Automated cancer diagnosis based on histopathological images: a systematic survey

Cigdem Demir and Bülent Yener

Abstract

In traditional cancer diagnosis, pathologists examine biopsies to make diagnostic assessments largely based on cell morphology and tissue distribution. However, this is subjective and often leads to considerable variability. On the other hand, computational diagnostic tools enable objective judgments by making use of quantitative measures. This paper presents a systematic survey of the computational steps in automated cancer diagnosis based on histopathology. These computational steps are: *1.)* image preprocessing to determine the focal areas, *2.)* feature extraction to quantify the properties of these focal areas, and *3.)* classifying the focal areas as malignant or not or identifying their malignancy levels. In Step 1, the focal area determination is usually preceded by noise reduction to improve its success. In the case of cellular-level diagnosis, this step also comprises nucleus/cell segmentation. Step 2 defines appropriate representations of the focal areas that provide distinctive objective measures. In Step 3, automated diagnostic systems that operate on quantitative measures are designed. After the design, this step also estimates the accuracy of the system. In this paper, we detail these computational steps, address their challenges, and discuss the remedies to overcome the challenges, emphasizing the importance of constituting benchmark data sets. Such benchmark data sets allow comparing the different features and system designs and prevent misleading accuracy estimation of the systems. Therefore, this allows determining the subsets of distinguishing features, devise new features, and improve the success of automated cancer diagnosis.

Index Terms

Clinical decision making, automated cancer diagnosis, biomedical image analysis, segmentation, feature extraction, classification

I. INTRODUCTION

Today, cancer constitutes a major health problem. In the United States, it is the second leading cause of death. Approximately one out of every two men and one out of every three women get cancer at some point during their lifetime. Furthermore, the risk of getting cancer has been further increasing due to the change in the habits of people in our century such as the increase in tobacco use, deterioration of dietary habits, and lack of activity.

Fortunately, the recent advances in medicine have significantly increased the possibility of curing cancer. However, the chance of curing cancer primarily relies on its early diagnosis and the selection of its treatment depends on its malignancy level. Therefore, it is critical for us to detect cancer, distinguish cancerous structures from the benign and healthy ones and identify its malignancy level.

Traditionally, pathologists use histopathological images of biopsy samples removed from patients, examine them under a microscope, and make judgments based on their personal experience. While examining such images, a pathologist typically assesses the deviations in the cell structures and/or the change in the distribution of the cells across the tissue under examination. However, this judgment is subjective, and often leads to considerable variability [1], [2]. To circumvent this problem and improve the reliability of cancer diagnosis, it is important to develop computational tools for automated cancer diagnosis that operate on quantitative measures. Such automated cancer diagnosis facilitates objective mathematical judgment complementary to that of a pathologist, providing a second opinion for patients.

Over the last two decades, a tremendous amount of research work has been conducted for automated cancer diagnosis. This is partly because automated cancer diagnosis holds great promise for large-scale use in the advanced cancer treatment and partly because automated cancer diagnosis is not a straightforward task, with a number of challenges to be overcome. The first challenge is the *noise elimination* in the task of determining the focal areas in the image. The noise arises from staining the biopsy samples; uneven distribution of stain usually cause problems

C. Demir is with the Department of Computer Science, Rensselaer Polytechnic Institute, Troy, NY 12180, USA (email: demir@cs.rpi.edu)
B. Yener is with the Department of Computer Science, Rensselaer Polytechnic Institute, Troy, NY 12180, USA (email: yener@cs.rpi.edu)

in processing the stained material. In the case of focusing on the properties of nuclei/cells in the image, the second challenge is the *nucleus/cell segmentation*. This is challenging because of the complex nature of the image scenes (e.g., touching and overlapping cells) and the noise (e.g., stain artifacts). The third challenge is the *feature selection* to represent a cell/tissue in the task of cellular or tissue-level property quantification. The features should provide distinguishing quantitative measures to automatically diagnose the cancer. The last important challenge is the *system evaluation* in the task of diagnosis. Due to the limited amount of available data, there might be a considerable amount of bias if the system evaluation is not conducted properly.

In this paper, we present a systematic survey on the computational steps to automatically diagnose cancer by using histopathological images. In each step, we explain the techniques, address the challenges, and discuss the remedies offered by these techniques to overcome the challenges.

II. OVERVIEW

In this paper, we focus on the following problem. We are given an image of a tissue sample. The goal is to automatically decide on the existence of cancer in the tissue and/or determine the malignancy level of cancer by examining the histopathological properties of the tissue. These properties enable to capture the deviations in the cell structures and the changes in the cell distribution across the tissue, which are possibly caused by cancer (thereafter, we refer these deviations and changes as "**cellular-level**" and "**tissue-level**" changes, respectively). The automated cancer diagnosis, both at the cellular and tissue-levels, is based on (i) extracting information from the histopathological images of stained biopsies and (ii) examining this information by using either statistical analysis or machine learning algorithms. The sample histopathological images of brain and breast biopsies (stained using hematoxylin-and-eosin technique) are shown in Fig.1. As illustrated in this figure, the cell distribution is completely different for the cancerous and the healthy. Moreover, a difference in the cell structures (e.g., the cell sizes are different) is observed especially in the case of brain cancer.

The automated cancer diagnosis consists of three main computational steps: *preprocessing*, *feature extraction*, and *diagnosis*. The aim of the preprocessing step is to eliminate the background noise and improve the image quality for the purpose of determining the focal areas in the image. This step also comprises nucleus/cell segmentation in the case of extracting cellular-level information. The preprocessing becomes the most important yet difficult step for a successful feature extraction and diagnosis.

After preprocessing the image, features are extracted either at the cellular or at the tissue-level. The cellular-level feature extraction focuses on quantifying the properties of individual cells without considering spatial dependency between them. For a single cell, the morphological, textural, fractal, and/or intensity-based features can be extracted. The tissue-level feature extraction quantifies the distribution of the cells across the tissue; for that, it primarily makes use of either the spatial dependency of the cells or the gray-level dependency of the pixels. For a tissue, the textural, fractal, and/or topological features can be extracted.

The aim of the diagnosis step is (i) to distinguish benignity and malignancy or (ii) to classify different malignancy levels by making use of extracted features. This step uses statistical analysis of the features and machine learning algorithms to reach a decision. An overview of these three steps is given in Fig.2. In the following sections, we will study each of these steps in detail.

There are also other diagnosis approaches that extract information from biological data at molecular and organ levels. At the molecular level, the information is obtained either from gene expression signatures using microarrays [3], [4], [5], [6], [7], or from protein biomarkers using mass spectrometers [8], [9], [10], [11], [12]. At the organ level, screening techniques such as mammography [13], [14], [15], [16], [17] are employed. The automated cancer diagnosis at these levels, however, is beyond the scope of this survey and will not be discussed further.

