Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/burns

Classification of burn injury using Raman spectroscopy and optical coherence tomography: An *ex-vivo* study on porcine skin

Lakshmi Priya Rangaraju [a,*], *Gautam Kunapuli* [a,b], *Dayna Every* [c], *Oscar D. Ayala* [c], *Priya Ganapathy* [a], *Anita Mahadevan-Jansen* [c]

[a] *Utopia Compression Corporation, 11150 W Olympic Blvd, Los Angeles,CA 90064, USA*
[b] *Department of Computer Science, The University of Texas at Dallas, 800 W Campbell Rd, Richardson, TX 75080, USA*
[c] *Department of Biomedical Engineering, Vanderbilt University, 2201 West End Ave, Nashville, TN 37235, USA*

ARTICLE INFO

Article history:
Accepted 4 October 2018

Keywords:
Raman spectroscopy
OCT
Classification
Burn depth

ABSTRACT

Accurate depth assessment of burn wounds is a critical task to provide the right treatment and care. Currently, laser Doppler imaging is able to provide better accuracy compared to the standard clinical evaluation. However, its clinical applicability is limited by factors like scanning distance, time, and cost. Precise diagnosis of burns requires adequate structural and functional details. In this work, we evaluated the combined potential of two non-invasive optical modalities, optical coherence tomography (OCT) and Raman spectroscopy (RS), to identify degrees of burn wounds (superficial partial-thickness (SPT), deep partial-thickness (DPT), and full-thickness (FT)). OCT provides morphological information, whereas, RS provides biochemical aspects. OCT images and Raman spectra were obtained from burns created on *ex-vivo* porcine skin. Algorithms were developed to segment skin region and extract textural features from OCT images, and derive spectral wave features from RS. These computed features were fed into machine learning classifiers for categorization of burns. Histological results obtained from trichrome staining were used as ground-truth. The combined performance of RS-OCT reported an overall average accuracy of 85% and ROC-AUC=0.94, in distinguishing the burn wounds. The significant performance on *ex vivo* skin motivates to assess the feasibility of combined RS-OCT in *in vivo* models.

1. Introduction

1.1. Background

Burn injuries are a major cause of death in the United States [1]. Every year, nearly 450,000 patients receive hospital and emergency room treatment for burns, of which, approximately 3400 burn injury deaths are reported [2]. Early decision on treatments has shown to shorten stay in hospitals and reduce infection [3,4]. Accurate assessment and in-time treatment can greatly reduce morbidity and mortality due to burn injuries.

Identification of burn wound magnitude plays a crucial role in deciding the treatment to be given to patient. Clinically, burn wound depth is assessed based on visual appearance, amount of bleeding, capillary refill, and sensitivity remaining in the burn area. Based on the extent of burn depth and healing

* *Corresponding author.*
E-mail address: lakshmi@utopiacompression.com (L.P. Rangaraju).
https://doi.org/10.1016/j.burns.2018.10.007
0305-4179/

time, burn wounds are classified mainly into four categories: superficial, superficial partial-, deep partial-, and full-thickness burns [5]. Superficial and superficial-partial burn wounds heal in 1–2 weeks and require basic procedures, such as application of antimicrobial creams, analgesic, and proper monitoring of the wound. Deep partial- to full-thickness wounds may require surgical intervention to perform excision and skin grafting. However, identifying burn degree, even by surgeons, is limited to 60–75% accuracy [6,7]. Furthermore, it is subjective and can result in variations among multiple clinicians [8,9]. Although clinical recognition of superficial- and full-thickness burns can be reliable, this method is limited for intermediate burns such as superficial partial- and deep partial-thickness burns. Additionally, the dynamic nature of wounds, which aggravates the extent of burn depth, is a hindrance in making proper treatment decisions [10]. Therefore, there is an essential need for a diagnostic approach to accurately assess burn depths, which will provide precise insights to determine appropriate treatment plans and burn care management.

1.2. *Current diagnostic technique and its limitations*

The necessity for accurate assessment of burn depth and intensity has stimulated evolution of a wide range of diagnostic methods and modalities to assist evaluation of wounds and provide proper burn care management. However, among various techniques, laser Doppler imaging (LDI) has been shown to stand out due to its accuracy and clinical applicability [10]. It provides better accuracy in prediction of wound healing compared to clinical assessment, especially for superficial-partial and deep-partial wounds [11]. The technique works by measuring perfusion index (PI) of skin based on Doppler shift of the reflected laser light penetrating through the sample. Using LDI, varying burn depths were identified in terms of PI, distinguishing superficial and deep burns [7,12]. In addition, LDI could successfully identify burns that may or may not heal by re-epithelialization within three weeks [13].

Despite positive results obtained through the use of LDI, the modality is still vulnerable to certain physiological factors. Changes in blood perfusion measured by LDI can substantially vary in individual patients, as it is sensitive to changes in blood pressure, skin temperature, anemia, and presence of infection [14]. This makes interpretation of PI maps unclear. Further, external factors such as scanning distance, thickness of antimicrobial dressings, and curvature of the tissue surface alter LDI outputs [15]. Limited accuracy of LDI alone to diagnose burn depth within the first 48 h has been reported as well [16]. The time taken by an LDI device to scan over large surface areas and obtain a higher resolution takes minutes and requires patients to remain motionless for the entire duration [17]. Moreover, the commercial cost of LDI devices is expensive, limiting its availability to only a few burn centers [18].

1.3. *Proposed diagnostic techniques*

1.3.1. *Optical coherence tomography (OCT)*

It is a non-invasive imaging technique based on low-coherence interferometry, which produces micro-resolution, two-dimensional, cross-sectional images of biological tissue. The images are produced by capturing reflected or scattered light from the target tissue, a principle that is analogous to ultrasonic pulse-echo imaging. In highly scattering tissues such as skin, OCT can image small blood vessels and other structures up to a depth of 1–2 mm below the surface [19]. Light reflected from a point on target surface constitutes one single A-scan (Z-axis; axial depth); combining a series of A-scans obtained through lateral scanning across the target surface produces a B-scan (X–Z plane). Acquisition of multiple B-scans over an area of target surface allows for three-dimensional visualization of the tissue [20].

OCT of healthy skin allows visualization of skin layers, mainly, *stratum corneum*, epidermis, and dermis depending on the thickness of skin area imaged. The technique has been used in studies for anatomical, physiological, and pathological evaluation of skin with different conditions [21]. In skin abnormalities, such as skin tumors and inflamed skin, OCT allows identification of the diseased regions, which are characterized by homogeneous signal distribution, thickening of epidermis, and signal attenuation in dermis [22,23]. The birefringence property of skin reduces due to loss of collagen after sustaining a burn, which results in lower penetration depth of light. Consequently, reduction in signal corresponds to low intensity regions on OCT images. Hence, the inverse correlation between depth estimate derived from OCT and the incident burn degree can be exploited for diagnostic purposes [24]. Polarization-sensitive OCT (PS-OCT) is an extension of conventional OCT which utilizes information contained in polarized light. Reduction in collagen has been quantified using PS-OCT in burned skin with a decreased phase retardation rate, and also correlated with burn depth [25,26]. Extending its capability of providing structural morphology of the tissue, OCT has also been used to capture the flow of blood in blood vessel networks. Such non-invasive OCT angiography (OCTA) techniques acquire images at a very high frame rate and distinguish the blood vessel from static tissue by analyzing the changes in OCT signal due to the blood flow [27–29]. *In vivo* studies on various skin conditions such as, actinic keratoses, squamous cell and basal carcinomas, and Bowen's disease, have distinguished the vascular patterns using OCTA [30,31]. Further, OCTA was also used in cross-sectional and longitudinal studies for assessment of burn scars [32,33], and monitoring the effect of laser fractionation [34]. While OCT images can provide structural information, they do not provide details on compositional aspects of tissue [35].

1.3.2. *Raman spectroscopy (RS)*

It is a non-invasive technique that provides information about molecular vibrations of a sample; specifically, the chemical framework of a sample is obtained in terms of shift in wavelength (Raman shift) of the inelastically scattered incident light. The shift in wavelength is due to interactions between incident electromagnetic waves and vibrational energy levels in molecules of a given sample, and hence, can be used for sample identification [36]. Biological molecules such as nucleic acids, proteins, and lipids are manifested with specific spectral sensitivity; these spectral features can be quantified and used for various applications. One such spectral sensitivity is noted in burn wounds, where the constituent proteins of skin tissue are denatured [37]. Hence,

it may be speculated that these changes correspond to the variations in the acquired spectra.

