Burn wound classification model using spatial frequency-domain imaging and machine learning

Rebecca Rowland
Adrien Ponticorvo
Melissa Baldado
Gordon T. Kennedy
David M. Burmeister
Robert J. Christy
Nicole P. Bernal
Anthony J. Durkin

SPIE.

Burn wound classification model using spatial frequency-domain imaging and machine learning

Rebecca Rowland,[a] Adrien Ponticorvo,[a] Melissa Baldado,[a] Gordon T. Kennedy,[a] David M. Burmeister,[b] Robert J. Christy,[b] Nicole P. Bernal,[c] and Anthony J. Durkin[a,d,*]
[a]University of California, Irvine, Beckman Laser Institute and Medical Clinic, Irvine, California, United States
[b]United States Army Institute of Surgical Research, San Antonio, Texas, United States
[c]UC Irvine Regional Burn Center, Department of Surgery, Orange, California, United States
[d]University of California, Irvine, Department of Biomedical Engineering, Irvine, California, United States

Abstract. Accurate assessment of burn severity is critical for wound care and the course of treatment. Delays in classification translate to delays in burn management, increasing the risk of scarring and infection. To this end, numerous imaging techniques have been used to examine tissue properties to infer burn severity. Spatial frequency-domain imaging (SFDI) has also been used to characterize burns based on the relationships between histologic observations and changes in tissue properties. Recently, machine learning has been used to classify burns by combining optical features from multispectral or hyperspectral imaging. Rather than employ models of light propagation to deduce tissue optical properties, we investigated the feasibility of using SFDI reflectance data at multiple spatial frequencies, with a support vector machine (SVM) classifier, to predict severity in a porcine model of graded burns. Calibrated reflectance images were collected using SFDI at eight wavelengths (471 to 851 nm) and five spatial frequencies (0 to 0.2 mm^{-1}). Three models were built from subsets of this initial dataset. The first subset included data taken at all wavelengths with the planar (0 mm^{-1}) spatial frequency, the second comprised data at all wavelengths and spatial frequencies, and the third used all collected data at values relative to unburned tissue. These data subsets were used to train and test cubic SVM models, and compared against burn status 28 days after injury. Model accuracy was established through leave-one-out cross-validation testing. The model based on images obtained at all wavelengths and spatial frequencies predicted burn severity at 24 h with 92.5% accuracy. The model composed of all values relative to unburned skin was 94.4% accurate. By comparison, the model that employed only planar illumination was 88.8% accurate. This investigation suggests that the combination of SFDI with machine learning has potential for accurately predicting burn severity. © The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: 10.1117/1.JBO.24.5.056007]

Keywords: machine learning; spatial frequency-domain imaging; multispectral and hyperspectral imaging; support vector machine; burns; spectroscopy.

Paper 180588RR received Oct. 17, 2018; accepted for publication May 2, 2019; published online May 27, 2019.

1 Introduction

Although a variety of optical imaging modalities have been investigated to assess burn wound depth, visual inspection by an experienced clinician is still the standard method for determining burn wound severity.[1,2] With about 500,000 patients admitted to burn units in the United States annually,[3,4] accurate and expedient treatment of burn wounds is integral to minimizing hospital stay and the risk of scarring or infection.[5] Superficial burns require monitoring and minimal dressing, as tissue damage only extends to the epidermis and the papillary dermis. Deep-partial and full thickness burns involve more aggressive treatment (i.e., debridement and grafting) because tissue damage extends into the reticular dermis and underlying subcutaneous fat. At postburn time points earlier than 24 h, the accuracy of clinical impression in differentiating between superficial-partial thickness and deep-partial thickness burn severity is between 60% and 80%.[6] While accuracy improves the longer the clinician waits to make a diagnosis, outcomes are improved the sooner debridement and grafting is performed.[4,5] Additionally, failure to graft regions of full or deep-partial burns

has the potential to result in infection, hypertrophic scarring, contraction, tissue necrosis, delayed wound healing, and a longer hospital stay.[7]

The need to accurately and quickly determine which burns require grafting has led to the investigation of various optical imaging modalities for identifying burn severity. For example, multispectral imaging has been used in preclinical models by implementing spatially uniform (planar) illumination to measure reflectance at different wavelengths.[8,9] Each wavelength samples different depths within tissue and subsequently shows a change in signal in the presence of tissue damage.[10–12]

Spatial frequency-domain imaging (SFDI) has also been used to determine burn severity through the measurement of absorption and scattering properties of burn tissue in animal models.[13–16] SFDI is a unique, noncontact, wide-field imaging modality that provides quantitative, spatial maps of tissue optical properties based on diffuse optical spectroscopic principles. Unlike multispectral imaging, which only uses planar (0-mm^{-1} spatial frequency) illumination, this modality utilizes a range of spatial frequencies to interrogate different depths and calculate tissue scattering. The measured reduced scattering coefficient indicates the structure and density of underlying tissue, whereas the wavelength-dependent absorption coefficient can be

converted to concentrations of oxyhemoglobin, deoxyhemoglobin, and water.

Previously, we have used SFDI to quantitatively and noninvasively characterize burn wound severity, and track burn wound healing over time.[13–17] SFDI was used on a rat model to track wound healing by monitoring the changes in scattering, water fraction, oxyhemoglobin concentration, and deoxyhemoglobin concentration as the vasculature and collagen reformed in damaged regions.[15] Preceding work in both rat and swine model established a correlation between burn depth as determined by histology and reduced scattering determined using SFDI. This is likely a consequence of thermal denaturation of collagen that affects the arrangement of collagen fibers.[13,16] The results of these investigations have suggested that scattering can be used to accurately differentiate superficial-partial, deep-partial, and full thickness burns at the 24-h postburn time point, which is commonly used as a temporal reference point when comparing predictive performance of various burn wound severity assessment tools.[17] While the approach of visualizing SFDI burn data at multiple parameters appears to be promising, initial investigations into machine learning have indicated an improved way of employing SFDI data within a burn severity assessment context that enables rapid and more nuanced visualization of burns.[18]

Machine learning has been utilized in previous SFDI studies.[19,20] Regressive machine learning models are used to estimate optical properties and trained with SFDI-derived reflectance values. Datasets made with varying absorption and reduced scattering values are built through computationally intensive Monte Carlo modeling, and the machine learning model outputs are compared against other proven quantitative modeling techniques, such as diffuse approximation and Monte Carlo-derived look-up-tables.

