RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances

RSC Advances RSCPublishing

ARTICLE

Cite this: DOI: 10.1039/x0xx00000x

Genetic algorithm spectral feature selection coupled with quadratic discriminant analysis for ATR-FTIR spectrometric diagnosis of basal cell carcinoma via blood sample analysis

Received 00th January 2012,
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

Mohammadreza Khanmohammadi,*[a] Keyvan Ghasemi[a] and Amir Bagheri Garmarudi[a]

Keywords:

BCC skin cancer, FTIR Spectroscopy, Quadratic Discriminant Analysis, Genetic Algorithm

A diagnostic approach for basal cell carcinoma (BCC) has been developed based on investigation of infrared spectra of blood samples. Different predictive procedures were developed using quadratic discriminant analysis (QDA) combined with simple filtered method and genetic algorithm (GA) as a feature subset and wavelength selection strategy. Results showed 94.74% and 100 % of accuracy for Filtering-QDA and GA-QDA models respectively.

Introduction

Basal cell carcinoma (BCC) is about 45% abundant in skin cancer known as the most common malignant disease in human. According to literature most of skin cancer patients are more than 40 years old. High rate of this disease has cost sever diagnosis related researches enabling the fast, correct in useful detection of malignancy. Nowadays, the most common diagnostic methods for cancer diagnosis rely on pathologic biopsies. During the last decade spectroscopy has been introduced as an efficient tool for such an aim. The main drawbacks of biopsy based detection approaches are in relation with:

- Heterogeneity of samples
- Difficult preparation procedure.
- Being harmful for the organ
- Probable leading to spread of cancer.
- Being a time consuming procedure
- Non-useful quality of the prepared sample for analysis due the physical damages.

Taking into account the dependence of obtained results to the experience of pathologist, it seems necessary to suppose more reliable approaches. Fourier-transform infrared spectroscopy (FTIR) is one of the recent introduced methods which have been employed to detect cancer tissues from non-cancerous ones in different types of cancer [1-7].

As mentioned before, within the last decade the use of Fourier transform infrared (FTIR) spectroscopy, has been much to the forefront in the development of practical diagnostic tools in cancer detection studies. It has been demonstrated that FTIR spectroscopy can reliably distinguish multiple types of carcinoma from healthy tissue but preparation of tissue samples is so time consuming and it needs a very sensitive keeping condition. The same as cancer, many other diseases have been studied by different techniques of vibration spectroscopy, using blood samples as the analyte [8-10].

Edwards and Duntley have studied skin samples by optical spectroscopy [11]. Lack of an easier procedure for malignancy diagnosis studies e.g. analyzing the samples which are provided easier, with more availability and providing more reliable results has led to utilizing blood samples for FTIR analysis. Previous studies confirm that chemometric processing of the data obtained from FTIR spectroscopic analysis of blood samples could be a useful method for this aim. Spectroscopic investigation of whole blood samples has been reported as a novel, accurate, precise and easy to perform method for cancer diagnosis [12-13].

Previous experiences have confirmed the role of chemometrics in quantitative and qualitative analysis. Several chemometric techniques are applied while a research is performed in order to combine the statistical skills with chemical ones, achieving the aimed analytical goals [14]. in case of BCC ,application of soft independent modeling of class analogy (SIMCA) and artificial neural networks (ANN) have been practically useful for diagnosis [13,15] .Research activities is faced to development of more comfortable and reliable chemometrics models which may usually provide more powerful analyzing system for the

[a]Chemistry Department, faculty of Science, IKIU, Qazvin, Iran. Address, Tel-Fax: +098 281 3780040
*Corresponding Author: E-mail: mrkhanmohammadi@gmail.com

researchers. Thus efforts for development of more specific models are always continued.

In this research quadratic discriminant analysis (QDA) coupled with variables selection method have been applied to improve the output of model. Nowadays, the application of discriminant function for prognostic and diagnostic classification in clinical medicine has become very common in different aspects. In particular, wavelength selection method could usually modify the output of analysis. In the present work, FTIR absorbance spectra of blood samples from healthy people and those with BCC skin cancer has been considered. An exploratory analysis was performed on spectrometric data in order to evaluate the feasibility of QDA. Filtered spectra were coupled by QDA and GA-QDA models for the diagnosis of BCC skin cancer. The efficiency of QDA with abovementioned modifiers was compared in automated classification of BCC.