An *in vitro* study has shown variation in Raman spectra (RS) associated with changes in lipid content of *stratum corneum* in human skin [38]. RS also successfully classified malignant and benign skin lesions with sensitivity between 95%-99% [39]. Confocal RS has demonstrated the ability to differentiate types of malignant melanoma and basal cell carcinoma [40]; additionally it provided differential diagnosis of non-melanoma skin cancers [41]. RS was also successful in extracting hydration information from skin, which can be essential in skin burn assessment [42]. A previous study has identified burn-induced heterotopic ossification using RS [43], thereby, demonstrating RS as a potential modality in the assessment of burn skin and burn degree categorization.

Accurate diagnosis and systemic burn care management entails availability of considerable structural and functional information. The ability of OCT to image microstructures at high-frame rates when combined with the caliber of RS to provide biochemical specificity might overcome constraints of each other, and provide more reliable diagnosis. Previous work has shown that fusion of two complementary technologies, such as frequency domain imaging with laser speckle imaging and OCT imaging with pulse speckle imaging, confocal microscopy with RS, yielded successful results for burn wounds [24,44,45]. Studies have also demonstrated the effective implementation of combined RS-OCT in identifying pre-cancer and cancer skin lesions, and classifying different tissues such as kidney, liver, and small intestine [46,47].

Here, we report our assessment on the combined power of RS and OCT modalities to classify burn wounds. OCT images were obtained using a commercial OCT system (Thorlabs Inc., Newton, NJ); RS data was obtained using a prototype device built at Vanderbilt University (Nashville, TN). We developed processing algorithms for OCT images and RS spectra to identify regions of interest (ROI) and extract specific features. In addition, we implemented machine learning (ML) algorithms to classify burn degree/type depending on collective features acquired and/or derived from RS and OCT data. The results from this work may encourage the concept of a unified RS-OCT device for non-invasive, reliable diagnosis of burn injury, and is a cost-effective method, making it affordable for clinical applications.

2. Methods

2.1. Burn wound experiment

Burn wounds were created on cadaver skin samples of a pig. Prior to creation of burn wounds pig skin samples were thawed using phosphate-buffered saline (PBS) for at least one hour. Burn severity is related to the burn temperature, contact duration, and resistance offered by the skin [48,49]. To produce consistent burn wounds, a protocol similar to the one adapted by Ponticorvo et al. was followed [44]. The tip of a soldering rod at a temperature of 140°C (the lowest temperature that could be set on the soldering rod) was used to create different degrees of burn by holding the rod against the pig skin for 5, 10, and 15 s. To ensure sufficient sample size for data analysis and validation of algorithms, 26 burn wounds for each burn time were created (5 s: N=26; 10 s: N=26; 15 s: N=26). An example of burn wounds created for all burn times at 4 different spots is shown in Fig. 1.

The histology of burn wounds was obtained using trichrome staining and the degree of each burn was determined. Exposure to heat denatures the collagen in the skin (loss of original structure), thus, leading to coagulation (dense-packed structure like appearance). The amount of coagulation is proportional to the burn intensity and duration of burn contact [49] and, can be observed using histological staining [50]. The degree of burn was measured as the ratio of coagulated collagen to total collagen thickness (or skin sample thickness) [44]. The histological image of a 10 s burn wound is represented in Fig. 2. The coagulated to total collagen ratio was obtained manually using calipers on each histology slide and an average of up to 20 points per slide was taken, based on the quality of the slide. Wounds with typical ratio values of below 0.35 were classified as superficial partial-thickness (SPT); wounds with collagen coagulation ratio between 0.35 and 0.65, were classified as deep-partial thickness (DPT), whereas, full-thickness (FT) burn wounds were identified as those with ratios greater than 0.65.

2.2. Data acquisition

2.2.1. OCT scanning

The images of each burn wound were acquired using the OCT device, *Thorlabs Telesto (TEL1300V2-BU)*. Images were obtained at an acquisition time of 0.201 s and 76 kHz imaging speed. Each image was averaged over 20 horizontal B-scans of 9 mm length and had a depth resolution of 5.5 μm in air (4.2 μm in water). After creation of burn wounds, the respective skin regions were imaged, such that the acquired scan consisted of small

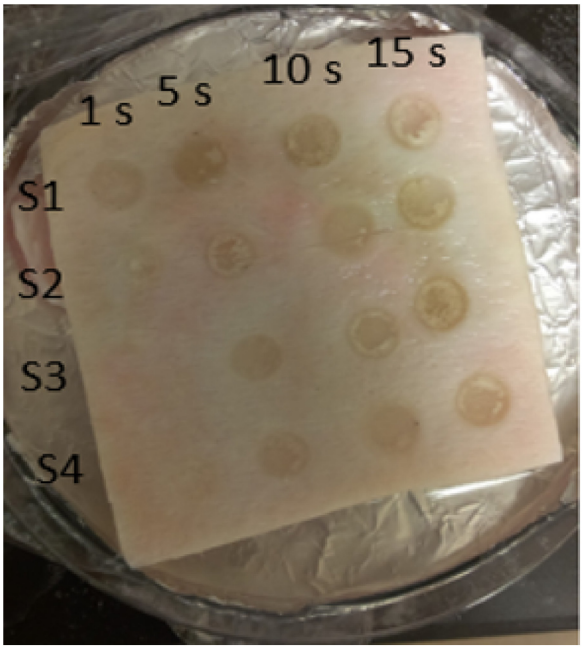

Fig. 1 – Burn wounds created on porcine skin sample at 4 spots (S1:S4) for 4 contact times (1 s, 5 s, 10 s and 15 s). (NOTE: 1 s wounds were not used for analysis).

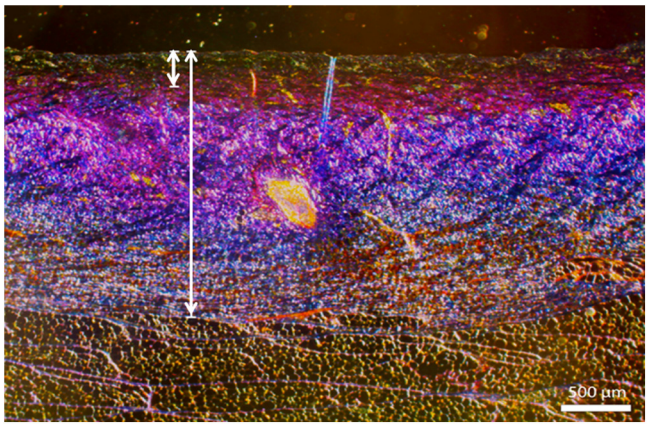

Fig. 2 – Histology of 10s burn wound (Superficial partial-thickness) (Scale: 500 μm). Shorter double-arrow line indicates coagulated collagen (top brown colored region) and the longer double-arrow line indicates total collagen thickness (ratio < 0.35). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

portion of adjacent normal (control) skin as well. Fig. 3 shows a sample white-light image and corresponding OCT image.

2.2.2. *Raman spectra*

The spectra were captured using the prototype RS device built at Vanderbilt University [51]. It consisted of a 785 nm laser source (at 80 mW) and a fiber optics probe (EmVision LLC, USA) with an excitation fiber of 200 μm and 10 collection fibers of 300 μm. The wavenumber was calibrated using a neon argon lamp, acetaminophen, and naphthalene. Green glass and white light were used for calibration of intensity to account for the response of the detector. The effective area on the skin sample that can be scanned by the RS device was 19.6 mm^2. The locations for acquiring the spectra were drawn on the petri dish containing the skin sample, and RS was approximately obtained over these marked regions. Raman spectra was collected four times from each spot with an integration time of 3s and 20 accumulations. RS data was collected at each point, both before (control spectra) and after (burn spectra) burn wound creation. The acquired spectra was post-processed with fluorescence background subtraction and noise reduction techniques, following methods developed by Lieber et al. [52]. These methods involve binning data to half the spectral resolution and noise filtering using a 2nd order Savitzky-Golay filter [53], with window size set to twice the spectral resolution, and background auto-fluorescence subtraction using a modified polynomial-fitting algorithm. Resolving for any intensity variation the final spectra were also mean normalized.

