Cancer Diagnosis Using N_2 Laser Excited Autofluorescence Spectroscopy of Formalin-Fixed Human Tissue

S. K. Majumder*, R. Kumar, N. Ghosh and P. K. Gupta
Biomedical Applications Section, Centre for Advanced Technology, Indore, India 452 013

ABSTRACT

We report results of an in-vitro autofluorescence spectroscopic study on formalin-fixed human breast tissue samples. The study involved tissue samples from 20 patients with breast cancer who underwent radical mastectomy at Choithram Hospital and Research Centre, Indore. A diagnostic algorithm making use of principal component analysis (PCA) as the feature extractor and artificial neural network (ANN) as the classifier was used to classify the samples as normal or cancerous. The algorithm based on combined PCA-ANN provided sensitivity and specificity values of 100% towards cancer for the training set data based on leave-one-out cross validation and a sensitivity of 97% and a specificity of 100% towards cancer for the independent validation set data. These results suggest that autofluorescence spectroscopy can provide a valuable alternate, in-vitro diagnostic modality in clinical pathology setting for discriminating cancerous tissue sites from normal sites.

Keywords: Histopathology, Autofluorescence spectroscopy, Formalin-fixed tissue, Diagnostic algorithm, Artificial neural network.

1. INTRODUCTION

In current medical practice, excisional biopsy followed by histology is the gold standard for definitive diagnosis of cancer[1]. The approach makes use of high resolution image information of the cellular and the sub-cellular structures of properly preprocessed and appropriately stained specimens of tissue kept in the field of view of a microscope. Despite its widespread success, a major drawback of this traditional approach is that the diagnosis requires expert interpretation of the microscopically derived histopathological information and is prone to human errors. This is particularly serious in the context of diagnosing tumors that are atypical or lack morphological features useful for differential diagnosis. An alternate technique that has shown promise to overcome these limitations is the technique based on autofluorescence spectroscopy of human tissues[2,3]. The underlying principle of the approach is that the onset and the progression of a disease like cancer is often accompanied by biochemical and morphological changes that are reflected in the measured autofluorescence spectra of tissue. Considerable work carried out over the last decade has shown that autofluorescence from human tissue can be used for quantitative, in-situ, near-real time and non-invasive diagnosis of cancer[2,3]. It is pertinent to note here that this spectroscopic approach may also provide an alternate, in-vitro diagnostic modality in clinical pathology setting. This may provide to a pathologist an additional quantitative diagnostic feedback. However, this necessitates validating the applicability of the method on formalin fixed resected tissue samples that are used for histopathology. This is important because formalin-fixation is known to dehydrate the tissue[4] and change its hemoglobin content[5] by draining out blood from it both of which are expected to lead to significant changes autofluorescence spectra of the tissue. Since most of the earlier in-vitro studies[2,6-8] were motivated by the objective of establishing the potential of the approach for in-situ, non-invasive diagnosis, specific care was taken in these studies to use fresh, unfixed tissue samples to simulate the in-vivo condition as close as possible. The concern for tissue handling protocols used in these studies made a few groups investigate autofluorescence from formalin fixed tissue[5, 9], but no systematic and comprehensive report on autofluorescence spectroscopy of formalin fixed tissue is available. The studies carried out by Fillipidis et al[5] on peripheral vascular tissue showed significant differences in measured autofluorescence with visible wavelength excitation from formalin fixed samples as compared to the unfixed ones. The effect of formalin fixation on Raman spectra has received relatively more attention[4]. Whereas changes in Raman spectra due to formalin fixation were significant for human bronchial tissue for breast tissue the change was not significant[4].

We report in this paper the results of an autofluorescence spectroscopic study carried out on formalin fixed breast tissue samples. We also report, for the first time to our knowledge, a comprehensive evaluation of the usefulness of the autofluorescence spectroscopy for discriminating cancerous from normal tissues in formalin-fixed tissue sample.