III. PREPROCESSING STEP

The main aim of the preprocessing step is to determine the focal areas in the image. Due to a considerable amount of noise that arises from the staining process, it is usually necessary to reduce the noise prior to the focal area identification. In some studies, focal area identification and noise reduction are carried out at the same time. For instance, in the case of tissue-level feature extraction, the preprocessing step thresholds the image to identify the focal areas by eliminating the noisy regions and those with little content [18], [19]. We will discuss the noise reduction techniques in Section III-A.
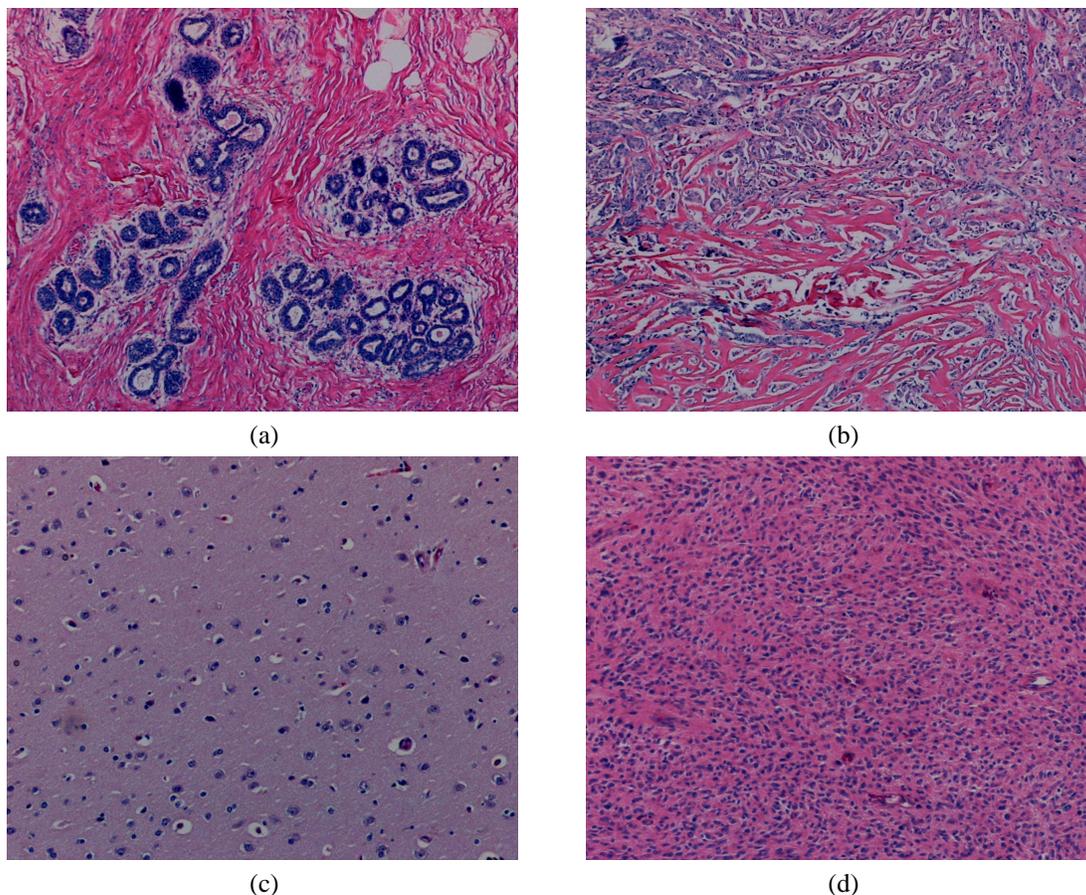
Fig. 1. The sample histopathological images of biopsy samples: (a) a healthy breast tissue, (b) a cancerous breast tissue, (c) a healthy brain tissue, and (d) a cancerous brain tissue. The biopsy samples are stained using hematoxylin-and-eosin technique.

In case of cellular-level feature extraction, noise reduction is followed by the segmentation process to determine the locations of the nuclei/cells[1] in a tissue. Depending on the type of the feature extraction method to be deployed, this may include determining the exact boundary points of cells [20] or determining their coarse locations [21]. In the former case, segmentation requires higher magnification images to resolve the exact details of cells and the success of the next steps becomes more sensitive to the success of the segmentation. In segmentation, one difficulty is the complex nature of image scenes; a typical tissue consists of touching and overlapping cells. Another difficulty comes from the stain related problems including lack of dark separation lines between a nucleus and its surroundings, inhomogeneity of the interior of a nucleus, and occurrence of non-nuclei stain artifacts in a tissue [22]. In Section III-B, we will discuss the segmentation techniques in detail.

A. Noise reduction

The most trivial method for noise reduction is thresholding the pixels of an image, which usually follows background correction and filtering. Background correction standardizes the images by making use of an empty image. Thus, it lessens the effects of different image acquisition conditions such as different lighting conditions [23]. In filtering, the value of a pixel is transformed to a new value which is computed as a function of the values

---

[1]Thereafter, we use "cell segmentation" to refer both segmenting a nucleus and a cell. Similarly, we use "cell" to refer both a nucleus and a cell in that context.
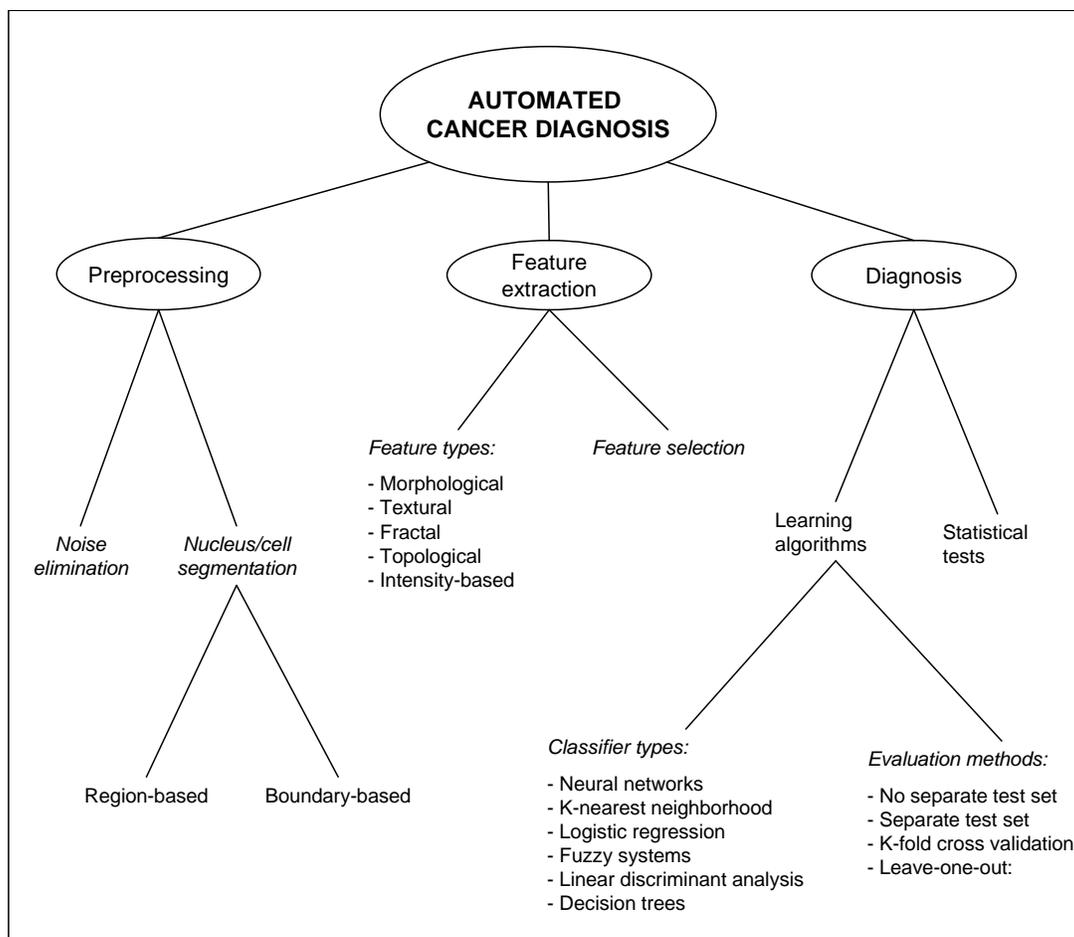
Fig. 2. Overview of the computational steps in automated cancer diagnosis

of pixels located in a selected neighborhood around this particular pixel. Filters especially reduces the random noise and, hence, improves the results of thresholding [24], [25], [26].