Machine learning algorithms have recently been employed in conjunction with wide-field imaging modalities to segment and classify the severity of burn wounds. In supervised machine learning-based approaches, models are trained using image features that correspond to a known classification output. Once trained on a representative set of data, the model can then be applied to data acquired in the same way, but where the true output is unknown, and examined for its potential to inform the output. The support vector machine (SVM) is one such model, which separates data within a multidimensional space by determining the plane that optimally separates classes at adjacent values.[21] Using an approach based on machine learning, parameters can be combined to predict the burn severity for each individual pixel within the image of a burn region. Previously, machine learning models have been trained and tested on data from multispectral and hyperspectral imaging devices that employ conventional planar imaging. These methods resulted in classification accuracies between 63% and 76.9%.[9,22] Various filtering techniques have also been applied to these datasets, increasing the accuracy of burn severity prediction to 77.8%.[23]

In the investigation presented herein, we examine the ability of a cubic SVM classification model to predict burn wound severity using calibrated reflectance data from multiple wavelengths and spatial frequencies obtained via SFDI. Machine learning-based classification potentially enables rapid clinical interpretation of SFDI data by compiling multiple parameters into a single output that does not require differentiation into scattering at multiple wavelengths or individual chromophores.

Tangentially, we compare the predictive performance of an SVM model that uses only conventional planar multispectral image data to a model that uses data at multiple spatial frequencies.

2 Methodology

2.1 Animals

The data employed in this investigation were obtained from imaging studies that were carried out in a porcine model of graded burn wounds in compliance with the UC Irvine Institutional Animal Care and Use Committee (IACUC protocol #2015-3154). Two Yorkshire pigs were used through the duration of the study. Each animal was allowed to acclimate for 7 days after their arrival to the facility, prior to experimentation. They were initially anesthetized with tiletamine-zolazepam (Telazol, 6 mg/kg). The animals were intubated and anesthesia was maintained throughout the experiment with 1% to 3% isofluorane. The respiration rate was controlled to 10 breaths per min. Before imaging, hair on the animal's dorsum was clipped, and the area cleaned of any debris. At the end of each imaging period, burn wounds were covered with saline-soaked nonadherent gauze (Telfa, Tyco Healthcare, Mansfield, Massachusetts) and held in place with an Ioban™ dressing (3M, St. Paul, Minnesota).

2.2 Creation of Controlled, Graded Burn Wounds

Controlled, graded burn wounds were created using a custom burn tool, as described in previous studies.[13,14] This burn tool composed of 3-cm-diameter brass rods, was heated to 100°C within an aluminum block inside an Isotemp™ dry bath incubator (Thermo Fisher Scientific Inc., Pittsburgh, Pennsylvania). Fiducial markers were placed with a surgical pen at 2.5 cm from the center of each burn region, to assist with the placement of the burn tool. To create the burns, one rod was held in a spring loaded apparatus that enables the user to safely apply a controlled, constant pressure while creating each burn. The burn wounds were placed in two rows along the prepared region on the pig's dorsum. Each pig received 16 burns, two of each contact time, 1 cm from the spine, and 3 cm from each adjacent burn. Contact times of the tool to with skin were 5, 10, 15, 20, 25, 30, 35, and 40 s, creating a range of superficial-partial, deep-partial, and full thickness burns. No debridement or grafting was performed, and the burns were allowed to heal for 28 days without surgical intervention besides biopsy collection.

2.3 Clinical Assessment and Burn Classification

A clinical evaluation of each burn was used to define the true classifications of each burn, before any data were compiled with a machine learning model. A board certified burn surgeon with 15 years of burn-related experience used visual and tactile examination to assess burn severity 28 days after the burn. At 28 days, wound contraction and hypertrophic scarring were apparent in regions of full and deep-partial thickness burns, and the regions of tissue that received superficial-partial burns re-epithelialized without contraction or scarring. The clinician judged each burn and identified severe burn areas that, in normal practice, would have been debrided and grafted within the first week postburn. They also defined other burn areas as regions where grafting would not be required. When making this
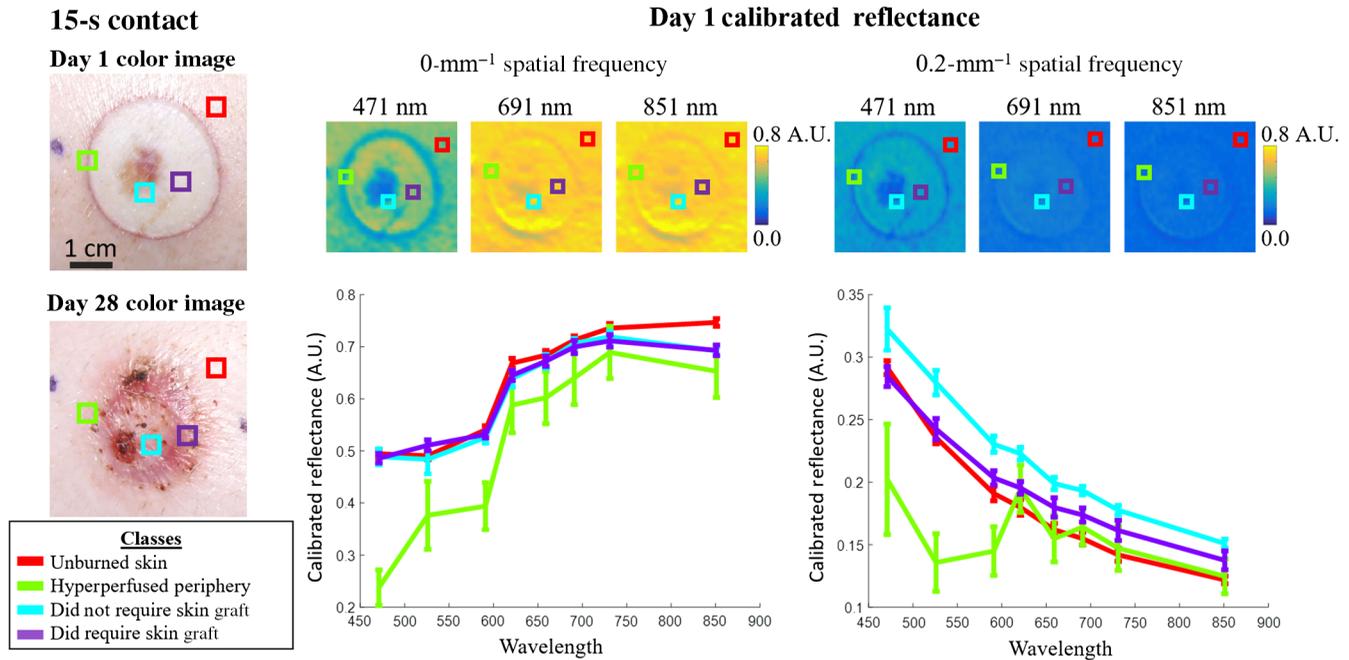
Fig. 1 ROIs corresponding to four classification regions (unburned skin, hyperperfused burn periphery, burn that did not require grafting, and burn that should have been grafted) were chosen from calibrated reflectance images collected 24 h after the burn. Displayed are the images taken with SFDI at 0.00- (multispectral imaging) and 0.20-mm⁻¹ spatial frequencies with 471-, 691-, and 851-nm wavelengths, as well as graphs plotting the training regions over all wavelengths.

