the diagnostic pathology journal DIAGNOSTIC PATHOLOGY

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

A review of graph-based methods for image analysis in digital histopathology

Harshita Sharma[1*], Norman Zerbe[2], Sebastian Lohmann[2], Klaus Kayser[2], Olaf Hellwich[1] and Peter Hufnagl[2]

*Corresponding author, e-mail: harshita.sharma@cv.tu-berlin.de

[1]Computer Vision and Remote Sensing, Technical University Berlin, Berlin, Germany

[2]Department of Digital Pathology and IT, Institute of Pathology, Charité - Universitätsmedizin Berlin, Berlin, Germany

Abstract:

Digital image analysis of histological datasets is a currently expanding field of research. With different stains, magnifications and types of tissues, histological images are inherently complex in nature and contain a wide variety of visual information. Several image analysis techniques are being explored in this direction. However, graph-based methods have recently gained immense popularity, as these methods can effectively describe tissue architecture and provide adequate numeric information for subsequent computer-based analysis. Graphs have the ability to represent spatial arrangements and neighborhood relationships of different tissue components, which are essential characteristics observed visually by pathologists during investigation of specimens. In this paper, we present a comprehensive review of the graph-based methods explored so far in digital histopathology. We also discuss the current limitations and suggest future directions in graph-based tissue image analysis.

Keywords:

Digital histopathology, graph-based methods, whole slide images, medical image analysis, image understanding, tissue architecture, spatial arrangement.

1 Introduction

Histological images are the magnified images of the tissues of an organism. They are a special kind of medical images, which contain various complex structures and information pertaining to different biological conditions of the organism. Pathology is

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

the branch of medicine dealing with causal study of diseases. The term 'histopathology' is derived from the combined field of histology and pathology. For humans, histopathological information is sought by medical specialists for determining the presence or extent of abnormal conditions, especially tumors and cancers. Generally, a pathological test is recommended after observing physiological symptoms of patients. There are diverse ways to acquire histopathological specimens from the body, such as biopsy and surgical resection [1]. The tissue specimens are mounted on glass slides and treated with detailed preparation procedures before becoming ready to be viewed by observers [2]. The specimens are stained with special chemicals following a staining process in order to improve contrast and visibility of specific tissue structures as required for observation. Some commonly used stains for this purpose are Haematoxylin and Eosin (H&E), Feulgen, Schiff, Wright, Ki-67 protein and Toluidine blue [3]. Conventional viewing and observation of tissues is performed using various microscopy techniques [4].

Digital pathology involves the use of digitized histological images for observation and analysis. Glass slides are scanned using sophisticated equipment like advanced microscopes with cameras or whole slide scanners to obtain high resolution digital images called whole slide images (WSI), which can be viewed on computer screens using image viewers, stored and archived into databases for future reference. Whole slide images usually provide a comprehensive view of the tissue appearance and can be used for determining underlying disease condition(s). Several benefits of digital pathology have been identified, some of the most apparent being better viewing and handling of slides, parallel viewing at distinct locations (telepathology), parallel viewing of different stains and positions, reduction in glass archives and glass transportation, and simplicity to handle annotations. These benefits are explained in detail in [5].

Computer-based analysis of histological images is a prospective challenge in the field of pathology and is being actively explored. Analysis techniques are being developed to automate the process of extracting useful information from tissue images in order to provide several benefits in clinical routine practice and research. Maximum work in this direction has been aimed towards developing diagnosis-related classification techniques in order to classify tissue regions into healthy or cancerous, benign or malignant or into several grades or types. But their popularity is limited mainly due to the error rates involved, so they are considered less reliable compared to pathologists' verdicts. Pathologists' experience has been acquired through a specialized long-term training process, and they expertly diagnose diverse cases on a day-to-day basis. Along with assessment of visual appearance of tissues, pathologists also consider additional clinical information and underlying causes like patient history and etiological agents, which

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

makes the decision-making process more complicated and difficult to formalize. Also, the image analysis software developed till date is specific to certain pathological disorders and doesn't consider exceptional cases. So the extent of usability of image analysis applications for routine work still remains unclear. However, most visual observations are currently subjective and vary from person to person and time to time, causing inter- and intra-observer variability. Hence, there is a need for developing more reliable computer-aided retrieval and classification methods, which can act as assisting tools, saving time and efforts, and reducing subjectivity. Pathologists' experience still remains most valuable and is required for development of such analysis software as well as for validation of obtained results. Some less critical applications can also be developed to improve the current state of technology in this field, for example, smart archiving and content-based image retrieval of interest regions from WSI specimens, bio-banking of large databases, providing second opinions to pathologists, marker quantification, assistance in detecting malignant changes over time and for educational purpose.

Histological images have typical characteristics which set them apart from other images. These include high resolutions, complex appearances, diverse magnifications, specific stains and the corresponding differences in semantic interpretations. The acquisition process is not ideal in most cases, hence the quality of acquired digital images is also a current question of discussion among pathologists worldwide. Considering the complexity involved into evaluating such image data, different analysis methods have been developed. We believe that one of the most promising approaches for histological image analysis is the use of graph-based techniques. Graphs are effective and flexible representation tools and have lately been of major interest to the computer vision and image analysis fields due to their expressive ability to model topological and relational information between image components. Moreover, histological image data is visually observed and interpreted by pathologists by considering the morphological changes, neighborhood relationships and spatial arrangements between tissue structures, and graphs have proved to be able to quantitatively represent these visual information cues in histological images. Several authors have also used the term *syntactic structure analysis* for describing a subset of graph-theoretic methods to quantitatively analyze tissue architecture, some of which we describe later in the paper.

The organization of the paper is as follows. Section 2 provides a mathematical overview of the graph-based techniques used in this field and most often used graph metrics. Section 3 describes the various graph-theoretic methods employed so far in the direction of histological image analysis. Section 4 concludes the paper and mentions future directions to be explored using graph-based methods in histology.

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

2 Overview of Graph-based Methods

A graph is a data structure consisting of a finite set of points called vertices (or nodes) and a set of edges (or arcs) which link the vertices with each other based on predefined criteria. Mathematically, a graph is a tuple $G(V,E)$ containing the set of vertices V and edges $E \subseteq V \times V$. An edge in E connects two vertices in V. A graph is planar if it can be embedded in the plane, in other words, it can be depicted on the plane such that its edges do not cross one another except only at their endpoints, otherwise it is a non-planar graph. A graph is undirected if the edges have no orientation. A directed graph (digraph) is the one where each edge has a direction associated with it, connecting an ordered pair of vertices. Other definitions associated with graphs are explained in [6].

2.1 Voronoi Diagram, Delaunay Triangulation and Related Graphs

The first formal definitions of Dirichlet tessellation and Voronoi diagram were proposed by Dirichlet [7] and Voronoi [8] respectively. Let $V = \{v_1, v_2 .... v_n\}$ be a set of n points (or vertices) in a plane and d denotes distance between two given points. It is assumed that any three points are non-collinear and any four points are not co-circular. We define the planar graph representations Voronoi diagram, Delaunay triangulation and related graphs using these definitions and assumptions. These graphs are included in the larger group called proximity graphs, also called neighborhood graphs, in which two vertices are linked by an edge if and only specific geometric requirements are satisfied by the vertices [9]. Proximity graphs are defined with reference to various metrics, however Euclidean metric is the most commonly used.

2.1.1 Voronoi Diagram

The Voronoi diagram has been described in detail in [10]. We summarize it as follows. Let two points (called sites) v_i and v_j be connected by the line segment $\overline{v_i v_j}$ and its perpendicular bisector $B(v_i, v_j)$ divides the plane between the two sites into two half

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

planes. The half plane of site v_i with respect to v_j is denoted by $H(v_i, v_j)$ and contains the set of points in the plane given by:

$$H(v_i, v_j) = \{x \mid d(v_i, x) < d(v_j, x)\} \quad (1)$$

The Voronoi region (or Voronoi polygon) $V_P(v_i)$ for the site v_i is given by equation 2. $V_P(v_i)$ is a convex polygon which may be unbounded.

$$V_P(v_i) = \cap_{j=1, j\neq i}^{n} H(v_i, v_j) \quad (2)$$

Hence, $V_P(v_i)$ contains all such points in the plane closer to v_i than to any other site v_j. The Voronoi diagram of the set of V sites is obtained by using equation 3.

$$VD(V) = \cup_{i=1}^{n} V_P(v_i) \quad (3)$$

Voronoi diagrams perform a nearest site proximity partitioning of the plane. This is illustrated in Figure 1. Using planarity and Euler's formula, the Voronoi polygons always follow the conditions given in equations 4 [11].

$$Number of\ Voronoi\ polygons = n$$

$$Number of\ polygon\ edges \leq 3n - 6$$

$$Number of\ polygon\ vertices \leq 2n - 5 \quad (4)$$

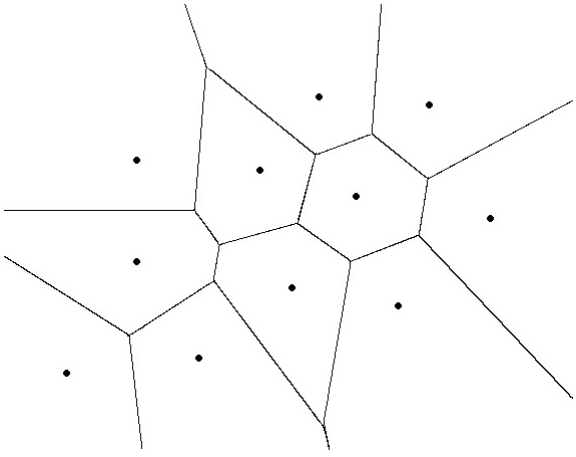


Figure 1: Voronoi Diagram of a set of random points

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

2.1.2 Delaunay Triangulation

The Delaunay triangulation (also called Delaunay graph) was first defined by Delaunay in 1934 [12]. It is obtained by connecting pair of points v_i and v_j in the plane such that the triangle formed by joining three non-collinear points with one side as $\overline{v_i v_j}$ is a Delaunay triangle, which means it is enclosed within a circumcircle with no other point $v_k \in V - \{v_i, v_j\}$ inside this circle. This property is also called “empty circle" property, and depicted in Figure 2. Delaunay triangulation is the dual of Voronoi diagram, as the centroids of Voronoi polygons correspond to the vertices in Delaunay triangulation. The duality property is also shown in Figure 2, and as a result of this property, the conditions in equation 4 are also satisfied by Delaunay triangulation with the following modification given in equation 5 [11]. Subgraphs of Delaunay triangulation can be generated from Delaunay graph and explained in next sections.

$$\begin{aligned} &Number\,of\ Delaunay\,vertices = n \\ &Number\,of\ Delaunay\,edges \le 3n-6 \\ &Number\,of\ Delaunay\,triangles \le 2n-5 \end{aligned} \tag{5}$$

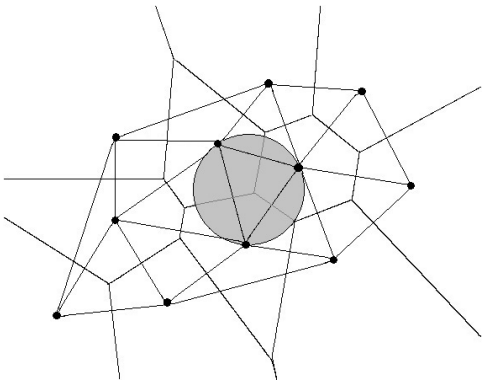


Figure 2: Delaunay Triangulation corresponding to the Voronoi diagram in Figure 1 showing empty circle and duality properties

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

2.1.3 Gabriel Graph

The Gabriel graph [13] is a subgraph of Delaunay graph and an edge exists between vertices v_i and v_j if they are 'least square adjacent' i.e. if for all other vertices $v_k \in V - \{v_i, v_j\}$ the condition given in equation 6 is satisfied [14].

$$d^2(v_i, v_j) < d^2(v_i, v_k) + d^2(v_j, v_k) \quad (6)$$

The Gabriel graph can be derived from Delaunay graph by retaining all the edges of the graph such that each edge is the diameter of a circle centered on the point halfway between its endpoints that has empty circle property. Formally, line segment $\overline{v_i v_j}$ is a Gabriel graph edge for all points $v_k \in V - \{v_i, v_j\}$ if the circle with diameter $\overline{v_i v_j}$ does not contain v_k. It is depicted in Figure 3.

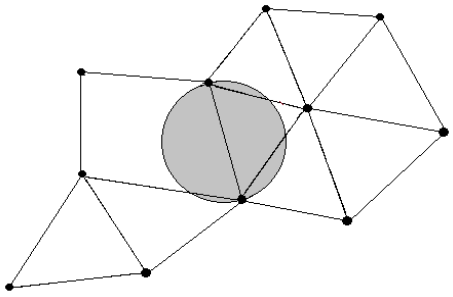


Figure 3: Gabriel graph corresponding to Delaunay graph in Figure 2 showing empty circle property

2.1.4 Relative Neighborhood Graph

The relative neighborhood graph (RNG) is a subgraph of the Delaunay graph and Gabriel graph. In this graph, two vertices are connected by an edge if there is no other vertex closer to both than they are to each other. This graph was proposed by Godfried Toussaint in [15]. An edge exists between vertices v_i and v_j if for all other vertices $v_k \in V - \{v_i, v_j\}$ the condition in equation 7 is satisfied.

$$d(v_i, v_j) \leq max\{d(v_i, v_k), d(v_j, v_k)\} \quad (7)$$

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

In other words, $\overline{v_i v_j}$ is an RNG edge, if for all points $v_k \in V - \{v_i, v_j\}$, $\overline{v_i v_j}$ is not the longest edge of triangle (v_i, v_j, v_k). It can be derived from Gabriel graph by retaining only those edges $\overline{v_i v_j}$ for which $lune(v_i, v_j)$ is empty, where $lune(v_i, v_j)$ is the intersection of two circles centered at v_i and v_j, each with radius $\overline{v_i v_j}$. RNG corresponding to the given set of points along with the empty lune property is represented in Figure 4.

