

Gland context networks: a novel approach for improving prostate cancer identification

Rodrigo de P. Mendes¹, Xin Yuan², Elizabeth M. Genega³, Xiaoyin Xu⁴, Luciano da F. Costa⁵, and Cesar H. Comin¹

¹*Department of Computer Science, Federal University of São Carlos, São Carlos, SP, Brazil*

²*Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA*

³*Department of Pathology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA**

⁴*Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA and*

⁵*São Carlos Institute of Physics, University of São Paulo, São Carlos, SP, Brazil*

Prostate cancer (PCa) is a pervasive condition that is manifested in a wide range of histologic patterns in biopsy samples. Given the importance of identifying abnormal prostate tissue to improve prognosis, many computerized methodologies aimed at assisting pathologists in diagnosis have been developed. It is often argued that improved diagnosis of a tissue region can be obtained by considering measurements that can take into account several properties of its surroundings, therefore providing a more robust context for the analysis. Here we propose a novel methodology that can be used for systematically defining contextual features regarding prostate glands. This is done by defining a Gland Context Network (GCN), a representation of the prostate sample containing information about the spatial relationship between glands as well as the similarity between their appearance. We show that such a network can be used for establishing contextual features at any spatial scale, therefore providing information that is not easily obtained from traditional shape and textural features. Furthermore, it is shown that even basic features derived from a GCN can lead to state-of-the-art classification performance regarding PCa. All in all, GCNs can assist in defining more effective approaches for PCa grading.

I. INTRODUCTION

Prostate cancer (PCa) is one of the most common types of cancer in men, representing one of the leading causes of cancer related deaths worldwide [1]. Early detection of PCa can improve prognosis and significantly reduce the risk of death [2]. Common approaches for PCa screening include rectal examination, prostate-specific antigen (PSA) testing and MRI [3–5]. If PCa is suspected, a biopsy of the prostate can be performed to confirm the diagnosis [6]. Prostate samples extracted from the biopsy can be used for PCa grading and prognosis using the Gleason Grading System (GGS), a methodology developed by Donald Gleason [7] that has been refined by the International Society of Urological Pathology (ISUP) [8–11].

The GGS provides a set of patterns regarding the prostate tissue that a pathologist can use as reference for PCa prognosis. The patterns are usually related to how well-differentiated the prostate glands are. Non-differentiated glands are usually associated with PCa and thus with higher GGS scores. Currently, GGS scores are evaluated visually by pathologists, which may lead to varying scores between different pathologists, misdiagnosis and can also be time-consuming. Recently, there have been many efforts aimed at providing systematic and unbiased computerized segmentation and classification of prostate tissues for aiding in PCa analysis [12–17].

When identifying PCa in prostate tissue, context can be particularly important. A well-differentiated gland surrounded by non-differentiated glands is probably still

associated with PCa. Thus, it is usual to define properties that are associated with the *context* of a given gland or tissue region. For instance Nguyen et al. [18] considered not only the shapes of the glands for classification, but also the similarity between the shapes of nearby glands. Doyle et al. [19] defined *architectural features* based on graphs constructed from cell nuclei as well as gland centroids. The features included average edge length, edge length variability, as well as Haralick's texture features [20]. Despite their graph construction methodology being similar to our approach, the graphs were not used for providing context for individual gland features, but for defining additional features related only to connectivity. Monaco et al. [21] used Markov random fields for propagating the labeling of cancerous glands in order to identify cancerous regions.

The definition of *ad hoc* contextual features may neglect important information for tissue classification. For instance, if elongated glands tended to appear together in PCa tissue, but gland elongation was measured only individually, their proximity could not be considered for classifying the tissue samples. Defining specific contextual features can be avoided by using, for instance, end-to-end Convolutional Neural Networks (CNN) [12, 15]. **However, the results obtained by CNNs are usually more difficult to interpret than those obtained using handcrafted features.**

In the current work we describe a methodology for systematically assigning context to any feature used for prostate tissue characterization. The context is defined using what we call a *gland context network* (GCN), a weighted graph representing prostate glands and the relationship between their features. We show that many properties calculated based on the generated networks lead to tissue characterization that is not captured by

* Current affiliation: Department of Pathology and laboratory Medicine, Tufts Medical Center, Boston, MA, USA

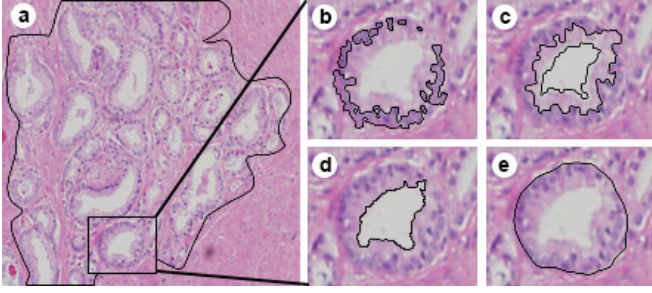


FIG. 1. Glandular structure. (a) A region of prostate tissue and respective glands. For one of the glands, the (b) nuclei, (c) cytoplasm (d) lumen and (e) whole gland are shown.

traditional properties adopted in the literature. Thus, GCNs can be used not only for automated identification of PCa, but also for identifying novel features associated with PCa that may be useful for future revisions of the GGS. To show the potential of the methodology, we create and characterize GCNs for *five* whole-mount prostate samples and use the generated structures to identify cancerous regions. For the considered case study, we identify that altered glands can be detected with high accuracy even by using only a simple network feature, the degree of the nodes.

The remainder of the article is organized as follows. In Section II we present the main features commonly used for PCa classification as well as some basic concepts regarding network theory. The motivation and definition of GCNs is presented in Section III. Section IV contains examples and a discussion about important parameters of GCNs as well as a case study of PCa classification for whole-mount prostate samples. In Section V we present the conclusions.

II. BASIC CONCEPTS

A. Characterizing Gland Shape

The images considered in our analysis were obtained from staining models with hematoxylin and eosin (H&E) from prostate tissues. The main structures to be analyzed in the images are the several glands in the prostate tissue, which contain three main components: nuclei, cytoplasm and lumen. Figure 1 illustrates these structures. Figure 1(a) shows a H&E stained prostate tissue image with one gland highlighted. The nuclei (figure 1(b)) tends to be the darkest structure inside the gland, while the lumen (figure 1(c)) is usually close to the white color. The cytoplasm (figure 1(d)) tends to be stained in a light pink color. The gland is mainly the union of these three structures (figure 1(e)).

The grading system described by Gleason [7] and modified and accepted by the ISUP considers, among other characteristics, the shape and size of the glands to iden-

tify the architectural growth patterns [8, 10, 11]. In the revisions of the GGS made by the ISUP in 2014 and 2019, the shape is mentioned as a predictor for several of the scores. For instance, benign patterns tend to have large circular and oval glands, while Gleason grade 3 patterns usually have smaller and more circular glands when compared to benign patterns. Furthermore, irregular glandular structures can be observed on Gleason grade 4 patterns since some glands are undifferentiated, fused with other glands, or contain many small holes (cribriform patterns) [10, 17].

The GGS provides a comprehensive reference for the many different patterns that can be observed in prostate tissues. Still, the scores are evaluated qualitatively. Thus, it is relevant to define quantitative values for the shape characteristics considered by the specialists in PCa prognosis. Many shape features were used in previous automated PCa prognosis studies [17, 18, 22, 23]. In order to show the potential of the GCC methodology, a set of shape features identified as relevant for PCa prognosis in the aforementioned studies was selected. The considered features and the respective definitions used in this work are:

Area: quantifies the size of the gland. It is calculated as the number of pixels representing the gland. The area is considered an important feature in automated PCa prognosis since it tends to change as the cancer grows [10, 11, 18].

