

Deep learning in digital pathology image analysis: a survey

Shujian Deng^{1,2,3,*}, Xin Zhang^{1,2,3,*}, Wen Yan^{1,2,3}, Eric I-Chao Chang⁴, Yubo Fan^{1,2,3}, Maode Lai⁵, Yan Xu (✉)^{1,2,3,4}

¹School of Biological Science and Medical Engineering, Beihang University, Beijing 100191, China; ²Key Laboratory of Biomechanics and Mechanobiology of Ministry of Education and State Key Laboratory of Software Development Environment, Beihang University, Beijing 100191, China; ³Beijing Advanced Innovation Center for Biomedical Engineering, Beihang University, Beijing 100191, China; ⁴Microsoft Research Asia, Beijing 100080, China; ⁵Department of Pathology, School of Medicine, Zhejiang University, Hangzhou 310007, China

© Higher Education Press 2020

Abstract deep learning (DL) has achieved state-of-the-art performance in many digital pathology analysis tasks. Traditional methods usually require hand-crafted domain-specific features, and DL methods can learn representations without manually designed features. In terms of feature extraction, DL approaches are less labor intensive compared with conventional machine learning methods. In this paper, we comprehensively summarize recent DL-based image analysis studies in histopathology, including different tasks (e.g., classification, semantic segmentation, detection, and instance segmentation) and various applications (e.g., stain normalization, cell/gland/region structure analysis). DL methods can provide consistent and accurate outcomes. DL is a promising tool to assist pathologists in clinical diagnosis.

Keywords pathology; deep learning; segmentation; detection; classification

Introduction

In clinical medicine, pathological examination has been regarded as a gold standard for cancer diagnosis for more than 100 years [1]. Pathologists use a microscope to observe a histological section. Many advanced technologies, including hematoxylin and eosin (H&E) staining and spectral methods, have been applied in the preparation of tissue slides to improve imaging quality. However, intra- and interobserver disagreement cannot be avoided through visual observation and subjective interpretation, especially for experienced pathologists [2,3]. The limited agreement has resulted in the necessity of computational methods for pathological diagnosis [4–16] because automatic approaches can attain robust performance. The first step for computer-aided analysis is digital imaging.

Digital imaging is the process of acquiring, compressing, storing, and displaying scenes digitally. Whole slide imaging is a more advanced and frequently used technology in pathology compared with traditional digital imaging technologies that process static images through cameras [17–19]. This technology involves two processes.

A specialized scanner is utilized to convert an entire glass histopathology or cytopathology slide into a digital slide. A virtual slide viewer is used to visualize the collected digital files. This technique can efficiently generate high-resolution whole slides. A whole slide image (WSI) typically contains trillions of pixels ranging from 200 MB to 1 GB [20]. The size of one patient's data generated from a biopsy can reach TB level because a WSI records the information of the entire tissue section. The WSI immensely promotes systematic image analysis toward microscopic morphology.

Digital pathology (DP) analysis aims to determine the degree of tissue cancerization and predict its prognosis. Traditional computational methods objectively evaluate disease-related tissue changes by extracting mathematical features, such as textural [6,7,9,12], morphological [4,5], structural [6,13–16], and fractal features [8,11]. However, staining, fixation, and sectioning difference can cause huge variances, thereby causing difficulties to these hand-crafted features with limited generalization. Deep learning (DL) methods provide domain agnostic solutions that are ideally suitable for DP image analysis tasks.

DL is a popular machine learning method based on a deep neural network. A deep neural network [21] is composed of multiple nonlinear modules and can learn representations from raw data without domain expertise. As a representation learning framework, multilevel

Received August 19, 2019; accepted March 5, 2020

Correspondence: Yan Xu, xuyan04@gmail.com

*Equal co-authorship, ordered alphabetically.

features are extracted during the multilayer propagation. Network internal parameters are randomly initialized at the beginning and updated through backpropagation (backward propagation of errors). DL can discover intricate structures in large data sets. Many training data usually result in enhanced performance. A convolutional neural network (CNN) is a popular deep neural network [22]. It was first proposed by LeCun *et al.* [23]. CNN is suitable for processing images in the form of 2D or 3D arrays. Convolutional and pooling layers are two typical components. The convolutional layer extracts feature maps from previous layers through the convolutional kernel. Every unit of the output feature map is the multiply-accumulate of the convolutional kernel and the local input area. The same kernel is shared among all the units in a feature map, making the network invariant to target locations. The concept of weight sharing and receptive field can reduce model complexity and the number of weights, thereby deepening the depth of CNN. The pooling layer is used to merge neighboring features from the previous convolutional layer, thereby immensely reducing the feature dimension and increasing the robustness to small shifts. Nonlinearity, such as a rectified linear unit module, is another important component that enables the network to emulate complex functions [24]. CNN takes raw images (or large patches) as input to avoid the complex feature extraction in conventional recognition algorithms. CNN is highly invariant to translation, scaling, inclination, and other forms of deformation. It has brought great breakthroughs in image processing and is widely applied in a vast variety of computer vision tasks. Histopathology images are characterized by data complexity, making deep architectures extremely suitable for complex feature learning in DP analysis [25].

This paper systematically reviews the research directions of DL in DP research: (1) Image preprocessing—stain normalization. Stain variation can seriously deteriorate training results. Color normalization is a prerequisite step for automatic algorithms. (2) Obtaining clinical/biological structure information. Qualitative and quantitative histology characteristics are the key indicators used in cancer evaluation [26,27]. The extraction and analysis of low-(cell), middle-(gland), and high-level (region) objects

(Fig. 1) through classification, semantic segmentation, detection, and instance segmentation are extensively discussed. (3) Grading and prognosis. DL methods that either extract valid pathological information to facilitate subsequent survival prediction and treatment suggestions or conduct cancer grading on the WSI are comprehensively presented. We aim to provide readers with either medical or engineering backgrounds with thorough and detailed cases in the intersection of DL and medical image processing.

Stain normalization

Original tissue sections are visually transparent under the microscope. Efficient examination requires dyeing tissue sections with colored histochemical stains. Stains enhance the contrast among various structures by selectively binding to particular cellular components. However, color variation exists in histopathology images because of the uncertainty of multiple factors, such as (1) scanner type, (2) stain concentration, (3) manufacturer, (4) time elapse, (5) environmental temperatures upon staining, and (6) digitization. Undesired color variations will affect the visual examinations of tissues by pathologists and automatic analysis of DP images by software. Stain normalization (also known as color normalization) that basically transfers the mean color from one source image to others is developed to alleviate interimage biases. Although DL algorithms can partially mitigate color variation through proper data augmentation, the result performance deteriorates without stain normalization because of a limited amount of data. Therefore, stain normalization acts as a prerequisite for pathological image analysis.

Various approaches [28–31] have been proposed for stain normalization. Khan *et al.* [28] presented a supervised method based on nonlinear mapping using a representation derived from color deconvolution. Principal color histograms were derived from a set of quantized histograms. A global stain color descriptor was then computed. Thus, a supervised classification framework was learned to calculate image-specific stain matrices.

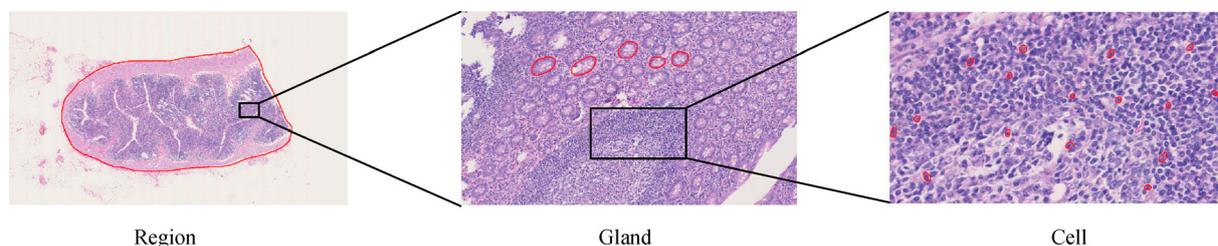


Fig. 1 Multilevel WSI objects (delineated in red lines).

Vahadane *et al.* [29] proposed a structure-preserving method for stain separation and color normalization to preserve biological structure information by modeling stain density maps on the basis of nonnegativity, sparsity, and soft classification. Janowczyk *et al.* [30] presented an algorithm based on an unsupervised neural network. They explored sparse autoencoders (SAEs) through an iterative process to partition similar tissue types in source and target images. The learned filters can optimally reconstruct the original image. The color of the target image was then altered to match the source image tissue through histogram specification. Bentaieb and Hamarneh [31] proposed a different approach based on generative adversarial networks. They cast stain normalization as a style transfer problem to transfer the staining appearance of tissue images across data sets. A recent review summarized the pros and cons of various stain normalization methods for histopathology images [32].

Cell level

Cellular objects are the frequently used biomarkers for cancer histology diagnosis [33,34]. In accordance with the widely-used Nottingham–Bloom–Richardson grade for breast cancer screening [33], tubule formation, nuclear pleomorphism, and the number of mitotic figures are essential for the assessment of breast cancer. Nuclei features, such as spatial distribution and morphological appearance, play an important role in identifying mitotic index and nuclear pleomorphism level [2,33–35]. The number of mitoses is a critical predictor of tumor aggressiveness and is of great importance for cancer screening and assessment. Manual cell segmentation and mitosis counting are time consuming and labor intensive for pathologists. The results are usually subjective, exhibiting intra- and interindividual variability. Thus, the development of automatic segmentation and detection methods is essential for efficient and reliable pathological diagnosis. However, accurate detection and segmentation of nuclei experience the following difficulties: (1) many cells touch each other with weak nuclear boundaries, making them difficult to discriminate; (2) variability exhibits in sizes, shapes, textures, and appearances of individual nuclei; (3) background clutter, stain imbalance, and image artifacts unavoidably exist. In addition to the general difficulties of nuclei analysis in the histopathology image, the mitosis analysis shows some unique challenges. The mitosis exhibits diverse shape configurations in different growth stages. Apoptotic nuclei look similar to mitoses, thereby causing false positive during detection [36]. A number of DL-based approaches have been proposed to achieve enhanced performance. Here, we first review the application of DL in cell and nucleus semantic segmentation. We then summarize the develop-

ment of cell and mitosis detection. We discuss nuclei instance segmentation, which is a combined task of nuclei semantic segmentation and detection. Table 1 presents an overview of each task in the cell-level analysis.