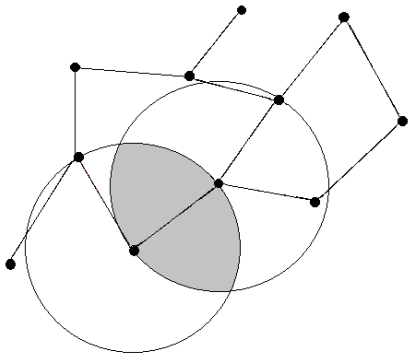


Figure 4: Relative neighborhood graph corresponding to Delaunay graph in Figure 2 showing empty lune property

2.1.5 Euclidean Minimum Spanning Tree

A tree is a obtained from an undirected graph by eliminating cycles, such that any two vertices are connected by exactly one path. A spanning tree of a graph is a tree including all its vertices. A minimum spanning tree (MST) [16] is the spanning tree of the graph whose sum of edge weights is less than or equal to the sum of edge weights of every other spanning tree. It is called Euclidean minimum spanning tree (EMST) when each edge weight is the Euclidean distance between the two vertex-coordinates [17]. EMST can be derived from the relative neighborhood graph such that edge $\overline{v_i v_j}$ is retained which is not the longest edge of a cycle in the RNG. It is shown in Figure 5.

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

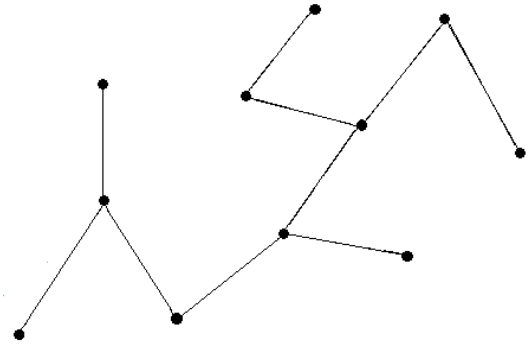


Figure 5: Euclidean minimum spanning tree corresponding to Delaunay graph in Figure 2

2.1.6 Nearest Neighbor Graph

The nearest neighbor graph [18] can be defined as a directed graph where $\overrightarrow{v_i v_j}$ is a directed edge from v_i to v_j if for all vertices $v_k \in V - \{v_i, v_j\}$, the condition in equation 8 is satisfied. It is depicted in Figure 6 for the given set of points.

$$d(v_i, v_j) \leq d(v_i, v_k) \tag{8}$$

The nearest neighbor graph, when considered as an undirected graph is a subgraph of Delaunay triangulation and can be obtained from Euclidean minimum spanning tree by retaining the edges to the closest neighbors.

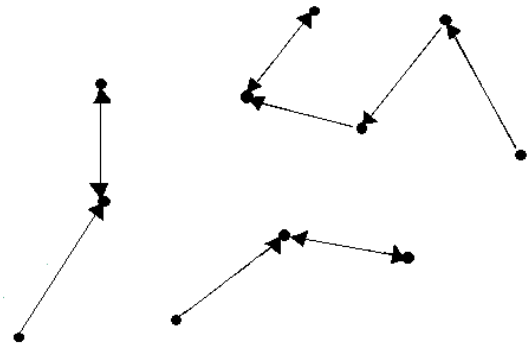


Figure 6: Nearest neighbor graph corresponding to Delaunay graph in Figure 2

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

Several algorithms exist in literature for generation of Voronoi diagram, Delaunay triangulation and its subgraphs. A summary of the popular algorithms, with their associated complexities is given in Table 1, where a graph with n vertices are considered. Many authors have also proposed time-efficient versions of the existing algorithms with improvement in complexities.

Table 1: Summary of popular algorithms for constructing Voronoi diagram, Delaunay triangulation and its subgraphs

Graph type	**Associated algorithms**	**Time complexity**
Voronoi diagram	Fortune's sweepline algorithm [19]	$O(nlog(n))$
	Divide and conquer algorithm [20]	$O(nlog(n))$
	Incremental algorithm [21]	$O(n^2)$
	Naïve algorithm [22]	$O(n^2log(n))$
	Lloyd's algorithm [23]	$O(nlog(n))$
Delaunay triangulation	Bowyer-Watson algorithm [24], [25]	$O(nlog(n))to O(n^2)$
	Lawson's flip algorithm [26]	$O(n^2)$
	Lifting or projection algorithm [27]	$O(nlog(n))$
Gabriel graph	From Delaunay graph [14]	$O(n)$
Relative neighborhood graph	From Delaunay graph [28]	$O(n)$
Euclidean minimum spanning tree	From Delaunay graph [29]	$O(nlog(n))$
Nearest neighbor graph	From Delaunay graph [30]	$O(nlog(n))$

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

2.2 β-Skeleton

In [31], Kirkpatrick and Radke have defined a parameterized group of proximity graphs known as β skeletons, β being the parameter. We define the neighborhood of vertices v_i and v_j by $N_{i,j}(\beta)$ for a fixed β. The β skeleton is an undirected graph, where two vertices v_i and v_j are connected by an edge if and only if no other point is located in $N_{i,j}(\beta)$. We denote a sphere centered at point p and radius r as $S(p,r)$. For $1 \leq \beta < \infty$, there are two definitions for β skeletons:

2.2.1 Lune-based β Skeleton

Points v_i and v_j are lune-based β-neighbors if an only if the lune $N_{i,j}(\beta)$ defined by the intersection of the spheres in equation 9 contains no other point $v_k \in V - \{v_i, v_j\}$.

$$N_{i,j}(\beta) = S((1-\frac{\beta}{2})v_i + \frac{\beta}{2}v_j, \frac{\beta}{2}d(v_i,v_j)) \cap S((1-\frac{\beta}{2})v_j + \frac{\beta}{2}v_i, \frac{\beta}{2}d(v_i,v_j)) \quad (9)$$

2.2.2 Circle-based β Skeleton

Points v_i and v_j are circle-based β-neighbors if and only if the neighborhood defined by the union of the two spheres each of radius $\frac{\beta}{2}d(v_i,v_j)$ passing through points v_i and v_j in their boundary, contains no point $v_k \in V - \{v_i, v_j\}$.

For $0 < \beta < 1$, vertices v_i and v_j are β-neighbors if and only if the neighborhood defined by the intersection of the two spheres each of radius $d(v_i,v_j)/2\beta$ which pass through points v_i and v_j in their boundary, contains no point $v_k \in V - \{v_i, v_j\}$. In this case, the lune-based and circle-based neighborhoods are identical.

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

β neighborhoods with diverse values of β are shown in Figure 7. It can be seen that β -skeletons contain both relative neighborhood graphs and Gabriel graphs as special cases. Gabriel graph can be defined as lune-based 1-skeleton ($\beta = 1$) and the relative neighborhood graph as lune-based 2-skeleton ($\beta = 2$) [32].

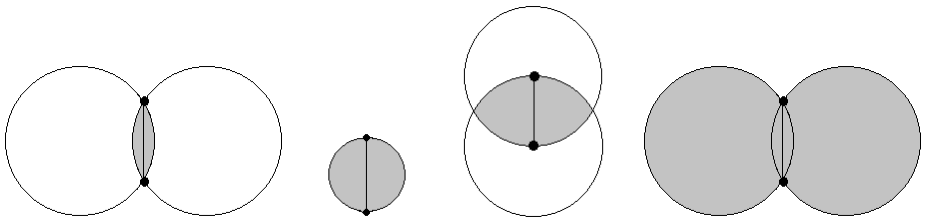


Figure 7: β neighborhoods (shaded regions) with different values of β. (a) Lune-based and circle-based $\beta = 0.5$ (b) Lune-based $\beta = 1$ (c) Lune-based $\beta = 2$ (d) Circle-based $\beta = 2$

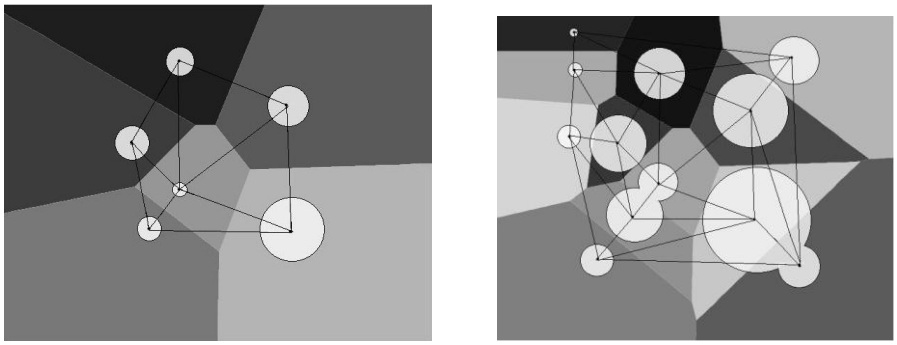
2.3 Johnson-Mehl Tessellation

Visualization of growth models is essential in many technical processes, including histological changes. For representing such phenomena, spatial patterns obtained from simple growth processes can be used. The Johnson-Mehl model [33] is defined as a Poisson-Voronoi growth model showing growth of particle aggregates, where a Poisson point process is applied to generate nuclei asynchronously, and the nuclei grow at the same radial speed [34]. The i^{th} generator $P_i = (p_i, t_i)$ is defined by a planar position vector p_i and associated appearance time t_i. The Johnson-Mehl tessellation can be viewed as comparable to a dynamic version of an additively weighted Voronoi diagram [35], in which the weights indicate the associated appearance times of particles in R^2 [22]. In stochastic geometry, the Johnsons-Mehl model is constructed to measure arbitrarily distributed geometrical properties.

The generation of Johnson-Mehl tessellation is explained in [34]. After appearance of a new nucleus the tessellation changes, as the incoming nucleus is inserted with a new Voronoi region and the neighboring Voronoi polygons are changed. The sizes of the associated spheres of nuclei are increased by the growth proportional to the time interval between the previous appearance and current one ($t_i - t_j$). This type of spatial

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

growth uses a Poisson point process, defined according to two cases of radial speed in [34], namely, time homogeneous Poisson point process and time inhomogeneous Poisson point process. An example of Johnson-mehl tessellation for a random set of points at two time instants is shown in Figure 8.



(a) (b)

Figure 8: Johnson-mehl tessellation for a set of random points (a) growth of particles at $t = t_1$ (b) growth of particles at $t = t_1 + 50$

2.4 O'Callaghan Neighbourhood Graph

An alternate definition of the neighborhood of a point given by O'Callaghan in [36] has been applied to histopathological images. For the histological context, O'Callaghan neighbourhood graphs have been discussed in detail in [37], where vertices are defined as "structures". Two types of constraints apply to the neighborhood, namely distance constraint and direction constraint. Two structures are neighbors if they are located within a certain distance (distance constraint), and not hidden behind other points classified as neighbors (direction constraint). It is shown in Figure 9, where v_1, v_2, v_3 are neighbors of v_i, however v_4 is not a neighbor of v_i. All distances d_{i1} to d_{i3} are below the upper distance threshold. Direction constraint is not fulfilled by v_4, as v_4 is hidden behind v_2. Also, v_4 is not a neighbor of v_1 as distance constraint is not fulfilled because d_{14} is above the upper distance threshold.

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

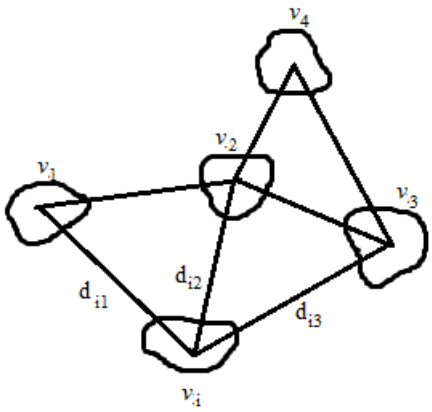


Figure 9: O'Callaghan neighborhood graph

2.5 Cell Graph

As the name suggests, cell graphs are formed by considering cells or cell clusters as vertices, and relationships between them as edges using certain predefined linking rules. Cell graphs are non-planar graphs as crossing of edges are allowed. The authors introducing cell graphs initially define the linking probabilities using the Waxman model [38]. However, in subsequent works like [39], three different variations of cell graphs are defined. These are depicted in Figure 10. Figures 10a and Figure 10b are created with 400 times magnification, whereas Figure 10c with 100 times magnification for clarity.

2.5.1 Simple Cell Graph

In a simple cell graph, an edge exists between two given vertices v_i and v_j if and only if the Euclidean distance d_e between the vertices is less than a predefined threshold D, which is defined according to the characteristics of the tissue architecture.

$$d_e(v_i, v_j) < D \tag{10}$$

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

2.5.2 Probabilistic Cell Graph

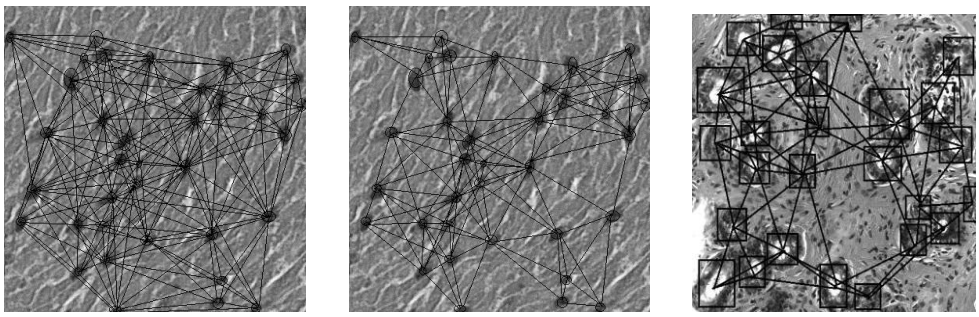
In a probabilistic cell graph, edges are assigned to vertices depending on the distance between vertices, by using a probability function which can be defined according to the tissue being analyzed. One such probability function used in literature is given by:

$$P(v_i, v_j) = d_e(v_i, v_j)^{-\alpha} \quad (11)$$

where v_i and v_j are two vertices and d_e denotes the Euclidean distance between them. α is an experimental parameter which determines the rate at which probability of a link decreases with the increase in distance between vertices, hence, controls the graph density. The edges are defined by a binary relation E on V for all pair of vertices v_i and v_j, such that $E = \{(v_i, v_j) : P(v_i, v_j) > r\}$, r being a real number between 0 and 1. With such a construction, the vertices closer to each other are more likely to be linked compared to vertices farther away, however it is not necessary that a link will be formed even if the distance between vertices is small, and will depend on the parameter α.

2.5.3 Hierarchical Cell Graph

A hierarchical graph is generated by connecting smaller subgraphs representing cell clusters together to form the larger graph on the image. Clusters are identified as cell groups exceeding a threshold of the number of cells in a given region (defined by placing a grid on top of the image). Hierarchical graphs are built using these clusters as vertices, where each vertex represents a simple cell graph on the cells included in the cluster. The linking rules applied are those used for probabilistic cell graphs.