2. MATERIALS AND METHODS

2.1 Instrumentation

Autofluorescence spectra from the breast tissue samples were recorded at 337nm excitation using a N_2 laser based portable fluorimeter developed earlier[10, 11]. The system comprised of a sealed-off pulsed N_2 laser, a spectrograph (Acton Research Corporation, USA), an optical fiber probe and a gateable intensified CCD detector (4 Quik 05A, Stanford computer optics, Inc, USA). The diagnostic probe, developed in-house, was a fiber bundle, which had two legs; one contained a single quartz fiber (400µm core diameter, 0.22 NA) and the other contained six quartz fibers (400µm core diameter, 0.22 NA). The two legs merged to form a common fiber bundle that consisted of a central fiber, surrounded by a circular array of six fibers. The central fiber delivered excitation light to the tissue surface and the six fibers surrounding the central fiber collected tissue fluorescence from the surface area directly illuminated by the excitation light. The proximal ends of the collection fibers were arranged in a vertical array and the light coming from the distal end was imaged at the entrance slit of the spectrograph (Acton Research Corporation, USA) coupled to the intensified CCD detector. An additional quartz fiber was also coupled to the spectrograph along with the six fluorescence collection fibers in order to monitor the energy of each nitrogen laser pulse by monitoring the luminescence from phosphor material coated at the tip of this fiber. This luminescence from the fiber tip served as the reference for calibration of the measured fluorescence. All the fibers used had core diameter of 400µm and NA of 0.22. The common end of the fiber bundle was enclosed in an SS tube (9mm outer diameter and 60mm long). The tip of the probe was shielded by a quartz optical flat 2mm thick to provide a fixed distance between tissues and the fibers for improved collection of fluorescence and also to protect contamination of the fiber tips with body fluids. The spectral data acquisition was computer controlled. The delay between N_2 laser pulse and the intensified CCD camera shutter was adjusted to pick up tissue fluorescence. The gate-width used was 100ns. The autofluorescence spectra were recorded with the tip of the fiber-optic probe placed in contact with the tissue surface.

2.2 Tissue samples

The breast tissue samples were obtained from Choithram Hospital and Research Centre, Indore, after resection at surgery from patients already diagnosed of having breast cancer and undergoing radical mastectomy. The study involved tissue samples from 20 patients with breast cancer. The tissue samples were kept preserved in formalin. A section of the tissue sample provided by the histopathologist was taken for fluorescence studies. The histopathological report provided by the histopathologist was taken as the "Gold Standard". The cancerous tissue samples were invasive ductal carcinomas. The uninvolved areas of the resected cancerous specimens were treated as normal.

2.3 Spectroscopic Measurements

In-vitro autofluorescence spectra were acquired from a total of 105 tissue sites, of which 63 were from cancerous and the rest were from uninvolved breast tissues. Each site was treated separately and classified via the diagnostic algorithm developed. For recording autofluorescence spectra, the tissue samples kept at room temperature in formalin were taken out in a petridish and the tip of the fiber-optic probe was placed in contact with the tissue site to be investigated. Prior to recording spectra from a tissue sample, the probe was washed with ethyl alcohol and cleaned dry with a piece of cotton. A background spectrum was acquired with the probe placed in air. It was subtracted from all subsequently acquired spectra. From each site, spectra were recorded for 3s (i.e. 30 pulses) and were averaged to yield a single spectrum per site. From each site, spectra were recorded in the 375-700 nm spectral range with a resolution of ~2nm. During each measurement of tissue fluorescence, a reference spectrum was also acquired simultaneously from the phosphor-coated tip of an additional fiber illuminated with N_2 laser radiation leaking from the other end of the N_2 laser cavity. The peak of this reference spectrum was used to normalize the acquired tissue spectra and thus account for the observed pulse to

pulse variation of the N_2 laser power. The intensity of fluorescence from each tissue site is reported in this calibrated unit. The minimum signal–to-noise ratio was ~100:1 at the fluorescence maximum.

3. SPECTRAL DATA

The N_2 laser excited autofluorescence spectra averaged over all the cancerous and uninvolved sites of breast tissue samples investigated are shown in Fig.1. In Fig.2 we show the corresponding spectra normalized with respect to the spectrally integrated fluorescence intensity values. Each tissue fluorescence spectrum consisted of 768 intensity values (corresponding to 768 pixels of the ICCD) spanning the wavelength range of 375-700nm. While it is apparent from Fig1 that the fluorescence yield of cancerous tissue sites is much higher as compared to that of the surrounding uninvolved tissue sites, the significant difference in the spectral intensity distribution between the tissue types is evident from Fig2. In Fig.3 we show a histogram for the spectrally integrated fluorescence intensity values for paired cancerous and surrounding uninvolved tissue samples from each of the 20 patients. The considerable difference in the values of the cancerous and surrounding uninvolved tissue spectra is apparent. No significant difference was observed in the values for the ratio of standard deviation to the mean values for the tissue types.

