

## Cell detection in pathology and microscopy images with multi-scale fully convolutional neural networks

Xipeng Pan<sup>1</sup> · Dengxian Yang<sup>2</sup> · Lingqiao Li<sup>1,3</sup> ·  
Zhenbing Liu<sup>3</sup> · Huihua Yang<sup>1,3</sup> · Zhiwei Cao<sup>1</sup> ·  
Yubei He<sup>4</sup> · Zhen Ma<sup>1</sup> · Yiyi Chen<sup>1</sup>

Received: 19 August 2017 / Revised: 6 November 2017 / Accepted: 6 December 2017  
© Springer Science+Business Media, LLC, part of Springer Nature 2018

**Abstract** Automated nucleus/cell detection is usually considered as the basis and a critical prerequisite step of computer assisted pathology and microscopy image analysis. However, due to the enormous variability (cell types, stains and different microscopes) and data complexity (cell overlapping, inhomogeneous intensities, background clutters and image artifacts), robust and accurate nucleus/cell detection is usually a difficult problem. To address this issue, we propose a novel multi-scale fully convolutional neural networks approach for regression of a density map to robustly detect the nuclei of pathology and microscopy images. The procedure can be divided into three main stages. Initially, instead of working on the simple dot label space, regression on the proposed structured proximity space for patches is performed so that centers of image patches are explicitly forced to produce larger values than their adjacent areas. Then, several multi-scale fully convolutional regression networks are developed for this task; this will enlarge the receptive field and not only can detect the single, small size cells, but also benefit to detecting cells with big size and overlapping states. In this stage, we copy the full feature maps from the contracting path and merge with the feature maps of the expansive path. This operation will make full use of shallow and deep semantic information of the networks. The networks do not have any fully connected layers; this strategy allows the seamless probability map prediction of arbitrarily large images. At the same time, data

---

This article belongs to the Topical Collection: *Special Issue on Deep Mining Big Social Data*  
Guest Editors: Xiaofeng Zhu, Gerard Sanroma, Jilian Zhang, and Brent C. Munsell

✉ Huihua Yang  
yhh@bupt.edu.cn

<sup>1</sup> School of Automation, Beijing University of Posts and Telecommunications, Beijing, China

<sup>2</sup> College of Arts and Sciences, University of Washington – Seattle, Seattle, Washington, USA

<sup>3</sup> School of Computer Science and Information Security, Guilin University of Electronic Technology, Guilin, China

<sup>4</sup> School of Computing and Information Systems, University of Melbourne, Melbourne, Australia

augmentations (e.g., small range shift, zoom and randomly flip) are carefully used to enhance the robustness of detection. Finally, morphological operations and suitable filters are employed and some prior information is introduced to find the centers of the cells more robustly. Our method achieves about 99.25% detection precision and the F1-measure is 0.9924 on fluorescence microscopy cell images; about 85.90% detection precision and the F1-measure is 0.9020 on Lymphocyte cell images and about 78.41% detection precision and the F1-measure is 0.8440 on breast histopathological images. This result leads to a promising detection performance that equates and sometimes exceeds the recently published leading detection approaches with the same benchmark datasets.

**Keywords** Cell detection · Fully convolutional neural networks · Pathology and microscopy images · Multi-scale

## 1 Introduction

Digital pathological image analysis plays an important role in decision making for disease diagnosis, since they can provide a large amount of useful information for computer-aided diagnosis (CAD). Nowadays automatic digital pathological image analysis has great benefits for pathologists and patients, and attracts a lot of attentions in both research and clinical practices [16, 51–53]. Manual examination of pathological images is an extremely tedious and time-consuming operation, and the results are also subject to intra- and inter-individual variability [28]. On the other hand, CAD can greatly reduce the bias and provide accurate characterization of diseases [15]. In addition, it allows individualized treatment, which is significantly beneficial for the patients.

A key prerequisite for computer-assisted diagnosis is nucleus/cell detection and segmentation, which are generally considered to be the bases of automatic image analysis. It supports various quantitative analyses including calculating cell morphology, such as shape, texture, and so on. While detection can determine the presence (and quantity) of an object of interest, such as cancer cells in a pathological image, in addition, detection can also be used as seeds for further segmentation, counting, tracking or morphological analysis. So cell detection is often the basis and key step in pathological image diagnosis. However, robust and accurate nucleus/cell detection is a difficult problem due to the enormous variability (cell types, stains and different microscopes) and data complexity (cell overlapping, inhomogeneous intensities, background clutters and image artifacts) [18, 40, 47].

In recent years, more and more people are interested in the application of “deep learning” strategy in natural and pathological image analysis. Pathology, in view of its diversity and complexity, is an excellent use case for deep learning strategies. One of the main challenges of computational techniques is to analyze all individual cells for accurate diagnosis, because the differentiation of most disease grades is highly dependent on cell-level information. In order to achieve this goal, deep convolutional neural network has been investigated to detect and segment cells from histopathological images [26, 29, 41, 44], which is of great significance to the cell level analysis of tumor diagnosis.

In this paper, we propose a novel multi-scale fully convolutional neural networks approach for regression of a density map to robustly detect the nuclei of pathology and microscopy images. The main contributions of this work include three parts. First, we devise and compare architectures for several multi-scale fully convolutional regression networks for this task; this

will enlarge the receptive field and not only can detect the single, small size cells, but also benefit to detecting cells with big size and overlapping state. Since the networks are fully convolutional, they can predict a density map for an input image with arbitrary size, and we exploit this to improve efficiency by end-to-end training on image patches. Second, instead of working on the simple dot label space, regression on the proposed structured proximity space for patches is performed so that centers of image patches are explicitly forced to produce larger values than their adjacent areas. At the same time, data augmentations (e.g., small range shift, zoom and randomly flip) are carefully used to enhance the robustness of detection. Finally, extensive experiments and their results lead to a promising cell detection performance that equates and sometimes exceeds the recently published leading detection approaches with the same benchmark datasets.

The rest of the work is organized as follow. In Section 2, we will review the recent deep learning methods for the nuclei detection of microscopy and pathology images. Section 3 presents our methodological contributions in detail, while Section 4 describes the experimental setup, validation results and comparisons with the previous methods. Section 5 draws a conclusion to the paper.

## 2 Related work

In this section, we review three major categories of relevant work: convolutional neural networks (CNN) for cell detection, CNN combined with other methods for cell detection, and other deep learning methods for cell detection.

