PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Classification of skin-cancer lesions based on Fluorescence Lifetime Imaging

Vasanthakumari, Priyanka, Romano, Renan, Rosa, Ramon, Salvio, Ana, Yakovlev, Vladislav, et al.

Priyanka Vasanthakumari, Renan A. Romano, Ramon G. T. Rosa, Ana G. Salvio, Vladislav Yakovlev, Cristina Kurachi, Javier A. Jo, "Classification of skin-cancer lesions based on Fluorescence Lifetime Imaging," Proc. SPIE 11317, Medical Imaging 2020: Biomedical Applications in Molecular, Structural, and Functional Imaging, 113170Z (28 February 2020); doi: 10.1117/12.2548625

SPIE.

Event: SPIE Medical Imaging, 2020, Houston, Texas, United States

Classification of skin-cancer lesions based on Fluorescence Lifetime Imaging

Priyanka Vasanthakumari[a], Renan A. Romano[b], Ramon G. T. Rosa[b], Ana G. Salvio[d], Vladislav Yakovlev[a], Cristina Kurachi[b], Javier A. Jo[c]
[a]Department of Biomedical Engineering, Texas A&M University, College Station, TX
[b] São Carlos Institute of Physics, University of São Paulo, São Paulo, Brazil
[c]School of Electrical and Computer Engineering, University of Oklahoma, Norman, OK
[d]Skin Department of Amaral Carvalho Hospital, Jau, São Paulo, Brazil

ABSTRACT

Every year more than 5.4 million new cases of skin cancer are reported in the US. Melanoma is the most lethal type with only 5% occurrence rate, but accounts for over 75% of all skin cancer deaths. Non-melanoma skin cancer, especially basal cell carcinoma (BCC) is the most commonly occurring and often curable type that affects more than 3 million people and causes about 2000 deaths in the US annually. The current diagnosis involves visual inspection, followed by biopsy of the lesions. The major drawbacks of this practice include difficulty in border detection causing incomplete treatment and, the inability to distinguish between clinically similar lesions. Melanoma is often mistaken for the benign lesion pigmented seborrheic keratosis (pSK), making it extremely important to differentiate benign and malignant lesions. In this work, a novel feature extraction algorithm based on phasors was performed on the Fluorescence Lifetime Imaging (FLIM) images of the skin to reliably distinguish between benign and malignant lesions. This approach, unlike the standard FLIM data processing method that requires time-deconvolution of the instrument response from the measured time-resolved fluorescence signal, is computationally much simpler and provides a unique set of features for classification. Subsequently, FLIM derived features were selected using a double step cross validation approach that assesses the reliability and the performance of the resultant trained classifier. Promising FLIM-based classification performance was attained for detecting benign from malignant pigmented (sensitivity: ~80%, specificity: 79%, overall accuracy: ~79%) and non-pigmented (sensitivity: ~88%, specificity: 83%, overall accuracy: ~87%) lesions.

Keywords: FLIM, Skin cancer, Classification, Phasors, Cross-validation, Machine learning

1. INTRODUCTION

In 2019, approximately 606,880 people are estimated to die of cancer in the United States [1]. Of all the types of cancer, skin cancer is the most common with melanoma being the fifth most prevalent among men and women. The 5-year survival rate of patients with early stage skin melanoma is ~98%; however, ~22% of skin melanoma patients are diagnosed with lesions already at intermediate or advance stages [2], associated with a 5-year survival rate of ~64% and ~23% respectively[1]. Basal cell carcinoma (BCC) is the most commonly occurring type of skin cancer, affecting 3 million people in the US each year. Although BCC is considered curable, delayed diagnosis and incomplete surgical resection are two major factors associated to the ~2,000 deaths caused by BCC reported each year in the US. The most common diagnosis for skin cancer lesions is visual examination followed by biopsy tissue resection and histopathological evaluations. One major drawback of this practice is the inability to visually distinguish between clinically similar lesions; in particular, melanoma is often mistaken for benign lesion pigmented seborrheic keratosis (pSK), and BCC for intradermic nevus. In addition, it is known that the accuracy of melanoma diagnosis with unaided eye is only about 60% [3]. Therefore, clinical tools that could provide objective, in situ and accurate noninvasive discrimination between cancerous and benign lesions could significantly facilitate early detection of skin cancer. One of the most common tools used by physicians to diagnose skin cancer lesions is a dermoscope [3,4] which helps the unaided eye by magnifying the features on the skin. This allows the doctors to conduct a much-detailed examination on the morphological features of the skin. Although dermoscopy has known to improve the diagnostic sensitivity of skin lesions by ~10-30%, its performance largely depends on the level of experience of the dermatologists and the type of lesions. Therefore, the highly subjective nature and poor reproducibility of this method led to the emergence of several computer aided diagnostic (CAD) systems [4,5]. CAD systems are becoming largely popular in