2.3. *Data analysis*

2.3.1. *OCT processing*

OCT images were processed using a customized algorithm developed in MATLAB (MathWorks, Inc., Natick, MA). The main objective of the algorithm was to extract skin region and estimate depth (thickness) of skin. All the processing steps were performed on grayscale OCT images of skin. Initially, to enhance contrast of the image, histogram equalization was performed. Histogram equalization enhances low contrast areas by spreading the most frequently occurring intensity values globally across the image. Next, the enhanced image was converted to binary image using Otsu's thresholding method [54]. This step replaces all pixel values above a certain threshold to ones and others to zeros. The selection of an

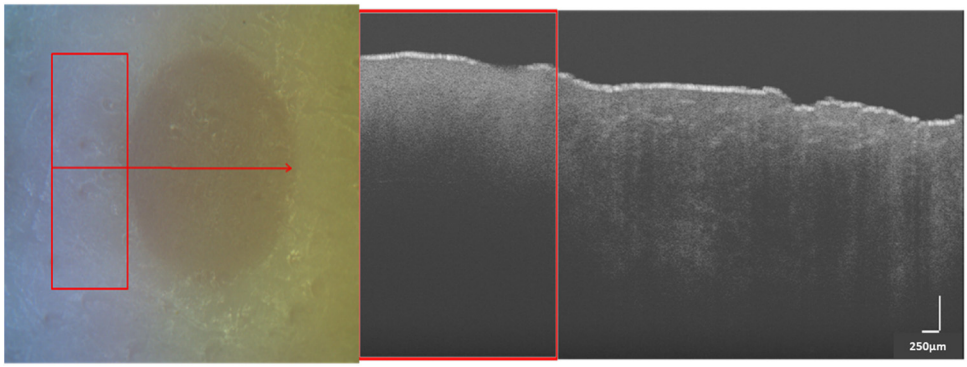

Fig. 3 – LEFT White light image of superficial partial-thickness (SPT) burn; arrow indicates region and direction of scan; red-rectangular box indicates control skin region. RIGHT: corresponding OCT B-scan (Scale: 250 μm); region inside the red box is normal skin, and that outside is burn region. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

adequate threshold is important to extract a region of interest. Otsu's method selects an optimum threshold value by maximizing the measure of separability of resulting classes in gray-level.

In the next step of segmentation, in order to segregate skin region from background, the flood-fill algorithm was used to identify maximum number of connected pixels. The algorithm considers an initial node or pixel, and recursively searches for adjacent pixels connected to it and marks them. Skin region, with these labeled pixels in the binary image were subtracted from background region (unlabeled pixels) having least connectivity. The extracted skin region was further refined to eliminate any isolated regions. Morphological operations were used to fill in any spaces and small discontinuous regions with area less than 10 pixels were identified and set to zero. The final binary image was combined, using logical AND operation, with the initial grayscale image to retain the original intensity values. The logical AND operator retains locations of only those pixels which contain non-zero values in both grayscale and binary images, thus retaining only the region of interest.

The segmented images of the skin were used to estimate the OCT-derived visible thickness [addressed as 'depth' throughout the paper] of the extracted skin region [24]. As the OCT image consisted of both control (normal) and burn regions, two depth measures were obtained per image. The depth for burn region was calculated at the central portion of the burn wound. A 300-column range was selected for both regions (control and burn) in each wound. The total number of non-zero pixels in each column was counted and averaged over 300 columns. The average measure was then multiplied by 2.46 μm (depth resolution) and converted to millimeters.

In addition to depth measurement, the texture of the OCT skin image was also evaluated. Statistical texture metrics like mean and variance of pixel intensities, entropy, and smoothness were calculated. Entropy is the estimate of the randomness of pixel intensities and is zero for a constant image. Smoothness is similar to variance; it gives a relative estimate of texture of the image (smoother or coarser). Smoothness and entropy estimates were made based on formulas presented in Gonzalez and Woods [55]. The statistical metrics were obtained over 300 columns for both control and burn regions. Mean intensity and variance were calculated across two regions considering only the non-zero pixels.

2.3.2. *Raman spectral analysis*

The RS data contains details of tissue composition at different wavenumbers. The wavenumbers in the range 800–1800 cm^{-1} were considered for analysis, as they contain useful information. Mean normalized spectra obtained four times from each burn wound/spot were averaged, producing one spectrum per spot. The averaged spectra were further divided into segments using a peak-detection algorithm resulting in 18 segments for each burn wound. The algorithm was applied to obtain the local minima across the spectrum, and the band of wave numbers lying between two successive minima (valley points) was grouped into one segment. To adjust for any spectral artefacts introduced by skin heterogeneity, quantitative metrics and features to analyze the spectra for burn wounds were obtained *relative* to control spectra (Method I, relative RS features). Specifically, three main metrics were estimated for each segment: (1) difference of area under intensity curve (AUC); (2) cosine similarity, measure of angle between two vectors and (3) peak-intensity ratio (maximum intensity in a given segment) between corresponding control and burn segments. These metrics were obtained for each of the 18 segments, resulting in 54 features per burn wound.

Although 54 features obtained from Raman spectra compensated for skin inhomogeneity, this may not be achievable in a real-time scenario. Therefore, we also obtained an additional set of *absolute* features considering only burn spectra and not difference with respect to control (Method II, absolute RS features). Previous studies on skin cancer, have observed changes in certain lipid-specific bands [39,56]. Peak intensity (maximum intensity) ratios at such significant wavenumbers were estimated (3 features). In addition, areas under the curves (AUC) for each of the 18 segments were also calculated (3+18=21 features per wound).

2.4. *Classification of burn types*

Classification of burn types was conducted for three sets, based on features derived from: OCT alone, RS alone and combined RS-OCT. Classification performance was evaluated using three supervised machine learning (ML)-classifiers: Logistic Regression (LR), Linear Support Vector Machines (SVM) and Random Forest (RF); based on the number of features and training data size, the classifiers were chosen among various ML algorithms, to classify burn types with high reliability and robustness [57–59]. Training models built using these classifiers were validated using 10-fold cross validation. All the classifiers were implemented using Waikato Environment for Knowledge Analysis (WEKA) software (The University of Waikato, Hamilton, New Zealand) [60].

The candidate classifiers were trained on RS features extracted from two methods: (I) using relative RS features (RS: 54 features; OCT: 5 features; RS-OCT: 59 features), (II) using absolute RS features (RS: 21 features; OCT: 5 features; RS-OCT: 26 features). Fig. 4 presents the input features derived from each modality, and the order of inputs provided to each of the three candidate ML classifiers.

3. Results

All the burn wounds were classified as SPT, DPT, and FT burn types based on histology. Two of the burn wounds were inconclusive due to degradation in the quality of histology slide, and hence, were excluded from analysis. Thus, number of samples (burn wounds) for each burn type, based on histology, was 28 -SPT, 22 -DPT, and 26 -FT. The correlation between burn spots created at different durations (5 s, 10 s, 15 s) for a constant temperature and the ground truth as assessed by histological grading, was statistically significant ($r=0.87$; $p<0.01$). No OCT image was discarded due to image quality issues.

