Spectral analysis of cell-graphs for automated cancer diagnosis

Cigdem Demir[1], S. Humayun Gultekin[2,*], Bülent Yener[1]
[1]Rensselaer Polytechnic Institute, [2]Mount Sinai School of Medicine
[*] Present address: Oregon Health and Science University
demir@cs.rpi.edu, gultekin@ohsu.edu, yener@cs.rpi.edu

Abstract

The **cell-graph** *approach captures the information encoded in tissue samples by capturing the spatial distribution of cells and their cluster formations. In a cell-graph, nodes and edges represent the cell clusters and pairwise relationships between them, respectively. It is shown in [1] that the features of cell-graphs of cancerous tissues are significantly different from those of healthy tissues and benign reactive/inflammatory processes. Thus, cell-graphs can be used for the purpose of automated cancer diagnosis [1]. In this paper, we present a new set of features and show how effectively they aid to automated cancer diagnosis. In particular, we investigate the spectral decomposition of adjacency matrices and normalized Laplacian of the cell-graphs. To the best of our knowledge, this is the first use of the spectral graph theory in tissue-based automated cancer diagnosis. Working with 646 brain biopsy samples of 60 different human patients, we demonstrate that the spectra of the cell-graphs of cancerous tissues are unique and distinguishable from those of the healthy and inflamed tissues with accuracy >90% (with a sensitivity of 91.57% and specificities of 93.75% and 98.15% for the inflamed and healthy tissues, respectively).*

Keywords: Automated cancer diagnosis, spectral graph theory, feature extraction

1. Introduction

In traditional cancer diagnosis, pathologists manually examine biopsies to make diagnostic assessments. The assessments are largely based on visual interpretation of cell morphology and tissue distribution, lacking of quantitative measures. Therefore, they are subject to considerable inter-observer variability [2-4]. To circumvent this problem, numerous studies aim at quantifying the characteristics of cancerous cells and tissues that distinguish them from their counterparts. Such quantification facilitates to design automated systems that operate on quantitative measures and, in turn, to reduce the inter-observer variability. However, this is not a straightforward task because of the complex nature of the image scenes; a typical tissue image consists of touching and overlapping cells. Additionally, in the tissue image, there is usually a considerable amount of noise that arises from staining the tissue; uneven distribution of stain usually causes problems in processing the stained material [5].

In literature, there are different approaches to quantify the characteristics (features) of cells and tissues: morphological, textural, intensity-based, and topological. The *morphological* approach quantifies the size and shape of a cell or its nucleus [6-9]. The *textural* approach makes use of spatial inter-relationships for the pixels to extract features [10-11] and quantifies properties such as the smoothness, regularity, and coarseness of the image [12-16]. The *intensity-based* approach employs the distribution of the intensity values of pixels to define its features [17-19]. The *topological* approach quantifies the spatial distribution of the cells within a tissue [20-22]. Although these approaches lead to promising results in automated cancer diagnosis, they suffer from one or both of the two problems: (i) the difficulty of determining the exact locations of cells/nuclei in the biopsy image or (ii) the noise that arises from its staining process.

The ***cell-graph*** approach relies on cluster formation in cancerous cells to define their distinctive features. In this method, we identify the cell clusters on a tissue image as the nodes and compute the spatial dependency between every pair of these nodes to probabilistically assign the edges [1]. Unlike the previous demonstrations, the cell-graph approach makes use of the cell clusters instead of the individual cells. Therefore, it eliminates the necessity of determining the exact locations of cells/nuclei on a tissue image, which allows using the low-magnification images. Furthermore, this approach relies on the dependency between the cell clusters rather than the pixels themselves and does not directly use the pixel values in feature extraction. Because of that, it is likely immune to noise inherit in a biopsy image.