Cell and nucleus semantic segmentation

The segmentation task, which aims to assign a class label to each pixel of an image, is a common task in pathology image analysis. Neural membrane segmentation in electron microscopy (EM) images is the primary task for automated pipeline reconstruction and mapping neural connections [65]. Cellular object segmentation is a prerequisite step for cellular morphology computation, characteristic quantification, and cell recognition. In glioblastoma multiforme, round oligodendroglioma can be distinguished from elongated and irregular morphology of astrocytoma with the assistance of nuclear segmentation [39]. In cervical cytology diagnosis, nuclear segmentation is necessary to discover all types of cytological abnormalities that are usually identified by certain nuclear abnormality [66]. Therefore, the development of accurate automatic nuclear segmentation methods is essential for computer-assisted diagnosis to ease pathologists' burden of manual inspection and alleviate the missed diagnosis and misdiagnosis. Previous studies [67–69] for nucleus or cell segmentation focused on region growth, threshold, clustering, level set, supervised color-texture-based, and watershed-based methods. These conventional approaches depend on elaborately designed features or representations that require intense domain knowledge and are unadaptable to different circumstances. DL approaches has been increasingly applied for automated nuclear segmentation. CNN is initially used in classification tasks, where the input is an image, and the corresponding output is a single class label. Hence, some studies have regarded segmentation tasks as pixelwise classification tasks [37,39,40]. These methods are based on fixed pipeline steps. First, a group of windows is densely sampled from the WSI via a sliding window. The central pixel is then predicted to a certain class at a probability by utilizing the rich context information of the window. This pixel is assigned a label, such as “foreground” or “background” in accordance with the fixed threshold. Finally, the whole image is segmented in accordance with the labels of all the pixels. DL methods function as the feature extractor and classifier during the pixel-level classification. Cireşan *et al.* [37] first applied a CNN-based method to neural membrane segmentation in EM stacks using the above strategy. They achieved the best performance at *ISBI 2012 EM Segmentation Challenge* [65]. Similarly, Zhou *et al.* [39] proposed an approach for nuclear segmentation. They randomly cropped image patches at nuclear centers and jointly learned a set of convolutional filters and a sparse linear regressor. The segmentation results were obtained by applying a threshold

Table 1 Overview of papers using deep learning for cell-level DP image analysis

Reference	Year	Object	Organ	Staining/modality	Architecture	Methods
Semantic segmentation of cellular objects						
[37]	2012	Neuronal membranes	Brain	EM	CNN	Ensemble of several CNNs with different architectures
[38]	2015	Neuronal cell membranes	Brain cervix	EM LM	U-net	U-net with deformation augmentation
[39]	2015	Nuclei	Brain	H&E	CNN	Sparsity constrained convolutional regression
[40]	2015	Cytoplasm and nuclei	Cervix	H&E	CNN	Multi-scale CNN and graph partitioning-based method
[41]	2016	Nuclei	Brain/pancreas/breast	H&E	CNN	CNN and selection-based sparse shape model
[42]	2016	Nuclei	Breast	H&E	CNN	Sparse reconstruction and CNN and morphological operations
[43]	2017	Nuclei	Cervix	H&E	FCN	FCN and graph-based approach
[44]	2018	Nuclei	-	FL	U-net	U-net with recurrent residual convolutional operations
Cell/nuclei detection						
[45]	2015	Cell	Breast cervix	H&E PC	CNN	CNN-based structured regression model
[46]	2016	Nuclei	Colorectal adenocarcinomas	H&E	CNN	Spatially constrained CNN for regression
[47]	2016	Nuclei	Breast	H&E	SAE	Stack SAE
[48]	2018	Cell	Breast cervix neuroendocrine tumor	H&E PC Ki-67	FCN	Structured regression model based on residual FCN
Mitosis detection						
[49]	2013	Mitosis	Breast	H&E	CNN	CNN-based pixel classifier
[50]	2014	Mitosis	Breast	H&E	CNN	Combination of CNN and hand-crafted features
[36]	2016	Mitosis	Breast	H&E	FCN	Deep regression network
[51]	2016	Mitosis	Breast	H&E	FCN	Deep cascaded networks
[52]	2016	Mitosis	Breast	H&E	CNN	Incorporated "crowdsourcing" layer into the CNN frame work
[53]	2018	Mitosis	Breast	H&E	FCN	Proposal-based deep detection network
[54]	2018	Mitosis	Breast	H&E	CNN	Multi-scale and similarity learning CNN
[55]	2019	Mitosis	Breast	H&E	FCN	Weakly supervised FCN using concentric loss
Cell/nuclei instance segmentation						
[56]	2016	Cell	Mouse brain	TPM	U-net	U-net with an iterative k-terminal cut algorithm
[57,58]	2016	Cell	Cervix	FL	FCN	Proposal-based CNN model
[59]	2017	Nuclei	Brain	H&E	FCN	Deep contour-aware network
[60]	2017	Cell	Cervix	H&E	CNN	Multi-scale CNN and multiple cell labeling and dynamic multitemplate deformation model
[61]	2017	Nuclei	Multiple organs	H&E	CNN	Three-class CNN with nuclear boundaries emphasized
[62]	2018	Nuclei	Rat kidney	FL	CNN	3D distance transform and CNN
[63]	2019	Nuclei	Multiple organs	H&E	U-net	U-net with distance map regression
[64]	2019	Nuclei	Multiple organs	H&E	FCN	Contour-aware informative aggregation

The staining and imaging modality used in the table are abbreviated as follows: H&E, hematoxylin and eosin staining; EM, electron microscopy; LM, light microscopy; PC, phase contrast; FL, fluorescent; TPM, two-photon microscopy.

to the regressor prediction in performing pixelwise binary classification. For most single-scale CNN-based methods, a small patch size cannot provide sufficient context information for learning effective features. Windows with large sizes are also unfeasible to achieve the goal because of the high computational cost. Song *et al.* [40] implemented a multi-scale CNN framework to segment cervical cytoplasm for solving the scale variations in the nuclei. They first extracted feature representations via multi-scale CNN and obtained coarse segmentation by inputting the features into a two-layer neural network. A graph partitioning model was then implemented on the basis of coarse segmentation and superpixels for fine segmentation. Finally, they computed nuclear markers to split touching nuclei. A similar approach that combines CNN features with morphological operations for nuclei segmentation was presented in Reference [42]. Xing *et al.* [41] used a CNN model to learn the feature representation of raw data and to generate a probability map. Each pixel was assigned a probability indicating its probability that it belongs to a nucleus. An iterative region merging approach was then performed to initialize the contours. Accurate nucleus segmentation was achieved via a selection-based sparse shape model. Although the above pixel-based methods [37,39–42] have shown promising performance, obvious limitations are found. First, the patch size in pixel-based methods is small. Although this property can increase the locating accuracy, rich context information cannot be fully utilized. Second, densely selected patches increase the calculation burden, making the algorithm cumbersome and time consuming. Finally, these methods usually rely on prior assumptions on the target structure, thereby reducing the generalization to targets with different morphologies. Semantic segmentation tasks have become asymptotically and absolutely efficient with the success of fully convolutional network (FCN) [70]. This end-to-end network consists of a downsampling path and an upsampling path. The downsampling path extends the classification network structure, which is composed of convolutional and pooling operations. The upsampling path includes convolutional and deconvolutional layers that can produce dense classification scores for a whole image. The full image is fed as input rather than multiple dense patches when conducting training and testing. A pixelwise probability map will be outputted with one single forward propagation. Zhang *et al.* [43] combined FCN and graph-based approach for cervical nuclei segmentation. The FCN acted as a nucleus high-level feature learner and generated a nucleus label mask and a nucleus probabilistic map. The former was used to guide the graph construction, and the latter was formulated into an in-region cost function. They achieved state-of-the-art Zijdenbos similarity index of 0.92 ± 0.09 . Ronneberger *et al.* [38] innovatively modified and extended the FCN by

adding a symmetric expanding path along with a usual contracting path. High-resolution features from the contracting path and the upsampled output from the expanding path were combined to enable precise localization. The resulted architecture was named U-net, and many varieties were subsequently invented for multiple medical image tasks [44,71,72]. Without pre- and postprocessing, FCN [70] and its variants, such as U-net [38], are proved to be well-suited in medical image segmentation tasks.

Cell and mitosis detection

A detection task aims to find and locate the regions of interests (ROIs), such as nuclei or mitoses in a tissue slide and is of great importance to cancer screening. Cell spatial distribution analysis and mitosis counting can help in distinguishing differentiation degrees [73]. Automatic cell/nuclei detection serves as an essential prerequisite for a series of subsequent tasks, such as cell/nuclei instance segmentation, tracking, and morphological measurements [74]. Object detection is necessary to automatically acquire spatial distribution and quantity information. A detection task requires classifying and positioning the targets. Many traditional methods based on hand-crafted features have been proposed [75–78] to solve this problem. A number of DL-based studies have been conducted on this field in recent years.

In recent studies toward nuclei detection, feature representations learned by DL-based methods have been demonstrated more effective than hand-crafted features [79]. DL-based nuclei detection approaches can be divided into two categories: (1) conduct per-pixel classification via a sliding window to find local maxima on a probability map. Xu *et al.* [47] applied a stack SAE to learn high-level features using such strategy and fed into a softmax classifier to classify each patch as nuclear or nonnuclear. Their method achieved promising performance on breast cancer histopathology images. Sirinukunwattana *et al.* [46] proposed a spatially constrained CNN to generate a probability mask for spatial regression. Their model predicted the centroid locations of nuclei and their confidence whether they corresponded to true centroids; (2) regress the proximity value for each pixel. Xie *et al.* [45] presented a CNN-based structured regression approach. Their method generated proximity patches that show high values for pixels near cellular centers rather than a single class label assigned to an input image patch. In this way, each cell was detected. However, the model with a large-size proximity patch was difficult to train because the subsampling operation resulted in information loss. In their another work [48], they extended the previous method with a fully residual CNN to directly output a dense proximity prediction that had the same size as the input image. The experimental results showed the superiority of their

method in terms of detection accuracy and speed.

Given the importance of mitosis counting for cancer prognosis and treatment, some challenges or contests have been encountered in developing robust algorithms on mitosis detection [80–82]. Most methods regarded this task as follows: (1) as a classification problem that relies on sliding window to classify whether an image patch is centered with the mitosis; (2) as a semantic segmentation task that first segments the images to find candidates of mitoses and then classifies these candidates to obtain the final detections; (3) as a proposal-based detection issue that produces a number of potential proposals on the image patch and eliminates the most possible ones.

(1) As a classification problem. Cireřan *et al.* [49] utilized CNN as a pixelwise classifier via a sliding window manner to detect the mitoses in breast histology images. Their method achieved the best performance in the *2012 ICPR Mitosis Detection Challenge* [81]. Wang *et al.* [50] performed mitosis detection in breast cancer pathology images with the combination of CNN and hand-crafted features. They presented a cascaded approach that possibly maximized the exploiting of the two distinct feature sets. Such a strategy exhibited higher accuracy and less demanding computational cost compared with each individual feature extraction technique. To acquire sufficient training images, crowdsourcing was introduced in the learning process of CNN to exploit additional data sources annotated by nonexpert users for mitosis detection [52]. They trained a multi-scale CNN model with the same basic architecture on different image scales to perform mitosis detection and provided the crowds with mitosis candidates for annotation. The collected annotations were then passed to the existing CNN with the specially designed aggregation layer attached for model refinement and ground truth generation. These methods are usually computationally demanding.