Medical Imaging 2020: Biomedical Applications in Molecular, Structural, and Functional Imaging, edited by Andrzej Krol, Barjor S. Gimi, Proc. of SPIE Vol. 11317, 113170Z · © 2020 SPIE · CCC code: 1605-7422/20/$21 · doi: 10.1117/12.2548625

both diagnosis and prognosis of various diseases as they allow automated and non-invasive analysis of the tissue conditions. Numerous predictive models are developed based on training datasets obtained from imaging modalities such as ultrasound imaging, MRI [6], coherence tomography (CT) [7], confocal microscopy [8] and even total body photographs [9]. Although CAD systems for skin cancer diagnosis based on several imaging modalities have been investigated in previous works, dermoscopy [10] remains the most popular one. However, the amount of information that can be extracted from dermoscopic images are limited as they essentially contain only the enhanced morphological information of the skin lesions. CAD systems require extraction of reliable feature sets and robust classification models to provide good diagnostic performance.

The presence of endogenous fluorophores in the tissue enables the monitoring of autofluorescence from them upon excitation with specific wavelengths. This eliminates the need for staining the target area with dyes or exogenous fluorophores. Pathological conditions and neoplastic progressions in the tissue are associated with epithelial changes such as an increase in nuclear size and texture, resulting in changes in the absorption and scattering characteristics of incident light. Increase in metabolism is also observed in cancerous tissue, whereby the relative concentrations of intrinsic fluorophores [11] such as NADH, FAD, collagen and elastin are different from that of the surrounding healthy tissue. These changes can also be utilized in differentiating between benign and malignant lesions. Several works have successfully demonstrated the capability of autofluorescence imaging in distinguishing between lesion and the surrounding healthy tissue [12]. However, their potential in differentiating between benign and malignant lesions are not yet fully explored. This is because, measuring the fluorescence intensity alone is not enough to distinguish between the types of lesions and results in very low specificities. For example, all benign and malignant lesions exhibit a decrease in the collagen content and thus a corresponding decrease in the fluorescence intensity. Monitoring the changes in the temporal characteristics of the fluorophores, particularly NADH and FAD could provide reliable biomarkers to identify malignant transformation in the lesions. FLIM enables the measurement of fluorescence lifetimes of multiple fluorophores upon excitation. Moreover, the broad emission spectral bandwidth of the fluorescence intensity signal has an intrinsic disadvantage that it is difficult to multiplex between the intensities of overlapping emissions from multiple fluorophores. FLIM overcomes this challenge by enabling the measurement of fluorescence decay lifetimes in addition to their intensities.

In this work, we develop an automated, non-invasive clinical diagnostic tool to classify skin lesions from FLIM datasets. This work introduces frequency domain deconvolution of FLIM data and phasor-based feature extraction for classification of benign and malignant skin lesions. In comparison to traditional iterative FLIM processing techniques, this method considerably reduces the computational complexity and provides reliable features for classification. Further, an unbiased classification model is designed by validating the classifier with dual level cross-validation strategies. Finally, the performance of the classifiers and the reliability of the extracted features are evaluated.

2. METHODS

2.1 FLIM dermoscopy imaging of skin lesions

A total of 88 patients were recruited for this study from the Dermatology Department of the Amaral Carvalho Cancer Hospital (Jahu, Sao Paulo, Brazil), following a human study protocol approved by the Internal Review Board of that institution (CAAE: 71208817.5.00005434). Only patients presenting at least one skin lesion undergoing biopsy examination for skin cancer diagnosis were recruited for this study. Multispectral FLIM images were obtained from clinically suspicious lesions using an in-house developed multispectral time-domain FLIM dermoscope previously described [13]. With this FLIM dermoscope, skin tissue autofluorescence is simultaneously imaged at three emission bands (390±20 nm, 452±22 nm, and >496 nm, preferentially targeting collagen, NADH and FAD autofluorescence emission, respectively) with a temporal resolution of 0.4 ns, field-of-view (FOV) of 8.65×8.65 mm^2, and lateral resolution of 120 μm. After signing the corresponding written informed consent form, each patient underwent the following imaging protocol right before the scheduled biopsy examination procedure. First, the lesion was gently cleaned with a gaze soaked in a saline solution. Then, the tip of the FLIM dermoscope, previously disinfected using a gaze soaked in ethanol 70%, was placed in contact with the lesion, and a multispectral FLIM image was acquired. The imaging site was selected so that regions of both lesion and healthy skin tissue were present within the FOV of the FLIM dermoscope. Right after FLIM imaging, lesion tissue biopsy was performed following standard procedures. Each FLIM image was labeled based on the histopathological evaluation of the lesion biopsy, which was considered the gold standard in this study. All images were acquired with an average laser excitation power of 10 mW measured at the sample, 140×140 pixels per image, and at a