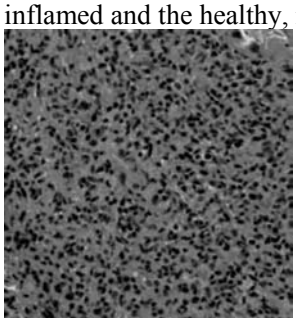
1.1 Contribution of this paper

In this paper, we investigate the *spectral decomposition* (i.e., eigenvalue decomposition) of the cell-graphs extracted from biopsy images and define a new set of

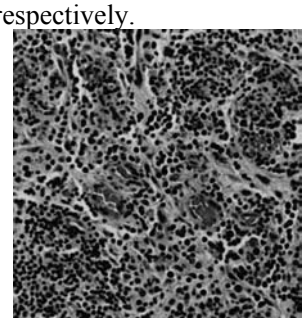
features derived from the spectra of the cell-graphs to distinguish the cancerous tissues from their counterparts. To the best of our knowledge, this is the first proposal for the use of the spectral graph theory for the purpose of automated cancer diagnosis.

Our motivation in this work is the fact that the spectral decomposition of a graph is strongly related with its structural properties [23, 24]. Besides this relationship, there is a strong theoretical work and a large literature behind the spectral graph theory which is applied in several research fields such as chemistry and communication networking. We believe that, this work will open up the possibility to explore the characteristics of cancerous tissues with a strong and well-studied mathematical tool and, hence, this paper could stimulate further investigation of the spectral graph theory into the area of automated cancer diagnosis.

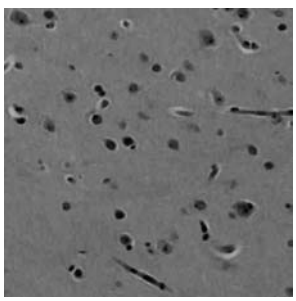
In this work, we conduct our experiments on a total of 646 human brain biopsy samples of 60 different patients. This data set includes cancerous (*malignant glioma*) samples, benign inflammatory processes (thereafter referred to as *"inflamed"*), and healthy samples. Although, the cancerous and healthy tissues can easily be distinguished even with untrained eyes (as shown in Fig.1(a) and Fig.1(c)), it is not straightforward to differentiate the cancerous and inflamed tissues (as shown in Fig.1(a) and Fig.1(b)). Our experiments demonstrate that, the graph spectra of cancerous and non-cancerous (inflamed and healthy) tissues exhibit different characteristics, yielding a sensitivity of 91.57% and specificities of 93.75% and 98.15% for the inflamed and the healthy, respectively.



(a)



(b)



(c)

Figure 2. Microscopic images of brain biopsy samples: (a) a cancerous (malignant glioma) tissue, (b) a benign reactive/inflammatory process, (c) a healthy tissue.

2. Background

2.1 Cell-graph generation

A cell-graph $G = (V, E)$ is an undirected and unweighted graph without loops, with V and E being the nodes and edges of the cell-graph G. In generating the cell-graph from a tissue image, we first determine the node set V by identifying the cell clusters in the image and then probabilistically establish the edges between the nodes in V [1].

In *node identification*, we first classify the pixels of the image as "cell" or "background". To do so, we first learn how to classify them by using the *k*-means algorithm [25]. On the training samples, the *k*-means algorithm clusters the pixels into *k* clusters based on their color information. Subsequently, a human expert assigns these clusters to either "cell" or "background" class. Once learned, the clusters and their corresponding class assignments are used for classifying the pixels of the other images. Next, we embed a grid on the image and compute the probability of being a cell cluster for each grid entry. For a grid entry, the probability is computed as the average of the values of the pixels located in this particular grid entry, assigning 1 to the pixels of cell class and 0 to the pixels of background class. Then, the grid entries with a probability greater than a threshold are identified as the nodes of the cell-graph.

In *edge establishing*, we probabilistically set an edge between any pairs of nodes using a probability function that decays with the increasing Euclidean distance between the nodes. In this work, we use a probability function of $P(u,v) = d(u,v)^{\alpha}$, with *d(u,v)* being the Euclidean distance between the nodes *u* and *v*, and α is the exponent that controls the connectivity level of the graph ($\alpha \leq 0$).

2.2 Spectral graph theory

The spectrum of a graph is the set of all eigenvalues of its adjacency matrix or its Laplacian. The eigenvalues of a graph are known to be closely related to major graph invariants [23, 24].