In thresholding, the intensity histogram of an image is employed and the pixels under a threshold value are considered to be noise. The threshold values can be determined automatically by using computational methods. For example, the Otsu method determines an optimal threshold which minimizes the within-class variance [27]. For a given image with L different gray levels, the Otsu method computes the within-class variance for a threshold T as follows:

$$\sigma_W^2(T) = \omega_0\sigma_0^2 + \omega_1\sigma_1^2 \tag{1}$$

where σ_0^2 and σ_1^2 are the variances of the pixels below and above the threshold, respectively. With $p(i)$ indicating the probability of the occurrence of gray level i in the image, ω_0 and ω_1 are defined as,

$$\omega_0 = \sum_{i=0}^{T-1} p(i)$$

$$\omega_1 = \sum_{i=T}^{L} p(i)$$

$$\tag{2}$$

It is also possible to apply a threshold function to a group of pixels instead of an individual pixel. By doing so, the regions with little content can be eliminated. For example, Hamilton et al. eliminate a region if the sum of pixel

values in the region falls below a threshold [19]. Similarly, Esgiar et al. consider a region only if at least 10% of its pixels are greater than a threshold [18]. While such techniques are successful to eliminate the regions with little content, they fail eliminating the noisy regions that consist of large stain artifacts.

Another method for noise reduction is to use mathematical morphology, which is a set-theoretic approach that considers pixels in an image as the elements of a set [28]. The basic morphological operators are the erosion and dilation of the set with a structuring element. Dilation of the image A by the structuring element S in $2D$ Euclidean space $\mathcal{E}^2$ is given by

$$A \oplus S = \{x \in \mathcal{E}^2 : x = a + s, \text{ for } a \in A, s \in S\} \tag{3}$$

Erosion is the morphological dual to dilation. Erosion of A by S is given by

$$A \ominus S = \{x \in \mathcal{E}^2 : \text{ for every } s \in S, \exists a \in A \text{ such that } x = a - s\} \tag{4}$$

Consecutive application of these two basic transformations gives two other transformations known as opening and closing. Opening is the erosion of an image followed by the dilation; it breaks narrow isthmuses and eliminates small objects and sharp peaks in the image. On the other hand, closing is the dilation of an image followed by the erosion; it fuses narrow breaks and fills small holes and gaps in the image.

While thresholding and filtering reduce the noise by making use of the pixel intensities, mathematical morphology reduces the noise based on the shape characteristics of the input image, which are characterized by the structuring element. Mathematical morphology can not distinguish the cellular areas and the stain artifacts with similar shapes but different intensity values. Thresholding the image (prior or subsequent to applying the morphological operations) reduces the effect of such an undesired condition.

B. Cell segmentation

There are mainly two different approaches in cell segmentation: the region-based and boundary-based approaches. The region-based approach is based on determining whether a pixel belongs to a cell or not, whereas the boundary-based approach is based on finding the boundary points of a cell. The next subsections explain these approaches in detail.

1) The region-based approach: Thresholding is also used for the purpose of cell segmentation. Although thresholding separates the cells from the background, it does not separate the overlapping cells from each other. The mathematical morphology [28] and watershed algorithms [29] offer solution to this problem. Successive application of opening and closing operators is used to find the separate centers of gravity of the cells. Watershed algorithms are useful to detect the boundary lines between the touching cells [30], [31], [32]. An important limitation of the watershed algorithms is over-segmentation. Applying low-pass filters prior to segmentation and letting water rise only from the marked seeds lessens the over-segmentation problem.

After thresholding the pixels, deleting small objects and large areas and filling small holes improve the segmentation results [31], [33]. For that, it is also possible to apply the morphological operators with different types of structuring elements such as square [34] and octagonal [35] ones. Here, selecting the size of the structuring element is important. The size should be smaller than the minimum size of a cell that will be determined; all the objects smaller than the structuring element are eliminated. On the other hand, it must be large enough to eliminate the noisy areas.

Another region-based approach is making use of a learning algorithm to determine whether a pixel belongs to a cell or not. The learning algorithm employs the features extracted for the pixels. Such features can be based on color information of the pixels or their textural properties. For example, Gunduz et al. automatically cluster the pixels according to their color information using the k-means algorithm [36], where these clusters were previously assigned to either "cell" or "non-cell" class by a human expert [21]. In [37], the textural features of pixels are extracted using autocorrelation function [38]; subsequently, minimum distance classifier is used in learning. In such approaches, the main difficulty is to assign a class of a training sample or a cluster, which requires human control.

2) The boundary-based approach: The trivial method to determine the boundary points is manual segmentation [39], [40]. A number of points are taken from the user and a closed curve is approximated from these points. For example, Einstein et al. approximate a closed curve by joining the arcs characterized by every three successive points [41]. This approach leads to successful segmentation since it requires extensive user interaction. However, its large scale use (for a large number of cells) is not feasible.

Active contour models ("snakes") are proposed to automatically determine the boundary points. A snake is a deformable spline that seeks to minimize an energy function [42]. In cell segmentation, the contour points that yield the minimum energy level form the boundary of a cell. The energy function can be defined to penalize the discontinuity in the curve, discontinuity in the curvature of the snake, and gray-level discontinuity along the snake [43], [20]. In [20], Street et al. define the energy function E over the snake points s,

$$E = \int_s (\alpha E_{cont}(s) + \beta E_{curv}(s) + \gamma E_{image}(s))ds \tag{5}$$

where E_{cont} measures discontinuities in the curve, E_{curv} measures discontinuities in the curvature of the snake, and E_{image} measures the gray-level discontinuities along the snake and the weights of α, β, and γ are empirically derived constants.

In active contour models, the accuracy of a final boundary mainly depends on the initialization of the contour points. However, this task is not easy and often requires user interaction. For example, Wu and Barba define an initial curve as the composite of four quarter ellipses whose end points are specified by the user [44]. For fully-automated segmentation, a preset shape (e.g., an elliptic shape [43]) is imposed for the initial snake.

The boundary-based and region-based approaches have their own advantages and drawbacks. The boundary-based approaches determine more precise locations of a cell compared to the region-based approaches. On the other hand, they necessitate resolving the boundary points and, hence, require higher magnification images. Furthermore, the initialization of the boundary points may be crucial in the success of segmentation, and it might necessitate user interaction.

The segmentation method should be chosen depending on the type of the features to be extracted. For example, in the case of morphological feature extraction, determining the exact locations of cells is more important, and thus, boundary-based approaches are more suitable than the region-based approaches. On the other hand, in the case of topological feature extraction, it might be sufficient to determine the coarse locations of the cellular areas without resolving the exact boundary points. In this case, the region-based approaches are more suitable.

IV. FEATURE EXTRACTION

Automated cancer diagnosis relies on capturing (i) the deviations in the cell structures (*cellular-level*), and (ii) the changes in the cell distribution across the tissue (*tissue-level*). The features are extracted to quantify these changes in a given tissue. To measure the deviations at the cellular-level, morphological, textural, fractal, and/or intensity-based features can be used. In extracting such kinds of features at the cellular-level, the exact locations of the cells should be determined beforehand. On the other hand, to measure the changes at the tissue-level, textural, fractal, and/or topological features can be extracted. The use of these features at the tissue-level does not always necessitate cell segmentation (e.g., textural feature extraction at the tissue level). Even if the segmentation is necessary, determining the coarse locations of the cells is usually sufficient (e.g., topological feature extraction).

In order to extract the features, there are two different types of information available in the image: (i) the intensity values of the pixels and (ii) their spatial interdependency. Although all feature extraction methods use the information on the intensity values, only a few use the spatial dependency between them. Employing only the intensity values, however, results in more sensitivity to the noise that arises from the stain artifacts and the image acquiring conditions.