Perimeter: it is associated with the size, and thus with the area, of a gland, but it provides complementary information since two shapes having similar areas can have very distinct perimeters. It is calculated as the arc length of the parametric contour of the gland [24].

Solidity: quantifies how close the shape of the gland is to a convex polygon. It is calculated as the ratio between the area of the gland and the area of the convex hull of the gland [24]. Glands having irregular contours tend to have low solidity.

Eccentricity: represents the elongation of a gland. In this work it is calculated as the ratio between the eigenvalues of the covariance matrix calculated from the positions of the pixels representing the shape [24]. Interestingly, patterns with a score of 5 in the GGS tend to contain glands having cylindrical shapes [10]. The cross-section of such glands can appear as elongated shapes in tissue samples.

Diameter: quantifies the maximum distance between any two points in the contour of a gland.

B. Complex Networks

In the past 20 years complex networks became an important tool for analyzing interconnected systems such as protein-protein interactions [25, 26], social networks [27] and cities [28]. A network or graph can be understood as a topological abstraction of the system in which the elements are represented by nodes [29, 30], with are in-

terconnected by edges if they are related according to some criterion. Such an abstraction provides a common representation for many distinct systems, allowing the respective application of methodologies from network theory for characterizing and modeling their dynamics. For the purpose of defining context networks, two main aspects need to be considered. They are the methodology used for building the network and the properties used for characterizing their nodes. Those aspects are presented and discussed below.

1. Defining Connectivity

A network can be mathematically represented as $G = (V, E)$, where V is the set of nodes and E the set of edges, reflecting the connectivity between pairs of nodes. Therefore, two nodes a, b of V are said to be adjacent, or connected, if $(a, b) \in E$. A network can also be weighted, in which case a value is associated to each of its edges. Usually, the weight is associated with the degree of relationship between the nodes, quantified by some similarity or dissimilarity metric. Networks can also be categorized as being spatial or non-spatial. In spatial networks, each node has a specific position in some metric space, usually the Euclidean space, and the distance between pairs of nodes influences their connectivity. In the case of non-spatial networks, other methodologies for connecting the nodes are considered. The most common approaches include connecting the nodes randomly [31], with preference to large-degree nodes [32] or according to a fixed sequence of the number of connections made by each node [33]. In the following, we only consider spatial networks. Still, we will show that GCN can greatly benefit from defining connectivity from both the spatial distance and the shape and texture similarity of the glands they represent.

Many different approaches can be used for taking into account the distance between nodes when deciding if they should be connected or not. In the experiments we only consider the Geometric Graph criterion [34], explained below, but relevant alternatives are: i) connecting nodes according to a probability that decays exponentially with the distance between the nodes, known as the Waxman criterion [35]; ii) connecting nodes having adjacent Voronoi cells [36]; and iii) connecting pairs of nodes only if a disk that is centered at the midpoint of the two nodes and have a diameter equal to the distance between the nodes do not contain any other node [37]. We will show that the methodology used for defining connectivity can be easily changed when generating context graphs.

For creating a geometric graph, nodes are connected by an edge if they are separated by a distance less than $2r$, where r is a parameter of the model. Therefore, each node can be seen as having a region of influence, given by a disk of radius r centered at the node. Two nodes are connected if their corresponding disks overlap. Parameter k has a large impact on many properties of the

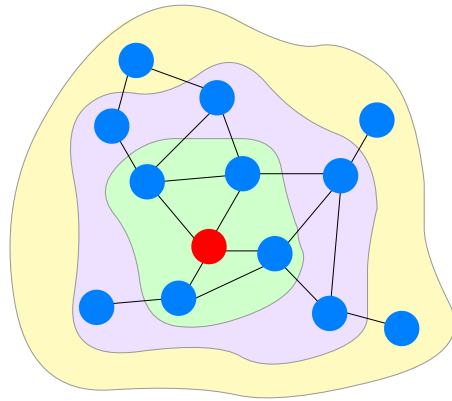


FIG. 2. Illustration of the different scales that can be considered for characterizing a network node, which is shown in red. The green region indicates immediate neighbors of the reference node. The purple and yellow regions contain nodes that are, respectively, second and third neighborhoods of the reference node.

generated network. Particularly, the expected number of connections of a node is proportional to r .

2. Network Properties

Given a network, it is relevant to characterize the nodes according to the respective topology, that is, the pattern of connections made by the nodes. In order to do so, we first define two key concepts. The *shortest path length* between two nodes V_i and V_j is calculated as the number of edges in the smallest path between the two nodes, where a path is defined as a sequence of edges in the network. The *k-th neighborhood* of a given node V_i is defined as the set of nodes that can be reached from node V_i using paths with length smaller than or equal to k .

When characterizing the connectivity of nodes, not only the immediate connections of the node need to be considered. For instance, it might be interesting to measure the number of connections among the neighbors of a node, or among the neighbors of the neighbors of that node. If only nodes in the first neighborhood of a node are considered, the respective characterization will be very local, that is, it will concern only nodes directly associated with the reference node (see green region in Figure 2). Contrariwise, if a large neighborhood is considered, nodes indirectly associated with the reference node will be also taken into account (see yellow region in Figure 2). This will be a key concept in Context Networks, since the neighborhood used for characterizing the node will set the desired context.

Many topological features can be used for characterizing the connectivity of a network [38]. In the following, we describe some basic features that will be considered for evaluating the potential of GCNs.

Degree: The number of connections made by a node.

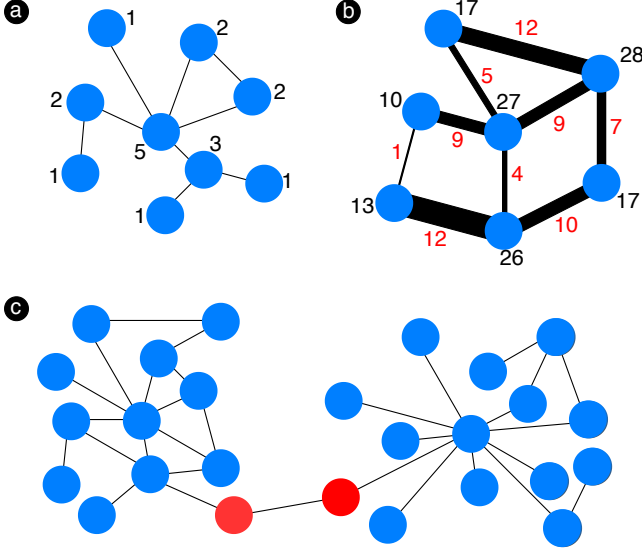


FIG. 3. Example of network properties: (a) *node degree*; (b) *edge weight* (in red) and *node strength*; and (c) *betweenness centrality*, red nodes have large betweenness.

It is usually associated with the importance, or centrality, of a node since nodes with large degrees are associated with many other nodes in the network. The probability distribution of the node degrees, known as the degree distribution, can reveal many important properties about the network [38]. An illustration of this property is shown in Figure 3(a).

Strength: Defined for weighted networks, it is calculated as the sum of weights of the edges connected to a node. It can be considered as the degree counterpart for weighted networks. Thus, it can also be interpreted as a measure of importance of a node [38]. Figure 3(b) shows a weighted network and the corresponding edge weights and node strengths.

Betweenness centrality: The shortest path length between two nodes is of great importance since it underlies the concept of proximity for nodes in a network. The betweenness centrality of a node V_i is associated with the number of shortest paths between any two nodes in the network that pass through node V_i [38]. Nodes with large betweenness can usually be seen as “bridges” between different groups of nodes in the network. Figure 3(c) shows an example of nodes with large betweenness.

