Hyperspectral index-based metric for burn depth assessment

Sorin Viorel Parasca,1,2 Mihaela Antonina Calin,3,* Dragos Manea,3 Sorin Miclos,3 and Roxana Savastru3

1Plastic and Reconstructive Surgery Department, Carol Davila University of Medicine and Pharmacy, 37 Dionisie Lupu Street, Bucharest 020022, Romania
2Emergency Clinical Hospital for Plastic, Reconstructive Surgery and Burns, 218 Grivitei Street, Bucharest, Romania
3National Institute of Research and Development for Optoelectronics INOE 2000, 409 Atomistilor Street, P.O. Box MG5, Magurele, Ilfov 077125, Romania

* mantonina_calin@yahoo.com

Abstract: Burn depth objective classification is of paramount importance for decision making and treatment. Despite the wide variety of burn depth assessment methods tested so far, none of them have gained wide clinical application. Here, we introduce a new approach for burn depth assessment based on hyperspectral imaging combined with a spectral index-based technique that exploits specific spectral bands to map skin areas with different burn degrees. The spectral index amplifies the contrast between normal skin and areas with different degrees of burn, taking advantage of the differences in spectral amplitudes that occur as a result of the morphological and physiological changes occurring in burned skin. We demonstrate that by using the new measurable spectral index, it is possible to generate accurate burn classification maps showing spatial distribution of burn types in the affected body areas, facilitating the decision-making process and prognosis evaluation. The results highlight the potential of the new hyperspectral metric in the field of burn depth classification and its applicability in hospital settings seems promising.

© 2018 Optical Society of America under the terms of the OSA Open Access Publishing Agreement

1. Introduction

Burns are one of the most serious and unpredictable skin injuries caused by a variety of environmental factors (thermal, chemical, electrical). Burns can be of minor, moderate and major severity depending on their size, the depth of damaged skin, age and medical history of the patient, involved anatomical zone and associated injuries [1]. An initial classification of a patient in one of these three classes of severity is the first and most important step that leads to correct treatment decisions and reliable estimates of the outcome. The initial classification of burns can be achieved based on the assessment of burns size or burns depth. Classification of burns based on burns size assessment rises, however, considerable problems of confusion even for those who are experts in this area [2], since the methods commonly used for the calculation of burn size (Wallace rule of nines [3], Lund and Browder chart [4] and palm-size estimation [5]) use different diagrams and charts and include only partial and full thickness injury.

Burn depth assessment is an essential part of burn classification. The method allows the clinicians to classify burns, depending on the depth of the damaged skin, into one of three categories [6]: first-degree burns (superficial burns), second-degree burns (partial-thickness burns: superficial partial-thickness and deep partial thickness burns), and third-degree burns (full-thickness burns). The superficial burn is a wound in which only epidermal layer of the skin (with mean values of thickness ranging from 430 µm to 38 µm [7]) is affected. The superficial partial-thickness burn is a wound in which both epidermal and superficial (papillary) dermal layers of the skin (~200 µm thick [8]) are damaged, whereas the deep
partial thickness burn refers to a wound that extends into deep (reticular) dermis (2-3 mm thick [8]). The full-thickness burn involves both epidermal and dermal skin layer. In a clinical situation, the early and accurate classification of a burn into one of these categories can be made with some difficulties due to mainly the evolution of burn over time into deeper skin layers and the inability of current methods for burn depth assessment to distinguish clearly one category from other (i.e. deep partial thickness from full thickness).

At present, the most widely used method for burn depth assessment is physical examination [9]. The accuracy of burns classification by physical examination was shown to be low (60 - 75%), being dependent on the experience of the clinicians [10]. The limitations of physical examination have motivated the necessity for the development of invasive or non-invasive paraclinical methods. Currently four categories of methods for depth burn assessment are now either in use or under development [11,12]: (1) optical methods (indocyanine green dye fluorescence imaging [13], laser Doppler imaging [14], infrared thermal imaging [15], spatial frequency domain imaging [16], near-infrared spectroscopy [17], confocal-laser-scanning microscopy [18], optical coherence tomography [19]), (2) ultrasound methods (ultrasonic pulse-echo method [20], ultrasound B-mode imaging [21] and high-frequency ultrasound method [22]), (3) photoacoustic methods [23] and, (4) nuclear imaging methods (radiolabeled tracers [24] and nuclear magnetic resonance spectroscopy [25]). Despite the variety of new methods, none has gained wide clinical application. Therefore, the development of new and more reliable methods remains still an open research field.