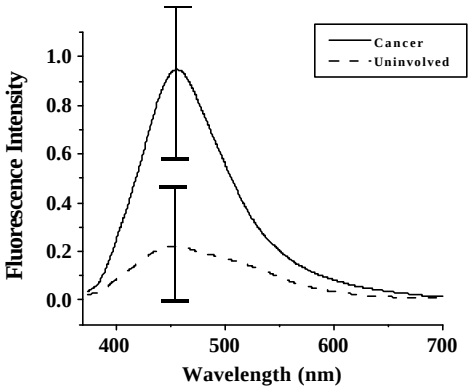

Figure 1: Site-averaged N_2 laser-excited autofluorescence spectra from cancerous (solid line) and adjoining uninvolved breast tissue

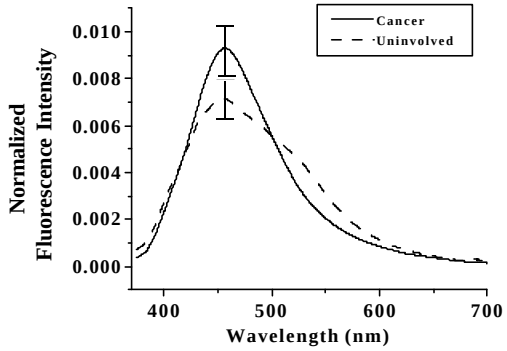

Figure 2: Area-normalized site-averaged autofluorescence spectra from cancerous (solid line) and adjoining uninvolved breast tissue (dashed line).

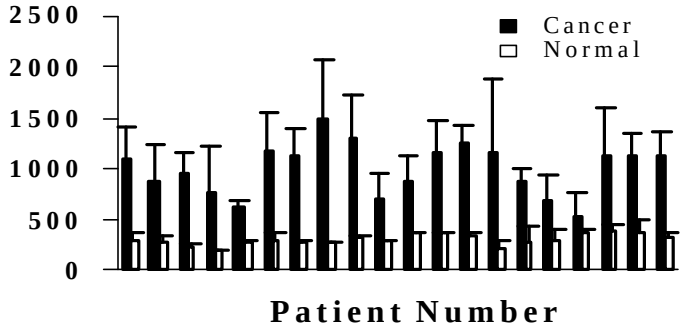

Figure 3: Histogram for the spectrally integrated intensity for paired cancerous and adjoining uninvolved breast tissue samples from each of the 20 patients. The values shown represent the average over the sites investigated in the given tissue samples.

3.1 Discrimination Analysis

The use of autofluorescence spectra to classify the tissue sites as normal or cancerous requires an appropriate diagnostic algorithm that can best classify the measured spectra by using a stored database of spectra of tissues of known histopathologic classification. Over the years a variety of diagnostic algorithms of varying rigor have been developed for optical diagnosis of cancer[2,3]. Most of the earlier algorithms are based on empirically selected indices like absolute or normalized fluorescence intensities[2, 6-8, 12], ratio of intensities at selected pairs of emission wavelengths[13] or ratio of integrated intensities over appropriately chosen wavelength bands[14] that are used either directly for discrimination[6-7,13] or as inputs to statistical analytical techniques like multivariate linear regression (MVLR) analysis to form a discrimination function[12,14]. Recent efforts are directed towards using statistical pattern recognition techniques[15] to exploit the entire spectral information content of the full range of spectral data for extracting the best diagnostic features and accurately classifying them into the corresponding histopathologic categories[10-11, 16-19]. We have used principal component analysis (PCA)[20] as the feature extractor and artificial neural network (ANN)[21] as the classifier to develop a diagnostic algorithm for discriminating the cancerous from the surrounding uninvolved sites of formalin fixed breast tissue samples. For the development of the algorithms we made use of an available database of recorded spectra from tissue sites whose class membership was already known. The algorithm was trained to not only separate this set of labeled input data into its constituent classes, but also, to predict the true class-membership of a tissue spectrum that is not part of the database using the approach of leave-one-out cross validation as well as validation over independent test set.