### 2.1 CNN for cell detection

Cruz-Roa et al. [9] put forward a kind of deep learning method that integrated image representations, basal-cell carcinoma cancer detection and visual interpretation for the skin histopathology images. Their approach obtained much better performance than the canonical methods, such as bag of features, discrete cosine transform and wavelet transform. Janowczyk et al. [2] employed an open source framework (Caffe) with an improved AlexNet-based architecture to address seven digital pathology (DP) image analysis tasks. In many cases, the performance of the paper was superior to results from the state-of-the-art hand-crafted feature-based [19, 55] classification approaches. Sirinukunwattana et al. [34, 35] introduced the spatial constraint strategy into the deep convolution network, which makes the nucleus center obtain higher probability value. This will be beneficial to the accurate detection of epithelial tumor nuclei, but the sliding window strategy with overlapping windows is very time consuming. In [8], a deep convolution neural network was trained to detect mitotic cells in breast cancer histology images by pixel-wise style. The specific method was to take a pixel point as the center, select a batch, and predict the patch to achieve the pixel classification. Finally, a smooth operation of the probability map was carried out, and the center of mitotic cells was obtained by combining the non-maximal suppression. This method had won the ICPR 2012 mitosis detection contest, which greatly exceeded the other algorithms. Dong et al. [11] employed a nine-layer CNN to extract features, and fed them to the SVM classifier. The proposed method outperformed the hand-crafted algorithms and obtained a promising result for cell detection in wide-field microscopy zebrafish images. Xing et al. [48] proposed a general pathologic cell detection and segmentation framework. The main contributions of this paper were divided into

two parts. First, CNN model was employed to obtain the reliable probability density map, which can benefit the subsequent cell shape inference. Second, a deformation model based on sparse shape constraint was proposed to effectively deal with the nuclei with weak or missing boundaries. However, if the CNN model was deployed to test images in a sliding window manner for pixel-wise classification, it would be computationally expensive for large-scale images. In [17], a fast scanning strategy was introduced to a deep CNN, which sharply sped up CNN on sliding windows. As the authors declared, the simple MATLAB implementation of the scanning strategy yielded more than 30-fold speedup over patch-based GPU implementation for neuronal membranes segmentation in electron microscopy images. Xie et al. [45] proposed a structural regression model for robust nuclei detection. This approach differed from traditional CNN in that it introduced the topological information for acquiring training data and adopted fast scanning technology. The proposed method obtained state-of-the-art performance on three modalities of cell images. Ciresan et al. [10] adopted CNN as the pixel-wise classifier and won the best performance on the three metrics (rand error, warping the error and pixel error) through plain stochastic gradient descent training strategy and non-special post-processing. However, the algorithm had a lot of computational redundancy, which greatly reduced the prediction speed. At the same time, the selection of the size of the patches must compromise the localization accuracy and the context size. Ronneberger et al. [32] presented a novel network architecture which consisted of a contracting path and a symmetric expanding path. The proposed method takes advantage of the available annotated samples more efficiently and achieves good performance for biomedical images segmentation.

## 2.2 CNN combined with other methods for cell detection

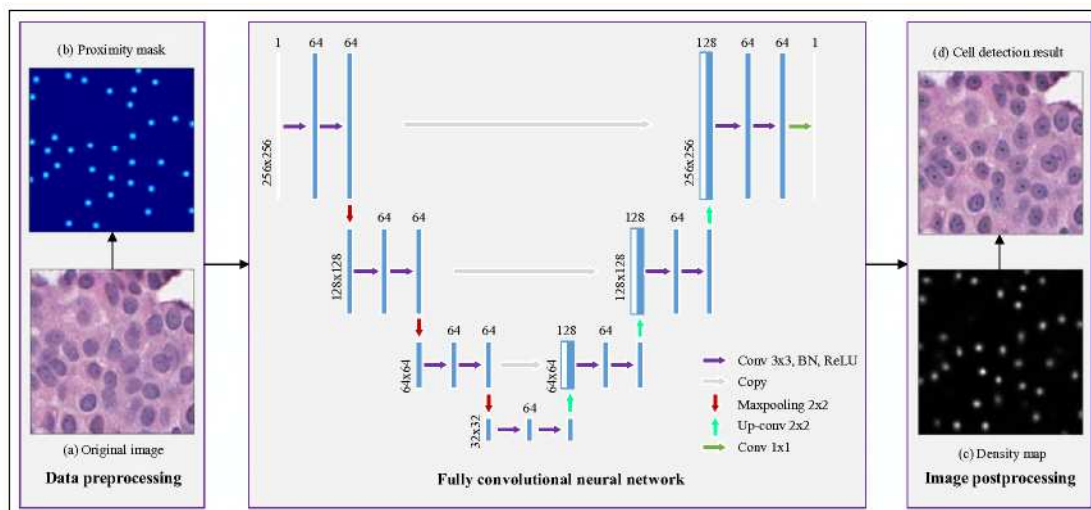
Liu et al. [26] formulated cancer nucleus detection into a combinatorial optimization problem. Specifically, it consisted of three steps. First, a set of nucleus candidates were produced by using different methods including LoG, MSER [3], and iterative voting [30]. Second, each candidate was assigned a score using a seven layer CNN and a weighted graph was constructed. Third, a subset of mutually non-adjacent graph nodes was chosen to maximize the sum of the weights of the selected nodes. Xie et al. [44] proposed a novel method that integrates CNN and Hough voting to accurately locate the nucleus of microscopy images. In the test phase, a fast scanning technique and the stride manner were introduced to boost the speed. Both of them are proved to be efficient. Song et al. [37] proposed a multiscale convolutional network(MSCN) and graph-partitioning-based method for accurate segmentation of cervical cytoplasm and nuclei, which was composed of 1) extracting feature representation and conducting pixel-wise coarse segmentation with a CNN, 2) building an undirected graph based on the coarse segmentation and superpixels generated from SLIC [1] and partitioning the graph with a fast min-cut/max-flow algorithm [6], and 3) finally calculating nucleus markers based on the marker seeking strategy. Sommer et al. [36] designed a composite framework for high-content cellular phenotype detection. The authors employed convolutional autoencoder to learn the representation features for a collection of images, and the features served as the input for novelty detection and the conventional supervised machine learning classifier. Wang et al. [42] combined the handcrafted and CNN features to boost the performance for mitosis cell detection. Khoshdeli et al. [21] employed CNN and Laplacian of Gaussian (LoG) filter to accentuate blob-shape objects, which was benefit for the detection of hyperchromatic and vesicular phenotype cells in histology sections of brain and breast images.