Fig. 5 illustrates raw and segmented OCT scans of an SPT burn wound; bottom image represents the region of interest, extracted using our customized algorithm. No visible shadow regions due to presence of hair were observed in the OCT

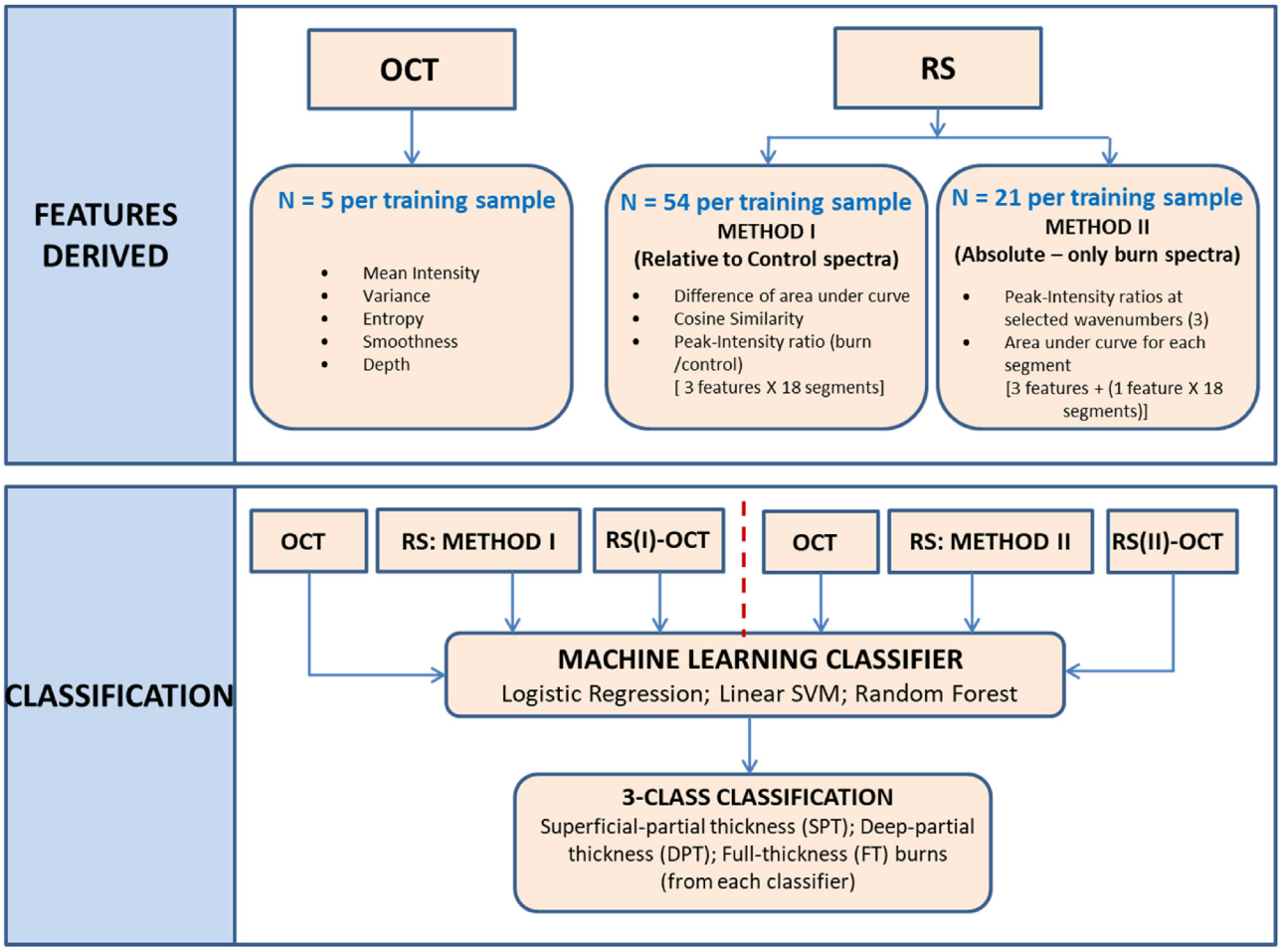

Fig. 4 – Flowchart indicating the features derived from each modality and the workflow using machine Learning (ML) classifiers to identify the burn types. Red-dashed line in the "classification" block shows segregation of input features to the respective ML classifiers. These classifiers take features from OCT alone, RS (Method I) alone, RS (Method II) alone, RS (I)-OCT and RS (II)-OCT, respectively to perform a 3-class classification.

images as the porcine skin samples were devoid of any hair (samples obtained from a local meat store). As visually perceived, a slight decrease in depth penetration is observed over the burn region. Control region exhibits a uniform distribution of intensities, compared to burn region; in addition to decrease in intensity, change in texture is also noticed across burn. Depth and texture estimates were obtained across the regions, control (300-columns) and burn (300-columns), which are indicated by red-rectangular boxes. Mean depth and texture metrics, thus, computed over control and burn regions for the three burn types SPT, DPT and FT are listed in Table 1. Mean depth for control and burn regions were significantly different in SPT and FT burns ($p<0.01$), but not in DPT burns ($p=0.23$). Depth and entropy values decreased from SPT to FT in the burn region ($p<0.01$); whereas, other texture measures were not significantly different ($p>0.05$).

The mean normalized spectra averaged across all burn regions and corresponding control spectra is illustrated in Fig. 6. As it can be observed the corresponding spectra for each burn wound/type follow the control (pre-burn) spectra closely. This may indicate that RS is sensitive to heterogeneities in skin tissue, not just across individuals, but also skin tissues at different parts of the body in specific individual. Fig. 7 shows mean normalized and averaged RS for control and burns across all burn types, over one spot each. The dotted lines indicate end of each segment (total 18 segments), which were divided using local minima (valley points), obtained from peak-detection algorithm. Metrics relative to control (Method I) were obtained from these 18 segments; peak intensity points (maximum intensity values in a given segment) are indicated with circular marks. Averaged spectra for only burn wounds, with peak intensity values at wavenumbers relevant to skin tissue (943, 971, 1268, 1300, 1450 and 1660 cm^{-1}), indicated by dashed-lines, is represented in Fig. 8. Peak intensity ratios of wavenumbers, 943/971 cm^{-1} ($NC_{\alpha}C$/CC proline ring), 1300/1268 cm^{-1} (CH bending/Amide III), and 1450/1660 cm^{-1} (CH_2 bending/Amide ICO stretch), were recorded and served as 3 of the 21 features per wound, which were eventually given as inputs to ML classifiers.

Tables 2 and 3 show the accuracy and AUC-ROC for individual classes obtained from all three classifiers on test sets, using relative and absolute RS features (Methods- I & II,

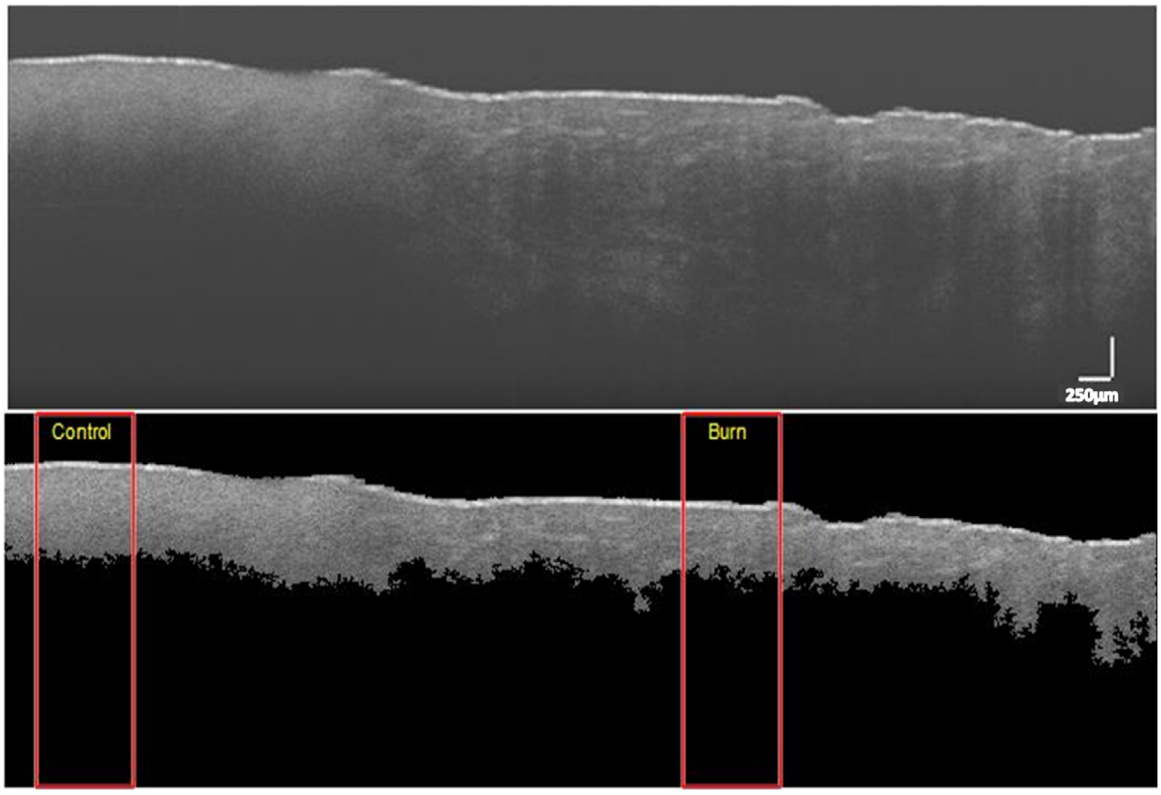

Fig. 5 – TOP- Raw unprocessed OCT image of superficial-partial thickness burn. BOTTOM- Segmented image. Average depth values of control and burn regions are 0.66 mm and 0.58 mm, respectively. Entropy values are 2.56 and 2.27, respectively (Scale: 250 μm).

respectively). OCT-features for both the methods remain same and provide AUC-ROC and accuracy values that range between 0.76–0.85 and 67–69%, respectively across all the three classifiers [61].