III. GLAND CONTEXT NETWORKS

As presented in Section II A, glands can be characterized using shape and texture features. However, in the GGS it is stated that gland appearance should not be considered isolatedly [11, 18, 39], that is, the context of the gland is also important. For instance, a single gland having an appearance that might be associated

with healthy glands, if surrounded by undifferentiated glands, will still be associated with a high Gleason score. For such reasons, some methodologies for defining context have been developed [18, 19, 21]. Yet, the context is usually not taken into account in a systematic fashion, or it is only considered for specific properties.

Here we define a systematic approach for setting the context of a gland using what we call a *Context Network*. Glands are represented as nodes, and connections indicate some relationship between glands. This relationship can be spatial, in the sense that the distances between the glands are used when defining the connectivity, may consider only the shape or texture of the glands, that is, similar glands tend to be connected, or it can be a mixture of both aspects. The obtained network can then be characterized using complex network properties [38], which are used as additional features for PCa prognosis.

Let S represent the set of glands in a sample, where the i -th gland is referred as s_i . Suppose m shape and texture features $f_1^i, f_2^i, \dots, f_m^i$ have been measured for each gland s_i . These features can be represented by a vector $\tilde{f}^i$. Each gland also has a position $\vec{p} = (p_1^i, p_2^i)$. In the following, we consider that the position is given by the centroid of the gland.

The spatial distance between two glands i and j is calculated as

$$d_{ij}^s = \sqrt{\sum_{k=1}^2 (\tilde{p}_k^i - \tilde{p}_k^j)^2} \quad (1)$$

where $\tilde{p}_k^i$ is the standard score of the k -th coordinate of gland i , that is,

$$\tilde{p}_k^i = \frac{p_k^i - \mu_k}{\sigma_k} \quad (2)$$

where μ_k and σ_k are, respectively, the average and standard deviation of the k -th coordinate of all the glands. The purpose of the normalization is to allow the definition of a general distance involving both the spatial coordinates of the glands and their shape and texture features.

The difference between the features of two glands is calculated as

$$d_{ij}^f = \sqrt{\sum_{k=1}^m (\tilde{f}_k^i - \tilde{f}_k^j)^2} \quad (3)$$

where $\tilde{f}_k^i$ represents the standard score of f_k^i , defined using the same approach describe above.

A combination of the distance between the glands and their shape and texture similarity is then calculated as

$$h_{ij} = \sqrt{\alpha (d_{ij}^s)^2 + (1 - \alpha) (d_{ij}^f)^2}, \quad (4)$$

where $\alpha \in [0, 1]$.

The weight of the connection between glands i and j is then given by

$$w_{ij} = e^{-h_{ij}}. \quad (5)$$

Parameter α sets the relative importance between the spatial distance and the feature similarity between the glands. When $\alpha = 1$, only the spatial distance is taken into account, while $\alpha = 0$ creates a context network in which edge weights are defined using only the appearance of the glands. It is important to note that the context network can be created for all features or using only a subset of them. For instance, it is possible to create an area context network, considering only this feature of the glands.

A GCN can be created using different combinations of distances d_{ij}^s and d_{ij}^f . For instance, it is possible to use one value of α for defining the connectivity between nodes and another value for defining edge weights. Here we consider an approach for generating a network that is easy to visualize and interpret. The connections between nodes are defined using the geometric graph approach, presented in Section II B 1. Thus, the connectivity is defined using $\alpha = 1$ and a given radius r that is a parameter of the method. Next, edge weights are defined using $\alpha = 0$, that is, only considering shape and texture features of the glands. This generates a network where only the distances between the glands define the existence or absence of edges and the weights of the edges depends only on the similarity between the appearance of the glands. Still, it is important to keep in mind that GCNs allow alternative approaches for building the network.

Having obtained the network, the nodes can be characterized using the properties discussed in Section II B 2. Therefore, the original feature vectors used for classifying glands or samples into healthy or unhealthy can be systematically enhanced by considering the context of the glands. Figure 4 illustrates the advantage of using GCNs for adding context for PCa analysis. The usual steps applied for PCa classification or GGS assessment are represented in green. GCNs can be added to the analysis for augmenting the original shape and texture features.

It is important to note that GCNs cannot be used for characterizing latter stages of PCa since the glands become undifferentiated. Thus, GCNs are aimed at providing early prognosis and also in revealing novel properties for tissues containing differentiated or partially differentiated glands.

IV. EXPERIMENTS

In the following we provide examples of GCNs created using prostate tissue samples. We describe important properties that can be obtained from the generated networks and also provide an example of identifying PCa.

A. Dataset

The dataset used in the examples consists of sections from [five](#) whole-mount prostate samples. The slides were imaged by a brightfield automatic microscope, which was configured to scan over the whole specimen and form a large mosaic picture composed of 40 to 50 individual images. Figure 5 shows one of the images in the dataset. The original sample is shown in Figure 5(a). A specialist analyzed the sample and identified two regions evaluated at a score of 3+3 in the GGS for PCa, which are indicated in red in the image. Figure 5(b) shows a binary image containing manually segmented glands. We note that the focus of the methodology is on classifying individual glands. The [five](#) samples together contain 9491 gland units, which was deemed enough for assessing the potential of GCNs in representing and characterizing prostate glands.

B. Gland Context Network Creation

Following the procedure described in Section III, GCNs were created based on manual segmentations of the glands in whole-mount prostate samples. Figure 6 shows an example of GCN visualized without considering edges weights, only the connectivity between glands. The connectivity alone can already reveal many interesting facts about the glands and their context. First, it is clear that glands tend to organize into clusters, defining subgraphs containing nodes that are more connected between themselves than with nodes that are not in the subgraph. Such clusters are commonly called *communities* in network theory [40]. Communities tend to be associated with components of a system that carry out similar functions. Many community detection methods have been developed in the literature [40] and could be used for identifying clusters of glands. This procedure could bring additional insights regarding gland organization when compared to the identification of connected components. The latter has been used for defining context in prostate tissues [18].

The generated GCNs are also very heterogeneous. Some regions contain a large number of connections while others contain many nodes having small degree. Examples of two highly distinct regions can be seen in Figures 6(a) and (b). It is important to note that variations on edge density tend to be the most noticeable characteristic when visualizing GCNs, but these networks may contain many other important information that is not immediately identified by visual inspection. For instance, the clustering coefficient [38], which, roughly speaking, is associated to the number of triangles connected to a node, varies throughout the network.

Besides connectivity, GCNs can also represent similarity between glands. Figure 7 shows two regions of a weighted GCN created using the method presented in Section III. Thicker edges are associated to edges having