pixel rate of 10kHz. These image acquisition parameters corresponded to an acquisition time of 1.96 s per image, an excitation energy exposure of 1.52 mJ and a radiant exposure of 2.03 mJ cm^{-2} at the sample, which is significantly lower than the maximum permissible exposure (MPE) levels for skin based on guidelines from the American National Standards Institute – ANSI (thermal MPE: 663 mJ cm^{-2}) [14].

2.2 FLIM data pre-processing

The multispectral FLIM data is composed of fluorescence intensity temporal decay signals, $y_\lambda(x,y,t)$, measured at each emission spectral band (λ) and each spatial location or image pixel (x,y). Figure 1A illustrates the steps involved in pre-processing of the FLIM data. First, offset and background subtraction was applied to the temporal signal at each pixel of the FLIM image. Second, pixels presenting temporal signal saturation were detected by setting a threshold on the maximum signal amplitude, and masked. Third, pixel masking based on SNR was also performed with an SNR threshold value of 15 dB. Fourth, spatial averaging (order 5x5) was applied to increase the temporal signal-to-noise ratio (SNR) at each spatial location. All the channels are then adjusted to be of equal time intervals by zero-padding them if necessary. To facilitate further analysis, the FLIM signal intensities at each pixel location are normalized to sum one. Since each FLIM image contains a lesion region that is surrounded by healthy region, it is necessary to separate them before any further processing is carried out. This is done by first creating a cluster mask with an unsupervised K-means clustering algorithm using the entire time vector at each pixel location as the feature. Subsequently, the mask is applied on the images to separate the lesion from the healthy pixels. It may be noted that since K-means clustering algorithm involves random initialization of cluster centroids, it is difficult to identify which cluster belongs to healthy or lesion category. Figure 1B shows an example FLIM image along with the generated cluster mask and the two separated regions. The transformations in a single pixel of multispectral FLIM data during pre-processing is shown in figure 1C.