Experimental

Apparatus and software

Spectroscopic data were obtained by a Bomem (Quebec, Canada) MB series FTIR spectrometer equipped with a DTGS mid-range detector; a KBr: Ge/Sb2S3 coated beam splitter and a SiC source, using SpectraTech (Warrington,UK) in-compartment contact with sampler horizontal attenuated total reflector with a 45° ZnSe trough plate. Pls-plus / Iq button to the GRAMS/32 version 5 and higher software Galactic industries corporation was used to process the absorbance data. All FTIR absorbance spectra were recorded in 900-2000 cm-1 spectral region with a data point spacing of 3.85 cm-1, being processed by MATLAB (Ver. 7.11) based m.files.

Sample preparation and typical modeling

Spectroscopic studies have been performed in first 48 hours after sampling to reduce the variations in blood structure. In all spectrometric investigations the whole blood samples were used. EDTA (5% concentrated, Merck) was used an anticoagulant and avoiding the spectral overlapping and shielding effect of water as the main ingredient of blood sample, it was set as the background in all of recorded spectra (totally 108 samples).

Training step was the 38 samples were considered for both sex, male and female (18 samples in normal case and 20 samples in cancer case). Then 15 known samples (7 normal and 8 cancer) were introduced to training model as validation. Finally, 75 unknown samples (35 normal and 40 cancer) were predicted by the models, being compared with pathologic clinical results. Patients were 40-60 years old, who had passed pathologic cancer diagnosis tests and all were informed about our research aim, participating in research tests voluntarily. The pathologic diagnostic studies were performed in Shahid Rajaee Hospital, Qazvin, Iran (Associated with Qazvin University of Medical Sciences).

Result and discussion

Spectroscopic features of FTIR spectral from BCC cases

Significant variations have been observed in dermal FTIR spectra comparing to the other skin components. These are mainly due to absorptions from collagen in dermis. Subtle differences in protein structure and nucleic acid content of BCC cases would cause severe spectral variations [16]. Accordingly, FTIR spectroscopy as a finger print monitoring method is useful to explore distinctive characteristics of basal cell carcinoma versus normal skin samples. The main spectral characteristics of BCC are:

- increased hydrogen bonding of the phosphodiester group of nucleic acids
- decreased hydrogen bonding of the C-OH groups of proteins
- increased intensity of the band at 972 cm-1
- decreased intensity ratio between the CH3 stretching and CH2 stretching bands
- accumulation of unidentified carbohydrates

which some of them are specifically observed in BCC [17]. Table 1 demonstrates the main spectral changes which have observed in BCCs.

Table 1- Main spectral changes in IR spectra due to BCC

Spectral Region (cm-1)	vibration	Functional Group	Biochemical structure
1680-1640	stretching	amide I	Protein
1300-1220	stretching	amide III	Protein
940-928	stretching	C-C	Amino acids
1450-1420	scissoring	CH2	Lipid
1305-1295	in-plane twisting	-(CH2)n-	Lipid
640–800	out-of-plane	NH bending	Protein

Precise study of multi signal spectral features in 1220-1360, 900-990 and 830-900 cm-1 spectral region would enable a discriminative analysis BCC and normal skin Spectra [18-20]. BCC is a multi-factorial disease in which both environmental factors and host genetic factors are implicated in tumorigenesis. Although the molecular genetic pathogenesis of BCC is not clear, p53 mutation has been reported and may form an important part of the pathogenetic sequence in a majority of BCC cases [21-25]. In case of biochemical variations during the disease progression, some of the most noticed signals are expression of simple epithelial keratins and high level of kinase D. Another biomarker is a cytoskeletal protein named keratin 17 (K17). K17 normally functions to provide mechanical support, and participates in the regulation of programmed cell death as well as protein synthesis, in various types of skin epithelial cells [26-30]. Researchers are engaged with discovering the structural changes associated with BCC, which encompass several components of the cell, in agreement with the expected complexity of the malignant phenotype. The recent developed experimental approaches for early detection of skin cancer which consists of a melanoma gene detection blood test, would detect the probable circulation of melanoma cells. The next step is to discover the certain proteins which do exist mostly on cancerous, called tumor antigen 90 (TA-90). The abovementioned blood test would check for antibodies to TA-90. Considering the general aspects of these investigations, blood analysis is a useful trend for proposing novel diagnosis procedures. Successful diagnosis of BCC via FTIR spectroscopy depends on the power of the chemometric model. More powerful model with more specific functioning causes more reliable results.

