Multi-scale learning based segmentation of glands in digital colonrectal pathology images

Yi Gao[a,b,c], **William Liu**[d], **Shipra Arjun**[e], **Liangjia Zhu**, **Vadim Ratner**, **Tahsin Kurc**[a,c], **Joel Saltz**[a,c], and **Allen Tannenbaum**[b,c,a]

[a]Department of Biomedical Informatics, Stony Brook University, NY, U.S.A

[b]Department of Applied Mathematics & Statistics, Stony Brook University, NY, U.S.A

[c]Department of Computer Science, Stony Brook University, NY, U.S.A

[d]The Upper School of Buckingham Browne & Nichols, MA, U.S.A

[e]Department of Biomedical Engineering, Stony Brook University, NY, U.S.A

Abstract

Digital histopathological images provide detailed spatial information of the tissue at micrometer resolution. Among the available contents in the pathology images, meso-scale information, such as the gland morphology, texture, and distribution, are useful diagnostic features. In this work, focusing on the colon-rectal cancer tissue samples, we propose a multi-scale learning based segmentation scheme for the glands in the colon-rectal digital pathology slides. The algorithm learns the gland and non-gland textures from a set of training images in various scales through a sparse dictionary representation. After the learning step, the dictionaries are used collectively to perform the classification and segmentation for the new image.

Keywords

digital pathology; gland segmentation; texture; dictionary learning

---

1. DESCRIPTION OF PURPOSE

Digital histopathological images provide detailed spatial information of the tissue at micrometer resolution.[1] Among the available contents in the pathology images, meso-scale information, such as the gland morphology, texture, and distribution, are useful diagnostic features.

Researchers in[2–6] design gland and nuclei segmentation algorithms that aid in determining the grading and/or sub-typing of the prostate, breast, and brain cancers. Gland morphology is studied in.[7–9] In gland extraction, one common assumption is that the contours of the glands appear as conic curves in 2D imaging planes.[10, 11] While this may be a satisfying

assumption for normal tissue, for the malignant tissue, however, such a morphological prior does not hold.

In this work, focusing on colon-rectal cancer tissue samples, we propose a multi-scale learning based segmentation scheme for the glands in the colon-rectal digital pathology slides. The algorithm learns the gland and non-gland textures from a set of training images in various scales through a sparse dictionary representation, regardless of the tissue type (benign vs malignant). After the learning step, the dictionaries are used collectively to perform the classification and segmentation for the new image.

2. METHOD

In,[12] a sparse dictionary approach is employed for extracting texture features from the foreground and background, for the purpose of interactive segmentation. In digital pathology images, it is noted that certain images features can be captured in a multi-scale fashion. As a result, we extend the sparse dictionary representation based segmentation into a multi-scale framework.

Specifically, denote the training images as

$$I_i : \mathbb{R}^2 \to \mathbb{R}^3, \quad i=1,\dots,N \quad (1)$$

Their corresponding ground truth segmentations are

$$L_i : \mathbb{R}^2 \to \{0,1\} \quad (2)$$

where 1 indicates the gland region. Then, M_0 image patches of the size $m \times m$ are sampled, with replacement and overlapping, from the training images.[13] A 3-channel RGB patch is denoted as

$$g_i, t_i \in \mathbb{R}^{m \times m \times 3}$$
$$\text{with } i=1,\dots,M_0 \quad (3)$$

where g_i is from gland region and t_i is non-gland. Such a sample-with-replacement strategy successfully enlarges the sample set and reduces the estimation variance. All the patches are grouped into two categories: gland and non-gland. Denote the patch list of the gland tissue as

$$G_0 = \{g_i \in \mathbb{R}^{3m^2}; i=1,\dots,M_0\} \quad (4)$$

and the non-gland tissue has the list

$$T_0 = \{t_i \in \mathbb{R}^{3m^2}; i = 1, \ldots, M_0\} \quad (5)$$

Here, we slightly abuse the notation such that g_i, $t_i \in \mathbb{R}^{3m^2}$ are vectors.

With the two lists of patches (vectors), the K-SVD algorithm is used to learn the underlining structure of the texture space as in Algorithm 1[14]

Algorithm 1

K-SVD Dictionary Construction

1:      Initialize D_0^g (T_0 resp.)

2:   **repeat**

3:      Find sparse coefficients Γ (γ_i's) using any pursuit algorithm.