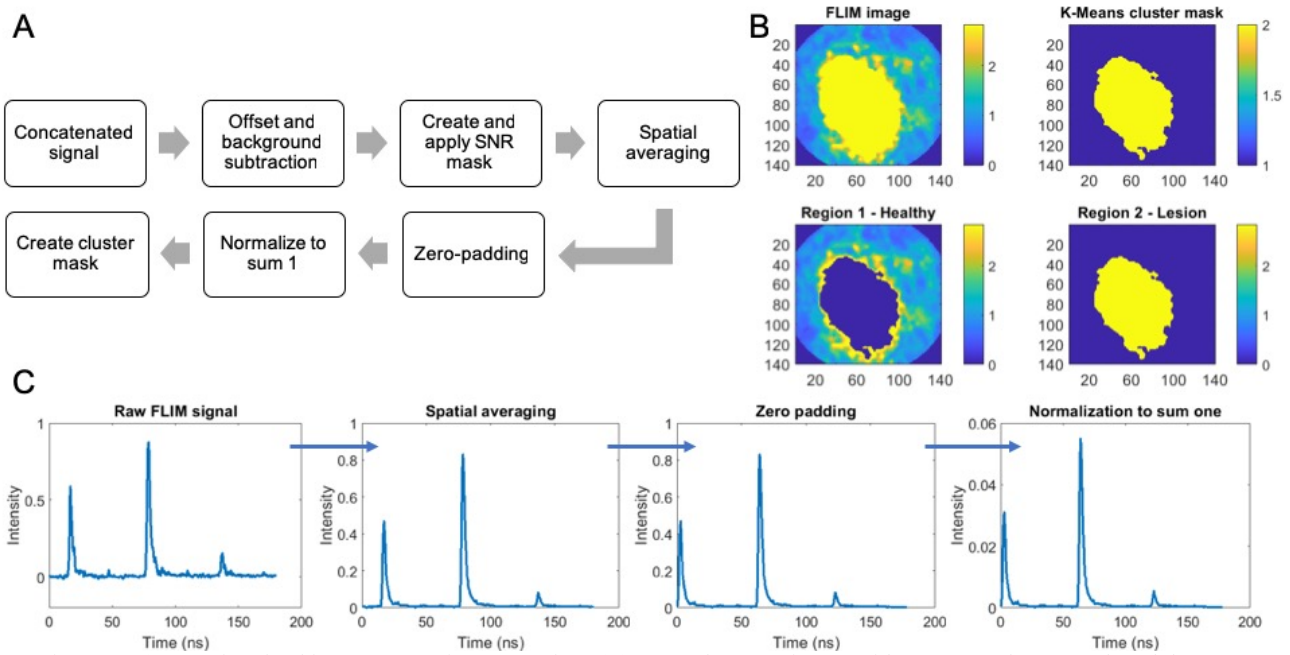

Figure 1: A) Steps involved in pre-processing FLIM data. (B) Example FLIM image with K-means cluster mask and the separated healthy and lesion pixel regions. (C) Transformations in a single pixel multispectral FLIM data during pre-processing

2.3 Feature Extraction

The traditional method for FLIM data processing proceeds by first deconvolving the measured time-resolved fluorescence signal at each pixel location in a spectral channel with the instrument response to obtain the tissue fluorescence impulse response [15]. The temporal characteristics like fluorescence decay lifetimes and coefficients are estimated by performing nonlinear least squares iterative reconvolution on the multiexponential fit model. This iterative process is computationally complex, time consuming and often results in approximations of the signal. This brings about the need to develop a much

simpler algorithm for extracting FLIM features and achieving reliable classification. We aim to explore an alternate fitting-free strategy, where a time domain deconvolution could be translated to a simple division operation in the frequency domain. The computational overload is further reduced by processing all the three spectral channels together unlike in the traditional method where the channels must be processed separately. Subsequently, several features are extracted from the phasor analysis [16] of the frequency domain tissue impulse response. In comparison to the traditional time domain feature analysis, phasor analysis provides a means to graphically represent the fluorescent decay at each pixel location.

This method proceeds in three steps: 1) frequency domain deconvolution of the measured signal with the instrument response, 2) constructing phasor plots for the data, and 3) feature extraction from the phasor plots. In comparison to the traditional approach, the advantages of this technique are as follows: 1) eliminates iterative calculations thereby reduces computational complexity; 2) the time-resolved signals of all the spectral channels are processed simultaneously, thus making the method computationally efficient and more informative; 3) provides distinctive features for FLIM-based classification.

In the context of time-domain FLIM data analysis, the fluorescence decay $y_\lambda(x,y,t)$ measured at each spatial location (x,y) can be modeled[15] as the convolution of the fluorescence impulse response (FIR) $h_\lambda(x,y,t)$ of the sample and the measured instrument response function (IRF) $u_\lambda(t)$:

$$y_\lambda(x,y,t) = u_\lambda(t) * h_\lambda(x,y,t) \quad (1)$$

The pre-processed signal at each pixel location contains the concatenated frequency decays at the three spectral channels and can be mathematically represented as:

$$y\,(x,y,t) = \sum_{n=0}^{2} y_{\,\lambda_{n+1}}(x,y,(t-Mn)) \quad (2)$$

Where, y (x,y,t) is the concatenated FLIM signal, $y_{\,\lambda_1}$, $y_{\,\lambda_2}$ and $y_{\,\lambda_3}$ are the FLIM signals from the three spectral channels, M is the temporal spacing between the signals from the three channels and (x, y) are the pixel locations. Similarly, the concatenated impulse response function can be mathematically represented as:

$$u\,(t) = \sum_{n=0}^{2} u_{\,\lambda_{n+1}}(t-Mn) \quad (3)$$

The first step of the algorithm is to compute the Fourier transform of the pre-processed concatenated signal and the instrument response function. Equation 1 in the frequency domain can be represented as:

$$\frac{Y(x,y,\omega)}{Y(x,y,0)} = \frac{H(x,y,\omega)\,U(\omega)}{H(x,y,0)\,U(0)} \quad (4)$$

Where H(x,y,ω), U(ω) and Y(x,y,ω) are the Fourier transforms of the tissue impulse response, concatenated instrument response u(t) and the measured FLIM signal y(x,y,t) respectively.

The tissue impulse response in the frequency domain can be estimated as:

$$P(x,y,\omega) = \frac{H(x,y,\omega)}{H(x,y,0)} = \frac{Y(x,y,\omega)\,U(0)}{U(\omega)\,Y(x,y,0)} \quad (5)$$

This is followed by the construction of a phasor, which is achieved by plotting the real part of P(x,y,ω) against its imaginary part. The phasor locations are computed at specific frequency indices. A total of 36 features are extracted from the phasors, containing 4 feature sets, each computed at 9 frequency indices. The features are computed in such a way that their values of the lesion pixels (region 2 in Figure 2A) are measured with respect to that of the healthy pixels (region 1 in Figure 2A). The feature sets and their description are as follows:

Distance between phasor distributions: Figure 2A shows the phasor histograms of lesion and healthy regions of one FLIM image. The distance between the phasor distributions will be the distance between their histogram maximum values along the horizontal plane as denoted by 'd'.

