DCAN: Deep contour-aware networks for object instance segmentation from histology images

Hao Chen [a,*], Xiaojuan Qi [a], Lequan Yu [a], Qi Dou [a], Jing Qin [b], Pheng-Ann Heng [a]

[a] *Department of Computer Science and Engineering, The Chinese University of Hong Kong, Hong Kong, China*
[b] *School of Nursing, The Hong Kong Polytechnic University, Hong Kong, China*

A R T I C L E   I N F O

A B S T R A C T

In histopathological image analysis, the morphology of histological structures, such as glands and nuclei, has been routinely adopted by pathologists to assess the malignancy degree of adenocarcinomas. Accurate detection and segmentation of these objects of interest from histology images is an essential prerequisite to obtain reliable morphological statistics for quantitative diagnosis. While manual annotation is error-prone, time-consuming and operator-dependant, automated detection and segmentation of objects of interest from histology images can be very challenging due to the large appearance variation, existence of strong mimics, and serious degeneration of histological structures. In order to meet these challenges, we propose a novel deep contour-aware network (DCAN) under a unified multi-task learning framework for more accurate detection and segmentation. In the proposed network, multi-level contextual features are explored based on an end-to-end fully convolutional network (FCN) to deal with the large appearance variation. We further propose to employ an auxiliary supervision mechanism to overcome the problem of vanishing gradients when training such a deep network. More importantly, our network can not only output accurate probability maps of histological objects, but also depict clear contours simultaneously for separating clustered object instances, which further boosts the segmentation performance. Our method ranked the first in two histological object segmentation challenges, including *2015 MICCAI Gland Segmentation Challenge* and *2015 MICCAI Nuclei Segmentation Challenge*. Extensive experiments on these two challenging datasets demonstrate the superior performance of our method, surpassing all the other methods by a significant margin.

1. Introduction

With the advent of whole slide imaging scanners, tissue histopathology slides can be digitized and stored in the form of digital images. Meanwhile, histopathological analysis performed on these digital images has been demonstrated as an effective and reliable tool for cancer diagnosis and prognosis (Gurcan et al., 2009). In the routine of histopathological examination, accurate detection and segmentation of certain histological structures, such as glands and cancer nuclei, is one of crucial prerequisite steps to obtain reliable morphological statistics that characterize the aggressiveness of tumors. Take the gland for example, a typical gland normally is composed of a lumen area forming the interior tubular structure and epithelial cell nuclei surrounding the cytoplasm, as illustrated in Fig. 1 (left). In contrast, malignant tumours arising from glandular epithelium, also known as adenocarcinomas, exhibit irregularly degenerated form, see Fig. 1 (right). Therefore, the morphology of glands has been widely used in clinical practice for assessing the malignancy degree of several adenocarcinomas, including breast (Elston et al., 1991), prostate (Gleason, 1992), and colon (Fleming et al., 2012). Another example is the counting of object instances such as cell nuclei, which has diagnostic significance for some cancerous diseases (Naik et al., 2008; Xu et al., 2015). This requires an accurate detection and segmentation of cell nuclei, as examples shown in Fig. 2. The nucleus morphism has an important diagnostic value for cancer grading (Stierer et al., 1991; Dunne and Going, 2001; Elston et al., 1991); and the count of mitosis has been regarded as one of the most prognostic factors in breast cancer (Roux et al., 2013; Veta et al., 2015).

The assessments derived from pathological examinations serve as the gold standard for cancer diagnosis in many clinical protocols. With the recent advances of techniques in digitalized scanning, large-scale histopathology images become readily accessible and need to be analyzed promptly. However, manual searching and segmenting histopathological structures from large-scale histopathology images in a conventional way can be
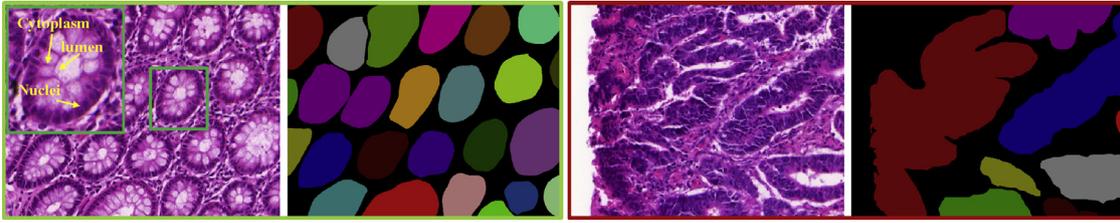
Fig. 1. Examples of gland segmentation in benign (left) and malignant (right) cases: original images (stained with hematoxylin and eosin) and corresponding annotations (individual objects are denoted with different colors) by pathologists. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).
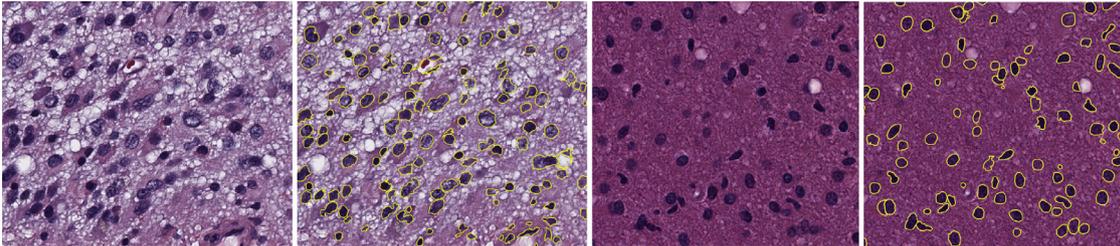

Fig. 2. Examples of nuclei segmentation: original images and corresponding annotations (overlaid on the original images) by pathologists.

expensive, error-prone, and time-consuming. Furthermore, it often suffers from a high inter and intra-observer variability, which results in a limited reproducibility. Therefore, automatic detection and segmentation methods are highly demanded in clinical practice to improve the efficiency, reliability as well as scalability for large-scale histopathological image analysis.

However, automated detection and segmentation of histological structures of interest from histology images can be quite challenging for several reasons. First, there is a huge variation of object appearance depending on the histologic grade as well as the type of the disease. Fig. 1 shows the large difference of glandular structures between benign and malignant cases from colon tissues. Second, the existence of touching clustered objects in tissue samples makes it quite hard for automated segmentation methods to separate them into individual ones. Third, in the malignant cases such as moderately and poorly differentiated adenocarcinomas, the structure of objects such as glands are seriously degenerated, as shown in Fig. 1 (right). Therefore, methods utilizing the prior knowledge with glandular regularity are prone to fail in such cases (Sirinukunwattana et al., 2015a). In addition, the variation of tissue preparation procedures such as sectioning and staining can cause deformation, artifacts and inconsistency of tissue appearance, which can impede the segmentation process as well.

1.1. Related work

Although limited by computational resources and digital imaging equipments, the analysis of histological structures from histology images can date back to 90s from early studies of Bartels et al. (1992), Hamilton et al. (1994) and Weind et al. (1998). Over the past few decades, dramatic advance in computational power, image scanning techniques, and automated analysis algorithms have promoted considerable progress in histopathological image analysis. However, obvious gap is still observed between the results obtained by the automated algorithms and annotations from pathologists.

Previous studies in the literature can be broadly categorized into two classes based on the employed features:

Methods based on hand-crafted features. Regarding the gland detection and segmentation, various hand-crafted features including texture (Farjam et al., 2007; Doyle et al., 2006; Sirinukunwattana et al., 2015b), color information (Tabesh et al., 2007; Jacobs

et al., 2014), morphological cues (Diamond et al., 2004; WU et al., 2005), structural information (Nguyen et al., 2012), and Haar-like features (Sabata et al., 2010) were utilized to analyze the glandular structure in histology images. Similarly, for the nuclei detection and segmentation, various methods have been proposed to tackle this problem ranging from relatively simple approaches, such as thresholding and morphological operations (Irshad et al., 2013; Jung and Kim, 2010), to more sophisticated methods based on hand-crafted features derived from boundaries/contours (Naik et al., 2008; Wienert et al., 2012), gradients (Veta et al., 2011), Laplacian-of-Gaussian (Al-Kofahi et al., 2010), cytological and textural features (Nguyen et al., 2011), etc. Then different classifiers (e.g., Support Vector Machine (SVM), Adaboost and Bayesian) have been employed in the literature to detect and segment nuclei from histology images (Irshad et al., 2014). However, the hand-crafted features suffer from limited representation capabilities, and hence they can be vulnerable to different variations. Furthermore, the piece-wise learning system separating feature extraction and classification may not be optimal as well as efficient for generating precise probability maps of histological structures. Aside from approaches devoting to generating more precise probability maps, several methods have been developed to take advantage of prior shape information of histological structures. For example, the glandular structure for gland segmentation was exploited in graph based methods (Altunbay et al., 2010; Gunduz-Demir et al., 2010), glandular boundary delineation with geodesic distance transform (Fakhrzadeh et al., 2012), polar space random field model (Fu et al., 2014), stochastic polygons model (Sirinukunwattana et al., 2015a), etc. Although these methods achieved promising results in cases of adenoma and well differentiated (low grade) adenocarcinoma, they may fail to achieve satisfying performance in subjects where the glandular structures are seriously deformed.

