposed system that consists of five processing stages. First, the preprocessing stage is implemented to remove noise, eliminate illumination effects, and improve the contrast of the processed images. Second, the feature extraction stage is conducted to retrieve different significant features from CWs images. Third, the feature reduction stage is implemented to reduce the number of processed features, which decreases the computation time of the proposed system. Then, the classification stage is developed to classify different types of wound tissues in addition to the background objects and skin. Finally, the healing assessment stage is implemented to assess the healing process of the images depending on calculating the variation of the area of each extracted tissue type. In the following subsections, the stages of the proposed system will be discussed in detail.

Basic Mathematical Notations: n represents a voxel in the CW RGB colored image. $\mathbb R$ indicates a 3D finite arithmetic lattice of $A \times B \times C$ voxels, where $\mathbb{R} = {\mathbf{n} = (a, b, c)}$: 0 ≤ *a* ≤ *A* − 1, 0 ≤ *b* ≤ *B* − 1, 0 ≤ *c* ≤ *C* − 1}, and *C* is the number of color channels. In a CW RGB colored image, $C =$ 3 and $l \in \mathbb{L} = \{0, 1, \ldots, L-1\}$ presents the integer intensity levels in each RGB color channel, $\mathbf{g} = \{g_{\mathbf{n}} : \mathbf{n} \in \mathbb{R}, g_{\mathbf{n}} \in \mathbb{L}\}.$ \mathbb{S}_{seg} indicates the segmentation map of the CW image, where 0 presents the background objects and skin, 1 presents the necrotic eschar tissue, 2 indicates the granulation tissue, 3 represnts the healing epithelial tissue, and 4 represnts the slough tissue. Finally, $s \in \mathbb{S}_{\text{seg}} = \{0, 1, 2, 3, 4\}$ presents the binary labels for the segmented CW RGB colored images.

A. DATA PREPROCESSING

The preprocessing stage is implemented to enhance the contrast of the RGB images (g_n) in addition to reducing the effect of the light reflection. We applied a contrast-limited adaptive histogram equalization (CLAHE) technique with 8×8 tiles [15]. A bell-shaped histogram was used with distribution parameter $\alpha = 0.4$. CLAHE divides the image into tiles with a predefined size. Then, it computes the contrast transform function for each tile separately. Finally, it combines neighboring tiles by using bilinear interpolation to avoid including artifacts in the boundaries. Therefore, the generated contrastenhanced image (g_{con}) is supplied to the next stage to extract the features from the CW images.

B. FEATURE EXTRACTION

In this stage, we extracted different significant features to classify various tissue types from the contrast-enhanced CW colored images *g*con. The most significant information elicited from the processed images comprises color, texture, SIFT [16], and prior appearance features. These four feature categories are discussed in the following paragraphs.

Color Features: We chose five different color spaces to handle some common issues in the captured RGB images. The used color spaces are the original RGB, normalized RGB, HSV, CIE 1976 L*a*b*, and CIE 1931 XYZ. In addition to the original RGB color space, the normalized RGB color space (*gnRGB*) is calculated to eliminate the distortions

in RGB images that are generated from lights and shadows [17]. It separates chromatic from achromatic information, while it keeps the proportional relationships of the components of the CWs RGB images. The HSV color space (g_{HSV}) is used to remove the effect of the changes in the illumination. The hue (H) component has only the intensity value of the color image that is not affected by illumination changes. Also, the saturation (*S*) component provides a higher contrast to the processed image [18]. The two CIE color spaces (*gLab* and *gXYZ*) are used to get better perceptual accuracy [5]. Also, they can handle the differences in the skin colors [6].

Texture Features: We extracted four different features, which are local entropy (*gent*), local standard deviation (*gstd*), local range (*grng*), and GLCM (*gGLCM*). The preprocessed color images g_{con} were converted to 8-bit grayscale images (g_r) to calculate various texture features. Local entropy of the grayscale CWs image calculates the entropy value for $n \times n$ neighborhood pixels around the processed pixel. We chose a window of size 9×9 to calculate the local entropy. Local range of the grayscale images compute the range values of a 3×3 neighboring pixels around the processed pixel. The range value is calculated by subtracting the maximum value from the minimum value. These values are computed by applying dilation and erosion morphological filters to the processed window. Similarly, the local standard deviation of the grayscale image is calculated by computing the standard deviation of 3 neighboring pixels around the processed pixel.