(a) (b) (c)

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

Figure 10: Representative cell graphs for small regions of H&E stained images (a) Simple cell graph with a $D = 200$ for gastric tissue region (b) Probabilistic cell graph with $\alpha = 0.2$ for gastric tissue region (c) Hierarchical cell graph for breast tissue where each vertex represents a cell cluster with a corresponding simple cell graph built on the constituent cells (denoted by rectangle)

2.6 Attributed Relational Graph

A graph G is an Attributed Relational Graph (ARG) when both the vertices and the edges contain associated attributes [40]. The vertex attributes for vertex v_i are denoted as a vector $\mathbf{a_i} = [a_i^{(k)}], (k = 1,2,3,...,K)$, where K is the number of vertex attributes in the vector $\mathbf{a_i}$, and the edge attributes (or weights) for edge e_j by the vector denoted as $\mathbf{b_j} = [b_j^{(m)}], (m = 1,2,3,...,M)$, where M is the number of edge attributes in the vector $\mathbf{b_j}$ [41]. In Figure 11, ARG with vertex attribute vectors $\mathbf{a_i}, i = 1,2,3$ and edge attribute vectors $\mathbf{b_j}, j = 1,2,3$ is shown. Vertex attributes represent object properties like size, position, shape and color whereas edge attributes define relationships between vertices like the distance, common boundary and dissimilarity between objects.

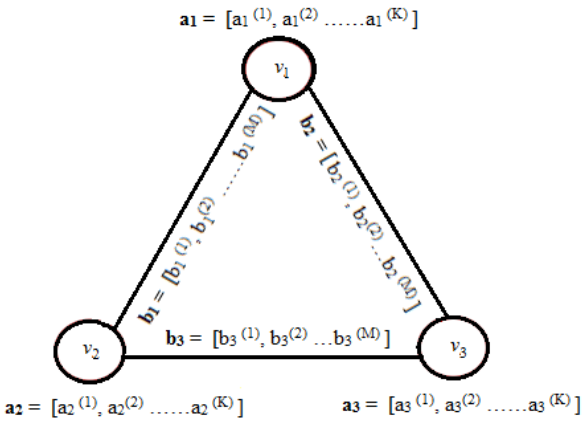


Figure 11: Attributed Relational Graph

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

2.7 Global Graph Measures: Few Examples

Once the graphs have been constructed using the above described methods or other approaches, the next step in image analysis is to extract appropriate quantities from these graphs. One way of achieving this is to define a set of global graph measures. There are many graph-based metrics which can be derived from graph representations. Definitions of these measures can be found in [42], some of the most commonly used measures in image analysis are briefly described below.

2.7.1 Graph Size

Graph size represents the span and extent of a graph, and can be measured by counting the total number of vertices, total number of edges and total number of trees in a graph. For Voronoi diagram, size can be measured in terms of total number of Voronoi polygons, their vertices or corresponding Delaunay triangles.

2.7.2 Degree

The degree of a vertex is defined as the number of neighbors of a vertex, or number of vertices linked to the given vertex by edges. Average degree D_{av} of graph G with n vertices and degree D_i of the i^{th} vertex v_i is calculated as below, and suitable for representing the relative density or sparsity of the graph.

$$D_{av} = \frac{1}{n}\sum_{i=1}^{n} D_i \tag{12}$$

2.7.3 Clustering Coefficient

Clustering coefficient is the measure of the extent to which vertices in a graph have a tendency to cluster together. For a given vertex v_i, let the neighboring vertices be contained in the neighborhood set N_i defined as:

$$N_i = \{v_j \in V \ \forall \ v_j \neq v_i, (v_i, v_j) \in E\} \tag{13}$$

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

For an undirected graph, clustering coefficient C_i for vertex v_i is defined as:

$$C_i = \frac{2|e_{jk}|}{D_i(D_i-1)} : v_j, v_k \in N_i; e_{jk} \in E \quad (14)$$

where, D_i is the degree of vertex v_i and $|e_{jk}|$ is the total number of edges between the neighbors of v_i. It can be noted that $0 \le C_i \le 1$. The average clustering coefficient C_{av} for the graph can be calculated as:

$$C_{av} = \frac{1}{n}\sum_{i=1}^{n} C_i \quad (15)$$

2.7.4 Eccentricity

The eccentricity ε_i of a graph vertex v_i is the maximum distance between v_i and any other vertex v_j in V. It can be considered as the measure of how far a vertex is from the vertex most distant from it in the graph. The vertices of a disconnected graph are said to have infinite eccentricity [43]. The minimum graph eccentricity is called the graph radius and the maximum eccentricity of graph is called the graph diameter. Average eccentricity of the graph can be calculated as:

$$\varepsilon_{av} = \frac{1}{n}\sum_{i=1}^{n} \varepsilon_i \quad (16)$$

2.7.5 Path Lengths

A path in a graph or tree is a finite or infinite sequence of edges which connects a sequence of vertices. Edges can be part of a path only once. The minimum, average and maximum path lengths of the graph can be calculated. Let $d(v_i, v_j)$, where $v_i, v_j \in V$ denote the shortest distance between v_i and v_j, assuming that $d(v_i, v_j) = 0$ if v_i cannot be reached from v_j. Average path length l_{av} can be defined as:

$$l_{av} = \frac{2}{n(n-1)}\sum_{i \ne j} d(v_i, v_j) \quad (17)$$

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

2.7.6 Cyclomatic Number

A connected component of an undirected graph is a subgraph containing vertices connected to each other by paths, but it is not connected to any other additional vertex in the supergraph. A cycle is a path where the start vertex is also the end vertex. Cyclomatic number or circuit rank γ of a graph is the minimum number of edges which must be removed in order to eliminate all cycles from the graph [44]. Considering a graph with n vertices, m edges and k connected components, γ is given by:

$$\gamma = m - n + k \tag{18}$$

2.7.7 Statistical Descriptors

For Voronoi diagrams and Delaunay triangulations, there can be certain measures like area and perimeter associated with the individual Voronoi polygons and Delaunay triangles respectively. For ARGs, there can be several attributes associated with vertices and edges. Statistical descriptors can be used for globally describing these attributes of individual vertices and edges. These include mean, standard deviation, minimum to maximum ratio, disorder, skewness, kurtosis and higher-order descriptors. For these graphs, co-occurrence matrix can also be constructed and Haralick features [45] can also be extracted from them. These statistical descriptors are capable to represent several diverse properties of entire graphs and their variations quantitatively.

2.7.8 Spectral Descriptors

Spectral features include properties extracted from the matrices used to describe the graph, such as adjacency matrix and Laplacian matrix, and their study is called spectral graph theory [46]. Some of the frequently used features include eigenvalues, spectral radius, eigen exponent, Cheeger constant, eigenmode perimeter, eigenmode volume and so on [47], [48]. These spectral features have the ability to indicate various fundamental properties of the graph.

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

3 Graph-based Methods in Histopathological Image Analysis

As stated before, graphs have been an interesting area of research for analysis of histological image data, due to their ability to efficiently represent topology of tissue structures in suitable ways. High level or structural feature extraction methods are increasingly used to characterize a tissue using spatial distribution and neighborhood properties of its components by constructing representative graphs and calculating a set of local and global features which provide quantitative information about tissues.

Graph-based techniques have been applied successfully to model several biological phenomena. Biological structures, especially biological membranes have been represented using minimum spanning trees in [49] and using a variant of MST called angular diagram with angular properties in [50]. [14] shows that Gabriel graph can be efficiently used for modeling in the field of geographic variation research in biology. Protein structures are also explained using three dimensional Delaunay tessellation and residue contact networks in [51], concluding that such naturally formed networks can be modeled by graphs. Other graph theoretical algorithms have been developed to model and identify protein structures in [52] and [53]. Random cellular structures in materials like bones, wood, cork, foams, plants, food etc. are studied utilizing planar graph theory, and topological properties of groups of cells with trivalent vertices in the plane are analyzed in [54]. Coronary arteries in human body have also been represented by graph structures in [55], followed by graph matching to label the coronary segments in arteriograms. It has been shown in [56] that cellular metabolism of 43 organisms can be identified with a network architecture as metabolic networks are composed of many small, densely connected topological modules that combine in a hierarchical manner into larger, less cohesive units. Multicellular tumor spheroids have been simulated and studied in [57] by presenting a novel three-dimensional agent based Voronoi-Delaunay hybrid model. Molecules and their interactions inside cells have also been represented as networks with functional organizational characteristics, and techniques have been reviewed in [58]. Another review containing recent developments in the field of graph-based methods to analyze complex cellular networks is given in [59]. In this paper, we focus on the graph-theoretic description used for representing histological architecture, as tissues are the main building blocks of organisms.

Graph-based methods to analyze histological images were initially explored in the direction of syntactic analysis. Syntactic methods based on neighborhood conditions in pattern recognition were introduced by Fu [60]. Following this, a relatively new technique was introduced in diagnostic quantitative pathology called *syntactic structure*

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

analysis, which was believed to provide quantitative information on tissue architecture [61]. The earliest work in this direction is presented by Prewitt et al. [62] for characterizing epithelial tissues of urinary bladder mucosa, where measurements are performed using an interactive digital image processing and decision making system called PEEP/DECIDE/GRAPH [63] to automate diagnostic procedures. Features are extracted from haematoxylin stained cells representing intensity, texture, shape, differentiation and structure, and from tissue regions quantifying spatial distributions and tissue organization. Furthermore, graphs and grammars in histology are formally defined by Prewitt in [64]. She introduces a graph-theoretic model for tissues and syntactic pattern recognition concepts using two random spatial structures: Dirichlet tessellation and Delaunay tessellation defined as unique topological planar dual graph of the Dirichlet tessellation. These structures are used to obtain a unique, invertible, relational and attributed graph representation of histologic sections. She also suggests that Dirichlet's domain can be extended to a more general Johnsons-Mehl domain. Another early application of the technique is found in [65] for analyzing muscle tissue patterns using a distance measure between graphs. In this paper we give an account of the various approaches using syntactic structure analysis in histology, however we emphasize that this approach though theoretically powerful, has limited usability in practical applications due to its extensive computational requirements, which was later also observed experimentally. Subsequently, with the development of digital pathology techniques, some more efficient graph-based methods were analyzed.

The commonly used graph-theoretic methods in histopathology are Voronoi diagrams, Delaunay triangulations and related graphs, O' Callaghan neighborhood graphs, cell graphs and attributed relational graphs. It has been stated in [66], and we are in agreement, that a total of approximately 150 spatial-relational features have been extracted from all graph structures for histological images. This section reviews the various graph-theoretic methods used in histopathological image analysis.

3.1 Voronoi Diagram, Delaunay Triangulation and Related Graphs

The general Johnsons-Mehl model was considered effective for representing tissue architecture, however the main drawback of this method is the computational time involved for its construction, as it requires working with individual pixels in the whole slide image. It was realized that Voronoi-based tessellations are computationally more manageable. In [37], the advantages of representing tissue structures using Voronoi tessellations are listed, that the construction of a complementary tessellation aids in

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

visualization of cell boundaries and introduces a metric system according to neighboring structures. Hence, the most abundantly applied graph-theoretic techniques for representation and analysis of histological images till date are the Voronoi diagram, Delaunay triangulation and related graphs.

Voronoi diagrams and Dirichlet domains have been explored in [67] for growth assessment of degrees of differentiation in terms of cell population dynamics in Feulgen stained colorectal neoplastic cell colonies growing on histological slides, by studying the structure of clones in 11 different media. The authors in [68] describe a graph-based approach for analysis of H&E stained images of Cervical Intraepithelial Neoplasia (CIN), where Delaunay triangulation mesh is first computed on a region and 18 quantitative features of the triangulation are extracted. To determine which characteristics contribute to classification, Kruskal Wallis Test is used. Discriminant analysis is applied for classifying image regions into one of the classes normal, CIN1, CIN2 and CIN3. Another work using Delaunay graph representation for H&E stained colon tissue images is given in [69]. A Delaunay graph is constructed on tissue components and their edges are colored based on relationships between end vertices. Structural graph features like average degree, average clustering coefficient and diameter are extracted from these color graphs. SVM is used to classify images into normal, low-grade and high-grade cancer, and the performance of this approach is compared with intensity based features, texture based features and simpler graphs like colorless graphs and Delaunay triangualations.

Neighborhood relationships between cells in germinal centers from lymph nodes are studied in [70] for H&E stained lymphoma biopsy images, by representing them with Gabriel graphs. Subgraphs are also formed for each cell type by assigning label to each node, where a node represents a cell. Two categories of graphs, namely binary and grey tone graphs are generated, and features are extracted from graphs using morphological operations like erosion, dilation, opening and closing. Additionally, distance transform is applied to these graphs. Tissue images of follicular hyperplasia and small cleaved cell follicular lymphoma are analyzed and compared using the extracted quantitative information.

The minimum spanning tree has been used abundantly for quantitative representation of tissue architecture in syntactic structure analysis tasks. MST can be constructed using different neighborhood conditions, the most frequently used being Voronoi neighborhood. H&E stained tissue sections of colorectal adenomas are evaluated in [71] using MST features to differentiate between three grades of dysplasia. In [72], the authors analyze invasive breast cancer tissue images using 10 syntactic structure

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

features extracted from the corresponding MST of each H&E stained breast tissue region. They conclude that this technique can contribute towards improvement in prediction of prognosis for breast cancer patients and consider it a promising method to analyze tissue architecture in breast cancer. Syntactic structural analysis using 10 features of the MST is carried out in [73] for studying H&E stained uveal melanoma specimens. A significant correlation between Callender cell type and certain MST parameters is reported, showing that features can be used for mathematically describing melanoma architecture across the spectrum of the Callender classes. Syntactic structure analysis using MST constructed on H&E stained low magnification gastric tissue images is performed in [74], where the authors observe how the 11 extracted features vary with different grades of gastric atrophy. Kayser et al. [75] discuss three different information recognition algorithms that can be used for field of view detection in virtual microscopy to measure diagnosis-related information. One of the algorithms uses graph-based representations of tissues by construction of associated MST. Measurements based on MST and Voronoi neighborhood condition are used for selection of the slide areas containing the most significant information to derive associated diagnosis. They perform experiments on large sets of histological images comprising organs like colon, lung, pleura, stomach and thyroid.