Prior to the development of diagnostic algorithms the entire set of spectral data from the cancerous and the uninvolved tissue sites of the patients was randomly split into two groups: training data set and validation data set ensuring that both sets contain roughly equal number of spectral data from each histopathologic category. The purpose of the training data set was to develop and optimize the diagnostic method and the purpose of validation set was to prospectively test its accuracy in an unbiased manner. The random assignment was carried out to ensure that not all the spectral data from a single individual were contained in the same data set. Next, the spectral data of the training set were used as inputs for the development of the diagnostic algorithms.

In order to develop diagnostic algorithms two different approaches were taken. In one approach, only the spectrally integrated fluorescence intensity scores computed from the raw spectral data were used as inputs to the classifier, and in the second approach, fluorescence intensities recorded over all the wavelengths normalized with respect to the spectrally integrated fluorescence intensities were used for classification. The objective of the first approach was to evaluate the discrimination efficacy of the fluorescence yield that was observed to be in general higher for cancerous as compared to the uninvolved tissue sites. The advantage of using such a diagnostic feature based only on the spectrally integrated intensity is that it could considerably simplify the experimental arrangement since no spectral resolution is required. The second approach exploits the total spectral information content of the full range of spectral data without any absolute intensity information. This would be helpful in situations where absolute intensities fail to show any significant difference between the tissue types. The spectral data of the training set was used to develop the algorithms and the validation set data was used for validation of the developed algorithms. The mathematical formulation and associated theoretical background of PCA and ANNs have been discussed in references[20, 21] . In the following, we will briefly present the basic ideas for the purpose of description of our algorithm.

3.2 Principal component analysis (PCA)

PCA[20] is a transformation used for representing high dimensional data in fewer dimensions such that maximum information about the data is present in the transformed space. In the case of LIF spectral data x with dimension N=714, the objective of PCA was to find the transformation Φ_M such the new M-dimensional (M<N) random vector $y_M = \Phi_M^T x$ contained the maximum information about x. To find Φ_M it is customary to minimize the mean square error between x and the approximation $x' = \Phi_M y_M$ i.e. we selected Φ_M such that $E[(x-x')^2]=E[(x-\Phi_M y_M)^2]$ was minimized. This solution for Φ_M satisfied the eigenvalue-eigenvector equation $C\Phi_M = \Lambda\Phi_M$, where Λ was a diagonal matrix whose elements were the eigenvalues of the covariance matrix $C=E[xx^T]-\mu\mu^T$, μ being the mean of the input spectral data matrix. The columns of Φ_M solution were the M eigenvectors of C with largest eigenvalues and are called the principal components (PC). This

solution also maximized the variance of the output random vector or the principal component scores in the $y_M = \Phi_M{}^T x$ transformed space, i.e. it maximized $E[yy^T]$.

The average values of principal components were calculated for each principal component for cancerous and uninvolved breast tissue sites. An unpaired two-tailed Student's t-test[22] was employed to test the statistical significance of the differences between the means of the principal component scores for cancerous and uninvolved tissue. The statistically significant principal components were then used for classification of the tissue types.

3.2 Artificial neural network (ANN)

An ANN[21] is a self-adaptive massively parallel machine-learning system composed of layers of processing elements (PEs). A PE is a construct composed of a set of inputs and corresponding weights (input connection strengths) that are combined to produce a result that is passed to a transfer function ultimately generating an output value that may be used by other PEs. This transfer function is used to constrain output to a particular range.

Typically, an ANN is composed of three types of layers: an input layer that passes data vectors to other layers; an output layer that produces an output vector that often represents the classification outcome for the corresponding input vector; and the hidden layers that take data from an input layer or a previous hidden layer and pass the transformed data to an output layer or a subsequent hidden layer. A learning strategy is used to make incremental changes to the weights in order to optimize some error criterion.