Deep learning based methods. Deep neural networks are driving advances in image recognition related tasks with powerful feature representation capability. They achieved state-of-the-art performance in many detection and segmentation tasks in medical image computing (Zheng et al., 2015; Dhungel et al., 2015; Ronneberger et al., 2015; Roth et al., 2015; Chen et al., 2015b; Dou et al., 2016a; 2016b). In the field of histopathological image analysis, a deep learning architecture with convolutional autoencoder was utilized in Cruz-Roa et al. (2013) for histopathology

image classification. Stacked sparse autoencoders (SSAE) were exploited with unsupervised pre-training and following fine-tuning for nuclei detection from breast cancer histopathology images in Xu et al. (2015). Although along with merit of unsupervised pre-training, which can handle the situation of limited medical training data, the auto-encoders usually achieved inferior performance on image recognition tasks compared to convolutional neural networks (CNNs). The success of latter one is mostly attributed to the more elegant structures for dealing with images. Regarding the convolutional neural networks, Cireşan et al. (2013) employed deep convolutional neural networks for mitosis detection and achieved the best performance in two grand challenges (Roux et al., 2013; Veta et al., 2015). To further improve the efficiency and effectiveness, Chen et al. (2016a) developed a cascaded deep learning framework, i.e., a coarse model for retrieving candidates and a fine-discrimination model for singling out mitoses from hard mimics. A spatially constrained convolutional neural network was present in Sirinukunwattana et al. (2016b) incorporated with neighboring ensemble prediction, demonstrating the efficacy of deep learning based features from CNNs. Recently, holistically-nested networks were exploited for object segmentation problems in the medical domain, which have demonstrated excellent performance (Xu et al., 2016; Roth et al., 2016; Isabella et al., 2016). Another relevant study to our work is the *U-net* that employed a U-shape deep convolutional network for biomedical image segmentation problems, which set state-of-the-art performance on several grand challenges lately (Ronneberger et al., 2015). Although hierarchical features from low-level to high-level were well explored in Ronneberger et al. (2015), the very deep U-shape architecture could have the issue of vanishing gradients during the training, hence decrease the segmentation performance. Moreover, the implicit utilization of boundaries by weighting losses among different classes may not handle the clustered objects properly. Several methods have been developed for efficient boundary detection explicitly (Bertasius et al., 2015a; 2015b). Different from above studies, we formulated the object segmentation problem under a unified multi-task learning framework, which harnessed the complementary object and contour information simultaneously for better segmentation performance.

1.2. Contributions

In this paper, we present a novel deep contour-aware network (DCAN) to meet the aforementioned challenges in object segmentation from histology images. We first propose a multi-level contextual fully convolutional network (FCN), which can effectively generate multi-scale feature representations in an end-to-end way in order to deal with the large variations of histological structures. The proposed FCN enables us to take an image as input and output the probability map with the same resolution of the input image directly using one single forward propagation. In this regard, it is quite efficient and has great potential to be applied to analyze large-scale histology images. Furthermore, to overcome the problem of vanishing gradients when training such a deep neural network employing limited training samples, we propose to add weighted auxiliary classifiers to encourage the back-propagation of gradient flow. Such an auxiliary supervision mechanism can efficiently alleviate vanishing gradients and hence speedup the training process as well as enhance the discrimination capability of the deep neural network.

More importantly, to handle the challenging cases of touching and clustered objects, we propose to elegantly integrate object appearance (e.g., textures, colors, etc.) and contour information into a multi-task learning framework to form a deep contour-aware network, in which the complementary appearance and contour information can further improve the discriminative ca-

pability of intermediate features, and hence more accurately separate the touching or clustered objects into individual ones. Note that our method does not make any assumptions on the prior shape of histological objects; it can be applied to biopsy samples with various histopathological structures with different grades, including both benign and malignant cases. In addition, in order to further mitigate the issue of insufficient training data, which is a common challenge for medical image analysis tasks, a cross-domain transfer learning approach is developed to improve the training performance. We evaluated the proposed DCAN on two challenging object segmentation tasks from histology images: gland segmentation and nuclei segmentation, and ranked the first in both challenges held conjunction with MICCAI 2015 (*Gland Segmentation Challenge* and *Nuclei Segmentation Challenge*), yielding much better performance than state-of-the-art methods.

A preliminary version of this work was presented in Chen et al. (2016c).[1] In this paper, we provided comprehensive literature review on segmentation related tasks from histology images and refined the methods by giving much more details on the underlying principles as well as running mechanisms. We reported the results of our method on another dataset of 2015 MICCAI Nuclei Segmentation Challenge. We also provided detailed discussions to thoroughly analyze the advantages as well as limitations of our methods.

1.3. Overview

The rest of the article is organized as follows: In Section 2, we first describe the architecture of the multi-level contextual FCN with auxiliary supervision, and then elaborate the proposed contour-aware component and the transfer learning strategy. We report the experiments and results in Section 3. In Section 4, we provide the implementation details and computation cost of the proposed DCAN. We further discuss and analyze our work in Section 5. Finally, conclusions are drawn in Section 6.

2. Methods

In this section, we present the formulation of the proposed contour-aware network in detail. We first introduce the end-to-end fully convolutional network, in which we elegantly fuse multi-level contextual information with auxiliary supervision in order to efficiently generate accurate likelihood maps of objects. Next, we elaborate the deep contour-aware network derived from the FCN by fusing the complementary information of object appearances and contours for more robust and precise segmentation. In addition, in order to mitigate the challenge of inadequate training samples, we leverage the transfer learning strategy to exploit the knowledge learned from cross domains to further improve the performance of the proposed network.

2.1. Multi-level contextual FCN with auxiliary supervision

Fully convolutional networks greatly advanced the state-of-the-art performance of image segmentation related tasks (Chen et al., 2015d; Long et al., 2015). These achievements are mainly attributed to the powerful capability of deep neural networks in feature representations for dense inference. In a FCN, the whole network can be trained in an end-to-end (image-to-image) way, which takes an image as input and outputs the probability map directly with the same resolution of the input image. Its architecture basically contains two components: downsampling path and upsampling path.

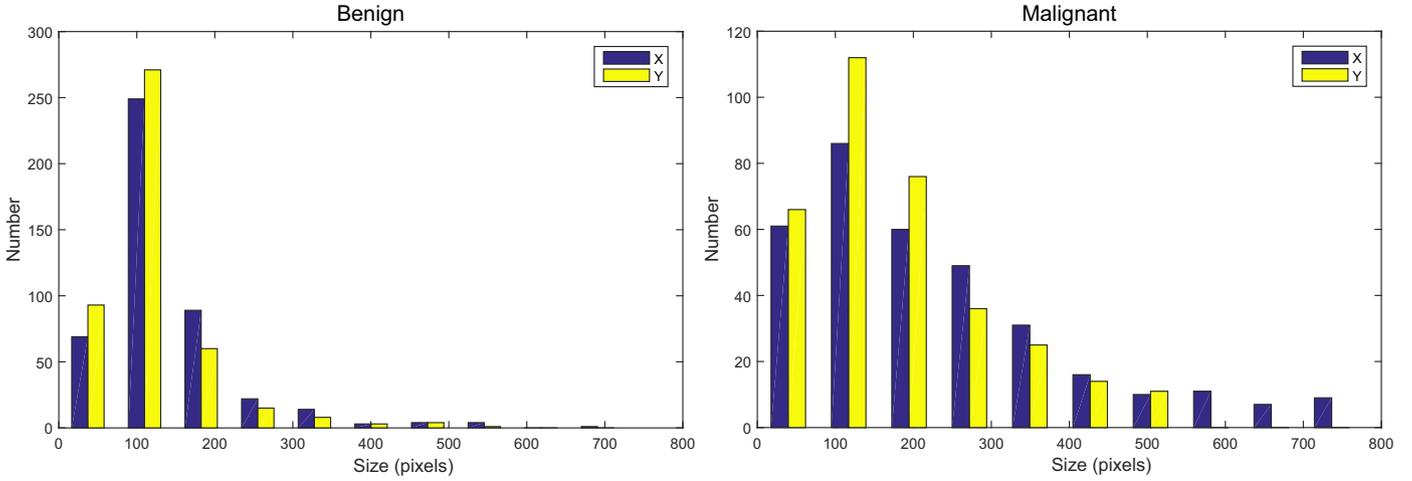[1] Project page: http://www.cse.cuhk.edu.hk/~hchen/research/2015miccai_gland.html.

Fig. 3. The large variation of gland size (the size denotes the smallest rectangle that encloses the gland object).
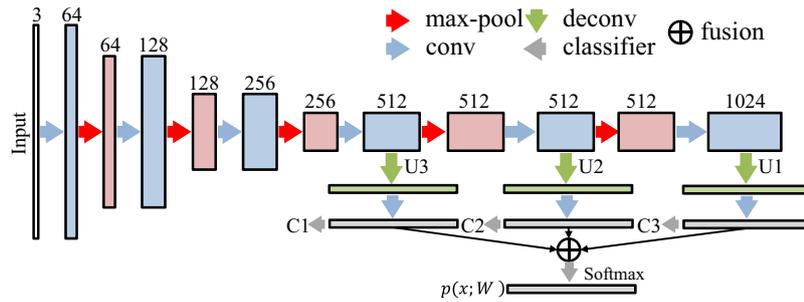

Fig. 4. The schematic illustration of multi-level contextual FCN with auxiliary supervsion.