assessment, the clinician was blinded to the contact times used when creating the burns. This assessment was treated as the ground truth and used as the known outcome while training the machine learning models in order to gauge a model's classification accuracy. The two burn outcomes served as two of the classes for the machine learning model to predict. The hyperperfused border of burn and the unburned skin were also treated as separate classes. The color images in Fig. 1 indicate examples of these four classes at day 1 and day 28.

2.4 Spatial Frequency-Domain Imaging

SFDI combines projections of sinusoidal patterns at multiple wavelengths with multispectral imaging to measure depth-resolved optical properties over a wide field of view.[24,25] The addition of spatially modulating light at different wavelengths changes the volume of tissue interrogated, and results in the measurement of absorption and reduced scattering. In this study, we employed the Reflect RS® (Modulated Imaging, Inc., Irvine, California), a commercial SFDI instrument capable of imaging optical properties of tissues over large fields of view $(20 \times 15 \text{ cm})$[24] (shown in Fig. 2). The measurements were taken 24 h after the placement of the burns. In addition to imaging each burn, a measurement was made of a tissue-simulating optical phantom having known optical properties in order to calibrate SFDI data. SFDI data for each region of tissue were acquired within 35 s. The imaging device was placed at a working height of 32 cm and centered so each measurement captured two neighboring burn regions. Illumination was provided by eight light-emitting diodes (LEDs) at 471-, 526-, 591-, 621-, 659-, 691-, 731-, and 851-nm wavelengths and projected at spatial frequencies of 0 (conventional planar illumination), 0.05, 0.10, 0.15, and 0.20 mm⁻¹, as described in our previous

work.[13,14] Each nonzero spatial frequency pattern was projected at three phases, separated by 120 deg. All told, a single measurement comprised of 120 images, each corresponding to a combination of one of eight wavelengths, five spatial frequencies, and three phases.

SFDI data acquisition and analysis were performed within the MI Analysis v1.14.21 software suite, provided with the Reflect RS™. Using this analysis software, raw measurement reflectance data and raw calibration phantom reflectance data at multiple phases were demodulated at each combination of the eight wavelengths and five spatial frequencies into 40 total images of calibrated reflectance, as described previously.[26] In previous studies, models of light propagation in turbid media were used to convert calibrated reflectance data into optical properties and chromophore concentrations.[27–30] For this study, only calibrated reflectance data were used in the training and testing of cubic SVM classification algorithms.

2.5 Color Photography

After each SFDI measurement, color photographs were taken of each burn with a 14-megapixel digital camera (NEX-3, Sony Corporation of America, New York, New York).

2.6 Development of Classifier based on Support Vector Machine

A cubic SVM model was used to classify burn regions. This algorithm segmented data by identifying a hyperplane that optimally separated classes within the data at the point where classes were most dissimilar.[31–33] In the case of the cubic SVM, the hyperplane is described with a cubic polynomial, which could account for nonlinear separation between
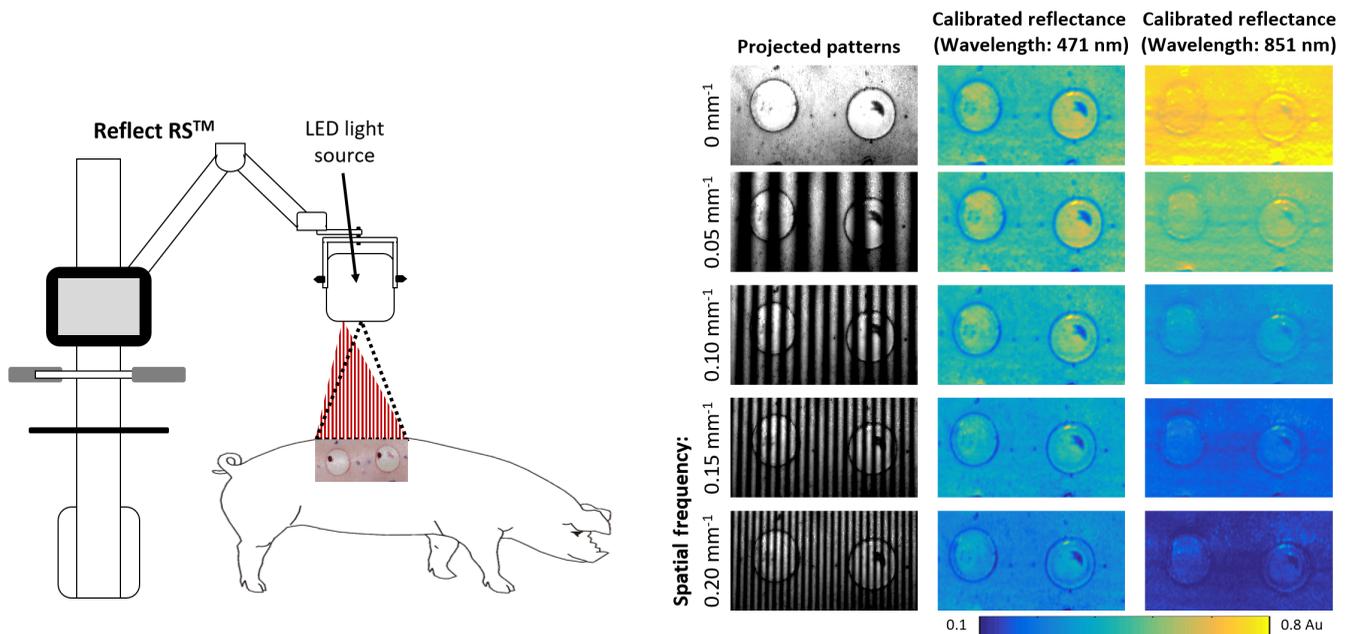
Fig. 2 A schematic of the SFDI device. Eight LEDs (wavelengths 471 to 851 nm) project sinusoidal patterns at five spatial frequencies (0 to 0.2 mm^{-1}). Raw images are converted to calibrated reflectance.