With relative RS features as inputs (Method I), RS alone provides accuracy of 80–84% and AUC-ROC of 0.88–0.95; whereas, combined RS-OCT features provide accuracy of about 78–88% and ROC area of 0.90–0.95. When absolute RS features (Method II) were given as inputs to the classifiers, RS's performance (accuracy=83–87%; AUC-ROC=0.90–0.95) was consistent with that of Method I (not significantly different at p=0.05). Accuracy and AUC-ROC for RS-OCT was recorded as 82–85% and 0.92–0.96, respectively, for Method II (not significantly different from Method I at p=0.05). In addition, for both methods, RS-OCT and RS alone offered similar performance. The results demonstrate that features derived from the Raman spectra (with and without taking a difference with respective control spectra) provide robust discrimination across all three burn categories.

4. Discussion

RS and OCT techniques have previously been combined to improve diagnostic accuracy of skin lesions (mainly, cancer) [39,56,35]. In this study, as a primary goal we extended RS-OCT

Table 1 – OCT depth and textural metrics of control and burn skin regions averaged over all wounds in each type.

Burn type	Mean intensity (Mean ± SD)		Variance (Mean ± SD)		Entropy (Mean ± SD)		Smoothness (Mean ± SD)		Depth [mm] (Mean ± SD)	
	Control	Burn	Control	Burn	Control	Burn	Control	Burn	Control	Burn
Full-thickness (FT) (N=26)	126.90 ± 7.29	110.81 ± 6.34	0.48 ± 0.04	0.37 ± 0.06	2.48 ± 0.14	1.86 ± 0.40	0.19 ± 0.03	0.12 ± 0.03	0.60 ± 0.05	0.44 ± 0.12
Deep partial-thickness (DPT) (N=22)	124.30 ± 5.78	109.18 ± 6.39	0.48 ± 0.03	0.31 ± 0.09	2.33 ± 0.21	2.09 ± 0.51	0.19 ± 0.02	0.09 ± 0.05	0.56 ± 0.06	0.52 ± 0.15
Superficial partial-thickness (SPT) (N=28)	126.41 ± 7.42	112.55 ± 3.98	0.49 ± 0.05	0.35 ± 0.05	2.28 ± 0.14	2.68 ± 0.22	0.19 ± 0.03	0.11 ± 0.03	0.55 ± 0.04	0.69 ± 0.07

Bold values signifies Average AUC-ROC and overall accuracy.

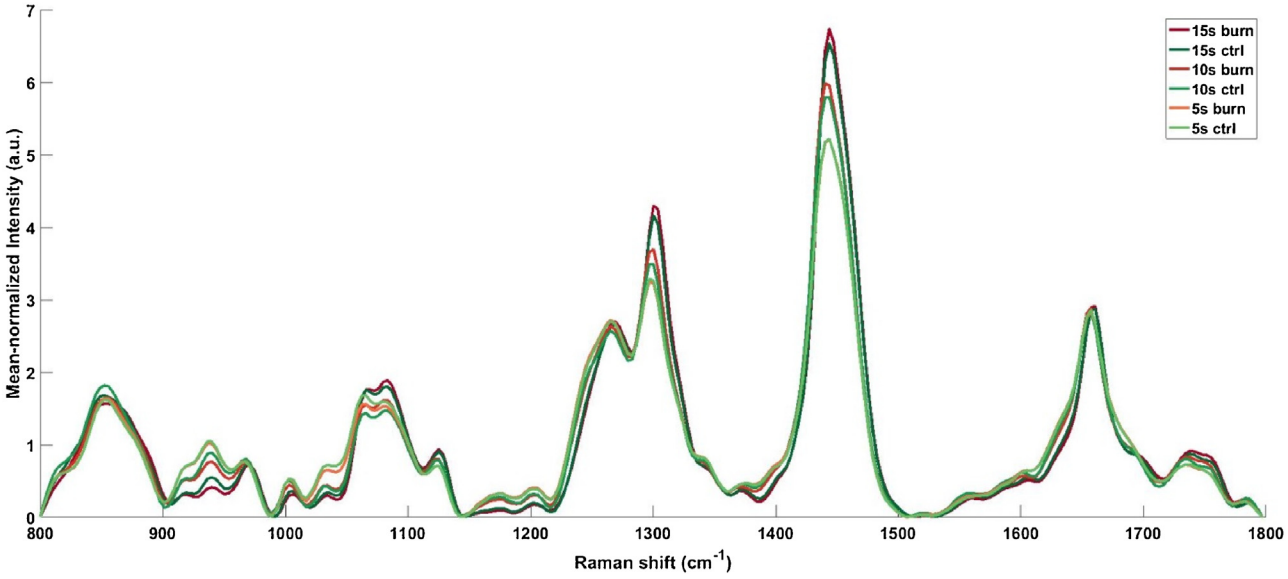

Fig. 6 – Mean-normalized Raman spectra averaged across all wounds in each burn type. The closeness of burn and corresponding control spectra indicates skin heterogeneity captured by RS.

application for *ex vivo* skin burns. We have developed an analytical pipeline to perform image processing and classification to diagnose burn wounds using RS and OCT data. Our prime focus is to demonstrate a proof-of-concept by leveraging appropriate, existing ML algorithms to perform burn wound classification.

Creation of consistent burn wounds at various spots, as evident from significant correlation ($r=0.87$, $p<0.01$), resulted in proper classification of burn wounds into SPT, DPT, and FT, which were identified through histological staining. The ground truth, thus, obtained were used in training the supervised ML classifiers. We tested Method I of extracting RS features (relative), because it was evident (Fig. 6) that acquired RS spectra pairs (control and burn) vary between different skin regions and the burn spectra closely follows the control spectra if acquired from the same spot. The difference between control and corresponding burn spectra helps us to capture only changes due to the inflicted burn. Method II (absolute RS features) is more practical as it may not be feasible to find co-located normal skin in burn-injured patients to implement Method I. Method II, based on inputs from experts, uses only 21 RS features compared to 54 features of Method I, and yields similar performance which is evident from the average AUC-ROC values across all 3 ML classifiers (RS (I): AUC-ROC=0.92; RS (II): AUC-ROC=0.93). These results demonstrate that Method II (using only the burn spectra) can be adopted for clinical use.