GLCM is a statistical technique for examining the texture by revealing the spatial distribution of the gray levels in the processed images. It considers the spatial relationship of pixels by computing how frequently a pair of pixels occurs in an image with particular values in a specified spatial relationship. Therefore, the generated GLCM matrix contains the number of times for each pixel values in grayscale with all grayscale levels. Fig. [3](#page-5-0) shows how the GLCM matrix is calculated from the gray image. For our purpose, we used a 256×256 matrix for each tissue type, and a neighborhood system with eight different displacements $\{(a, b + 1), (a - 1)\}$ $1, b + 1$, $(a - 1, b)$, $(a - 1, b - 1)$, $(a, b + 2)$, $(a - 2, b + 1)$ 2), (*a* − 2, *b*), (*a* − 2, *b* − 2)}. We selected 100 images from different types of CWs that had been labeled by three different experts to calculate the GLCM matrices for each tissue type. Finally, we calculated 40 probability values for each pixel in the testing images according to eight various neighborhood distances and five different tissues including the background objects and patient's skin. Algorithm [1](#page-5-1) lists the main steps of creating the GLCM features for CWs images.

SIFT Features: SIFT feature is calculated for each pixel in the processed CWs images to have a scale-invariant representation. First, SIFT searches for stable features in multiple scales (σ) of the processed image using a continuous function of scale. It divides scale space into octaves. In each octave, a Gaussian kernel is convolved with the processed image to produce a set of scale space images.

FIGURE 3. The GLCM matrix calculation.

Algorithm 1 The GLCM Features Generation for CWs Images

```
Data: The labeled RGB images, testing RGB images,
      and neighborhood distances.
```

```
Result: The GLCM tissue matrices and the probability
        values for pixels in the tested RGB images.
```

```
for Labeled RGB Images do
```

```
Smoothing the RGB images and eliminate the noise;
   Generate the equivalent grayscale image;
   foreach Neighborhood Distance do
       foreach Tissue Type do
          Calcuate the GLCM;
       end
   end
end
foreach Testing Image do
   Smooth the RGB image;
   Generate the equivalent grayscale image;
   foreach Pixel do
       Retrieve the GLCM values for each tissue type
       and each neighborhood distance;
       Calculate the probabilities of the pixel;
   end
```
Then, the difference of Gaussian (DOG) is calculated by subtracting adjacent Gaussians. Second, a histogram of local gradient directions for each pixel is calculated at the selected scale. Each pixel's feature represents stable 2D coordinates with scale and orientation. Finally, the descriptor for each pixel is calculated as a 128-element vector. A histogram of gradient directions was calculated for locations in the 16×16 neighborhood of the processed pixel. The most frequent value of the gradient directions in the histogram presents the local

end

pixel orientation. Finally, the features were organized in a 128-elements feature vector that is normalized to make the features independent from intensity variations (g_{SIFT}) [19].

Prior Appearance Feature: The prior appearance map (*gp*) for the CW image is generated based on the prior color information of various tissue types of the CWs. The prior probability of each pixel is calculated based on its color value and Euclidean distance to different tissue classes. The prior appearance feature is computed by applying three essential stages. First, 100 CWs labeled preprocessed RGB images were used to generate the databases (DB_T , $T \in \{1, 2, 3, 4\}$) of four CW tissue types. Each tissue database contains the highest 200 repeated color values. Second, we calculate the Euclidean distance $(D_{n,T})$ between each pixel in the processed images and the generated four tissue databases. Then, the lowest ten values are chosen from each tissue class. Finally, the probabilities of the currently processed pixel with different tissue types are calculated by Eq. [1](#page-5-2) and used as the prior visual appearance feature. All the extracted features are aggregated in one feature vector and supplied to the feature reduction stage. Fig. [4](#page-6-0) shows a graphical representation of the prior appearance feature calculation.

$$
\mathbf{P}_{\rm p} = \frac{N_T}{\sum_{T \in \{1,2,3,4\}} N_T} \tag{1}
$$

where **T** is the type of the CW tissue, which equals 1 for eschar, 2 for granulation, 3 for healing, and 4 for slough tissues. N_T is the total number of pixels of a specific tissue that have a Euclidean distance greater than or equal to a threshold.

After extracting all features, the feature vector for each pixel is generated by concatenating the output from each extracted feature. The generated feature vector has 182 elements, which will be time-consuming if all of these elements will be processed. Therefore, we applied NMF to reduce the number of elements in the feature vectors.