A supervised ANN requires the desired output from each input vector in order that it may be compared to the actual output generated by the ANN (Fig.4). The learning strategy attempts to minimize a global error function for the set of training data. Local errors are computed for each PE in order to make adjustments to the weights. This process is repeated for each input vector in the training set and the ANN continues iteration through the set until an acceptable minimization of the error is achieved.

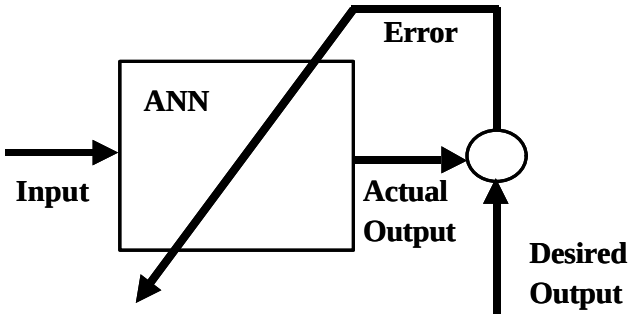

Figure 4: A supervised artificial neural network.

A feed-forward ANN has unidirectional data flow from the input layer, through each hidden layer, and finally to the output layer. In other words, no PE is allowed to pass its output to a PE in a previous layer or back to itself (feedback).

An RBF network is an ANN whose hidden PEs are radially symmetric as shown in Fig.5. Although many RBF variants exist, we used a traditional one[23]. A radially symmetric PE, j, has a center $\mathbf{c}_j$, stored as its corresponding weights from the input layer. An Euclidean distance measure is used to determine how close an input vector is to each $\boldsymbol{c}_j$. The transfer function, φ, used is an RBF, i.e. its output increases or decreases monotonically with distance from center. Typical RBFs are Gaussian, multiqudratic, or inverse multiquadratic. We have used an inverse multiquadratic RBF, as the transfer function where the output decreases monotonically with distance from center i.e. the output is large when the distance is small. Thus the output for PE_j and the input vector $\boldsymbol{x}$ is $\varphi(\|\boldsymbol{x}-\boldsymbol{c}_j\|)$. As such, the hidden PEs are prototypes for a cluster

of data points. The hidden layer is linearly connected to an output layer. Standard *k*-means clustering is used to determine the $\boldsymbol{c}_j$'s and a nearest-neighbor heuristic is used to determine the widths of the inverse-multiquadratic. A gradient decent iterative strategy[24] is used to determine the weights from the hidden layer to the output layer.

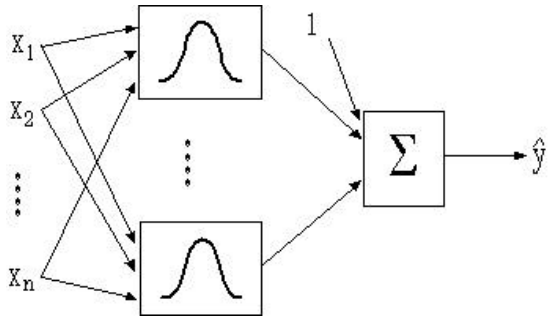

Figure 5: A typical RBF Architecture.

4. RESULTS AND DISCUSSIONS

Fixation of excised tissue samples in formalin is a routine process in any clinical pathological laboratory due to the time lapse between tissue excision and histopathological examination. The objective is to prevent autolysis and keep the properties and structure of the samples as close to their original state as possible so as to keep the state of microenvironment intact after excision and fixation. Formalin fixation meets this objective by promoting the cross-linkage of the amine groups in collagen[25] thereby promptly interrupting its metabolism as well as stabilizing its structure. However, from spectroscopic point of view, this particular mechanism of fixation is expected to lead to significant differences in the autofluorescence signatures of fresh and formalin fixed tissues particularly in the amino acid fingerprinting regions of proteins. In Fig.6 we show the N_2 laser excited autofluorescence spectra of fresh and formalin-fixed cancerous and uninvolved breast tissue. Each spectrum is the average of the spectra over the respective number of tissue sites investigated. As expected, significant differences in spectral line shape as well as intensities between the fresh and formalin-fixed tissue samples are apparent. While the spectra of fresh breast tissues are characterized by two major wavelength bands centered around 390nm and 470nm characteristic of the connective tissue proteins collagen and the co-enzyme NADH respectively[6-7], the prominent 390nm band is practically washed out for both the cancerous and normal breast tissues after formalin fixation effectively leading to spectra with a single broad band (centered around 455nm) for formalin fixed tissues. Further, in formalin fixed tissues, the overall fluorescence yield is observed to get considerably enhanced. Although a direct comparison is not possible, similar increase in fluorescence intensity was also observed by Filippidis et al[5] in formalin fixed peripheral vascular tissues. They observed this increase during the first 24 hours of fixation after which no noticeable change in intensity occurred. The increase in intensity can be attributed to extraction of blood from the tissue samples due to prolonged immersion in formalin.

