Automatic Glandular and Tubule Region Segmentation in Histological Grading of Breast Cancer

Kien Nguyen, Michael Barnes, Chukka Srinivas, Christophe Chefd’hotel
Ventana Medical Systems, Inc. a member of Roche Group, Mountain View, USA

ABSTRACT
In the popular Nottingham histologic score system for breast cancer grading, the pathologist analyzes the H&E tissue slides and assigns a score, in the range of 1-3, for tubule formation, nuclear pleomorphism and mitotic activity in the tumor regions. The scores from these three factors are added to give a final score, ranging from 3-9 to grade the cancer. Tubule score (TS), which reflects tubular formation, is a value in 1-3 given by manually estimating the percentage of glandular regions in the tumor that form tubules. In this paper, given an H&E tissue image representing a tumor region, we propose an automated algorithm to segment glandular regions and detect the presence of tubules in these regions. The algorithm first detects all nuclei and lumen candidates in the input image, followed by identifying tumor nuclei from the detected nuclei and identifying true lumina from the lumen candidates using a random forest classifier. Finally, it forms the glandular regions by grouping the closely located tumor nuclei and lumina using a graph-cut-based method. The glandular regions containing true lumina are considered as the ones that form tubules (tubule regions). To evaluate the proposed method, we calculate the tubule percentage (TP), i.e., the ratio of the tubule area to the total glandular area for 353 H&E images of the three TSs, and plot the distribution of these TP values. This plot shows the clear separation among these three scores, suggesting that the proposed algorithm is useful in distinguishing images of these TSs.

1. INTRODUCTION
To perform breast cancer grading, the pathologist usually uses the Nottingham histologic score system applied on H&E tissue slides. This system considers three different factors regarding the tissue appearance, namely tubule formation, nuclear pleomorphism and mitotic activity in the tumor regions, and gives a score (in the range of 1-3) to each factor. The total score from these three factors (ranging from 3-9) is used to grade the cancer, i.e., a grade of 1, 2 or 3 is given to the cancer if the total scores are between 3-5, 6-7 or 8-9, respectively. In this paper, we mainly focus on the first factor, the tubule formation of the tissue. The tubule formation is represented by the TS (ranging from 1-3), which is determined based on the percentage of glandular regions in the tumor that form tubules. Typically, a TS 1 is given when most of the glandular regions (> 75%) form tubules, while a TS 3 is given when almost no tubules are formed (< 10%) in the glandular regions. The remaining tissues are given a TS 2. In this work, we present an automated algorithm to process an H&E tissue image (which represents a tumor region) to segment glandular regions and detect the presence of tubules in these regions. Note that we do not perform automatic analysis of nuclear pleomorphism and mitotic activity in this study.

2. RELATED WORK
Automatic breast cancer grading using H&E tissue images has been previously addressed in the literature. To distinguish between low and high grades of breast cancer, Doyle et al. combine both textural features and nuclear architectural features, followed by a spectral clustering procedure to reduce the dimension of the feature space. In Chekkoury et al., a large set of features including topological, morphometric and textural features are used to classify the tissue samples into cancerous and non-cancerous. The tubule scoring problem, which was mentioned above, is addressed in Ojansivu et al. by using textural features and support vector machine classifier. Finally, the tubule segmentation problem, which is similar to our work, is addressed in Dalle et al. and Maqlin et al. However, compared to the previous work, our method differs in that:

- Similar to the methods in Dalle et al. and Maqlin et al., we first detect all nuclei from the image. However, unlike these methods, we conduct an additional nuclei classification step to identify tumor nuclei from non-tumor nuclei. This is because non-tumor nuclei are irrelevant when segmenting glandular regions.
In order to identify true lumina from other white regions (e.g., fatty regions, broken tissue), we employ the learning-classification scheme (rather than using heuristic rules as in Maqlin et al.). By doing this, we are able to incorporate a large number of relevant features for lumen identification. Further, we let the classifier make the decision automatically by combining all these features, rather than manually imposing subjective rules for this task.

Since the methods in Dalle et al. and Maqlin et al. only associate the lumen and its closest nuclei, they cannot handle the case where a lumen is surrounded by multiple layers of nuclei. However, by constructing the graph of lumina and nuclei and using the normalized cut method to perform the segmentation, we take into account the global distribution of the nuclei and lumina in the image, hence, we are able to form the tubule which contains one lumen and multiple layers of nuclei.