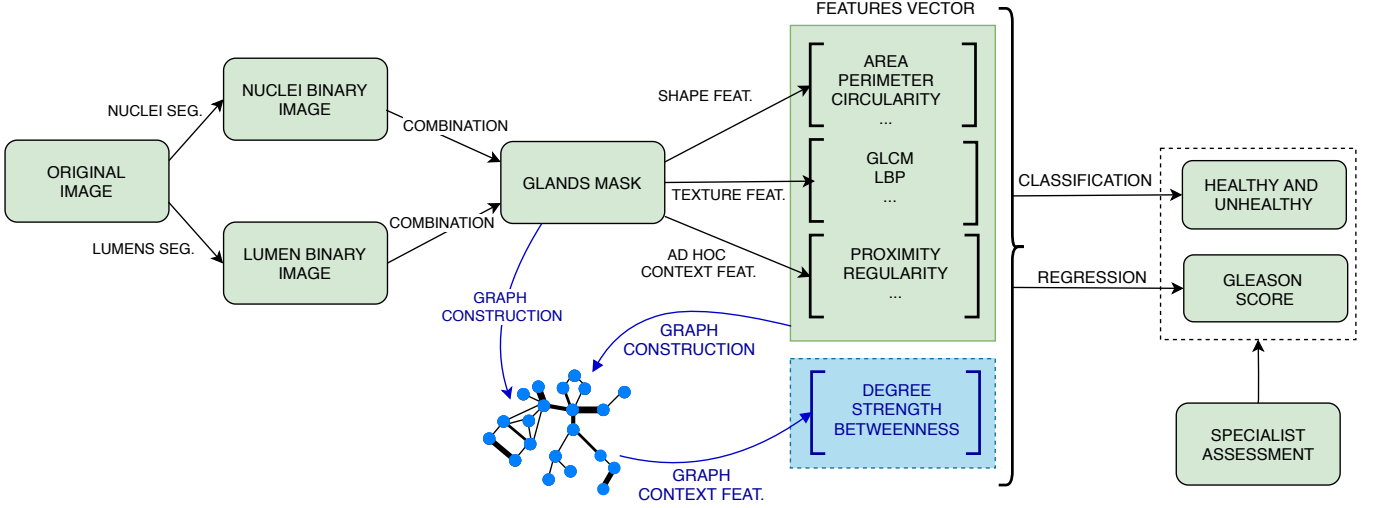


FIG. 4. Usual procedure for characterizing prostate glands, augmented by a context network. Given an input sample, the nuclei and lumina of the glands are segmented and then combined to define a binary image containing the detected glands. A number of shape, texture and ad hoc context features can then be calculated and used for classifying the tissue into healthy or unhealthy or, alternatively, the features can be used for identifying the Gleason Score of the sample. The Gland Context Network can be used for improving this procedure. From the constructed network, additional context features (indicated in blue) can be derived from the shape and texture features, and may lead to better PCa prognosis.

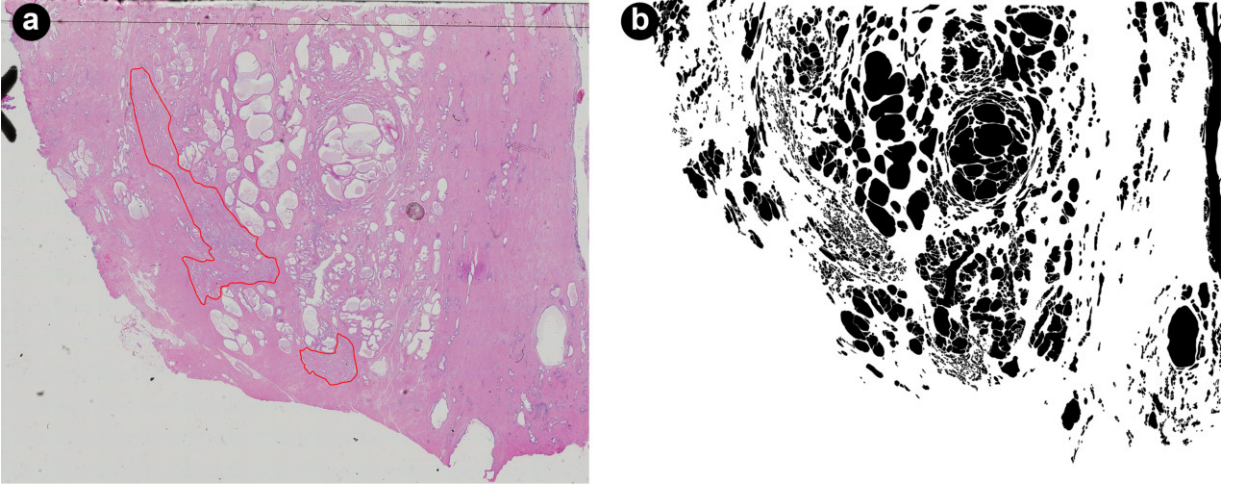


FIG. 5. Example of images used in the analysis. (a) Original whole mount prostate sample with 3+3 graded regions, according to the GGS, marked in red. (b) Binary image showing manually segmented glands.

larger weight, and thus to glands having similar shapes. In this case, all shape properties described in Section II A were used for defining the weights according to Equation 5. Among other properties, the GCN reveals clusters of glands that are close to each other and have similar shapes.

As mentioned before, GCNs can be created using any set of gland properties. The main purpose being to allow a systematic incorporation of context to the considered properties. The context is created by calculating complex networks features based on the generated GCN. For instance, the degree of a node incorporates a local con-

text regarding neighboring glands. Figure 8 shows the degrees of the nodes in the GCN created from the sample in Figure 5. Interestingly, larger degree nodes tend to be located inside or near the region demarcated by the specialist (shown in Figure 5), suggesting a relationship between node degree and PCa.

The degree can reveal local context, but other network measurements could be used for augmenting glands properties at medium, global, or even specific scales [41]. Such measurements may reveal additional information about gland properties that are not easily quantifiable using traditional methods. As a consequence, GCN derived

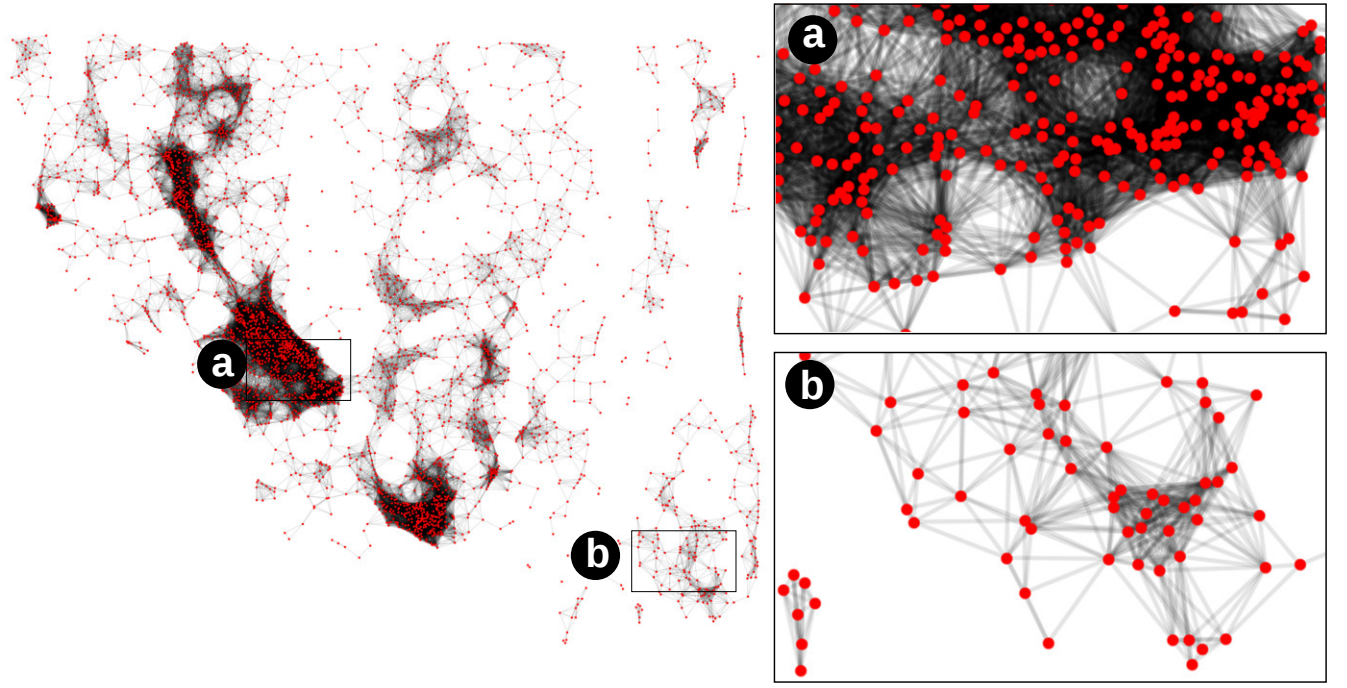


FIG. 6. (a) Geometric network generated from the sample shown in Figure 5. Red points indicate the centroids of glands. Zoomed in views of the network are shown in (a) and (b). A radius of $r = 350$ pixels was used for generating the network.