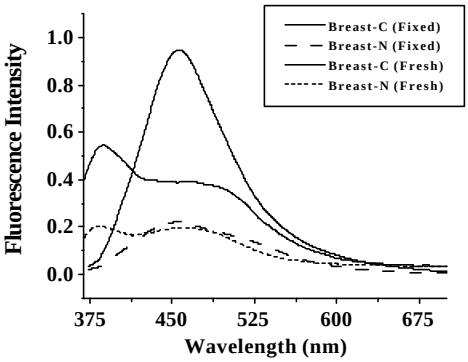

Figure 6: Site-averaged N_2 laser-excited autofluorescence spectra from fresh and formalin-fixed breast tissue. Dark solid and dashed lines represent fixed cancerous and fixed adjoining uninvolved breast tissue respectively, and the light solid and dashed lines represent fresh cancerous and fresh adjoining uninvolved breast tissue respectively.

Although formalin fixation introduces considerable variation in the autofluorescence spectra between fresh and fixed tissue sample the results of this study show that differences in fluorescence from cancerous and uninvolved sites of the breast tissue samples are still preserved. In fact, similar to fresh breast tissues, the formalin-fixed cancerous breast tissue sites have been found to be significantly more fluorescent as compared to the surrounding uninvolved breast tissue sites. This suggests that formalin fixation has minimal influence on the nature of changes introduced in the cancerous and uninvolved region of the breast tissue.

PCA of the area-normalized spectra resulted in four principal components that collectively accounted for 99.9% of the total variance of the spectral data. Of the four principal components, only four (PC-1, PC-2, and PC-3) were found to have significantly different ($p<0.001$) values for cancerous and uninvolved breast tissue. The difference in the values for the other component (PC-4) for the two diagnostic categories was not statistically significant ($p>0.05$). Therefore, classification was carried out using these three principal components which together accounted for 99% of the total variance, PC-1 accounting for 97%, PC-2 the 1% and PC-3 the remaining 1%.

In table-1 we summarize the performance of the diagnostic algorithms for the data sets comprising the spectrally integrated fluorescence intensity scores as well as the area-normalized spectral intensity values over the different wavelengths. For ANN based algorithms classification results were obtained for two cases. In one, ANN was used as a classifier with PCA providing the diagnostically relevant features (ANN-PCA) and in the second case ANN was used for classification with the full set of spectral features. For comparison's sake, a conventional nearest mean-classifier (NMC), based on least Euclidean distance of the test features from the means of the prototype features of the corresponding tissue types in the training set, was also used for classification. These results are also listed in the same table. In all the cases, the sensitivity and specificity values for the training set data were obtained on the basis of leave-one-out cross validation.

Table-1 Classification results of the RVM and nearest-mean classifier based diagnostic algorithms for the training and the independent validation data sets. Sensitivity and specificity values in the training set data represent leave-one-out cross validation values. The number of relevance vectors (RV) generated by the classifiers is also listed in the table.