The *adjacency* matrix of a cell-graph $G = (V, E)$ is defined by

$$A(u,v) = \begin{cases} 1 & \text{if u and v are adjacent} \\ 0 & \text{otherwise} \end{cases}$$

With d_v indicating the degree of a vertex *v*, the *normalized Laplacian* of G is defined by

$$L(u,v) = \begin{cases} 1 & \text{if u = v and } d_v \neq 0, \\ -\dfrac{1}{\sqrt{d_u d_v}} & \text{if u and v are adjacent,} \\ 0 & \text{otherwise.} \end{cases}$$

Although the adjacency matrix of a graph was much more studied in the past, the eigenvalues of the Laplacian relate to the graph invariants better than the eigenvalues of the adjacency matrix [24, 26]. For instance, the number of the connected components in a graph is equal to the number of eigenvalues of its normalized Laplacian that have a value of 0. Moreover, as the eigenvalues of the normalized Laplacian lie in the range of [0,2], it is easier to quantify the properties on the normalized Laplacian spectra of the graphs with different sizes and, in turn, to compare the normalized Laplacian spectra of such graphs.

In this work, by using the adjacency and normalized Laplacian spectra of the cell-graphs, we define new sets of features as detailed below.

Features derived from the adjacency spectrum

Let $\lambda_0 \leq \lambda_1 \leq \mathrm{K} \leq \lambda_{n-1}$ be the eigenvalues of the adjacency matrix of a graph G with n nodes. The range of these eigenvalues can vary according to the graph in contrast with those of the normalized Laplacian matrix. We define the following spectral features for adjacency spectrum:

- The *spectral radius*, which is defined as a maximum absolute value of eigenvalues in the spectrum ($\max_{1 \leq i \leq n} |\lambda_i|$).
- The *second largest* absolute value of the eigenvalues in the spectrum.
- The *eigen exponent* which is defined as the slope of the sorted eigenvalues as a function of their orders in log-log scale [27]. In this work, we use the slope between the first largest and its next largest 50 eigenvalues.
- The *sum* of the eigenvalues.
- The sum of the squared eigenvalues (thereafter referred as to "*energy*").

Features derived from the Laplacian spectrum

Let $0 = \lambda_0 \leq \lambda_1 \leq \mathrm{K} \leq \lambda_{n-1} \leq 2$ be the eigenvalues of the normalized Laplacian of a graph G with n nodes. We define the following spectral features for normalized Laplacian spectrum; in Fig.2, the first 5 of them are illustrated on an exemplary cell-graph:

- The number of the eigenvalues with a value of 0, which gives the *number of connected components* in the graph.
- The slope of the line segment representing the eigenvalues that have a value between 0 and 1. For its computation, we first fit a line on these eigenvalues by using linear regression, and then compute the slope of this line (thereafter referred as to "*lower-slope*").
- The *number of the eigenvalues with a value of 1*. This feature was previously defined for the analysis of the Internet graphs [28]; it was used to compare the empirical data collected from the Internet and the data generated by the synthetic data generators.
- The slope of the line segment representing the eigenvalues that have a value between 1 and 2 (thereafter referred as to "*upper-slope*").
- The *number of the eigenvalues with a value of 2*. Its value is only greater than 0 if and only if a connected component of the graph is bipartite and nontrivial [24].
- The *sum* of the eigenvalues. For the graph spectrum, $\sum_i \lambda_i \leq n$ with n being the number of the nodes in the graph. The equality holds for the graphs that have no isolated vertices (the vertices with a degree of 0).
- The sum of the squared eigenvalues (thereafter referred as to "*energy*").