The downsampling path consists of convolutional and max-pooling layers, which are extensively used in CNNs for image classification tasks (Ciresan et al., 2012; Krizhevsky et al., 2012). The upsampling path is composed of convolutional and deconvolutional layers (backwards strided convolution Long et al., 2015), which upsample the feature maps and output the score masks. The underlying principle of FCN is that the downsampling path is responsible for extracting the high level abstraction information, while the upsampling path aims to predict the score map in a pixel-wise way, i.e., dense inference.

However, in traditional FCN, the classification scores are usually figured out based on the intensity information from a given receptive field (i.e., the region in the original input image that influences the recognition results LeCun et al., 1998). The size of receptive field should contain the targeting objects along with some contextual information that contributes to the recognition. Unfortunately, the network with single receptive field size cannot satisfactorily deal with the large variation of histological structures. For example, as shown in Fig. 3, a small receptive field (e.g., 150 × 150) is suitable for most glands of benign cases, while malignant cases usually demand a relatively larger receptive field since the gland shape in adenocarcinomas is degenerated and elongated. In this regard, enclosing multi-level contextual information ranging from small receptive fields to larger ones can help to recognize structures with large variations and hence improve the segmentation performance.

We propose to improve the FCN by harnessing multi-level contextual feature representations, which include different levels of contextual information, i.e., intensities appearing in various sizes of receptive field. Fig. 4 presents the schematic illustration of FCN with multi-level contextual feature representations. Specifically, our FCN contains a number of convolutional layers, 5 max-pooling layers for downsampling and 3 deconvolutional layers for upsampling. A non-linear mapping layer, i.e., element-wise rectified linear unit (ReLU) (Krizhevsky et al., 2012), is followed for each layer with trained parameters. As the network goes deeper, the size of global receptive field is becoming larger. Built upon this characteristic, the upsampling layers are designed by taking the requirement of different receptive field sizes into account. They upsample the feature maps and make predictions based on the contextual appearance information extracted from the given receptive field. Then these predictions are fused together by a summing operation and final segmentation results based on multi-level contextual features can be obtained from the softmax classification layer.

Directly training such a deep network, however, may face the difficulty of optimization due to the issue of vanishing gradients. Inspired by previous studies on training neural networks with deep supervision (Lee et al., 2015; Xie and Tu, 2015; Chen et al., 2016b), we added three weighted auxiliary classifiers C1–C3 (see Fig. 4 for illustration) into the network to further strengthen the training optimization process. This mechanism can effectively alleviate the problem of vanishing gradients, as the auxiliary supervision can encourage the back-propagation of gradient flow (Lee et al., 2015). Furthermore, the incorporated losses on auxiliary supervision from mid-level to high-level layers are helpful to enhance the feature representation capability throughout the network.

Finally, the FCN with multi-level contextual features extracted from input I can be trained by minimizing the overall loss $\mathcal{L}$, i.e., a combination of auxiliary loss $\mathcal{L}_a(I; W)$ with corresponding discount weights w_a and main loss $\mathcal{L}_e(I; W)$ between the predicted results and ground truth annotations:

$$\mathcal{L}(I; W) = \lambda \psi(W) + \sum_a w_a \mathcal{L}_a(I; W) + \mathcal{L}_e(I; W) \tag{1}$$
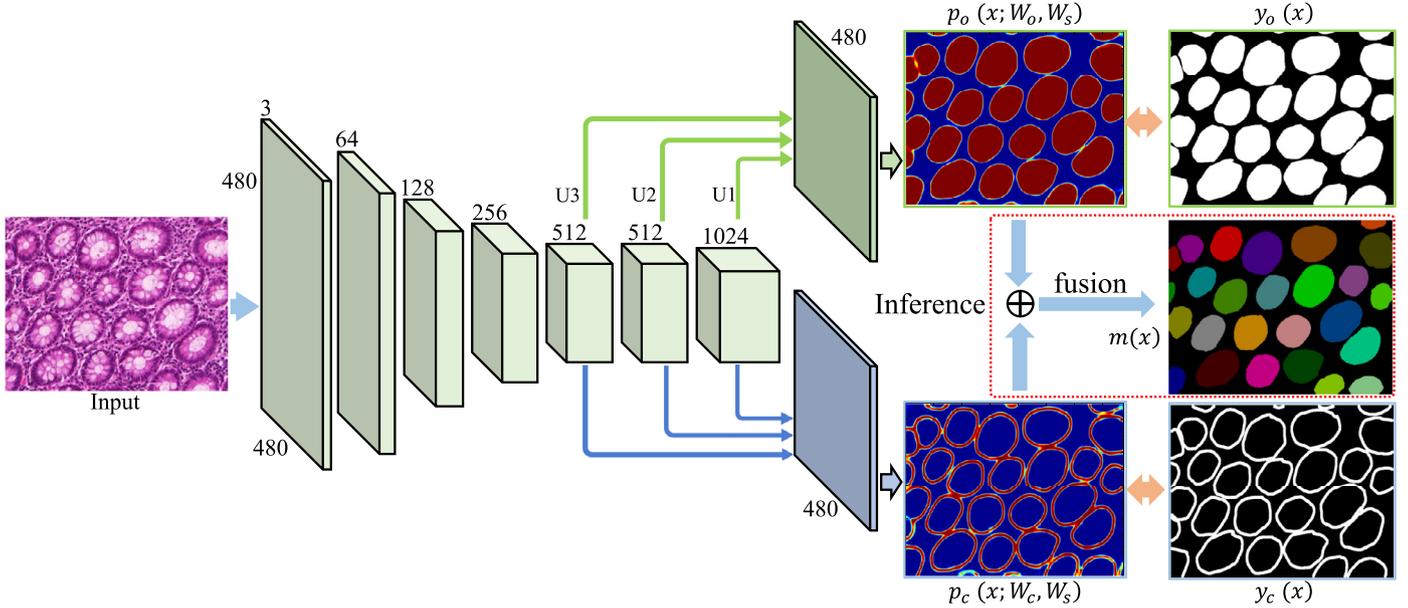
Fig. 5. The overview of the proposed deep contour-aware network. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article).

where W denotes the parameters of neural network including weights and biases; $\psi(W)$ is the regularization term (L_2 norm) with hyperparameter λ for balancing the tradeoff with other terms.

2.2. Deep contour-aware network

By leveraging the multi-level contextual features with auxiliary supervision (Xie and Tu, 2015; Chen et al., 2016b), the network can produce satisfactory probability maps of histological structures. However, it is still quite hard to separate the touching and overlapped object instances by relying only on the likelihood of objects due to the essential ambiguity in clustered regions. This is because while the downsampling path can acquire high level abstraction features, it leads to spatial information loss alone with the abstraction. As we know, in segmentation tasks, the boundary information of objects provides good complementary cues for splitting objects. To this end, we propose to integrate contour information into the FCN to form a deep contour-aware network to more accurately segment the objects in histology images, and, in particular, separate clustered objects into individual ones.

Fig. 5 shows an overview of the proposed deep contour-aware network. Instead of treating the histological structure segmentation and contour detection as single and independent tasks, we formulate them as a multi-task learning framework, which can infer the information of objects and contours simultaneously. Specifically, the feature maps are upsampled with two different branches (as indicated by the green and blue arrows shown in Fig. 5) in order to output the segmentation masks of objects and contours, respectively. In each branch, the mask is predicted by the FCN with multi-level contextual features as introduced in Section 2.1. During the training process, the parameters of downsampling path W_s are shared and updated for both of these two tasks, while the parameters of upsampling layers for two individual branches (denoted as W_o and W_c) are updated independently for inferring the probability of objects and contours, respectively. Therefore, the feature representations through the hierarchical structure can encode the information of segmented objects and contours at the same time. Note that multi-task network is still trained in an end-to-end way.

This joint multi-task learning process has several advantages. First, it can increase the discriminative capability of intermedi-

ate feature representations with multiple regularizations on disentangling subtly correlated tasks (Zhang et al., 2014), hence improve the robustness of the segmentation. Second, in the application of histological structure segmentation, the multi-task learning framework can provide the complementary contour information that serves well to further separate the clustered objects. Last but not least, when dealing with large-scale histopathological data, this unified framework can be quite efficient. With one forward propagation, it can generate the results of objects and contours simultaneously instead of resorting to additional post-separating steps by generating contours based on low-level cues (Gunduz-Demir et al., 2010; WU et al., 2005).

In the training process, the discount weights w_a from auxiliary classifiers are decreased until marginal values with the number of iterations increasing. Therefore, we dropped these terms in the final loss for simplicity. Finally, the training of network is formulated as a per-pixel classification problem regarding the ground truth segmentation masks including both objects and contours:

$$\mathcal{L}_{\text{total}}(x;\theta) = \lambda\psi(\theta) - \sum_{x\in\mathcal{X}}\log p_o(x, y_o(x); W_o, W_s)$$
$$- \sum_{x\in\mathcal{X}}\log p_c(x, y_c(x); W_c, W_s) \tag{2}$$

where the first part is the L_2 regularization term and latter two are the data error loss terms defined with cross-entropy functions; x is the pixel position in image space $\mathcal{X}$; $p_o(x, y_o(x); W_o, W_s)$ denotes the predicted probability for true label $y_o(x)$ (i.e., the index of 1 in one-hot vector encoding) of objects after softmax classification; and $p_c(x, y_c(x); W_c, W_s)$ is the predicted probability for true label $y_c(x)$ of object contours. The parameters $\theta = \{W_s, W_o, W_c\}$ of network are optimized by minimizing the total loss function $\mathcal{L}_{\text{total}}$ with stochastic gradient descent (Williams and Hinton, 1986).