classes. Optimization and analysis were performed using the Statistics and Machine Learning Toolbox™ within MATLAB 2017a (Mathworks, Inc., Natick, Massachusetts). The cubic SVM model is a supervised classification that requires training data in order to learn how to correctly classify data, in this case, regions of tissue that had undergone thermal insult. First, the clinical assessment and day 28 color photography were evaluated for areas that represented one of four tissue health classes describing burn severity. The first class (marked in red in Fig. 1) characterized unburned skin that never received thermal injury. The second class (marked in green in Fig. 1) referred to the hyperperfused region bordering the burn. The third class (marked in blue in Fig. 1) categorized regions of the burn wound that re-epithelialized without scarring at the day 28 time point and would not have required a skin graft. The fourth class (marked in violet in Fig. 1) indicated areas where scarring, wound contraction, or lack of re-epithelialization were apparent, signifying a need for debridement and grafting.

A 5 × 5 pixel (1.4 × 1.4 mm) region of interest (ROI) was selected in the day 1 calibrated reflectance image, then averaged and labeled with the corresponding tissue health class as established at the 28-day postburn time point. ROIs were necessary since accurate point-by-point severity assessment of day 1 calibrated reflectance images could not be performed due to wound contraction by day 28. The ROI size was chosen as the largest size that could definitively spatially encompass all burn severity classes. The limiting case was the hyperperfused regions, which were typically only 2-mm wide. Finally, this same ROI was selected in every iteration of wavelength and spatial frequency. Examples of these classes and the regions of interest chosen for the training set can be seen in Fig. 1. For this model, 40 regions were labeled for each class type, resulting in 160 regions for the training dataset. All data regions from the training set were selected from a single pig. Data from the second pig were used to test the pretrained model.

Three iterations of the cubic SVM model were compared, in order to assess how the addition of spatially modulated light affects the accuracy of classifying burn wounds. The first version used a training set that included only a subset of the SFDI data, where images were taken with 0-mm^{-1} spatial frequency illumination at all eight wavelengths. This subset of the SFDI data represented a dataset that would be obtained from conventional planar multispectral imaging. The second model used the complete set of SFDI data, five spatial frequencies and eight wavelengths. The third model also included the complete set of SFDI data, but each calibrated reflectance value was normalized relative to unburned skin. These values were divided by an average calibrated reflectance value of all the ROIs classified as unburned skin. This model was used to compensate for relative differences in calibrated reflectance values between animal subjects.

The accuracy of each model was determined using k-fold cross-validation for 10 folds,[34–37] a leave-one-out validation method. Sixteen regions were removed from the training set, and a model was created using the remaining 144 data regions. The 16 regions were treated as a testing set, containing four regions from each class. The trained algorithm was applied to the testing set, and the outcome compared to the true classes of each test region. This was repeated 10 times for different training and testing sets, such that each data region was included in a testing subset once. The overall accuracy of each model was the average of these 10 repetitions and is summarized in Table 1 and the confusion matrices of Figs. 4(a)–6(a). These confusion matrices also describe the sensitivity and precision that each model predicts individual classes. Here, sensitivity is defined as proportion of correct positive predictions to all positive outcomes, and precision is defined as the proportion of correct positive predictions to the sum of all positive predictions.

Finally, each of these models was tested against day 1 calibration reflection data from 16 burns created on a separate pig. Figure 3 shows the results from each model on a single burn.

Table 1 Model accuracy for all four classes was determined by a 10-fold cross-validation of 160 ROIs.

Input data spatial frequencies	0 mm^{-1} (planar)	All spatial frequencies	All spatial frequencies (relative to unburned skin)
Cross-validation accuracy for training set (%)	88.8	92.5	94.4

3 Results

3.1 Clinical Assessment and Color Photography

Out of the set of burns used to derive the training set, seven out of the 16 burns were considered healed at the 28-day time point and would not have required a skin graft. The clinician noted that regions within the remaining nine burns were either deep-partial or full thickness burns. In normal practice, these regions would have been debrided and then received a skin graft. The color images were used to associate the classification of training set regions with the clinical assessment at day 28. The tissue health classification prediction provided by the cubic SVM models was also compared to the assessment on day 28.

3.2 Model Based on Conventional Planar Wide-Field Multispectral Imaging

The first model incorporated reflectance images where the illuminaiton was projected with 0-mm^{-1} spatial frequency (conventional planar illumination). The training set k-fold cross-validation accuracy was 88.8% across all classes. Within this

model, the prediction of hyperperfused burn regions was the most accurate. As seen in the confusion matrix, Fig. 4(a), this model identified 40 out of 40 points that described the hyperperfused border, with a precision of 97.6% and a sensitivity of 100%. This model was least accurate at predicting regions of superficial and superficial-partial thickness burn regions. The model correcty identified 32 of the 40 points of superficial burns within the training set, with a precision of 76.2% and a sensitivity of 80%.