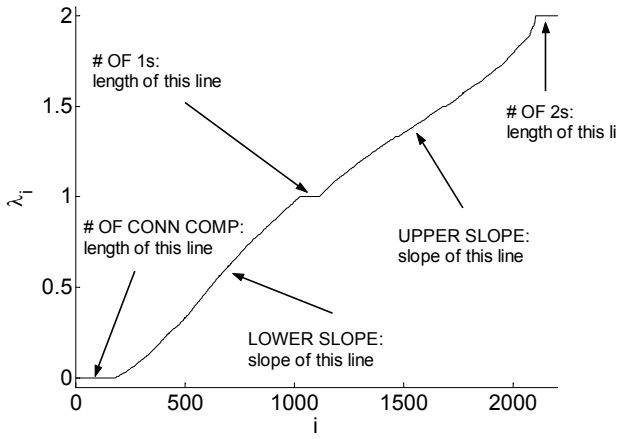


Figure 2. Laplacian spectrum of an exemplary cell-graph. First 5 spectral features are illustrated.

3. Experiments

3.1 Data set preparation

We conduct our experiments on a data set that consists of 646 microscopic images of brain biopsy

samples of 60 different patients[1]. This data set consists of 329 cancerous samples of 41 patients, 210 healthy samples of 14 patients, and 107 inflamed samples of 9 patients; for 4 of the patients, we have both the cancerous and healthy tissue samples. These samples consist of a 5-6 micron-thick tissue section stained with hematoxylin and eosin technique and mounted on a glass slide.

We take the images of samples with a magnification of 100× by using a digital camera. The images are stored in the RGB color space. Prior to segmentation, we convert the RGB values to their corresponding La*b* values since the La*b* is a uniform color space that provides separate color and detail information [29]. The images used in this data set consist of 480×480 pixels.

We divide this data set into training and test data sets. We use half of the patients of each type in the training set and the rest in the test set. The test set includes approximately 8 tissue images from each patient. The training set still includes 8 tissue images from each cancerous patient. Since the number of the patients of other types is less than that of the cancerous patients, we use larger amounts of the tissue images for the healthy and inflamed. Additionally, since the number of available inflamed tissues is less than those of healthy and cancerous, we replicate the inflamed tissues in the training set. In summary, in the training set, we use 163 cancerous tissues of 20 patients, 150 inflamed tissues of 5 patients (the data set included 75 inflamed tissues prior to the replication), and 156 healthy tissues of 7 patients. In the test set, we use 166 cancerous tissues of 21 patients, 32 inflamed tissues of 4 patients, and 54 healthy tissues of 7 patients. Note that the tissues that come from the same patient are either in the training or the test set, but not in both. This prevents over-optimistic test accuracies.

3.2 Parameter selection

In the generation of cell-graphs, we have 4 control parameters: (i) value of k, (ii) grid size, (iii) node threshold, and (iv) exponent value of α. The value of k in k-means algorithm should be large enough to represent all different parts of a tissue such as nuclei, cytoplasm, and blood vessels. However, its value should be small enough so that the human expert can differentiate the clusters and reproducibly assign the corresponding classes to these clustering vectors. Considering this issue, we use a value of k as 16 in this work.

The grid size determines the size of the cell-clusters. Depending on the grid size, a cell-cluster consists of a single cell, a part of a cell or a bunch of cells. In this work, we use a grid size of 6 which matches the size of a typical cell. The node threshold eliminates the noise that arises from staining the tissues. In this work, we use a threshold value of 0.25 which eliminates the noise inherit in an image without resulting in significant information lost on the cell clusters for the selected grid size.

The value of exponent α determines the connectivity level of the cell-graph. Increasing the values of α towards 0 yields denser graphs, whereas decreasing it towards $-\infty$ produces sparser graphs. In this work, we conduct our experiments with different values of α, ranging between -4.8 and -2.0.

3.3 Results

In this work, we use the support vector machines (SVM) [30] with a kernel of the radial basis functions as our classifiers. As our feature sets, we separately use the spectral properties derived from the adjacency matrix of a graph and its normalized Laplacian.

Table 1. Classification results obtained on the test set when the features extracted from the *normalized Laplacian spectrum* are used.