The results obtained using ML-based classifiers show value in combining RS-OCT (average AUC-ROC=0.94), although RS (average AUC-ROC=0.93) by itself provides highly accurate

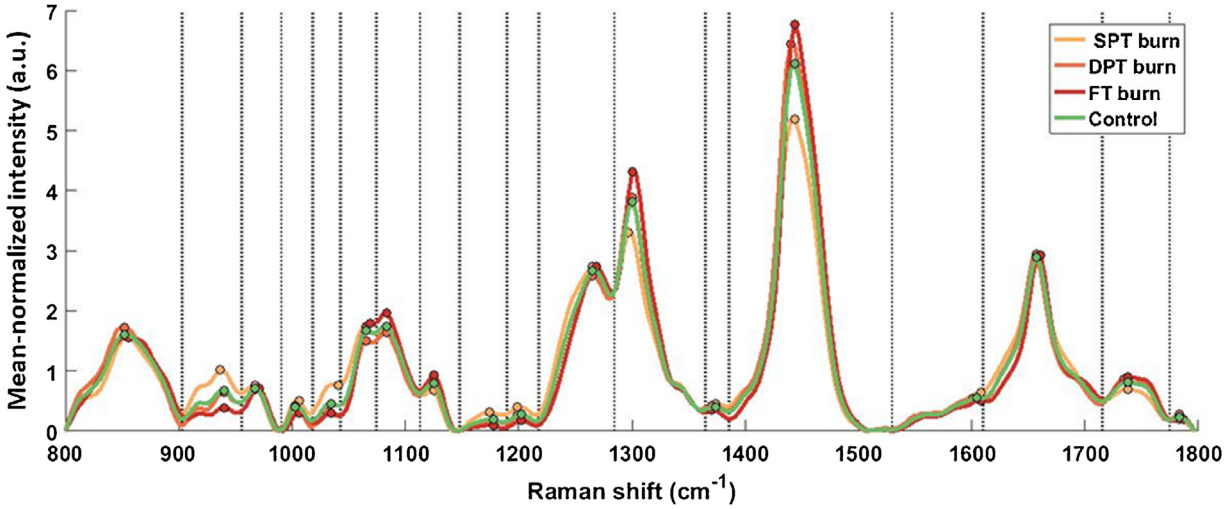

Fig. 7 – Mean normalized and averaged Raman spectra across control and all types of burns. Dotted delimiters indicate segment boundaries (18 segments), which were divided using local minima (valley points), obtained from peak detection algorithm.

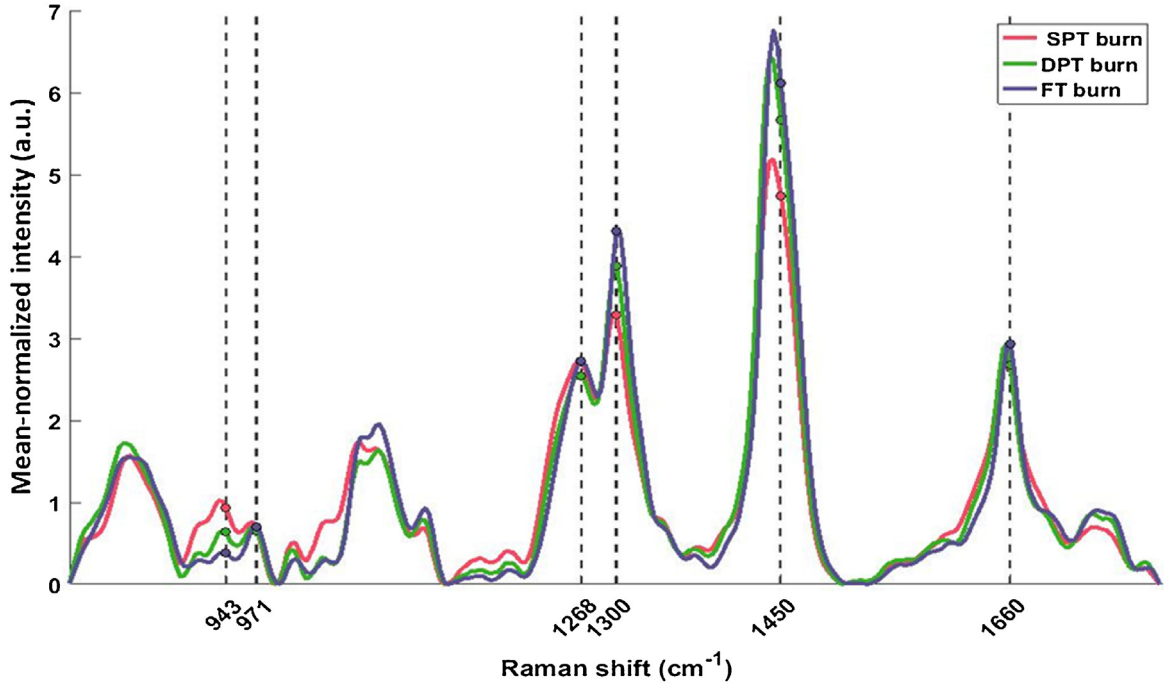

Fig. 8 – Mean normalized Raman spectra for only burn wounds. Dashed-lines indicate clinically-relevant wavenumbers specific to skin tissue. Peak intensity ratios, *i.e.*, 943/971 cm^{-1} (NCαC/CC proline ring), 1300/1268 cm^{-1} (CH bending/Amide III), and 1450/1660 cm^{-1} (CH_2 bending/Amide ICO stretch), were computed. The relevant peaks are marked as colored circles on the plot.

Table 2 – Method I: Classification performance metrics (area under ROC curve [AUC-ROC] and accuracy) of three classifiers on OCT, RS (relative to control), and RS-OCT features. The classifiers are implemented to solve a 3-class problem (Class 1: FT; Class 2: DPT and Class 3: ST).

Performance measures	Logistic regression			Linear SVM			Random forest		
	OCT	RS	RS-OCT	OCT	RS	RS-OCT	OCT	RS	RS-OCT
AUC-ROC: Class-1 (FT burns)	0.90	0.95	0.96	0.84	0.88	0.91	0.92	0.96	0.98
AUC-ROC: Class-2 (DPT burns)	0.80	0.87	0.89	0.64	0.82	0.85	0.76	0.90	0.89
AUC-ROC: Class-3 (ST burns)	0.84	0.98	0.99	0.76	0.93	0.94	0.85	0.97	0.97
Average AUC-ROC (FT, DPT and ST burns)	**0.87**	**0.94**	**0.95**	**0.76**	**0.88**	**0.90**	**0.85**	**0.95**	**0.95**
Overall accuracy (%)	**67.11**	**80.26**	**86.84**	**68.42**	**84.21**	**86.88**	**69.74**	**81.58**	**78.95**

Bold values signifies Average AUC-ROC and overall accuracy.

Table 3 – Method II: Classification performance metrics (area under ROC curve [AUC-ROC] and accuracy) of three classifiers on OCT, RS (absolute-only burn), and RS-OCT features. The classifiers are implemented to solve a 3-class problem (Class 1: FT; Class 2: DPT and Class 3: ST).

Performance measures	Logistic regression			Linear SVM			Random forest		
	OCT	RS	RS-OCT	OCT	RS	RS-OCT	OCT	RS	RS-OCT
AUC-ROC: Class-1 (FT burns)	0.90	0.92	0.93	0.84	0.89	0.91	0.92	0.95	0.97
AUC-ROC: Class-2 (DPT burns)	0.80	0.89	0.86	0.64	0.81	0.87	0.76	0.92	0.93
AUC-ROC: Class-3 (ST burns)	0.84	0.99	0.99	0.76	0.98	0.98	0.85	0.99	0.99
Average AUC-ROC (FT, DPT and ST burns)	**0.87**	**0.94**	**0.93**	**0.76**	**0.90**	**0.92**	**0.85**	**0.95**	**0.96**
Overall accuracy (%)	**67.11**	**82.89**	**84.21**	**68.42**	**86.84**	**89.47**	**69.74**	**85.53**	**85.53**

Table 4 – Top 10 RS-OCT features chosen most frequently by different feature selection techniques on WEKA. Entropy and Depth estimates are calculated from OCT images. The remaining features in the table are the features derived from Raman spectra of burn wounds relative to corresponding segments of control spectra (Method I).

RS-OCT feature selection
Entropy
Segment-2: cosine similarity
Segment-17: difference of AUC
Segment-9: cosine similarity
Depth
Segment-7: peak ratio
Segment-10: cosine similarity
Segment-5: difference of AUC
Segment-5: cosine similarity
Segment-16: cosine similari`ty

classification results compared to OCT alone (average AUC-ROC=0.83). The high accuracy provided by RS features alone, in both methods, implies the significance of biochemical information. To identify the relevance of OCT, feature selection methods *via* WEKA were implemented to identify top 10 features from RS-OCT pool of features. The 10 features that were frequently selected by all the feature selection methods are shown in Table 4. It is evident from Table 4, that depth and entropy computed from OCT images are two of the top 10 features selected when pooling RS and OCT features. Frequent selection of OCT-derived depth information may be justified, as depth variation among SPT (0.69 ± 0.07 mm), DPT (0.52 ± 0.14 mm) and FT (0.44 ± 0.12 mm) wounds was significant ($p<0.01$). Similarly, the difference in mean entropy values was also statistically significant ($p<0.01$) (SPT=2.68 ± 0.22; DPT=2.09 ± 0.51; FT=1.86 ± 0.40). These results indicate that despite superior performance from RS, features derived from OCT cannot be discounted.