3.3 Model from Combined Spatial Frequency Data

The following model combined images taken at all five spatial frequencies and eight wavelengths. This combined model was 92.5% accurate for all classifications. As shown in the confusion matrix in Fig. 5(a), this model accurately classified all 40 of the 40 training points for unburned skin. The precision was 95.2% and the sensitivity was 100%. The model was the least accurate when predicting burn wound regions that would not require a graft, correctly classifying 33 of 40 points for each class. For this classification, the precision was 89.2% and the sensitivity was 82.5%.

3.4 Model from Relative Dataset

The final model also used the combined dataset from all spatial frequency and wavelength images, but each point in the data was related to the average calibrated reflectance values for unburned skin, as shown in Figs. 6(a)–6(b). Across all classifications, this model was 94.4% accurate. As with the combined data model without normalization, this model accurately assessed all 40 of the training set points for unburned skin. The model's precision was 95.2% and sensitivity was 100%. The model also identified all 40 of the hyperperfused regions, with a precision of 97.6%
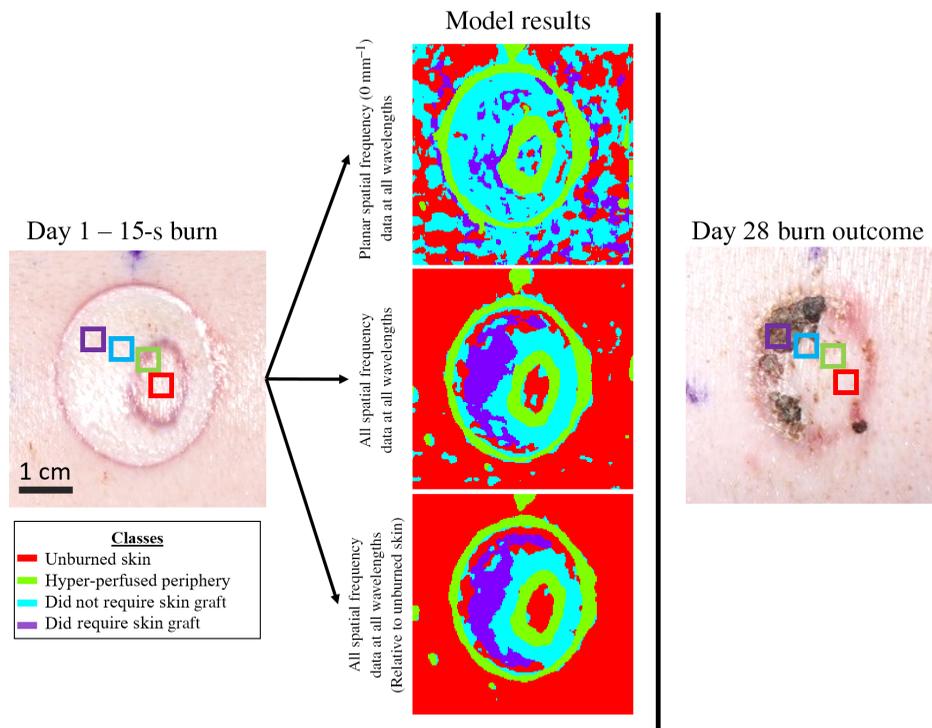

Fig. 3 Classification predictions from the three models on a 15-s burn. ROIs and day 28 color image indicate the true burn classification.
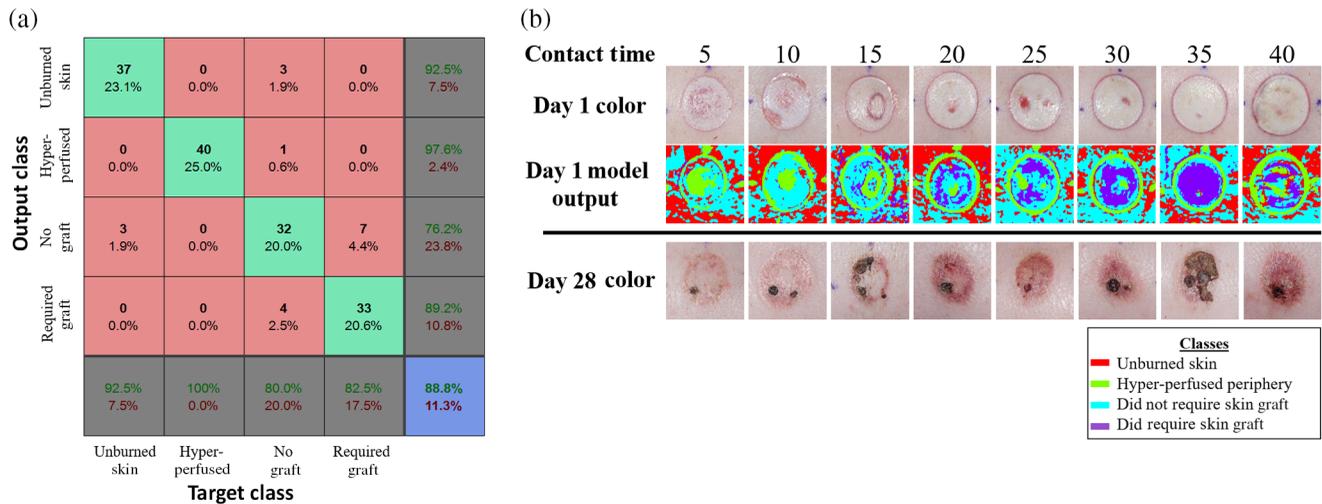
Fig. 4 (a) Confusion matrix of 10-fold cross-validation performed on the cubic SVM model created using 0-mm^{-1} spatial frequency data. (b) Cubic SVM model based on 0-mm^{-1} data, applied to day 1 calibrated reflectance imaging data of burns in a single pig (middle row). Corresponding day 28 color image (bottom row) shows the true outcome for each burn.
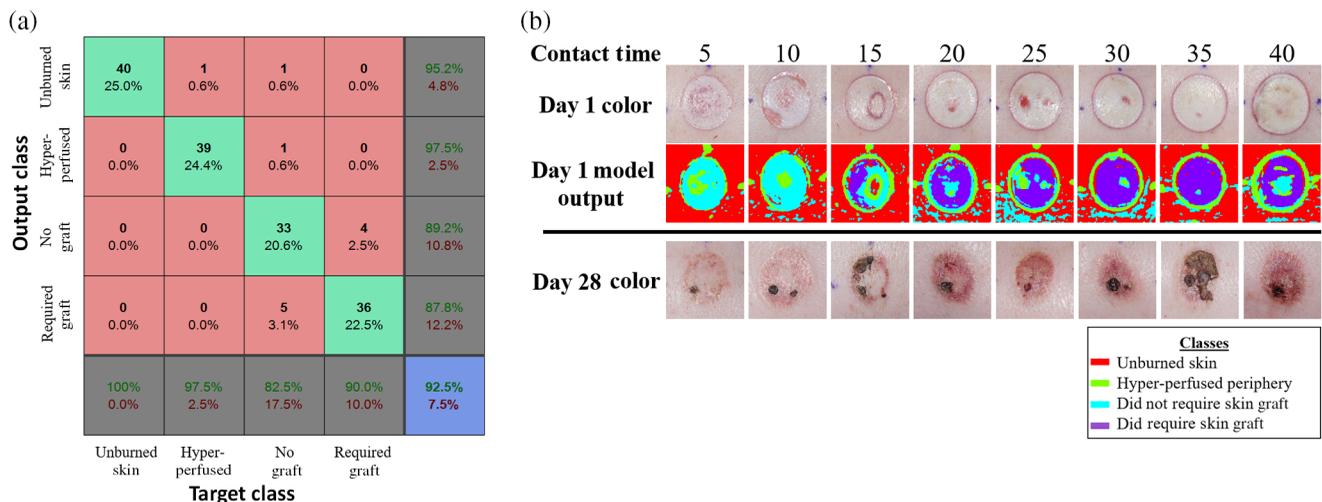