Difference in spread of the phasor distributions: The spread of a phasor distribution is measured by first fitting a multivariate gaussian function to the phasor point distribution, shown as an ellipse in Figure 2B. The determinant of the corresponding covariance matrix gives a measure of the spread of the distribution. The final feature value is given by $|\text{Spread}_{\text{Lesion}} - \text{Spread}_{\text{Healthy}}|$.

Difference in symmetry of the phasor distributions: The ratio of the lengths of major and minor axes of the ellipse gives a measure of its symmetry as given by 'y/x' in Figure 2B. This quantity is obtained by calculating the ratio of the diagonal elements of the covariance matrix, which are measures of variances along the two axes. The final feature value is given by $|\text{Symmetry}_{\text{Lesion}} - \text{Symmetry}_{\text{Healthy}}|$.

Angle between the phasor distributions: This feature is defined as the angle between the major axes of the fitted ellipses of lesion and healthy phasor distributions, given as 'θ' in Figure 2C. The value can be obtained by computing the acute angle between the eigen vectors of maximum variance in their multivariate gaussian distributions.

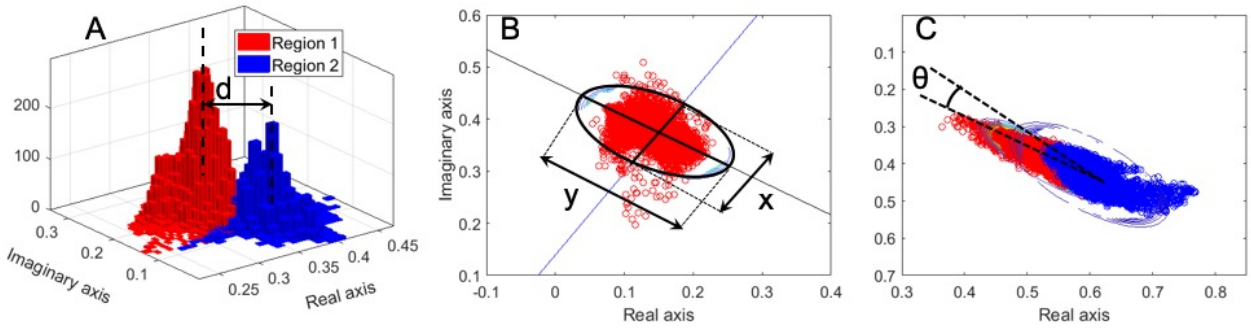

Figure 2: A) 2D histogram distributions of phasor from healthy and lesion region pixels (B) Multivariate gaussian fit on the phasor distribution of lesion pixels (C) Phasor distribution scatter plots of lesion and healthy pixels showing the angle between their major axes

Global features are exploited in this study, whereby one set of features (a single feature vector) is used to represent the whole image. Thus, the pixel wise features are converted into global features, which further reduces the computational complexity. In addition, since each feature is measured in reference to the healthy pixel values of the same patient, possible patient to patient variability in the extracted features is minimized. This is particularly important because, the color and texture of skin varies considerably with ethnicity and age.

2.4 Classifier Design

The classifier was designed in two approaches as shown by the flow diagrams in figure 3. The number of features in the initial feature set (n_{Total}) is 6, 9 and 15 for intensity, lifetime and combined feature sets respectively. Figure 3A illustrates the classifier design with one level of cross-validation (CV). A sequential forward search (SFS) algorithm [17] is performed on the initial feature set to identify the relevant features for optimum classifier performance. Unlike exhaustive search whereby every possible combination of features is examined to choose the best set, SFS is computationally simpler and requires less time to explore the feature domain. Although the best performance obtained with SFS may not be optimum, it provides a good enough means to investigate the importance of features and to obtain good classification. The feature set corresponding to the highest area under the curve (AUC) value in ROC analysis was selected. AUC in each iteration is calculated in combination with leave-one-out cross-validation (LOOCV) to minimize training bias and to provide reliability. The number of features in the selected feature set ($n_{Selected}$) was varied between 3 to 6. Subsequently, in addition to attaining high accuracy, it is also important to obtain a fair balance between the sensitivity and specificity of the classifier. This is particularly important for training sets with unbalanced classes as it minimizes the bias of the classifier towards the majority class. This is done by optimizing the threshold of the posterior probabilities by choosing the value corresponding to the point located closest to the ideal point (0,1) on the ROC curve ($Th_{selected}$). Finally, the classifier is trained on the selected features and the threshold value to obtain a final confusion matrix and the corresponding performance metrics such as accuracy, sensitivity and specificity.