FIGURE 4. A graphical representation for the calculation of the prior appearance feature.

C. NMF-BASED FEATURE REDUCTION

After generating the feature vector for each pixel, it is supplied to the NMF technique to reduce the feature vector dimensionality. NMF is a dimension reduction method, which is based on a low-rank approximation of the feature space, as shown in Eq. [2.](#page-6-1) Also, it guarantees that the features are non-negative by generating additive models [20].

$$
X(i,j) \approx W(i,k)H(k,j) \tag{2}
$$

where *X* is a $i \times j$ non-negative matrix that represents the extracted feature for all pixels in the processed image. K is a positive integer where $k < min(i, j)$. NMF calculates two non-negative matrices, which are $W(i, k)$ and $H(k, j)$. It minimizes the norm of the difference of $X - WH$. *W* can be considered as the reduced features and *H* as the importance of these features. Therefore, the product of *W* and *H* presents a reduced approximation of the data stored in the *X* matrix. To optimize the performance of our proposed system, we tested different values of *k* (from 2 to 100). We found that $k = 10$ gave the best performance among the other values.

D. TISSUE CLASSIFICATION

In this stage, the reduced feature matrix is supplied to a classifier to assign each pixel in the processed CW image a label corresponding to a tissue type. As mentioned previously, we want to detect four various types of tissues, which are eschar, slough, granulation, and healing tissues, in addition to the patient's skin and other background objects. The GBT classifier is used to classify the processed pixels into these five tissue groups. GBT is an ensemble boosting algorithm, which generates a prediction model by using an ensemble weak prediction models. It is based on independent sequential predictors that learn from previous ones. The main objective

of the GBT is to minimize the loss function ($O(o_i, o_i^p)$ $\binom{p}{i}$), which is calculated by the mean square error between the target (o_i) and predicted (o_i^p) i ^{p}) outputs (Eq. [3\)](#page-6-2). Therefore, the new predicted value (o_i^p) $\binom{p}{i_n}$ is calculated based on the previous predicted values and the learning rate (α) (Eq. [4\)](#page-6-2) [21].

$$
O(o_i, o_i^p) = \sum (o_i - o_i^p)^2 \tag{3}
$$

$$
o_{i_n}^p = o_i^p - 2\alpha \sum (o_i - o_i^p)
$$
 (4)

The highest performance is obtained with a number of trees equal to 140 and maximal depth equal to 7 at $\alpha = 0.1$. We compared the results of the GBT classifier with six other state-of-the-art techniques. The GBT classifier outperforms the results of the other tested techniques.

E. HEALING ASSESSMENT

The final phase of our proposed framework is the healing assessment stage. In general, the CW heals as it gradually changes from necrotic eschar (black) to slough (yellow) to granulation (red) to healing (pink) tissues [4]. The healing process is entirely dependent on the initial severity of the CW. In the late grades of CW $(3rd$ and $4th$ grades), it is entirely manifested with necrotic eschar tissue or superimposed layers of the slough and necrotic tissues. As CW progresses in the healing process, the granulation tissue begins to appear and fills the wound cavity. Finally, it transformed into a healing epithelial tissue with pink color, which gradually grows from the wound edges to close it.

We assessed the healing process of the CW by calculating the differences between the areas of each tissue type in two or more consecutive images for the same CW. A histogram of tissue class labels was constructed for each image. Finally, the final healing assessment (H_A) is provided as a cumulative percentage of the healing process in all extracted tissues,

FIGURE 5. Some examples of the output of the assessment process: (a) A PU case and (b) A DFU case.

as shown in Eq. [5.](#page-7-0)

$$
H_A = \frac{(E_1 + S_1 + G_2 + H_2) - (E_2 + S_2 + G_1 + H_1)}{4 * (E_1 + S_1 + G_1 + H_1)} * 100
$$
\n(5)

where E_1 and E_2 are the total necrotic eschar labeled pixels in the first and second CW images, respectively. Similarly, *S*¹ and *S*² indicates the change in the slough tissue, *G*¹ and G_2 presents the change in the granulation tissue, and H_1 and *H*² pointed out to the changes in the healing tissue. Fig. [5](#page-7-1) shows two examples of the healing assessment process, which include the original CW RGB images, segmented images, and the healing assessment histogram. Fig. [5](#page-7-1) (a) presents a PU case, whereas Fig. [5](#page-7-1) (b) presents a DFU case.