In search for new systems and methods for depth burns assessment, some researchers have paid attention to the multispectral (MSI) and hyperspectral imaging (HSI) techniques, taking into account their potential to classify the spatial/spectral data of burned tissues. Thus, Afromowitz et al. [26] showed that using a red/green/near-infrared multispectral imaging system, they could predict better than the attending physicians which burned areas were expected to heal in less than 3 weeks from injury. Eisenbeiss et al. [27] reported promising results after using a reflection optical multispectral imaging system for the determination of burn depth. In a study conducted by Chin et al. [28] on mice, HSI technique was shown to differentiate among different types of burn depth (intermediate-dermal, deep-dermal, and full-thickness) based on tissue oxygenation parameters (oxyhemoglobin, deoxyhemoglobin, total hemoglobin, or oxygen saturation). Recently, Paluchowski et al. [29] highlighted the role of HSI technique in differentiation of burns of different severity when combined with an unsupervised spectral - spatial segmentation method. In an earlier study [30] we also used HSI technique together with linear spectral unmixing model for burns characterization and proposed a methodology that has the ability to generate maps of injured tissues in burned skin, which can help clinicians to distinguish more accurately the extent of damage.

This paper proposes a new approach for burns depth assessment based on hyperspectral imaging combined with a spectral index-based technique that exploits specific spectral bands to map skin areas with different burn degrees. This approach, providing spatial distribution of burns types in the affected body areas, facilitates the decision-making process and eases the prognosis evaluation.

This main objective was achieved by addressing the following issues: (1) selecting of the appropriate methods for hyperspectral image processing; (2) defining a Skin Burns Spectral Index (SBSI) as a new metric for burn depth assessment based on the spectral properties of injured skin extracted from hyperspectral images; (3) generation of specific burn classification maps; (4) comparing the ability of SBSI-based technique to create burn classification maps with laser Doppler imaging (LDI) technique; (5) establishing future research directions to extend the applicability of the method to other skin pathologies.
2. Materials and methods

2.1 Patients

A total of 16 adult patients (13 men and 3 women), aged between 29 and 72 years with superficial partial-thickness, deep partial thickness and full-thickness burns, admitted in the Emergency Clinical Hospital for Plastic, Reconstructive Surgery and Burns, Bucharest from January 2017 to March 2017, were enrolled in this study. Two of them with a total burn surface area (TBSA) > 20% were excluded. Patients with special medical conditions (sepsis and its complications, inhalation injury, burns on esthetic or functional areas) were also excluded (five patients). Finally, 9 patients (8 men and 1 woman) with burn injuries on extremities and trunk were included in this study. Informed consent was obtained from each patient prior to participation in the study. All procedures performed in this study involving human participants were in accordance with the ethical standards of the Emergency Clinical Hospital for Plastic, Reconstructive Surgery and Burns Research Committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. All patients were examined clinically by physicians and information about age, sex, etiology, burn location and clinical diagnosis were recorded (Table 1).

| Patient | Age (years) | Sex | Etiology | Location | Clinical diagnosis |
|---------|-------------|-----|----------|----------|-------------------|
| 1       | 44          | M   | Hot liquid | Dorsal face of the hand | 2ndA, 2ndB degree burn |
| 2       | 29          | M   | Flame     | Lateral chest | 2ndB, 3rd degree burn |
| 3       | 39          | M   | Flame     | Buttocks and posterior thighs | 2ndA, 2ndB degree burn |
| 4       | 32          | M   | Hot liquid | Posterior leg | 2ndA, 2ndB degree burn |
| 5       | 33          | M   | Hot liquid | Dorsum of the foot | 2ndA, 2ndB degree burn |
| 6       | 36          | M   | Hot liquid | arm        | 2ndA, 2ndB degree burn |
| 7       | 29          | M   | Flame     | Anterior and medial forearm | 2ndA, 2ndB degree burn |
| 8       | 30          | M   | Flame     | arm        | 2ndA, 2ndB degree burn |
| 9       | 72          | W   | Hot liquid | Anterior arm | 2ndB, 3rd degree burn |

All recordings were done between days 3 and 5 after the accident in the local therapy room, after the wounds were cleansed with soap and saline and dried with sterile gauge. No cream or antiseptic solution were applied on the wound or the surrounding healthy skin. Pain medication was administered according to local standards. All the chosen areas were investigated using both hyperspectral imaging and laser Doppler imaging techniques.

2.2 Hyperspectral index-based metric for burn depth assessment

The spectral index-based method proposed in this study consists of four steps: (1) hyperspectral image acquisition; (2) hyperspectral image processing; (3) computing burns spectral index (SBSI) in each pixel from the hyperspectral image and (4) creating a burn classification map (BCM).