A. Morphological features

The morphological features provide information about the size and shape of a cell. The size is expressed by the radius, area, and perimeter of the cell. On the other hand, the shape is expressed by the compactness, roundness, smoothness, length of the major and minor axes, symmetry, concavity, and perimeter [20].

Suppose that $\mathcal{S} = \{s_1, ..., s_n\}$ is a set of the boundary points of a segmented cell/nucleus and C is the centroid of these boundary points; a sample of a nucleus with its boundary points and centroid is illustrated in Fig.3(a). The morphological features defined on the set of the boundary points, $\mathcal{S}$, are given as follows:

- *Radius* r is defined as the average length of the radial lines towards every boundary points. Mathematically,

$$r = \frac{\sum_{i=1}^{n} |s_i C|}{n} \tag{6}$$

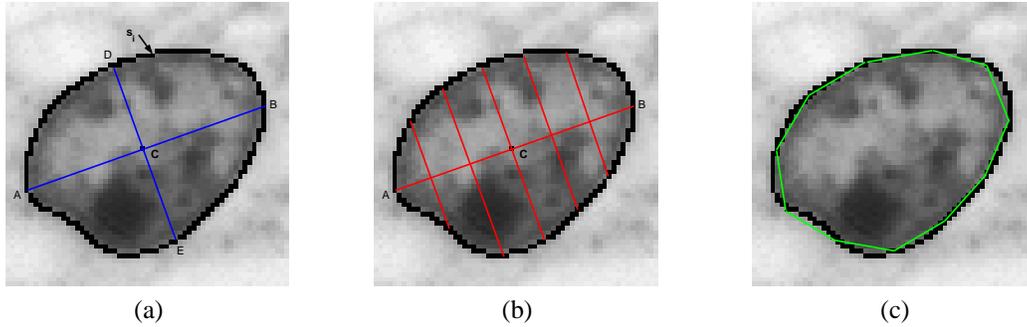- *Area* is the number of pixels within the boundary.

Fig. 3. (a) A sample of a nucleus with its boundary points and centroid, (b) line segments used in symmetry computation (adapted from [20]), and (c) chords used in concavity computation (adapted from [20]).

- *Perimeter* $\mathcal{P}$ is measured as the sum of the distances between every consecutive boundary points. Mathematically,

$$\mathcal{P} = |s_n s_1| + \sum_{i=1}^{n-1} |s_i s_{i+1}| \tag{7}$$

- *Compactness* is the ratio of the square of the perimeter to the area ($compactness = perimeter^2/area$). *Roundness* is defined as $compactness/4\pi$ [37].
- *Smoothness* of the boundary is the sum of the smoothness of each boundary points. For a boundary point s_i, the smoothness is defined as the difference between the radial line $|s_i C|$ and the average length of the radial lines surrounding s_i.
- *Major axis* is the longest chord that goes through the center and *minor axis* is the line that is perpendicular to the major axis and that goes through the center. In Fig.3(a), the major axis $|AB|$ and the minor axis $|DE|$ are illustrated.
- *Symmetry* is quantified by measuring the length difference between the line segments in opposite directions that are defined to a boundary point and that are perpendicular to the major axis. A set of such line segments is shown in Fig.3(b).
- *Concavity* is quantified by drawing chords between non-adjacent boundary points and checking whether or not the boundary points lie inside these chords. A set of such chords is shown in Fig.3(c).

The quantification of these properties enables to differentiate the malignant cells from those of benign and normal. Moreover, the statistics computed on these properties is used to detect cancer in a tissue [31], [34], [45]. For example, Thiran and Macq use the variation in the size and shape of nuclei as an indicator for the existence of cancer [34]. Similarly, Wolberg et al. calculate the average, standard deviation, and maximum [45] while Choi et al. calculate the median and percentile values [31] to define the morphological features.

Besides these features, the ratio of the same feature for different parts of a biological structure is used as another feature. For example, the nuclear area/cytoplasm area ratio [34] and the luminal area/ductal area ratio [46] are such kind of features. Obviously, to extract such features, the other types of biological structures (e.g., lumen, duct, etc.) should be segmented.

B. Textural features

Texture is a connected set of pixels that occurs repeatedly in an image. It provides information about the variation in the intensity of a surface by quantifying properties such as smoothness, coarseness, and regularity. To describe textural features, the two most widely accepted models are those that use the co-occurrence and run-length matrices.

The co-occurrence matrix quantifies the various textural features such as correlation, contrast and angular second moment [47] by making use of the spatial dependency between the gray-level pixel values. The co-occurrence matrix C, computed on a gray-level image P, is defined by a distance d and an angle θ. $C(i, j)$ is the number of times that the gray value i co-occurs with the gray value j in a particular spatial relationship defined by d and θ;

mathematically,

$$C(i,j) = |\{m,n\} : P(m,n) = i \ and \ P(m + d \cdot cos\theta, n + d \cdot sin\theta = j)| \tag{8}$$

The run-length matrix R measures the coarseness of a texture in a specific direction [48]. Given a direction of θ, $R(i,j)$ is the number of gray-level runs with a length of j and a gray value of i, where a gray-level run is defined as a set of consecutive, collinear pixels in the direction of θ that have the same gray value of i. Among the run-length matrix features are short run emphasis, long run emphasis and low gray-level run emphasis .

At the cellular-level, textural features extracted by using the co-occurrence matrix and/or the run-length matrix are computed for an individual cell [37], [20], [49]. It is also possible to extract textural features from the different regions of a cell. For instance, Albregtsen et al. extract features for the periphery and the center of a nucleus [39].

At the tissue level, textural features are computed for an entire image or its sub-images [18], [19], [50]. For a single image/sub-image or a cell, different textural features with different sets of parameters (e.g., with different values of d and θ in the case of the co-occurrence matrix) are first computed. Then, either a subset of features is selected or the same type of features are combined by using statistics measures (e.g., averaging, selecting the maximum, etc.).

For sub-image feature extraction, the feature sets of the sub-images can be combined to reach a single set for the entire image [18]; this set is to be used in the diagnosis of the entire tissue. It is also possible to directly use the features extracted from the sub-images; these features can be used to classify the sub-images. For example, Wiltgen et al. divide the image into squares of sub-images and extract textural features from these squares [51]. Subsequently, they use these textural features to differentiate the squares that include malignant cells from those that consist of their counterparts.

There are also other methods to extract textural features. One of such methods uses wavelets representation which discriminates several spatial orientations in the image [52]. For instance, Weyn et al. compute wavelets by passing the image through a set of iterative low/high-pass filters and use the energies of the filtered images as the textural features [53]. Another method makes use of the complexity curve of a binary image which is obtained by computing the number of black-to-white transitions observed in the image [24].

C. Fractal-based features

The fractal is an object with the self-similarity property, i.e., it appears the same at different magnifications, [54]. The fractal geometry provides information on the regularity and complexity of an object by quantifying its self-similarity level. For that, the fractal dimension of the object is computed; unlike the Euclidean geometry, this dimension can be fractional. Fractal analysis is used to understand different phenomena in different biomedical applications including the cancer diagnosis [55], [56].

For the fractal analysis of a cell or a tissue, the most common feature is the fractal dimension [57], [58], [59], [60], [61]. The fractal dimension d is defined as $d = log(N)/log(p)$, where N is the number of self similar pieces at the magnification scale of p. The fractal dimension can be measured by using several methods including the box-counting, perimeter-stepping, and pixel dilation methods. Among those, the most common and easy-to-implement one is the box-counting method. This method divides the binary image (black and white) into boxes with different sizes (r) and counts the number of black pixels in each box$(N(r))$. Then, it computes the fractal dimension as the slope of $1/r$ vs. $N(r)$ line in log-log scale [62].