Fig. 5 (a) Confusion matrix of 10-fold cross-validation performed on the cubic SVM model created using all spatial frequency data. (b) Combined data cubic SVM model applied to day 1 calibrated reflection data of burns in a test pig (middle row). Corresponding day 28 color image (bottom row) shows the true outcome for each burn.

and a sensitivity of 100%. The model indicated burn regions that did not require grafting with a precision of 94.4% and sensitivity of 85%.

3.5 Summary

The *k*-fold cross-validation outcomes for each model are shown in Table 1.

4 Discussion

4.1 Planar (0 mm^{-1}) Model and Spectroscopy

Previous instances of machine learning as a means with which to predict burn wound severity have used multispectral imaging data to train an algorithm. The SFDI dataset of calibrated reflectance values taken only at the 0.00-mm^{-1} spatial frequency is

comparable to conventional multispectral imaging in previous studies. These modalities collect images at multiple wavelengths of light but do not utilize spatially modulated illumination. A previous burn imaging study used multispectral imaging in a Hanford pig model.[9] Reflectance images taken at eight wavelengths with the range of 420 to 972 nm were compiled with the linear SVM and k-nearest neighbors methods, and included an additional outlier detection algorithm. These model outcomes were between 63% and 76% accurate for all classifications. Another series of studies was performed on Hanford pigs using a combination of reflectance measurements from multispectral imaging at eight wavelengths between 420 and 850 nm, parameters from the digital color images such as gradient and skew, and photoplethysmography (PPG).[22,23] These studies described four classes for unburned skin, shallow burn, deep burn, and exposed wound bed due to debridement. Quadratic discriminant analysis models were created from measurements with each
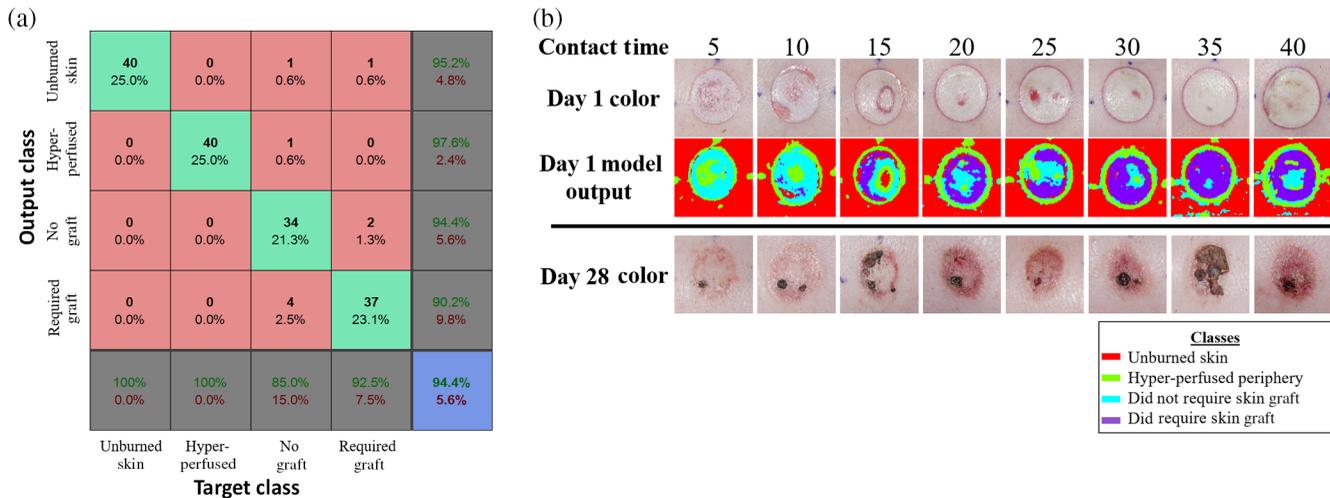
(a)

(b)

Fig. 6 (a) Confusion matrix of 10-fold cross-validation performed on the cubic SVM model created using all spatial frequency data, relative to values of the unburned skin regions. (b) Combined data cubic SVM model applied to day 1 calibrated reflection data of burns in a test pig (middle row). Corresponding day 28 color image (bottom row) shows the true outcome for each burn.

individual modality and the combined data from all the modalities. These models predicted classifications at accuracies between 73.4% and 74.4% for all classes for the models created through multispectral imaging alone, and accuracies between 76.9% and 77.8% when multispectral data were combined with features from PPG and color image data.