4:      **for** $j = 1, \ldots, d_0$, update f_j, the j-th column of D_0^g (T_0 resp.), by the following process **do**

5:        Find the group of vectors that use this atom: $\zeta_j = \{i : 1 \quad i \quad d_0, \gamma_i(j) \quad 0\}$

6:        Compute $E_j := Q - \sum_{i \neq j} f_i \Gamma_T^i$ where Γ_T^i is the i-th row of Γ

7:        Extract the i-th columns in E_j, where $i \in \zeta_j$, to form E_j^R

8:        Apply SVD to get $E_j^R = U \Delta V$

9:        f_j is updated with the first column of U

10:       The non-zeros elements in Γ_T^j is updated with the first column of $V \times (1, 1)$

11:      **end for**

12:   **until** Convergence criteria is met

The over-complete basis for G_0 (T_0 resp.) is denoted as

$$D_0^g \in \mathbb{R}^{3m^2 \times d_0}, D_0^t \in \mathbb{R}^{3m^2 \times d_0}$$
$$\text{with } d_0 \gg 3m^2 \quad (6)$$

Such a dictionary construction is performed in the highest (native) resolution of the image. In order to better capture the texture information at various scales, wavelet decomposition is performed on the I_j's, and the learning is performed at each decomposition level. More explicitly, let the original image be at level 0, then level i 1 is the 2^i down-sampled version of the original image. In all the levels, the same patch size is used for learning. Equivalently, the same patch will cover bigger tissue region and therefore capture larger scale information

of the image texture. The dictionaries (gland and non-gland) in the i-th level are denoted as D_i^g and D_i^t, with $i \geq 0$.

After the gland and non-gland dictionaries at all levels have been constructed, we perform classification on the patches in the given image to be segmented. The classification is also performed in a multi-scale fashion, which is detailed as follows.

At the original resolution, given a new image patch $g: \mathbb{R}^{m \times m} \to \mathbb{R}^3$, by slightly abuse of notation, we also consider g as a $3m^2$ dimensional vector. Then, we compute the sparse coefficient $\boldsymbol{\eta}$ of g, with respect to the dictionary D_0^g.

$$\min_{\boldsymbol{\eta}} \|g - D_0^g \cdot \boldsymbol{\eta}\|_2 \leq \varepsilon$$
$$\text{s. t. } \|\boldsymbol{\eta}\|_0 \leq k \qquad (7)$$

where k is the sparsity constraint. The reconstruction error is defined as

$$e_0^g := \|g - D_0^g \cdot \boldsymbol{\eta}\|_2 \quad (8)$$

This problem is solved by the Orthogonal Matching Pursuit (OMP) method.[15] Similarly, the sparse coefficient $\boldsymbol{\xi}$ of g under the dictionary D_0^t is also computed as:

$$\min_{\boldsymbol{\xi}} \|g - D_0^t \cdot \boldsymbol{\xi}\|_2 \leq \varepsilon$$
$$\text{s. t. } \|\boldsymbol{\xi}\|_0 \leq k \qquad (9)$$

The reconstruction error is defined as:

$$e_0^t := \|g - D_0^t \cdot \boldsymbol{\xi}\|_2 \quad (10)$$

Moreover, we define

$$e_0 := e_0^g - e_0^t \quad (11)$$

It is noted that the errors e_0^g and e_0^t indicate how well (or how bad) the new image patch can be reconstructed from information learned from the gland and non-gland region. Hence, the higher the e_0 is, the more this patch is considered as the non-gland region and vice versa. Furthermore, for each pixel $(x, y) \in \Omega$, a patch $p(x, y)$ is extracted as the sub-region of I defined on $[x, x + m] \times [y, y + m]$. Following the procedure above, we compute the image $e_0(x, y)$ accordingly. A probability map $P_0(x, y)$ for the gland region is defined as

$$P_0(x,y) = \frac{1}{1+\exp(-e_0(x,y))} \quad (12)$$

Such procedure is performed for all the scales and a series of probability maps $P_i(x, y)$ with $i = 0, 1, \ldots, L$ are constructed. Note that each $P_i(x, y)$ is reconstructed to the original scale (with all detail coefficients set to 0) and is therefore of the same size as P_0. Finally, define

$$P(x,y) = \max_{l=0}^{L} P_l(x,y) \quad (13)$$

as the maximum response along the scale dimension. The final probability is threshold at 0.5 for the final mask of gland region.

3. EXPERIMENTS AND RESULTS

80 images are extracted from The Cancer Genome Atlas (TCGA) colon-rectal digital pathology diagnosis images. The resolution is 0.5μm/pixel. The image sizes are all 1024×1024 pixels.