2.2.1 Hyperspectral image acquisition

Hyperspectral images of the each burned area were acquired using a push-broom hyperspectral imaging system covering a (380 – 800) nm spectral range. It consisted of an imaging spectrograph (ImSpector V8E, Specim, Oulu, Finland), a 12-bit monochrome camera (DX4, Kappa, Gleichen, Germany), an illumination unit containing two 300 W halogen lamps (OSRAM, Munich, Germany) equipped with diffusion filters (Kaiser Fototechnik GmbH and Co. KG, Buchen, Germany), a single-axis galvanometer scanning mirror system (GVS211, Thorlabs, New Jersey, USA) equipped with a broadband dielectric mirror (E02) with average reflection > 95% in (400 – 750) nm spectral range, a computer for acquisition, processing and analysis of hyperspectral data and a tripod (Manfrotto, Cassola, Italy). The image acquisition process was controlled by the SpectralDAQ data acquisition software (Specim, Oulu, Finland) and processing and analysis of hyperspectral data were performed with ENVI v.5.1
software (Exelix Visual Information Solutions, Boulder, Colorado, USA). This hyperspectral system allows for simultaneous acquisition of 258 spectral bands with a spectral resolution of 1.63 nm. The frequency of the input signal of the mirror scanner system was adjusted to an average of 0.23 Hz to match with the monochrome camera acquisition speed, whose exposure time is 24.2 ms. The halogen lamps were fixed above the burned area from both sides at an angle of about 45° to assure a uniform illumination of the zone of interest. The dimension of the acquired hyperspectral image was 348 x 350 x 258 (spatial x spatial x spectral) pixels. The size of the object pixel was 1.03 mm at a working distance of 0.7 m.

2.2.2 Hyperspectral image processing

Prior to determining the SSBI index, all hyperspectral images were processed to improve their quality by removing different categories of artefacts and aberrations associated with the pushbroom scan [31]. In this respect, the hyperspectral images were first calibrated with dark and white reference images in order to minimize the inherent influences of the dark current in the monochrome camera and spatial nonuniformity of the light intensity on the scene [32]. The dark reference image was acquired by covering the hyperspectral system lens with its opaque cap and turning off the light sources. The image of a white polytetrafluoroethylene (PTFE) reference tile (model WS-2, Avantes, Apeldoorn, Netherlands) with approximately 98% reflectance in spectral range 350-1800 nm was used as the white reference image. The PTFE reference tile was located on the patient’s body in the same illumination conditions as the investigated area. Each acquired hyperspectral image was calibrated by using the following equation [33]:

\[ I_{\text{calibrated}} = \frac{I_{\text{sample}} - I_{\text{dark}}}{I_{\text{white}} - I_{\text{dark}}} \]  

where: \( I_{\text{calibrated}} \) is the calibrated hyperspectral image of burn wound, \( I_{\text{sample}} \) is the original hyperspectral image of the wound, \( I_{\text{dark}} \) is the dark reference image and \( I_{\text{white}} \) is the white reference image.

In the second processing step, redundant elements from each calibrated hyperspectral image were eliminated by selecting a region of interest (ROI) containing only the burned areas using the polygon drawing mode of ENVI v.5.1 software (Exelix Visual Information Solutions, Boulder, Colorado, USA) and masking the background pixels inside the ROI (Fig. 1). All subsequent analyses were performed only on these ROIs.

![Fig. 1. Selection of a region of interest (ROI) from the calibrated hyperspectral image.](image)

2.2.3 Skin Burn Spectral Index (SBSI) development

Spectral indices have been considered over the last 30 years as good candidates for monitoring the state and evolution of vegetation [34], droughts [35,36], floods [37], desertification [38], and vegetation burn severity [39]. The successful use of these spectral indices in agriculture has led us to believe that specific spectral indices defined in form of mathematical combinations of different spectral bands can also be developed for the medical field.
Starting from this hypothesis, in this study we developed a specific spectral index, named Skin Burns Spectral Index (SBSI), intended for precise skin burn classification.

The SBSI index was defined, as a modified version of relative delta normalized burn ratio index proposed by Miller and Thode [39] for satellite-inferred burn severity assessment, by the following expression:

\[
SBSI = \frac{NSI - BSI}{\sqrt{NSI}}
\]  

(2)

where: NSI and BSI are normalized difference indices for normal skin and burned skin respectively. The NSI and BSI indices are defined as follows:

\[
NSI = \frac{R_{NSi} - R_{NSj}}{R_{NSi} + R_{NSj}}
\]  

(3)

\[
BSI = \frac{R_{BSi} - R_{BSj}}{R_{BSi} + R_{BSj}}
\]  

(4)

where \(R_{NSi}, R_{NSj}, R_{BSi}\) and \(R_{BSj}\) are spectral reflectance values of normal skin and burned skin respectively, corresponding to two wavelengths \(i\) and \(j\) in the spectral range (400 - 800) nm.

The combination of wavelengths \(i\) and \(j\) used to calculate the SBSI was derived from an analysis of the normal skin and burned skin reflectance spectra (Fig. 2).