With the predicted probability maps of objects $p_o(x; W_o, W_s)$ and contours $p_c(x; W_c, W_s)$ from the deep contour-aware network, the complementary information are fused together to generate the final segmentation masks $m(x)$:

$$m(x) = \begin{cases} 1 & \text{if } p_o(x; W_o, W_s) \geq t_o \text{ and } p_c(x; W_c, W_s) < t_c \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where t_o and t_c are the thresholds (both set as 0.5 in our experiments empirically). Then, post-processing steps including smoothing with a disk filter (the radius was set as 3), filling holes and removing small spurious segments are performed on the fused segmentation results. Finally, each connected component is labeled with a unique value for representing one segmented object.

2.3. Transfer learning with rich feature hierarchies from cross-domains

When employing deep neural networks in medical image analysis tasks, one of the main challenges is the insufficiency of the high quality training data with accurate annotations (Greenspan et al., 2016), due to the expensive cost and complicated acquisition procedures. Compared with the limited data in medical domain, much more training data can be obtained in natural image recognition tasks. Previous studies (Chen et al., 2015a; Shin et al., 2016; Tajbakhsh et al., 2016; Chen et al., 2016d) have evidenced that transfer learning can alleviate the problem of insufficient training data in deep convolutional networks. The learned parameters (convolutional filters) in the lower layers of convolutional neural networks are inherently generic, while those in higher layers are more specific to different tasks (Yosinski et al., 2014). Thus, transferring the rich feature hierarchies with embedded knowledge learned from cross domains could help to reduce the issue of overfitting on limited medical dataset and further boost the performance.

Therefore, we utilized an off-the-shelf model from DeepLab of Chen et al. (2015d), which was trained on the PASCAL VOC 2012 dataset (Everingham et al., 2010). Compared to the small scale dataset (a few hundred images) in the histopathological image analysis, the PASCAL VOC dataset contains more than ten thousand images with pixel-level annotations. Leveraging the effective generalization ability of transfer learning in deep neural networks, we initialized the layers in downsampling path with pretrained weights from the DeepLab model while the rest layers randomly initialized with Gaussian distribution. Then we fine tuned the whole network on our medical task in an end-to-end way with stochastic gradient descent. In our experiments, we observed the training process converged much faster (about four hours) by virtue of the knowledge learned from rich dataset than the random initialization setting trained from scratch.

3. Experiments and results

We evaluated the proposed method on two challenging object segmentation tasks from histology images: gland segmentation and nuclei segmentation. We participated two challenges held conjunction with MICCAI 2015: *Gland Segmentation (GlaS) Challenge Contest*[2] and *Segmentation of Nuclei in Digital Pathology Images*[3], and ranked the first in both of the challenges. We first introduce the datasets employed in these two challenges and then report the experimental results in detail.

3.1. Datasets and pre-processing

Dataset of gland segmentation challenge. The histology images were acquired by a Zeiss MIRAX MIDI slide scanner from colorectal cancer tissues with a resolution of 0.62 μm/pixel (Sirinukunwattana et al., 2016a). The dataset consisted of a wide range of histologic grades from benign to malignant subjects. It is

worth noting that poorly-differentiated cases were also included to more comprehensively evaluate the performance of participated algorithms. The training data were composed of 85 (benign/malignant = 37/48) images with ground truth annotations provided by experienced pathologists. The testing data contained a total of 80 (benign/malignant = 37/43) images for evaluation. Participants were required to detect and segment the glands from histology images into individual ones. The ground truths of testing data were held out by the challenge organizers for independent evaluation.

Dataset of nuclei segmentation challenge. The training data contained more than 500 manually segmented nuclei in 15 image tiles, which were extracted from a set of Glioblastoma and lower-grade Glioma whole slide tissue images. The testing data included 18 images for evaluation and the ground truths were held out by the challenge organizers for independent evaluation. Participants were asked to detect and segment all the nuclei in the testing images and the results were compared with the ground truth annotations provided by experienced pathologists.

In order to increase the robustness and reduce overfitting in training process, we utilized the strategy of data augmentation to enlarge the training dataset. The augmentation transformations included translation, rotation, and elastic distortion (e.g., pincushion and barrel distortions).

3.2. Results of the gland segmentation challenge

3.2.1. Qualitative evaluation

Some typical segmentation results of the testing images, including both benign cases and malignant cases, are shown in Fig. 6. For diagnosing the role of complementary contour information, i.e., the contour-aware branch network, we performed an ablation study and compared the performance of the proposed DCAN and that of the network relying only on the prediction of gland objects without employing contour information. From the segmentation results, we can observe that the FCN leveraging the multi-level contextual features without contour-aware component can also effectively segment the gland objects in both benign and malignant cases. However, compared with the results of DCAN, there are some touching gland objects that cannot be separated very well. The situation is further deteriorated when the touching objects are clustered together, as shown in the third column of Fig. 6. In contrast, thanks to the contributions of the contour-aware component, the DCAN is capable of separating these touching gland objects into individual instances clearly (as the yellow arrows indicate in the figures), and hence achieves more accurate segmentation results. These examples qualitatively demonstrate the superiority of the DCAN obtained by exploring the complementary information of object appearances and contours under a unified multi-task learning framework.

3.2.2. Quantitative evaluation and comparison

The evaluation metrics employed in the grand challenge include F1 score, object-level Dice index and Hausdorff distance, which reflect the performance of gland detection, segmentation and shape similarity, respectively. We submitted two entries to assess the performance of the proposed methods. The first one, namely *CUMedVision1*, is the multi-level contextual FCN without integrating contour-aware component, as shown in Fig. 4. The second one, namely *CUMedVision2*, is the proposed DCAN, as shown in Fig. 5. In order to probe the performance of submitted methods on histology images with different histologic grades, the results are reported on three categories of testing data: overall data, benign data and malignant data.

Gland detection. For the gland detection evaluation, the metric F1 score is utilized, which is the harmonic mean of precision P and
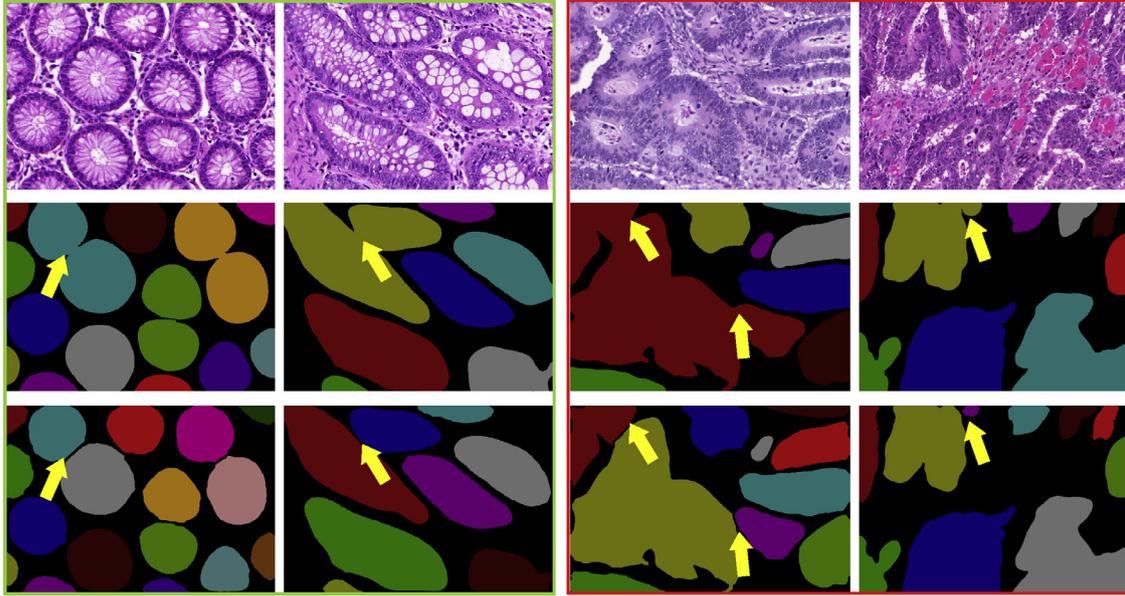
Fig. 6. Segmentation results of benign cases (left two columns) and malignant cases (right two columns). From top to bottom shows the original images, segmentation results of our network without integrating contour information, and segmentation results of our contour-aware network, respectively. Different colors denote individual gland objects. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

Table 1
The detection results (F1 score) of different methods in 2015 MICCAI GlaS Challenge (only top 10 entries are shown here).