Three randomly picked results are shown in Figure 1. The top row shows the image with the shading indicating the manual segmentation. The corresponding panel below shows the proposed automatic segmentation results.

In addition to the benign tissue, the algorithm also performs consistently on malignant tissue, as shown in Figure 2. Here two cases are shown from left to the right: The first/third are the manual segmentation and the second/forth are the algorithm outputs.

Quantitatively, a leave-one-out scheme is performed for testing. Specifically, the 79 images and their manual segmentations are used for learning. The learned model is then applied on the left-out image. The Dice coefficient is measured for all the 80 cases and a mean of 0.81 accuracy (0.056 standard deviation) is achieved. Moreover, the Hausdorff distance is measured and a mean of 70μm (10.0μm standard deviation) is achieved.

4. CONCLUSION AND FUTURE RESEARCH DIRECTIONS

In this work, we propose a multi-scale learning based segmentation scheme for the glands in the colon-rectal digital pathology slides. The algorithm learns the gland and non-gland textures from a set of training images in various scales through a sparse dictionary representation. After the learning, the dictionaries are used collectively to perform the classification and segmentation for the new image.

More recently, the deep neural network (DNN) framework is becoming more and more widely adopted, including for the task of gland segmentation.[16, 17] The DNN framework has the advantage that minimal manually curated features are needed for the classification

process. The proposed framework utilizes the sparse dictionary for the learning and feature extraction. However, the signal sparse reconstruction based on dictionary is a linear combination. DNN harnesses the hierarchical nonlinear transformation and may result in better performance. The ongoing research include combining the manually designed features with the DNN based features, for a better overall performance.

The work has not been submitted for publication or presentation elsewhere.

References

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

2. Naik, S.; Doyle, S.; Agner, S.; Madabhushi, A.; Feldman, M.; Tomaszewski, J. Automated gland and nuclei segmentation for grading of prostate and breast cancer histopathology. 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro; IEEE; 2008. p. 284-287.

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

4. Kong H, Gurcan M, Belkacem-Boussaid K. Partitioning histopathological images: an integrated framework for supervised color-texture segmentation and cell splitting. Medical Imaging, IEEE Transactions on. 2011; 30(9):1661–1677.

5. Cooper L, Gutman DA, Long Q, Johnson BA, Cholleti SR, Kurc T, Saltz JH, Brat DJ, Moreno CS. The proneural molecular signature is enriched in oligodendrogliomas and predicts improved survival among diffuse gliomas. PloS one. 2010; 5(9):e12548. [PubMed: 20838435]

6. Kong J, Cooper LA, Wang F, Gao J, Teodoro G, Scarpace L, Mikkelsen T, Schniederjan MJ, Moreno CS, Saltz JH, Brat DJ. Machine-based morphologic analysis of glioblastoma using whole-slide pathology images uncovers clinically relevant molecular correlates. PLoS ONE. 2013; 8(11)

7. Lopez, CM.; Agaian, S.; Sanchez, I.; Almuntashri, A.; Zinalabdin, O.; Rikabi, A.; Thompson, I. Exploration of efficacy of gland morphology and architectural features in prostate cancer gleason grading. Systems, Man, and Cybernetics (SMC), 2012 IEEE International Conference on; IEEE; 2012. p. 2849-2854.

8. Cohen A, Rivlin E, Shimshoni I, Sabo E. Memory based active contour algorithm using pixel-level classified images for colon crypt segmentation. Computerized Medical Imaging and Graphics. 2015

9. Xu, J.; Janowczyk, A.; Chandran, S.; Madabhushi, A. SPIE Medical Imaging. International Society for Optics and Photonics; 2010. A weighted mean shift, normalized cuts initialized color gradient based geodesic active contour model: applications to histopathology image segmentation; p. 76230Y–76230Y

10. Fu H, Qiu G, Shu J, Ilyas M. A novel polar space random field model for the detection of glandular structures. Medical Imaging, IEEE Transactions on. 2014; 33(3):764–776.

11. Gunduz-Demir C, Kandemir M, Tosun AB, Sokmensuer C. Automatic segmentation of colon glands using object-graphs. Medical image analysis. 2010; 14(1):1–12. [PubMed: 19819181]

12. Gao Y, Bouix S, Shenton M, Tannenbaum A. Sparse texture active contour. Image Processing, IEEE Transactions on. 2013; 22(10):3866–3878.

13. Efron, B.; Tibshirani, R.; Tibshirani, R. An introduction to the bootstrap. Chapman & Hall/CRC; 1993.

14. Aharon M, Elad M, Bruckstein A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. Signal Processing, IEEE Transactions on. 2006; 54(11):4311–4322.