## 2.3 Other deep learning methods for cell detection

Su et al. [41] presented a cell detection and segmentation method in histopathological image, which is divided into two steps. First, the sparse reconstruction [56, 58] model was employed to deal with shape variations, and cell overlapping. Second, a stacked denoising autoencoder, trained with original data and their structural labels, was used for cell segmentation. Xu et al. [49] employed stacked sparse autoencoder (SSAE) to detect the nuclei of breast cancer histopathological images. It mainly consisted of two parts. First, plenty of nuclei and non-nuclei patches were sent to the SSAE to train a robust classifier. Then, the classifier was employed to scan on the high-resolution histopathological images to detect the nuclei. Yellin et al. [50] plugged convolution operation into sparse dictionary learning and coding [54, 57], which was beneficial to the detection of red blood cell (RBC) and white blood cell (WBC) in the holographic lens-free image. In [38], the authors proposed a deep autoencoder network to detect and classify erythroid and myeloid cells simultaneously. The proposed network not only decreased the training time, but also kept almost the same performance with the conventional deep learning methods. Song et al. [39] devised an unsupervised hybrid network with curvature Gaussian model for detecting various types of cells in bone marrow hematopoietic stem cells. The method attained better F1-score than some state-of-the-art algorithms and could detect some U-shaped cells, but the size of the training cell patches was small (e.g.,  $29 \times 29$ ).

## 3 Methodology

We propose a fully convolutional neural networks method to detect the cells in fluorescence microscopic images and pathological images. We modify and extend the “fully convolutional network” architecture such that it works with very few training images and gains more precise probability map. A flowchart of the proposed method is illustrated in Figure 1. The method mainly consists of three parts: 1) Data preprocessing. The benchmark datasets only give the centers of cells (one dot per cell) as ground truth. Instead of working on the simple dot label



**Figure 1** A flowchart of the proposed method

space, the regression on the proposed structured proximity space for patches is performed so that centers of image patches are explicitly forced to produce larger values than their adjacent areas. 2) Multi-scale fully convolutional neural networks. To increase the robustness and the detection performance, multi-scale feature maps generated from the scratch input images are merged together (See section 3.3 for detailed description). 3) Image postprocessing. In order to improve the detection performance and reduce the errors due to background clutters, stains and image artifacts, morphological operations and suitable filters are employed and some prior information is introduced to find the centers of the cells more robust.

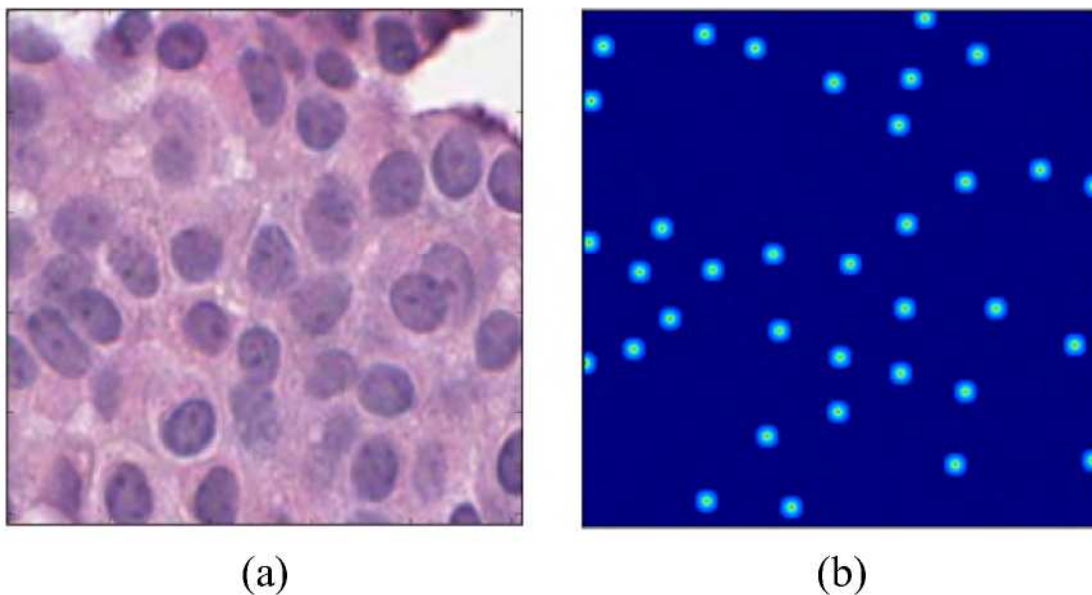
### 3.1 Data preprocessing

The label image, which is corresponding to original cell image, contains only the center point of the cell as its label. Our method will process the label image and generate structured outputs (referred to as proximity patches). In these patches, the cell center has a larger value and the maximum intensity value is the cell center. In the training phase, proximity patches as the training sets that will explicitly force the center of the adjacent patches to produce larger values than their adjacent areas. The maximum value of the proximity map generated by our method will be more robust as a cell center.

Let  $I$  denotes the input image, we define  $M$  as the proximity mask corresponding to image  $I$ , and compute the value of the  $ij$ -th entry in  $M$  in formula

$$M_{ij} = \begin{cases} 1/(1 + \alpha D(i, j)) & \text{if } D(i, j) \leq r, \\ 0 & \text{otherwise,} \end{cases}$$

where  $D(i, j)$  represents the Euclidean distance from pixel  $(i, j)$  to the nearest pathologist annotated cell center.  $r$  is a distance threshold which has been set to a small value (e.g., 5 pixels).  $\alpha$  is the decay ration which has been set to 0.8. The value of  $M_{ij}$  is a number within the interval  $V = [0, 1]$ . The details are shown in Figure 2.