The detection and classification of nuclei in this work is similar to the work in Nguyen et al. However, we use a richer set of nuclei features here compared to Nguyen et al. Moreover, to stop the recursive cut procedure we use a pre-defined cut threshold, instead of computing the gland-score for the connected components as done in Nguyen et al. because we do not look for components with close-chain structure or elliptical shape.

3. METHOD

The proposed algorithm (i) detects all nuclei and lumen candidates in the image (ii) uses a random forest classifier to identify tumor nuclei from the detected nuclei and identify true lumina from the lumen candidates, and (iii) forms the glandular regions by grouping closely located nuclei and lumina using a graph-cut-based method. If a glandular region contains lumina, we consider that it forms a tubule structure. The flowchart of the proposed method is shown in Fig. 1.

3.1 Nuclei detection and classification

Nuclei detection (nuclei center detection) is performed by applying the radial-symmetry-based method on the Hematoxylin channel (obtained using color deconvolution). Next, we identify tumor nuclei from all the detected nuclei (which also contain lymphocytes, stroma nuclei, etc) using a classification-based approach, i.e., we (i) create a training set of tumor and non-tumor nuclei, (ii) extract features for each nucleus, (iii) train a random forest classifier using these features and (iv) use it to classify the nuclei in a test image into the two nuclei types. See Fig. 2 for an example.

Nuclei feature extraction: We first apply the Otsu’s thresholding method locally on the region surrounding each nuclei center (detected above) to segment the nuclei region. The features computed for each nucleus $n_0$ include:

- Morphology features: area, circularity, and solidity of the nuclei.
- Density of the neighboring nuclei in the neighborhood area (a patch of size $S_N \times S_N$ around $n_0$).
- Texture features: histogram of intensity, histogram of gradient magnitude and orientation, co-occurrence features in the neighborhood area.
3.2 Lumen candidate detection and true lumen identification

The presence of lumen, the white region surrounding by tumor nuclei, is the sign of tubule formation in the glandular regions. To detect lumen, we first (i) find all lumen candidates (LCs), the white regions, by applying a simple thresholding operation in the grayscale image. Besides true lumina, the LCs also contain artifacts such as fat regions, broken tissue regions, which also appear white in the image. To detect true lumina from these LCs, we again use the classification-based approach, i.e., extract features from the LCs and classify them into true lumina vs artifacts. See Fig. 2 for an example.

Lumen feature extraction:
A true lumen is usually surrounded by a ring of nuclei, while artifacts do not have this property. We first identify nuclei associated with each LC, i.e., nuclei within a distance d from the closest pixel on the LC boundary. Note that we use all detected nuclei (without classification result) for this feature extraction since we do not want the errors from the nuclei classification task to propagate here. The following features are computed for each LC.

- Nuclei distribution around LC: we divide the area around the LC into N angular bins, and compute (i) the number of bins where nuclei are present, and (ii) the largest number of consecutive bins where nuclei are present.
- Morphology features: area, circularity, solidity, and curvature of the LC, grayscale intensity in the LC area.
- Texture features: histogram of intensity, histogram of gradient magnitude and orientation, co-occurrence features \( S_L \times S_L \) in the neighborhood area around the LC. A random forest classifier is again selected for this true lumina vs artifacts classification task.
3.3 Glandular and tubule region segmentation

Once tumor nuclei and lumina are found, we aim at grouping them together to generate glandular regions, since glandular regions are usually formed by either a group tumor nuclei or a group of both lumina and the surrounding tumor nuclei (Fig. 2). To perform this task, we use a similar method to the work presented in Nguyen et al.\(^7\) In this method, we need to build a nuclei-lumen-graph for the image, in which each vertex represents a nucleus or a lumen, while each edge represents a link between a nucleus and a nucleus or between a nucleus and a lumen. A link is created for two nuclei if their distance is less than a threshold \(\delta_n\) and no stroma is present in the line connecting them. A link is created for a lumen and a nucleus if the distance from the nucleus to the closest pixel on the lumen boundary is less than \(\delta_l\) and no stroma is present in the line connecting the nucleus and the pixel. A weight of 1 is assigned to all the edges. Once the graph is created, we use the normalized cut method to partition the graph into different connected components. This is done by recursively partitioning the graph (or components) into two smaller components, until the cost of the cut exceeds a threshold \(\delta_c\). Since the normalized cut method aims to remove the set of weakest links (sparse links), the resulted components are likely to represent the groups of closely located nuclei and lumina (with dense links between them). If a segmented component contains lumen, it is considered as a tubule region, otherwise it is considered as a non-tubule glandular region. Components with too few nuclei (less than 3) are discarded. See Fig. 2.