Hyperspectral imaging data have also been used to train machine learning models to segment burns. A study on Noroc pigs used hyperspectral data from 400 to 2500 nm and models compiled by the k-means algorithm and spectral–spatial image segmentation.[8] The outcomes from these methods indicated considerable heterogeneity, even in regions of unburned skin. Additionally, classifications through this method cannot be related back to physiologic differences since the segments were determined by unsupervised learning and not trained according to known burn severity.

The model using a subset of the SFDI data that contained only the planar (0-mm^{-1} spatial frequency) calibrated reflectance data resulted in an accuracy of 88.8% across all classes. While this subset of data is similar to that of previous groups using multispectral imaging, differences in model types, classification labels, and methods of assessing the classification accuracy make it difficult to directly compare model accuracy between studies.

4.2 Machine Learning with Spatially Modulated Light

In the models that used the complete set of SFDI data, accuracy across all classes was 92.5%, and 94.4% for the model relative to unburned skin. This improvement in accuracy of these two models over the model using the subset of planar (0 mm^{-1}) SFDI data demonstrates how the addition of spatially modulated light improves the machine learning model. The source of these differences in model output and overall accuracy can be seen in the graphs of calibrated reflectance plotted against wavelength in Fig. 2. The graphs for data taken at the planar (0 mm^{-1}) spatial frequency show some variation at lower wavelengths between regions of unburned skin, superficial burns that would not have required surgical intervention, and deep or

full thickness burns that should have received a skin graft. Alternately, there is variation in calibrated reflectance between these same regions at multiple wavelengths at the 0.20-mm^{-1} spatial frequency. The addition of this spatial frequency data to the model can account for an increase in the model accuracy, as there is more contrast between regions of different classifications.

Ideally, the outcomes predicted by the application of these models to our testing set would be further validated against another quantitative method that provides similar outputs. Such validation methods can be seen in previous SFDI works, where machine learning regression models are used to interpret optical properties.[19,20] Here, the machine learning-derived optical properties are compared against gold-standard methods of derivation, such as diffuse approximation and Monte Carlo-based look-up-tables. However, a quantitative classification model of burn wounds through optical imaging and independent of machine learning has not been established. To this end, model validation is assessed through k-fold accuracy and qualitative visual comparison of the model predictions.

This model could conceivably be simplified into a binary predictor, separating regions that require grafting with regions that do not require any kind of intervention. While the identification of the hyperperfused and unburned skin may not seem relevant to the clinical outcome, marking these two classes provides landmarks by which to compare color images to the final model prediction. Also, as seen by the graphs in Fig. 1, inclusion of these regions in the training set is necessary to take into account the low calibrated reflectance in the hyperperfused region. Additionally, even though this area will re-epithelialize, this zone is involved in the inflammatory response during burn wound healing.[38]

Compiling all calibrated reflectance data into a single output also simplifies the final outputs previously gained in burn studies with SFDI. In previous works, SFDI was performed on burns in a porcine model.[13–17] These earlier experiments reported values for scattering and absorption at every wavelength, as well as oxyhemoglobin and deoxyhemoglobin concentrations. While it is more difficult to equate classification values with physiological effects, the classifications are still based in real tissue optical

properties. For example, deoxyhemoglobin and oxyhemoglobin concentrations speak to vessel patency, and reduced scattering coefficients are related to the extent of collagen coagulation. While these different parameters have meaning, the machine learning outcome has the advantage of being a single intuitive result, rather than a series of multiple datasets that require further interpretation.

5 Conclusion

Machine learning has previously been used with multispectral imaging data to classify burn wound severity. Here, we show that the planar (0 mm^{-1}) spatial frequency subset of SFDI data can similarly be used to create an accurate model. However, the addition of modulated illumination at more spatial frequencies improves the accuracy of machine learning models to classify burns. The final outcome of this model creates a single image that can be easily interpreted by clinicians to assess burn wound severity. Ultimately, a faster diagnosis will allow for more appropriate and expedient treatment decision improving outcomes and reducing costs associated with burn care.

This work has shown that the addition of calibrated reflectance data collected with spatial modulation adds predictive power to a classification model for burns. Further investigation can now be made into how individual wavelengths and spatial frequencies improve the model outcome, and how these aspects may correlate with physiological differences between regions of the burn wound.

Disclosures

Dr. Durkin has a financial interest in Modulated Imaging, Inc., which developed the Reflect RS™. However, Dr. Durkin does not participate in the management of Modulated Imaging and has not shared these results with that company. Conflicts of interest have been disclosed and managed in accordance with the University of California and NIH policies. The other authors have no financial interests or commercial associations that might pose or create a conflict of interest with the information presented in this article.