In addition to the fractal dimension, the lacunarity is used for the purpose of fractal analysis of a cell or a tissue [63], [40]. The lacunarity quantifies the deviation from homogeneity in the texture. It is measured by computing the size of the holes on the fractal, i.e., counting the white pixels in the fractal image.

D. Topological features

The topological features provide information on the structure of a tissue by quantifying the spatial distribution of its cells. For that, this approach encodes the spatial interdependency of the cells prior to the feature extraction.

The Voronoi diagrams and their Delaunay triangulations can be used to encode the dependency between adjacent cells [31], [26], [64]. On a tissue image, the Voronoi diagram constitutes convex polygons for each cell. For a particular cell, every point in its polygon is closer to itself than to another cell in the tissue. The dual graph of the Voronoi diagram is the Delaunay triangulation. The Voronoi diagram of a sample tissue image and its Delaunay triangulation are shown in Fig.4(b). Among the properties defined for Voronoi diagrams are the area and shape of their polygons. Among the properties defined for Delaunay triangulations are the number of connections of a cell
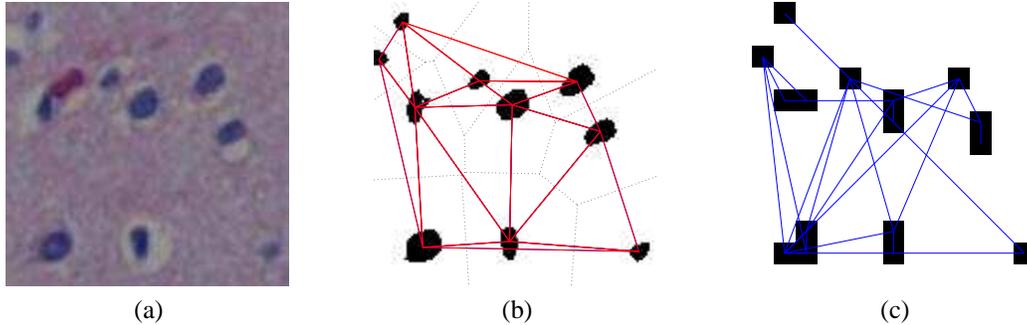
Fig. 4. (a) A sample image, (b) the Voronoi diagram of the image (black dotted lines) and its Delaunay triangulation (red solid lines), and (c) a cell-graph of the image.

and the average length of these connections. It is also possible to find a minimum spanning tree of the Delaunay triangulation and use its properties such as the total tree length to identify the topological features.

To encode the dependency between every pair of cells, Gunduz et al. propose to construct a graph for an image. In this graph, vertices are cell clusters and edges are probabilistically assigned between the vertices; the probability of being an edge between a pair of vertices decays with the increasing Euclidean distance between them [21]. In [21], by making use of the Waxman model [65], the probability of being an edge between the vertices u and v is defined as follows,

$$P(u,v) = \alpha \cdot exp(-d(u,v)/\beta L) \qquad (9)$$

where $d(u,v)$ is the Euclidean distance between the vertices u and v, L is the largest possible Euclidean distance between any of the two nodes, and α and β are the Waxman model parameters that control the number of the links and connectivity of these cell-graphs. Fig.4(c) illustrates a cell-graph of the sample image shown in Fig.4(a). After the graph construction, local and global graph features are extracted. Among the local graph features are the degree, clustering coefficient, and eccentricity of a vertex. In this approach, it is possible to consider the cell clusters as sub-images that consist of a bunch of cells. Therefore, the local metrics correspond to the properties of these sub-images. Among the global ones are the giant connected component ratio, spectral radius, and eigen exponent. Furthermore, the distributions of the local features are used to define the global graph features such as the average degree [66].

E. Intensity-based features

The intensity-based features are extracted from the gray-level or color histogram of the image. This type of features does not provide any information about the spatial distribution of the pixels. The intensity histogram in a cell is employed to define features. For example, Weyn et al. define the optical density of a pixel by making use of its gray-level value [53]; for the background corrected images, they convert the gray value of each pixel in a nucleus to its corresponding optical density (OD) as follows:

$$OD = log_{10}\left(\frac{gray\ value\ of\ the\ background}{gray\ value\ of\ the\ pixel}\right) \qquad (10)$$

Subsequently, they compute the sum and the mean of the optical densities of the pixels located in a nucleus to define its intensity-based features.

The features that are extracted from the color histograms are computed by making use of either the pixel values in a single color channel or the relationship between the color values in different channels [67], [51], [68]. For example, Schnorrenberg et al. use the YIQ color system and compute their intensity-based features as the average of the pixel values in each color channel (i.e., in Y, I, and Q channels) [67]. Similarly, Wiltgen et al. compute the average, standard deviation, skewness, and kurtosis of the pixel values in each RGB color channel [51]. In [68], in addition to the features extracted in each RGB color channel, Zhou et al. use the difference between the red and blue components and proportion of the blue component as the intensity-based features.

F. Feature selection

Although it is possible to extract a large set of features, only a small subset of them is used in the classification due to the curse of dimensionality. The curse of dimensionality states that as the dimensionality increases, the amount of required training data increases exponentially [69]. Moreover, there might be a strong correlation between different features. Therefore, there is an incentive to reduce the size of the feature set.

The feature set reduction is typically achieved either (i) by transforming the features into a new space and designing new features as the combinations of the original ones (e.g., by using principal component analysis [70]) or (ii) by selecting a subset of the features as to maximize an objective function. One way to define this objective function is to use the predictive accuracy of each subset. For example, Baheerathan et al. estimate the classification accuracy of a subset using leave-one-out method on the training set [24]. Another way of defining an objective function for the evaluation of the subsets is to use their information content. The information content is measured by making use of the interclass distance measures such as Euclidean and Mahalanobis measures (e.g., [39]), or the statistical dependency of the features (e.g., [19]). In the latter case, the features that show significant differences between different classes are selected to be included in the subset.

In feature selection, a subset of features can be selected either using exhaustive search or sequential search algorithm. The best possible subset is determined by an exhaustive search, i.e., by trying all possible feature subsets. However, the number of all possible subsets grows exponentially with the number of features, making such an exhaustive search impractical for even a moderate number. To overcome this difficulty, the contents and the size of the subsets can be restricted [63], [37]. Sequential search algorithms add or remove features sequentially. Representative examples of this kind are sequential forward selection, sequential backward selection, plus-l minus-r selection, and sequential floating selection. For example, to reduce the size of a feature set that is to be used in automated cancer diagnosis, plus-2 minus-1 selection and sequential floating selection techniques are used in [35] and [53], respectively. Sequential forward selection starts with an empty set of features and incrementally expands the subset by adding a feature. These additional features are selected so that the subsequent subsets lead to the best objective function. The search terminates if no additional feature yields a better objective function. Likewise, sequential backward search starts with a complete set of features and incrementally removes the useless features one at a time. Both of these search methods are examples of greedy algorithms and they do not guarantee the best possible solution. Besides, both of these algorithms select the features assuming that they are independent. However, there might be a feature(s) that does not yield a good result when it is used alone, and the objective function gives a better result when this feature is used in conjunction with other features. Sequential forward and backward selections cannot capture such complimentary features since they work incrementally. To circumvent this problem, plus-l minus-r selection is proposed. This algorithm starts with an empty set and it adds l features and removes r of them in each step. The main difficulty in this method is, however, to select the optimal values of l and r. Rather than using fixed l and r, sequential floating selection allows using different values of l and r in each step. All these selection algorithms have possibility to trap in local minima and, hence, to find suboptimal solutions.

V. Diagnosis

After determining an appropriate set of features, the next step is to distinguish the malignant structures from their counterparts. In this step, a cell or a tissue is assigned to one of the classes of cancerous, benign, or healthy. As a part of diagnosis, it is also possible to classify the malignancy level of the tissues (i.e., *grading*). In this case, the classes are the possible grades of the cancer of interest.