Method	Overall	Benign	Malignant
CUMedVision1	0.856	0.872	**0.788**
CUMedVision2	**0.887**	**0.919**	**0.771**
CVML	0.637	0.691	0.439
ExB1	0.868	0.915	0.698
ExB2	0.868	**0.920**	0.686
ExB3	0.875	0.919	0.714
Freiburg2 (Ronneberger et al., 2015)	0.849	0.916	0.616
Freiburg1 (Ronneberger et al., 2015)	0.804	0.893	0.521
LIB	0.704	0.800	0.388
vision4GlaS (Kainz et al., 2015)	0.620	0.651	0.478

recall R, defined as

$$F1 = \frac{2PR}{P + R}, \quad P = \frac{N_{tp}}{N_{tp} + N_{fp}}, \quad R = \frac{N_{tp}}{N_{tp} + N_{fn}} \tag{4}$$

where N_{tp}, N_{fp}, and N_{fn} denote the number of true positives, false positives, and false negatives, respectively. According to the challenge evaluation, the ground truth for each segmented object is the object in the manual annotation that has maximum overlap with that segmented object. A segmented gland object that intersects with at least 50% of its ground truth is considered as a true positive, otherwise it's considered as a false positive. A ground truth gland object that has no corresponding segmented object or has less than 50% of its area overlapped by its corresponding segmented object is considered as a false negative.

Table 1 shows the detection results of various methods participated in the challenge. We just listed the top 10 entries in the table. Note that all the top 5 entries utilized CNN-based methods. For examples, the team *ExB* utilized a two-path network by cropping global and local patches from the image and the segmentation results were generated by fusing these features from these two branches. The team *Freiburg* designed a very deep U-shaped network and has achieved the best results in several other grand detection and segmentation challenges (Ronneberger et al., 2015). Other teams, such as *LIB* and *SUTECH*, utilized hand-crafted structural and textural features. However, the performance of these

methods is much worse than deep learning based methods, as illustrated in Table 1.

Our submitted entry *CUMedVision2* with the proposed DCAN achieved the best results on overall testing data, and competitive performance (ours vs. the first = 0.919 vs. 0.920) on benign testing data, which demonstrated the effectiveness of the proposed DCAN in gland detection. Meanwhile, the other entry *CUMedVision1* employing the multi-level contextual FCN without integrating contour-aware component achieved the best results on malignant testing data, highlighting the advantages of the multi-level contextual FCN with auxiliary supervision. From Table 1, readers can notice that all methods yielded a relatively lower performance on the malignant testing data than the benign data. This is mainly due to the difficulties incurred by the seriously degenerated structures in malignant subjects. Nevertheless, our two submitted entries ranked first and second on the malignant testing data with an obvious advantage over other methods, demonstrating the multi-level contextual features learned from our architecture can more efficiently deal with the large variations of malignant cases. The *CUMedVision2* achieved lower but comparable performance (0.771 vs. 0.788) compared to *CUMedVision1* on malignant testing data. By carefully investigating the results, especially those of malignant cases, we observed that some inaccurate contours in interior structures may cause the deformed glands fragmented in some seriously degenerated cases. Overall, the proposed DCAN achieved the best performance on the evaluation of gland detection according to the standard competitive ranking,[4] which was employed by the challenge organizers.

Gland segmentation. Given a set of pixels G annotated as a ground truth object and a set of pixels S as a segmented gland object, Dice index is often employed for segmentation evaluation, which is defined as $D(G, S) = 2(|G \cap S|)/(|G| + |S|)$. However, Dice index is not suitable for evaluation of segmentation results of individual objects. To the end, an object-level Dice index is utilized in the challenge. It is defined as

$$D_{\text{object}}(G, S) = \frac{1}{2}\left[\sum_{i=1}^{n_S} \omega_i D(G_i, S_i) + \sum_{j=1}^{n_G} \tilde{\omega}_j D(\tilde{G}_j, \tilde{S}_j)\right] \tag{5}$$

---

[4] https://en.wikipedia.org/wiki/Ranking.

Table 2

The segmentation results (object-level Dice index) of different methods in 2015 MICCAI GlaS Challenge.

Method	Overall	Benign	Malignant
CUMedVision1	0.850	0.873	0.826
CUMedVision2	**0.868**	**0.907**	**0.827**
CVML	0.647	0.663	0.630
ExB1	0.858	0.892	0.823
ExB2	0.851	0.902	0.799
ExB3	0.855	0.900	0.810
Freiburg2 (Ronneberger et al., 2015)	0.853	0.903	0.801
Freiburg1 (Ronneberger et al., 2015)	0.852	0.903	0.798
LIB	0.744	0.817	0.668
vision4GlaS (Kainz et al., 2015)	0.705	0.738	0.673

Table 3

The shape similarity results (Hausdorff distance) of different methods in 2015 MICCAI GlaS Challenge.

Method	Overall	Benign	Malignant
CUMedVision1	94.515	67.862	122.187
CUMedVision2	**74.731**	**43.025**	106.979
CVML	160.506	128.250	192.555
ExB1	79.761	53.979	**105.986**
ExB2	89.180	49.330	129.422
ExB3	83.483	52.617	114.722
Freiburg2 (Ronneberger et al., 2015)	79.842	49.115	111.959
Freiburg1 (Ronneberger et al., 2015)	79.326	49.118	111.037
LIB	130.345	78.436	184.510
vision4GlaS (Kainz et al., 2015)	133.158	102.277	164.005

where S_i denotes the ith segmented object, G_i denotes a ground truth object that maximally overlaps S_i, $\tilde{G}_j$ denotes the jth ground truth object, $\tilde{S}_j$ denotes a segmented object that maximally overlaps $\tilde{G}_j$, $\omega_i = |S_i|/\sum_{m=1}^{n_S} |S_m|$, $\tilde{\omega}_j = |\tilde{G}_j|/\sum_{n=1}^{n_G} |\tilde{G}_n|$, n_S and n_G are the total number of segmented objects and ground truth objects, respectively.

Table 2 reports the segmentation results of various methods participating in the challenge according to the object-level Dice index. It is observed that our results of *CUMedVision2* achieved the best performance on all three categories of testing data, outperforming all the other methods.

Shape similarity. The shape similarity is measured by using the Hausdorff distance between the shape of segmented object and that of the ground truth object. It is defined as

$$H(G, S) = \max \left\{ \sup_{x \in G} \inf_{y \in S} \|x - y\|, \sup_{y \in S} \inf_{x \in G} \|x - y\| \right\} \quad (6)$$

Likewise, an object-level Hausdorff is defined for the challenge

$$H_{\text{object}}(G, S) = \frac{1}{2} \left[\sum_{i=1}^{n_S} \omega_i H(G_i, S_i) + \sum_{j=1}^{n_G} \tilde{\omega}_j H(\tilde{G}_j, \tilde{S}_j) \right] \quad (7)$$

The shape similarity results of different methods are shown in Table 3. In comparison to the *CUMedVision1*, the *CUMedVision2*

with contour-aware branch can dramatically reduce the Hausdorff distances (around 20, 25, and 15 pixels decrement for category of overall, benign and malignant testing data, respectively). This highlights the superiority of contour-aware branch in DCAN for separating touching glands into individual object instances. Our results of *CUMedVision2* achieved the smallest Hausdorff distances on both the overall and benign testing data, outperforming other methods by a significant margin. Meanwhile, the results of *CUMedVision2* are comparable to the best results achieved by *ExB1* on malignant testing data (106.979 vs. 105.986 as shown in Table 3).

Final ranking. For the final ranking, each team is assigned one ranking number for each category of testing data based on the three metrics mentioned above using a standard competition ranking.[4] The sum score of these numbers is used for the final ranking, i.e., a smaller score stands for better overall segmentation performance. The final ranking is reported in Table 4 (only top 10 entries are shown). Our deep contour-aware network yielded the best performance in terms of whole results out of 13 teams, outperforming all the other methods by a significant margin. This corroborates the superiority of our method by harnessing object appearance and contour information explicitly under a unified multi-task learning framework.

3.3. Results of the nuclei segmentation challenge

3.3.1. Qualitative evaluation

Some typical nuclei segmentation results of the testing images are shown in Fig. 7. It is observed that both the FCN without integrating contour information and the DCAN can effectively segment most nuclei from the testing images. However, compared with the results of DCAN, there are some touching nuclei that cannot be separated into individual ones in the results of FCN without contour information, as indicated by the yellow arrows overlaid in the images. By integrating the object and contour prediction maps, the DCAN clearly separated the touching nuclei into individual nucleus, and achieved more precise segmentation results.

We also carefully studied the errors in the results. We found that the errors were mostly observed in following cases: high-degree of clustering resulting in under-segmentation, irregular nuclei, noisy background and poor edge information. Although there are still some errors in our results, our method can achieve good detection and segmentation performance on most testing images.

3.3.2. Quantitative evaluation and comparison

The nuclei segmentation challenge employed two metrics for evaluation: traditional Dice coefficient (D_1) and object-level Dice coefficient (D_2). The D_1 metric was applied to measure the amount of overlap between the results of algorithms and human annotations in terms of the nuclei regions that was detected and segmented. D_1 does not take into account the cases of split and merge. A split is the case in which the human segments a region

Table 4

The final ranking of participants in 2015 MICCAI GlaS Challenge (top 10 entries are shown here).