IV. EXPERIMENTAL RESULTS

We used 377 CW images to train and test our proposed framework from two different sources. First, the Medetec wound database [22] was used, which contains 341 RGB images of various types of open CWs. Medetec includes 158 images for PUs, 44 images for DFUs, 120 images for VAUs, and 19 images for burns. In addition to Medetec wound database, we obtained 36 PU images from the national pressure ulcer advisory panel website (https://www.npuap.org/). Therefore, we have 377 CW images in total from both datasets for four various types of open CWs.

On the other hand, we had 22 image sets for assessing the healing processing for different types of CWs. Each image set had two or more images of the same CW taken at different periods during the treatment process. All processed images have dimensions from 560×400 to 1024×1024 pixels and were taken by using a regular camera at a 30 cm distance. The GT for different types of tissues was generated by three graduate students. Three dermatologists refined and approved the GT for all images. The majority vote technique is used to choose the final GT images for CWs. The proposed system is implemented by using Matlab 2018a on a Dell workstation with 64 GB RAM and dual Intel Xeon CPU E5-2620V4 quadcore processors at 2.10 GHz.

We utilized three different metrics to evaluate the performance of the proposed healing assessment framework. These metrics are accuracy (ACC), true positive rate (TPR)/ sensitivity, and positive predictive value (PPV)/precision, as listed in Eqs. [6,](#page-7-2) [7,](#page-7-2) and [8](#page-7-2) [23], [24]. ACC indicates the ratio between the correctly classified pixels to the total number of all processed pixels. TPR indicates the correctly classified positive pixels to the total number of all positive pixels. PPV indicates the correctly classified positive pixels to the total number of predicted positive pixels.

$$
ACC = \frac{TP + TN}{TP + TN + FP + FN}
$$
 (6)

FIGURE 6. The resulting ACC by testing different values of the k parameter in NMF technique for different types of CW.

Classifiers	NB	GLM	DL	DT	RF	-SVM	GBT
PU	$88.3 + 0.2$	$87.9 + 0.1$	$93.0 + 0.2$	$90.7 + 0.3$	$90.0 + 0.1$	$94.1 + 0.7$	$94.8 + 0.2$
DFU	$89.2 + 0.2$	$90.6 + 0.1$	$95.4 + 0.1$	$92.7 + 0.0$	$75.8 + 0.3$	$96.3 + 0.3$	$96.3 + 0.1$
VAU	$84.7 + 0.1$	$92.4 + 0.1$	$96.3 + 0.3$	$88.2 + 0.2$	81.4 ± 0.1	$93.4 + 0.4$	95.9 ± 0.1
Burns	$93.5 + 0.0$	$94.9 + 0.0$	$97.0 + 0.1$	$94.4 + 0.2$	$87.9 + 0.2$	95.7 ± 0.3	$97.1 + 0.2$
Average	89.1 ± 0.1	91.9 ± 0.1	95.7 ± 0.2	$91.9 + 0.1$	$83.7 + 0.2$	94.8 ± 0.3	$96.2 + 0.1$

TABLE 3. The accuracy (%) of different types of tested CWs images by using NMF (k = 10) and holdout validation (40% for testing).

$$
TPR = \frac{TP}{TP + FN} \tag{7}
$$

$$
PPV = \frac{TP}{TP + FP}
$$
 (8)

where *TP* presents the true positive value, *TN* presents the true negative value, *FP* presents the false positive value, and *FN* presents the false negative value.

For NMF feature reduction, we tested many values for k ($k = 2, 4, \ldots, 100$) to achieve the highest performance. Here, *k* represents the number of elements in the feature vectors. We found that $k = 10$ gives us the best results for all types of CWs, which reduces feature vector dimensions from 182 to 10 elements. Fig. [6](#page-8-0) shows the resulting ACC with respect to choosing different values of the *k* parameter in NMF technique for different types of CW.