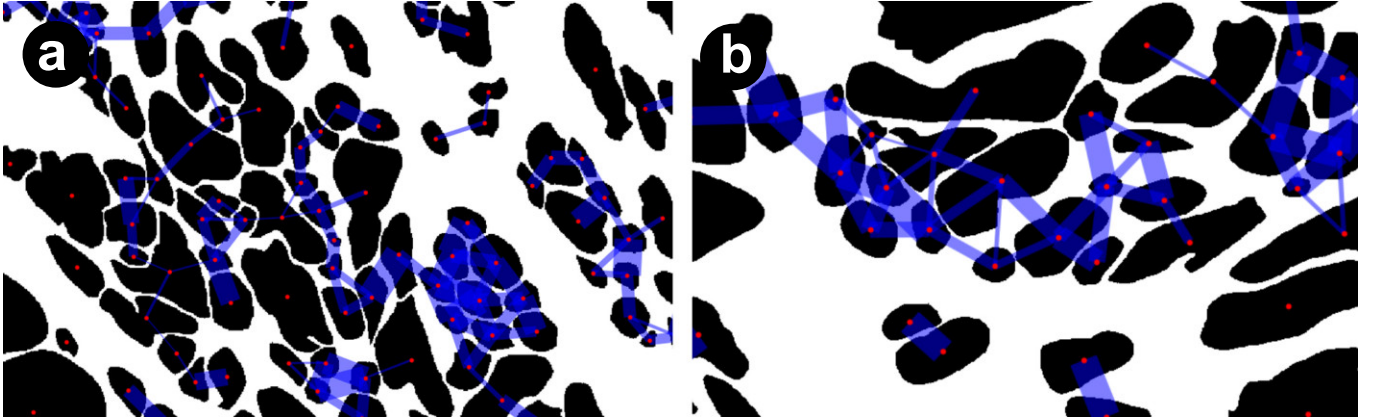


FIG. 7. Visualization of representative regions of the Gland Context Network generated from the sample shown in Figure 5. All gland shape properties described in Section II.A.1 were used for defining edges weights. Notice that thicker edges correspond to glands that have similar shapes.

properties may improve the accuracy of PCa identification and also of automated Gleason score assignment.

C. A case study: Gland Classification Using Context Networks

As an example of the potential of using GCNs for identifying PCa, we employed GCN properties for classifying glands into healthy and unhealthy (GGS score of 3+3, as indicated by the specialist). The classification, done

at the individual gland level, can then be used for identifying abnormal tissue regions. Different sets of properties were obtained and used for classification: i) shape features (SF), ii) network features (NF), iii) all features (AF), and iv) only the degree feature (DF).

In order to quantify the performance of the method, we measured the precision, recall, specificity and accuracy of the classification [42]. Glands belonging to the region demarcated by the specialist and that were classified as unhealthy defined the true positive (TP) samples. The FN, TN and FP samples were defined accord-

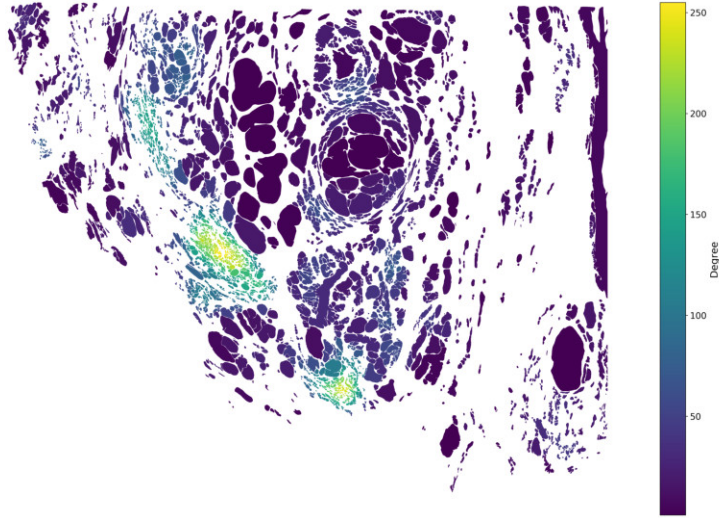


FIG. 8. Visualization of the degrees of the nodes in the Gland Context Network. Each gland is colored according to the degree of its respective node. Notice that the majority of high-degree glands are located inside the cancer region delimited in Figure 5.

ingly. For classifying the glands, we considered the K -nearest neighbors, Support Vector Machine (SVM) and Multilayer Perceptron classifiers. Each method was applied using 5-fold cross-validation. Each fold of the cross-validation contained the same number of healthy and unhealthy glands. A grid search of the methods parameters was applied for searching their respective optimal configurations. The three methods provided similar classification performance. Thus, in the following we present only the results obtained for the K -nearest neighbors classifier with $K = 5$.

Figure 9 shows a typical result obtained from the cross-validation using the whole set of features. The colors differentiate between TP, FN, TN and FP glands. It is clear that the vast majority of glands are correctly classified. Interestingly, most of the incorrectly classified glands are near the border of the region demarcated by the specialist, indicating that these glands might already have been affected by PCa.

The performance metrics of gland classification were measured for five whole-mount samples and the four sets of properties described above. The average value obtained for each metric is shown in Figure 10. It is clear that features derived from the GCNs improved the classification when compared to using the gland shape features alone. The average accuracy obtained when using the network feature set was around $93\% \pm 2.5\%$, which is better or equivalent to state-of-the-art results obtained by other approaches such as 91.9% by Nir et al. [13],

85.4% by Doyle et al. [19], 95.14% by Naik et al. [22], and 79% by Nguyen et al. [18]. However, it is important to note that, with the exception of Naik et al., the other approaches used automatically segmented glands in the classification.

Figure 10 also shows that the degree alone led to a good classification performance. The degree tends to be higher in regions containing small glands in large numbers. Thus, the GCN methodology revealed, in a straightforward fashion, that these two properties were important for identifying tissue regions affected by PCa in the analyzed samples.

Segmenting glands is a hard task. Changes in the segmentation might influence the generation of the GCN, which in turn will influence the performance of classification tasks. In order to verify the impact that changes in the segmentation can have in the gland classification, four types of perturbations were applied to the manual annotations of the glands: i) binary erosion; ii) binary dilation; iii) binary dilation while avoiding mergings of the glands and iv) random removal of glands. The erosion and dilation procedures represent situations where the segmented glands are smaller or larger than they should be. Procedure ii) tends to generate very large connected components due to the dilation procedure, which is unrealistic in real situations. On the other hand, procedure iii) leads to dilated glands but the number of glands do not change. The random removal of glands represents situations where glands are missed either by an expert