For diagnosis, one group of studies employs a statistical test on the features [57], [63], [26], [59], [60], [61]. All these studies examine whether or not a significant difference exists in the value of at least one feature of interest for different classes. However, for the histopathological images, the results of statistical tests should be interpreted with an extra caution for the following reason. Statistical tests assume independent samples and lead to conclusions accordingly. On the other hand, the data set consists of different tissue images taken from the same patient, which are not independent, and this may cause misleading and confusing results.

Another group of studies uses machine learning algorithms to learn (from data) how to distinguish the different classes from each other. Among those algorithms are the neural networks, k-nearest neighborhood algorithm, logistic regression method, fuzzy systems, linear discriminate functions, and decision trees. The list of the studies that use these algorithms is given in Fig.5

Neural networks:
> Choi et al. (1997), Demir et al. (2004), Einstein et al. (1998), Gunduz et al. (2004)
> Spyridonos et al. (2002), Tasoulis et al. (2003), Zhou et al. (2002)

K-nearest neighborhood:
> Esgiar et al. (2002), Schnorrenberg et al. (1996), Weyn et al. (1999)

Logistic regression:
> Einstein et al. (1998), Wolberg et al. (1995)

Fuzzy systems:
> Blekas et al. (1998)

Linear discriminant analysis:
> Esgiar et al. (1998), Hamilton et al. (1997), Smolle (2000)

Decision trees:
> Wiltgen et al. (2003)

Fig. 5.   The list of classifiers that are used by different studies.

No separate evaluation set:
> Anderson et al. (1997), Smolle (2000), Thiran and Macq (1996)

Separate training and test sets:
> Blekas et al. (1998), Choi et al. (1997),Demir et al. (2004), Diamond et al. (2004)
> Esgiar et al. (1998), Gunduz et al. (2004), Pena-Reyes and Sipper (1999),
> Spyridonos et al. (2002), Wiltgen et al. (2003)

K-fold cross-validation:
> Street et al. (1993), Tasoulis et al. (2003), Wolberg et al. (1995), Zhou et al. (2002)

Leave-one-out:
> Albregtsen et al. (2000), Einstein et al. (1998), Schnorrenberg et al. (1996),
> Spyridonos et al. (2001), Tasoulis et al. (2003), Weyn et al. (1999)

Fig. 6.   The list of techniques for classifier evaluation that are used by different studies.

A. Evaluation of the classification system

In general, a classification system should have two stages: (i) training the classifier to learn the system parameters and (ii) testing the system to evaluate the success of the classifier. Since there is a limited amount of available data in training, it is very important to test the system with extra data. However, it is an issue how to use this limited amount of data in both training and testing. More data used in training lead to better system designs, whereas more data used in testing lead to more reliable evaluation of the system.

Evaluating the system according to the success obtained on the training set brings the risk of memorization of data and obtaining over-optimistic error rates. To circumvent the memorization problem, the system should be evaluated on a separate data set that is not used in training the system. For that, one approach is to split the data into two disjoint sets and and use these sets to train and test the system. In the case that it is not feasible to use a significant portion of the data as the testing set, k-fold cross-validation can be used. This approach randomly partitions the data set into k groups. Then, it uses $k-1$ groups to train the system and uses the remaining group to estimate an error rate. This procedure is repeated k times such that each group is used for testing the system. Leave-one-out is a special case of the k-fold cross validation where k is selected to be the size of the data; therefore only a single sample is used to estimate the error rate in each step. In Fig.6, we provide the list of the system evaluation methods used in different studies of automated cancer diagnosis.

Since the testing stage should measure how well the system will work on unknown samples in the future, the test set should also consist of the samples that are independent from those used in the training. However, in the case of k-fold cross-validation, random partitioning may result in using the test sets that do not include such independent

samples. Therefore, over-optimistic results may be obtained. An illustrative example that shows the effects of each approach on the system success can be found in [74]. In this example, by using the same data, 95% accuracy is achieved when the entire data is used in both training and testing; 87% testing accuracy is obtained in the case of k-fold cross-validation; but, only 60% testing accuracy is obtained when separate training and test sets are used.

For a given sample, a diagnostic system can lead to one of the four possible categories:

- True positive (TP): the diagnostic system yields positive test result for the sample and the sample actually has the disease,
- False positive (FP): the diagnostic system yields positive test result for the sample but the sample does not actually have the disease,
- True negative (TN): the diagnostic system yields negative test result for the sample and the sample does not actually have the disease,
- False negative (FN): the diagnostic system yields negative test result for the sample but the sample actually has the disease,

By using the number of samples that fall into these categories, *sensitivity* and *specificity* are defined to assess the success of the diagnostic system. Sensitivity is the probability of a positive diagnosis test among persons that have the disease and it is defined as,

$$sensitivity = \frac{number\ of\ TP}{number\ of\ TP + number\ of\ FN} \tag{11}$$

Specificity is the probability of a negative diagnosis test among persons that do not have the disease and it is defined as,

$$specificity = \frac{number\ of\ TN}{number\ of\ TN + number\ of\ FP} \tag{12}$$

The aim is to design diagnostic systems that lead to both high sensitivity and high specificity. The numerical measure in terms of both sensitivity and specificity is obtained by calculating the area under a Receiver Operating Characteristic (ROC) curve [75]. ROC curves plot the sensitivity versus one minus the specificity; the greater the area under these curves, the better the systems.

B. Comparison of the reliability of the studies

The previous demonstrations of the reported studies have led to very promising results. However, we believe that it is not straightforward to evaluate and numerically compare these studies solely based on their reported results for a number of reasons: First, these studies use different evaluation methods to estimate their system performances. This, however, leads to different evaluation results as we discussed in Section V-A.

Second, different studies work on different types of cancer. They extract different types of features to represent the cells and the tissues of those cancer types (Fig.7). Since the cell and tissue structures may be different for different organs; a method that works on one cancer type may not work equally well on another. Moreover, some of these studies focus on distinguishing these different cancerous structures from their non-cancerous counterparts while some of them aim at classifying them into different grades.

Third, the experimental methods may vary: (i) different lightening conditions and magnifications are used and/or (ii) different numbers of samples are collected from different numbers of patients.

Fourth, the histological grading and diagnosis are subjective due to variations in expert evaluations and may result in poor reproducibility, i.e., may result in poor agreement among the repeated measurements. Reproducibility is measured by using the variation in the classification of the same image by the same pathologist at different times (*intra-observer*) and the variation in the classification of the same image by different pathologists (*inter-observer*).

For numerical comparison of the studies, it is important for us to constitute benchmark data sets. These data sets should consist of samples that are taken from a large number of patients and examined by different pathologists. Such an effort makes it possible to numerically compare the results obtained by different studies and, in turn, to identify the distinguishing features.

VI. Conclusion

In this review, we investigate the computational steps to automatically diagnose cancer by making use of histopathological images. There are primarily three steps: preprocessing, feature extraction, and diagnosis.