15. Pati, YC.; Rezaiifar, R.; Krishnaprasad, P. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on; IEEE; 1993. p. 40-44.

16. Kainz P, Pfeiffer M, Urschler M. Semantic segmentation of colon glands with deep convolutional neural networks and total variation segmentation. 2015 arXiv preprint arXiv:1511.06919.

17. Chen, H.; Qi, X.; Yu, L.; Heng, P-A. DCAN: Deep contour-aware network for accurate gland segmentation. GlaS challenge in Medical Image Computing and Computer-Assisted Intervention; 2015; Springer; 2015.
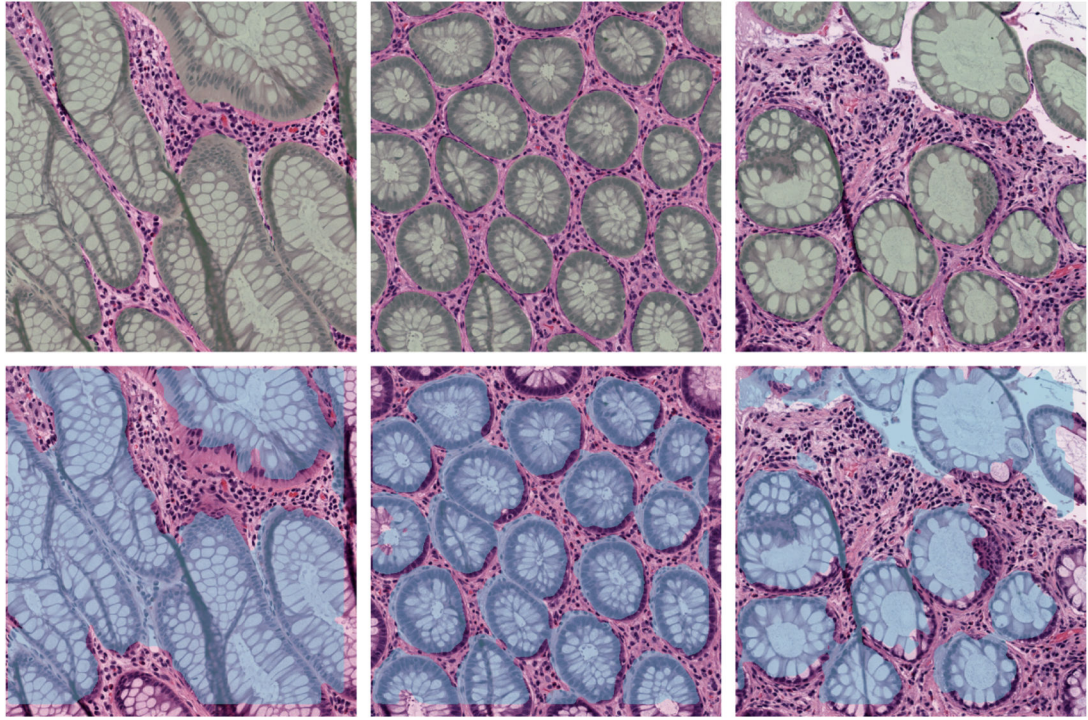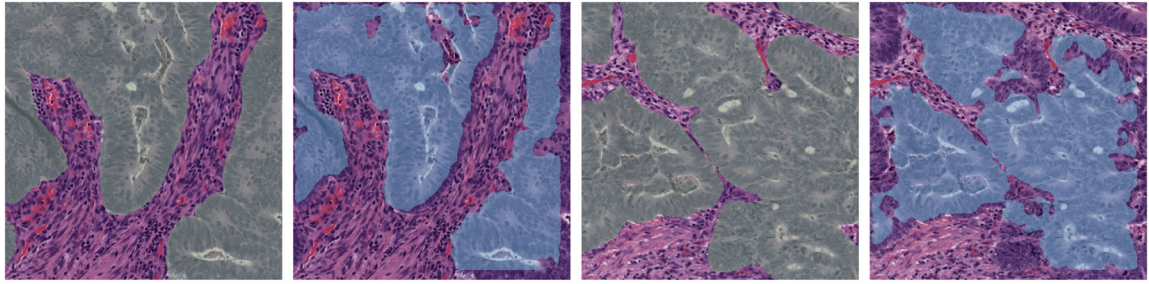
Figure 1.
Segmentation on three benign tissues patches.

Figure 2.
Segmentation on two malignant tissues patches.