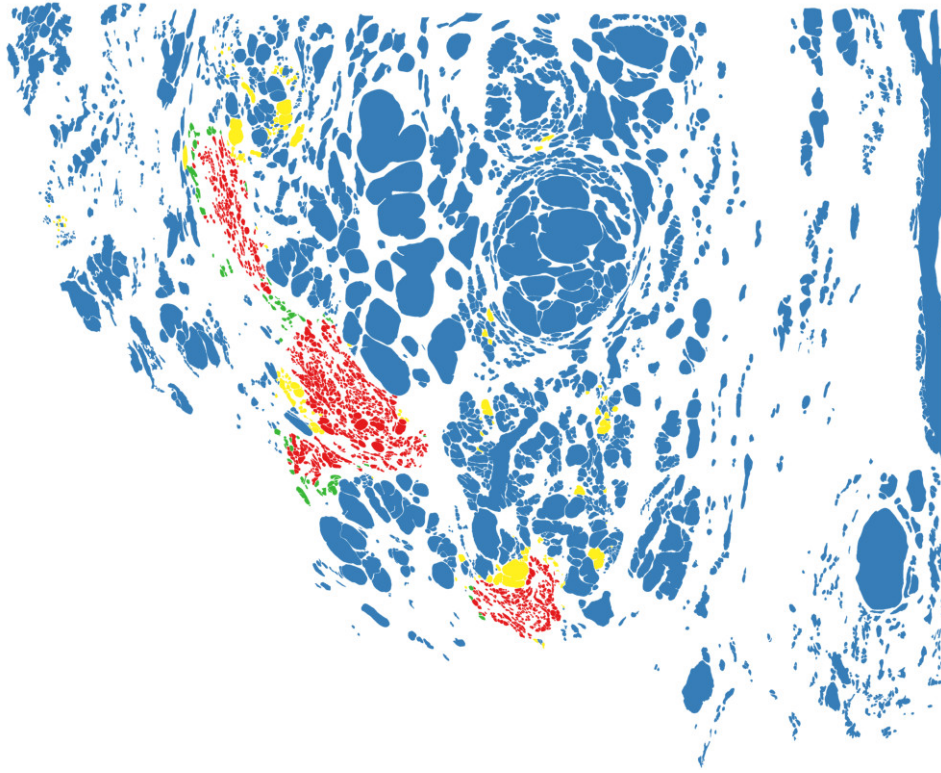


FIG. 9. Results of gland classification into the healthy and unhealthy categories using network context features. Glands are colored based on correct or incorrect classifications. Blue: true negatives, red: true positives, yellow: false positives, green: false negatives.

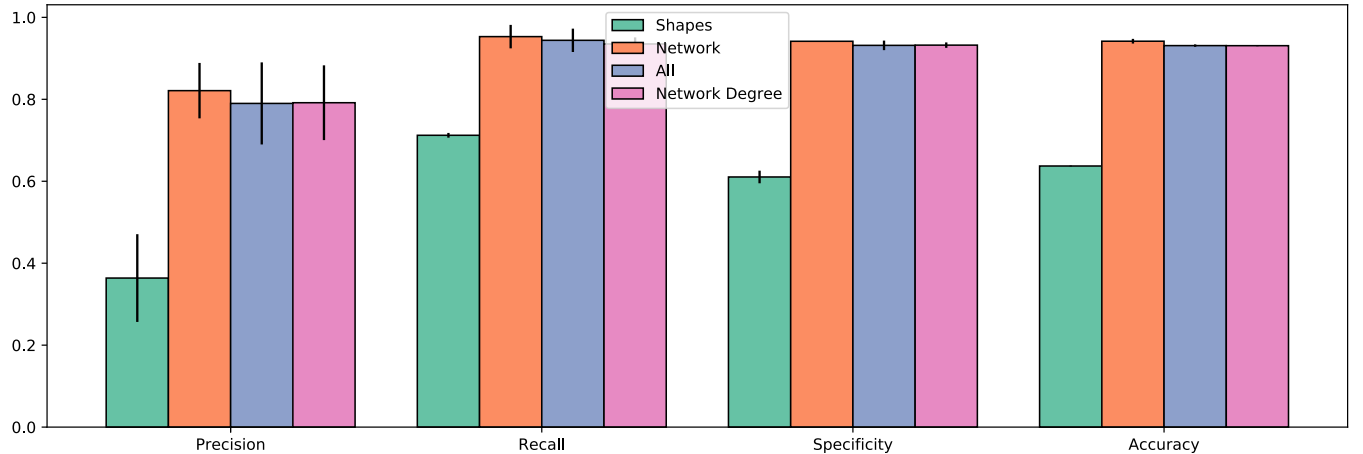


FIG. 10. Average gland classification precision, recall, specificity and accuracy when using different sets of features for characterizing the glands. Vertical lines indicate the standard deviation of the performance metrics.

annotator or by automated segmentation procedures.

Figure 11 shows the performance metrics obtained for different amounts of erosion, dilation and gland removal iterations applied to the segmented glands. For the erosion and dilation, the performance is measured as a function of the radius of the structuring element used in the

respective morphological operation. Regarding the erosion (Figure 11(a)), the method is robust to erosions of up to a radius of 13 pixels. We observed that, for this radius, the segmented glands are much smaller than the actual glands in the image. This can be verified by the disappearance of many glands for radii larger than 13

pixels, as seen in the plot. For instance, for a radius of 15 pixels roughly 15% of the glands disappear. Given that the method shows good performance even when the segmentation of the glands has low quality (e.g., for a radius of 13 pixels), we consider that the method is robust to the erosion perturbation.

Regarding the robustness of the method with respect to the dilation operation (Figure 11(b)), it is clear that small amounts of dilation can significantly decrease the performance of the classification. In particular, the precision becomes much lower after only 5 dilations. On the other hand, as explained above, since the glands are very close to each other, the dilation procedure leads to large connected components covering almost the whole sample, which do not realistically represent glands or possible errors in manual annotations. This can be seen by the rapid decrease in gland number as the dilation becomes larger. Figure 11(c) shows the result for the dilation procedure while avoiding gland merges. In this case, there is almost no change in the classification performance. This happens because, as mentioned before, features related to the spatial distribution of the glands tended to be more relevant for classification than those associated with their shapes. Since the dilation without gland merges does not change the spatial distribution of the glands, the performance remains high.

The results obtained for the random removal of glands are shown in Figure 11(d). Surprisingly, a good performance is obtained even when around 60% of the glands are removed. This is so because when glands are randomly removed with uniform probability, their overall spatial distribution does not change. For instance, after the removal, regions having a high density of glands will still have relatively large density when compared to other regions in the sample. The information regarding the spatial distribution of the remaining glands is enough to correctly classify them.

The results of the perturbation analysis indicate that, at least for the considered data, GCNs are robust with respect to the considered changes in the shape of the glands as well as to situations where some glands are not correctly segmented.

A challenging aspect for quantifying the performance of tumor detection methodologies is that it is not straightforward to define precise boundaries of a tumor. This can be a source of variation on the ground truth gland classes used for measuring the performance of the method. In order to verify to which extent the manual annotation of the tumor regions influenced our results, an intra-expert evaluation procedure was employed. The results of the methodology were shown to the specialist that originally identified the tumor regions in our samples. The annotations were re-evaluated by the specialist considering the TP, FN, TN and FP glands identified by our analysis. The specialist's assessment was that TP and TN glands were indeed correctly classified. Interestingly, the specialist found that the majority of FN glands, that is, glands inside the delimited tumor region

but that were classified by our method as being normal (green glands in Figure 9), were indeed normal glands. Those glands should be considered normal because, despite being inside or close to the tumor region, they had a more disperse spatial distribution and still had basal cells. Furthermore, many glands that were close to, but outside, the cancerous region and were classified as being abnormal by the method (yellow glands in Figure 9) were suspicious of being tumor acini (atypical acini), given their small sizes and small distances to each other. An absence of basal cells on these glands was also observed. Thus, the method successfully identified abnormal glands that were previously not included in the tumor region.

V. CONCLUSIONS

The automated or semi-automated characterization of prostate tissue can provide important support for pathologists to identify abnormal tissues. Furthermore, analyzing prostate images at a large scale, using whole-mount samples, may reveal important factors for grading PCa that have not been accounted for in the most up-to-date revision of the ISUP guidelines [8]. As a consequence, some automated methods have been developed for segmenting, characterizing and classifying prostate tissue [12–15, 17]. Since the developed methods tend to have an automated segmentation as a first step, the remaining analyses becomes highly dependent on the quality of the segmentation. For instance, it may be difficult to identify if a poor performance on identifying abnormal tissue was due to inadequate features or if it was caused by tissue not being correctly segmented.