In several works, a combination employing more than one Delaunay related graph is used. For example, syntactic structure analysis of Feulgen-Schiff stained mesothelioma tissue sections of lungs is described in [76], where tissues are represented using Voronoi diagrams, Gabriel graphs and MST. A set of 48 global graph features are extracted and discriminative features are selected using sequential floating forward selection method. The performance of these graph-based parameters is evaluated with respect to differential diagnosis of lesions and grading of tumors using k-nearest neighbors (kNN) classification algorithm. Architectural features of bronchial intraepithelial neoplasia are evaluated in [77], in which around 30 features are derived from Voronoi diagram, Delaunay tessellation and MST constructed on nuclei within a predefined region of interest. A discriminant analysis is then performed to differentiate between normal lung biopsies and carcinoma in-situ (CIS). An architectural index is also calculated and analyzed for each of the bronchial biopsies interpreted as hyperplasia, metaplasia, mild, moderate or severe dysplasia by conventional histopathology criteria. In [78], H&E stained oral cancer epithelium images are first segmented using color deconvolution for haematoxylin. The cell profiles are then used for construction of Voronoi regions (V-cells), and the corresponding Delaunay triangulation, Gabriel graph, relative neighborhood graph and MST. 29 architectural features including statistics of cell areas are extracted from graph representations. Discriminant analysis classify images into normal, dysplastic and neoplastic. The authors also assess that the graph features are

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

more discriminative between normal and malignant tissue, and less between normal and premalignant tissue and between premalignant and malignant tissue. In [79], the authors distinguish high grade breast cancer in H&E stained biopsy specimens based on the presence or absence of lymphocytic infiltration, which is determined by performing a lymphocyte segmentation followed by representation of images using Voronoi diagram, Delaunay triangulation and MST. Graph-based features along with nuclear features are used for training SVM classifiers. The authors compare their approach to the Varma-Zisserman texton approach, confirming higher classification performance.

In some works, another structure called Ulam tree has been applied along with Voronoi diagram and related graphs, for tissue representation. To our knowledge, there is limited documentation available defining specific algorithms for generation of an Ulam tree, however, it can generally be described as a mathematical object growing in space and time according to certain recursive rules. It is based on the tessellation described in [80], where the plane is divided into regular structures by using some predefined generation rules that define growth of structures. Delaunay triangulation and Ulam tree have been used for describing Feulgen-Schiff stained oral squamous cell carcinoma images in [81]. Geometric centers of cell nuclei are marked as graph vertices, and two global graph features namely average Delaunay edge length and average homogeneity of Ulam tree are extracted from the graphs. Prognostic values are evaluated suggesting whether significant prognostic information is available or not, for invasive front, superficial part and putatively normal mucosa using the two features. In [82], authors perform graph-based study on H&E stained prostate cancer, cervical carcinoma, tongue cancer and normal oral mucosa images. Voronoi diagrams, Delaunay graphs, Gabriel graphs, MST and Ulam trees are computed for these images following a simple cell segmentation based on thresholding, and 27 structural features are obtained from these graphs. 10 discriminative features have been demonstrated to have a diagnostic or prognostic potential, showing the ability of Voronoi diagrams and related graphs in quantitative tissue architecture analysis. In [83], the authors present a continuation to this work, where they use H&E stained images of normal oral mucosa, carcinoma of lungs, larynx, prostate and cervix and the same methodology as in [82], for graph-based representation and feature extraction, followed by study of variation in graph parameters with number of cells. They study the stability in values of graph-based features as function of number of cells for different features in same tissue, for same feature in different tissues and for same feature in different cases of cervical carcinoma. They discuss the numerical requirements i.e. the number of objects needed to obtain acceptable degree of reproducibility for a given set of features. They find that coefficient of variation depends severely on the number of cells considered, and on increasing the number, the features obtain a higher prognostic value. In [84], the authors state that

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

graph-based methods in [81] and [82] for quantitative evaluation of tissue architecture can be a feasible approach for chemoprevention of oral cancer to monitor treatment effects. They suggest that by employing Voronoi diagrams, Delaunay graphs, MST, Gabriel graph and Ulam tree, structural manifestations of the interaction between cells in tissue can be quantified.

3.2 O'Callaghan Neighborhood Graphs

Some of the earliest works have explored non-oriented complete graphs, constructed using the neighborhood condition defined by O'Callaghan [36], as explained in Section 2.4. Kayser stated in [37], that Voronoi neighborhood condition is useful for representing solid tissue structures, however, not suitable for adenomatous structures (tubular structures). The reasons are that firstly, Voronoi neighborhood is independent of the size of structures and cannot be influenced by additional external parameters, and secondly the empty space inside tubular structures is considered to belong to the basic metric space. Hence, some studies of syntactic structure analysis apply the O'Callaghan neighborhood conditions for graph construction for analysis of glandular structures in histopathological images. Some global features derived from the graphs include number of vertices, number of edges, cyclomatic number, mean distances and distribution of nearest neighbors, second nearest neighbors, third nearest neighbors etc. and frequency distribution of n-stars and n-closed paths. Feature calculation is followed by quantitative discriminant analysis showing which structural features are more representative than others to represent a specific disease condition.

In [85], low magnification H&E stained colon tissue images of healthy mucosa, tubulovillous adenoma and highly to moderately differentiated adenocarcinoma are analyzed for measuring histomorphological structures. Images are represented using graphs based on O'Callaghan's neighborhood, with glands as vertices and coherence of neighboring glands as edges. More study in this direction is shown in [86], performing measurements on H&E stained and immuno-stained colon tissue sections. The authors construct graphs with centers of interesting structures like glands, epithelial cells and positively immuno-stained cells as vertices and neighboring structures as edges using O'Callaghan's definition. Similar approach for differentiating healthy mucosa and adenomatous structures of colon is also described in [87], analyzing networks formed by single cells and by tubulopapillary formations in adenomatous tissue. In [88], H&E sections of healthy lung parenchyma and four major types of bronchus carcinomas namely, epidermoid, adeno, small cell and large cell carcinomas are studied. Graphs are

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

formed by epithelial cells using O'Callaghan neighborhood condition, and features are extracted from the graphs. Discriminant analysis is performed to classify these sections. Reproducibility of graph theory approach in previous studies is tested and found high. Related work for small cell anaplastic lung carcinoma is presented in [89].

A study describing three dimensional morphometric analysis of epithelium sections is described in [90]. Histological sections are represented using three dimensional MST and O'Callaghan neighborhood trees. Features are extracted from the trees and are quantitatively analyzed for differentiating H&E stained tissue sections of normal squamous epithelium, dysplasia, CIS and carcinoma. It is reported that 3D morphometric analysis performs better and provides significant information compared to 2D analysis.

3.3 Cell Graphs

In the cell graph technique, the authors have used low magnification (80-100 times) tissue images of brain, breast, colon and bone tissues. As explained in detail in Section 2.5, authors have constructed these graphs using cells or cell clusters as graph nodes and different linking conditions for edges. Global graph metrics like average degree, clustering coefficient, average eccentricity, eigenvalues of graph, average node weight, most frequent edge weight etc. are extracted from the cell graphs. Using these features, the images are classified with the help of ANN or SVM classification methods.

The authors introducing cell graphs point out some advantages of the cell graph approach over Delaunay triangulation and its subgraphs in [91], specifically the relaxation of geometric constraints. They state that, since the Delaunay triangulation is planar and consists of edges only between adjacent vertices, it can capture spatial relationships between nuclei within a small vicinity, however, no evidence justifies such a restriction in architecture of tissues. In contrast, cell graphs are non-planar and can allow crossing of edges, overcoming this restriction. Secondly, unlike Delaunay graphs, cell graphs are not always a single connected component, hence, can also be used for representing sparse tissues. Thirdly, the length of shortest path in the cell graph is not always three as in Delaunay triangulation, and depends on spatial distribution in the tissue. Further, thresholds for linking vertices are not fixed in cell graphs as in the Delaunay graphs. These properties provide a greater flexibility to cell graphs for modeling diverse types of tissues based on the tissue topology.

Cell graphs for brain cancer analysis have been introduced in the following studies. In [38], the authors first segment surgically removed tissue sample images of brain cancer

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

(glioma) using k-means clustering, followed by cell graph generation with linking between nodes defined according to Waxman model [92]. They extract global graph metrics and perform classification at cellular level into cancerous, healthy and unhealthy (inflamed) cells, and tissue level into cancerous, healthy and non-neoplastic inflamed tissue using ANN classification method. In [93], the technique for simple cell graphs is extended to develop augmented cell graphs (ACG) for H&E stained brain cancer (glioma) tissues. ACG are fully-connected graphs containing weights associated with nodes and edges. Node weight is defined as size of the cluster corresponding to a node, and edge weight is the Euclidean distance between end nodes. The authors evaluate the performance of the method and compare with that of simple cell graphs, reporting an improvement using ACG. In [94], H&E stained brain biopsy samples are studied by constructing cell graphs similar to [38] but using a modified probability function over Waxman model for link establishment, and extracting spectral features based on the eigenvalues of the adjacency matrix and normalized Laplacian. Hierarchical classifier is designed to classify samples as healthy or unhealthy in first layer, then as benign or malignant in the second layer of classifier. In [95], the authors construct probabilistic cell graphs for H&E stained brain biopsy samples, followed by extraction of a set of global graph metrics from the constructed graphs. Samples are classified into cancerous, healthy or inflamed tissue using multilayer perceptrons. They also perform a comparison of the graph-based method with spatial cell distribution and texture-based approaches, and establish better recognition rates for the graph-based method. In [96], H&E stained brain tissue photomicrograph samples are analyzed by generating probabilistic cell graphs. The graph evolution process is analyzed by identifying three phases that a graph passes through during its evolution, and strong relation between the malignancy of cancer and the phase of its graph is demonstrated. Connectivity is controlled using edge exponent, and a set of 11 graph features are quantitatively analyzed for each phase. It is observed that normal, low grade cancerous and high grade cancerous cell graphs evolve differently. Their approach is also compared with texture-based approach and leads to better results. They propose that their method opens up new possibilities for dynamic modeling of cancer, an ability that is not exhibited using the texture approach.

In [39], the authors use cell graphs to model H&E stained breast tissues. Hierarchical cell graphs are introduced and generated for the images, and a set of global graph features are extracted from hierarchical graphs. SVM classification is applied to classify images into benign, invasive and in-situ. Performance accuracy is compared with intensity based features, Delaunay graphs, probabilistic graphs, simple cell graphs and hybrid approach combining all these features. Hierarchical cell graphs obtain the best learning ratio compared to the other techniques.

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

H&E stained histopathological images of bone tissue samples are analyzed using advanced version of cell graphs called *ECM-aware cell graphs* in [91]. The images are first segmented using eigenvalues of Hessian matrix and spatial coordinates of cell nuclei are calculated to obtain corresponding nodes of cell graphs. A color code is allotted to each node depending on composition of Extra-Cellular Matrix (ECM) around the node. An edge exists between nodes if they have same color code and if they are neighbors (i.e. have physical contact). Topological and spectral features calculated from these graphs, and SVM classification is applied to classify the images into healthy, fractured or cancerous. The approach can distinguish between different types of cells, and is compared with simple cell graphs and Delaunay triangulation methods, showing better performance than these methods.

Three dimensional cell graphs for spatio-temporal analysis of tissues are explored in [97] and [98]. In the first study, the authors analyze 3D collagen-I cultures of human mesenchymal stem cells (hMSCs). Images are first acquired using fluorescence microscopy and SYTOX Green Dye for staining. Cell graphs are constructed on the images using three dimensional Euclidean space and features are extracted from these graphs. The features are quantitatively analyzed to track the underlying structural changes and biological events. The authors compare their cell graph approach to Voronoi diagrams and Delaunay graphs, and show that cell graphs outperform the other two graph-based methods. In the second study, 11 different cell types (non-cancerous, pre-cancerous and malignant states of different tissue origins) are discriminated in collagen-I hydrogel cell culture images using metrics extracted from 3D cell graphs. 20 cell graph features are initially extracted, following which a three-mode tensor analysis is applied to determine the most representative features. The authors conclude that five features are sufficient to easily distinguish between the cell types, as they provide important quantitative image information and can potentially improve accuracy of disease diagnosis and detection.

A study based on the idea of *cell map* and *cell web*, similar to the cell graphs described above, also requires special mention in this section. In [99], the authors perform analysis of H&E stained histological sections of colon of the types normal mucosa, aspecific colitis, ulcerative colitis and Crohn's disease. They first describe three separate algorithms for detection of objects, namely, cell nuclei, glands and epithelium, for their dataset. Cell nuclei are determined by a segmentation procedure involving a color-thresholding algorithm, after which cell maps are generated where cells are represented by single spots. Cell webs are constructed on these cell maps where cells act as nodes and edges connect each cell to its 5- nearest neighbors. They employ cell webs and their properties for detecting glands and epithelium in images. This is followed by extraction

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

of 38 morphometric parameters from their objects, and 14 tissue cytometric parameters describing ratios of tissue compartments. These parameters are analyzed using discriminant analysis, and they conclude that their parameter set can classify images of normal mucosa, aspecific colitis and ulcerative colitis with reasonable accuracy, however, Crohn's disease identification requires development of more advanced techniques. In [100], the authors use the technique described in [99] for classification of H&E stained gastric biopsies into normal mucosa, gastritis and adenocarcinoma. They conclude that tissue architecture and cytometric parameters are effective in the classification of gastric samples.

Recently, a graph-based method similar to cell graph, namely, cell cluster graph (CCG) has been explored in [101] for investigating prostate cancer tissue microarray images. In this work, cell nuclei clusters at low magnification are first identified by using a method based on concavity detection [102]. Cell clusters are treated as nodes, and a probability function is defined for establishing links between nodes. Following the graph construction, local and global features are computed, capturing the morphology of tumor tissue. Support vector machine classification is used to determine biochemical recurrence. The authors also report a superior performance of CCG compared to Voronoi and Delaunay methods. They state the advantage of CCG method is that nuclei boundaries need not be located accurately, so the method can be applied at low magnifications. Further, this approach identifies exact nuclei clusters unlike the cell graph approach described earlier in this section.