Figure 3B illustrates the flow diagram of the classifier design with double level of cross-validation, in which the previous approach is modified to provide an additional outer level of LOOCV. In this strategy, one data point which is left out at the initial step remains completely unbiased of the feature selection and threshold optimization steps. The left-out data is then utilized in prediction on the trained classifier. This process is repeated until the entire data set is used for prediction, resulting in a final confusion matrix. Performance metrics such as accuracy, sensitivity and specificity are then evaluated from the confusion matrix. The algorithm provides an added advantage of estimating the completely unbiased performance

of the classifier with its dual level cross-validation strategy. It is very important to test the reliability of the predictors and the classification model with more than one level of cross-validation. The main idea behind this is that if the classification performance remains fairly stable with the additional level of cross-validation, the classification model and the selected predictors are robust and reliable for real word applications. The two approaches for classifier design are implemented to demonstrate the importance of more than one level of cross-validation, especially in medical applications where the size of the dataset is not that big. Two classification algorithms, linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA) are investigated to explore both linear and non-linear decision boundaries.

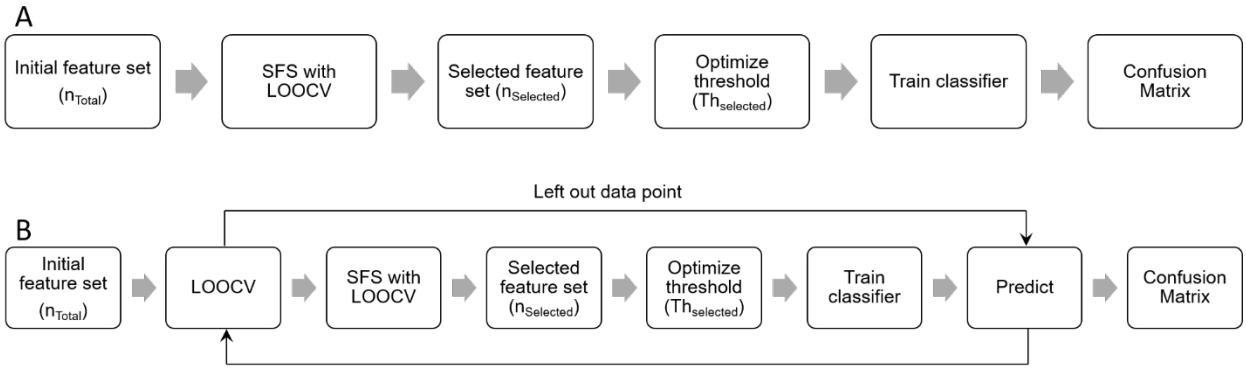

Figure 3: Flow diagram showing classifier design with (A) one level and (B) double level cross-validation

3. RESULTS

The skin lesions are broadly categorized into two groups - pigmented and non-pigmented - based on their color. As benign and malignant lesions are present in both the groups, classification is performed on them separately. The database consists of 96 pigmented (56 benign, 40 malignant) and 121 non-pigmented (23 benign, 98 malignant) FLIM images. It may be noted that multiple images of the lesion were obtained from the same patient. Within the pigmented category, benign lesions include solar lentigo (SL) and pigmented seborrheic keratosis (pSK), while malignant lesions include melanoma (Mel), pigmented superficial BCC (pSBCC) and pigmented nodular BCC (pNBCC). Similarly, within the non-pigmented category, benign lesions include intradermic nevus (IN), while malignant lesions include superficial BCC (SBCC) and nodular BCC (NBCC). The distribution of the lesions is provided in table 1.

Table 2 shows the best performance metrics including accuracy, sensitivity (Sn) and specificity (Sp) obtained when classifying pigmented and non-pigmented data sets with 3, 4, 5 and 6 phasor feature sets and with LDA and QDA classification algorithms. Table 3 shows the confusion matrices for the best classification performances. Figure 4 shows the frequency at which the features are selected during the double level cross-validation approach.