To evaluate the performance of our system, we calculated the overall ACC of the classification of the CW tissues and compared the results with six different state-of-the-art classifiers, which are naïve Bayes (NB), deep learning (DL),

generalized linear model (GLM), Random Forest (RF), decision tree (DT), and support vector machine (SVM). All the results are obtained by using 10-fold cross-validation and holdout validation (60% for training and 40% for testing) to avoid overfitting. Table [2](#page-8-1) lists the results of the average accuracy of each CW type for 10-fold cross-validation. The proposed system achieved an average overall accuracy of 96.2% for all types of tested CW images. On the other hand, NB, GLM, DL, DT, RF, and SVM achieved 89.1%, 91.9%, 95.7%, 91.9%, 83.7%, and 94.8%, respectively. Table [3](#page-8-2) lists the results of the average accuracy of each CW type for holdout validation (40% for testing). The proposed system achieved an average overall accuracy of 96.1% for all types of tested CW images. On the other hand, NB, GLM, DL, DT, RF, and SVM achieved 88.9%, 92%, 95.2%, 91.6%, 84.8%, and 94.8%, respectively. The results show that our proposed system, which is based on the GBT technique, outperforms other state-of-the-art techniques. Also, all the listed results in both tables are done by using $k = 10$ for

FIGURE 7. The average computation time for processing one CW image by using different classifiers.

TABLE 4. The confusion matrix of the proposed system based on GBT classifier for the PU images with holdout validation technique.

Tissues	Eschar	Slough	Granulation	Healing
Eschar	10769	272	47	33
Slough	407	53796	4695	45
Granulation	15	2619	31564	2095
Healing	23	27	1203	29275

NMF technique. Table [4](#page-9-0) shows an example of the confusion matrix of the proposed system by using GBT classifier for the PU images with holdout validation technique.

For the computation time, we measured the elapsed time for each tested classifier. Fig. [7](#page-9-1) shows the average consumed time in ms for all tested classifiers on PU, DFU, VAU, and burn images. The results shown represent the average time to process a single image. The GBT did not present the smallest time as it consumes on average less than 5 seconds for processing only one CW image. We chose GBT to be the classifier of the proposed CW healing assessment framework because it achieved the highest accuracy as compared to the state-of-the-art classifiers with reasonable computation time.

Second, we calculated the TPR and PPV for each CW tissue type. Table [5](#page-9-2) lists the results of both PPV and TPR for each tissue types for different CW types by using the same value of $k = 10$ for NMF and with 10-fold cross-validation technique. The table shows the capability of the system to distinguish between different types of tissues.

On the other hand, we evaluated the performance of our proposed tissue classification stage by comparing its results with three different state-of-the-art image segmentation methods. These segmentation techniques are the 3D

TABLE 5. The PPV (%) and TPR (%) for different tissue types by using GBT and NMF $(k = 10)$.

Types	Metrics	Others	Eschar	Slough	Gran.	Heal.
PU	PPV	98.74	88.45	96.05	91.34	87.01
	TPR	99.55	75.21	93.32	94.90	84.26
DFU	PPV	96.42	91.76	95.84	98.59	95.34
	TPR	97.52	80.54	96.23	96.00	92.37
VAU	PPV	95.30	92.86	96.93	97.79	
	TPR	98.89	74.72	98.64	95.05	
Burns	PPV	97.51	91.23	100	100	
	TPR	98.72	84.08	100	100	

DeepMedic CNN [25], Fuzzy C-Means (FCM) [26], and Otsu [27] segmentation method. The DeepMedic CNN is developed initially to segment a 3D medical image volume. In our case, it is used to segment the colored RGB CW images into four different tissue types in addition to the background skin. The RGB CW image is supplied as the input to the CNN. The CNN network consists of only one path with 11 layers. It has 8 convolutional layers (2 layers of 30×21^3 , 2 layers of 40×17^3 , 2 layers of 40×13^3 , and 2 layers of 50×9^3), 2 fully connected layers (150 \times 9³), and 1 classification layer (2 \times 9³). Its learning rate equals 0.001, momentum rate equals 0.7, and regularization coefficients equal $L_1 = 0.000001$ and $L_2 =$ 0.0001. The tissue classification stage is evaluated by using two common metrics, which are Dice similarity coefficient (DSC) and area under the curve (AUC). The DSC calculates the relevant correspondence between two areas with respect to their true/false positive and negative values (Eq. [9\)](#page-10-0). The AUC presents the expectations of a uniformly drawn random positive is ranked before a uniformly drawn random negative (Eq. [10\)](#page-10-0). The higher the value of DSC and AUC, the more

accurate the tissue classification stage. Table [6](#page-10-1) lists the comparison results of the proposed tissue classification stage and the other state-of-the-art segmentation techniques.

$$
DSC = \frac{2TP}{2TP + FP + FN} \tag{9}
$$

$$
AUC = 0.5 \left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP} \right) \tag{10}
$$

As stated previously, three various dermatologists refined and approved the GT for all CW images. To show the consistency of each generated GT to the other two, the Bland-Altman technique [28] is utilized to compute the agreement degree between the resulting GTs. To have a good agreement, the mean difference (bias) between two GTs of two dermatologists is near zero. Also, most of the data points should fall within 95% limits of agreement with the ± 1.96 standard deviation (SD). This statistical technique is calculated for each CW tissue by two various dermatologists. The Bland-Altman analysis confirmed the robustness of the GTs of the three dermatologists.