As it can be seen in Fig. 2, the reflectance spectrum of normal skin shows three main peaks located around 525 nm (\(R_{max1} = 0.3071\)), 556 nm (\(R_{max2} = 0.2929\)) and 760 nm (\(R_{max3} = 0.5826\)) respectively. The first and second peaks have relatively the same height (\(\Delta R = R_{max2} - R_{max1} = 0.0058\)). Significant difference can be seen between the second and the third peak (\(\Delta R = 0.2897\)). In the burned skin, the reflection spectrum has the same shape, but the reflectance values increase in all spectral range. The difference between the height of the second peak and the third peak becomes almost twice as great as in normal skin (\(\Delta R = 0.4744\)) and can become an indicator of burn degree. Therefore, taking into account this important difference in the spectral characteristics of normal and burned skin, the wavelengths \(\lambda_i = 556\) nm and \(\lambda_j = 760\) nm were considered as being optimal for calculating the SBSI index. Moreover, these wavelengths correspond to the absorption maxima of deoxyhemoglobin which correlates the above defined index with the pathophysiology of the burned skin. This SBSI index amplifies the contrast between normal skin and areas with different degrees of burn, taking advantage of the differences in spectral amplitudes that occur as a result of morphological and
physiological changes occurring in burned skin. The SBSI calculated in each pixel of the hyperspectral image is displayed as a greyscale image varying from black at the weakest intensity to white at the strongest. Darker pixels indicate higher burn degrees (deeper burns).

2.2.4 Burn classification map (BCM) development

By converting the SBSI image from the continuous scale of burn degree to distinct burn degree classes a burn classification map (BCM) can be created. In order to do that, K-means clustering was used [40] as an unsupervised classification method which can perform partition of the SBSI index data set into K clusters. K-means clustering is an iterative two-steps classification process: (1) calculation of the initial cluster means distributed in the input data set, (2) clustering the pixels into the nearest cluster using a minimum distance method. During each iteration, cluster means are recalculated and the pixels are reassigned accordingly. The iterations are continued until no further reassignments occur, or the preset number of iterations has been completed. In this study, an initial number of six clusters was chosen for clustering the SBSI data set aiming for an easier comparison with blood flow images generated by the laser Doppler imaging system. The iterations number was set at 100. The result of clustering process is a burn classification map (BCM) in which similar pixels are grouped together based on their SBSI index value in six groups that correspond to: 3rd degree burn (classes 1 and 2), 2nd B degree burn (class 3 and 4), 2nd A degree burn (classes 5 and 6). The values of the SBSI associated with each burning class are shown in Table 2. The color coding was set up based on the SBSI values of the six clusters and on burn degree.

Table 2. SBSI index values corresponding to different category of burns

| Class No | Name Class | SBSI index |
|----------|------------|------------|
| 1 [dark blue] | 3rd degree | < -0.15 |
| 2 [blue] | 3rd degree | -0.15 ÷ -0.03 |
| 3 [green] | 2nd B degree | -0.03 ÷ 0.06 |
| 4 [yellow] | 2nd B degree | 0.06 ÷ 0.25 |
| 5 [magenta] | 2nd A degree | 0.25 ÷ 0.35 |
| 6 [red] | 2nd A degree | > 0.35 |

2.3 Laser Doppler imaging (LDI)

A laser Doppler blood flow imaging system (MoorLDI2-BI system, Moor Instruments Ltd. Axminster, Devon, UK) was used in this study for mapping blood flow in the areas of burned skin. This system consists of a scan head, scan controller and a computer, all components being mounted on a mobile unit. The system uses a narrow laser beam (λ = 633 nm, maximum output power = 2.5 mW, Φ = 1 mm) that is directed to the burn areas using a scan mirror. When laser light is scattered by the moving blood cells it undergoes a Doppler frequency shift proportional to the average blood cell rate. The scattered laser light is captured by a detector and converted into an electrical signal. This signal is then processed by the computer using the incorporated MoorLDI software v3.1 (Moor Instruments Ltd. Axminster, Devon, UK) which calculates the blood flow in the burned areas and displays the results as a colour-coded blood flow image. The color coding is done in correspondence with the perfusion unit (PU) values of the different parts of the burn. Thus, the dark blue color is assigned for PU <140 (3rd degree burn), light blue for PU <200 (3rd degree burn), green for PU between 200 and 260 (2ndB degree burn), yellow for PU between 260 and 440 (2ndB degree burn), pink color for PU between (440-600) (2ndA degree burn), and red for PU> 600 (2ndA degree burn).

In this study, the default parameters of MoorLDI2-BI system (bandwidth = 250-15KHz, background threshold = 10 and scan speed = 4 ms/pixel) were used for scanning all burn wounds from a distance of 50 cm with a resolution of 256 × 256 pixels. The blood flow images were used to demonstrate the performances of index based - hyperspectral imaging method in burn depth assessment.
2.4 Accuracy assessment

The accuracy assessment of the classification results produced by SBSI and LDI methods was performed based on the confusion matrix analysis. The confusion matrix is a specific table that allows visualization of the classification results compared to ground truth data. Each row of the matrix represents the pixels in a predicted class while each column represents the ground truth pixels in an actual class. According to the confusion matrix, the overall accuracy (OA), user’s (UA) and producer’s (PA) accuracies and Kappa coefficient (Kc) were computed for each of the two classification methods applied to each individual case. The overall accuracy is expressed as the sum of the number of pixels correctly classified divided by the total number of pixels in the image, the user accuracy indicates the probability that a pixel predicted to be in a certain class really is that class, the producer accuracy represents the probability that a pixel in a given class is correctly classified, and Kappa coefficient indicate the effectiveness of the overall classification.