Table 1: Distribution of lesions in pigmented and non-pigmented lesion categories

	Pigmented		**Non-pigmented**	
	Type	**Count**	**Type**	**Count**
Benign	Pigmented seborrheic keratosis	46	Intradermic nevus	23
	Solar lentigo	10		
Malignant	Melanoma	24	Nodular BCC	81
	Pigmented nodular BCC	6	Superficial BCC	17
	Pigmented superficial BCC	10		

Table 2: Performance metrics of the classifier with single and double level cross-validation trained on both pigmented and non-pigmented data sets with the best performance for phasor features

Category	Single level CV				Double level CV			
	Classifier ($n_{Selected}$)	*Accuracy (%)*	*Sn (%)*	*Sp (%)*	*Classifier* ($n_{Selected}$)	*Accuracy (%)*	*Sn (%)*	*Sp (%)*
Pigmented	QDA (4)	81.25	82.50	80.36	QDA (4)	79.16	80.00	78.57
Non-Pigmented	QDA (6)	88.42	89.80	82.61	QDA (5)	86.77	87.76	82.61

Table 3: Confusion matrices with best performance using single and double level cross-validation for both pigmented (left) and non-pigmented (right) datasets using phasor features

Single level CV QDA (4)			Predicted	
			Benign	**Malignant**
True	**Benign**	pSK	38	8
		SL	7	3
	Malignant	Mel	5	19
		pNBCC	1	5
		pSBCC	1	9

Single level CV QDA (6)			Predicted	
			Benign	**Malignant**
True	**Benign**	IN	19	4
	Malignant	NBCC	8	73
		SBCC	2	15

Double level CV QDA (4)			Predicted	
			Benign	**Malignant**
True	**Benign**	pSK	37	9
		SL	7	3
	Malignant	Mel	5	19
		pNBCC	2	4
		pSBCC	1	9

Double level CV QDA (5)			Predicted	
			Benign	**Malignant**
True	**Benign**	IN	19	4
	Malignant	NBCC	10	71
		SBCC	2	15

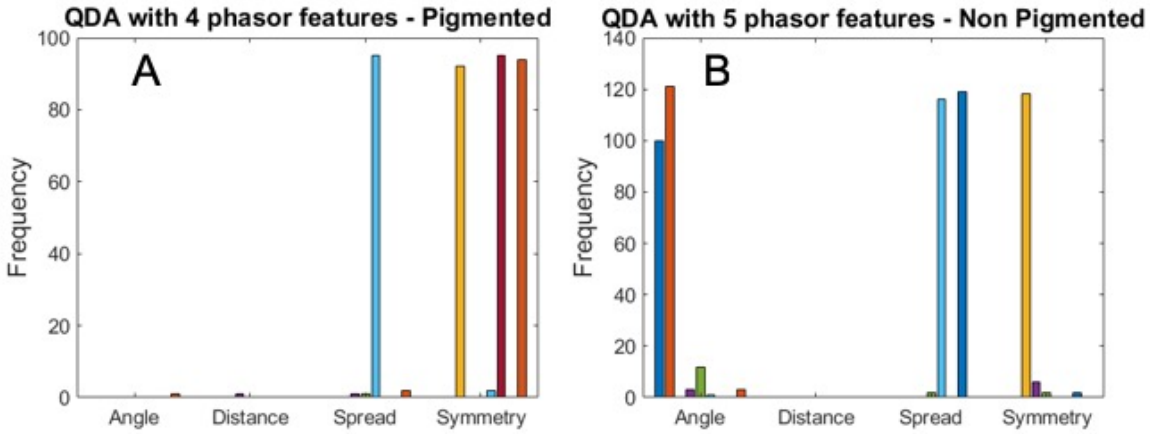

Figure 4: Frequency at which each feature is selected during double level cross-validation for (A) pigmented and (B) non-pigmented data sets using phasor features.

4. CONCLUSION AND DISCUSSION

Differential diagnoses of skin lesions play a crucial role in determining the survival rates of patients exhibiting symptoms that may be similar to multiple skin conditions. This is also very important in determining adequate treatment strategies at an early stage to considerably improve the chances of survival. In this work, a number of benign and malignant skin lesions in both pigmented and non-pigmented categories are considered as shown in table 1. The feature extraction based on frequency domain deconvolution and phasor representation is much simpler and computationally efficient in comparison to the traditional iterative time domain analysis. The reliability of the features and the robustness of the classifier is evaluated using a double level cross-validation approach. As can be seen in table 2, the performance metrics with the double level cross-validation remains stable and is only slightly less than that of single level cross-validation. An overall accuracy of ~79% and ~87% was obtained with double level cross-validation for the pigmented and non-pigmented skin lesions respectively. This shows that the phasor based FLIM features are reliable for generalizing the classifier to real time applications. It can also be seen from table 2 that the best performance for both pigmented and non-pigmented skin lesions is obtained with QDA classification models. This is because the true decision boundary separating benign and malignant skin lesions could be non-linear, resulting in relatively poor performance with LDA classifiers. As can be seen in figure 4, 'spread' and 'symmetry' features are the main contributors in discriminating pigmented lesions, while 'angle', 'spread' and 'symmetry' are the main contributors for discriminating non-pigmented lesions. As future work, we propose to optimize this algorithm to include more skin lesions and also to improve the performance metrics by exploring more features.