Investigating the main informative spectral region of blood samples (1800-900 cm-1) the most significant features are realized (Table 2).

Table 2- Spectral features of blood IR spectra in case of BCC

Spectral Region (cm^{-1})	Type of vibration	Functional Group	Structure
1680-1630	stretching	C=O	Amide RCONHR
1640-1620	bending	C=O	Amide RCONHR
1570-1540	Stretching bending	C-N -N-H	$RCONR_2$

In many situations, spectral signals at these regions would overlap and their identification is impossible. According to previous experiences, in order to evaluate the effect of signals due to all biochemicals of blood, the FTIR absorbance spectra in total spectral region (1800-900 cm-1) was selected to carry out the data processing. Figure 1 shows a typical average IR spectrum of cancer and normal cases.

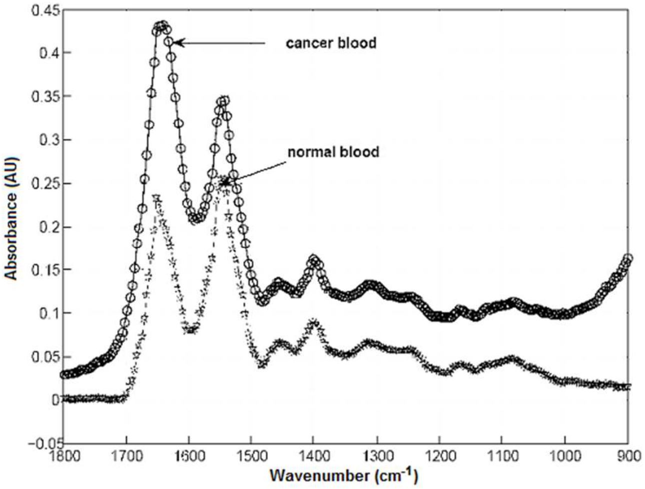


Figure 1-Typical average spectra of blood samples from normal and cancer cases

Linear Discriminant Analysis (LDA)

Basic expression

LDA is a well-studied useful method of pattern recognition. It has been originally proposed by Fisher [31-32] and is applied very often in chemometrics. LDA [33,34] aims to produce a linear classifier which can successfully differentiate classes. A group of objects belong exactly to one out of k similar classes and the class membership is known for each object. Each object is defined by some characteristic parameters. LDA uses this information to calculate (k-1) linear discriminant functions describing a separation hyper plane which would optimally discriminate k classes. These functions are applied to assign the unknown objects to classes. The normal vector of this separation plane is the direction that maximizes the ratio of the difference between classes (interclass variance) to the differences within the classes (intra class variance). This direction of the vector is simply the direction that connects the class means if the intraclass variance is one in all directions, i.e. if the intraclass covariance matrix is the unity matrix. In this research LDA was applied using spectral band ratios as parameters to distinguish FTIR spectra of normal blood sample and cancerous cases. Applications of LDA with diagonal covariance matrix estimation (naive bayes) have also been reported [35-37]. LDA classification is based on the Mahalanobis distance which is derived from a common covariance matrix for all classes, while QDA classification is based on a Mahalanobis distance that is based on class-specific covariance matrices. QDA is less subject to constraints in case of the distribution of objects in space. The Naive Bayes classifier is designed to be used when features are independent of one another within each class, but it appears to work well in practice even in the situation by which the independence assumption is not valid.

In this work QDA was applied to 38 samples as a training set and performance of validation set was compared to models with a reduction strategy in model for prediction improvement. Error rate, correction rate, sensitivity and specificity are some of the figures of merit have been reported in table 3. Obtained results confirm the reliability of model for training set while the validation results are not noticeable enough. In other words, sensitivity of the results for training is about 90.00% while this point is about 75.00% for validation which is not acceptable for clinical purpose.