Diagnostic Algorithm	Data set	Training Set data		Validation set data	
		Sensitivity(%)	Specificity(%)	Sensitivity(%)	Specificity(%)
ANN	Spectrally integrated intensity	88	100	97	100
NMC		81	100	81	100
ANN	Area-normalized spectra	97	100	97	100
PCA+ANN		100	100	97	100
PCA+NMC		81	100	84	100

A perusal of the diagnostic results listed in table-1 demonstrates that the ANN based diagnostic algorithm based on the full set of spectral features of the area-normalized spectral data as well as the diagnostically relevant features obtained via PCA of the area-normalized spectral data resulted in a slightly better classification performance over the algorithm based on the spectrally integrated intensity scores derived from the raw spectral data. However, the performance of both the algorithms was significantly improved as compared to that based on the nearest-mean classifier. The superior classification performance of the ANN classifier originates from the built-in capability of the ANN approach[21] to separate classes, which are not linearly separable in the original parametric space. Further, it is pertinent to note that since the nearest-mean classifier considers the least Euclidean distance of the test feature from the means of each class as the classification criterion, it is not expected to perform well on non-symmetric data like the autofluorescence spectral data that may have multiple clusters per class. This follows because when an input data has multiple clusters per class it might so happen that the mean for a class of two clusters can lie close to the mean of another class.

From the viewpoint of pattern recognition, the task of tissue classification based on LIF spectral data is a pattern classification problem, and the feature vector for classification is the measured intensities corresponding to the different

pixels (of the detector) that specify the dimension of the features. If directly working with all these spectral features whose dimension is much higher (N=768 in this case) as compared to the size (53 in this case) of the training samples, the classifier might suffer from the so called "curse of dimensionality" causing it to have poor generalization in classification performance[15]. The use of PCA prior to ANN reduces the feature dimension by solving the "curse of dimensionality" problem. This is evident from the observed improvements in the classification performance of the diagnostic algorithm based on ANN for the diagnostically relevant PCs as compared to the full set of spectral features leading to an increase of 3% in the sensitivity values for leave-one-out cross validation.

5. CONCLUSIONS

To conclude, an in-vitro study on N_2 laser excited autofluorescence spectroscopy of formalin fixed breast tissue samples are reported. Significant spectral differences were observed between the cancerous and the surrounding uninvolved tissue sites of formalin fixed breast tissue samples similar to that observed in the case of unfixed, fresh breast tissue samples. Use of a diagnostic algorithm developed using PCA as the feature extractor and the ANN as the classifier could correctly classify 30 out of 31 cancerous and 21 out 21 uninvolved breast tissue sites in the validation data set and 32 out of 32 cancerous and 21 out of 21 uninvolved breast tissue sites in the training set based on leave-one-out cross-validation. This demonstrates the potential of the autofluorescence technique as an in-vitro diagnostic modality in clinical pathology setting.

ACKNOWLEDGEMENTS

The authors thank Dr. S Pathak of Choithram Hospital and Research Centre, Indore for providing the pathologically characterized formalin-fixed tissue samples and several fruitful discussions. They would also like to thank Mr. S Shyam Sunder, Mr. C Rajan and Mr. A G Bhujle for their contribution to the development of the experimental system.

REFERENCES

1. Macfarlane S P, Reid R, Callander R, Pathology Illustrated, Churchill Livingstone, http://www.ontariopathologists.org/resources/cancer
2. "In vivo fluorescence spectroscopy and imaging for oncological applications", G A Wagnieres, W M Star, B C Wilson, *Photochem Photobiol*; **68**, 603-632 (1998).
3. "Fluorescence spectroscopy of neoplastic and non-neoplastic tissues", N Ramanujam, *Neoplasia* **2(1)**, 1-29 (2000).
4. "Effect of formalin fixation on the near-infrared Raman spectroscopy of normal and cancerous human bronchial tissue", Z Huang, A McWilliams, S Lam, J English, D I McLean, H Lui and H Zeng, International J. Oncology, **23**, 649 (2003).
5. "Effect of liquid-nitrogen and formalin-based conservation in the in vitro mesurement of laser-induced fluorescence from peripheral vascular tissue", G Filippidis, G Zacharakis, A Katsamouris, A Giannoukas, M Mouktzela and T G Papazoglou, *J. Photochem. Photobiol. B: Biol.*, **47**, 109 (1998).
6. "N_2 laser excited autofluorescence spectroscopy for human breast cancer diagnosis", P K Gupta, S K Majumder, A Uppal, *Lasers Surg Med*, **21**, 417 (1997).
7. "UV excited autofluorescence spectroscopy of human breast tissues for discriminating cancerous tissues from benign tumor and normal tissues", S K Majumder, A Uppal and P K Gupta, *Lasers Life Sci.*, **8**, 249 (1999).
8. "Autofluorescence spectroscopy of tissues from human oral cavity for discriminating malignant from normal", S K Majumder, A Uppal, P K Gupta, *Lasers Life Sci.* **8**, 211 (1999).
9. "Effect of handling and fixation processes on fluorescence spectroscopy of mouse skeletal muscles under two-photon excitation", M G Xu, E D Williams, E W Thompson and M Gu, *Appl. Opt.*, **39(34)**, 6312 (2000).