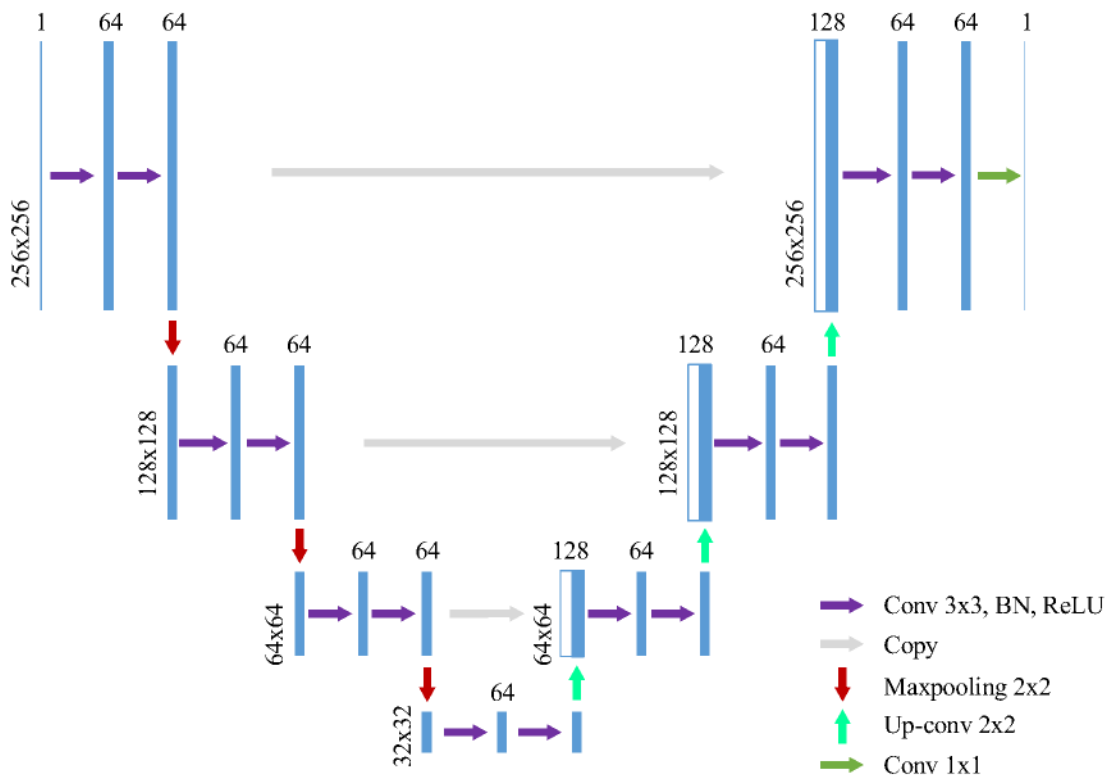
**Figure 2** The original image and the proximity masks. (a) Original RGB image with simple dot labels. (b) Proximity masks corresponding to the original image with the same size

The proposed approach is different from the conventional CNN classifier, which introduces a new structured regression method to capture the topologic information implying in the training data. The image also maintains spatial consistency.

## 3.2 Fully convolutional neural networks

### 3.2.1 Architecture design

The typical application of the convolution network is the classification task, where the output of the image is a class label. However, in many visual tasks, especially in biomedical image processing, the expected output should include localization. That is to say, a class label should be assigned to each pixel. Inspired by literature [27], we design a more elegant architecture, which works with very few training images and obtains more precise probability map. As shown in Figure 3, it consists of a contracting path and an expansive path. The contracting path follows a convolutional network and consists of the repeated application of two  $3 \times 3$  convolutions (padded convolutions). Each convolution operation followed a  $2 \times 2$  max pooling operation for down-sampling, and the activation function is rectified linear unit (ReLU). To make the training easier, we add batch normalization [20]. Every step in the expansive path consists of three details. First, an up-sampling operation enlarges the size of every feature map twice and keeps the number of feature channels. Second, the feature maps concatenate with the corresponding full feature maps from the contracting path. Third, it conducts two  $3 \times 3$  convolution operations, each followed by a ReLU.



**Figure 3** The proposed architecture (named Unet-like). Every blue box denotes the multi-channel feature maps. The number on top of the box is the number of feature maps. The size of the feature map is shown on the left edge of the box. The arrows represent the different operations in the network

The proposed structure of the networks is similar to U-net [32], but there are three differences. First, U-net crops the feature maps from the contracting path and concatenate with the feature maps of the expansive path, however, we copy the full feature maps from the contracting path and merge with the feature maps of the expansive path. Second, the size of the density map of our proposed model is the same as the size of input image, which is beneficial for the determination of the spatial position of the cell images. Third, U-net doubles the number of feature channels at each down-sampling step that will significantly increase the parameters of the model and reduce the training speed. In the meantime, we employed batch normalization to make the training easier. At the same time, the network does not have any fully connected layers; this strategy allows the seamless probability map prediction of arbitrarily large images. To predict the pixels in the border region of the image, the missing context is extrapolated by mirroring the input image. The proposed architecture is shown in Figure 3.

### 3.2.2 Implementation details

The implementation is based on keras (<https://github.com/fchollet/keras>), and the backend is tensorflow (<https://github.com/tensorflow/tensorflow>). During the training period, we crop patches from the large images, for instance, we randomly sample patches of size  $256 \times 256$  pixels from big size ( $1600 \times 1200$ ) breast cancer pathological images. Some data augmentation techniques (e.g., small rotations, zoom, horizontal flipping and vertical flipping) are also applied. Before training, each patch is normalized by subtracting its own mean intensity value and then dividing by the standard deviation.

We denote the function of the convolutional network from the input  $X$  to the output  $Y$  as  $Y = F_{w,b}(X)$ , where  $w$  and  $b$  are free parameters of the network.  $Y^*$  is the ground truth annotation with respect to  $Y$ . The cost function is given by

$$E(X, F_{w,b}, Y^*) = \|F_{w,b}(X) - Y^*\|_2^2 \quad (1)$$

The learning process of the convolution network can be defined as the minimization of the cost function with respect to all free parameters  $w$  and  $b$ .

$$\operatorname{argmin}_{w,b} E(X, F_{w,b}, Y^*) \quad (2)$$

Stochastic gradient descent (SGD) with momentum is used for optimization. Then, the parameters  $w$  are updated by:

$$\nu = \beta \times \nu - \eta \times \nabla_w E \quad (3)$$

$$w \leftarrow w + \nu \quad (4)$$

where  $\eta$  is some sufficiently small value called learning rate, which control the convergence speed. In this paper, the learning rate  $\eta$  is initiated as 0.001, and gradually decreases by a factor of 10. Before learning, the parameters  $w$  are initialized with an orthogonal basis [33].

The momentum  $\beta$  is set to 0.9 and weight decay value is 0.0005. Since most of the pixel values in the label are zeros, moreover, the peak value of non-zero region is 1; network will be

especially difficult to train. In order to alleviate this problem, we simply scale the structured proximity space-annotated ground truth (Figure 2) by a factor of 100, which forcing the network to fit structured proximity space rather than the background zeros.