Table 3- Results of QDA for training and validation set by total range of spectra.

Parameter	Training	Validation
Correction Rate	92.11	80.00
Error Rate	7.89	20.00
Sensitivity	88.89	85.71
Specificity	95.00	75.00

Filtering of spectra

Previously presented model would illustrate the usefulness of selecting subsets of variables which provide acceptable predictive power together, as opposed to ranking variables according to their individual predictive power. This problem is investigated outlining the main directions which have been taken to tackle it. The main aim is to reduce the dimension of the data by finding a small set of important features which can give good classification performance. Filtering methods rely on general characteristics of the data to evaluate and to select the feature subsets without involving the chosen learning algorithm [38-39].

Filters are usually used as a pre-processing step since they are simple and fast. A widely-used filtering approach is to apply a univariate criterion separately on each feature and assuming that there is no interaction between features. This method relies on simple statistic hypothesis tests which assume that the data are independently sampled from a normal distribution. As a result, t-test is applied on each feature and p-value (or the absolute values of t-statistics) is compared for each feature as a measure of how effective it is at separating groups. T-test with a null hypothesis on data in the vectors x and y that were independent random samples from normal distributions was performed- with equal means and equal but unknown variances- on the other hand the alternative was that the means are not equal. The result of the test is returned in h. h = 1 indicating a rejection of the null hypothesis at the 5% significance level. h = 0 indicates a failure to reject the null hypothesis at the 5% significance level. So when the p-value for the under studying feature was below the critical values of hypothesis (0.05 in this case) we reject the null hypothesis and conclude that the two population means are different at the 0.05 significance level. It means the feature has significant effect if selected for classification. In order to get a general idea of how well-separated the two groups are by each feature, empirical cumulative distribution function (CDF) of p-values is plotted (Figure 2). Cumulative distribution functions are used to specify the distribution of multivariate random variables. There are about 25 % of features having p-values close to zero and 54% of features having p-values smaller than 0.05, meaning there are about 138 features among the original 260 features that have strong discrimination power. One can sort these features according to their p-values (or the absolute values of the t-statistic) and select some features from the sorted list. However, it is usually difficult to decide how many features are needed unless one has some domain knowledge or the maximum number of features that can be considered has been dictated in advance based on outside constraints.

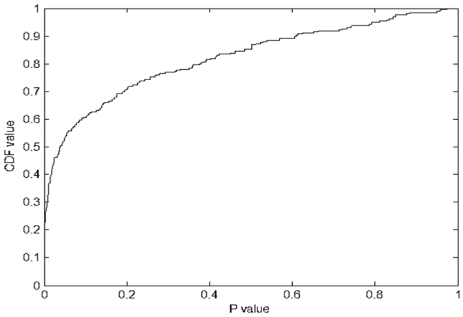


Figure 2- Empirical Cumulative distribution Function plot after filtering.

Accordingly it was decided to use the features with p-value smaller than 0.05 as significant features. Results of QDA model for training and test set have been reported in table 4.

Table 4- QDA results after filtering

Parameter	Training	Validation
Correction Rate	94.74	80.00
Error Rate	5.26	20.00
Sensitivity	94.44	85.71
Specificity	95.00	75.00

During the filtering approach, there are 138 features with highest significance in model. The main goal here is to reduce the dimension of the data by finding a small set of most important features which can give reliable classification performance. As a result, after evaluation of selected feature by filtering, about 23% of selected features are in critical infrared spectral region of BCC about 1500-1700 cm-1.

This selection is compatible to the assigned spectral features of blood IR spectra in case of BCC which have been shown in table 2. Results show some improvement in prediction performance. Clearly this modification is not reliable enough to improve the model because filtering relies on general characteristics of the data to evaluate and to select the feature subsets without involving the chosen learning algorithm method (for this case classification). Selected features are shown in figure 3.

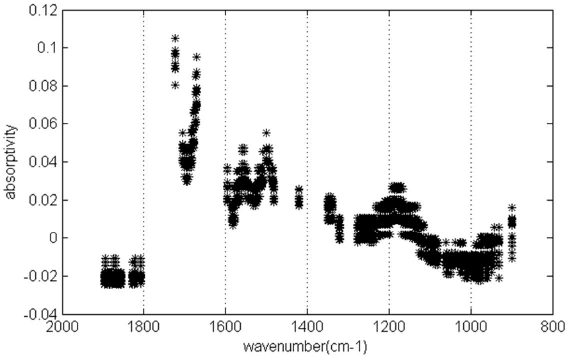


Figure 3- selected wavelength after filtering.