4. RESULTS

4.1 Quantitative evaluation by true lumen identification

Since the identification of true lumina is critical in determining the presence of tubules in the tissue image (recall that a tubule is considered as a group of a true lumen and the surrounding nuclei), we first evaluated the proposed method by evaluating the true lumina vs artifacts classification performance. To this end, we obtained a database of 479 true lumina and 1,342 artifacts from 36 tissue images\(^*\) (which is referred to as DB1), and performed classification by 10-fold cross-validation on this database (we used random forest as the classifier as mentioned above). The average accuracy (standard deviation) obtained is 0.91 (0.03), which shows the effectiveness of the proposed method.

4.2 Quantitative evaluation by computing TP and performing tubule scoring

a. Computing TP:

The glandular and tubule segmentation results can be used in computing TP, which is the criterion to distinguish tissue images w.r.t. tubule formation (images of different TSs). As a result, an additional useful experiment is to obtain a database of tissue images with known TS, compute TPs from the glandular and tubule segmentation results, and observe the TPs obtained from these images.

The database we used in this experiment includes 229 images with TS 1, 28 images with TS 2 and 96 images with TS 3 (these scores were given by a pathologist). We referred to this database as DB2. When applying the proposed method on this database, we used DB1 as the training database for true lumina vs artifacts classification. Examples of the results of the proposed method on these images are shown in Fig. 3. The TP of an image can be computed by calculating the ratio of the total tubule area to the total glandular area (which includes both tubule regions and non-tubule glandular regions). See Fig. 2 for an example. The normalized distribution of TPs of all images in DB2 are plotted in Fig. 4. As can be seen in this figure, there is a clear separation among the TP distributions of the three TSs, i.e., images of TS 1 obtain the highest TPs, while images of TS 3 obtain the lowest TPs. This result agrees with the Nottingham histologic score system, showing the robustness of the proposed tubule segmentation method.

b. Perform tubule scoring:

In this tubule scoring problem, we aimed to automatically classify a tissue image into one of the three TSs by using TP as an image feature. For comparison purposes, we further extracted textural features (Gabor features were used due to its popularity in tissue image processing), and nuclear architectural features (as described in Doyle et al.\(^1\)) from the images, and used them for classification. We conducted the 10-fold cross-validation on

\(^*\)The selection of true lumina and artifacts in these images are done by a pathologist.
Figure 3. Examples of glandular and tubule region segmentation for images of TS 1 (a), TS 2 (b) and TS 3 (c). More tubules (cyan contours) are found in TS 1 than TS 2 images, while no tubules are found in TS3 image.
Figure 4. The normalized distribution of TPs for images of the three TSs.

<table>
<thead>
<tr>
<th>Features</th>
<th>TP as the only feature</th>
<th>Nuclei architectural features (NAF)</th>
<th>NAF + TP</th>
<th>Gabor features</th>
<th>Gabor features + TP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy (std)</td>
<td>0.84 (0.06)</td>
<td>0.68 (0.07)</td>
<td>0.87 (0.04)</td>
<td>0.73 (0.09)</td>
<td>0.83 (0.08)</td>
</tr>
</tbody>
</table>

Table 1. Cross-validation classification accuracies showing the usefulness of using TP as an image feature for the tubule scoring problem.

DB2 and reported the average classification accuracy (with standard deviation) in Table 1. In this table, we showed the results using TP as the only feature, Gabor features, nuclei architectural features and their combinations with TP. A significant improvement in accuracy was obtained when TP was used, demonstrating the usefulness of the proposed method in the automatic tubule scoring problem.

5. CONCLUSIONS

We proposed a method to segment glandular regions and detect tubules from tumor area in the H&E breast tissue image. Unlike the published methods, our method used the classification-based approach to detect tumor nuclei and true lumen in the image. This helps to discard the irrelevant objects (non-tumor nuclei or non-lumen white regions) without the need of designing specific rules as done in the literature. Moreover, the normalized cut method, which involves an optimization strategy, is used for glandular region segmentation. The tubule-percent computed based on the segmentation result is useful in classifying tissue images into the three tubule scores.

REFERENCES