(2) As a semantic segmentation task. Chen *et al.* [36] proposed a deep regression network based on FCNs. Their method can produce a dense score map with the same size as the input image and can be trained in an end-to-end fashion. The regression layer was added at the end of the FCN architecture, and the proximity score map was defined to efficiently locate the mitotic centroid on the map. Their method was concise and general to be applied to other similar tasks. They also utilized a transfer learning strategy from cross-domain to alleviate the insufficiency of medical data. In their another study [51], they used two networks to build a cascade detection system. They first exploited FCN to locate the candidates of mitoses. A second fine discrimination model was then applied to classify the candidate patches. Li *et al.* [55] expanded the mitosis labels that were usually single pixel to labels with concentric circles, where the inside circle was a mitotic region, and the outside ring was a “middle ground.” Thus,

a concentric loss was defined to detect mitosis from a segmentation model.

(3) As a proposal-based issue. Li *et al.* [53] applied a proposal-based deep detection network to the mitosis detection task and utilized a patch-based deep verification network to improve the accuracy.

Strategy 1 (classification-based) is largely applied in previous DL-based studies. The advancement of computational power and DL technologies have promoted Strategy 2 (semantic segmentation-based) to be the mainstream. Strategy 3 (proposal-based) is becoming increasingly popular with the success of detection algorithms developed in natural scenes.

Nuclei instance segmentation

Instance segmentation is a unified work that groups pixels into semantic classes and object instances [83]. Compared with detection, instance segmentation provided finer segmentation rather than a box for each independent object. Unlike semantic segmentation, foreground pixels are separated into different individuals [84]. Nuclei instance segmentation, which aims to differentiate nuclear regions from the background and separate them into different individuals, is fundamental to acquire many grading indexes. For example, the average nuclear size estimated by nuclei instance segmentation results is an important indicator used for nuclear pleomorphism scoring [85] in manual [86,87] and automatic [88] measurements. Increasing explorations have been conducted for nuclei instance segmentation.

Some public data sets [61,89–91], such as the Fluorescence-2D-HeLa data set from ISBI cell tracking challenge [89], the Computational Precision Medicine Nuclear Segmentation Challenge at MICCAI 2015 and MICCAI 2017 [90], the Triple Negative Breast Cancer [91], and the Multiorgan Nuclei Segmentation (MoNuSeg) Data set [61], are available to facilitate the development of nuclei instance segmentation algorithms. The cell/nucleus instance segmentation approaches are classified into two strategies, namely, (1) semantic segmentation-based: apply semantic segmentation with traditional image processing methods (image labeling, watershed, etc.) to separate nuclei into different individuals; (2) detection-based: present potential seeds (centers of each nucleus) or proposals (bounding boxes surrounding each nucleus) initially and predict segmented masks on the basis of detection outputs.

(1) Semantic segmentation-based. To separate touching nuclei into individual ones, Chen *et al.* [59] proposed to integrate contour information into a multilevel FCN for developing a deep contour-aware network. Their method won the *2015 MICCAI Nuclei Segmentation Challenge* [90]. Kumar *et al.* [61] explicitly introduced a third class of

nuclear boundary (besides foreground that is inside any nucleus and background that is outside every nucleus) to discriminate different nuclear instances. Song *et al.* [60] first relied on a semantic FCN to classify each pixel into the class of nuclei, cytoplasm, and background. They then defined overlapped cell splitting as a discrete labeling task and designed a suitable cost function to split crowded cells. They applied a dynamic multitemplate deformation model for cell boundary refinement. Similarly, Yang *et al.* [56] achieved instance-level segmentation of glial cells in 3D images by first obtaining voxel-level segmentation via FCN. Subsequently, they adopted a k -terminal cut algorithm to separate touching cells. Naylor *et al.* [63] predicted the distance map of nuclei with U-net and a regression loss. Postprocessing was conducted on the basis of the output of the regression network. Zhou *et al.* [64] achieved the highest ranking in the *2018 MICCAI MoNuSeg Challenge* [92]. Their method aggregated multilevel information between two task-specific decoders. With bidirectionally combined task-specific features, they leveraged the spatial and texture dependencies between the nuclei and contours. Zeng *et al.* [72] applied residual-inception-channel attention-U-net. Their model output contour and nuclei masks to achieve instance segmentation. The above methods tend to obtain semantic segmentation results as a premise and implicitly or explicitly introduce contour priors (e.g., as the additional label, encoded in the distance loss, etc.) to disentangle nucleus-to-nucleus connections. Such methods show simplicity and fastness. However, some disadvantages are found. Oversegmentation will lead to additional identified nuclei instances. Undersegmentation will fail to split touching nuclei accurately. These methods usually experience heavily-engineered post-processing, making them difficult to be generalized. Nuclei with irregular boundaries cannot be well-handled.

(2) Detection-based. Akram *et al.* [57,58] designed a two-stage network for nuclei instance segmentation. In the first stage, a FCN model proposed possible bounding boxes for cell proposals and their scores. Nonmaxima suppression [93] was then implemented to remove low-scored duplicate proposals. In the second stage, segmentation masks were obtained through thresholding and morphological operations [57] or predicted using a second CNN for each proposed bounding box [58]. Ho *et al.* [62] proposed a 3D CNN framework for nuclei detection and segmentation in fluorescence microscopy images. They first used the 3D distance transform [94] to individually detect seeds and a 3D CNN segmented nuclei within a subvolume centered at each seed. Region-based object detection algorithms [95,96] have been increasingly used for nuclei instance segmentation [97,98]. Detection-based approaches usually exhibit high performance but cost considerable time and may include certain false positives or false negatives.

Gland level

A gland is a group of cells that can synthesize and release proteins and carbohydrates [99,100]. It is included in many types of organs, such as colon, breast, and lung. Cancerization of these organs usually causes organizational and structural changes toward their glands [101]. Adenocarcinoma is a common form of carcinoma [102]. Among all the colon cancers, colorectal adenocarcinoma is the most common form that originates in intestinal glands [100]. Most of the pancreatic cancers are adenocarcinomas [100,103]. The quantitatively measured glandular formation is a crucial indicator for the degree of differentiation [104], and architectural features (such as size and shape) are equally important for the prediction of prognosis [105]. Instance segmentation [106] is a prerequisite step to obtain the morphological and statistical features of glands in digital histology images. In this section, we review the application of DL in gland instance segmentation.

Instance segmentation

Gland instance segmentation is a challenging task because of the huge variance of gland morphology [101]. First, the size and shape of glands vary with different sectioning orientations. Glands within the same tissue image can possibly have different sizes when the orientation is different (Fig. 2A and 2B). Second, the density difference between the connective tissue and glands can cause artifacts to the boundaries of glandular tissue (Fig. 2C). Third, section thickness and dye freshness or uniformity will lead to intensity variations (Fig. 2B and 2D). Finally, glands in neoplastic tissue usually show heterogeneous appearances (Fig. 2E and 2F). The glandular structures will degenerate with the increase in the grading of cancer. Separating touching glands, which is a crucial step in instance segmentation, is challenging.

The automatic segmentation of glands in histology images has been explored by many studies [104]. Traditional methods rely on the extraction of hand-crafted features, conventional classifiers, and a large amount of prior knowledge [67,101,112–125]. Although these studies perform well for tissues with many regular glands (usually healthy and benign cases), they yield unsatisfactory results for cancer cases, where the glands show substantial variation and deformation. DL algorithms have enabled accurate and general gland segmentation.

In the *2015 MICCAI Colon Gland Segmentation Challenge (GlaS)* [100], DL methods were introduced in the segmentation and classification tasks of the gland and showed superior performance over traditional methods. Chen *et al.* [59] proposed an FCN-based deep contour-aware network (DCAN). This method combined gland segmentation and contour prediction into a unified multi-task structure. Multiple regularizations can provide

Table 2 Overview of papers using deep learning for gland-level DP image analysis

Reference	Year	Architecture	Methods	Postprocessing
[107]	2017	CNN	Object-Net to predict foreground and Separator-Net to segment individuals	Yes
[108]	2016	CNN	Pixel-level classification	Yes
[59]	2017	FCN	FCN-based semantic segmentation and contour prediction	Yes
[106,109]	2016, 2017	FCN	FCN-based semantic segmentation and HED side convolution channel for contour detection	No
[110]	2016	FCN	Combine two losses into a unified topology-aware loss	No
[111]	2016	CNN	Multitask learning for gland grading and segmentation	No

CNN, convolutional neural network; FCN, fully convolutional network; HED, holistically-nested edge detector. All studies use the Warwick-QU data set of histology colon glands stained by H&E.

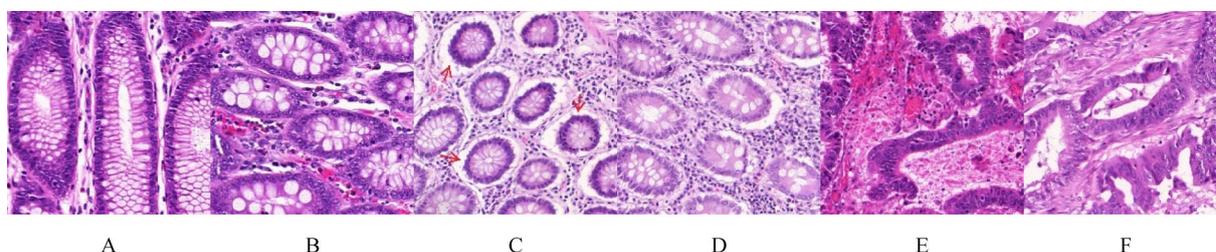


Fig. 2 H&E stained pathology images of colon glands [100, 116]. (A–D) Benign cases. (E and F) Malignant cases. All the slides are under the same magnification and lighting conditions.

complementary information for the two tasks. The final segmentation mask, with individual glands separated from each other, can be generated through this single forward network by fusing the predicted probability map and predicted contour. In the semantic segmentation branch, DCAN combined multilevel contextual features to handle variant gland shapes. This method achieved the best performance in the contest. Xu *et al.* [106,109] conducted gland instance segmentation on this data set. They designed a deep multichannel side supervision system (DMCS), where region and boundary cues were fused with side supervision to solve the instance segmentation issue. The DMCS adopted one channel for semantic segmentation and another channel for contour prediction. The region feature channel followed the FCN structure. The holistically-nested edge detector (HED)-side convolution channel was inspired by HED [126], where five multi-scale outputs were generated from the FCN channel to combine into a final edge map. The boundary information loss during the FCN downsampling can be compensated by adding an additional contour segmentation channel. Different from DCAN, DMCS alleviated the burden of postprocessing by generating the final instance segmentation result through fully convolutional layers. Their method exceeded the previously reported performance on this data set, including the competition champion [59].