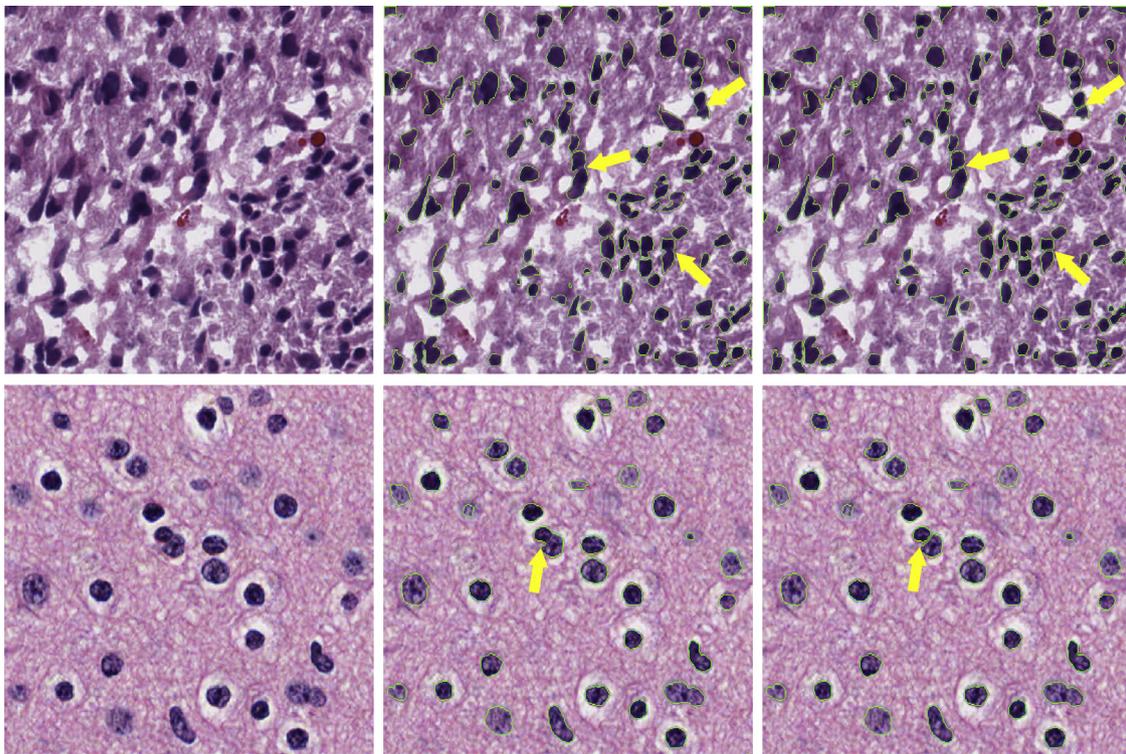| Method	Ranking score			Sum score	Final ranking
	Overall	Benign	Malignant		
CUMedVision2	**1**	**1**	**1**	**3**	**1**
ExB1	2	6	2	10	2
Freiburg2 (Ronneberger et al., 2015)	5	2	5	12	3
ExB3	3	5	4	12	3
Freiburg1 (Ronneberger et al., 2015)	4	4	6	14	5
ExB2	6	3	7	16	6
CUMedVision1	7	7	3	17	7
LIB	8	8	9	25	8
vision4GlaS (Kainz et al., 2015)	9	9	8	26	9
CVML	10	10	10	30	10

Fig. 7. Segmentation results of nuclei. From left to right in each row: original image, segmentation result of our network without integrating contour information, and segmentation result of our contour-aware network, respectively. The boundaries of segmentation masks are overlaid on the original images, and best viewed in color. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

Table 5
Results in 2015 MICCAI nuclei segmentation challenge.

Team	D_1	D_2	Score	Ranking
Ours (with contour-aware component)	**0.876**	**0.748**	**0.812**	**1**
Ours (w/o contour-aware component)	**0.877**	0.722	0.799	2
Team2 (Jbarker)	0.826	0.694	0.760	3
Team3 (Sofiev)	0.792	0.642	0.717	4
Team4 (Czakon)	0.813	0.617	0.715	5

in a single nucleus, but the algorithm segments the same region in multiple nuclei. A merge is the case in which the algorithm segments a region into a single nucleus, but the human segments the same region in multiple nuclei. D_2 is calculated based on the object-level segmentation, which provides a measure of splits and merges. Readers can refer to the challenge website[3] to learn more details of D_1 and D_2. The final ranking score was made by considering both of D_1 and D_2: $score = \frac{D_1+D_2}{2}$.

Table 5 listed the results of the challenge. Compared with other teams, our method without contour-aware component outweighed other teams by a great margin on the performance of D_1 (5% higher than Team2) and D_2 (3% higher than Team2). By incorporating the contour information into the unified framework, we further improved the D_2 by 2.6% (0.748 vs. 0.722) while obtaining similar D_1 (0.876 vs. 0.877). Our DCAN achieved the highest score of 0.812 in the challenge with the proposed multi-level contextual FCN ranking second. These results corroborated the effectiveness of the proposed methods, especially the contour-aware network, for object segmentation tasks in histology images.

4. Implementation details and computation cost

Our framework was implemented based on the open-source deep learning library of Caffe (Jia et al., 2014). The network ran-

domly cropped a 480 × 480 region from the original image as input and output the prediction masks of individual objects and contours. The score masks of whole testing image were generated with an overlap-tile strategy. For the label of contours, we extracted the boundaries of connected components based on the object annotations from pathologists, then dilated them with a disk filter (radius was set as 3 empirically) to enlarge the number of pixels for contours. In the training phase, the learning rate was set as 0.001 initially and divided by a factor of 10 every 10,000 iterations till 10^{-7} (around 40,000 iterations). The discount weight w_a was set as 1 initially and decreased by a factor of 10 every 10,000 iterations until a marginal value 10^{-3}. In addition, dropout layers in Hinton et al. (2012) (dropout rate was set as 0.5) were incorporated in the last two convolutional layers for preventing the co-adaption of intermediate features.

In order to investigate the time performance of the DCAN, we report the computation cost on the gland segmentation challenge. The time performance of the nuclei segmentation is similar to that of the gland segmentation. Generally, it took about four hours to train the deep contour-aware network on a workstation with 2.50 GHz Intel(R) Xeon(R) E5-1620 CPU and a NVIDIA GeForce GTX Titan X GPU. Leveraging the efficient inference of fully convolutional architecture, the average time for processing one testing image with size 755 × 522 was about 1.5 s, which was much faster than other methods in the literature, such as Sirinukunwattana et al. (2015a) and Gunduz-Demir et al. (2010). Considering the large-scale histology images are demanded for prompt analysis with the advent of whole slide imaging, the fast speed implies that our method has great potential to be employed in clinical practice.

5. Discussion

Pathological examinations are regarded as the gold standard in many medical protocols and, in particular, play a key role in can-

cer diagnosis process (Gurcan et al., 2009). Aiming at improving the efficiency and robustness of automated histopathological image analysis, we present a novel deep contour-aware network. Extensive experiments conducted on two challenging histopathological object detection and segmentation tasks demonstrate the effectiveness of the deep contour-aware network as well as the contributions of its components. It is worthwhile to note that the proposed CNN-based method does not employ any prior on object shape along with the cancer status of histopathological images. Therefore, it is general enough to be applied to various objects and cancerous tissues of different histologic grades. This is evidenced by the successes of the proposed method on two different object segmentation tasks with both benign and malignance cases.

It is observed from the detection results on GlaS Challenge that our method without the contour-aware component achieves a little better detection performance than that with the contour-aware component on malignant testing data. Through a careful study, we find it probably arises from the fact that irregular structures in some seriously degenerated malignant cases may reduce the effect of contour information. For example, the interior structures with dense nuclei as a result of high proliferation may lead to false contours, and hence the method more likely fails in such cases. Nevertheless, on the detection evaluation of overall test data as well as benign data, the results of *CUMedVision2* were much better than those of *CUMedVision1*, which validated the effectiveness of introducing contour-aware branch. In addition, the *CUMedVision2* achieved better performance than the *CUMedVision1* in terms of all metrics (including Dice index and Hausdorff distance) in the segmentation task. These results clearly highlighted the superiority of our unified framework by incorporating contour-aware branch. In general, as the histological structures of glands are seriously degenerated in some malignant cases, the segmentation task can be much more challenging. We will further investigate this as part of our future work with more histology images.

The detection and segmentation of clustered object instances are crucially important and have a wide range of applications, such as gland segmentation for colon cancer diagnosis (Sirinukunwattana et al., 2015a), nuclei segmentation for cancer grading (Irshad et al., 2014), and cell segmentation and tracking for understanding cellular mechanisms (Meijering, 2012), etc. Previous methods such as *U-net* in Ronneberger et al. (2015) and SDS in Hariharan et al. (2014) focusing only on appearance information may not perform well in the situation where many clustered object instances exist. Unfortunately, this is a common phenomenon in histology images. It is worth noting that contour detection and object segmentation are two complementary tasks and previous studies showed preliminarily promising results by exploring semantic contours on PASCAL dataset (Bertasius et al., 2015a; Chen et al., 2015c). In this paper, we propose a more efficient and effective solution by harnessing the information of object appearances and contours simultaneously. Furthermore, the efficiency is greatly boosted with our unified framework within one forward propagation (about 1.5 s), which was much faster than state-of-the-art methods (for instance, 200+ s in Sirinukunwattana et al., 2015a).

The separation of clustered histological structures (e.g., nuclei) in high-level density is a long-standing problem, which usually requires to incorporate domain-specific shape prior. While different methods have been developed to segment clustered or overlapping nuclei with great improvements, this problem has not been completely solved yet. Some high-degree clustered nuclei can be further separated with other advanced post-separating steps by incorporating specific shape prior of nuclei, such as maker-controlled watershed algorithm in Cheng and Rajapakse (2009) and spatially adaptive active physical model in Plissiti and Nikou (2012). This topic is out of scope of this paper since our main aim is to segment touching objects and provide initially good segmentation results for high-degree clustered cases. As Al-Kofahi et al. (2010) indicated that the segmentation performance depends crucially on the accuracy and reliability of the initial seed points or shape markers, our method can provide such good probability maps, which can serve as 'seeds' or 'markers' in the subsequent algorithms for delineating the spatial extent of each nucleus.