10. "Nonlinear pattern recognition for laser-induced fluorescence diagnosis of cancer", S K Majumder, N Ghosh, S Kataria and P K Gupta, *Lasers Surg. Med.* **33**, 48 (2003).
11. "A pilot study on the use of autofluorescence spectroscopy for diagnosis of the cancer of human oral cavity", S K Majumder, S K Mohanty, N Ghosh, P K Gupta, D K Jain, F Khan, *Current Science* **79(8)**, 1089 (2000).
12. "Ultraviolet laser-induced fluorescence of colonic tissue: Basic biology and diagnostic potential", K T Schomacker, J K Frisoli, C C Compton, T J Flotte, J M Richter, N S Nishioka, T F Deutsch, *Lasers Surg. Med.*, **12,** 63 (1992).
13. "Optical spectroscopy of benign and malignant breast tissues", Y Yang, A Katz, E J Celmer, M Zurawska-Szczepaniak, R R Alfano, *Lasers Life Sci.*, **7(2)**, 115 (1987).
14. "In-vitro diagnosis of human uterine malignancy using N_2 laser-induced autofluorescence spectroscopy", S K Majumder, A Uppal, P K Gupta, Current Science, **70(9)**, 833 (1996).
15. "Statistical pattern recognition: A review", A K Jain, R P W Duin, J Mao, *IEEE Trans Pattern Anal.* **22(1)**: 4 (2000).
16. "Cervical precancer detection using multivariate statistical algorithm based on laser-induced fluorescence spectra at multiple excitation wavelengths", N Ramanujam, M Follen-Mitchell, A Mahadevan-Jansen, S Thomson, G Staerkel, A Malpica, T Wright, N Atkinson, R Richards-Kortum, *Photochem. Photobiol.*,**64**, 720 (1996).
17. "Ensembles of radial basis function networks for spectroscopic detection of cervical pre-cancer", K Tumer, N Ramanujam, J Ghosh, R Richards-Kortum, *IEEE Trans* BME., **45(8)**, 953 (2001).
18. "Comparison of multiexcitation fluorescence and diffuse reflectance spectroscopy for the diagnosis of breast cancer (March 2003)", G M Palmer, C Zhu, T M Breslin, F Xu, K W Gilchrist and N Ramanujam, *IEEE Transactions on Biomedical Engineering*, **50(11)**, 1233 (2003).
19. "Classification of in-vivo autofluorescence spectra using support vector machines", W M Lin, X Yuan, P Yuen, W I Wei, J Sham, P Shi C and J Qu, *Journal of Biomedical Optics*, **9(1)**, 180 (2004).
20. R A Johnson, D W Wichern. eds. Principal Component Analysis in *Applied Multivariate statistical Analysis*, Second edition, Prentice-Hall International, New Jersy, 1988.
21. "Artificial neural networks: A tutorial", A K Jain, J Mao, K M Mohiuddin, *Computer*, 628 (1996).
22. Kleinbaum D G, Kupper L L, Muller K E eds. Chapter-3, *Applied Regression and other multivariable methods*, Second edition, Duxbury Press, Belmont, California, 1988.
23. "Classification and Regression using Linear Networks, Multilayer Perceptrons and Radial Basis Functions", V.K Rayker., *ENEE 739q Spring 2002 Course Assignment* 2 Report.
24. C.M Bishop, *Neural Networks for Pattern Recognition*, Oxford, Clarendon press, 1995.
25. W Bloom and D W Fawcett, *A Textbook of Histology*, W. B. Saunders, Philadelphia, 1980.

* shkm@cat.ernet.in ; phone 091-731-2488452; fax 091-731-2488430.