DCAN and DMCS aim to preserve object boundaries by

adding an additional branch. BenTaieb *et al.* [110] incorporated topology and smoothness information into the training of FCN by designing a new loss function. In this work, two losses, namely, traditional pixel-level and topology-aware losses, were combined. Hierarchical relations and boundary smoothness constraints were encoded into a unified topology-aware loss function. The label hierarchy should be followed in accordance with the proposed algorithm. Epithelium and lumen regions were labeled as foregrounds, and surrounding stroma was labeled as background. The boundary was smoothed by limiting the neighborhood pixels to output similar probabilities. This end-to-end network did not require postprocessing nor cause additional computational burden during testing. BenTaieb *et al.* [111] creatively proposed another method to solve the gland instance segmentation problem. On the basis of the thought that classification and segmentation tasks can provide compensation information for each other, they designed a symmetrically joint combined classification and segmentation network. Two loss functions were utilized, where the first function penalized the classification of gland grading, and the other was used for pixel-level segmentation. The pretrained classification output can provide spatial information of multiple glands that assists in the discrimination of different objects. In this way, improved performance was achieved for the two tasks.

In addition to the abovementioned FCN-based approaches, some studies have viewed gland segmentation as a pixel-based classification problem [107,108]. One of the participant teams in 2015 GlaS [100], Kainz *et al.* [107], applied two distinct CNNs as pixelwise classifiers. The first network called Object-Net performed four-class classification to distinguish benign background, benign gland, malignant background, and malignant gland. The second network called Separator-Net was trained to predict the separating structures among glands for differentiating touching glands. The postprocessing method based on weighted total variation transformed the combination of two outputs into a final segmentation result. This method successfully segmented glands from the background and classified whether the tissue was benign or malignant. Li *et al.* [108] used Alexnet [126] and Googlenet [127] (with weights pretrained on ImageNet [128]) for window classification rather than using specifically designed CNN structures. Their result outperformed the previous state-of-the-art method that utilized hand-crafted features with support vector machine (SVM) (HC-SVM) [129]. They also validated that the fusion between CNN and HC-SVM will result in additional improvement, indicating that the two contrasting methods are complementary (Table 3).

Region level

The DP analysis is regarded as a region-level task when the target ROI is a larger tissue compared with the cell and gland. Region-level information is an important diagnostic factor. The biological behavior and morphological characteristics of tissue cells and contextual information should be considered. Pathologists need to identify tumor areas in WSIs, determine carcinoma types (e.g., small cell and non-small types for lung cancer), and assess their aggressiveness in the following treatment. Several challenges exist in the automatic analysis of region-level tasks [130]: (1) the complexity of clinical feature representation: histopathology characteristics, such as morphology, scale, texture, and color distribution, can be remarkably heterogeneous for different cancer types, making it difficult to find a general pattern for different cancers; (2) the insufficiency of training images: different from a natural scene image data set that can include millions of images, a pathological image data set usually contains only hundreds of images;

(3) the extremely large size of a single histopathology image: a gigapixel WSI typically possesses a size larger than 100 000 pixels and contains more than 1 million descriptive objects, thereby making effective feature extraction difficult. In this section, we will illustrate the application of DL in region-level structure analysis from two aspects. The first aspect is to classify whether the region is cancerous or distinguish different cancer subtypes. The second aspect is to segment certain structures associated with specific clinical features. Table 4 presents an overview of each task in the region-level analysis.

Classification

Previous studies [141,142] on histopathology image classification mainly focused on manual feature design. Automatic feature learning by CNNs has drawn much attention with the development of DL methods. Directly applying CNNs to region-level analysis is impractical because of the large scale of WSIs and high computational cost. Downsampling a WSI to fit into a neural network is unsuitable because essential details may be lost. Alternatively, most studies [130,132,134] choose to perform analysis locally on small patches that are densely cropped. The final results are acquired on the basis of the overall patch-level predictions. Knowledge transferred from cross-domain is largely utilized to alleviate data insufficiency. Xu *et al.* [130,132] exploited CNN activation features to achieve region-level classification. They first divided each region into a set of overlapping patches. A transfer learning strategy was then explored by pretraining CNN with ImageNet [126]. Each patch was transformed into a 4096-dimensional feature vector. Feature pooling and feature selection were implemented to yield a region-level feature vector for reducing redundancy and selecting a subset of many relevant features. Region-level features were passed to a linear SVM [143] for classification. Their method achieved a state-of-the-art accuracy of 97.5% in the *MICCAI 2014 Brain Tumor Digital Pathology Challenge* [144]. Similarly, Källén *et al.* [134] developed a region-level classification algorithm to report the Gleason score. They used the pretrained OverFeat [145] to extract multilayer features for each patch. Random forest (RF) [146] or SVM was used for patch classification. They applied a voting strategy on all patches for different classes

Table 3 Performance comparison of the three methods based on pure deep learning, hand-crafted features, and their fusion [108], respectively

Method	Jaccard index	Dice index
HC-SVM	0.71 ± 0.11	0.83 ± 0.09
Alexnet + Googlenet	0.74 ± 0.14	0.84 ± 0.10
HC-SVM + Alexnet + Googlenet	0.77 ± 0.11	0.87 ± 0.08

Table 4 Overview of papers using deep learning for region-level DP image analysis

Reference	Year	Architecture	Region/WSI	Organ	Task	Methods
[131]	2014	CNN	Region	Colon	Classification and segmentation	CNN with MIL
[132]	2015	CNN	Region	Brain/colon	Classification and segmentation	CNN with feature pooling and selection
[133]	2016	CNN	WSI	Lung	Classification	EM-based model with CNN
[134]	2016	CNN	Region	Prostate	Classification	Pretrained OverFeat with multilayer feature learning
[135]	2016	CNN	Region	Breast/colon	Classification and segmentation	Supapixel-based scheme for generating patches
[136]	2017	CNN	WSI	Colon	Segmentation	Combine persistent homology features with CNN features for ensemble learning
[137]	2017	FCN	Region	Colon	Segmentation	FCN with constrained deep weak supervision
[138]	2018	CNN	WSI	Lung/lymph nodes	Classification and segmentation	ResNet-50 with MIL
[139]	2019	FCN	WSI	Lung	Classification	FCN with context-aware feature selection
[140]	2019	CNN	Region	Breast	Classification	Pretrained VGG-16 and average pooling of patch-level features for region-level ones

CNN, convolutional neural network; FCN, fully convolutional network; WSI, whole slide image; MIL, multiple instance learning. The data sets in all included studies are stained by H&E.

to classify the whole region image. The class with the highest votes was assigned. However, this approach may miss small lesion areas in a cancerous WSI. An average pooling strategy was exploited for aggregating patch-level features to region-level ones [140]. In particular, they first fine-tuned a pretrained VGG-16 network [147] using fixed-size patches sampled from the lesion spots. The CNN was then used to extract a convolutional feature vector for each patch. The softmax probabilities outputted by the CNN were used as weights of patch-level feature representations. Region-level features were obtained through the average pooling of patch-level ones. Their method may be limited because they only considered ROIs with the most severe diagnosis within a WSI.

The abovementioned supervised patch-level prediction strategies have several inherent drawbacks. First, these methods require WSIs with well-annotated cancer regions. However, pixel wise annotations are seldom available because of the large sizes of WSIs. Comparatively, many WSIs with only image-level ground truth labels can be utilized. Second, patch-level labels do not consistently correspond with the region label [133]. Simple decision fusion approaches (e.g., voting and max pooling) are insufficiently robust to make the right region-level prediction. Weakly supervised methods, especially MIL-based ones [131,133,139,148–151], are extensively investigated to solve the above issues. Xu *et al.* [131] extracted feature representation via deep neural networks and applied the MIL framework for classification. Their MIL performance outweighed that of supervised DL features.

Hou *et al.* [133] proposed an expectation–maximization (EM)-based method that combines patch-based CNN with supervised decision fusion. First, they identified discriminative patches in WSIs by combining an EM-based method with CNN. The histograms of patch-level predictions were then passed to a multiclass logistic regression or SVM to predict the WSI-level labels. They verified their method on two WSI data sets. However, their algorithm was computationally intensive, and its performance improvement was slight. Wang *et al.* [139] first collected discriminative patches using FCN. The spatially contextual information was then utilized for feature selection. A globally holistic region descriptor was generated after aggregating the features from multiple representative instances and fed into a RF for WSI-level classification.

Segmentation

Considering the large sizes and small quantity of tissue sections, the region-level segmentation problem is typically formulated as a patch-level classification task. In the *MICCAI 2014 Brain Tumor Digital Pathology Challenge* [144], Xu *et al.* [130,132] extracted patch features using CNN (pretrained on ImageNet [126]). A linear SVM was then applied to classify the patches as necrosis or nonnecrosis. The method achieved state-of-the-art performance in the challenge. Qaiser *et al.* [136] computed the persistent homology and CNN features of each patch. Two types of features were separately fed into the RF regression model. The output prediction was obtained via a multistage

ensemble strategy. Although their approach was elaborately designed, the improvement was small. A superpixel-based scheme was used to oversegment the region into atomic areas rather than cropping patches using a tiling strategy, indicating that this scheme performed better than fixed-size square window-based approaches [135].

Some studies have applied weakly supervised methods for segmentation, where only region-level labels are available. Courtiol *et al.* [138] adopted a MIL technique for localizing disease areas in WSIs. They first cropped all patches from nonoverlapping grids. The pretrained ResNet-50 [152] architecture was then utilized to extract patch-level feature vectors, followed by feature embedding. Top instances and negative evidence were selected. A multilayer perceptron classifier was used for classification. Several studies have combined FCN and MIL for weakly supervised segmentation. Jia *et al.* [137] developed a new MIL algorithm under a deep weak supervision (DWS) formulation. They utilized the first three stages of the VGGNet and connected side-output layers with DWS under MIL. Area constraints were presented to regularize the sizes of predicted positive instances. Subsequently, the side-outputs were merged to exploit the multi-scale predictions. The method was general and had potential applications for various medical image analysis tasks.

Grading and prognosis

Prognosis provides the prediction of possible disease development, such as the likelihood of disease deterioration or amelioration, chance of survival, and expectations of life quality [153]. However, recent advancements of DL methods in histopathology make DL a potential tool to enhance the accuracy of prognosis [154].

DL methods are primarily applied to extract cancer-related features or structures during prognosis. Non-small cell lung cancer (NSCLC) is typically treated with surgical resection. However, disease recurrence usually occurs for patients with early-stage NSCLC. Validated biomarkers are scarce for the prediction of the recurrence rate [155]. To assist subsequent therapy decisions (whether or not to adopt adjuvant chemotherapy), Wang *et al.* [156] first used CNN and watershed-based method for nuclear segmentation. They then extracted and selected massive features from previously segmented nuclei pixels. Three different classifiers were used to distinguish binary classes that corresponded to recurrence or nonrecurrence. In their experiment, DL showed slightly better performance compared with conventional watershed algorithms. Vaidya *et al.* [157] presented a radiology–pathology fusion approach to predict recurrence-free survival. These methods also used CNN to segment nuclei. The pathological features extracted from the nuclei and intratumoral/peritumoral features from computed tomography scans

were combined, which was an innovative concept. After feature selection using a minimum redundancy maximum relevance algorithm, three classifiers were used for recurrence risk stratification.