Another notoriously difficult issue in histopathological image analysis rests with the large variation of tissue appearance, which is regarded as one of the major obstacles for developing a robust automated analysis tool (Veta et al., 2015). The variation is subjected to several factors, including different scanners for image acquisition, different sectioning and staining operations, etc, which are quite common when tissue samples are acquired from different patients or at different time slots. Pre-processing normalization steps (Khan et al., 2014) can potentially address the inconsistent issue of tissue appearance and further improve the performance.

Currently, our method has been evaluated on two applications with hundreds of histology images. In the future, we shall assess it on more large-scale histopathological datasets acquired from various scanners under different conditions.

6. Conclusions

We present a deep contour-aware network that integrates multi-level contextual features to accurately detect and segment histological objects from histology images. We formulate this challenging segmentation problem as a unified multi-task learning process by harnessing the complementary appearance information (such as textures and colors) and contour information explicitly, which further boost the object instance segmentation performance. Extensive experimental results on two challenging object segmentation tasks from histology images demonstrate the superior performance of our method, surpassing state-of-the-art methods by a great margin. The proposed DCAN is inherently general and can be applied to other similar problems in histopathological image analysis. Future investigations include evaluating our method on more histology images and promoting its applications in clinical practice.

Acknowledgments

References

Al-Kofahi, Y., Lassoued, W., Lee, W., Roysam, B., 2010. Improved automatic detection and segmentation of cell nuclei in histopathology images. IEEE Trans. Biomed. Eng. 57 (4), 841–852.

Altunbay, D., Cigir, C., Sokmensuer, C., Gunduz-Demir, C., 2010. Color graphs for automated cancer diagnosis and grading. IEEE Trans. Biomed. Eng. 57 (3), 665–674.

Bartels, P., Thompson, D., Bibbo, M., Weber, J., 1992. Bayesian belief networks in quantitative histopathology. Anal. Quant. Cytol. Histol./Int. Acad. Cytol. Am. Soc. Cytol. 14 (6), 459–473.

Bertasius, G., Shi, J., Torresani, L., 2015a. High-for-low and low-for-high: efficient boundary detection from deep object features and its applications to high-level vision. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 504–512.

Bertasius, G., Shi, J., Torresani, L., 2015. Semantic segmentation with boundary neural fields. CoRR. arXiv preprint arXiv: 1511.02674

Chen, H., Dou, Q., Ni, D., Cheng, J.-Z., Qin, J., Li, S., Heng, P.-A., 2015a. Automatic fetal ultrasound standard plane detection using knowledge transferred recurrent neural networks. In: Proceedings of Medical Image Computing and Computer Assisted Intervention, MICCAI. Springer, pp. 507–514.

Chen, H., Dou, Q., Wang, X., Qin, J., Heng, P.A., 2016a. Mitosis detection in breast cancer histology images via deep cascaded networks. In: Proceedings of thirtieth AAAI Conference on Artificial Intelligence.

Chen, H., Qi, X., Cheng, J.-Z., Heng, P.-A., 2016b. Deep contextual networks for neuronal structure segmentation. In: Proceedings of thirtieth AAAI Conference on Artificial Intelligence.

Chen, H., Qi, X., Yu, L., Heng, P.-A., 2016c. Dcan: deep contour-aware networks for accurate gland segmentation. In: Proceedings of Computer Vision and Pattern Recognition, CVPR.

Chen, H., Shen, C., Qin, J., Ni, D., Shi, L., Cheng, J.C., Heng, P.-A., 2015b. Automatic localization and identification of vertebrae in spine CT via a joint learning model with deep neural networks. In: Proceedings of Medical Image Computing and Computer Assisted Intervention, MICCAI. Springer, pp. 515–522.

Chen, H., Zheng, Y., Park, J.-H., Heng, P.-A., Zhou, S.K., 2016d. Iterative multi-domain regularized deep learning for anatomical structure detection and segmentation from ultrasound images. In: Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 487–495.

Chen, L.-C., Barron, J.T., Papandreou, G., Murphy, K., Yuille, A.L., 2015c. Semantic image segmentation with task-specific edge detection using cnns and a discriminatively trained domain transform. CoRR arXiv preprint arXiv:1511.03328.

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2015d. Semantic image segmentation with deep convolutional nets and fully connected CRFs. In: Proceedings of International Conference on Learning Representations, ICLR.

Cheng, J., Rajapakse, J.C., 2009. Segmentation of clustered nuclei with shape markers and marking function. IEEE Trans. Biomed. Eng. 56 (3), 741–748.

Ciresan, D., Giusti, A., Gambardella, L.M., Schmidhuber, J., 2012. Deep neural networks segment neuronal membranes in electron microscopy images. In: Proceedings of Neural Information Processing Systems, NIPS, pp. 2843–2851.

Cireşan, D.C., Giusti, A., Gambardella, L.M., Schmidhuber, J., 2013. Mitosis detection in breast cancer histology images with deep neural networks. In: Proceedings of Medical Image Computing and Computer-Assisted Intervention, MICCAI 2013. Springer, pp. 411–418.

Cruz-Roa, A.A., Ovalle, J.E.A., Madabhushi, A., Osorio, F.A.G., 2013. A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection. In: Proceedings of Medical Image Computing and Computer-Assisted Intervention, MICCAI 2013. Springer, pp. 403–410.

Dhungel, N., Carneiro, G., Bradley, A.P., 2015. Deep learning and structured prediction for the segmentation of mass in mammograms. In: Proceedings of Medical Image Computing and Computer Assisted Intervention, MICCAI. Springer, pp. 605–612.

Diamond, J., Anderson, N.H., Bartels, P.H., Montironi, R., Hamilton, P.W., 2004. The use of morphological characteristics and texture analysis in the identification of tissue composition in prostatic neoplasia. Hum. Pathol. 35 (9), 1121–1131.

Dou, Q., Chen, H., Jin, Y., Yu, L., Qin, J., Heng, P.-A., 2016a. 3d deeply supervised network for automatic liver segmentation from ct volumes. In: Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 149–157.

Dou, Q., Chen, H., Lequan, Y., Zhao, L., Qin, J., Defeng, W., Vincent, M., Shi, L., Heng, P.A., 2016b. Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks. IEEE Trans. Med. Imaging 35 (5), 1182–1195.

Doyle, S., Madabhushi, A., Feldman, M., Tomaszeweski, J., 2006. A boosting cascade for automated detection of prostate cancer from digitized histology. In: Proceedings of Medical Image Computing and Computer Assisted Intervention, MICCAI. Springer, pp. 504–511.

Dunne, B., Going, J., 2001. Scoring nuclear pleomorphism in breast cancer. Histopathology 39 (3), 259–265.

Elston, C.W., Ellis, I.O., et al., 1991. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology 19 (5), 403–410.

Everingham, M., Van Gool, L., Williams, C.K., Winn, J., Zisserman, A., 2010. The pascal visual object classes (voc) challenge. Int. J. Comput. Vis. 88 (2), 303–338.

Fakhrzadeh, A., Sporndly-Nees, E., Holm, L., Hendriks, C.L.L., 2012. Analyzing tubular tissue in histopathological thin sections. In: Proceedings of 2012 International Conference on Digital Image Computing Techniques and Applications, DICTA. IEEE, pp. 1–6.

Farjam, R., Soltanian-Zadeh, H., Jafari-Khouzani, K., Zoroofi, R.A., 2007. An image analysis approach for automatic malignancy determination of prostate pathological images. Cytometry Part B: Clin. Cytom. 72 (4), 227–240.

Fleming, M., Ravula, S., Tatishchev, S.F., Wang, H.L., 2012. Colorectal carcinoma: pathologic aspects. J. Gastrointest. Oncol. 3 (3), 153–173.

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

Gleason, D.F., 1992. Histologic grading of prostate cancer: a perspective. Hum. Pathol. 23 (3), 273–279.

Greenspan, H., van Ginneken, B., Summers, R.M., 2016. Guest editorial deep learning in medical imaging: overview and future promise of an exciting new technique. IEEE Trans. Med. Imaging 35 (5), 1153–1159.

Gunduz-Demir, C., Kandemir, M., Tosun, A.B., Sokmensuer, C., 2010. Automatic segmentation of colon glands using object-graphs. Med. Image Anal. 14 (1), 1–12.

Gurcan, M.N., Boucheron, L.E., Can, A., Madabhushi, A., Rajpoot, N.M., Yener, B., 2009. Histopathological image analysis: a review. IEEE Rev. Biomed. Eng. 2, 147–171.

Hamilton, P., Anderson, N., Bartels, P., Thompson, D., 1994. Expert system support using bayesian belief networks in the diagnosis of fine needle aspiration biopsy specimens of the breast. J. Clin. Pathol. 47 (4), 329–336.

Hariharan, B., Arbeláez, P., Girshick, R., Malik, J., 2014. Simultaneous detection and segmentation. In: Proceedings of European Conference on Computer Vision, ECCV. Springer, pp. 297–312.