3.4 Attributed Relational Graphs

Attributed relational graphs are emerging as a topic of interest for researchers in this field. Attributed minimum spanning trees are explored for representing Feulgen stained soft tissue tumors (malignant fibrous histiocytoma, fibrosarcoma, rhabdomyosarcoma, osteosarcoma and Askin tumor) in [103], where a basic graph is first constructed according to the neighborhood condition of O'Callaghan. Nuclear features related to morphometry and DNA-content are attributed to the vertices, and the differences or distances between features of connected vertices are attributed to the corresponding edges. Resulting MST is decomposed into clusters using a suitable decomposition function on the edges, and clusters of distinct nuclear orientation are detected. A cluster tree is constructed by defining the geometric center of a cluster as a new vertex and by computing the neighborhood of the cluster vertices. This generates an attributed MST

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

containing characteristic structural properties of the histological image which are analyzed for the mentioned tumor types.

Another study employing ARG for content-based retrieval of H&E stained breast tumor tissue images is given in [41]. The authors have represented segmented tissue images using ARG, where nodes are centroids of tissue regions having attributes as a label from the segmentation and some associated morphological features of the associated segment. Edges connect neighboring segments and are characterized by edge attributes which are Euclidean distance and common boundary between connecting nodes. A* based graph matching algorithm is employed to retrieve the image regions most similar to a given query region.

3.5 Other Graph-based Approaches

Some authors have developed and/or applied other graph-based methods used less frequently, for representation and analysis of histological images. A novel graph-based approach is given in [104], where the authors define Zone of Influence (ZOI) tessellation for haematoxylin stained cervical tissue section images. Images are first segmented using a two-stage procedure, and ZOI tessellation is formed by the skeletonization or medial axis transformation of the background image. Skeletonization is performed using an algorithm for sequential application of hit or miss transformation widely used in mathematical morphology [105]. The authors emphasize on some advantages of generating neighborhood graphs by this method as compared to O'Callaghan graphs and the Voronoi diagrams. Firstly, in O'Callaghan's method it is necessary to fix certain parameters like maximum possible distance and minimum angular separation of neighbors with respect to the node in question, which is not required in this method. Moreover, O'Callaghan graphs and Voronoi diagrams are based only on points and information about the cell shape is not considered in defining neighborhoods, however, shape information is considered in ZOI tessellation. After generating ZOI tessellations, they extract a set of global features using which they characterize tissue sections into normal and dysplastic. They also compare results of their approach with minimum spanning tree approach.

A similar morphological approach is described in [106], where H&E stained tick epithelium and mammary gland adipose tissue from mice are differentiated using a graph-theoretic description called neighborhood graph built using morphological operators. A segmentation process using mathematical morphology is first applied to

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

the tissue images. Watershed algorithm defining neighborhoods between image objects is applied on the segmented image and neighborhood graphs are created according to the method explained in [107], where vertices are defined by the centroids of cells. Global measures are derived from the neighborhood graphs and regions, and individual measures are derived from cells. SVM classification is used and performance of classifier is evaluated, along with the discriminative ability of each measure in characterizing the two tissues.

Another work in this category for classification of H&E stained prostate cancer images is found in [108]. Gland segmentation is performed and gland angle is calculated that reflects the dominant orientation of the gland. Following this, a subgraph is constructed to connect neighboring glands together and form a gland network. According to the authors, the resulting subgraphs are superior to Voronoi diagrams, Delaunay graphs and minimum spanning trees because they are able to characterize local gland arrangement. Co-occurrence gland angularity (CGA) matrix is computed to represent second-order statistics of gland orientations within each gland network and 39 CGA second order statistical features are extracted. A random forest classifier is used to classify images into benign or malignant and biochemical recurrence or non-biochemical recurrence. Comparative evaluation with other quantitative histomorphometric attributes namely gland morphology, Voronoi diagram, Delaunay triangulation, MST, gland density and texture-based approaches is also performed, and authors report better performance of the CGA approach compared to other approaches for the given problem.

3.6 Hybrid Methods: Graph-based Features with Other Feature Types

Besides graph-based methods, image analysis methods used abundantly in histology incorporate approaches for extracting information in the form of texture, intensity (or color), morphology (or shape) and frequency-based features from individual pixels or superpixels (sometimes predefined objects) in the image. One of the earliest works using these methods are explored by Hamilton et al. in [109], [110] and [111]. In [109], H&E stained colorectal epithelium images are analyzed using morphological features followed by discriminant analysis that classifies images as normal and malignant. In [110], cytometric features based on cell morphology are used for classifying fine needle aspiration cytology (FNAC) breast specimens into benign and malignant by constructing Bayesian belief networks. In [111], texture analysis is used to classify colorectal mucosa images into normal and adenomatous (dysplastic). More work using texture features for classification purpose is given in [112], [113], [114], [115], [116] and [117]. Color or

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

intensity-based features in histology have been explored in [118] using grey-level features in muscle tissues, and in [119] for analysis of stain components in skin tissues. Shape-based or morphological features have been extensively used in histological image analysis, for instance, in [120], [121] and [122]. Frequency-based features include the use of spectral bands, Fourier transforms and wavelets among others. Applications utilizing multiple spectral bands for analysis include [123] and [124]. Wavelets have also been proved as powerful features in [125], [126] and [127] for breast cancer, prostate cancer and meningioma classification respectively.

A part of the developed techniques use a combination of the graph-based characteristics with other feature types. Visual information extracted using both approaches followed by data fusion, promise an improvement in results than individual approaches. One of the earliest studies in this direction is presented in [128], where the authors report analysis of Feulgen stained primary lung carcinoma tissue using geometrical and morphometric parameters of cell nuclei along with texture parameters based on MST, Johnson-Mehl diagram and Voronoi tessellation. They study the discriminating power of combined features derived from graph representations and nuclear morphometric features for classifying adenocarcinoma, epidermoid, small cell anaplastic and large cell anaplastic carcinoma. In [129], Feulgen stained histological sections of lung carcinoma are quantitatively described by MST-based and morphometric features of tumor cell nuclei, allowing analysis of growth centers and micro-environment conditions in human lung cancer in relation to tumor texture. Another study in [130] presents analysis to differentiate Feulgen stained primary and metastatic lung carcinoma in the intrapulmonary lymph nodes using a combination of nuclear features measured using integrated optical density (IOD) related to DNA content of nuclei and MST derived features. Similar work in [131] measuring IOD and MST-based features for Feulgen stained human non-small cell lung carcinoma sections shows significant differences in the biopsy specimens as compared to surgically excised carcinomas and intrapulmonary lymph nodes (lymph node metastases). Cytometric features based on DNA-content and morphometry and histometric features based on MST are used in [132] for analyzing prenatal development of lungs, and show discriminative values for different periods of gestation in Feulgen-stained fetal lung specimens. Related studies using fusion of MST based features and other features using syntactic structure analysis for lung carcinomas are also presented in [133], [134], [135] and [136]. In [137], authors present a study to analyze localized fibrous tumors (LFTs) of the lung pleura using quantitative imaging techniques and syntactic structure analysis. Quantitative assessment includes quantification of asbestos intoxication, assessment of immuno-, glyco-, and lectin histochemistry followed by staining assessment, and syntactic structure analysis involves calculation of several parameters derived from MST of cells. This is followed by

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

a statistical analysis of features and finding a correlation with clinical data, including habits (smoking), asbestos exposure, survival, and tumor recurrence in patients. A recent study involving a combination of texture analysis and syntactic structure analysis is presented in [138], where the authors use entropy calculations dependent on grey value thresholds along with MST representations to evaluate image content information in histopathological images stained using different staining techniques such as anti Her2_new and glycohistochemical staining.

Feulgen stained tissue sections of malignant mesothelioma, hyperplastic mesothelium and pulmonary adenocarcinoma are analyzed in [139] where a total of 82 features including texture features, morphometric features, densitometry features and graph-based features based on Voronoi diagram, Gabriel graph and MST are extracted from histological sections. kNN classification is performed and it is observed that texture and graph-based features outperform the densitometry and morphometry features. Another related study for determination of tumour prognosis in three types of tumors namely, invasive cervical carcinoma, colorectal carcinoma and malignant mesothelioma for immunostained tissue sections is given in [140]. In this study, spatial arrangement of blood vessels is represented using Voronoi diagram, Gabriel graph and minimum spanning tree. A set of features including vessel-derived features (density and morphometry), fractal features using fractal analysis and graph-based features using syntactic structural analysis are extracted and automated kNN classifier is used to find most discriminative feature combinations. They conclude that for all three types of cancers, prediction of prognosis based on syntactic structure analysis yields higher recognition scores compared to vessel density and fractal dimension features.

In [141], the authors analyze H&E stained prostate cancer images for classification into different Gleason grades. They extract a hybrid set of 102 features including 57 texture features, 32 gland morphology and arrangement features and 13 nuclear arrangement features. Spatial arrangement of nuclei is studied by constructing Voronoi diagram, Delaunay graph and MST on nuclei, and spatial arrangement of glands is studied by constructing co-adjacency matrix, as a low number of glands per region does not yield informative graph representations. SVM is applied to discriminate between pairs of tissue types from Gleason grade 3 adenocarcinoma, Gleason grade 4 adenocarcinoma, benign epithelium and benign stroma and results are discussed. Another related work is [142], where H&E stained breast biopsy images are studied. A total of 3400 textural and architectural features are extracted. Texture features include gray level statistical features, Haralick features and Gabor features, and architectural features include Voronoi diagram, Delaunay graph and MST based features. Spectral clustering algorithm called graph embedding is used for dimensionality reduction followed by SVM

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

classification. This method facilitates classification of breast tissue images into cancerous or non-cancerous and low grade or high grade cancer. A study to discriminate between normal, abnormal and cofounder classes of H&E stained prostate biopsy samples is given in [143], where nuclei are first detected and segmented, after which features are extracted in the form of architectural features based on Voronoi diagram, Delaunay triangluation, MST and nuclear density, and texture features based on first-order statistics, co-occurence features and steerable filters. A cascaded approach consisting of a set of binary classifiers is used to classify images into seven classes namely epithelium, stroma, atrophy, prostatic intraepithelial neoplasia, Gleason grades 3, 4 and 5. Content-based image retrieval of prostate cancer images using hybrid features is described in [144]. It involves extraction of 576 features, including 483 texture features, 44 morphological features, and 49 graph features based on Voronoi diagram, Delaunay triangulation and MST, followed by manifold learning.

A method for quantification of extent of lymphocytic infiltration in H&E stained Her2-positive (Her2+) breast carcinoma images is mentioned in [145] and also explained in detail in [146]. Firstly, automatic lymphocyte detection is performed to identify lymphocyte nuclei from surrounding stroma and cancer cell nuclei using a combination of region growing and Markov random field algorithms. Subsequently, Voronoi diagram, Delaunay triangulation and MST are constructed on these images. 50 architectural features are extracted from the images including 25 graph-based and 25 nuclear features. Nuclear features represent nuclear statistics, such as lymphocyte density, nearest neighbor and distribution statistics. Graph embedding is used for dimensionality reduction and SVM classification method is applied. Comparative evaluation is performed with texture-based features (Varma-Zisserman texton-based features and global texture features), showing superiority of architectural feature set over both texture feature sets. Related work is proposed in [147] for distinguishing modified Bloom-Richardson grades from H&E stained estrogen receptor-positive (ER+) whole slides of breast cancer. Extracted feature set includes 50 nuclear architecture features with 25 graph-based features from Voronoi diagram, Delaunay triangulation and MST and 25 nuclear statistic features, and 13 Haralick texture features. Feature selection is performed using a minimum redundancy maximum relevance method and a whole slide classifier is introduced to extract features from multiple fields of view of varying sizes, followed by classification of large images. Evaluation results show that the grading system is able to distinguish low versus high, low versus intermediate and intermediate versus high grade specimens with reasonable accuracy.

In [148], Feulgen stained bladder carcinoma histological tissue sections are analyzed using object based, texture and graph based features. MST is generated from the

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

segmented image using Kruskal's algorithm connecting the nuclei and geometric and densitometric features are derived from the trees. The combined features are used for performing image classification based on grading (grades 1, 2A, 2B and 3) using Linear discriminant analysis and multi-layer back-propagation neural networks (BPNN) methods and results are evaluated.

H&E and Thionin-Feulgen stained cervical cancer biopsy specimens are analyzed in [149], where 103 nucleometric features from morphometry, photometry and texture, and 29 architectural features from graphs including Voronoi diagrams, Delaunay graphs and minimum spanning trees are extracted. Linear discriminant analysis is performed to classify nuclei into different types (negative, koilocytosis, CIN1, CIN2, CIN3) and to detect the presence of Human papillomavirus(HPV).

Breast biopsy whole slide images are analyzed using high and low resolution algorithms embedded in a multi-scale framework in [150]. The human visual system is simulated where receptive fields are selected using first order color features and second order color features based on Gaussian distributions. These receptive fields are used for training GPU-based SVM for classifying regions into normal or invasive. Voronoi diagrams are constructed incrementally, which are used in nuclear pleomorphism detection to select the most informative samples in their multi-scale dynamic sampling algorithm.

A system for interactive classification and retrieval of microscopic tissue images is proposed in [151] where the authors use low level features like color, texture and shape to describe regions and high level features like statistics and spatial relationships represented by fuzzy membership functions to define neighborhoods of regions. Their framework learns prototype regions in a collection of images with the help of model-based clustering and density estimation. Then it performs retrieval of tissue regions similar to user query regions with optional relevance feedback. The authors also present a Bayesian framework for classification of images using their feature set and models. This method does not use graphs, but uses a combination of different feature types to represent tissue regions and their relationships.

A hybrid approach using cell graphs already discussed in Section 3.3 is given in [39].

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

3.7 Graph-Theoretic Frameworks

In this section, we mention some of the commonly used and freely available frameworks and tools for developing graph-based applications. A standardized generic interface for graph traversal is provided by Boost Graph Library (BGL) to the C++ developers, which is open source and encourages the use of graph-based algorithms and data structures [152]. Quickgraph, an open source library based on BGL in C#, provides generic graph data structures and algorithms for the .NET platform [153]. OpenGM is another open source library in C++ for defining discrete graphical models and distributive operations on these models [154]. An additional open source graph template library in C++ is the Library for Efficient Modeling and Optimization in Networks (LEMON) for combinatorial optimization tasks, especially those involving graphs and networks [155]. An open source graph modeling and visualization framework written in Java is the Java Universal Network/ Graph Framework (JUNG) [156]. Computational Geometry utility in Mathematics feature [157], Bioinformatics Toolbox [158] and Symbolic Math Toolbox [159] in MATLAB also provide several functions used in graph theory, as a part of MATLAB proprietary software.