Gene status is an important indicator of prognosis. The mutation of specific genes reveals the level of tumor severity and the probability of healing [158–160]. However, modeling the relationship between pathology images and high-dimensional genetic data is challenging [161]. Many studies have focused on the prediction of specific gene mutation. Coudray *et al.* [162] used a modified Inception v3 to predict the mutation of six types of genes (STK11, EGFR, FAT1, SETBP1, KRAS, and TP53). In particular, the network uses an adenocarcinoma (LUAD) tile as input and outputs one of the six gene types.

Some research groups [163] studied the feasibility of substitute DL methods for complicated or expensive prognosis-related clinical examination. Gastrointestinal cancer patients with microsatellite instable (MSI) tumors considerably benefit from immunotherapy. However, the MSI test requires additional immunohistochemical or genetic analysis. Many patients do not receive MSI screening, thereby making potential responders to immune checkpoint inhibitors for missing timely treatment [164]. To address this problem, Kather *et al.* [163] successfully identified MSI from ubiquitously available H&E-stained histology images. This method first trained ResNet-18 [152] to detect tumors. Square tiles were then cut from tumor regions. Another ResNet-18 [152] was used for MSI classification after color normalization. With advanced DL, a broad target population can be excavated without additional laboratory testing.

Cancer grade describes the appearance of tumors and closely correlates with survival. It is an important determinant for prognosis, and has attracted the interest of many researchers. CNN can mine the hidden pattern of pathology characteristics. Therefore, cancer grading is directly conducted on raw WSIs, without intermediate feature extraction (e.g., tissue structures and cell morphology). Nagpal *et al.* [165] designed a DL system that outperforms human pathologists. The WSI was divided into dense patches, where each patch was classified as either nontumor or Gleason pattern (GP) 3/4/5 by InceptionV3 [166]. After calculating the GP quantitation of a slide, Gleason scoring was predicted. Ing *et al.* [167] demonstrated the suitability of semantic segmentation in prostate adenocarcinoma grading. A semantic segmentation network can delineate and classify tumor regions. In their method, foreground tissue areas were roughly located through gray-level thresholding at the beginning and partitioning multiple subtiles. Subsequently, four CNN architectures (FCN-8s [70], SegNet-Full [168], SegNet-Basic [168], and U-net [38]) were trained to divide the areas of stroma, benign epithelium, and Gleason Grade 3/4/5 tumors. Segmentation was conducted in differently

scaled images, which were combined and reassembled to whole slide grading maps. Li *et al.* [169] regarded automatic Gleason grading as an instance segmentation task. In this way, glands with the same grading will be separated into different individuals. They proposed a novel path R-CNN model. In path R-CNN, the ResNet model was used as a backbone to extract and feed the feature maps into two branches. The first branch was a cascade of region proposal network (RPN) and grading network head (GNH). In this classical two-stage branch, RPN worked as a proposal generator, and GNH predicted a binary mask, class, and box offset for each ROI. The second branch was called epithelial network head (ENH), which was a simple binary classification network. ENH can filter the images without tumors, where the network outputted the whole image as background. The network will output the results produced by GNH when ENH predicted the existence of tumors. A fully connected conditional random field model will reduce the artifacts on the edges of stitched patches. Although Path R-CNN adopted an end-to-end structure, the training process was two-stage. GNH and the high layers of the ResNet backbone were trained and fixed in the first place, and ENH was then trained.

Conclusions

The advent of DL has immensely improved the consistency, efficiency, and accuracy in pathology analysis. As a powerful tool, DL provides reliable support for diagnostic assessment and treatment decisions. In comparison with traditional machine learning methods, DL algorithms can uncover unrecognized features to assist prognosis. CNN benefits from transfer learning, indicating that the performance can be improved by pretraining the network on large nonpathology data sets (e.g., ImageNet [128]).

However, DL-based approaches are still in the exploratory stage and have several limitations. First, large high-quality data cohorts labeled by multiple authoritative pathologists are needed. The performance of CNN strongly depends on the quality and quantity of training data. The evaluation of different algorithms relies on the objectivity of the gold standard. However, the annotation is relatively expensive, thereby hindering the acquisition of adequate data. Second, intrinsic difficulties exist in histology annotations, especially for expert pathologists. Despite the decision discrepancy of experts, only one criterion is used for network training and testing, making it unreliable. Third, strong robustness for the clinical application cannot be achieved. The result will probably become inferior when a proposed algorithm in one research is used in another type of target (WSIs with different image magnification, dyeing situation, cancer type, etc.). Many reliable paradigms are required. Finally, DL methods are still “black boxes” and inside patterns cannot be

comprehended. Low interpretability creates a semantic gap, making these methods undependable in clinical practice. Apparently, many progresses need to be achieved in this field.

Despite all these challenges, DL has certainly driven a revolution on DP diagnosis. In the future, effective cancer prognosis, image registration, and biological target prediction on pathology images through DL methods will be valuable and are challenging research directions. Solving medical problems through unsupervised or weakly supervised learning, such as one-shot learning, remains to be explored. We envision that DL will serve as a decision support tool for human pathologists and immensely alleviate their workloads via high-throughput analysis. The combination of human and artificial intelligence shows a bright prospect.

Acknowledgements

This work was supported by the National Science and Technology Major Project of the Ministry of Science and Technology of China (No. 2017YFC0110903), Microsoft Research under the eHealth program, the National Natural Science Foundation of China (No. 81771910), the Beijing Natural Science Foundation in China (No. 4152033), the Technology and Innovation Commission of Shenzhen in China (No. shenfagai2016-627), the Beijing Young Talent Project in China, the Fundamental Research Funds for the Central Universities of China (No. SKLSDE-2017ZX-08) from the State Key Laboratory of Software Development Environment in Beihang University in China, and the 111 Project in China (No. B13003).

Compliance with ethics guidelines

Shujian Deng, Xin Zhang, Wen Yan, Eric I-Chao Chang, Yubo Fan, Maode Lai, and Yan Xu declare that they have no conflicts of interest. This manuscript is a review article and does not involve a research protocol requiring approval by the relevant institutional review board or ethics committee.

References

1. Cardesa A, Zidar N, Alos L, Nadal A, Gale N, Klöppel G. The Kaiser's cancer revisited: was Virchow totally wrong? *Virchows Arch* 2011; 458(6): 649–657
2. Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. Histopathological image analysis: a review. *IEEE Rev Biomed Eng* 2009; 2: 147–171
3. Andron A, Magnani C, Betta PG, Donna A, Mollo F, Scelsi M, Bernardi P, Botta M, Terracini B. Malignant mesothelioma of the pleura: interobserver variability. *J Clin Pathol* 1995; 48(9): 856–860
4. Wolberg WH, Street WN, Heisey DM, Mangasarian OL. Computer-derived nuclear features distinguish malignant from benign breast cytology. *Hum Pathol* 1995; 26(7): 792–796

5. Thiran JP, Macq B. Morphological feature extraction for the classification of digital images of cancerous tissues. *IEEE Trans Biomed Eng* 1996; 43(10): 1011–1020
6. Choi HK, Jarkrans T, Bengtsson E, Vasko J, Wester K, Malmström PU, Busch C. Image analysis based grading of bladder carcinoma. Comparison of object, texture and graph based methods and their reproducibility. *Anal Cell Pathol* 1997; 15(1): 1–18
7. Hamilton PW, Bartels PH, Thompson D, Anderson NH, Montironi R, Sloan JM. Automated location of dysplastic fields in colorectal histology using image texture analysis. *J Pathol* 1997; 182(1): 68–75
8. Esgiar AN, Naguib RN, Sharif BS, Bennett MK, Murray A. Fractal analysis in the detection of colonic cancer images. *IEEE Trans Inf Technol Biomed* 2002; 6(1): 54–58
9. Spyridonos P, Ravazoula P, Cavouras D, Berberidis K, Nikiforidis G. Computer-based grading of haematoxylin-eosin stained tissue sections of urinary bladder carcinomas. *Med Inform Internet Med* 2001; 26(3): 179–190
10. Wiltgen M, Gerger A, Smolle J. Tissue counter analysis of benign common nevi and malignant melanoma. *Int J Med Inform* 2003; 69(1): 17–28
11. Nielsen B, Albregtsen F, Danielsen HE. The use of fractal features from the periphery of cell nuclei as a classification tool. *Anal Cell Pathol* 1999; 19(1): 21–37
12. Esgiar AN, Naguib RN, Sharif BS, Bennett MK, Murray A. Microscopic image analysis for quantitative measurement and feature identification of normal and cancerous colonic mucosa. *IEEE Trans Inf Technol Biomed* 1998; 2(3): 197–203
13. Weyn B, van de Wouwer G, Kumar-Singh S, van Daele A, Scheunders P, van Marck E, Jacob W. Computer-assisted differential diagnosis of malignant mesothelioma based on syntactic structure analysis. *Cytometry* 1999; 35(1): 23–29
14. Keenan SJ, Diamond J, McCluggage WG, Bharucha H, Thompson D, Bartels PH, Hamilton PW. An automated machine vision system for the histological grading of cervical intraepithelial neoplasia (CIN). *J Pathol* 2000; 192(3): 351–362
15. Demir C, Gultekin SH, Yener B. Learning the topological properties of brain tumors. *IEEE/ACM Trans Comput Biol Bioinformatics* 2005; 2(3): 262–270
16. Gunduz-Demir C. Mathematical modeling of the malignancy of cancer using graph evolution. *Math Biosci* 2007; 209(2): 514–527
17. Weinstein RS, Graham AR, Richter LC, Barker GP, Krupinski EA, Lopez AM, Erps KA, Bhattacharyya AK, Yagi Y, Gilbertson JR. Overview of telepathology, virtual microscopy, and whole slide imaging: prospects for the future. *Hum Pathol* 2009; 40(8): 1057–1069
18. Krupinski EA, Bhattacharyya AK, Weinstein RS. *Telepathology and Digital Pathology Research*. Springer, Cham. 2016: 41–54
19. Farhani N, Parwani AV, Pantanowitz L. Whole slide imaging in pathology: advantages, limitations, and emerging perspectives. *Pathol Lab Med Int* 2015; 7: 23–33
20. Ying X, Monticello TM. Modern imaging technologies in toxicologic pathology: an overview. *Toxicol Pathol* 2006; 34(7): 815–826
21. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015; 521(7553): 436–444
22. LeCun Y, Boser BE, Denker JS, Henderson D, Howard RE, Hubbard WE, Jackel LD. Handwritten digit recognition with a back-propagation network. In: *Advances in Neural Information Processing Systems*. 1990: 396–404
23. LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proc IEEE* 1998; 86(11): 2278–2324
24. Wang P, Ge R, Xiao X, Cai Y, Wang G, Zhou F. Rectified-linear-unit-based deep learning for biomedical multi-label data. *Interdiscip Sci* 2017; 9(3): 419–422
25. Madabhushi A, Lee G. Image analysis and machine learning in digital pathology: challenges and opportunities. *Med Image Anal* 2016; 33: 170–175
26. Epstein JI. An update of the Gleason grading system. *J Urol* 2010; 183(2): 433–440
27. Frierson HF Jr, Wolber RA, Berean KW, Franquemont DW, Gaffey MJ, Boyd JC, Wilbur DC. Interobserver reproducibility of the Nottingham modification of the Bloom and Richardson histologic grading scheme for infiltrating ductal carcinoma. *Am J Clin Pathol* 1995; 103(2): 195–198
28. Khan AM, Rajpoot N, Treanor D, Magee D. A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution. *IEEE Trans Biomed Eng* 2014; 61(6): 1729–1738
29. Vahadane A, Peng T, Sethi A, Albarqouni S, Wang L, Baust M, Steiger K, Schlitter AM, Esposito I, Navab N. Structure-preserving color normalization and sparse stain separation for histological images. *IEEE Trans Med Imaging* 2016; 35(8): 1962–1971
30. Janowczyk A, Basavanthally A, Madabhushi A. Stain normalization using Sparse AutoEncoders (StaNoSA): application to digital pathology. *Comput Med Imaging Graph* 2017; 57: 50–61
31. Bentaieb A, Hamarneh G. Adversarial stain transfer for histopathology image analysis. *IEEE Trans Med Imaging* 2018; 37(3): 792–802
32. Roy S, Kumar Jain A, Lal S, Kini J. A study about color normalization methods for histopathology images. *Micron* 2018; 114: 42–61
33. Elston CW, Ellis IO. Author Commentary: “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. C. W. Elston & I. O. Ellis. *Histopathology* 1991; 19: 403–410.” *Histopathology* 2002; 41(3A): 151
34. Chow KH, Factor RE, Ullman KS. The nuclear envelope environment and its cancer connections. *Nat Rev Cancer* 2012; 12(3): 196–209
35. Dey P. Cancer nucleus: morphology and beyond. *Diagn Cytopathol* 2010; 38(5): 382–390
36. Chen H, Wang X, Heng PA. Automated mitosis detection with deep regression networks. 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). IEEE. 2016: 1204–1207
37. Cireşan D, Giusti A, Gambardella LM, Schmidhuber J. Deep neural networks segment neuronal membranes in electron microscopy images. *Adv Neural Inf Process Syst* 2012: 2843–2851
38. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham. 2015: 234–241