In order to perform the accuracy assessment of the classification results, two ground-truth data set representatives of the six classes of injured tissues identified in each individual burn wound were selected (an average of 332 pixels per class) directly from the original HSI and LDI images after a visual examination of the images by a clinician. Confusion matrix, OA, UA, PA and Kappa coefficient were calculated using MATLAB software (MathWorks, Inc., Natick, Mass., USA). Finally, the degree of agreement between the proposed SBSI method and LDI method was assessed using the statistical method proposed by Bland and Altman [40].

3. Results

3.1 Burn depth assessment using the SBSI approach

Here, we report the results of the proposed hyperspectral index-based metric in burn classification obtained on a lot of nine patients (8 man and 1 women) with a mean age of 38.22 years (± 13.61) with thermal burns on extremities and trunk. The resulting burn classification maps (BCM) are compared with blood flow images (LDI maps) acquired with a laser Doppler blood flow imaging system.

The results of burn classification in a 47-year old male patient with thoracic burn are shown in Fig. 3. The SBSI image (Fig. 3(a)) shows that the lower the SBSI index, the deeper the burn. Negative values of SBSI index correspond to the skin areas (dark pixels) with full thickness burns or normal skin. The bright pixels (positive SBSI values) in the image correspond to skin areas with superficial and deep partial thickness burns. It probably correlates with deoxyhemoglobin concentration, that has higher concentrations in more superficial burns [29]. It is noticeable that the central area of the burn in Fig. 3 looks like a 3rd degree burn on the digital photo (Fig. 3(d)) but proved to be a 2nd B on both blood flow laser Doppler image (Fig. 3(c)) and SBSI index image (Fig. 3(a)).

For easier interpretation of the SBSI image, the continuous value range of SBSI index was divided into six subdomains corresponding to the different categories of burns using the K-means clustering method. The resulting burn classification map (BCM) is shown in Fig. 3(b). The colors were assigned to the map as follows: dark blue/blue for areas with full thickness and normal skin (classes 1 and 2), green/yellow for areas with deep dermal (classes 3 and 4), magenta/red for areas with superficial dermal (classes 5 and 6).
The distribution of the SBSI index among the six burns classes in this particular case ranged from −0.29 to −0.15 (mean = −0.18), −0.15 to −0.03 (mean = −0.08), −0.30 to 0.06 (mean = 0.015), 0.06 to 0.25 (mean = 0.12), 0.25 to 0.35 (mean = 0.28), and 0.35 to 0.45 (mean = 0.38), respectively (Fig. 3(e)). It is noted that only classes 2 and 3 seems to have a symmetric distribution of SBSI, the median being roughly in the middle of the box. These classes, corresponding to deep or deep partial thickness burns, have the most homogenous distribution of the spectral index in this case. For other classes, the corresponding SBSI values are skewed, most of the data being located on the high (classes 4, 5 and 6) or low
(class 1) side of the box. This means that the higher values of the SBSI are more spread out than the lower values for classes 4, 5 and 6. These classes stand for more superficial burns, where the variability of deoxyhemoglobin distribution seems to be higher. For class 1, the higher values of the SBSI index are closer together, so these data are more condensed than the lower values. The variability in SBSI index for each burn class, measured by the interquartile interval (IQR), was: 0.032, 0.052, 0.045, 0.065, 0.045 and 0.032 respectively. Therefore, class 4 shows the largest variability among all burns classes. It should also be noted that the boxes for class 4, 5 and 6 are not centered between the whiskers, which means that in these cases there are many low values of the SBSI and only a few of the extremely high ones. Overall, the distribution of the SBSI within the six classes of burns looks as though it is generally non-uniform with a particular variability to the left, but there are clear borders between adjacent classes, which can lead to an accurate classification of burns. This is confirmed by a qualitative comparative analysis (by visual inspection) between the SBSI classification results (BCM map – Fig. 3(b)) and the LDI map (Fig. 3(c)) which shows that BCM provides more details (e.g. the right upper pole area which is more distinctively displayed by BCM in total correspondence with the digital photo).