		Specificity	
A	Sensitivity	Inflamed	Healthy
-2.0	100.00	34.38	98.15
-2.4	88.55	68.75	98.15
-2.8	89.76	78.12	98.15
-3.2	89.76	90.62	98.15
-3.6	91.57	93.75	98.15
-4.0	86.75	87.50	98.15
-4.4	89.76	90.62	98.15
-4.8	89.16	84.38	98.15

In Table 1, we report the sensitivity and specificity results obtained by using the features derived from the normalized Laplacian spectra of cell-graphs; the results reported in this table are all obtained on the test set. This table shows that the inflamed tissues cannot successfully be differentiated for larger values of α (i.e., relatively denser graphs). For the

[1] These 60 patients were randomly chosen from Pathology Department archives in Mount Sinai School of Medicine. The patients were adults with both sexes included. The identifiers were removed, and slides were numerically recoded corresponding to diagnostic categories by the pathologist, prior to obtaining digital images of the tissues. Therefore, the remaining two authors had access to images and diagnoses only, without retraceable personal identifiers.

graphs extracted by using a value of α smaller than -2.8, i.e., after the connectivity of a graph falls below a certain value, the inflamed tissues are differentiated successfully. For $\alpha = -3.6$, it leads to the best sensitivity and specificity results.

Table 2. Classification results obtained on the test set when the features extracted from the *adjacency spectrum* are used.

		Specificity	
A	Sensitivity	Inflamed	Healthy
-2.0	75.90	71.88	98.15
-2.4	79.52	65.62	98.15
-2.8	74.70	50.00	98.15
-3.2	81.33	53.12	98.15
-3.6	78.92	56.25	98.15
-4.0	82.53	59.38	98.15
-4.4	74.10	59.38	98.15
-4.8	72.89	65.62	98.15

Similarly, in Table 2, we report the classification results (on the test set) obtained by using the features derived from the adjacency spectra of cell-graphs. This table shows that the adjacency spectrum features yield successful results only for classifying the healthy tissues. The correct classification rate of the cancerous (sensitivity) and inflamed tissues are less compared to the correct classification in the case of the normalized Laplacian spectrum. These results are consistent with the fact that the eigenvalues of the normalized Laplacian relate to the graph invariants better than those of the adjacency matrix [24, 26]. For the normalized Laplacian, it is easier to define the distinctive graph features that yield better accuracy results.

Analysis of the individual normalized Laplacian features

In our experiments, we also analyze the spectral properties of the cell-graphs derived from their normalized Laplacian to identify the most distinctive features. For that, we train a SVM classifier by using a single feature. For the normalized Laplacian spectra, we report the sensitivity and specificity results in Table 3. In this table, we use the cell-graphs extracted by using the exponent value of -3.6, since this exponent yields both the best sensitivity and specificity results. Table 3 demonstrates that there exists no feature that can successfully differentiate all of the tissue types. Although there are many features that distinguish the healthy from the other types, only the number of the eigenvalues with a value of 0 (the number of connected components in the graph) can distinguish the cancerous and inflamed from each other.

Table 3. Classification results obtained on the test set when the classifier that employs of a single-feature of the *Laplacian spectrum* is used. The exponent is selected to be -3.6.

		Specificity	
Feature	Sensitivity	Inflamed	Healthy
# of 0s	87.35	87.50	5.56
lower slope	100.00	15.62	94.44
#of 1s	1.20	71.88	98.15
upper slope	69.28	43.75	98.15
#of 2s	55.22	61.96	73.33
Sum	56.63	40.62	98.15
energy	52.41	78.12	98.15

In Fig.3, we illustrate the distribution of the number of the connected components of the cell-graphs for the cancerous and inflamed tissues by using a box and whisker plot[2]. In this figure, we observe that the distribution of the number of the connected components in the cell-graphs of the cancerous and inflamed tissues is prominently separable for the edge exponents smaller than -2.8. This observation is closely related with the classification results reported in Table 1. In Table 1, the accuracy results for the inflamed tissues are higher in the case of the edge exponents greater than -2.8.