### 3.3 Multi-scale fully convolutional neural networks

Nevertheless, robust cell detection in pathological image is generally a difficult problem due to the large variability. For example, the size of a single cell varies greatly, very large cell clumps, where three or more cells overlap, especially the latter. In this case, larger clumps are more susceptible to deformation than individual cells and are harder to regress; In addition, regression of large cell clumps requires the network to have an even larger receptive field, which can cover important parts of the whole clump, such as concave information, or curved edges in specific directions. The math formula of receptive field (RF) is as follow:

$$l_k = l_{k-1} + \left( (f_k - 1) \times \prod_{i=1}^{k-1} s_i \right) \quad (5)$$

where  $l_{k-1}$  is the receptive field of layer  $k-1$ ,  $f_k$  is the filter size (height or width, but assuming they are the same here), and  $s_i$  is the stride of layer  $i$ .

The formula above calculates receptive field from bottom up (from layer 1). Intuitively, RF in layer  $k$  covers  $(f_k - 1) \times s_{k-1}$  more pixels relative with layer  $k-1$ . However, the increment needs to be translated to the first layer, so the increments is a factorial, that is to say, a stride in layer  $k-1$  is exponentially more strides in the lower layers.

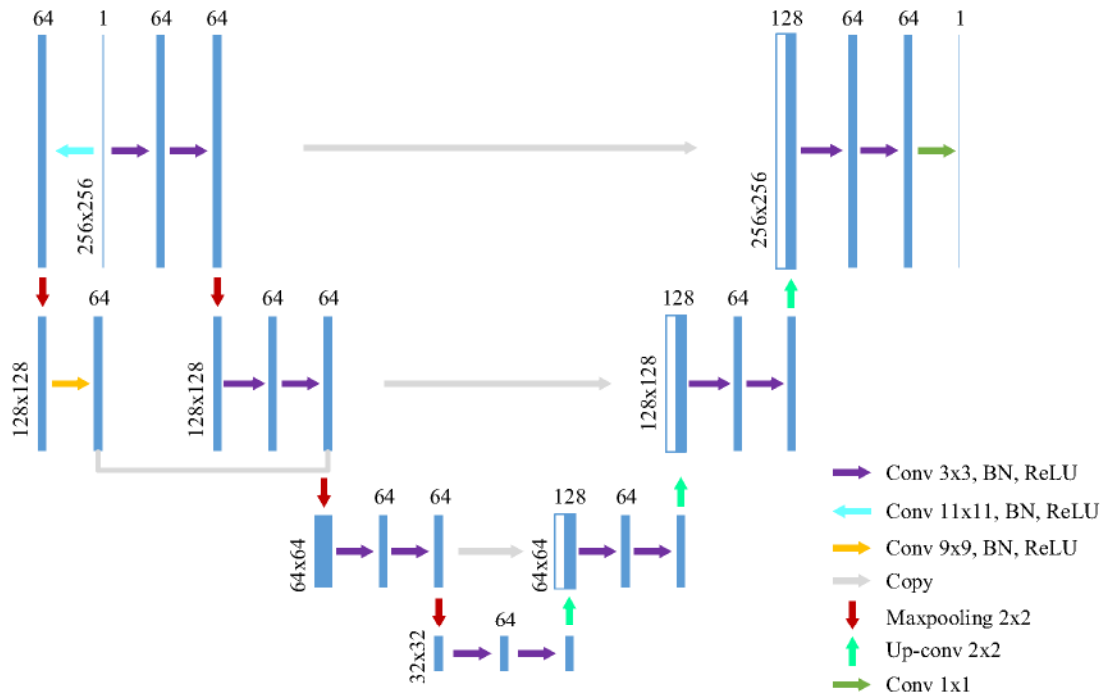
In [46], the proposed network is relatively shallow, with only  $38 \times 38$  or  $32 \times 32$  pixels of the receptive field. For small cell clumps, their curved edges are usually covered and can be predicted correctly. However, for a round cell clump with about three or more cells, it may be larger than their largest receiving field, which usually leads to a false prediction. Inspired by the literature [7, 31, 43], we build upon the multi-scale networks to deal with the intractable tasks.

The structure of our network is illustrated in Figure 4. Base on the original network (Figure 3), we add a branch network (named coarse branch network) from the input, and the output feature maps merge with the fourth convolution layer of the original network. The feature maps of the coarse branch network are used as the additional ones; by design, the size of the coarse branch network output feature maps is the same as the output of the fourth convolution layer. The size of convolutional kernel in the coarse branch network is  $11 \times 11$  and  $9 \times 9$  respectively, which is bigger than original network and will enlarge the size of receptive field of the whole network. In fact, the size of the Multiscale\_4layers architecture and the Multiscale\_x2\_4layers architecture is  $82 \times 82$  and  $108 \times 108$  respectively, which is benefit to detecting the large size cells and overlapping cells. As the proposed two architectures look “4 layer-like”, we name the models suffixed by “4layers”. Furthermore, we also train several networks with “3 layers” as shown in Figure 5, and performance comparisons with two “4 layer-like” architectures are shown in the sections below.