3.2 Classification accuracy

The SBSI and LDI classification results were evaluated in terms of overall accuracy, user’s and producer’s accuracies and Kappa coefficient (Kc), as detailed in Section 2.4 in order to test the performance of SBSI method in burn depth assessment. Classification accuracies of the two methods for the particular case presented in Fig. 3 are summarized in Table 3.

| Class | Class 1 | Class 2 | Class 3 | Class 4 | Class 5 | Class 6 | Total | UA (%) |
|-------|---------|---------|---------|---------|---------|---------|-------|--------|
| Class 1 | 240 | 0 | 0 | 186 | 2 | 1 | 429 | 55.94 |
| Class 2 | 0 | 244 | 0 | 1 | 0 | 0 | 245 | 95.90 |
| Class 3 | 0 | 14 | 514 | 0 | 0 | 0 | 528 | 93.75 |
| Class 4 | 0 | 0 | 0 | 663 | 0 | 0 | 663 | 100.00 |
| Class 5 | 0 | 5 | 0 | 22 | 289 | 0 | 316 | 91.45 |
| Class 6 | 240 | 0 | 0 | 1 | 0 | 141 | 142 | 99.29 |
| Total | 240 | 263 | 514 | 873 | 291 | 142 | 2323 | OA = 90.00% |
| PA (%) | 100.00 | 92.77 | 100.00 | 75.94 | 99.29 | 99.29 | K = 0.8736 |

| Class | Class 1 | Class 2 | Class 3 | Class 4 | Class 5 | Class 6 | Total | OA = 86.16% |
|-------|---------|---------|---------|---------|---------|---------|-------|------------|
| Class 1 | 228 | 7 | 1 | 98 | 0 | 0 | 334 | 68.26 |
| Class 2 | 15 | 247 | 9 | 3 | 0 | 0 | 274 | 90.14 |
| Class 3 | 13 | 20 | 426 | 25 | 20 | 1 | 487 | 87.47 |
| Class 4 | 0 | 3 | 23 | 620 | 25 | 7 | 678 | 91.45 |
| Class 5 | 0 | 0 | 0 | 19 | 275 | 13 | 307 | 89.58 |
| Class 6 | 0 | 0 | 0 | 0 | 5 | 129 | 134 | 36.27 |
| Total | 256 | 277 | 459 | 765 | 327 | 150 | 2234 | OA = 86.16% |
| PA (%) | 89.00 | 89.17 | 92.81 | 81.04 | 84.10 | 86.00 | K = 0.8259 |

In these confusion matrices, the elements located along the main diagonal show pixels correctly classified by the two classification methods (SBSI and LDI). Matrix elements located outside the main diagonal are pixels where mistakes occurred. The overall classification accuracy (OA) calculated as the ratio between the total number of correctly classified pixels and the total number of pixels in the matrices were 90.00% for SBSI and 86.16% for LDI. Moreover, the kappa coefficient (K) calculated for each method was found to be above 0.8 (SBSI method: K = 0.8736; LDI method: K = 0.8259) indicating that there is a very good agreement between the classification results and the ground truth image. However, the non-diagonal elements of the confusion matrices are either pixels that were assigned to a class that they don’t belong to or pixels that belong to a certain class but were
assigned to other classes. The measure of such errors is the producer’s accuracy and, respectively, user’s accuracy indicators. Typically, for any given class, these indicators have different values. For example, for SBSI method, the producer’s accuracy ranged from 75.94% (class 4) to 100% (classes 1 and 3), whereas the user’s accuracy was found to be between 55.94% (class 1) and 100% (class 4). The finding that producer’s accuracy for the class 1 was 100.00% while the user’s accuracy was 55.94% means that even though all the pixels belonging to class 1 were correctly identified as so, only 55.94% of the pixels identified as class 1 in the classification were in fact deep burn or normal skin. For LDI method, the value ranges of producer’s and user’s accuracies indicators were 81.04% - 92.81% and 36.27% - 91.45%, respectively. The lowest producer’s accuracy of 81.04% was found for the yellow class (3) and the highest of 92.81% for green class (4), whereas the lowest user’s accuracy was 36.27% for red class and the highest 91.45% for yellow class (3).

The producer’s and user’s accuracy for each class calculated for the particular case presented above for the two methods are shown in Fig. 4.

![Fig. 4](image)

Fig. 4. The producer and user accuracies for different burn classes. Error bars represent one standard deviation from the mean.

For all burn classes, SBSI method has both mean producer’s and user’s accuracy higher than the LDI method, recording the highest average values for classes 2 (blue) and 5 (magenta). This means that the full thickness burn and the superficial partial burn are best identified by the method (as it is the case with the clinical examination). However, the SBSI method performed in burn classification as well as the LDI method in terms of overall classification accuracy and Kappa coefficient (Fig. 5).

![Fig. 5](image)

Fig. 5. Classification accuracy of the SBSI method compared to the LDI method. a) overall accuracy; b) Kappa coefficient.
The average overall classification accuracy of the SBSI and LDI methods was found to be of $92.17 \pm 2.513\%$ and $89.42 \pm 2.276\%$ respectively. The average Kappa coefficient was $0.9022 \pm 0.0310$ for the SBSI method and $0.8689 \pm 0.0287$ for the LDI method.

By analyzing the differences between the OA and K coefficients of the two methods against their means according to the Bland and Altman method [40], the classification accuracy of SBSI method proved to be high enough compared to the one of the LDI method (Fig. 6).

![Fig. 6](image)

**Fig. 6.** Difference in the average values of accuracy indicators computed for SBSI and LDI methods for the 9 patients in the group: a) overall accuracy; b) Kappa coefficient.