Modern pathology information systems are often .NET applications, developed using C# programming. So far, most of the graph-theoretic analysis methods in histopathology are based on Voronoi diagram, Delaunay triangulation and their subgraphs, hence, we provide an overview of some past implementations and libraries for these methods with focus on C#. Fortune-voronoi [160] implementation follows Fortune's sweepline algorithm for Voronoi diagram described in [19]. EmguCV [161] is a wrapper for .NET to access the OpenCV library [162] and the triangulation in OpenCV is formed iteratively [163]. OpenCV functions can be used for implementing Voronoi diagram, Delaunay triangulation and related graphs like relative neighborhood graph and nearest neighbor graph. NetTopologySuite [164] is a part of the JTS Topology Suite [165], [166] adapted to .NET or C#. JTS Topology Suite is an application program interface (API) in Java for two-dimensional linear geometry, and can be used for implementing Voronoi diagram, Delaunay triangulation and constrained Delaunay triangulation (CDT), based on the method described in [167] and modifications from [168]. Nielsen's Bourke [169] is a .NET 2.0 library created by Morton Nielsen based on a method by Paul Bourke [170], to triangulate point data incrementally i.e. to an existing triangulation more nodes are added and the triangulation is updated. Boost.Polygon Voronoi library [171], a part of the Boost framework [172], provides implementations of Voronoi diagram and Delaunay triangulation in C++ using Fortune's algorithm. Poly2Tri [173] is a library for production of constrained Delaunay triangulation, and is available in C++ and Java. The triangulation

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

is carried out with a sweepline procedure described in [174]. Table 2 summarizes the various available implementations of Voronoi diagram, Delaunay triangulation and related graphs described above.

4 Conclusion and Recommendations

In the current practice of medicine, histopathological examination of tissue slides is the routinely used method for the diagnosis and prognosis of diseases, as the disease characteristics can be identified mainly from tissue appearance. Attempts have been made to automate the process of analyzing digital histopathological images using a wide range of analysis techniques. This paper presents a comprehensive account of the various architectural methods developed using graph-theoretic descriptions of histological images. We discuss the progress of exciting developments and applications

Table 2: Often used implementations of Voronoi diagram, Delaunay triangulation and related graphs

Framework	**Associated graphs**	**Base algorithms**	**Programming language**	**Provider**	**Open source**	**License**
Fortune-voronoi	Voronoi diagram	Fortune's sweepline algorithm	C#	https://code.google.com/p/fortune-voronoi/	Yes	Mozilla Public License 1.1
OpenCV planar subdivisions	Voronoi diagram, Delaunay triangulation and subgraphs	Delaunay's algorithm	C++,C# (EmguCV wrapper)	http://docs.opencv.org/modules/legacy/doc/planar_subdivisions.html	Yes	BSD License (EmguCV: GNU GPL License v3)
NetTopology Suite	Voronoi diagram, Delaunay triangulation and Constrained Delaunay triangulation	Guibas & Stolfi method and Dani Lischinki method	C#	https://github.com/NetTopologySuite/NetTopologySuite	Yes	GNU Lesser General Public License
JTS Topology Suite	Voronoi diagram, Delaunay triangulation and Constrained	Guibas & Stolfi method and	Java ^{TM}	http://www.vividsolutions.com/jts/JTSHome.htm	Yes	GNU Lesser General Public License

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

	Delaunay triangulation	Dani Lischinki method				
Nielsen's Bourke	Delaunay triangulation	Paul Bourke's method	C#	http://paulbourke.net/papers/triangulate/morten.html	Yes	None
Boost.Polygon Voronoi Library	Voronoi diagram and Delaunay triangulation	Fortune's sweepline algorithm	C++	http://www.boost.org/doc/libs/1_54_0/libs/polygon/doc/voronoi_main.htm	Yes	Boost Software License, Version 1.0
Poly2Tri	Constrained Delaunay triangulation	Sweepline algorithm for CDT	C++, Java ^{TM}	https://code.google.com/p/poly2tri/	Yes	New BSD License

of the graph-based techniques in histology, starting from the mathematically complex but less computationally efficient tessellations like Dirichlet and Johnson Mehl tessellations, followed by studies on proximity graphs, especially Delaunay graphs and their subgraphs. Subsequently, graphs based on alternative neighborhood conditions like O'Callaghan and zone of influence were explored, reporting their benefits over the Voronoi-based graphs. Meanwhile, there was also an increased interest in syntactic pattern recognition methods deriving quantitative information from graph representations, mostly utilizing minimum spanning trees. However, with further developments in virtual microscopy techniques, we note a paradigm shift from the 'constrained' or geometric methods utilizing proximity graphs to more 'relaxed' or flexible application-specific graphs, for example cell graphs and attributed relational graphs applied on different types of tissues. Many authors have also used a combination of the graph-based methods with other familiar image analysis methods, and report superior performance of the hybrid methods over classical approaches.

To the best of our knowledge, the graph-theoretic techniques explored for histological image analysis as mentioned in this paper, have been applied on standard sized individual regions of smaller dimensions and lower resolutions (usually previously selected regions of interest). With the rapid advancement in hardware and software equipment in the field of digital pathology in the past few years, slides are now available as high resolution whole slide images ready to be analyzed. The histopathological community has observed the need of developing analysis techniques applicable to the whole slide images and not only to some specific regions. The state-of-the art solutions

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

using graphs are too computationally expensive to be applied on large whole slide images. Another limitation of many of the current graph-based approaches is their semi-automatic nature, as cells or other interesting objects have been mostly interactively marked with their boundaries before graph construction, which also makes the process time consuming. For some other graph-based methods, the corresponding tessellations provide cell boundaries which may not coincide with the actual cell boundaries. Sometimes simpler and more general segmentation algorithms like thresholding and k-means algorithm have been applied, leading to blob-like appearance of cells and cell clusters which can clearly affect the accuracy of subsequent analysis. Hence, there is a need to develop tissue specific automatic segmentation algorithms for isolating the different tissue components in histological images.

For future research in this direction, we suggest that the current graph-based algorithms should be modified for analysis of whole slide images, for instance, by using advanced parallel processing techniques. Attempts towards algorithm optimization in histological image analysis for complete slides can prove pivotal in shaping the future of digital pathology. An approach based on parallel processing frameworks for sharpness adjustment of WSIs using distributed computing mechanisms is proposed in [175], and this system can be further explored to process, analyze and parallelize analysis of WSIs. Moreover, segmentation and object detection methods in histological images should be completely automated and optimized for histological image datasets. One recently developed example of fully automatic nuclei segmentation strategy as a minimum model approach, for H&E stained images is given in [176]. Additionally, there is a need to consider the information at different magnifications of histological image in a multi-resolution or hierarchical manner, in order to capture properties at both cellular and tissue levels. A multi-scale framework using parallel processing techniques is presented in [150], where sparse coding and GPU programming are explored and dynamic sampling techniques are used to rapidly identify fields of interest and analyze breast biopsy histopathological WSIs. The authors report their computational times on WSIs comparable to processing times of pathologists. Another recent multi-resolution approach to combine visually significant information at different magnifications to improve segmentation results in gastric cancer images is described in [177].

We can conclude that graph-based methods have proved their ability as potentially very powerful tools for describing and analyzing tissue architectures in digital histopathology, and need to be explored further. It can be stated that graph-based techniques comprise an important direction of study and research in the field of histopathogical image analysis, and can provide the basis for developing various automatic applications and tools to retrieve and classify tissue sections in a reliable way.

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

Acknowledgment

This work is supported with funds from the German Academic Exchange Service (DAAD).

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

[1] Weidner N, Cote RJ, Suster S, and Weiss LM. Modern Surgical Pathology: 2-Volume Set, Expert Consult-Online & Print. Elsevier Health Sciences, 2009.

[2] Rolls G. An introduction to specimen preparation. http://www.leicabiosystems.com/pathologyleaders/an-introduction-to-specimen-preparation/, May 2011.

[3] Pawlina W and Ross M. *Histology: A Text and Atlas*. Lippincott Williams & Wilkins, Baltimore, MD, 2006.

[4] Rochow TG and Tucker PA. Introduction to microscopy by means of light, electrons, X-rays, or acoustics. Springer Science & Business Media, 1994.

[5] Sucaet Y and Waelput W. *Digital Pathology*. SpringerBriefs in Computer Science. Springer, 2014.

[6] Trudeau RJ. *Introduction to Graph Theory*. Dover Publications, New York, 1993.

[7] Lejeune Dirichlet G. Über die Reduction der positiven quadratischen Formen mit drei unbestimmten ganzen Zahlen. *Journal für die reine und angewandte Mathematik*, 40:209–227, 1850.

[8] Voronoi G. Nouvelles applications des paramètres continus á la thèorie des formes quadratiques. *Journal für die reine und angewandte Mathematik (Crelle's Journal)*, 133:198-287, 1907.

[9] Toussaint GT. Some unsolved problems on proximity graphs. In Dearholt D and Harary F, editors, Proceedings of the First Workshop on Proximity Graphs. Memoranda in Computer and Cognitive Science MCCS, pages 91–224. Citeseer, 1991.

[10] Aurenhammer F and Klein R. Voronoi diagrams. Handbook of computational geometry, 5:201– 290, 2000.

[11] Rozenberg G and Salomaa A. *Current trends in theoretical computer science: essays and tutorials*, volume 40. World Scientific, 1993.

[12] Delaunay B. Sur la sphère vide. A la mémoire de Georges Voronoï. *Bulletin of Academy of Sciences of the USSR*, (6):793–800, 1934.

[13] Gabriel KR and Sokal RR. A new statistical approach to geographic variation analysis. *Systematic Biology*, 18(3):259–278, 1969.

[14] Matula DW and Sokal RR. Properties of Gabriel graphs relevant to geographic variation research and the clustering of points in the plane. *Geographical analysis*, 12(3):205–222, 1980.

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

[15] Toussaint GT. The relative neighbourhood graph of a finite planar set. *Pattern recognition*, 12(4):261–268, 1980.

[16] Graham RL and Hell P. On the history of the minimum spanning tree problem. *Annals of the History of Computing*, 7(1):43–57, 1985.

[17] Czumaj A and Sohler C. Testing euclidean minimum spanning trees in the plane. *ACM Transactions on Algorithms (TALG)*, 4(3):31, 2008.

[18] Preparata FP and Shamos MI. Introduction. In *Computational Geometry*, pages 1–35. Springer, 1985.

[19] Fortune S. A sweepline algorithm for voronoi diagrams. *Algorithmica*, 2(1-4):153–174, 1987.

[20] Shamos MI and Hoey D. Closest-point problems. In *Foundations of Computer Science, 1975. 16th Annual Symposium on*, pages 151–162. IEEE, 1975.

[21] Green PJ and Sibson R. Computing dirichlet tessellations in the plane. The Computer Journal, 21(2):168–173, 1978.

[22] Okabe A, Boots B, Sugihara K, and Chiu SN. *Spatial tessellations: concepts and applications of Voronoi diagrams*, volume 501. John Wiley & Sons, 2009.

[23] Lloyd S. Least squares quantization in PCM. *Information Theory, IEEE Transactions on*, 28(2):129–137, 1982.

[24] Bowyer A. Computing dirichlet tessellations. *The Computer Journal*, 24(2):162–166, 1981.

[25] Watson DF. Computing the n-dimensional delaunay tessellation with application to voronoi polytopes. *The computer journal*, 24(2):167–172, 1981.

[26] Lawson CL. Software for C^1 Surface Interpolation. In Mathematical Software III (J. Rice, editor), pages 161–194, 1977.

[27] Zimmer H. Voronoi and delaunay techniques. *Proceedings of Lecture Notes, Computer Sciences*, 8, 2005.

[28] Lingas A. A linear-time construction of the relative neighborhood graph from the delaunay triangulation. Computational Geometry, 4(4):199–208, 1994.

[29] Eppstein D. Spanning trees and spanners. *Handbook of computational geometry*, pages 425–461, 1999.

[30] Edelsbrunner H. *Algorithms in combinatorial geometry*, volume 10. Springer Science & Business Media, 1987.

[31] Kirkpatrick DG and Radke J. A framework for computational morphology. In Toussaint GT, editor, *Computational Geometry, Machine Intelligence and Pattern Recognition*, volume 2, pages 217–248. Elsevier, North-Holland, 1985.

[32] Jaromczyk JW and Toussaint GT. Relative neighborhood graphs and their relatives. *Proceedings of the IEEE*, 80(9):1502–1517, 1992.

[33] Johnson W and Mehl R. Reaction Kinetics in processes of nucleation and growth. *Transactions of American Institute of Mining and Mettalurgical Engineers*, 135:416–458, 1939.

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

[34] Anton F, Mioc D, and Gold CM. The voronoi diagram of circles and its application to the visualization of the growth of particles. In *Transactions on Computational Science III*, pages 20–54. Springer, 2009.

[35] Anton F, Mioc D, and Gold CM. Dynamic additively weighted voronoi diagrams made easy. In *Proceedings of the 10th Canadian Conference on Computational Geometry (CCCG'98)*, 1998.

[36] O'Callaghan JF. An alternative definition for neighborhood of a point. *IEEE Transactions on Computers*, 24(11):1121–1125, 1975.

[37] Kayser K. Neighborhood condition and application of syntactic structure analysis in histopathology. *Acta Stereol*, 6:373–384, 1987.

[38] Gunduz C, Yener B, and Gultekin S. The cell graphs of cancer. In *ISMB/ECCB (Supplement of Bioinformatics)*, pages 145–151, 2004.

[39] Bilgin C, Demir C, Nagi C, and Yener B. Cell-graph mining for breast tissue modeling and classification. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 5311–5314. IEEE, 2007.

[40] Sanfeliu A and Fu KS. A distance measure between attributed relational graphs for pattern recognition. *Systems, Man and Cybernetics, IEEE Transactions on*, (3):353–362, 1983.

[41] Sharma H, Alekseychuk A, Leskovsky P, Hellwich O, Anand R, Zerbe N, and Hufnagl P. Determining similarity in histological images using graph-theoretic description and matching methods for content-based image retrieval in medical diagnostics. *Diagnostic pathology*, 7(1):134, 2012.

[42] Wallis WD. *A beginner's guide to graph theory*. Springer Science & Business Media, 2010.

[43] West DB. *Introduction to Graph Theory*. Prentice Hall, 2 edition, September 2000.