39. Zhou Y, Chang H, Barner KE, Parvin B. Nuclei segmentation via sparsity constrained convolutional regression. *Proc IEEE Int Symp Biomed Imaging* 2015; 2015: 1284–1287
40. Song Y, Zhang L, Chen S, Ni D, Lei B, Wang T. Accurate segmentation of cervical cytoplasm and nuclei based on multiscale convolutional network and graph partitioning. *IEEE Trans Biomed Eng* 2015; 62(10): 2421–2433
41. Xing F, Xie Y, Yang L. An automatic learning-based framework for robust nucleus segmentation. *IEEE Trans Med Imaging* 2016; 35(2): 550–566
42. Pan X, Li L, Yang H, Liu Z, Yang J, Zhao L, Fan Y. Accurate segmentation of nuclei in pathological images via sparse reconstruction and deep convolutional networks. *Neurocomputing* 2016; 229: S0925231216313765
43. Zhang L, Sonka M, Lu L, Summers RM, Yao J. Combining fully convolutional networks and graph-based approach for automated segmentation of cervical cell nuclei. 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017). IEEE. 2017: 406–409
44. Alom MZ, Yakopcic C, Taha TM, Asari VK. Nuclei segmentation with recurrent residual convolutional neural networks based U-Net (R2U-Net). *NAECON 2018-IEEE National Aerospace and Electronics Conference*. IEEE. 2018: 228–233
45. Xie Y, Xing F, Kong X, Su H, Yang L. Beyond classification: structured regression for robust cell detection using convolutional neural network. *Med Image Comput Comput Assist Interv* 2015; 9351: 358–365
46. Sirinukunwattana K, Ahmed Raza SE, Tsang YW, Snead DRJ, Cree IA, Rajpoot NM. Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. *IEEE Trans Med Imaging* 2016; 35(5): 1196–1206
47. Xu J, Xiang L, Liu Q, Gilmore H, Wu J, Tang J, Madabhushi A. Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images. *IEEE Trans Med Imaging* 2016; 35(1): 119–130
48. Xie Y, Xing F, Shi X, Kong X, Su H, Yang L. Efficient and robust cell detection: a structured regression approach. *Med Image Anal* 2018; 44: 245–254
49. Cireşan DC, Giusti A, Gambardella LM, Schmidhuber J. Mitosis detection in breast cancer histology images with deep neural networks. *Med Image Comput Comput Assist Interv* 2013; 16(Pt 2): 411–418
50. Wang H, Cruz-Roa A, Basavanahally A, Gilmore H, Shih N, Feldman M, Tomaszewski J, Gonzalez F, Madabhushi A. Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features. *J Med Imaging (Bellingham)* 2014; 1(3): 034003
51. Chen H, Dou Q, Wang X, Qin J, Heng PA. Mitosis detection in breast cancer histology images via deep cascaded networks. *Thirtieth AAAI Conference on Artificial Intelligence*. 2016
52. Albarqouni S, Baur C, Achilles F, Belagiannis V, Demirci S, Navab N. Aggnet: deep learning from crowds for mitosis detection in breast cancer histology images. *IEEE Trans Med Imaging* 2016; 35(5): 1313–1321
53. Li C, Wang X, Liu W, Latecki LJ. DeepMitosis: mitosis detection via deep detection, verification and segmentation networks. *Med Image Anal* 2018; 45: 121–133
54. Ma M, Shi Y, Li W, Gao Y, Xu J. A novel two-stage deep method for mitosis detection in breast cancer histology images. 2018 24th International Conference on Pattern Recognition (ICPR). IEEE. 2018: 3892–3897
55. Li C, Wang X, Liu W, Latecki LJ, Wang B, Huang J. Weakly supervised mitosis detection in breast histopathology images using concentric loss. *Med Image Anal* 2019; 53: 165–178
56. Yang L, Zhang Y, Guldner IH, Zhang S, Chen DZ. 3D segmentation of glial cells using fully convolutional networks and k-terminal cut. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham. 2016: 658–666
57. Akram SU, Kannala J, Eklund L, Heikkilä J. Cell proposal network for microscopy image analysis. 2016 IEEE International Conference on Image Processing (ICIP). IEEE. 2016: 3199–3203
58. Akram SU, Kannala J, Eklund L, Heikkilä J. Cell segmentation proposal network for microscopy image analysis. *Deep Learning and Data Labeling for Medical Applications*. Springer, Cham. 2016: 21–29
59. Chen H, Qi X, Yu L, Dou Q, Qin J, Heng PA. DCAN: deep contour-aware networks for object instance segmentation from histology images. *Med Image Anal* 2017; 36: 135–146
60. Song Y, Tan EL, Jiang X, Cheng JZ, Ni D, Chen S, Lei B, Wang T. Accurate cervical cell segmentation from overlapping clumps in pap smear images. *IEEE Trans Med Imaging* 2017; 36(1): 288–300
61. Kumar N, Verma R, Sharma S, Bhargava S, Vahadane A, Sethi A. A dataset and a technique for generalized nuclear segmentation for computational pathology. *IEEE Trans Med Imaging* 2017; 36(7): 1550–1560
62. Ho DJ, Fu C, Salama P, Dunn KW, Delp EJ. Nuclei detection and segmentation of fluorescence microscopy images using three dimensional convolutional neural networks. 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). IEEE. 2018: 418–422
63. Naylor P, Laé M, Reyat F, Walter T. Segmentation of nuclei in histopathology images by deep regression of the distance map. *IEEE Trans Med Imaging* 2019; 38(2): 448–459
64. Zhou Y, Onder OF, Dou Q, Tsougenis E, Chen H, Heng PA. Cianet: Robust nuclei instance segmentation with contour-aware information aggregation. *International Conference on Information Processing in Medical Imaging*. Springer, Cham. 2019: 682–693
65. Arganda-Carreras I, Seung HS, Cardona A, Schindelin J. Segmentation of neuronal structures in EM stacks challenge–ISBI 2012. 2012
66. Oren A, Fernandes J. The Bethesda system for the reporting of cervical/vaginal cytology. *J Am Osteopath Assoc* 1991; 91(5): 476–479
67. Naik S, Doyle S, Feldman M, Tomaszewski J, Madabhushi A. Gland segmentation and computerized gleason grading of prostate histology by integrating low-, high-level and domain specific information. *MIAAB workshop*. 2007: 1–8
68. Karvelis PS, Fotiadis DI, Georgiou I, Syrrou M. A watershed based segmentation method for multispectral chromosome images classification. *Conf Proc IEEE Eng Med Biol Soc* 2006; 2006: 3009–3012
69. Petushi S, Garcia FU, Haber MM, Katsinis C, Tozeren A. Large-