In this work we assumed that the glands have already been correctly segmented, either manually or automatically, and focused on a methodology for characterizing the identified glands. The concept of a Gland Context Network (GCN), a weighted graph describing the spatial and visual relationship between the glands, was presented. We argued that one of the main benefits of using GCNs is that the vast literature regarding complex networks characterization becomes immediately available for augmenting individual gland measurements.

The proposed methodology can be used for a more comprehensive characterization of tissues by integrating information from varying spatial scales. As an illustration, we created GCNs for whole-mount samples and showed that heterogeneous networks with many interesting properties are created and can be further investigated. In particular, clusters of glands are revealed and might be identified using community detection algorithms. Furthermore, we showed that even in a straightforward gland classification procedure, using an off-the-shelf classifier, GCN properties led to the identification of the vast majority of abnormal glands. The classification performance was verified to be high even when the glands are not correctly segmented.

For future analyses, GCNs could be defined and ap-

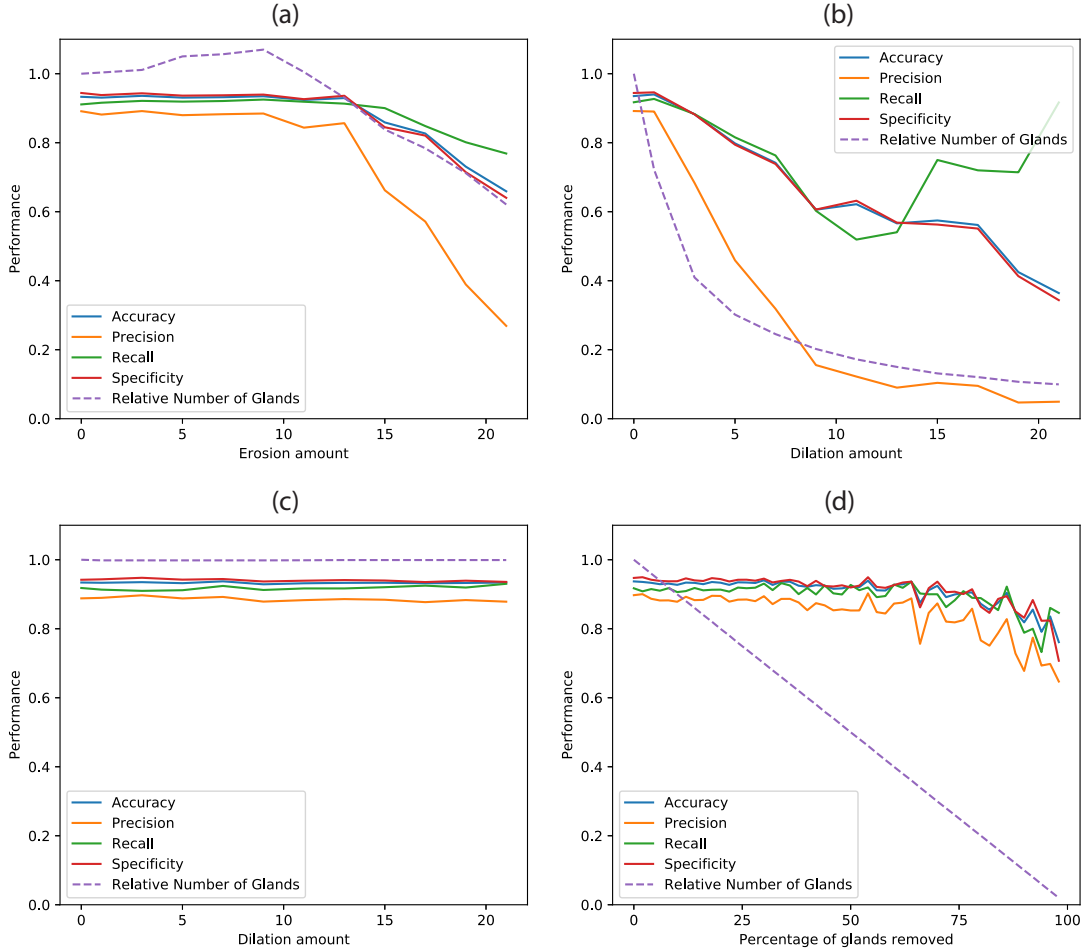


FIG. 11. Performance of the gland classification task as a function of the amount of perturbation applied to manual gland segmentations. The perturbations were: (a) erosion, (b) dilation, (c) dilation without gland merges, and (d) random removal of glands. The dashed line indicates the number of glands after the perturbation divided by the number of glands when no perturbation was applied to the images.

plied for a large prostate tissue dataset for the identification of novel contextual properties associated to different Gleason scores. Furthermore, GCNs could be used for further improving the classification of the glands using a two step classification scheme, applied as follows. As can be seen in Figure 9, glands that were incorrectly classified are usually surrounded by correctly classified glands. Thus, one could apply a consensus formation dynamics, such as opinion spreading [27, 43], to propagate gland categories throughout the GCN, which would correct most of the misclassifications. The methodology could also be applied to other types of tissues, provided

that appropriate features are used for building the GCN.

ACKNOWLEDGMENTS

Cesar H. Comin thanks FAPESP (grant no. 18/09125-4) for financial support. Luciano da F. Costa thanks CNPq (grant no. 307085/2018-0) for sponsorship. This work has also been supported by the FAPESP grant 15/22308-2.

[1] Freddie Bray, Jacques Ferlay, Isabelle Soerjomataram, Rebecca L Siegel, Lindsey A Torre, and Ahmedin Jemal. Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185

countries. *CA: a cancer journal for clinicians*, 68(6):394–424, 2018.

[2] Martin CS Wong, William B Goggins, Harry HX Wang, Franklin DH Fung, Colette Leung, Samuel YS Wong,