The average differences in OA and K coefficient are 2.7489 and 0.0332 respectively. This means that systematically the OA and K coefficient of the SBSI method are above those of LDI method. Moreover, the limits of agreement obtained for OA (0.3291 ± 5.1687) % and K coefficient (0.0017 ÷ 0.0648) are close enough, so the SBSI method can be considered fairly accurate for the classification of burns in comparison to LDI.

4. **Discussion**

Objective classification of burn depth is of paramount importance in choosing the surgical treatment of the lesion (conservative approach with dressing changes in superficial burns and excision and grafting for deeper burns that do not heal in three weeks). Clinical examination has major limitations and leads to over estimations of the areas to be grafted. Therefore, the need for objective data is a current subject of research. From the wide variety of methods investigated to date, only laser Doppler imaging has gained significant clinical applications.

This study describes a new metric for assessing the burns depth based on hyperspectral imaging. It uses the spectral properties of normal and burned skin to define a spectral index and generate burn classification maps. In defining the spectral index, the authors took into account two wavelengths ($\lambda_1 = 556$ nm and $\lambda_2 = 760$ nm) at which the greatest difference between reflection spectra of the normal and burned skin was identified. These wavelengths correspond to absorption maxima of deoxyhemoglobin. This finding is in accord with the results reported by Chin et al. [28] that highlighted that deoxyhemoglobin correlates better than oxyhemoglobin with burn depth. For generation of burn classification maps (BCM) a sequence of four steps was proposed starting with hyperspectral image acquisition, hyperspectral image processing, computing of the burns spectral index (SBSI) in each pixel from the hyperspectral image resulting in a grayscale image and finally creating a burn classification map. Since the grayscale images generated by SBSI approach are difficult to use in clinical decision making, K-means clustering was used to generate color-coded BCMs that are more easily compared to perfusion ones produced by LDI method. The results presented here reveal that BCM map based on SBSI index are easier to read and superpose...
better over digital photos being more helpful in clinical settings. Another way to generate burn severity maps has recently been proposed by Paluchowski et al. [29]. They reported an unsupervised spectral-spatial segmentation algorithm applied on hyperspectral images to discriminate burns of variable severity, but only in a porcine model. Earlier studies [26,27] report good results of HSI-based human burn classification, but despite long time spans from publication, there are no clinical applications of the methods.

Performances in burns classification of BCMs generated in the present study were compared to LDI perfusion maps using confusion matrices. The results showed that BCM correlates well with LDI map, being more detailed in certain areas (like the right upper region of the burn in Fig. 3). Like the LDI, SBSI method generates more classes for the same burn depth. Yet, the most important issue is to set up a border between burns that need excision and those that do not. Data gathered in this study suggested that green areas are to be grafted, while yellow class would heal through supervised epithelization. This conclusion is based mainly on comparison with well-established LDI indications. One concern is the low user’s accuracy that class 4 (green) displayed for both LDI and SBSI methods, because it means that burns that are not in this class are assigned to it in a significant proportion, which leads to over estimation of the area to be excised and grafted. The group tested in this study was rather small and showed large variations in differences of overall accuracy and Kappa coefficient (Fig. 5) between the two methods. Only 5 of the patients are clustered near the mean line (Fig. 6). Therefore, the conclusion that SBSI-generated BCM maps are fairly accurate in comparison to LDI maps is supported by the experimental findings.

The results presented here reveal, for the first time, that, by combining hyperspectral imaging with a suitable computational approach to determine the spectral index of the skin, it is possible to accurately assess burn depth. This new approach is an important progress in the fields of both HSI processing and burn depth classification.

However, our work raises several interesting questions. It would be important to consider a comparison of the SBSI classification results with histological exam or clinical outcome as preparatory steps for clinical use. Furthermore, validation of the method on a larger group of patients is required. It is also important to define other spectral indices using different combinations of wavelengths (associated with the optical properties of the skin) and to assess their performance in burn classification. In addition, new SBSI image conversion methods to the BCM map need to be investigated in order to reduce the misclassification errors. These questions will be addressed in the further studies.

5. Conclusions

In conclusion, the new hyperspectral index-based metric for burn depth assessment presented in this study has the potential to generate accurate burn classification maps showing spatial distribution of burns types in the affected body areas and its applicability in hospital settings seems promising.

Funding

The Romanian Ministry of Research and Innovation (No. PN 33N/16.03.2018).

Disclosures

The authors declare that there are no conflicts of interest related to this article.