Classification power is dependent on discriminating power of input features, training data size, and the characteristic of the classification method [62]. To avoid any overfitting while training the ML models, 10-fold cross validation was used and the resulting accuracies and AUC-ROC were reported [63]. In order to evaluate the efficacy of RS and OCT features, we employed simple to complex ML models (LR, Linear SVM, and RF). In our case, LR (AUC-ROC=0.92) and RF (AUC-ROC=0.92) slightly outperform linear SVM (AUC-ROC=0.86) method (Tables 2 and 3) across all types of input features (RS, OCT, RS-OCT) (Fig. 5). Here, we have performed a feature-level classification by concatenating the features acquired from RS and OCT modalities before classification by a ML algorithm. Decision-level classification involves fusing the classification results obtained separately on RS and OCT features [64]. This classification scheme will be explored in future *in vivo* animal studies.

The average AUC-ROC obtained in the current work for classification of burn types is 0.94 (across all 3 candidate classifiers for RS(I)-OCT and RS(II)-OCT), which is higher in comparison to a previous study that employed PSI and OCT imaging modalities [24]. The investigators obtained an aggregate AUC-ROC of 0.86 across three distinct burn wound categories (superficial thickness, partial thickness, and full thickness burns) with the PSI-OCT combination. However, the study was performed on *in vivo* porcine models. A good reproducibility in terms of performance by the current RS-OCT method across *in vivo* models will further strengthen the system's capability, and can serve as a lesser expensive option than PSI-OCT.

Features derived from Raman spectra, using both methods, yielded highly promising results on *ex vivo* skin. However, we need to further investigate its feasibility on *in vivo* skin, where the outcomes of the burnt skin may differ due to the presence of blood flow, and also address the influence of other factors such as occurrence of edema, inter- and intra-variation in skin thickness, skin color, presence of hair, and respiration. Furthermore, the weak nature of inelastic Raman scattering makes it a point-and-shoot method. This limitation of RS can be overcome by using OCT, which can acquire skin depth profile across a length/region, to guide the RS probe. In future, we intend to carry out the analysis for *in vivo* porcine studies, develop a unified RS-OCT probe and algorithms to perform OCT guided RS. Replacing manual marking of RS acquisition locations with advanced registration algorithms will be focused. We also aim to refine our algorithms adjusting for any variations due to physiological parameters in *in vivo* porcine models, and optimize ML methods, accordingly, to classify the degree of burns.

In summary, through this work on *ex vivo* porcine skin we demonstrated the usefulness of combined approach of RS-OCT to diagnose clinically relevant superficial partial-, deep partial- and full-thickness burns with an average accuracy of 85% and AUC-ROC=0.94. The high performance reported on *ex vivo* skin data provides encouragement to perform *in vivo* animal model studies to gather evidence on RS-OCT discriminating power for burn wound diagnosis (within 48 h post injury).

Role of the funding source

This work was supported by the US Army Medical Research and Materiel Command under Contract No. W81XWH-17-C-0060. The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation. US Army Medical Research and Materiel Command had no involvement in preparation or submission of this manuscript.

Conflict of interest

Authors do not have any conflict of interest.

Acknowledgments

We express our deep gratitude to the entire team of *Biophotonics Center* at Vanderbilt University (Nashville, TN), for acquiring and sharing RS and OCT data required for the assessment. We

greatly appreciate the support and assistance provided by the team during the course of the analytical work.

REFERENCES

[1] Centers for Disease Control and Prevention, U.S. Prevention. Department of Health and Human Services. Fire deaths and injuries: fact sheet. Web. 2014.
[2] American Burn Association. Burn incidence and treatment in the united states: fact sheet. Web. 2014.
[3] Gray DT, Pine RW, Harnar TJ, Marvin JA, Engrav LH, Heimbach DM. Early surgical excision versus conventional therapy in patients with 20 to 40 percent burns. Am J Surg 1982;144(1):76–80.
[4] Sjöberg F, Danielsson P, Andersson L, Steinwall I, Zdolsek J, Östrup L, Monafo W. Utility of an intervention scoring system in documenting effects of changes in burn treatment. Burns 2000;26(6):553–9.
[5] Papini Remo. Management of burn injuries of various depths. BMJ 2004;329(7458):158–60.
[6] Heimbach D, Afromowitz M, Engrav L, Marvin J, Perry B. Burn depth estimation—man or machine. J Trauma 1984;24 (5):373–8.
[7] Riordan Colin L, McDonough Mark, Davidson Jeffrey M, Corley Randall, Nanney Lillian B, et al. Noncontact laser doppler imaging in burn depth analysis of the extremities. J Burn Care Rehabil 2001;24(4):177–8.
[8] Godina M, Derganc M, Brcic A. The reliability of clinical assessment of the depth of burns. Burns 1978;4(2):92–6.
[9] Milseki W, Atiles L, Purdue G, Kagan R, Saffle J, Herndon D, et al. Serial measurements increase the accuracy of laser. J Burn Care Rehabil 2003;24(4):187–91.
[10] Monstrey S, Hoeksema H, Verbelen J, Pirayesh A, Blondeel P. Assessment of burn depth and burn wound healing potential. Burns 2008;34(6):761–9.
[11] Jan SN, Khan FA, Bashir MM, Nasir M, Ansari HH, Shami HB, et al. Hamid Hussain Ansari, Hussan Birkhez Shami, et al. Comparison of Laser Doppler Imaging (LDI) and clinical assessment in differentiating between superficial and deep partial thickness burn wounds. Burns 2018;44(2):405–13.
[12] Pape S, Skouras C, Byrne P. An audit of the use of laser Doppler imaging (LDI) in the assessment of burns of intermediate depth. Burns 2001;27(3):233–9.
[13] Monstrey S, Hoeksema H, Spence R, Jeng J, Wilson D, Pape S. Clinical assessment of burn wound healing potential by laser Doppler imaging (LDI): results of a prospective, multi centre study. Abstract Book 38th Annual Meeting Am. Burn Association 2006;27(2):4–7.
[14] Kaiser M, Yafi A, Cinat M, Choi B, Durkin AJ. Noninvasive assessment of burn wound severity using optical technology: a review of current and future modalities. Burns 2011;7(3):377–86.
[15] Droog EJ, Steenbergen W, Sjöberg F. Measurement of depth of burns by laser Doppler perfusion imaging. Burns 2001;27 (6):561–8.
[16] Hoeksema H, Van de Sijpe K, Tondu T, Hamdi M, Van Landuyt K, Blondeel P, et al. Accuracy of early burn depth assessment by laser Doppler imaging on different days post burn. Burns 2009;35(1):36–45.
[17] Moor Instruments, Laser Doppler blood flow assessment. MoorLDI2-BI FAQs. August 2017. https://www.nice.org.uk/guidance/mtg2.
[18] Thomas Resch R, Rachel Drake M, Stephen Helmer D, Gary Jost D, Jacqueline Osland S. Estimation of burn depth at burn centers in the united states: a survey. J Burn Care Res 2014;35 (6):491–7.
[19] Schmitt Joseph M. Optical coherence tomography (OCT): a review. IEEE J Sel Top Quantum Electron 1999;5(4):1205–15.
[20] Podoleanu Adrian Gh, Rogers John A, Jackson David A, Dunne Shane. Three dimensional OCT images from retina and skin. Opt Express 2000;7(9):292–8.
[21] Olsen J, Themstrup L, Jemec GB. Optical coherence tomography in dermatology. G Ital Dermatol Venereol 2015;150(5):603–15.
[22] Silver R, Helms A, Fu W, Wang H, Diaconu D, Loyd CM, et al. Using optical coherence tomography for the longitudinal non-invasive evaluation of epidermal thickness in a murine model of chronic skin inflammation. Skin Res Technol 2012;18(2):225–31.
[23] Welzel Julia. Optical coherence tomography in dermatology: a review. Skin Res Technol 2001;7(1):1–9.
[24] Ganapathy P, Tamminedi T, Qin Y, Nanney L, Cardwell N, Yadegar J, et al. Dual-imaging system for burn depth diagnosis. Burns 2014;40(1):67–81.
[25] Pierce, Mark C, Robert Sheridan L, Hyle Park B, Barry Cense, Johannes de Boer F. Collagen denaturation can be quantified in burned human skin using polarization-sensitive optical coherence tomography. Burns 2004;30(6):511–3.
[26] Srinivas SM, de Boer JF, Park H, Keikhanzadeh K, Huang HF, Zhang J, et al. Determination of burn depth by polarization-sensitive optical coherence tomography. J Biomed Opt 2004;9 (1):207–13.
[27] Mariampillai A, Standish BA, Moriyama EH, Khurana M, Munce NR, et al. Speckle variance detection of microvasculature using swept-source optical coherence tomography. Opt Lett 2008;33(13):1530–2.
[28] Zhang A, Zhang Q, Chen CL, Wang RK. Methods and algorithms for optical coherence tomography-based angiography: a review and comparison. J Biomed Opt 2015;20 (10):100901.
[29] Gao SS, Jia Y, Zhang M, Su JP, Liu G, Hwang TS, et al. Optical coherence tomography angiography. Invest Ophthalmol Vis Sci 2016;57(9):OCT27–36.
[30] Schuh S, Holmes J, Ulrich M, Themstrup L, Jemec GB, Welzel J, et al. Imaging blood vessel morphology in skin: dynamic optical coherence tomography as a novel potential diagnostic tool in dermatology. Dermatol Ther 2017;7(2):187–202.
[31] Themstrup L, De Carvalho N, Nielsen SM, Olsen J, Ciardo S, et al. In vivo differentiation of common basal cell carcinoma subtypes by microvascular and structural imaging using dynamic optical coherence tomography. Exp Dermatol 2018;27(2):156–65.
[32] Gong P, McLaughlin RA, Liew YM, Munro PR, Wood FM, Sampson DD. Assessment of human burn scars with optical coherence tomography by imaging the attenuation coefficient of tissue after vascular masking. J Biomed Opt 2013;19 (2):021111.
[33] Gong P, Es' haghian S, Wood FM, Sampson DD, McLaughlin RA. Optical coherence tomography angiography for longitudinal monitoring of vascular changes in human cutaneous burns. Exp Dermatol 2016;25(9):722–4.
[34] Gong P, Es' haghian S, Harms KA, Murray A, Rea S, McLaughlin RA, et al. Optical coherence tomography for longitudinal monitoring of vasculature in scars treated with laser fractionation. J Biophotonics 2016;9(6):626–36.
[35] Iftimia N, Ferguson RD, Mujat M, Patel AH, Zhang EZ, Fox W, et al. Combined reflectance confocal microscopy/optical coherence tomography imaging for skin burn assessment. Biomed Opt Express 2013;4(5):680–95.
[36] Choo-Smith LP, Edwards HG, Endtz HP, Kros JM, Heule F, Barr H, et al. Medical applications of Raman spectroscopy: from proof of principle to clinical implementation. Biopolymers 2002;67(1):1–9.
[37] Hettiaratchy S, Dziewulski P. Pathophysiology and types of burns. BMJ 2004;328(7453):1427–9.