Acknowledgments

References

1. D. Heimbach et al., "Burn depth: a review," *World J. Surg.* **16**(1), 10–15 (1992).
2. A. D. Jaskille et al., "Critical review of burn depth assessment techniques: part I. Historical review," *J. Burn Care Res.* **30**(6), 937–947 (2009).
3. A. H. E. Talwalkar and K. Palso, "National Ambulatory Medical Care Survey: 2011 summary tables," Centers for Disease Control and Prevention (2011).
4. C. E. White and E. M. Renz, "Advances in surgical care: management of severe burn injury," *Crit. Care Med.* **36**(7 Suppl.), S318–S324 (2008).
5. Z. Janzekovic, "A new concept in the early excision and immediate grafting of burns," *J. Trauma* **10**(12), 1103–1108 (1970).
6. S. Monstrey et al., "Assessment of burn depth and burn wound healing potential," *Burns* **34**(6), 761–769 (2008).
7. K. M. Bombaro et al., "What is the prevalence of hypertrophic scarring following burns?" *Burns* **29**(4), 299–302 (2003).
8. L. A. Paluchowski et al., "Can spectral-spatial image segmentation be used to discriminate experimental burn wounds?" *J. Biomed. Opt.* **21**(10), 101413 (2016).
9. W. Li et al., "Outlier detection and removal improves accuracy of machine learning approach to multispectral burn diagnostic imaging," *J. Biomed. Opt.* **20**(12), 121305 (2015).
10. T. Shi and C. A. DiMarzio, "Multispectral method for skin imaging: development and validation," *Appl. Opt.* **46**(36), 8619–8626 (2007).
11. M. S. Chin et al., "Hyperspectral imaging for burn depth assessment in an animal model," *Plast. Reconstr. Surg. Glob. Open.* **3**(12), e591 (2015).
12. M. G. Sowa et al., "Classification of burn injuries using near-infrared spectroscopy," *J. Biomed. Opt.* **11**(5), 054002 (2006).
13. A. Ponticorvo et al., "Quantitative assessment of graded burn wounds in a porcine model using spatial frequency domain imaging (SFDI) and laser speckle imaging (LSI)," *Biomed. Opt. Express* **5**(10), 3467–3481 (2014).
14. D. M. Burmeister et al., "Utility of spatial frequency domain imaging (SFDI) and laser speckle imaging (LSI) to non-invasively diagnose burn depth in a porcine model," *Burns* **41**(6), 1242–1252 (2015).
15. A. Ponticorvo et al., "Quantitative long-term measurements of burns in a rat model using spatial frequency domain imaging (SFDI) and laser speckle imaging (LSI)," *Lasers Surg. Med.* **49**(3), 293–304 (2017).
16. J. Q. Nguyen et al., "Spatial frequency domain imaging of burn wounds in a preclinical model of graded burn severity," *J. Biomed. Opt.* **18**(6), 066010 (2013).
17. A. Ponticorvo et al., "Evaluating clinical observation versus spatial frequency domain imaging (SFDI), laser speckle imaging (LSI) and thermal imaging for the assessment of burn depth," *Burns* **45**(2), 450–460 (2018).
18. R. Rowland et al., "A simple burn wound severity assessment classifier based on spatial frequency domain imaging (SFDI) and machine learning," *Proc SPIE* **10851**, 1085109 (2019).
19. Y. Zhao et al., "Deep learning model for ultrafast multifrequency optical property extractions for spatial frequency domain imaging," *Opt. Lett.* **43**(22), 5669–5672 (2018).
20. S. Panigrahi and S. Gioux, "Machine learning approach for rapid and accurate estimation of optical properties using spatial frequency domain imaging," *J. Biomed. Opt.* **24**(7), 1–6 (2018).
21. C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.* **20**(3), 273–297 (1995).
22. J. Heredia-Juesas et al., "Non-invasive optical imaging techniques for burn-injured tissue detection for debridement surgery," *IEEE Eng. Med. Biol.* **2016**, 2893–2896 (2016).
23. J. Heredia-Juesas et al., "Burn-injured tissue detection for debridement surgery through the combination of non-invasive optical imaging techniques," *Biomed. Opt. Express* **9**(4), 1809–1826 (2018).
24. D. J. Cuccia et al., "Modulated imaging: quantitative analysis and tomography of turbid media in the spatial-frequency domain," *Opt. Lett.* **30**(11), 1354–1356 (2005).
25. D. J. Cuccia, "Spatial frequency domain imaging (SFDI): a technology overview and validation of an LED-based clinic-friendly device," *Proc SPIE* **8254**, 825405 (2012).
26. D. J. Cuccia et al., "Quantitation and mapping of tissue optical properties using modulated imaging," *J. Biomed. Opt.* **14**(2), 024012 (2009).
27. R. H. Wilson et al., "High-speed spatial frequency domain imaging of rat cortex detects dynamic optical and physiological properties following cardiac arrest and resuscitation," *Neurophotonics* **4**(4), 045008 (2017).
28. M. Torabzadeh et al., "Compressed single pixel imaging in the spatial frequency domain," *J. Biomed. Opt.* **22**(3), 030501 (2017).
29. J. R. Weber et al., "Multispectral imaging of tissue absorption and scattering using spatial frequency domain imaging and a computed-tomography imaging spectrometer," *J. Biomed. Opt.* **16**(1), 011015 (2011).

30. D. J. Cuccia et al., "Quantitative in vivo imaging of tissue absorption, scattering, and hemoglobin concentration in rat cortex using spatially modulated structured light," in *In Vivo Optical Imaging of Brain Function*, 2nd ed., R. D. Frostig, Ed., CRC Press/Taylor & Francis, Boca Raton, Florida (2009).

31. W. S. Noble, "What is a support vector machine?" *Nat. Biotechnol.* **24**(12), 1565–1567 (2006).

32. Z. R. Yang, "Biological applications of support vector machines," *Brief Bioinform.* **5**(4), 328–338 (2004).

33. I. Inza et al., "Machine learning: an indispensable tool in bioinformatics," *Methods Mol. Biol.* **593**, 25–48 (2010).

34. G. Varoquaux, "Cross-validation failure: small sample sizes lead to large error bars," *Neuroimage* **180**, 68–77 (2017).

35. R. J. Meijer and J. J. Goeman, "Efficient approximate k-fold and leave-one-out cross-validation for ridge regression," *Biom. J.* **55**(2), 141–155 (2013).

36. M. W. Browne, "Cross-validation methods," *J. Math. Psychol.* **44**(1), 108–132 (2000).

37. R. L. Figueroa et al., "Predicting sample size required for classification performance," *BMC Med. Inf. Decis. Making* **12**, 8 (2012).

38. S. Hettiaratchy and P. Dziewulski, "ABC of burns: pathophysiology and types of burns," *BMJ* **328**(7453), 1427–1429 (2004).

Biographies of the authors are not available.