References

1. P. A. Brigham and A. R. Dimick, “The evolution of burn care facilities in the United States,” J. Burn Care Res. 29(1), 248–256 (2008).
2. S. Hettiaratchy and R. Papini, “Initial management of a major burn: II–assessment and resuscitation,” BMJ 329(7457), 101–103 (2004).
3. A. B. Wallace, “The exposure treatment of burns,” Lancet 1(6653), 501–504 (1951).
4. C. C. Lund and N. C. Browder, “The estimation of areas of burns,” Surg. Gynecol. Obstet. 79, 352–358 (1944).
1. J. M. Still, E. J. Law, K. G. Klavuhn, T. C. Island, and J. Z. Holtz, “Diagnosis of burn depth using laser-induced
10. S. Monstrey, H. Hoeksema, J. Verbelen, A. Pirayesh, and P. Blondeel, “Assessment of burn depth and burn
11. L. Devgan, S. Bhat, S. Aylward, and R. J. Spence, “Modalities for the assessment of burn wound depth,” J.
15. A. Renkielska, M. Kaczmarek, A. Nowakowski, J. Grudziński, and K. M. Cross, L. Leonardi, J. R. Payette, M. Gomez, M. A. Levasseur, B. J. Schattka, M. G. Sowa, and J. S. Fish, “Clinical utilization of near-infrared spectroscopy devices for burn depth assessment,” Wound Repair Regen. 15(3), 332–340 (2007).
22. J. A. Brink, P. W. Sheets, K. A. Dines, M. R. Etchison, C. W. Hanke, and A. M. Sadove, “Quantitative imaging (LSI),” Biomed. Opt. Express 5(10), 3467–3481 (2014).
25. M. J. Koruda, A. Zimblcr, R. G. Settle, D. O. Jacobs, R. H. Rolandez, G. L. Wolf, and J. L. Rombeau, “Assessing burn wound depth using in vitro nuclear magnetic resonance (NMR),” J. Surg. Res. 40(5), 475–481 (1986).
26. M. A. Afronzowitz, J. B. Callis, D. M. Heimbach, L. A. DeSoto, and M. K. Norton, “Multispectral imaging of burn wounds: a new clinical instrument for evaluating burn depth,” IEEE Trans. Biomed. Eng. 35(10), 842–850 (1988).
27. W. Eisenbeiß, J. Marotz, and J. P. Schrade, “Reflection-optical multispectral imaging method for objective determination of burn depth,” Burns 25(8), 697–704 (1999).
28. M. S. Chin, O. Babchenko, J. Lujan-Hernandez, L. Nobel, R. Ignotz, and J. F. Lalikos, “Hyperspectral imaging for burn depth assessment in an animal model,” Plast. Reconstr. Surg. Glob. Open 3(12), e591 (2015).
29. L. A. Paluchowski, H. B. Nordgaard, A. Bjorgan, H. Hov, S. M. Berget, and L. L. Randeberg, “Can spectral-spatial image segmentation be used to discriminate experimental burn wounds?” J. Biomed. Opt. 21(10), 101413 (2016).
30. M. A. Calin, S. V. Parascu, R. Savastru, and D. Manea, “Characterization of burns using hyperspectral imaging technique - a preliminary study,” Burns 41(1), 118–124 (2015).
31. J. Nieke, D. Schläpfer, F. Dell’Endice, J. Brazile, and K. I. Itten, “Uniformity of imaging spectrometry data products,” IEEE Trans. Geosci. Remote Sens. 46(10), 3326–3336 (2008).
32. G. Polder, G. W. A. M. van der Heijden, L. C. P. Keizer, and I. T. Young, “Calibration and characterization of imaging spectrographs,” J. Near Infrared 11(3), 193–210 (2003).
33. H. Akbari, L. V. Halig, D. M. Schuster, A. Osunkoya, V. Master, P. T. Nieh, G. Z. Chen, and B. Fei, “Hyperspectral imaging and quantitative analysis for prostate cancer detection,” J. Biomed. Opt. 17(7), 0760051 (2012).
34. A. Bannari, D. Morin, F. Bonn, and A. R. Huete, “A review of vegetation indices,” Remote Sens. Rev. 13(1-2), 95–120 (1995).
35. J. Imanishi, Y. Morimoto, A. Imanishi, K. Sugimoto, and K. Isoda, “The independent detection of drought stress and leaf density using hyperspectral resolution data,” Landscape Ecol. Eng. 3 (1), 55–65 (2007).
36. M. Boschetti, F. Nutini, G. Manfron, P. A. Brivio, and A. Nelson, “Comparative Analysis of normalised difference spectral indices derived from MODIS for detecting surface water in flooded rice cropping systems,” PLoS One 9(2), e88741 (2014).
37. S. Chabrillat, “Land Degradation Indicators: Spectral indices,” Ann. Arid Zone 45, 331–354 (2006).
38. L. Schepers, B. Haest, S. Veraverbeke, T. Spanhove, J. V. Borre, and R. Goossens, “Burned area detection and burn severity assessment of a heathland fire in Belgium using airborne imaging spectroscopy (APEX),” Remote Sens. 6(3), 1803–1826 (2014).
39. J. M. Bland and D. G. Altman, “Statistical methods for assessing agreement between two methods of clinical measurement,” Lancet 1(8476), 307–310 (1986).