[38] Caspers PJ, Lucassen GW, Wolthuis R, Bruining HA, Puppels GJ. In vitro and in vivo Raman spectroscopy of human skin. Biospectroscopy 1998;4(S5):S31–40.
[39] Lui Harvey, Zhao Jianhua, McLean David, Zeng Haishan. Real-time Raman spectroscopy for in vivo cancer diagnosis. Cancer Res 2012;72(10):2491–500.
[40] Caspers P, Bruining Puppels G, Ucassen G, Carter E. In vivo confocal raman microspectroscopy of the skin: noninvasive determination of molecular concentration profiles. J Invest Dermatol 2001;116(3):434–42.
[41] Lieber CA, Majumder SK, Ellis DL, Billheimer DD, Mahadevan-Jansen A. In vivo nonmelanoma skin cancer diagnosis using Raman microspectroscopy. Lasers Surg Med 2008;40(7):461–7.
[42] Boncheva M, de Sterke J, Caspers PJ, Puppels GJ. Depth profiling of stratum corneum hydration in vivo: a comparison between conductance and confocal Raman spectroscopic measurements. Exp Dermatol 2009;18(10):870–6.
[43] Peterson J, Okagbare P, Rosa S, Cilwa K, Persosky J, Eboda O, et al. Early detection of burn induced heterotopic ossification using transcutaneous Raman spectroscopy. Bone 2013;54(1):28–34.
[44] Ponticorvo A, Burmeister DM, Yang B, Choi B, Christy RJ, Durkin AJ. Quantitative assessment of graded burn wounds in a porcine model using spatial frequency domain imaging (SFDI) and laser speckle imaging (LSI). Biomed Opt Express 2014;5(10):3467–81.
[45] Caspers PJ, Lucassen GW, Puppels GJ. Combined in vivo confocal Raman spectroscopy and confocal microscopy of human skin. Biophys J 2003;85(1):572–80.
[46] Patil C, Krishnamoorthi H, Ellis D, Van Leeuwen T, Mahadevan-Jansen A. A clinical instrument for combined RS-OCT of skin cancers. Lasers Surg Med 2011;43(2):143–51.
[47] Liu C, Ji Qi, Lu J, Wang S, Wu C, Shih W, et al. Improvement of tissue analysis and classification using optical coherence tomography combined with Raman spectroscopy. J innov Opt Health Sci 20158(2) 1550006-1.
[48] Pathophysiology of thermal burn injury. Civic plus. 2007.
[49] Thomsen S. Pathologic analysis of photothermal and photomechanical effects of laser-tissue interactions. Photochem Photobiol 1991;53(6):825–35.
[50] Chvapil M, Speer DP, Owen JA, Chvapil TA. Identification of the depth of burn injury by collagen stainability. Plast Reconstr Surg 1984;73(3):438–41.
[51] Pence I, Mahadevan-Jansen A. Clinical instrumentation and applications of Raman spectroscopy. Chem Soc Rev 2016;45(7):1958–79.
[52] Lieber Chad A, Mahadevan-Jansen Anita. Automated method for subtraction of fluorescence from biological Raman spectra. Appl Spectrosc 2003;57(11):1363–7.
[53] Savitzky A, Golay MJE. Smoothing and differentiation of data by simplified least squares procedures. Anal Chem 1964;36(8):1627–39.
[54] Otsu Nobuyuki. A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern 1979;9(1):62–6.
[55] R.C. Gonzalez, R.E. Woods. Representation and description. In: Digital Image Processing. 3rd ed. Pearson Education. 2008 pp. 795–860..
[56] Gniadecka M, Philipsen PA, Sigurdsson S, Wessel S, Nielsen OF, Christensen DH, et al. Melanoma diagnosis by Raman spectroscopy and neural networks: structure alterations in proteins and lipids in intact cancer tissue. J Invest Dermatol 2004;122(2):443–9.
[57] le Cessie S, van Houwelingen JC. Ridge estimators in logistic regression. Appl Stat 1992;41(1):191–201.
[58] Chang Chih-Chung, Lin Chih-Jen. LIBSVM — a library for support vector machines. 2001 http://www.csie.ntu.edu.tw/?cjlin/libsvm/.
[59] Breiman Leo. Random forests. Mach Learn 2001;45(1):5–32.
[60] Eibe Frank, Mark A. Hall, and Ian H. Witten. The WEKA Workbench. Online Appendix for “Data Mining: Practical Machine Learning Tools and Techniques”, Morgan Kaufmann, Fourth Edition, 2016. https://www.cs.waikato.ac.nz/ml/weka/Witten_et_al_2016_appendix.pdf.
[61] Fawcett T. An introduction to ROC analysis. Pattern Recognit Lett 2006;27(8):861–74.
[62] Osisanwo FY, Akinsola JET, Awodele O, Hinmikaiye JO, Olakanmi O, Akinjobi J. Supervised machine learning algorithms: classification and comparison. Int J Comput Trends Technol 2017;48(3):128–38.
[63] Refaeilzadeh P, Tang L, Liu H. Cross-validation. In: Liu L, Özsu MT, editors. Encyclopedia of database systems. Boston, MA: Springer; 2009.
[64] Hall D, Llinas J. An introduction to multisensor data fusion. Proc IEEE 1997;85(1):6–23.