Finally, we used 22 image sets for assessing the healing process for different types of CWs by drawing the progress of healing histogram. Fig. [5](#page-7-1) shows the healing assessment histogram for two different image sets. Fig. [5](#page-7-1) (a) shows a PU image set that consists of two consecutive images. It is noticed that the slough tissue is completely disappeared from the second image. Also, the areas of the granulation and healing tissues are increased. The overall healing ratio is 48.67% as it is calculated by averaging all healing ratios. Fig. [5](#page-7-1) (b) shows an example of a DFU image set with a healing ratio equaling 61.75. We calculated the healing score for all 22 images sets and compared the results with the actual clinical diagnosis, which are provided by three dermatologists. The proposed CW healing assessment system achieved an accuracy of 95% as compared with actual clinical diagnosis, which shows promising results.

V. CONCLUSION

In this paper, we proposed a framework for wound healing assessment, which was based on identifying significant features of different types of CWs. The main idea of the proposed system was to segment a wound into various tissue types in order to calculate the healing ratio of the wound. NMF technique is used to retrieve the most significant features to reduce computation time. Then, GBT classifier is used to classify different tissue types. Finally, the healing assessment of the CWs is calculated, which depends on calculating the improvement in the areas of different CW tissues. By fusing significant features, the proposed framework proved an accurate tool for evaluation of the CW healing process. The proposed system achieved an average accuracy of 96% in the tissue classification stage, which is promising as compared with other state-of-the-art techniques. In future work, we will try to predict the grades of CWs by analyzing the image contents. Also, we will expand our image database to evaluate the assessment process comprehensively.

ACKNOWLEDGMENT

(Ashraf Khalil and Mohammed Elmogy contributed equally to this work.)