[44] Volkmann L. Estimations for the number of cycles in a graph. *Periodica Mathematica Hungarica*, 33(2):153–161, 1996.

[45] Haralick RM, Shanmugam K, and Dinstein IH. Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions on*, 3(6):610–621, 1973.

[46] Brouwer AE and Haemers WH. *Spectra of graphs*. Springer Science & Business Media, 2011.

[47] Wilson RC, Hancock ER, and Luo B. Pattern vectors from algebraic graph theory. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(7):1112–1124, 2005.

[48] Luo B, Wilson RC, and Hancock ER. Spectral embedding of graphs. *Pattern recognition*, 36(10):2213–2230, 2003.

[49] Dusser C, Rasigni M, Palmari J, Rasigni G, Llebaria A, and Marty F. Minimal spanning tree analysis of biological structures. *Journal of theoretical biology*, 125(3):317–323, 1987.

[50] Dussert C, Rasigni G, and Llebaria A. Quantization of directional properties in biological structures using the minimal spanning tree. *Journal of theoretical biology*, 135(3):295–302, 1988.

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

[51] Taylor TJ and Vaisman II. Graph theoretic properties of networks formed by the delaunay tessellation of protein structures. *Physical Review E*, 73(4):041925, 2006.

[52] Samudrala R and Moult J. A graph-theoretic algorithm for comparative modeling of protein structure. *Journal of molecular biology*, 279(1):287–302, 1998.

[53] Patra S and Vishveshwara S. Backbone cluster identification in proteins by a graph theoretical method. *Biophysical Chemistry*, 84(1):13–25, 2000.

[54] Godreche C, Kostov I, and Yekutieli I. Topological correlations in cellular structures and planar graph theory. *Physical review letters*, 69(18):2674, 1992.

[55] Dumay AC, van der Geest RJ, Gerbrands JJ, Jansen E, and Reiber J. Consistent inexact graph matching applied to labelling coronary segments in arteriograms. In Pattern Recognition, 1992. Vol. III. Conference C: Image, Speech and Signal Analysis, Proceedings, 11*th IAPR International Conference on*, pages 439–442. IEEE, 1992.

[56] Ravasz E, Somera AL, Mongru DA, Oltvai ZN, and Barabási AL. Hierarchical organization of modularity in metabolic networks. *science*, 297(5586):1551–1555, 2002.

[57] Schaller G and Meyer-Hermann M. Multicellular tumor spheroid in an off-lattice voronoi- delaunay cell model. *Physical Review E*, 71(5):051910, 2005.

[58] Barabasi AL and Oltvai ZN. Network biology: understanding the cell's functional organization. Nature Reviews Genetics, 5(2):101–113, 2004.

[59] Aittokallio T and Schwikowski B. Graph-based methods for analysing networks in cell biology. *Briefings in bioinformatics*, 7(3):243–255, 2006.

[60] Fu KS. *Syntactic methods in pattern recognition*. Elsevier, 1974.

[61] Van Diest P, Kayser K, Meijer G, and Baak J. Syntactic structure analysis. *Pathologica*, 87(3):255–262, 1995.

[62] Prewitt J and Wu S. An application of pattern recognition to epithelial tissues. In *Proceedings of the Annual Symposium on Computer Application in Medical Care*, page 15. American Medical Informatics Association, 1978.

[63] Prewitt JM. Interactive decision-making for picture processing. In *Decision and Control including the 16th Symposium on Adaptive Processes and A Special Symposium on Fuzzy Set Theory and Applications, 1977 IEEE Conference on*, pages 373–379. IEEE, 1977.

[64] Prewitt JM. Graphs and grammars for histology: An introduction. In Proceedings of the Annual Symposium on Computer Application in Medical Care, page 18. American Medical Informatics Association, 1979.

[65] Sanfeliu A. An application of a distance measure between graphs to the analysis of muscle tissue patterns. School of Electrical Engineering, Purdue University, 1981.

[66] Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, and Yener B. Histopathological image analysis: A review. Biomedical Engineering, IEEE Reviews in, 2:147–171, 2009.

[67] Darro F, Kruczynski A, Etievant C, Martinez J, Pasteels JL, and Kiss R. Characterization of the differentiation of human colorectal cancer cell lines by means of voronoi diagrams. Cytometry, 14(7):783–792, 1993.

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

[68] Keenan SJ, Diamond J, Glenn McCluggage W, Bharucha H, Thompson D, Bartels PH, and Hamilton PW. An automated machine vision system for the histological grading of cervical intraepithelial neoplasia (CIN). The Journal of pathology, 192(3):351–362, 2000.

[69] Altunbay D, Cigir C, Sokmensuer C, and Gunduz-Demir C. Color graphs for automated cancer diagnosis and grading. *Biomedical Engineering, IEEE Transactions on*, 57(3):665–674, 2010.

[70] Raymond E, Raphael M, Grimaud M, Vincent L, Binet JL, and Meyer F. Germinal center analysis with the tools of mathematical morphology on graphs. Cytometry, 14(8):848–861, 1993.

[71] Meijer G, Van Diest P, Fleege J, and Baak J. Syntactic structure analysis of the arrangement of nuclei in dysplastic epithelium of colorectal adenomatous polyps. Analytical and quantitative cytology and histology/the International Academy of Cytology [and] American Society of Cytology, 14(6):491–498, 1992.

[72] van Diest PJ, Fleege JC, and Baak J. Syntactic structure analysis in invasive breast cancer: analysis of reproducibility, biologic background, and prognostic value. Human pathology, 23(8):876–883, 1992.

[73] Coleman K, Van Diest PJ, Baak J, and Mullaney J. Syntactic structure analysis in uveal melanomas. British journal of ophthalmology, 78(11):871–874, 1994.

[74] Zaitoun A, Al Mardini H, and Record C. Quantitative assessment of gastric atrophy using the syntactic structure analysis. *Journal of clinical pathology*, 51(12):895–900, 1998.

[75] Kayser K, Görtler J, Borkenfeld S, and Kayser G. How to measure diagnosis-associated information in virtual slides. *Diagnostic Pathology*, 6(1):1–9, 2011.

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

[77] Guillaud M, MacAulay CE, Le Riche JC, Dawe C, Korbelik J, and Lam S. Quantitative architectural analysis of bronchial intraepithelial neoplasia. In *BiOS 2000 The International Symposium on Biomedical Optics*, pages 74–81. International Society for Optics and Photonics, 2000.

[78] Landini G and Othman IE. Architectural analysis of oral cancer, dysplastic, and normal epithelia. Cytometry Part A, 61(1):45–55, 2004.

[79] Basavanhally A, Agner S, Alexe G, Bhanot G, Ganesan S, and Madabhushi A. Manifold learning with graph-based features for identifying extent of lymphocytic infiltration from high grade, her2+ breast cancer histology. Image Anal. Appl. Biol.(in Conjunction MICCAI), New York [Online]. Available: http://www.miaab.org/miaab-2008-papers/27-miaab-2008-paper-21.pdf, 2008.

[80] Ulam S. Patterns of growth of figures: Mathematical aspects. In Kepes G, editor, Module, Proportion, Symmetry, Rhythm, pages 64–74. Braziller, 1966.

[81] Sudbø J, Bankfalvi A, Bryne M, Marcelpoil R, Boysen M, Piffko J, Hemmer J, Kraft K, and Reith A. Prognostic value of graph theory-based tissue architecture analysis in carcinomas of the tongue. *Laboratory investigation*, 80(12):1881–1889, 2000.

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

[82] Sudbø J, Marcelpoil R, and Reith A. New algorithms based on the Voronoi Diagram applied in a pilot study on normal mucosa and carcinomas. *Anal Cell Pathol.*, 21(2):71–86, 2000.

[83] Sudbø J, Marcelpoil R, and Reith A. Caveats: numerical requirements in graph theory based quantitation of tissue architecture. *Analytical Cellular Pathology*, 21(2):59–69, 2000.

[84] Sudbø J, Aamdal S, Reith A, and Sudbø A. Chemoprevention of oral cancer. In *Cancer Chemoprevention*, pages 383–399. Springer, 2005.

[85] Kayser K, Shaver M, Modlinger F, Postl K, and Moyers J. Neighborhood Analysis of Low Magnification Structures (Glands) in Healthy, Adenomatous, and Carcinomatous Colon Mucosa. Pathology-Research and Practice, 181(2):153–158, 1986.

[86] Kayser K. Application of structural pattern recognition in histopathology. In *Syntactic and structural pattern recognition*, pages 115–135. Springer, 1988.

[87] Kayser K, Kiefer B, Toomes H, and Burkhard H. Analysis of adenomatous structures in histopathology. *Analytical and quantitative cytology and histology/the International Academy of Cytology [and] American Society of Cytology*, 9(3):273–278, 1987.

[88] Kayser K, Kiefer B, Burkhardt H, and Shaver M. Syntactic structure analysis of bronchus carcinomas-first results. *Acta Stereol*, 4(2):249–253, 1985.

[89] Kayser K, Fitzer M, Bülzebruck H, Bosslet K, and Drings P. TNM stage, immunohistology, syntactic structure analysis and survival in patients with small cell anaplastic carcinoma of the lung. *Journal of cancer research and clinical oncology*, 113(5):473–480, 1987.

[90] Albert R, Schindewolf T, Baumann I, and Harms H. Three-dimensional image processing for morphometric analysis of epithelium sections. *Cytometry*, 13(7):759–765, 1992.

[91] Bilgin CC, Bullough P, Plopper GE, and Yener B. ECM-aware cell-graph mining for bone tissue modeling and classification. *Data mining and knowledge discovery*, 20(3):416–438, 2010.

[92] Waxman BM. Routing of multipoint connections. *Selected Areas in Communications, IEEE Journal on*, 6(9):1617–1622, 1988.

[93] Demir C, Gultekin SH, and Yener B. Augmented cell-graphs for automated cancer diagnosis. In *ECCB/JBI*, page 12, 2005.

[94] Demir C, Gultekin SH, and Yener B. Spectral analysis of cell-graphs of cancer. *Department of Computer Science, Rensselaer Polytechnic Institute, Troy, New York*, 2004.

[95] Demir C, Gultekin SH, and Yener B. Learning the topological properties of brain tumors. *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, 2(3):262–270, 2005.

[96] Gunduz-Demir C. Mathematical modeling of the malignancy of cancer using graph evolution. *Mathematical biosciences*, 209(2):514–527, 2007.

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

[97] Lund A, Bilgin C, Hasan M, McKeen L, Stegemann J, Yener B, Zaki M, and Plopper G. Quantification of spatial parameters in 3d cellular constructs using graph theory. *BioMed Research International*, 2009, 2009.

[98] McKeen-Polizzotti L, Henderson KM, Oztan B, Bilgin CC, Yener B, and Plopper GE. Quantitative metric profiles capture three-dimensional temporospatial architecture to discriminate cellular functional states. *BMC medical imaging*, 11(1):11, 2011.

[99] Ficsor L, Varga VS, Tagscherer A, Tulassay Z, and Molnar B. Automated classification of inflammation in colon histological sections based on digital microscopy and advanced image analysis. *Cytometry Part A*, 73(3):230–237, 2008.

[100] Ficsor L, Varga V, Berczi L, Miheller P, Tagscherer A, Wu MLc, Tulassay Z, and Molnar B. Automated virtual microscopy of gastric biopsies. *Cytometry Part B: Clinical Cytometry*, 70(6):423–431, 2006.

[101] Ali S, Veltri R, Epstein JA, Christudass C, and Madabhushi A. Cell cluster graph for prediction of biochemical recurrence in prostate cancer patients from tissue microarrays. In *SPIE Medical Imaging*, pages 86760H–86760H. International Society for Optics and Photonics, 2013.

[102] Fatakdawala H, Xu J, Basavanhally A, Bhanot G, Ganesan S, Feldman M, Tomaszewski JE, and Madabhushi A. Expectation–maximization-driven geodesic active contour with overlap resolution (emagacor): Application to lymphocyte segmentation on breast cancer histopathology. *Biomedical Engineering, IEEE Transactions on*, 57(7):1676–1689, 2010.

[103] Kayser K, Sandau K, Böhm G, Kunze KD, and Paul J. Analysis of soft tissue tumors by an attributed minimum spanning tree. *Analytical and quantitative cytology and histology/the International Academy of Cytology [and] American Society of Cytology*, 13(5):329–334, 1991.

[104] Chaudhuri B, Rodenacker K, and Burger G. Characterization and featuring of histological section images. *Pattern recognition letters*, 7(4):245–252, 1988.

[105] Dougherty ER. An introduction to morphological image processing. *Tutorial texts in optical engineering*, 1992.

[106] de Assis Zampirolli F, Stransky B, Lorena AC, and de Melo Paulon FL. Segmentation and classification of histological images-application of graph analysis and machine learning methods. In *Graphics, Patterns and Images (SIBGRAPI), 2010 23rd SIBGRAPI Conference on*, pages 331–338. IEEE, 2010.

[107] Zampirolli FdA. Neighborhood graphs built with morphological operators. *Revista de Informática Aplicada/Journal of Applied Computing*, 4(2), 2010.

[108] Lee G, Sparks R, Ali S, Shih NN, Feldman MD, Spangler E, Rebbeck T, Tomaszewski JE, and Madabhushi A. Co-Occurring Gland Angularity in Localized Subgraphs: Predicting Biochemical Recurrence in Intermediate-Risk Prostate Cancer Patients. *PloS one*, 9(5):e97954, 2014.

[109] Hamilton P, Allen D, Watt P, Patterson C, and Biggart J. Classification of normal colorectal mucosa and adenocarcinoma by morphometry. *Histopathology*, 11(9):901–911, 1987.

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

[110] Hamilton P, Anderson N, Bartels P, and Thompson D. Expert system support using bayesian belief networks in the diagnosis of fine needle aspiration biopsy specimens of the breast. J Clin Pathol, 47(4):329–36, 1994.

[111] Hamilton PW, Bartels PH, Thompson D, Anderson NH, Montironi R, and Sloan JM. Automated location of dysplastic fields in colorectal histology using image texture analysis. *The Journal of pathology*, 182(1):68–75, 1997.

[112] Walker RF, Jackway P, Lovell B, and Longstaff I. Classification of cervical cell nuclei using morphological segmentation and textural feature extraction. In Intelligent Information Systems, 1994. Proceedings of the 1994 Second Australian and New Zealand Conference on, pages 297–301. IEEE, 1994.