Hinton, G.E., Srivastava, N., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.R., 2012. Improving neural networks by preventing co-adaptation of feature detectors. CoRR arXiv preprint arXiv:1207.0580.

Irshad, H., Veillard, A., Roux, L., Racoceanu, D., 2014. Methods for nuclei detection, segmentation, and classification in digital histopathology: a reviewcurrent status and future potential. IEEE Rev. Biomed. Eng. 7, 97–114.

Irshad, H., et al., 2013. Automated mitosis detection in histopathology using morphological and multi-channel statistics features. J. Pathol. Inform. 4 (1), 10.

Isabella, N., Lu, L., Xiaosong, W., Roth, H.R., Nathan, L., Jianbo, S., Y., T., Summers, R.M., 2016. Automatic lymph node cluster segmentation using holistically-nested networks and structured optimization. In: Proceedings of Medical Image Computing and Computer-Assisted Intervention, MICCAI 2016.

Jacobs, J.G., Panagiotaki, E., Alexander, D.C., 2014. Gleason grading of prostate tumours with max-margin conditional random fields. In: Machine Learning in Medical Imaging. Springer, pp. 85–92.

Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T., 2014. Caffe: convolutional architecture for fast feature embedding. Proceedings of the 22nd ACM international conference on Multimedia. ACM, pp. 675–678.

Jung, C., Kim, C., 2010. Segmenting clustered nuclei using h-minima transform-based marker extraction and contour parameterization. IEEE Trans. Biomed. Eng. 57 (10), 2600–2604.

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

Khan, A.M., Rajpoot, N., Treanor, D., Magee, D., 2014. A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution. IEEE Trans. Biomed. Eng. 61 (6), 1729–1738.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In: Proceedings of Neural Information Processing Systems, NIPS, pp. 1097–1105.

LeCun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. Proc. IEEE 86 (11), 2278–2324.

Lee, C., Xie, S., Gallagher, P., Zhang, Z., Tu, Z., 2015. Deeply-supervised nets. In: Proceedings of International Conference on Artificial Intelligence and Statistics, AISTATS.

Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In: Proceedings of Computer Vision and Pattern Recognition, CVPR, pp. 3431–3440.

Meijering, E., 2012. Cell segmentation: 50 years down the road [life sciences]. IEEE Signal Process. Mag. 29 (5), 140–145.

Naik, S., Doyle, S., Agner, S., Madabhushi, A., Feldman, M., Tomaszewski, J., 2008. Automated gland and nuclei segmentation for grading of prostate and breast cancer histopathology. In: Proceedings of 5th IEEE International Symposium on Biomedical Imaging. IEEE, pp. 284–287.

Nguyen, K., Jain, A.K., Sabata, B., et al., 2011. Prostate cancer detection: fusion of cytological and textural features. J. Pathol. Inform. 2 (2), 3.

Nguyen, K., Sarkar, A., Jain, A.K., 2012. Structure and context in prostatic gland segmentation and classification. In: Proceedings of Medical Image Computing and Computer Assisted Intervention, MICCAI. Springer, pp. 115–123.

Plissiti, M.E., Nikou, C., 2012. Overlapping cell nuclei segmentation using a spatially adaptive active physical model. IEEE Trans. Image Process. 21 (11), 4568–4580.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: convolutional networks for biomedical image segmentation. In: Proceedings of Medical Image Computing and Computer Assisted Intervention, MICCAI. Springer, pp. 234–241.

Roth, H.R., Lu, L., Farag, A., Shin, H.-C., Liu, J., Turkbey, E.B., Summers, R.M., 2015. Deeporgan: multi-level deep convolutional networks for automated pancreas segmentation. In: Proceedings of Medical Image Computing and Computer Assisted Intervention, MICCAI. Springer, pp. 556–564.

Roth, H.R., Lu, L., Farag, A., Sohn, A., Summers, R.M., 2016. Spatial aggregation of holistically-nested networks for automated pancreas segmentation. In: Proceedings of Medical Image Computing and Computer-Assisted Intervention, MICCAI 2016.

Roux, L., Racoceanu, D., Loménie, N., Kulikova, M., Irshad, H., Klossa, J., Capron, F., Genestie, C., Le Naour, G., Gurcan, M.N., 2013. Mitosis detection in breast cancer histological images an icpr 2012 contest. J. Pathol. Inform. 4, 1–4.

Sabata, B., Babenko, B., Monroe, R., Srinivas, C., 2010. Automated analysis of pin-4 stained prostate needle biopsies. In: Prostate Cancer Imaging. Springer, pp. 89–100.

Shin, H.-C., Roth, H.R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., Summers, R.M., 2016. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Trans. Med. Imaging 35 (5), 1285–1298.

Sirinukunwattana, K., Pluim, J.P.W., Chen, H., Qi, X., Heng, P.-A., Guo, Y.B., Wang, L.Y., Matuszewski, B.J., Bruni, E., Sanchez, U., and others, 2016a. Gland segmentation in colon histology images: the GlaS challenge contest. arXiv preprint arXiv:1603.00275.

Sirinukunwattana, K., Raza, S.E.A., Tsang, Y.-W., Snead, D.R.J., Cree, I.A., Rajpoot, N.M., 2016b. Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. IEEE Trans. Med. Imaging 35 (5), 1196–1206.

Sirinukunwattana, K., Snead, D., Rajpoot, N., 2015a. A stochastic polygons model for glandular structures in colon histology images. IEEE Trans. Med. Imaging 34 (11), 2366–2378.

Sirinukunwattana, K., Snead, D.R., Rajpoot, N.M., 2015b. A novel texture descriptor for detection of glandular structures in colon histology images. SPIE Medical Imaging. International Society for Optics and Photonics 94200S-94200S.

Stierer, M., Rosen, H., Weber, R., 1991. Nuclear pleomorphism, a strong prognostic factor in axillary node-negative small invasive breast cancer. Breast Cancer Res. Treat. 20 (2), 109–116.

Tabesh, A., Teverovskiy, M., Pang, H.-Y., Kumar, V.P., Verbel, D., Kotsianti, A., Saidi, O., 2007. Multifeature prostate cancer diagnosis and gleason grading of histological images. IEEE Trans. Med. Imaging 26 (10), 1366–1378.

Tajbakhsh, N., Shin, J.Y., Gurudu, S.R., Hurst, R.T., Kendall, C.B., Gotway, M.B., Liang, J., 2016. Convolutional neural networks for medical image analysis: full training or fine tuning? IEEE Tran. Med. Imaging 35 (5), 1299–1312.

Veta, M., Huisman, A., Viergever, M.A., van Diest, P.J., Pluim, J.P., 2011. Marker-controlled watershed segmentation of nuclei in h&e stained breast cancer biopsy images. In: Proceedings of 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE, pp. 618–621.

Veta, M., Van Diest, P.J., Willems, S.M., Wang, H., Madabhushi, A., Cruz-Roa, A., Gonzalez, F., Larsen, A.B., Vestergaard, J.S., Dahl, A.B., et al., 2015. Assessment of algorithms for mitosis detection in breast cancer histopathology images. Med. Image Anal. 20 (1), 237–248.

Weind, K.L., Maier, C.F., Rutt, B.K., Moussa, M., 1998. Invasive carcinomas and fibroadenomas of the breast: comparison of microvessel distributions–implications for imaging modalities. Radiology 208 (2), 477–483.

Wienert, S., Heim, D., Saeger, K., Stenzinger, A., Beil, M., Hufnagl, P., Dietel, M., Denkert, C., Klauschen, F., 2012. Detection and segmentation of cell nuclei in virtual microscopy images: a minimum-model approach. Sci. Rep. 2, 503–510.

Williams, D.R.G.H.R., Hinton, G., 1986. Learning representations by back-propagating errors. Nature 323, 533–536.

WU, H.-S., Xu, R., Harpaz, N., Burstein, D., Gil, J., 2005. Segmentation of intestinal gland images with iterative region growing. J. Microsc. 220 (3), 190–204.

Xie, S., Tu, Z., 2015. Holistically-nested edge detection. In: Proceedings of International Conference on Computer Vision, ICCV, pp. 1395–1403.

Xu, J., Xiang, L., Liu, Q., Gilmore, H., Wu, J., Tang, J., Madabhushi, A., 2016. Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images. IEEE Trans. Med. Imaging 35 (1), 119–130.

Xu, Y., Li, Y., Liu, M., Wang, Y., Lai, M., Chang, E.I., et al., 2016. Gland instance segmentation by deep multichannel side supervision. In: Proceedings of Medical Image Computing and Computer-Assisted Intervention, MICCAI 2016.

Yosinski, J., Clune, J., Bengio, Y., Lipson, H., 2014. How transferable are features in deep neural networks? In: Proceedings of Neural Information Processing Systems, NIPS, pp. 3320–3328.

Zhang, Z., Luo, P., Loy, C.C., Tang, X., 2014. Facial landmark detection by deep multi-task learning. In: Proceedings of European Conference on Computer Vision, ECCV. Springer, pp. 94–108.

Zheng, Y., Liu, D., Georgescu, B., Nguyen, H., Comaniciu, D., 2015. 3d deep learning for efficient and robust landmark detection in volumetric data. In: Proceedings of Medical Image Computing and Computer-Assisted Intervention, MICCAI 2015. Springer, pp. 565–572.