REFERENCES

- [1] R. Mukherjee, S. Tewary, and A. Routray, ''Diagnostic and prognostic utility of non-invasive multimodal imaging in chronic wound monitoring: A systematic review,'' *J. Med. Syst.*, vol. 41, no. 3, p. 46, Feb. 2017.
- [2] C. K. Sen, Gayle M. Gordillo, S. Roy, R. Kirsner, L. Lambert, T. K. Hunt, F. Gottrup, G. C. Gurtner, and M. T. Longaker, ''Human skin wounds: A major and snowballing threat to public health and the economy,'' *Wound Repair Regener.*, vol. 17, no. 6, pp. 763–771, 2009.
- [3] B. Song and A. Sacan, ''Automated wound identification system based on image segmentation and artificial neural networks,'' in *Proc. IEEE Int. Conf. Bioinf. Biomed.*, Oct. 2012, pp. 1–4.
- [4] A. F. M. Hani, L. Arshad, A. S. Malik, A. Jamil, and F. Y. B. Bin, ''Haemoglobin distribution in ulcers for healing assessment,'' in *Proc. 4th Int. Conf. Intell. Adv. Syst. (ICIAS)*, vol. 1, Jun. 2012, pp. 362–367.
- [5] L. Wang, P. C. Pedersen, E. Agu, D. M. Strong, and B. Tulu, ''Area determination of diabetic foot ulcer images using a cascaded two-stage SVM-based classification,'' *IEEE Trans. Biomed. Eng.*, vol. 64, no. 9, pp. 2098–2109, Sep. 2017.
- [6] L. Wang, P. C. Pedersen, D. M. Strong, B. Tulu, E. Agu, and R. Ignotz, ''Smartphone-based wound assessment system for patients with diabetes,'' *IEEE Trans. Biomed. Eng.*, vol. 62, no. 2, pp. 477–488, Feb. 2015.
- [7] C. P. Loizou, T. Kasparis, O. Mitsi, and M. Polyviou, ''Evaluation of wound healing process based on texture analysis,'' in *Proc. IEEE 12th Int. Conf. Bioinf. Bioeng. (BIBE)*, Nov. 2012, pp. 709–714.
- [8] F. J. Veredas, H. Mesa, and L. Morente, "Efficient detection of woundbed and peripheral skin with statistical colour models,'' *Med. Biol. Eng. Comput.*, vol. 53, no. 4, pp. 345–359, Apr. 2015.
- [9] J. L. Seixas, S. Barbon, C. M. Siqueira, I. F. L. Dias, A. G. Castaldin, and A. S. Felinto, ''Color energy as a seed descriptor for image segmentation with region growing algorithms on skin wound images,'' in *Proc. IEEE 16th Int. Conf. e-Health Netw., Appl. Services (Healthcom)*, Oct. 2014, pp. 387–392.
- [10] C. Wang, X. Yan, M. Smith, K. Kochhar, M. Rubin, S. M. Warren, J. Wrobel, and H. Lee, ''A unified framework for automatic wound segmentation and analysis with deep convolutional neural networks,'' in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 2415–2418.
- [11] M.-S. Badea, C. Vertan, C. Florea, L. Florea, and S. Bădoiu, ''Automatic burn area identification in color images,'' in *Proc. Int. Conf. Commun. (COMM)*, Jun. 2016, pp. 65–68.
- [12] M. Goyal, M. H. Yap, N. D. Reeves, S. Rajbhandari, and J. Spragg, "Fully convolutional networks for diabetic foot ulcer segmentation,'' in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2017, pp. 618–623.
- [13] B. García-Zapirain, M. Elmogy, A. El-Baz, and A. S. Elmaghraby, "Classification of pressure ulcer tissues with 3D convolutional neural network,'' *Med. Biol. Eng. Comput.*, vol. 56, no. 12, pp. 2245–2258, Dec. 2018.
- [14] M. Elmogy, B. García-Zapirain, C. Burns, A. Elmaghraby, and A. El-Baz, ''Tissues classification for pressure ulcer images based on 3D convolutional neural network,'' in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 3139–3143.
- [15] G. Yadav, S. Maheshwari, and A. Agarwal, "Contrast limited adaptive histogram equalization based enhancement for real time video system." in *Proc. Int. Conf. Adv. Comput., Commun. Inform. (ICACCI)*, Sep. 2014, pp. 2392–2397.
- [16] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
- [17] M. Loesdau, S. Chabrier, and A. Gabillon, ''Chromatic indices in the normalized RGB color space,'' in *Proc. Int. Conf. Digit. Image Comput., Techn. Appl. (DICTA)*, Nov./Dec. 2017, pp. 1–8.
- [18] R. Mukherjee, D. D. Manohar, D. K. Das, A. Achar, A. Mitra, and C. Chakraborty, ''Automated tissue classification framework for reproducible chronic wound assessment,'' *Biomed. Res. Int.*, vol. 2014, pp. 1–9, Jul. 2014. [Online]. Available: https://www. hindawi.com/journals/bmri/2014/851582/
- [19] K. D. Toennies, *Guide to Medical Image Analysis: Methods and Algorithms* (Advances in Computer Vision and Pattern Recognition), 2nd ed. London, U.K.: Springer-Verlag, 2017.
- [20] N. Gillis, ''Regularization, optimization, kernels, and support vector machines,'' in *The Why and How of Nonnegative Matrix Factorization* (Machine Learning and Pattern Recognition Series). London, U.K.: Chapman & Hall, 2014, pp. 257–291.
- [21] M. Kuhn and K. Johnson, *Applied Predictive Modeling*, 2nd ed. Berlin, Germany: Springer, 2018.
- [22] S. Thomas. (2019). *Medetec Wound Database*. [Online]. Available: http://www.medetec.co.uk/files/medetec-image-databases.html
- [23] D. M. W. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation,'' *J. Mach. Learn. Technol.*, vol. 2, no. 1, pp. 37–63, 2011.
- [24] A. Tharwat, ''Classification assessment methods,'' *Appl. Comput. Inform.*, to be published. [Online]. Available: https://www. sciencedirect.com/science/article/pii/S2210832718301546. doi: [10.1016/](http://dx.doi.org/10.1016/j.aci.2018.08.003) [j.aci.2018.08.003.](http://dx.doi.org/10.1016/j.aci.2018.08.003)
- [25] K. Kamnitsas, C. Ledig, V. F. J. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, and B. Glocker, ''Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation,'' *Med. Image Anal.*, vol. 36, pp. 61–78, Feb. 2017.
- [26] Y. Yang and S. Huang, ''Image segmentation by fuzzy C-means clustering algorithm with a novel penalty term,'' *Comput. Inform.*, vol. 26, no. 1, pp. 17–31, 2007.
- [27] N. Otsu, ''A threshold selection method from gray-level histograms,'' *IEEE Trans. Syst., Man, Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979.
- [28] J. M. Bland and D. G. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement,'' *Lancet*, vol. 327, no. 8476, pp. 307–310, 1986.