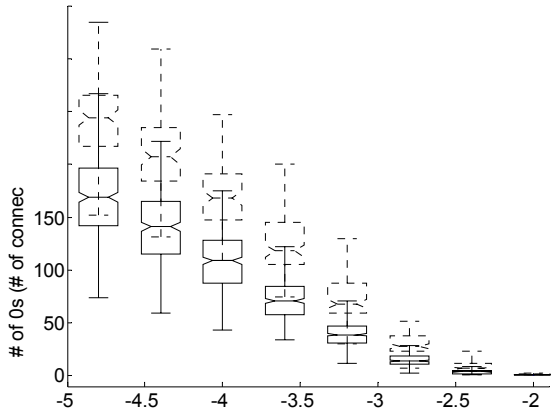


Figure 3. Distribution of the number of connected components for the cancerous (shown in *solid* lines) and inflamed (shown in *dotted* lines) samples.

[2] They are drawn by using *boxplot* function in Matlab. Each box in this figure shows the lower quartile, median, and upper quartile values and the whiskers show the extent of the rest of the data.

4. Conclusion

This work investigates the spectral decomposition (eigen decomposition) of the cell-graphs for the purpose of automated cancer diagnosis. In this work, we propose to use new sets of the spectral features derived from (i) the adjacency matrix and (ii) the normalized Laplacian of the cell-graphs to distinguish the cancerous tissues from their counterparts.

In this work, experiments are conducted on a data set consisting of 646 human brain biopsy samples of 60 different patients. These experiments demonstrate that the spectral features derived from the normalized Laplacian of the cell-graph successfully differentiate the cancerous from the healthy and inflamed, with a sensitivity of 91.57% and specificities of 98.15% and 93.75% for the healthy and the inflamed, respectively.

References

[1] C. Gunduz, B. Yener, and S.H. Gultekin, "The cell graphs of cancer," *Bioinformatics*, 20: i145-i151, Aug 2004.

[2] A. Andrion, C. Magnani, P.G. Betta, A. Donna, F. Mollo, M. Scelsi, P. Bernardi, M. Botta, and B. Terracini, "Malignant mesothelioma of the pleura: interobserver variability," *J Clin Pathol*., 48(9):856-860, Sep 1995.

[3] S.M. Ismail, A.B. Colclough, J.S. Dinnen, D. Eakins, D.M. Evans, E. Gradwell, J.P. O'Sullivan, J.M. Summerell, and R.G. Newcombe, "Observer variation in histopathological diagnosis and grading of cervical intraepithelial neoplasia," *Br Med J*., 298(6675):707-710, Apr 1989.

[4] P.J. Klinkhamer, G.P. Vooijs, and A.F. de Haan, "Intraobserver and interobserver variability in the diagnosis of epithelial abnormalities in cervical smears," *Acta Cytol.*, 32(6):794-800, Nov-Dec 1998.

[5] C. Demir, S.H. Gultekin, and B.Yener. Automated cancer diagnosis based on histopathological images: a systematic survey, Rensselaer Polytechnic Institute Technical Report TR-05-09, Troy, 2005.

[6] K. Blekas, A. Stafylopatis, D. Kontoravdis, A. Likas, and P. Karakitsos, "Cytological diagnosis based on fuzzy neural networks," *J Intelligent Systems*, 8:55-79, 1998.

[7] W.N. Street, W.H. Wolberg, and O.L. Mangasarian, "Nuclear feature extraction for breast tumor diagnosis," *IS&T/SPIE 1993 International Symposium on Electronic Imaging: Science and Technology*, 1905:861-870, San Jose, CA, 1993.

[8] J.-P. Thiran and B. Macq, "Morphological feature extraction for the classification of digital images of cancerous tissues," *IEEE T Bio-Med Eng*., 43(10):1011-1020, Oct 1996.

[9] W.H. Wolberg, W.N. Street, D.M. Heisey, and O.L. Mangasarian, "Computer-derived nuclear features distinguish malignant from benign breast cytology," *Hum. Pathol.*, 26(7):792-796, Jul 1995.

[10] R.M. Haralick, "Statistical and structural approaches to texture," *Proc of IEEE*, 67(5):786-804, 1979.

[11] M.M. Galloway, "Texture analysis using gray level run lengths," *Computer Graphics and Image Processing*, 4:172-179, 1975.