Genetic Algorithm (GA)

Since optimization problems arise frequently, this makes GAs quite useful for a great variety of tasks. As in all optimization problems, researchers are faced to the problem of maximizing/minimizing an objective function f(x) over a given space X of arbitrary dimension. Classification problem deals with associating a given input pattern with one of the distinct classes. Patterns are specified by a number of features (representing some measurements made on the objects that are being classified) so it is common to propose them as d-dimensional vectors, where d is the number of different features. Patterns are points in this d-dimensional space and classes are sub-spaces. Classification problem reduces to determining which region a given pattern falls into. If classes do not overlap they are said to be separable and, in principle, one can design a decision rule which will successfully classify any input pattern. A decision rule determines a decision boundary which partitions the feature space into regions associated with each class. It represents our best solution to the classification problem. [40-43]

Selected parameters used in GA are shown in table 5. These parameters have been selected after performing genetic algorithm procedure frequently, extracting the best set of them.

Table 5- Selected parameters for GA.

population size	10
generation	50
variable	16
fitness function	Quadratic Discriminant Analysis

mutation rate	0.01

Figure 4 shows the selected wavelength of GA, approving that these feature selection methods have selected some wavelength with high efficiency in classification algorithm.

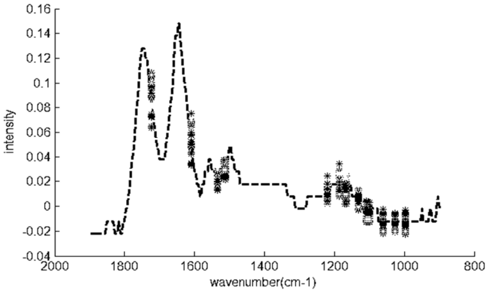


Figure 4- selected feature by GA

The ultimate goal of the genetic algorithms is the optimization of a given response function. These algorithms are inspired by the theory of evolution: in a living environment, the "best" individuals have a greater chance to survive and a greater probability to spread their genomes by reproduction. In this work our response function was QDA and all selected feature for being in best group should be examine with the classifier performance. As a results after application of GA-QDA ,there were about 16 exact features that were based on this selection criteria and also were best matching with range of spectral domain that were significantly responsible in BCC diagnostic. Output of GA–QDA approach shows great modification in calibration and test set which is reported in table 6. It is important to realize that the performance of GA is based on population of solution rather than on one specific solution. In this case the accuracy of GA was defined, so GA check several set of population (wavelengths in this case) which are so closer to solve our problem and when the QDA classifier was in the selection process these selected features cause best classification performance for diagnosis purposes. Not that this properties is opposite to filtering method. IN filtering the selection criteria is general and determines by significant in simple statistic test and shows poor capability rather than GA.

Table 6- QDA results after GA wavelength selection

Parameter	Training	Validation
Correction Rate	100.0	93.33
Error Rate	0.00	6.67
Sensitivity	100.00	85.71
Specificity	100.00	100.00

Method Comparing

In order to evaluate the role of each variable selection strategy in model performance target factor analysis (TFA) was performed. In TFA, some perfectly characterized target vectors with chemical meaning e.g., reference spectra, can be tested to observe whether they lie in the space spanned by the data set. If they are found to lie within the space, they can be identified as real sources of variation. Thus all of 15 samples of validation set were projected on calibration space before and after feature selection model and their distance and angle has been calculated. All of calculated distances and angles of GA-QDA dataset are closer to the subspace of calibration and shows capability of the selected method. On the other hand the results of projection of filtered data do not demonstrate significant change. Comparing the output of QDA and GA-QDA, GA is confirmed to be capable to reduce mean of distance of data from 0.071 to 0.049. The receiver operating characteristic curves (ROC) have been compared in figure 5 to test the quality of the discrimination and compare among the variable selection methods. Finally, all models analyzed in the present study have been evaluated by test samples and compared results are presented in Tables 7 and 8.