MOHAMMED ELMOGY (M'08) received the B.Sc. and M.Sc. degrees from the Faculty of Engineering, Mansoura University, Mansoura, Egypt, and the Ph.D. degree from Informatics Department, MIN Faculty, Hamburg University, Hamburg, Germany, in 2010. He is an Associate Professor with Information Technology Department, Faculty of Computers and Information, Mansoura University, Egypt. He is currently a Visiting Researcher with Bioengineering Department,

University of Louisville, Louisville, USA. He has authored/coauthored over 150 research publications in peer-reviewed reputed journals, book chapters, and conference proceedings. He advised more than 25 master and doctoral graduates. His current research interests are computer vision, medical image analysis, machine learning, pattern recognition, and biomedical engineering. He has served as a Reviewer for various international journals. He served as a technical program committee member in many workshops and conferences. He is a member of ACM societies.

MOHAMMED GHAZAL received the B.Sc. degree in computer engineering from the American University of Sharjah (AUS), in 2004, and the M.A.Sc. and Ph.D. degrees in electrical and computer engineering (ECE) from Concordia University, Montreal, Canada, in 2010 and 2006, respectively. He is currently an Associate Professor and the Chair of ECE with Abu Dhabi University and an Adjunct Associate Professor of bioengineering with the University of Louisville. He has authored

or co-authored over 40 publications in recognized international journals and conferences including the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON CONSUMER ELECTRONICS, *Renewable Energy Reviews* (Elsevier), and *Multimedia Tools and Applications* (Springer). His research interests include bioengineering, image and video processing, and smart systems. He was a recipient of multiple awards including the Distinguished Faculty Award of Abu Dhabi University, in 2017 and 2014, a Special Innovation Award of AED 1 000 000 by the Office of Prime Minister of the UAE, the NSERC's Alexander Graham Bill's Scholarship, AUS's Presidents Cup, and the Ministry of Education Shield for Creative Thinking. He is a member of ACM and BMES.

CONNOR BURNS is currently pursuing the under graduation degree in bioengineering with the University of Louisville. He has worked as a Bioengineering Intern with Cook Medical, Bloomington, IN, USA, and has coauthored a previous research publication in bioimage processing. He plans to attain a master's degree in bioengineering. He is a Grawemeyer Scholar, which is a prestigious research scholarship given from the University of Louisville.

AYMAN EL-BAZ received the bachelor's and master's degrees in electrical engineering, in 1997 and 2001, respectively, and the Ph.D. degree in electrical engineering from the University of Louisville, in 2006. He is a Professor with University Scholar, and the Chair of the Bioengineering Department, University of Louisville, KY. In 2009, he was named as a Coulter Fellow for his contributions to the field of biomedical translational research. He has 15 years of hands-on experience

in the fields of bio-imaging modeling and non-invasive computer-assisted diagnosis systems. He has authored or coauthored more than 450 technical articles (105 journals, 15 books, 50 book chapters, 175 refereed-conference papers, 100 abstracts, and 15 U.S. patents).

ASHRAF KHALIL is the Director of Research with Abu Dhabi University and an Associate Professor of computer science. His research interests include ubiquitous computing, social and mobile computing, persuasive computing, human–computer interaction, and applying diverse technological innovations in addressing pertinent problems. The results of his research studies are published in the top conferences in the field, such as CHI, CSCW, and INTERACT. His work has also been featured

in public media including the UAE's The National, the USA's National Public Radio, and Slashdot. Since, he moved to the UAE in 2006, he is an active member of the Gulf community of academicians. During his research career, he has accumulated vast knowledge in applying and developing usability testing techniques to mobile applications. He believes technology should not be confined to the work environment but should be woven into every aspect of our daily lives. He was a recipient of many awards and research grants including a second place in the national ''Made in UAE'' competition. He serves in the Steering Committee of the Mobile Application Contest (MAC) and numerous conference program committees.