5. ACKNOWLEDGEMENT

This work support was provided by Brazilian Funding Agencies: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001; CNPq-PVE (401150/2014-3 and 314533/2014-1); CNPq-PQ (305795/2016-3) and São Paulo Research Foundation (FAPESP) grants: 2013/07276-1 (CEPOF); 2014/50857-8 (INCT). This project was also supported by the National Institutes of Health (NIH/NCI grant 1R01CA218739) and the Cancer Prevention and Research Institute of Texas (CPRIT grant RP180588).

6. REFERENCES

[1] "Cancer Facts and Figures - American Cancer Society.", (2019).

[2] Siegel, R. L., Miller, K. D. and Jemal, A., "Cancer statistics, 2019 (US statistics)," CA. Cancer J. Clin. **69**(1), 7–34 (2019).

[3] Kittler, H., Pehamberger, H., Wolff, K. and Binder, M., "Diagnostic accuracy of dermoscopy," Lancet Oncol. **3**(3), 159–165 (2002).

[4] Masood, A. and Al-Jumaily, A. A., "Computer aided diagnostic support system for skin cancer: A review of techniques and algorithms," Int. J. Biomed. Imaging **2013** (2013).

[5] Ali, A.-R. A. and Deserno, T. M., "A systematic review of automated melanoma detection in dermatoscopic images and its ground truth data," Med. Imaging 2012 Image Perception, Obs. Performance, Technol. Assess. **8318**(February 2012), 83181I (2012).

[6] Kamnitsas, K., Ledig, C., Newcombe, V. F. J., Simpson, J. P., Kane, A. D., Menon, D. K., Rueckert, D. and Glocker, B., "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," Med. Image Anal. **36**, 61–78 (2017).

[7] Sahu, P., Yu, D., Dasari, M., Hou, F. and Qin, H., "A Lightweight Multi-Section CNN for Lung Nodule Classification and Malignancy Estimation," IEEE J. Biomed. Heal. Informatics **23**(3), 960–968 (2019).

[8] Calzavara-Pinton, P., Longo, C., Venturini, M., Sala, R. and Pellacani, G., "Reflectance confocal microscopy for in vivo skin imaging," Photochem. Photobiol. **84**(6), 1421–1430 (2008).

[9] Dengel, L. T., Petroni, G. R., Judge, J., Chen, D., Acton, S. T., Schroen, A. T. and Slingluff, C. L., "Total body photography for skin cancer screening," Int. J. Dermatol. **54**(11), 1250–1254 (2015).

[10] Rosado, B., Menzies, S., Harbauer, A., Pehamberger, H., Wolff, K., Binder, M., Kittler, H. and Corona, R., "Accuracy of computer diagnosis of melanoma: A quantitative meta-analysis," Arch. Dermatol. **139**(3), 361–367 (2003).

[11] Skala, M. C., Riching, K. M., Gendron-Fitzpatrick, A., Eickhoff, J., Eliceiri, K. W., White, J. G. and Ramanujam, N., "In vivo multiphoton microscopy of NADH and FAD redox states, fluorescence lifetimes, and cellular morphology in precancerous epithelia," Proc. Natl. Acad. Sci. U. S. A. **104**(49), 19494–19499 (2007).
[12] Galletly, N. P., McGinty, J., Dunsby, C., Teixeira, F., Requejo-Isidro, J., Munro, I., Elson, D. S., Neil, M. A. A., Chu, A. C., French, P. M. W. and Stamp, G. W., "Fluorescence lifetime imaging distinguishes basal cell carcinoma from surrounding uninvolved skin," Br. J. Dermatol. **159**(1), 152–161 (2008).
[13] Romano, R. A., Rosa, R. G. T., Jo, J. A. and Kurachi, C., "Label-free multispectral lifetime fluorescence to distinguish skin lesions," Proc. SPIE **10890** (2019).
[14] Henderson, R. and Schulmeister, K., [Laser safety] (2004).
[15] Lakowicz, J. R., [Principles of fluorescence spectroscopy] (2006).
[16] Kim, D., Hwang, W., Won, Y., Moon, S., Lee, S. Y., Kang, M. G., Han, W. S. and Kim, D. Y., "Phasor plot analysis using low pass filter for high-speed FLIM," 24 (2019).
[17] Pudil, P., Novovičová, J. and Kittler, J., "Floating search methods in feature selection," Pattern Recognit. Lett. **15**(94), 1119–1125 (1994).