- scale computations on histology images reveal grade-differentiating parameters for breast cancer. *BMC Med Imaging* 2006; 6(1): 14
70. Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Trans Pattern Anal Mach Intell* 2017; 39(4): 640–651
 71. Oktay O, Schlemper J, Folgoc LL, Lee M, Heinrich M, Misawa K, Mori K, McDonagh S, Hammerla NY, Kainz B, Glocker B, Rueckert D. Attention U-net: learning where to look for the pancreas. *arXiv* 2018; 1804.03999
 72. Zeng Z, Xie W, Zhang Y, Lu Y. RIC-Unet: an improved neural network based on Unet for nuclei segmentation in histology images. *IEEE Access* 2019; 7: 21420–21428
 73. Chen JM, Li Y, Xu J, Gong L, Wang LW, Liu WL, Liu J. Computer-aided prognosis on breast cancer with hematoxylin and eosin histopathology images: a review. *Tumour Biol* 2017; 39(3): 1010428317694550
 74. Veta M, Pluim JP, van Diest PJ, Viergever MA. Breast cancer histopathology image analysis: a review. *IEEE Trans Biomed Eng* 2014; 61(5): 1400–1411
 75. Sommer C, Fiaschi L, Hamprecht FA, Gerlich DW. Learning-based mitotic cell detection in histopathological images. *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*. IEEE. 2012: 2306–2309
 76. Veta M, van Diest PJ, Pluim JPW. Detecting mitotic figures in breast cancer histopathology images. *Medical Imaging 2013: Digital Pathology*. International Society for Optics and Photonics. 2013; 8676: 867607
 77. Khan AM, Eldaly H, Rajpoot NM. A gamma-Gaussian mixture model for detection of mitotic cells in breast cancer histopathology images. *J Pathol Inform* 2013; 4(4): 149–152
 78. Paul A, Dey A, Mukherjee DP, Sivaswamy J, Tourani V. Regenerative random forest with automatic feature selection to detect mitosis in histopathological breast cancer images. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham. 2015: 94–102
 79. Cruz-Roa AA, Arevalo Ovalle JE, Madabhushi A, González Osorio FA. A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection. *Med Image Comput Comput Assist Interv* 2013; 16(Pt 2): 403–410
 80. Veta M, van Diest PJ, Willems SM, Wang H, Madabhushi A, Cruz-Roa A, Gonzalez F, Larsen AB, Vestergaard JS, Dahl AB, Cireşan DC, Schmidhuber J, Giusti A, Gambardella LM, Tek FB, Walter T, Wang CW, Kondo S, Matuszewski BJ, Precioso F, Snell V, Kittler J, de Campos TE, Khan AM, Rajpoot NM, Arkoumani E, Lacle MM, Viergever MA, Pluim JP. Assessment of algorithms for mitosis detection in breast cancer histopathology images. *Med Image Anal* 2015; 20(1): 237–248
 81. Roux L, Racoceanu D, Loménie N, Kulikova M, Irshad H, Klossa J, Capron F, Genestie C, Naour GL, Gurcan MN. Mitosis detection in breast cancer histological images. *An ICPR 2012 contest*. *J Pathol Inform* 2013; 4: 8
 82. Yang F, Mackey MA, Ianzini F, Gallardo G, Sonka M. Cell segmentation, tracking, and mitosis detection using temporal context. *Med Image Comput Comput Assist Interv* 2005; 8(Pt 1): 302–309
 83. Payer C, Štern D, Feiner M, Bischof H, Urschler M. Segmenting and tracking cell instances with cosine embeddings and recurrent hourglass networks. *Med Image Anal* 2019; 57: 106–119
 84. Hariharan B, Arbeláez P, Girshick R, Malik J. Simultaneous detection and segmentation. *European Conference on Computer Vision*. Springer, Cham. 2014: 297–312
 85. Veta M, Van Diest PJ, Pluim JPW. Cutting out the middleman: measuring nuclear area in histopathology slides without segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham. 2016: 632–639
 86. Kronqvist P, Kuopio T, Collan Y. Morphometric grading of invasive ductal breast cancer. I. Thresholds for nuclear grade. *Br J Cancer* 1998; 78(6): 800–805
 87. Mommers EC, Page DL, Dupont WD, Schuyler P, Leonhart AM, Baak JP, Meijer CJ, van Diest PJ. Prognostic value of morphometry in patients with normal breast tissue or usual ductal hyperplasia of the breast. *Int J Cancer* 2001; 95(5): 282–285
 88. Veta M, Kornegoor R, Huisman A, Verschuur-Maes AH, Viergever MA, Pluim JP, van Diest PJ. Prognostic value of automatically extracted nuclear morphometric features in whole slide images of male breast cancer. *Mod Pathol* 2012; 25(12): 1559–1565
 89. Maška M, Ulman V, Svoboda D, Matula P, Matula P, Ederra C, Urbiola A, España T, Venkatesan S, Balak DMW, Karas P, Bolcková T, Štreitová M, Carthel C, Coraluppi S, Harder N, Rohr K, Magnusson KEG, Jaldén J, Blau HM, Dzyubachyk O, Křížek P, Hagen GM, Escuredo DP, Carretero DJ, Carbayo MJL, Barrutia AM, Meijering E, Kozubek M, Solorzano CO. A benchmark for comparison of cell tracking algorithms. *Bioinformatics* 2014; 30(11): 1609–1617
 90. Vu QD, Graham S, To NN, Minh, Muhammad S, Talha Q, Navid AK, Ali KS, Tahsin K, Keyvan F, Tianhao Z, Rajarsi G, Tae KJ, Nasir R, Joel S. Methods for segmentation and classification of digital microscopy tissue images. *Front Bioeng Biotechnol* 2019; 7: 53
 91. Naylor P, Laé M, Reyal F, *et al*. Nuclei segmentation in histopathology images using deep neural networks[C]//2017 IEEE 14th International Symposium On Biomedical Imaging (ISBI 2017). IEEE. 2017: 933–936
 92. Kumar N, Verma R, Anand D, Zhou Y, Onder OF, Tsougenis E, Chen H, Heng P, Li J, Hu Z, Wang Y, Koohbanani NA, Jahanifar M, Tajeddin NZ, Gooya A, Rajpoot N, Ren X, Zhou S, Wang Q, Shen D, Yang C, Weng C, Yu W, Yeh C, Yang S, Xu S, Yeung PH, Sun P, Mahbod A, Schaefer G, Ellinger I, Ecker R, Smedby O, Wang C, Chidester B, Ton T, Tran M, Ma J, Do MN, Graham S, Vu QD, Kwak JT, Gunda A, Chunduri R, Hu C, Zhou X, Lotfi D, Safdari R, Kascenas A, O'Neil A, Eschweiler D, Stegmaier J, Cui Y, Yin B, Chen K, Tian X, Gruening P, Barth E, Arbel E, Remer L, Ben-Dor A, Sirazitdinova E, Kohl M, Braunewell S, Li Y, Xie X, Shen L, Ma J, Baksi KD, Khan MA, Choo J, Colomer A, Naranjo V, Pei L, Iftekharuddin KM, Roy K, Bhattacharjee D, Pedraza A, Bueno MG, Devanathan S, Radhakrishnan S, Koduganty P, Wu Z, Cai G, Liu X, Wang Y, Sethi A. A multi-organ nucleus segmentation challenge. *IEEE Trans Med Imaging* 2020; 39(5): 1380–1391
 93. Neubeck A, Van Gool L. Efficient non-maximum suppression. *18th International Conference on Pattern Recognition (ICPR'06)*. IEEE. 2006, 3: 850–855

94. Maurer CR, Qi R, Raghavan V. A linear time algorithm for computing exact euclidean distance transforms of binary images in arbitrary dimensions. *IEEE Trans Pattern Anal Mach Intell* 2003; 25(2): 265–270
95. Ren S, He K, Girshick R, Sun J. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans Pattern Anal Mach Intell* 2017; 39(6): 1137–1149
96. Zhang R, Cheng C, Zhao X, Li X. Multiscale mask R-CNN-based lung tumor detection using PET imaging. *Mol Imaging* 2019; 18: 1536012119863531
97. Wang S, Rong R, Yang DM, Fujimoto J, Yan S, Cai L, Yang L, Luo D, Behrens C, Parra ER, Yao B, Xu L, Wang T, Zhan X, Wistuba II, Minna J, Xie Y, Xiao G. Computational staining of pathology images to study the tumor microenvironment in lung cancer. *Cancer Res* 2020; 80(10): 2056–2066
98. Johnson JW. Adapting mask-RCNN for automatic nucleus segmentation. *arXiv* 2018; 1805.00500
99. Hoda SA, Hoda RS. Rubin's pathology: clinicopathologic foundations of medicine. *JAMA* 2007; 298(17): 2070–2075
100. Sirinukunwattana K, Pluim JPW, Chen H, Qi X, Heng PA, Guo YB, Wang LY, Matuszewski BJ, Bruni E, Sanchez U, Böhm A, Ronneberger O, Cheikh BB, Racoceanu D, Kainz P, Pfeiffer M, Urschler M, Snead DRJ, Rajpoot NM. Gland segmentation in colon histology images: the GlaS challenge contest. *Med Image Anal* 2017; 35: 489–502
101. Gunduz-Demir C, Kandemir M, Tosun AB, Sokmensuer C. Automatic segmentation of colon glands using object-graphs. *Med Image Anal* 2010; 14(1): 1–12
102. Hess KR, Varadhachary GR, Taylor SH, Wei W, Raber MN, Lenzi R, Abbruzzese JL. Metastatic patterns in adenocarcinoma. *Cancer* 2006; 106(7): 1624–1633
103. Ryan DP, Hong TS, Bardeesy N. Pancreatic adenocarcinoma. *N Engl J Med* 2014; 371(11): 1039–1049
104. Fleming M, Ravula S, Tatishchev SF, Wang HL. Colorectal carcinoma: pathologic aspects. *J Gastrointest Oncol* 2012; 3(3): 153–173
105. Travis WD, Brambilla E, Geisinger KR. Histological grading in lung cancer: one system for all or separate systems for each histological type? *Eur Respir J* 2016; 47(3): 720–723
106. Xu Y, Li Y, Liu M, Wang Y, Lai M, Eric C. Gland instance segmentation by deep multichannel side supervision. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham. 2016: 496–504
107. Kainz P, Pfeiffer M, Urschler M. Segmentation and classification of colon glands with deep convolutional neural networks and total variation regularization. *PeerJ* 2017; 5: e3874
108. Li W, Manivannan S, Akbar S, Zhang J, Trucco E, McKenna SJ. Gland segmentation in colon histology images using hand-crafted features and convolutional neural networks. 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). IEEE. 2016: 1405–1408
109. Xu Y, Li Y, Wang Y, Liu M, Fan Y, Lai M, Chang EIC. Gland instance segmentation using deep multichannel neural networks. *IEEE Trans Biomed Eng* 2017; 64(12): 2901–2912
110. BenTaieb A, Hamameh G. Topology aware fully convolutional networks for histology gland segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham. 2016: 460–468
111. BenTaieb A, Kawahara J, Hamameh G. Multi-loss convolutional networks for gland analysis in microscopy. 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). IEEE. 2016: 642–645
112. Wu HS, Xu R, Harpaz N, Burstein D, Gil J. Segmentation of microscopic images of small intestinal glands with directional 2-D filters. *Anal Quant Cytol Histol* 2005; 27(5): 291–300
113. Farjam R, Soltanian-Zadeh H, Jafari-Khouzani K, Zoroofi RA. An image analysis approach for automatic malignancy determination of prostate pathological images. *Cytometry B Clin Cytom* 2007; 72(4): 227–240
114. Wu HS, Xu R, Harpaz N, Burstein D, Gil J. Segmentation of intestinal gland images with iterative region growing. *J Microsc* 2005; 220(Pt 3): 190–204
115. Fu H, Qiu G, Shu J, Ilyas M. A novel polar space random field model for the detection of glandular structures. *IEEE Trans Med Imaging* 2014; 33(3): 764–776
116. Sirinukunwattana K, Snead DR, Rajpoot NM. A stochastic polygons model for glandular structures in colon histology images. *IEEE Trans Med Imaging* 2015; 34(11): 2366–2378
117. Monaco JP, Tomaszewski JE, Feldman MD, Hagemann I, Moradi M, Mousavi P, Boag A, Davidson C, Abolmaesumi P, Madabhushi A. High-throughput detection of prostate cancer in histological sections using probabilistic pairwise Markov models. *Med Image Anal* 2010; 14(4): 617–629
118. Diamond J, Anderson NH, Bartels PH, Montironi R, Hamilton PW. The use of morphological characteristics and texture analysis in the identification of tissue composition in prostatic neoplasia. *Hum Pathol* 2004; 35(9): 1121–1131
119. Doyle S, Madabhushi A, Feldman M, Tomaszewski J. A boosting cascade for automated detection of prostate cancer from digitized histology. *Med Image Comput Comput Assist Interv* 2006; 9(Pt 2): 504–511
120. Tabesh A, Teverovskiy M, Pang HY, Kumar VP, Verbel D, Kotsianti A, Saidi O. Multifeature prostate cancer diagnosis and Gleason grading of histological images. *IEEE Trans Med Imaging* 2007; 26(10): 1366–1378
121. Nguyen K, Sarkar A, Jain AK. Structure and context in prostatic gland segmentation and classification. *Med Image Comput Comput Assist Interv* 2012; 15(Pt 1): 115–123
122. Jacobs JG, Panagiotaki E, Alexander DC. Gleason grading of prostate tumours with max-margin conditional random fields. *International Workshop on Machine Learning in Medical Imaging*. Springer, Cham. 2014: 85–92
123. Sabata B, Babenko B, Monroe R, Srinivas C. Automated analysis of pin-4 stained prostate needle biopsies. *International Workshop on Prostate Cancer Imaging*. Springer, Berlin, Heidelberg. 2010: 89–100
124. Altunbay D, Cigir C, Sokmensuer C, Gunduz-Demir C. Color graphs for automated cancer diagnosis and grading. *IEEE Trans Biomed Eng* 2010; 57(3): 665–674
125. Fakhzadeh A, Sporndly-Nees E, Holm L, Hendriks CLL. Analyzing tubular tissue in histopathological thin sections. 2012 International Conference on Digital Image Computing Techniques and Applications (DICTA). IEEE. 2012: 1–6
126. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification

- with deep convolutional neural networks. *Adv Neural Inf Process Syst* 2012; 1097–1105
127. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. *Proceedings of the IEEE Conference on computer Vision and Pattern Recognition*. 2015: 1–9
 128. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. Imagenet: a large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2009: 248–255
 129. Manivannan S, Li W, Akbar S, Wang R, Zhang J, McKenna SJ. An automated pattern recognition system for classifying indirect immunofluorescence images of hep-2 cells and specimens. *Pattern Recognit* 2016; 51: 12–26
 130. Xu Y, Jia Z, Wang LB, Ai Y, Zhang F, Lai M, Chang EIC. Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. *BMC Bioinformatics* 2017; 18(1): 281
 131. Xu Y, Mo T, Feng Q, Zhong P, Lai M, Chang EI. Deep learning of feature representation with multiple instance learning for medical image analysis. *IEEE International Conference on Acoustics, Speech and Signal Processing* 2014; 1626–1630
 132. Xu Y, Jia Z, Ai Y, Zhang F, Lai M, Change EI. Deep convolutional activation features for large scale brain tumor histopathology image classification and segmentation. *IEEE International Conference on Acoustics, Speech and Signal Processing*. 2015: 947–951
 133. Hou L, Samaras D, Kurc TM, Gao Y, Davis JE, Saltz JH. Patch-based convolutional neural network for whole slide tissue image classification. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit* 2016; 2016: 2424–2433
 134. Källén H, Molin J, Heyden A, Lundstrom C, Astrom K. Towards grading gleason score using generically trained deep convolutional neural networks. *IEEE International Symposium on Biomedical Imaging*. 2016: 1163–1167
 135. Xu J, Luo X, Wang G, Gilmore H, Madabhushi A. A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images. *Neurocomputing* 2016; 191: 214–223
 136. Qaiser T, Tsang YW, Epstein D, Rajpoot N. Tumor segmentation in whole slide images using persistent homology and deep convolutional features. *Annual Conference on Medical Image Understanding and Analysis*. 2017; 723: 320–329
 137. Jia Z, Huang X, Chang EI, Xu Y. Constrained deep weak supervision for histopathology image segmentation. *IEEE Trans Med Imaging* 2017; 36(11): 2376–2388
 138. Courtiol P, Tramel EW, Sanselme M, Wainrib G. Classification and disease localization in histopathology using only global labels: a weakly-supervised approach. 2018
 139. Wang X, Chen H, Gan C, Lin H, Dou Q, Tsougenis E, Huang Q, Cai M, Heng PA. Weakly supervised deep learning for whole slide lung cancer image analysis. *IEEE Trans Cybern* 2019; [Epub ahead of print] doi: 10.1109/TCYB.2019.2935141
 140. Mercan C, Aksoy S, Mercan E, Shapiro LG, Weaver DL, Elmore JG. From patch-level to roi-level deep feature representations for breast histopathology classification. *Medical Imaging 2019: Digital Pathology*. 2019. 109560H
 141. Xu Y, Jiao L, Wang S, Wei J, Fan Y, Lai M, Chang EI. Multi-label classification for colon cancer using histopathological images. *Microsc Res Tech* 2013; 76(12): 1266–1277
 142. Jiao L, Qi C, Li S, Yan X. Colon cancer detection using whole slide histopathological images. *IFMBE Proc* 2013; 39: 1283–1286
 143. Adankon MM, Cheriet M. Support vector machine. *Comput Sci* 2002; 1(4): 1–28
 144. Golland P, Hata N, Barillot C, Hornegger J, Howe R. Preface. *Medical image computing and computer-assisted intervention—MICCAI 2014*. *Med Image Comput Comput Assist Interv* 2014; 17(Pt 1): V–VI
 145. Sermanet P, Eigen D, Zhang X, Mathieu M, Fergus R, Lecun Y. Overfeat: integrated recognition, localization and detection using convolutional networks. *International Conference on Learning Representations*. 2014
 146. Liaw A, Wiener M. Classification and regression by random forest. *R News* 2007; 2: 18–22
 147. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv* 2014; 1409.1556v6
 148. Yan X, Zhu JY, Chang EI, Tu Z. Multiple clustered instance learning for histopathology cancer image classification, segmentation and clustering. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit* 2012; 964–971
 149. Xu Y, Zhang J, Chang EI, Lai M, Tu Z. Context-constrained multiple instance learning for histopathology image segmentation. *Med Image Comput Comput Assist Interv* 2012; 15(Pt 3): 623–630
 150. Xu Y, Zhu JY, Chang EI, Lai M, Tu Z. Weakly supervised histopathology cancer image segmentation and classification. *Med Image Anal* 2014; 18(3): 591–604
 151. Xu Y, Li Y, Shen Z, Wu Z, Gao T, Fan Y, Lai M, Chang EI. Parallel multiple instance learning for extremely large histopathology image analysis. *BMC Bioinformatics* 2017; 18(1): 360
 152. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*. 2016: 770–778
 153. Hagerty RG, Butow PN, Ellis PM, Dimitry S, Tattersall MH. Communicating prognosis in cancer care: a systematic review of the literature. *Ann Oncol* 2005; 16(7): 1005–1053
 154. Norgeot B, Glicksberg BS, Butte AJ. A call for deep-learning healthcare. *Nat Med* 2019; 25(1): 14–15
 155. Uramoto H, Tanaka F. Recurrence after surgery in patients with NSCLC. *Transl Lung Cancer Res* 2014; 3(4): 242–249
 156. Wang X, Janowczyk A, Zhou Y, Thawani R, Fu P, Schalper K, Velcheti V, Madabhushi A. Prediction of recurrence in early stage non-small cell lung cancer using computer extracted nuclear features from digital H&E images. *Sci Rep* 2017; 7(1): 13543
 157. Vaidya P, Wang X, Bera K, Khunger A, Choi H, Patil P, Velcheti V, Madabhushi A. Raptomics: integrating radiomic and pathomic features for predicting recurrence in early stage lung cancer. *Medical Imaging 2018: Digital Pathology International Society for Optics and Photonics*. 2018: 105810M
 158. Sanchez-Cespedes M, Parrella P, Esteller M, Nomoto S, Trink B, Engles JM, Westra WH, Herman JG, Sidransky D. Inactivation of LKB1/STK11 is a common event in adenocarcinomas of the lung. *Cancer Res* 2002; 62(13): 3659–3662
 159. Shackelford DB, Abt E, Gerken L, Vasquez DS, Seki A, Leblanc M, Wei L, Fishbein MC, Czernin J, Mischel PS, Shaw RJ. LKB1 inactivation dictates therapeutic response of non-small cell lung

- cancer to the metabolism drug phenformin. *Cancer Cell* 2013; 23(2): 143–158
160. Parums DV. Current status of targeted therapy in non-small cell lung cancer. *Drugs Today (Barc)* 2014; 50(7): 503–525
161. Wang S, Yang DM, Rong R, Zhan X, Fujimoto J, Liu H, Minna J, Wistuba II, Xie Y, Xiao G. Artificial intelligence in lung cancer pathology image analysis. *Cancers (Basel)* 2019; 11(11): 1673
162. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyö D, Moreira AL, Razavian N, Tsirigos A. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med* 2018; 24(10): 1559–1567
163. Kather JN, Pearson AT, Halama N, Jäger D, Krause J, Loosen SH, Marx A, Boor P, Tacke F, Neumann UP, Grabsch HI, Yoshikawa T, Brenner H, Chang-Claude J, Hoffmeister M, Trautwein C, Luedde T. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat Med* 2019; 25(7): 1054–1056
164. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, Skora AD, Luber BS, Azad NS, Laheru D, Biedrzycki B, Donehower RC, Zaheer A, Fisher GA, Crocenzi TS, Lee JJ, Duffy SM, Goldberg RM, de la Chapelle A, Koshiji M, Bhajee F, Huebner T, Hruban RH, Wood LD, Cuka N, Pardoll DM, Papadopoulos N, Kinzler KW, Zhou S, Cornish TC, Taube JM, Anders RA, Eshleman JR, Vogelstein B, Diaz LA Jr. Pd-1 blockade in tumors with mismatch-repair deficiency. *N Engl J Med* 2015; 372(26): 2509–2520
165. Nagpal K, Foote D, Liu Y, Chen PHC, Wulczyn E, Tan F, Olson N, Smith JL, Mohtashamian A, Wren JH, Corrado GS, MacDonald R, Peng LH, Amin MB, Evans AJ, Sangoi AR, Mermel CH, Hipp JD, Stumpe MC. Development and validation of a deep learning algorithm for improving gleason scoring of prostate cancer. *NPJ Digit Med* 2019; 2(1): 1–10
166. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit.* 2016: 2818–2826
167. Ing N, Ma Z, Li J, Salemi H, Arnold C, Knudsen BS, Gertych A. Semantic segmentation for prostate cancer grading by convolutional neural networks. *Medical Imaging 2018: Digital Pathology International Society for Optics and Photonics.* 2018: 105811B
168. Badrinarayanan V, Kendall A, Cipolla R. Segnet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans Pattern Anal Mach Intell* 2017; 39(12): 2481–2495
169. Li W, Li J, Sarma KV, Ho KC, Shen S, Knudsen BS, Gertych A, Arnold CW. Path R-CNN for prostate cancer diagnosis and gleason grading of histological images. *IEEE Trans Med Imaging* 2019; 38(4): 945–954