[12] F. Albregtsen, B. Nielsen, and H.E. Danielsen, "Adaptive gray level run length features from class distance matrices," *15th Int Conf on Pattern Recognition*, 3:746-749, Barcelona, 2000.

[13] A.N. Esgiar, R.N.G. Naguib, B.S. Sharif, M.K. Bennett, and A. Murray, "Microscopic image analysis for quantitative measurement and feature identification of normal and cancerous colonic mucosa," *IEEE T. Inf. Technol. B.*, 2(3):197-203, Sep 1998.

[14] A.N. Esgiar, R.N.G. Naguib, B.S. Sharif, M.K. Bennett, and A. Murray, "Fractal analysis in the detection of colonic cancer images," *IEEE T. Inf. Technol. B.*, 6(1):54-58, Mar 2002.

[15] P.W. Hamilton, P.H. Bartels, D. Thompson, N.H. Anderson, and R. Montironi, "Automated location of dysplastic fields in colorectal histology using image texture analysis," *J. Pathol.*, 182(1):68-75, May 1997.

[16] J. Smolle, "Computer recognition of skin structures using discriminant and cluster analysis," *Skin Res Technol.*, 6(2):58-63, May 2000.

[17] F. Schnorrenberg, C.S. Pattichis, C.N. Schizas, K. Kyriacou, and M. Vassiliou, "Computer-aided classification of breast cancer nuclei," *Technol. Health Care,* 4(2):147-161, Aug 1996.

[18] M. Wiltgen, A. Gerger, and J. Smolle, "Tissue counter analysis of benign common nevi and malignant melanoma," *Int J Med Inform.*, 69(1):17-28, Jan 2003.

[19] Z.H. Zhou, Y. Jiang, Y.B. Yang, and S.F. Chen, "Lung cancer cell identification based on artificial neural network ensembles," *Artif. Intell. Med.*, 24(1):25-36, Jan 2002.

[20] H.-K. Choi, T. Jarkrans, E. Bengtsson, J. Vasko, K. Wester, P.-U. Malmstrom, and C. Busch, "Image analysis based grading of bladder carcinoma. Comparison of object, texture and graph based methods and their reproducibility," *Anal Cell Pathol.*, 15(1):1-18, 1997.

[21] S.J. Keenan, J. Diamond, W.G. McCluggage, H. Bharucha, D. Thompson, B.H. Bartels, and P.W. Hamilton, "An automated machine vision system for the histological grading of cervical intraepithelial neoplasia (CIN)," *J Pathol.*, 192(3):351-362, Nov 2000.

[22] B. Weyn, G. van de Wouwer, S. Kumar-Singh, A. van Daele, P. Scheunders, E. van Marck, and W. Jacob, "Computer-assisted differential diagnosis of malignant mesothelioma based on syntactic structure analysis," *Cytometry*, 35(1):23-29, Jan 1999.

[23] D.M. Cvetkovic, M. Boob, and H. Sachs, *Spectra of Graph*, Academic Press, 1978.

[24] F.R.K. Chung, *Spectral Graph Theory*. CBMS Regional Conference Series in Mathematics, Providence, Rhode Island, American Mathematical Society, 1997.

[25] J.A. Hartigan and M.A. Wong, “A k-means clustering algorithm,” *Appl*. *Stat*., 28:100-108, 1979.

[26] B. Mohar, “The Laplacian spectrum of graphs,” *Graph Theory, Combinatorics and Applications,* Y. Alavi, G. Chartrand, O. Oellermann, and A. Schwenk, eds., (2):871-898, Wiley 1991.

[27] M. Faloutsos, P. Faloutsos, and C. Faloutsos, “On power-law relationships of the Internet topology,” *Proceedings of ACM/SIGCOMM,* 251-262, 1999.

[28] D. Vukadinovic, P. Huang, and T. Erlebach, “A spectral analysis of the Internet topology,” Technical report, ETH Zurich, 2001.

[29] G. Wyszecki and W.S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae,* 2nd edition, Wiley and Sons, 2000.

[30] C.J.C. Burges, “A tutorial on support vector machines for pattern recognition,” Data Mining and Knowledge Discovery, 2:121-167, 1998.