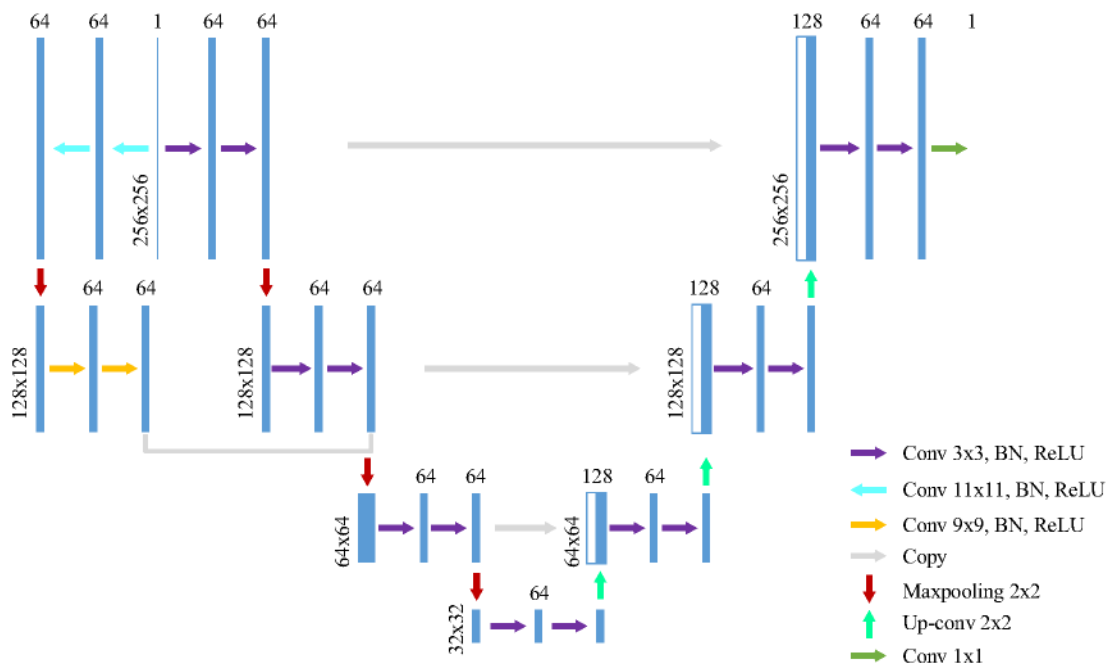
### 3.4 Postprocessing

In the postprocessing phase, in order to improve the detection performance and reduce the errors due to background clutters, stains, and image artifacts, morphological operations and suitable filters are employed and some prior information is introduced. First, the isolated pixels and small areas of noises can be eliminated by the morphological clean and opening operation.

Second, the center areas of nuclei can be obtained by maximum filter, which is the built-in function in Scipy (one of the famous open-source software for mathematics, science, and engineering). Third, we label the connected domains and calculate the center coordinates of the domains, and that are the geometric centers of nuclei. The details are shown in Figure 6.

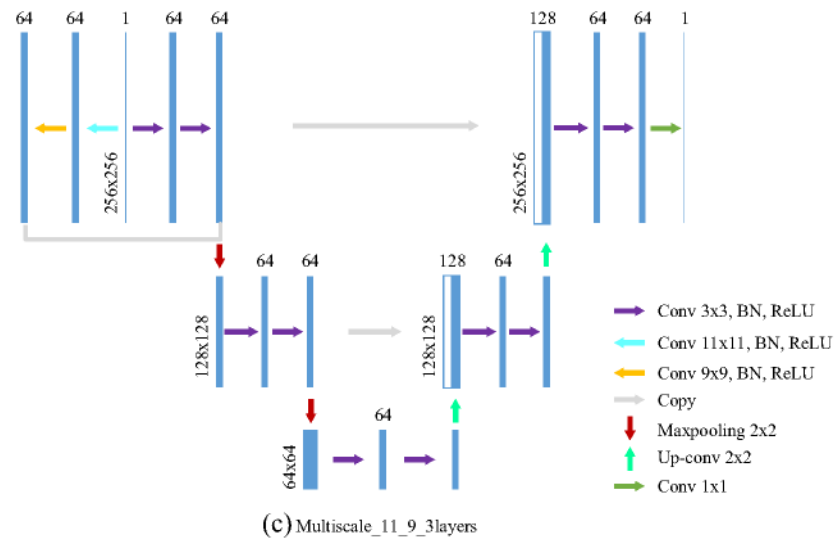
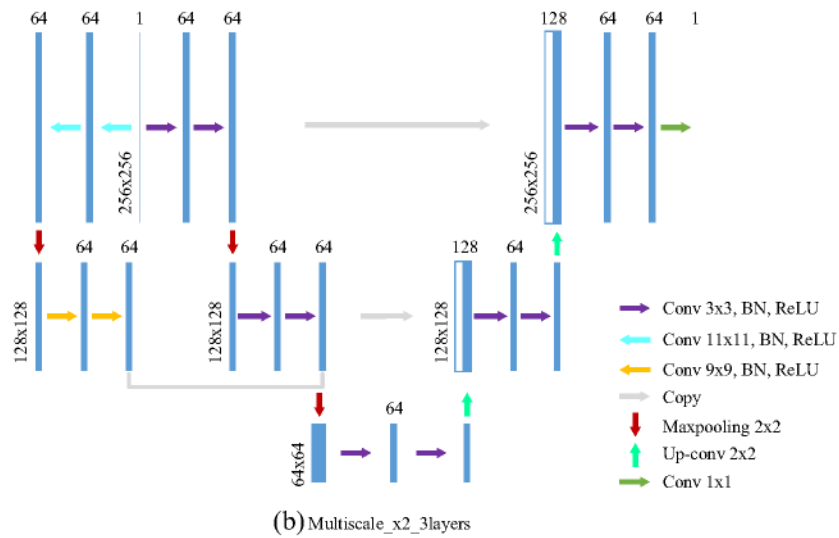
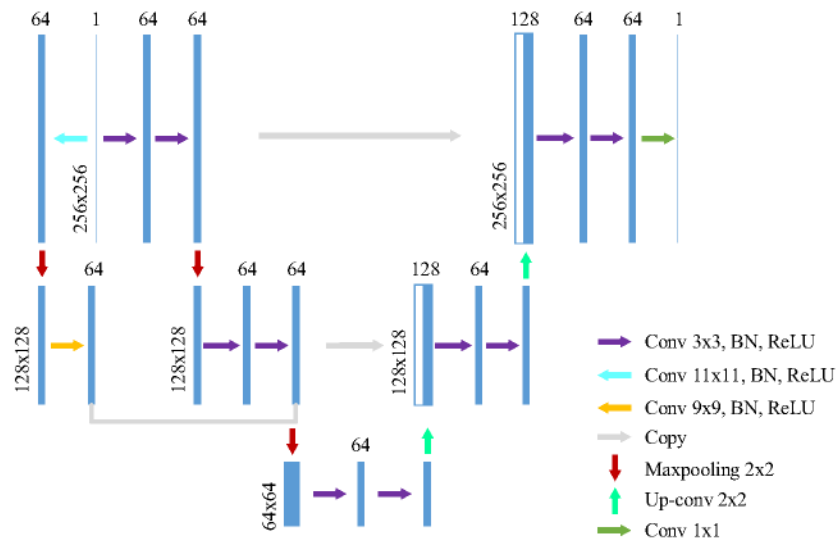


