Structure Prediction for Gland Segmentation With Hand-Crafted and Deep Convolutional Features

Siyamalan Manivannan, *Member, IEEE*, Wenqi Li, *Member, IEEE*, Jianguo Zhang, *Senior Member, IEEE*, Emanuele Trucco, and Stephen J. McKenna, *Senior Member, IEEE*

***Abstract*— We present a novel method to segment instances of glandular structures from colon histopathology images. We use a structure learning approach which represents local spatial configurations of class labels, capturing structural information normally ignored by sliding-window methods. This allows us to reveal different spatial structures of pixel labels (e.g., locations between adjacent glands, or far from glands), and to identify correctly neighboring glandular structures as separate instances. Exemplars of label structures are obtained via clustering and used to train support vector machine classifiers. The label structures predicted are then combined and post-processed to obtain segmentation maps. We combine hand-crafted, multi-scale image features with features computed by a deep convolutional network trained to map images to segmentation maps. We evaluate the proposed method on the public domain GlaS data set, which allows extensive comparisons with recent, alternative methods. Using the GlaS contest protocol, our method achieves the overall best performance.**

***Index Terms*— Molecular and cellular imaging, gastrointestinal tract, segmentation.**

I. Introduction

HISTOLOGICAL assessment of gland formation and morphology informs diagnosis, prognosis and treatment planning of patients [1]. It is useful for grading of adenocarcinomas in colon, breast, and prostate. Such assessment is labour intensive, performed by highly trained pathologists, and often has limited reproducibility. The emergence of whole-slide imaging is increasing the volume of digital histology image data to be analysed, exacerbating the problem. Algorithms capable of reliably segmenting glandular structures automatically would accelerate analysis and provide reproducible, quantitative measures of gland morphology. The development of such algorithms is challenging because malignancy results in irregular morphology and poorly differentiated gland boundaries, and because glandular structures can be closely packed together but need to be segmented as separate instances. Glands in healthy epithelial tissue have a clear structure with interior lumen surrounded by columnar epilthelial cells (Fig. 1 (a)). This structure degenerates in moderately or poorly differentiated adenocarcinomas (Fig. 1 (b)).

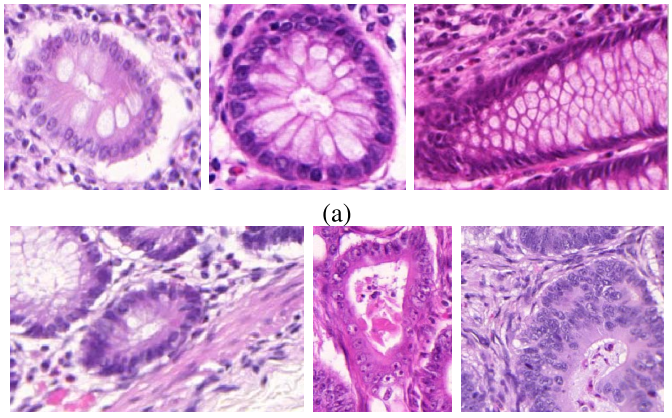

Fig. 1. Glandular structures in the Warwick-QU dataset [2]. (a) Glands in healthy tissue. (b) Left to right: adenoma, moderately differentiated, and poorly differentiated adenocarcinoma.

Gland segmentation, and more generally semantic pixel labelling, often incorporates a sliding window classification procedure based on features extracted from a local window centred at each image location (e.g. [3], [4]). Such a procedure ignores the class labels' spatial structure. Instead, we propose to learn discriminative models for segmentation in which local spatial structures are encoded in the label (output) space as well as in the feature (input) space. By directly employing label structure we can more reliably separate objects and thus improve instance segmentation. The number of possible label structures grows exponentially as the size of the local region considered increases, posing a challenge. We show how this large output space can be handled by combining small numbers of local structure exemplars obtained via clustering.

We combine hand-crafted features with learned deep convolutional features to capture image context information. We conduct experiments with the publicly available GlaS

Manuscript received June 14, 2017; revised August 29, 2017; accepted August 31, 2017. Date of publication September 8, 2017; date of current version December 29, 2017. This work was supported by U.K. Engineering and Physical Sciences Research Council under Grant EP/M005976/1. *(Corresponding author: Stephen J. McKenna.)*

S. Manivannan was with the University of Dundee, Dundee DD1 4HN, U.K. He is now with the University of Jaffna, Jaffna 44000, Sri Lanka.

W. Li was with the University of Dundee, Dundee DD1 4HN, U.K. He is now with the University College London, London WC1E 6BT, U.K.

J. Zhang, E. Trucco, and S. J. McKenna are with the Computer Vision and Image Processing, School of Science and Engineering, University of Dundee, Dundee DD1 4HN, U.K. (e-mail: stephen@computing.dundee.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMI.2017.2750210