Table 7- TFA evaluation for validation set.

method	Distance	angle
QDA	0.0719	4.1229
GA-QDA	0.0491	2.8149

Table 8- comparing between feature selection methods.

Parameter	QDA	FILTERED-QDA	GA-QDA
Correction Rate	90.67	94.67	97.33
Error Rate	9.33	5.33	2.67
Sensitivity	85.71	94.29	1.00
Specificity	95.00	95.00	95.00

On the other hand, ROC of GA-QDA at several thresholds around optimum value has been plotted in figure 6. Discrimination power of classifier in the test set is realized in suitable condition and can be verified by several cut off points according to table 9.

Table 9- Statistical parameters of different cut off points

	Accuracy (%)	Std. Error (%)	95% C.L. (cut off point)
1	94.4	2.7	0.890 0.997
2	94.2	2.8	0.888 0.996
3	94.2	2.8	0.887 0.996
4	92.5	3.2	0.863 0.987

In this work, target factor analysis has been employed for selected spectral data ,and results of spanned spectra to calibration set confirms that selected feature have a closer distance to the calibration set and this exhibited the quality of input data from obtained spectra has been improved. On the other hand, ROC curve for several unique thresholds have been plotted and results in other point than optimum value for sensitivity and specify (with a cutoff point 0.89 and 1 in plot) show that there are some tolerances in prediction of classifier. In table 9 results of several cut off point is illustrated. Accuracy has been measured by the area under the ROC curve. An area

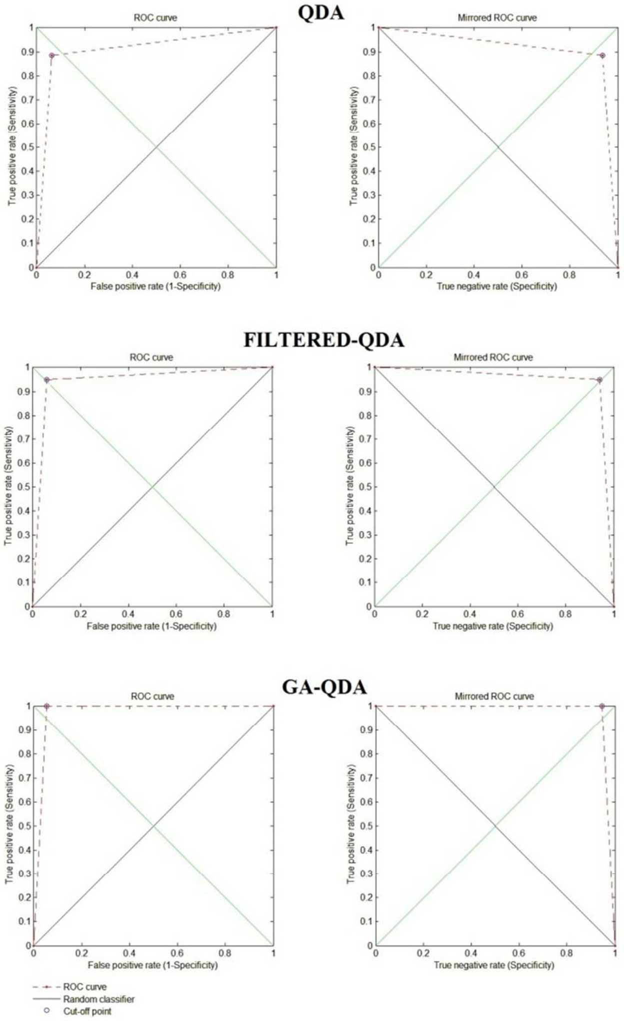
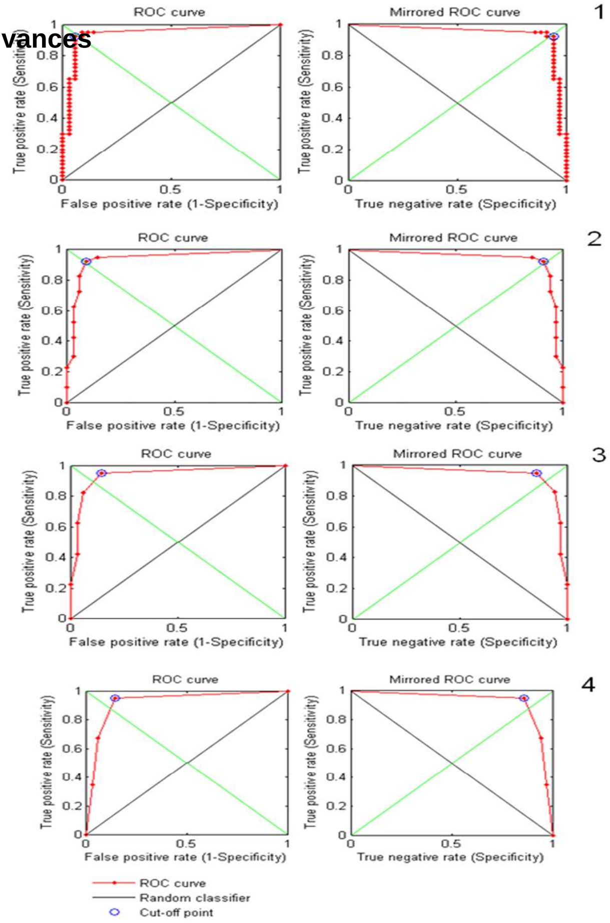