- Chi Fai Ng, and Joseph JY Sung. Global incidence and mortality for prostate cancer: analysis of temporal patterns and trends in 36 countries. *European Urology*, 70(5):862–874, 2016.
- [3] Fritz H Schröder, Jonas Hugosson, Monique J Roobol, Teuvo LJ Tammela, Stefano Ciatto, Vera Nelen, Maciej Kwiatkowski, Marcos Lujan, Hans Lilja, Marco Zappa, et al. Screening and prostate-cancer mortality in a randomized european study. *New England Journal of Medicine*, 360(13):1320–1328, 2009.
 - [4] Gerald L Andriole, E David Crawford, Robert L Grubb III, Sandra S Buys, David Chia, Timothy R Church, Mona N Fouad, Edward P Gelmann, Paul A Kvale, Douglas J Reding, et al. Mortality results from a randomized prostate-cancer screening trial. *New England Journal of Medicine*, 360(13):1310–1319, 2009.
 - [5] Saradwata Sarkar and Sudipta Das. A review of imaging methods for prostate cancer detection: supplementary issue: image and video acquisition and processing for clinical applications. *Biomedical engineering and computational biology*, 7:BECB–S34255, 2016.
 - [6] Amit R Patel and J Stephen Jones. Optimal biopsy strategies for the diagnosis and staging of prostate cancer. *Current Opinion in Urology*, 19(3):232–237, 2009.
 - [7] DF Gleason. The Veteran’s Administration Cooperative Urologic Research Group: histologic grading and clinical staging of prostatic carcinoma. In M Tannenbaum, editor, *Urologic Pathology: The Prostate*, pages 171–198. Philadelphia: Lea and Febiger, 1977.
 - [8] Geert JLH van Leenders, Theodorus H van der Kwast, David J Grignon, Andrew J Evans, Glen Kristiansen, Charlotte F Kweldam, Geert Litjens, Jesse K McKenney, Jonathan Melamed, Nicholas Mottet, et al. The 2019 international society of urological pathology (isup) consensus conference on grading of prostatic carcinoma. *The American Journal of Surgical Pathology*, 2020.
 - [9] Oleksandr N Kryvenko and Jonathan I Epstein. Prostate cancer grading: a decade after the 2005 modified Gleason grading system. *Archives of Pathology & Laboratory Medicine*, 140(10):1140–1152, 2016.
 - [10] Jonathan I Epstein, Lars Egevad, Mahul B Amin, Brett Delahunt, John R Srigley, and Peter A Humphrey. The 2014 international society of urological pathology (isup) consensus conference on gleason grading of prostatic carcinoma. *The American Journal of Surgical Pathology*, 40(2):244–252, 2016.
 - [11] Jonathan I Epstein, William C Allsbrook Jr, Mahul B Amin, Lars L Egevad, ISUP Grading Committee, et al. The 2005 international society of urological pathology (isup) consensus conference on gleason grading of prostatic carcinoma. *The American Journal of Surgical Pathology*, 29(9):1228–1242, 2005.
 - [12] S Larry Goldenberg, Guy Nir, and Septimiu E Salcudean. A new era: artificial intelligence and machine learning in prostate cancer. *Nature Reviews Urology*, 16(7):391–403, 2019.
 - [13] Guy Nir, Soheil Hor, Davood Karimi, Ladan Fazli, Brian F Skinnider, Peyman Tavassoli, Dmitry Turbin, Carlos F Villamil, Gang Wang, R Storey Wilson, et al. Automatic grading of prostate cancer in digitized histopathology images: Learning from multiple experts. *Medical Image Analysis*, 50:167–180, 2018.
 - [14] Malay Singh, Emarene Mationg Kalaw, Danilo Medina Giron, Kian-Tai Chong, Chew Lim Tan, and Hwee Kuan Lee. Gland segmentation in prostate histopathological images. *Journal of Medical Imaging*, 4(2):027501, 2017.
 - [15] Eirini Arvaniti, Kim S Fricker, Michael Moret, Niels Rupp, Thomas Hermanns, Christian Fankhauser, Norbert Wey, Peter J Wild, Jan H Rueschoff, and Manfred Claassen. Automated gleason grading of prostate cancer tissue microarrays via deep learning. *Scientific Reports*, 8(1):1–11, 2018.
 - [16] Scott Doyle, Michael Feldman, John Tomaszewski, and Anant Madabhushi. A boosted bayesian multiresolution classifier for prostate cancer detection from digitized needle biopsies. *IEEE Transactions on Biomedical Engineering*, 59(5):1205–1218, 2010.
 - [17] Kien Nguyen, Bikash Sabata, and Anil K Jain. Prostate cancer grading: Gland segmentation and structural features. *Pattern Recognition Letters*, 33(7):951–961, 2012.
 - [18] Kien Nguyen, Anindya Sarkar, and Anil K Jain. Structure and context in prostatic gland segmentation and classification. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 115–123. Springer, 2012.
 - [19] Scott Doyle, Mark Hwang, Kinsuk Shah, Anant Madabhushi, Michael Feldman, and John Tomaszewski. Automated grading of prostate cancer using architectural and textural image features. In *Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on*, pages 1284–1287. IEEE, 2007.
 - [20] Robert M Haralick, Karthikeyan Shanmugam, and Its’ Hak Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6):610–621, 1973.
 - [21] J Monaco, J Tomaszewski, M Feldman, Mehdi Moradi, Parvin Mousavi, Alexander Boag, Chris Davidson, Purang Abolmaesumi, and Anant Madabhushi. Detection of prostate cancer from whole-mount histology images using markov random fields. In *Workshop on Microscopic Image Analysis with Applications in Biology (in conjunction with MICCAI)*. Citeseer, 2008.
 - [22] Shivang Naik, Scott Doyle, Shannon Agner, Anant Madabhushi, Michael Feldman, and John Tomaszewski. Automated gland and nuclei segmentation for grading of prostate and breast cancer histopathology. In *2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 284–287. IEEE, 2008.
 - [23] Shivang Naik, Scott Doyle, Michael Feldman, John Tomaszewski, and Anant Madabhushi. Gland segmentation and computerized Gleason grading of prostate histology by integrating low-, high-level and domain specific information. In *MIAAB Workshop*, pages 1–8. Citeseer, 2007.
 - [24] Roberto Marcondes Cesar Jr and Luciano da Fontura Costa. *Shape classification and analysis: theory and practice*. CRC Press, 2009.
 - [25] Jing-Dong J Han, Nicolas Bertin, Tong Hao, Debra S Goldberg, Gabriel F Berriz, Lan V Zhang, Denis Dupuy, Albertha JM Walhout, Michael E Cusick, Frederick P Roth, et al. Evidence for dynamically organized modularity in the yeast protein–protein interaction network. *Nature*, 430(6995):88–93, 2004.
 - [26] Danila Vella, Simone Marini, Francesca Vitali, Dario Di Silvestre, Giancarlo Mauri, and Riccardo Bellazzi. Mtgo: Ppi network analysis via topological and functional module identification. *Scientific Reports*, 8(1):1–13, 2018.

- [27] Claudio Castellano, Santo Fortunato, and Vittorio Loreto. Statistical physics of social dynamics. *Reviews of Modern Physics*, 81(2):591, 2009.
- [28] Guilherme S Domingues, Filipi N Silva, Cesar H Comin, and L da F Costa. Topological characterization of world cities. *Journal of Statistical Mechanics: Theory and Experiment*, 2018(8):083212, 2018.
- [29] Luciano da Fontoura Costa. Complex networks, simple vision. *arXiv preprint cond-mat/0403346*, 2004.
- [30] Luciano da Fontoura Costa. Complex networks: New concepts and tools for real-time imaging and vision. *arXiv preprint cs.CV/0606060*, 2006.
- [31] Rényi A Erdős P. On random graphs. *Publicationes Mathematicae Debrecen*, 6(290-297), 1959.
- [32] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.
- [33] Mark Newman. *Networks*. Oxford University Press, 2018.
- [34] Jesper Dall and Michael Christensen. Random geometric graphs. *Physical Review E*, 66(1):016121, 2002.
- [35] Bernard M Waxman. Routing of multipoint connections. *IEEE Journal on Selected Areas in Communications*, 6(9):1617–1622, 1988.
- [36] Marc Barthélemy. Spatial networks. *Physics Reports*, 499(1-3):1–101, 2011.
- [37] K Ruben Gabriel and Robert R Sokal. A new statistical approach to geographic variation analysis. *Systematic Zoology*, 18(3):259–278, 1969.
- [38] L da F Costa, Francisco A Rodrigues, Gonzalo Travieso, and Paulino Ribeiro Villas Boas. Characterization of complex networks: A survey of measurements. *Advances in Physics*, 56(1):167–242, 2007.
- [39] Vinay Kumar, Abul K Abbas, Nelson Fausto, and Jon C Aster. *Robbins and Cotran pathologic basis of disease, professional edition e-book*. Elsevier Health Sciences, 2014.
- [40] Santo Fortunato. Community detection in graphs. *Physics Reports*, 486(3-5):75–174, 2010.
- [41] Luciano da Fontoura Costa and Filipi Nascimento Silva. Hierarchical characterization of complex networks. *Journal of Statistical Physics*, 125(4):841–872, 2006.
- [42] Tom Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874, 2006.
- [43] Katarzyna Sznajd-Weron and Jozef Sznajd. Opinion evolution in closed community. *International Journal of Modern Physics C*, 11(06):1157–1165, 2000.