[113] Wiltgen M, Gerger A, and Smolle J. Tissue counter analysis of benign common nevi and malignant melanoma. *International journal of medical informatics*, 69(1):17–28, 2003.

[114] Zhao D, Chen Y, and Correa N. Statistical categorization of human histological images. In *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, volume 3, pages III–628. IEEE, 2005.

[115] Yang L, Chen W, Meer P, Salaru G, Goodell LA, Berstis V, and Foran DJ. Virtual microscopy and grid-enabled decision support for large-scale analysis of imaged pathology specimens. *Information Technology in Biomedicine, IEEE Transactions on*, 13(4):636–644, 2009.

[116] Kong J, Sertel O, Shimada H, Boyer K, Saltz J, and Gurcan M. Computer-aided Evaluation of Neuroblastoma on Whole-slide Histology Images: Classifying Grade of Neuroblastic Differentiation. *Pattern Recogn.*, 42(6):1080–1092, June 2009.

[117] Sertel O, Kong J, Shimada H, Catalyurek U, Saltz JH, and Gurcan MN. Computer-aided prognosis of neuroblastoma on whole-slide images: Classification of stromal development. Pattern recognition, 42(6):1093–1103, 2009.

[118] Wong E and Fu K. A Parallel Algorithm for Muscle Tissue Images Classification. In Proceedings/the... Annual Symposium on Computer Application [sic] in Medical Care. Symposium on Computer Applications in Medical Care, pages 751–754. American Medical Informatics Association, 1983.

[119] Díaz G and Romero E. Histopathological image classification using stain component features on a PLSA model. In *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, pages 55–62. Springer, 2010.

[120] Deligdisch L, Kerner H, Cohen C, Dargent D, and Gil J. Morphometric differentiation between responsive tumor cells and mesothelial hyperplasia in second-look operations for ovarian cancer. Human pathology, 24(2):143–147, 1993.

[121] Thiran JP and Macq B. Morphological feature extraction for the classification of digital images of cancerous tissues. *Biomedical Engineering, IEEE Transactions on*, 43(10):1011–1020, 1996.

[122] Nedzved A, Belotserkovsky A, Lehmann T, and Ablameyko S. Morphometrical feature extraction on color histological images for oncological diagnostics. In 5th International Conference on Biomedical Engineering, pages 14–16, 2007.

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

[123] Roula M, Diamond J, Bouridane A, Miller P, and Amira A. A multispectral computer vision system for automatic grading of prostatic neoplasia. In Proceedings IEEE International Symposium on Biomedical Imaging, pages 193–196, 2002.

[124] Rajpoot K and Rajpoot NM. Hyperspectral colon tissue cell classification. MSc DSIP Dissertation, Faculty of Computing Sciences & Engineering, De Montfort University, UK, September 2003.

[125] De Wouwer V, Marck V, Dyck V, et al. Wavelets as chromatin texture descriptors for the automated identification of neoplastic nuclei. Journal of microscopy, 197(1):25–35, 2000.

[126] Jafari-Khouzani K and Soltanian-Zadeh H. Multiwavelet grading of pathological images of prostate. *Biomedical Engineering, IEEE Transactions on*, 50(6):697–704, 2003.

[127] Qureshi H, Sertel O, Rajpoot N, Wilson R, and Gurcan M. Adaptive discriminant wavelet packet transform and local binary patterns for meningioma subtype classification. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008*, pages 196–204. Springer, 2008.

[128] Kayser K and Stute H. Minimum Spanning Tree, Voronoi's Tesselation, Johnson-Mehl Diagrams in Human Lung Carcinoma. Pathology-Research and Practice, 185(5):729–734, 1989.

[129] Kayser K, Stute H, Bubenzer J, and Paul J. Combined morphometrical and syntactic structure analysis as tools for histomorphological insight into human lung carcinoma growth. *Analytical cellular pathology: the journal of the European Society for Analytical Cellular Pathology*, 2(3):167–178, 1990.

[130] Kayser K, Liewald F, Kremer K, and Tacke M. Integrated optical density (IOD), syntactic structure analysis, and survival in operated lung carcinoma patients. *Pathology-Research and Practice*, 190(11):1031–1038, 1994.

[131] Kayser K, Liewald F, Kremer K, Tacke M, Storck M, Faber P, and Bonomi P. Alteration of integrated optical density and intercellular structure after induction chemotherapy and survival in lung carcinoma patients treated surgically. *Analytical and quantitative cytology and histology/the International Academy of Cytology [and] American Society of Cytology*, 16(1):18–24, 1994.

[132] Kayser K, Jacinto S, Böhm G, Fritz P, Kunze W, Nehrlich A, and Gabius H. Application of Computer-assisted Morphometry to the Analysis of Prenatal Development of Human Lung. *Anatomia, histologia, embryologia*, 26(2):135–139, 1997.

[133] Kayser K, Bovin N, Zeng F, Zeilinger C, and Gabius H. Binding capacities to blood-group antigens A, B and H, DNA and MST measurements, and survival in bronchial carcinoma. *Radiol. Oncol*, 28:282–286, 1994.

[134] Kayser K, Kayser C, Rahn W, Bovin NV, and Gabius HJ. Carcinoid tumors of the lung: Immuno-and ligandohistochemistry, analysis of integrated optical density, syntactic structure analysis, clinical data, and prognosis of patients treated surgically. *Journal of surgical oncology*, 63(2):99–106, 1996.

[135] Kayser K, Kayser C, Rahn W, Bovin NV, and Gabius HJ. Carcinoid tumors of the lung: Immuno-and ligandohistochemistry, analysis of integrated optical density, syntactic structure analysis, clinical data, and prognosis of patients treated surgically. *Journal of surgical oncology*, 63(2):99–106, 1996.

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

[136] Kayser K, Richter B, Stryciak R, and Gabius HJ. Parameters derived from integrated nuclear fluorescence, syntactic structure analysis, and vascularization in human lung carcinomas. *Analytical Cellular Pathology*, 15(2):73–83, 1997.

[137] Kayser K, Trott J, Böhm G, Huber M, Kaltner H, André S, and Gabius HJ. Localized fibrous tumors (LFTs) of the pleura: Clinical data, asbestos burden, and syntactic structure analysis applied to newly defined angiogenic/growth-regulatory effectors. *Pathology-Research and Practice*, 201(12):791–801, 2005.

[138] Kayser K, Borkenfeld S, Djenouni A, and Kayser G. Texture and object related image analysis in microscopic images. *Diagnostic Pathology*, 1(1), 2015.

[139] Weyn B, Van De Wouwer G, Koprowski M, Van Daele A, Dhaene K, Scheunders P, Jacob W, and Van Marck E. Value of morphometry, texture analysis, densitometry, and histometry in the differential diagnosis and prognosis of malignant mesothelioma. *The Journal of pathology*, 189(4):581–589, 1999.

[140] Weyn B, Tjalma W, Vermeylen P, Van Daele A, Van Marck E, and Jacob W. Determination of tumour prognosis based on angiogenesis-related vascular patterns measured by fractal and syntactic structure analysis. *Clinical Oncology*, 16(4):307–316, 2004.

[141] Doyle S, Hwang M, Shah K, Madabhushi A, Feldman M, and Tomaszeweski J. Automated Grading of Prostate Cancer using architectural and textural image Features. 4th IEEE International Symposium on Biomedical Imaging, pages 1284–1287, 2007.

[142] Doyle S, Agner S, Madabhushi A, Feldman M, and Tomaszewski J. Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features. In Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on, pages 496–499. IEEE, 2008.

[143] Doyle S, Feldman MD, Shih N, Tomaszewski J, and Madabhushi A. Cascaded discrimination of normal, abnormal, and confounder classes in histopathology: Gleason grading of prostate cancer. BMC bioinformatics, 13(1):282, 2012.

[144] Doyle S, Hwang M, Naik S, Feldman M, Tomaszeweski J, and Madabhushi A. Using manifold learning for content-based image retrieval of prostate histopathology. In *MICCAI 2007 Workshop on Content-based Image Retrieval for Biomedical Image Archives: Achievements, Problems, and Prospects*, pages 53–62. Citeseer, 2007.

[145] Madabhushi A, Doyle S, Lee G, Basavanhally A, Monaco JP, Master SR, Tomaszewski JE, and Feldman MD. Review: Integrated diagnostics: a conceptual framework with examples. *Clinical Chemistry and Laboratory Medicine*, 48:989–998, 2010.

[146] Basavanhally AN, Ganesan S, Agner S, Monaco JP, Feldman MD, Tomaszewski JE, Bhanot G, and Madabhushi A. Computerized image-based detection and grading of lymphocytic infiltration in her2+ breast cancer histopathology. Biomedical Engineering, IEEE Transactions on, 57(3):642–653, 2010.

[147] Basavanhally A, Ganesan S, Feldman MD, Shih N, Mies C, Tomaszewski J, and Madabhushi A. Multi-field-of-view framework for distinguishing tumor grade in ER+ breast cancer from entire histopathology slides. *IEEE Trans. Biomed. Engineering*, 60(8):2089–2099, 2013.

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

[148] Choi HK, Jarkrans T, Bengtsson E, Vasko J, Wester K, Malmstro¨m PU, and Busch C. Image analysis based grading of bladder carcinoma. Comparison of object, texture and graph based methods and their reproducibility. Analytical Cellular Pathology, 15(1):1–18, 1997.

[149] Guillaud M, Cox D, Adler-Storthz K, Malpica A, Staerkel G, Matisic J, Van Niekerk D, Poulin N, Follen M, and MacAulay C. Exploratory analysis of quantitative histopathology of cervical intraepithelial neoplasia: Objectivity, reproducibility, malignancy-associated changes, and human papillomavirus. Cytometry Part A, 60(1):81–89, 2004.

[150] Huang CH, Veillard A, Roux L, Lom´enie N, and Racoceanu D. Time-efficient sparse analysis of histopathological whole slide images. Computerized medical imaging and graphics, 35(7):579– 591, 2011.

[151] Aksoy S, Marchisio GB, Tusk C, and Koperski K. Interactive classification and content-based retrieval of tissue images. In International Symposium on Optical Science and Technology, pages 71–81. International Society for Optics and Photonics, 2002.

[152] Lee LQ and Lumsdaine A. The Boost Graph Library: User Guide and Reference Manual. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2002.

[153] Microsoft. QuickGraph, Graph Data Structures And Algorithms for .NET. https://quickgraph.codeplex.com/documentation, 2011.

[154] Andres B, Beier T, and Kappes JH. OpenGM: A C++ Library for Discrete Graphical Models. CoRR, abs/1206.0111, 2012.

[155] Dezso B, Jüttner A, and Kov´acs P. LEMON–an open source C++ graph template library. Electronic Notes in Theoretical Computer Science, 264(5):23–45, 2011.

[156] Fisher D, Omadadhain J, Smyth P, White S, and Boey YB. Analysis and visualization of network data using JUNG. Journal of Statistical Software, 10(2):1–35, 2005.

[157] MathWorks. Computational Geometry. http://de.mathworks.com/help/matlab/computational-geometry.html/, 1994-2015.

[158] MathWorks. Bioinformatics Toolbox. http://de.mathworks.com/products/bioinfo/, 1994-2015.

[159] MathWorks. Symbolic Math Toolbox. http://de.mathworks.com/products/symbolic/, 1994- 2015.

[160] BenDi. Fortune's Voronoi algorithm implemented in C#. http://www.codeproject.com/KB/ recipes/fortunevoronoi.aspx, 2013.

[161] EmguCV. Emgu CV: OpenCV in .NET (C#, VB, C++ and more), 2013.

[162] Bradski G and Kaehler A. *Learning OpenCV: Computer vision with the OpenCV library*. "O'Reilly Media, Inc.", 2008.

[163] OpenCV. Planar Subdivisions. http://docs.opencv.org/modules/legacy/doc/planar_subdivisions.html, 2010.

[164] NetTopologySuite. NetTopologySuite. https://github.com/NetTopologySuite/NetTopologySuite, 2006.

[165] Davis M. JTS Topology Suite. http://tsusiatsoftware.net/jts/main.html, 2012.

H. Sharma et. al. diagnostic pathology 2015, 1:61
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61

[166] Davis M. Secrets of the JTS Topology Suite. Free and Open Source Software for Geospatial, 2007.

[167] Guibas L and Stolfi J. Primitives for the manipulation of general subdivisions and the computation of voronoi. ACM Transactions on Graphics (TOG), 4(2):74–123, 1985.

[168] Lischinski D. Incremental delaunay triangulation. In Heckbert PS, editor, *Graphics gems IV*, volume 4, pages 47–59. Morgan Kaufmann, 1994.

[169] Nielsen M. Delaunay Triangulation in C#. http://paulbourke.net/papers/triangulate/morten.html, 2006.

[170] Bourke P. Efficient triangulation algorithm suitable for terrain modelling. In Pan Pacific Computer Conference, Beijing, China, 1989.

[171] Sydorchuk A. Boost.Polygon Voronoi Library. http://www.boost.org/doc/libs/1_54_0/libs/polygon/doc/index.htm, 2012.

[172] Demming R and Duffy DJ. Introduction to the Boost C++ Libraries; Volume I-Foundations. Datasim Education BV, 2010.

[173] Liang W. Poly2Tri: Fast and Robust Simple Polygon Triangulation With/Without Holes by Sweep Line Algorithm, 2005.

[174] Domiter V and Zcalik B. Sweep-line algorithm for constrained delaunay triangulation. International Journal of Geographical Information Science, 22(4):449–462, 2008.

[175] Zerbe N, Hufnagl P, and Schlu¨ns K. Distributed computing in image analysis using open source frameworks and application to image sharpness assessment of histological whole slide images. Diagn Pathol, 6(Suppl 1):S16, 2011.

[176] Wienert S, Heim D, Saeger K, Stenzinger A, Beil M, Hufnagl P, Dietel M, Denkert C, and Klauschen F. Detection and Segmentation of Cell Nuclei in Virtual Microscopy Images: A Minimum-Model Approach. *Scientific Reports*, 2, 2012.

[177] Sharma H, Zerbe N, Heim D, Wienert S, Behrens HM, Hellwich O, and Hufnagl P. A Multi- resolution Approach for combining Visual Information Using Nuclei Segmentation and Classification in Histopathological Images. In *Proceedings of the International Conference on Computer Vision Theory and Applications*, pages 37–46, 2015.