(a) Multiscale\_4layers

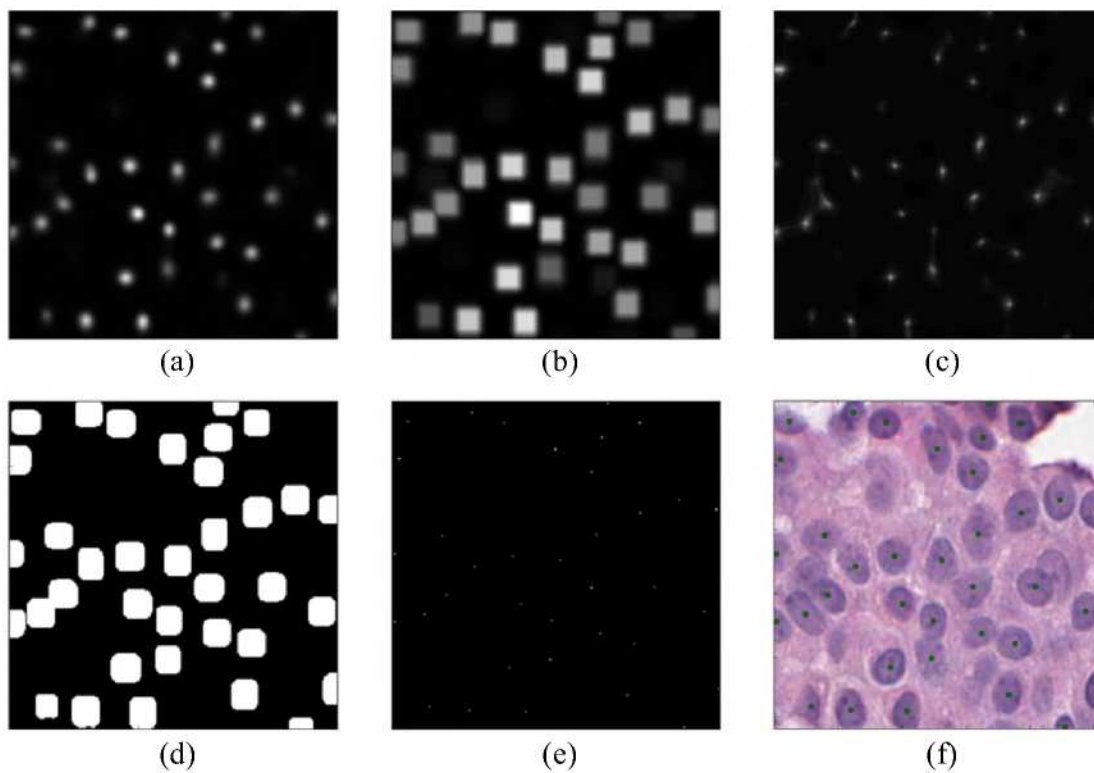


(b) Multiscale\_x2\_4layers

**Figure 4** The proposed four layer architectures. (a) Multiscale\_4layers. (b) Multiscale\_x2\_4layers



**Figure 5** The proposed three layer architectures. (a) Multiscale\_3layers. (b) Multiscale\_x2\_3layers. (c) Multiscale\_11\_9\_3layers

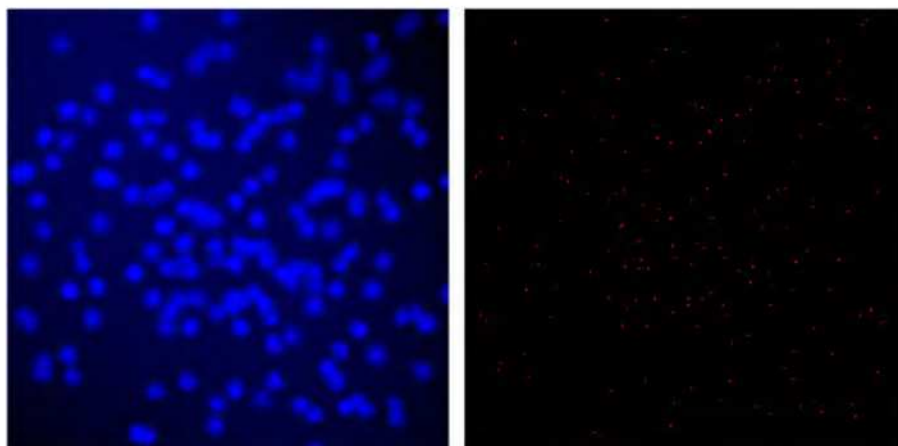


**Figure 6** Postprocessing process. (a) Density map of the test image. (b) Image after the maximum filter of the density map. (c) Image after the minimum filter of the density map. (d) Difference between images after maximum filter and minimum filter. (e) The maxima areas of the density map. (f) Detection result

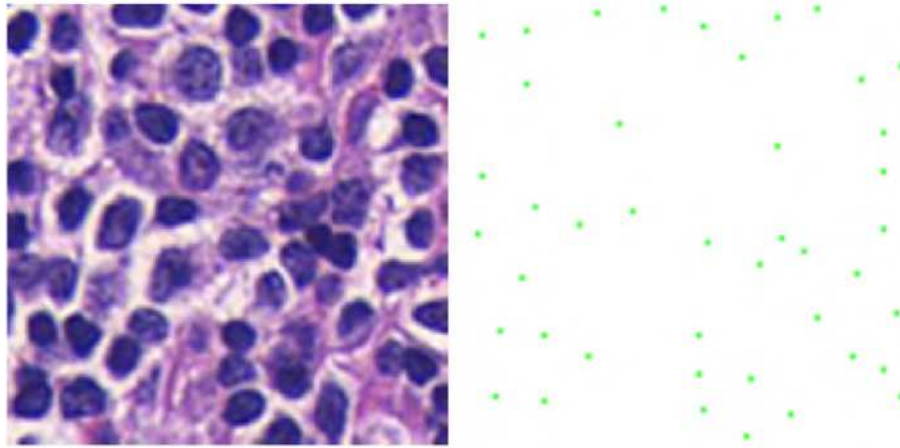
## 4 Experiments and results

### 4.1 Dataset description

Three datasets are used to validate our proposed models. The first dataset is synthetic dataset which contains 200 cell images of fluorescence microscopy. The second and third datasets are two pathological images datasets: lymphocyte cell images and breast cancer histopathological images. More detail will see the next subsection.