Figure 5- ROCs for different chemometric methods



of 100% represents a perfect test and decrement in this value shows lack of reliability in classification power. Fortunately, the reduction in these values is not significant in this work.

Figure 6- ROCs of GA-QDA at several thresholds

Conclusions

A diagnostic approach for basal cell carcinoma (BCC) has been developed based on investigation of infrared spectra of blood samples. Chemometric and multivariate statistical analyses were utilized to generate an automated IR-based histology without any chemical preparation. Different predictive procedures were developed using Quadratic Discriminant Analysis (QDA) combined with simple filtered method and genetic algorithm (GA) as a feature subset and wavelength selection strategy. Spectroscopic studies were performed in 900-2000 cm-1 spectral region with 3.85 cm-1 data space. The main object was to discriminate between the spectra of healthy and malignant cases. Results showed 94.74% and 100 % of accuracy for Filtering-QDA and GA-QDA models respectively. In the first step, 38 blood samples were applied to construct the model. In order to modify the capability of QDA in prediction of test samples, two different feature selection methods have been applied. It was concluded that ATR-FTIR spectroscopy and QDA-GA chemometric technique would be a reliable approach for detection of BCC via blood sample analysis.

References

1 P. Lasch , W. Wasche ,WJ. Mc Carthy ,G. Muller and D.Naumann , Cell Mol Biol ,1998, **44**,189.

2 S. Argov , J. Ramesh and A. Salman, J Biomed Opt ,2002, **7**, 1.

3 J. Ramesh , A.Salman ,S. Argov and et al, Subsurf Sens Technol Appl ,2001, **2**,99.

4 A.Salman ,S. Argov ,RK. Shau ,E. Bernshtain ,S. Walfisch and S. Mordechai , Vib Spectrosc, 2004, **34**, 301.

5 MA. Cohenford and B. Rigas , Proc Natl Acad Sci,1998, **95**, 15327.

6 S.Argov ,RK. Sahu,E. Bernshtain and et al., Biopolymers 2004, **75**, 384.

7 N.Fujioka ,Y Morimoto,T Arai and M. Kikuchi , Cancer Detect Prev ,2004, **28**, 32.

8 RA.Shaw ,M. Leroux ,M Paraskevas ,FB. Guijon ,S. Kotowich and HH Mantsch, Washington DC: SPIE, 1998,**3257**, 42.

9 GH.Werner,J. Früh ,F. Keller and et al., Washington DC: SPIE, 1998, **3257**, 35.

10 AJ.Berge,TW. Koo ,I Itzkan , G Horowitz and MS Feld , Appl Opt, 1999, **38**, 2916.

11 EA.Edwards and SQ. Duntley, Am J Anat 1939, **65**, 1.

12 M.Khanmohammadi,MA. Ansari,A. Bagheri Garmarudi ,G. Hassanzadeh and G. Garoosi,Cancer Invest ,2007, **25**, 397.

13 M.Khanmohammadi ,R. Nasiri ,K. Ghasemi ,S. Samani and A. Bagheri Garmarudi , J Cancer Res Clin 2007, **133**, 1001.