Bladder	*Morphological, textural,fractal-based, topological*
	Choi et al. (1997), Rajesh and Dey (2003), Spyridonos et al. (2001, 2002), Tasoulis et al. (2003)
Brain	*Textural, topological*
	Demir et al. (2004), Gunduz et al. (2004), Spyridonos et al. (2002)
Breast	*Morphological, textural, fractal-based, intensity-based*
	Anderson et al (1997), Dey and Mohanty (2003), Einstein et al. (1998) Schnorrenberg et al. (1996), Street et al. (1993), Weyn et al. (1998), Wolberg et al. (1995)
Cervical	*Textural, fractal-based, topological*
	Sedivy et al. (1999), Keenan et al. (2000), Walker et al. (1994)
Colorectal	*Textural, fractal-based, intensity-based*
	Esgiar et al. (1998, 2000), Hamilton et al. (1997)
Gastric	*Morphological, textural, intensity-based*
	Blekas et al. (1998)
Liver	*Textural, fractal-based*
	Albregtsen et al. (2000), Baheerathan (1999), Kerenji et al. (2000), Nielsen et al. (1999)
Lung	*Morphological, intensity-based*
	Thiran and Macq (1996),Zhou et al. (2002)
Mesothelioma	*Morphological, textural, topological, intensity-based*
	Weyn et al. (1999)
Prostate	*Textural*
	Diamond et al. (2004)
Skin	*Textural, intensity-based*
	Smolle (2000), Wiltgen et al. (2003)

Fig. 7. The type of features that are used in the diagnosis of different type of cancers

In the preprocessing step, the focal areas determined. This comprises to eliminate the noise and improve the image quality. Although, different techniques such as filtering and mathematical transformations show different levels of success in noise reduction, the problem of noise elimination has not been entirely solved. In the case of diagnosis at the cellular-level, this step also includes cell segmentation. The segmentation is achieved either by finding their boundary points or by differentiating the pixels of cells. The important challenge in segmentation is to separate the touching cells from each other.

Next is the feature extraction step. This step quantifies the properties of the biological structures of interest, extracting features either at the cellular-level or at the tissue-level. While cellular-level features focus on capturing the deviations in the cell structures, tissue-level features focus on capturing the changes in the cell distribution across the tissue. The features can be grouped into five based on the information they provide; the list of these features are given below. The important challenge in this step is to find the most proper cell/tissue representation(s) and select a subset of the features extracted from this representation(s).

- The **morphological** features provide information about the size and the shape of a nucleus/cell.
- The **textural** features provide information about the variation in the intensity of a surface and quantify properties such as smoothness, coarseness, and regularity.
- The **fractal-based** features provide information on the regularity and complexity of a cell/tissue by quantifying its self-similarity level.
- The **topological** features provide information on the cellular structure of a tissue by quantifying the spatial distribution of its cells.
- The **intensity-based** features provide information on the intensity (gray-level or color) histogram of the pixels located in a nucleus/cell.

After feature extraction, the next step is to distinguish benign and malignant structures as well as to classify the malignancy level. The important challenge is to evaluate the reliability of the designed diagnostic systems because of a limited amount of available data. This limited amount of data should be used both to learn the system parameters and to estimate the system reliability. The improper use of an evaluation method, however, may lead to biased and misleading results. For example, if the system performance evaluation is not done by using independent samples, overoptimistic results might be obtained. The numerical comparison of different studies is important to identify and avoid such biased and misleading results. For that, it is essential to form a benchmark of data sets that include biopsy samples taken from a large number of patients and examined by different pathologists.

REFERENCES

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

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

[3] A. Ben-Dor, L. Bruhn, N. Friedman, I. Nachman, M. Schummer, Z. Yakhini, Tissue classification with gene expression profiles, J. Comput. Biol. 7(2000) 559-583.

[4] T.S. Furey, N. Cristianini, N. Duffy, D.W. Bednarski, M. Schummer, D. Haussler, Support vector machine classification and validation of cancer tissue samples using microarray expression data, Bioinformatics 16(2000) 906-914.

[5] I. Guyon, J. Weston, S. Barnhill, V. Vapnik, Gene selection for cancer classification using support vector machines, Mach. Learn. 46(2002) 389-422.

[6] J. Khan, J.S. Wei, M. Ringner, L.H. Saal, M. Ladanyi, F. Westermann, F. Berthold, M. Schwab, C.R. Antonescu, C. Peterson, P.S. Meltzer, Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks, Nat. Med. 7(2001) 673-679.

[7] C.H. Yeang, S. Ramaswamy, P. Tamayo, S. Mukherjee, R.M. Rifkin, M. Angelo, M. Reich, E. Lander, J. Mesirov, T. Golub, Molecular classification of multiple tumor types, Bioinformatics 17(2001) 316-322.

[8] B.L. Adam, A. Vlahou, O.J. Semmes, G. L. Wright, Jr., Proteomic approaches to biomarker discovery in prostate and bladder cancers, Proteomics 1(2001) 1264-1270.

[9] A.A. Alaiya, B. Franzen, G. Auer, S. Linder, Cancer proteomics: from identification of novel markers to creation of artificial learning models for tumor classification, Electrophoresis 21(2000) 1210-1217.

[10] G. Ball, S. Mian, F. Holding, R.O. Allibone, J. Lowe, S. Ali, G. Li, S. McCardle, I.O. Ellis, C. Creaser, and R.C. Rees, An integrated approach utilizing artificial neural networks and SELDI mass spectrometry for the classification of human tumours and rapid identification of potential biomarkers, Bioinformatics 18(2002) 395-404.

[11] J. Li, Z. Zhang, J. Rosenzweig, Y.Y. Wang, D.W. Chan, Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer, Clin. Chem. 48(2002) 1296-1304.

[12] B. Wu, T. Abbott, D. Fishman, W. McMurray, G. Mor, K. Stone, D. Ward, K. Williams, H. Zhao, Comparison of statistical methods for classification of ovarian cancer using mass spectrometry data, Bioinformatics 19(2003) 1636-1643.

[13] M. Kallergi, Computer-aided diagnosis of mammographic microcalcification clusters, Med. Phys. 31(2004) 314-326.

[14] B. Sahiner, H.-P. Chan, N. Petrick, D. Wei, M.A. Helvie, D.D. Adler, M.M. Goodsitt, Classification of mass and normal breast tissue: A convolution neural network classifier with spatial domain and texture images, IEEE Trans. Med. Imaging 15(1996) 598-610.

[15] L. Shen, R.M. Rangayyan, J.E.L. Desautels, Application of shape analysis to mammographic calcifications, IEEE Trans. Med. Imaging 13(1994) 263-274.

[16] A.A. Tzacheva, K. Najarian, J.P. Brockway, Breast cancer detection in gadolinium-enhanced MR images by static region descriptors and neural networks, J. Magn. Reson. Imaging 17(2003) 337-342.

[17] U. Wedegartner, U. Bick, K. Wortler, E. Rummeny, G. Bongartz, Differentiation between benign and malignant findings on MR-mammography: usefulness of morphological criteria, Eur. Radiol. 11(2001) 1645-1650.

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

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

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

[21] C. Gunduz, B. Yener, S.H. Gultekin, The cell graphs of cancer, Bioinformatics 20(2004) i145-i151.

[22] J. Gil, H. Wu, B.Y. Wang, Image analysis and morphometry in the diagnosis of breast cancer, Microsc. Res. Techniq. 59(2002) 109-118.

[23] K. Rodenacker, E. Bengtsson, A feature set for cytometry on digitized microscopic images, Anal. Cell. Pathol. 25(2003) 1-36.

[24] S. Baheerathan, F. Albregtsen, H.E. Danielsen, New texture features based on complexity curve, Pattern Recogn. 32(1999) 605-618.

[25] K. Blekas, A. Stafylopatis, D. Kontoravdis, A. Likas, P. Karakitsos, Cytological diagnosis based on fuzzy neural networks, J. Intelligent Systems. 8(1998) 55-79.

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

[27] N. Otsu, A threshold selection method from gray level histograms, IEEE T. Syst. Man Cy. 9(1979) 62-66.

[28] J. Serra, Image Analysis and Mathematical Morphology (Academic, London, 1982).

[29] L. Vincent, P. Soille, Watersheds in digital spaces: An efficient algorithm based on immersion simulations, IEEE T. Pattern Anal. 13(1991) 583-598.