14 Wackernagel H. Multivariate Geostatistic: An Introduction with Application. 3rd ed. Heidelberg: Springer, 2003.

15 M. Khanmohammadi, A. Bagheri Garmarudi, and K. Ghasemi, J. Chemometrics,2009, **23**, 538.

16 LM.McIntosh,M. Jackson,HH. Mantsch,JR. MansWeld and AN. Crowson, Vib Spectrosc,2002, **28**,53.

17 PT.Wong,SM. Goldstein,RC. Grekin ,TA. Godwin,C. Pivik and B. Rigas , Cancer Res ,1993,**53**,762.

18 LM.McIntosh,M. Jackson,HH. Mantsch,MF. Stranc,D. Pilavdzic and AN. Crowson , *J Invest Dermatol* ,1999, **112,** 951.

19 M.Gniadecka,HC. Wulf ,NN.Mortensen ,OF. Nielsen and DH. Christensen, *J Raman Spectrosc,* 1997, **28**, 125.

20 R.Montesano, P.Hainaut and J. Hall*, IARC Sci Publ,* 1997, **142,** 291.

21 JF.Doré ,R. Pedeux,M. Boniol,MC. Chignol and P .Autier, IARC Sci Publ, 2001, **154**, 81.

22 M.Barbareschi,S. Girlando ,P. Cristofolini,M. Cristofolini and R. Togni , Boi S, 1992,**21**, 53.

23 EV.Salomatina,B. Jiang,J. Novak and AN. Yaroslavsky, *J Biomed Opt,* 2006, **11,** 064026.

24 M. D'Errico,AS. Calcagnile,R. Corona and et al, *Cancer Res,* 1997 , **57,** 53.

25 CR.Shea ,NS. McNutt,M. Volkenandt,J. Lugo,PG. Prioleau and AP. Albino , *Am. J. Pathol.,* 1992,**141,** 25.

26 N. Calli Demirkan,N. Colakoglu and E. Düzcan , Pat Oncol Res ,2000, **6**, 272.

27 L. Yan ,M. Chen and G. Yu , Bulletin of Hunan Medical University, 1999, **24**, 179.

28 C.Cazal ,MR. Ely ,APV. Sobral and DNN. Padilha , Clínica Integrada, 2006, **6**, 267.

29 WJ.Wang ,JY. Huang ,CK. Wong and YT. Chang, *Arch. Dermatol Res,* 2000, **292,** 379.

30 G.Krekels ,M. Verhaegh ,SS. Wagenaar,F. Ramaekers and H.Neumann , E*ur J Dermatol,* 1997, **7,** 158.

31 R.A. Fisher, Annual Eugenics ,1936,**7**, 179.

32 D.Coomans, M. Jonckheer, D.L. Massart,I. Broeckaert and P.Blockx, Anal. Chim. Acta,1978,**103**, 409.

33 R.O. Duda, P.E. Hart and D.G. Stork, Wiley, New York, 2001.

34 J.H. Friedman, Journal of the American Statistical Association 1989,**84**, 165.

35 T.Mitchell, Machine Learning, McGraw Hill,1997.

36 M.Vangelis,A. Ion and P. Geogios, Third Conference on Email and Anti-Spam,2006.

37 H.George, John and Pat Langley, The Eleventh Conference on Uncertainty in Artificial Intelligence,1995.

38 G. John ,R. Kohavi , Artificial Intelligence, 1997,**97**,272

39 Nina Zhou and Lipo Wang , Genomics, Proteomics & Bioinformatics, 2007,**5**, 249

40 C.W.M. Yuena, W.K. Wonga, S.Q. Qiana, L.K. Chana and E.H.K. Fungb, Expert Systems with Applications ,2009, **2**, 2037.

41 M.Tsenga, S. Chenb, G.Hwangc and M.Shend, ISPRS Journal of Photogrammetry and Remote Sensing , 2008,**63**,202.

42 S. Gunala and R. Edizkan, Information Sciences ,2008, **178**, 3716.

43 L. Rokach, Pattern Recognition ,2008,**41**, 1676.