[30] P. Bamford, B.C. Lovell, A water immersion algorithm for cytological image segmentation, Proc. APRS Image Segmentation Workshop (1996) 75-79, Sydney, Australia.

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

[32] N. Malpica, C. Ortiz de Solorzano, J.J. Vaquero, A.Santos, I. Vallcorba, J.M. Garcia-Sagredo, F. del Pozo, Applying watershed algorithms to the segmentation of clustered nuclei, Cytometry 28(1997) 289-297.

[33] B. Weyn, G. Van de Wouwer, M. Koprowski, A. Van Daele, K. Dhaene, P. Scheunders, W. Jacob, E. Van Marck, Value of morphometry, texture analysis, densitometry, and histometry in the differential diagnosis and prognosis of malignant mesothelioma, J. Pathol. 189(1999) 581-589.

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

[35] R.F. Walker, P.T. Jackway, B. Lovell, Classification of cervical cell nuclei using morphological segmentation and textural feature extraction, Proc of the 2nd Australian and New Zealand Conference on Intelligent Information Systems (1994) 297-301.

[36] J.A. Hartigan, M.A. Wong, A k-means clustering algorithm, Appl. Stat. 28(1979) 100-108.

[37] P. Spyridonos, P. Ravazoula, D. Cavouras, K. Berberidis, G. Nikiforidis, Computer-based grading of haematoxylin-eosin stained tissue sections of urinary bladder carcinomas, Med. Inform. Internet Med. 26(2001) 179-90.

[38] O. Faugeras, W. Pratt, Decorrelation methods of texture feature extraction, IEEE T. Pattern Anal. 14(1980) 323-332.

[39] F. Albregtsen, B. Nielsen, H.E. Danielsen, Adaptive gray level run length features from class distance matrices, Int. Conf. on Pattern Recognition 3(2000) 3746-3749.

[40] B. Nielsen, F. Albregtsen, H.E. Danielsen, The use of fractal features from the periphery of cell nuclei as a classification tool, Anal. Cell. Pathol. 19(1999) 21-37.

[41] A.J. Einstein, J. Gil, S. Wallenstein, C.A. Bodian, M. Sanchez, D.E. Burstein, H.S. Wu, Z. Liu, Reproducibility and accuracy of interactive segmentation procedures for image analysis in cytology, J. Microsc. 188(1997) 136-148.

[42] M. Kass, A. Witkin, D. Terzopoulos, Snakes: Active contour models. International Journal of Computer Vision 1(1988) 321-331.

[43] K.-M. Lee, W.N. Street, A fast and robust approach for automated segmentation of breast cancer nuclei, Proc of the Second IASTED International Conference on Computer Graphics and Imaging (1999) 42-47, Palm Springs, CA.

[44] H.-S. Wu, J. Barba, An efficient semi-automatic algorithm for cell contour extraction, J. Microsc. 179(1995) 270-276.

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

[46] N.H. Anderson, P.W. Hamilton, P.H. Bartels, D. Thompson, R. Montironi, J.M. Sloan, Computerized scene segmentation for the discrimination of architectural features in ductal proliferative lesions of the breast, J. Pathol. 181(1997) 374-380.

[47] R.M. Haralick, Statistical and structural approaches to texture, Proc. of IEEE. 67(1979) 786-804.

[48] M.M. Galloway, Texture analysis using gray level run lengths, Computer Graphics and Image Processing 4(1975) 172-179.

[49] D.K. Tasoulis, P. Spyridonos, N.G. Pavlidis, D. Cavouras, P. Ravazoula, G. Nikiforidis, M.N. Vrahatis, Urinary bladder tumor grade diagnosis using on-line trained neural networks, Proc. Knowl. Based Intell. Inform. Eng. Syst. (2003) 199-206.

[50] J. Smolle, Computer recognition of skin structures using discriminant and cluster analysis, Skin Res. Technol. 6(2000) 58-63.

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

[52] S.G. Mallat, A theory for multiresolution signal decomposition: The wavelet representation, IEEE T. Pattern Anal. 11(1989) 674-693.

[53] B. Weyn, G. Van de Wouwer, A. Van Daele, P. Scheunders, D. Van Dyck, E. Van Marck, W. Jacob, Automated breast tumor diagnosis and grading based on wavelet chromatin texture description, Cytometry 33(1998) 32-40.

[54] B.B. Mandelbrot, The Fractal Geometry of Nature, W. H. Freeman, New York, 1982.

[55] S.C. Cross, Fractals in Pathology, J. Pathol. 182(1997) 1-8.

[56] O. Heymans, J. Fissette, P. Vico, S. Blacher, D. Masset, F. Brouers, Is fractal geometry useful in medicine and biomedical sciences? Med. Hypotheses. 54(2000) 360-366.

[57] P. Dey, S.K. Mohanty, Fractal dimensions of breast lesions on cytology smears, Diagn. Cytopathol. 29(2003) 85-86.

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

[59] A.S. Kerenji, Z. Bozovic, M.M. Tasic, Z.M. Budimlija, I.A. Klem, A.F. Polzovic, Fractal dimension of hepatocytes' nuclei in normal liver vs hepatocellular carcinoma (hcc) in human subjects - preliminary results, Archive of Oncology 8(2000) 47-50.

[60] L. Rajesh, P. Dey, Fractal dimensions in urine smears: a comparison between benign and malignant cells, Anal. Quant. Cytol. Histol. 25(2003) 181-182.

[61] R. Sedivy, C. Windischberger, K. Svozil, E. Moser, G. Breitenecker, Fractal analysis: an objective method for identifying atypical nuclei in dysplastic lesions of the cervix uteri, Gynecol. Oncol. 75(1999) 78-83.

[62] H. Peitgen, H. Jurgens, D. Saupe, Fractals for the Classroom: Part One, Introduction to Fractals and Chaos, Springer-Verlag, New York, 1991.

[63] A. J. Einstein, H.S. Wu, M. Sanchez, J. Gil, Fractal characterization of chromatin appearance for diagnosis in breast cytology, J. Pathol. 185(1998) 366-381.

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

[65] B.M. Waxman, Routing of multipoint connections, IEEE J. Sel. Area Commun., 6(1988) 1617-1622.

[66] C. Demir, S.H. Gultekin, B.Yener, Learning the topological properties of brain tumors, Rensselaer Polytechnic Institute Technical Report TR-04-14, Troy, 2004.

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

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

[69] R. Bellman, Adaptive Control Processes: A Guided Tour, Princeton University Press, Princeton, New Jersey, 1961.

[70] I.T. Jolliffe, Principal Component Analysis, Springer-Verlag, New York, 1986.

[71] J. Diamond, N.H. Anderson, P.H. Bartels, R. Montironi, P.W. Hamilton, The use of morphological characteristics and texture analysis in the identification of tissue composition in prostatic neoplasia, Hum. Pathol. 35(2004) 1121-1131.

[72] C.A. Pena-Reyes, M. Sipper, A fuzzy genetic approach to breast cancer diagnosis, Artif. Intell. Med. 17(1999) 131-155.

[73] P. Spyridonos, D. Glotsos, D. Cavouras, P. Ravazoula, V. Zolota, G. Nikiforidis, Pattern recognition based segmentation method of cell nuclei in tissue section analysis, Proc of 14th IEEE International Conference on Digital Signal Processing (2002) 1121-1124, Santorini, Greece.

[74] H. Schulerud, G.B. Kristensen, K. Liestol, L. Vlatkovic, A. Reith, F. Albregtsen, H.E. Danielsen, A review of caveats in statistical nuclear image analysis, Anal. Cell. Pathol. 16(1998) 63-82.

[75] C.E.Metz, Basic principles of ROC analysis, Seminars Nucl. Med., 8(1978) 283298.