**Figure 7** Synthetic data (fluorescence microscopy images with cell nuclei)



**Figure 8** Lymphocyte cell image and its associated ground truth cell detection labeling image

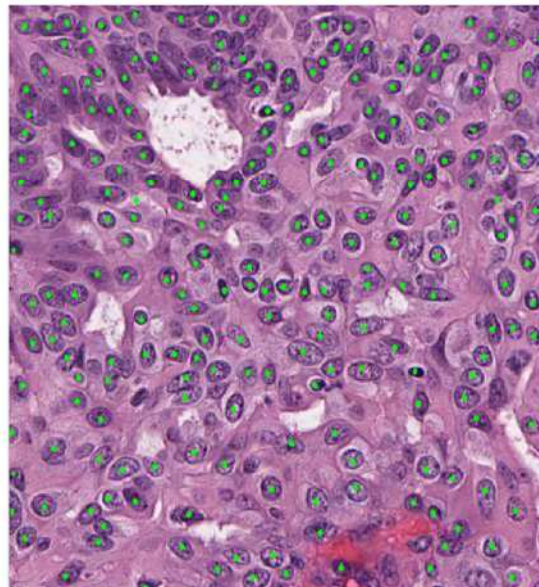
#### 4.1.1 Synthetic data

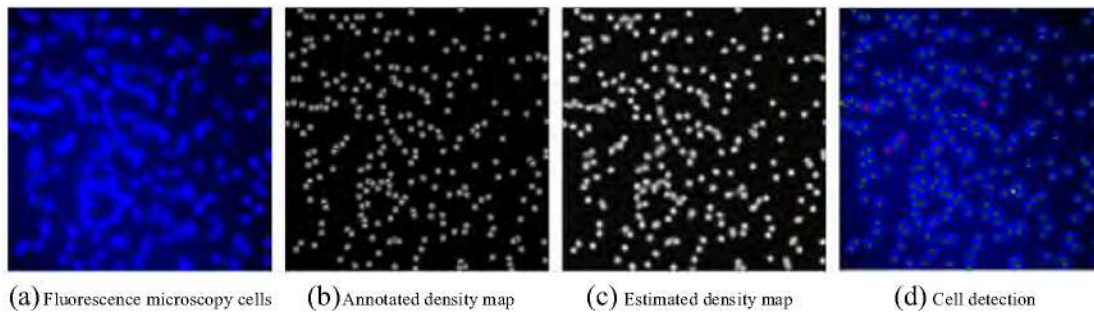
The synthetic datasets consist of 200 cell images of fluorescence microscopy that generated in Lehmuussola et al. [23]. As shown in Figure 7, each synthetic image has about  $174 \pm 64$  cells. Serious overlap between instances is often observed in the dataset, which makes detection challenging. As Figure 4 in literature [46] shows, human experts can't even tell the difference between the overlapping cells and a single one. We report the mean precision, recall, F1-measure and their standard deviations for the proposed models.

#### 4.1.2 Real data

We evaluated our proposed models on other two different kinds of data; (1) Lymphocyte cell images. Lymphocyte, a subtype of white blood cells, is an important component of the immune system. Lymphocytic infiltration is the process where the density of lymphocytes greatly increases at sites of disease or foreign bodies, indicating an immune

**Figure 9** Breast cancer histopathological image and its associated ground truth cell detection labeling image (green dots)





**Figure 10** Multiscale\_x2\_4layers applied on fluorescence microscopy cells. (a) Fluorescence microscopy cells. (b) Annotated density map of the label image. (c) Estimated density map. (d) Cell detection for the fluorescence microscopy cell image

response. A stronger immune response has been highly correlated to better outcomes in many forms of cancer. As a result, identifying and quantifying the density and location of lymphocytes has gained great interest recently, particularly in the context of identifying which cancer patients to place on immunotherapy [2]. Lymphocytes present with a blue tint from the absorption of hematoxylin, their appearance similar in hue to nuclei, making them difficult to differentiate in some cases. As shown in Figure 8, lymphocytes tend to be smaller, more chromatically dense, and circular. Our goal was to identify the center of lymphocytes, making this a detection problem. (2) Breast cancer histopathological images. Breast cancer is the second most common cancer in a woman's lifetime. Classical examination methods include screening and biopsy. Compared to mammography, histopathologic slides provide more comprehensive diagnostic information [12, 52]; nevertheless, the histopathological image is much more complicated. As Figure 9 shows, the size and shape of the nuclei are different, some nuclei are clustered closely together and overlap, and there exist some image artifacts in the HE (hematoxylin and eosin) histopathological image, which will increase the challenge of nuclei detection.

## 4.2 Performance metrics

Precision (P), recall (R), and *F1-measure* ( $F_1$ ) are adopted as the metrics of cell detection performance, which are given as:

$$P = \frac{TP}{TP + FP} \quad (6)$$

**Table 1** The comparison results of the cell detection on the standard synthetic cell images

Methods	Precision (%)	Recall (%)	F1-measure	RF
FCRN-A [46]	99.06 + -0.04	96.71 + -0.67	0.9790 + -0.0033	38
FCRN-B [46]	98.73 + -0.04	96.56 + -0.56	0.9764 + -0.0029	32
Unet-like	99.10 + -0.04	97.15 + -0.55	0.9813 + -0.0027	68
Multiscale_4layers	99.24 + -0.04	99.00 + -0.31	0.9911 + -0.0014	82
Multiscale_x2_4layers	99.25 + -0.04	99.23 + -0.24	0.9924 + -0.0011	108
Multiscale_3layers	99.25 + -0.09	99.02 + -0.28	0.9913 + -0.0010	46
Multiscale_x2_3layers	99.23 + -0.09	99.00 + -0.31	0.9911 + -0.0012	72
Multiscale_11_9_3layers	99.33 + -0.08	98.89 + -0.23	0.9911 + -0.0008	46

**Table 2** The comparison results of the cell detection on Lymphocyte cell images

Methods	Precision (%)	Recall (%)	F1-measure	RF
EMaGACOR [13]	86.00	64.00	–	–
LIPSyM [22]	70.21	70.08	0.6984	–
MSER-detector [4]	86.99	90.03	0.8848	–
Janowczyk et al. [2]	87.00 + -2.00	93.00 + -1.00	0.9000 + -0.0100	–
FCRN-A [46]	85.37 + -0.42	96.04 + -0.15	0.9002 + -0.0028	38
FCRN-B [46]	83.93 + -0.91	96.43 + -0.24	0.8934 + -0.0061	32
Unet-like	85.13 + -0.41	96.43 + -0.38	0.9006 + -0.0030	68
Multiscale_4layers	86.30 + -0.76	95.34 + -0.14	0.9017 + -0.0041	82
Multiscale_x2_4layers	85.90 + -0.86	95.85 + -0.69	0.9020 + -0.0045	108
Multiscale_3layers	84.82 + -0.56	96.11 + -0.20	0.8974 + -0.0030	46
Multiscale_x2_3layers	85.20 + -0.52	95.97 + -0.36	0.8987 + -0.0017	72
Multiscale_11_9_3layers	84.71 + -0.21	96.31 + -0.22	0.8977 + -0.0022	46

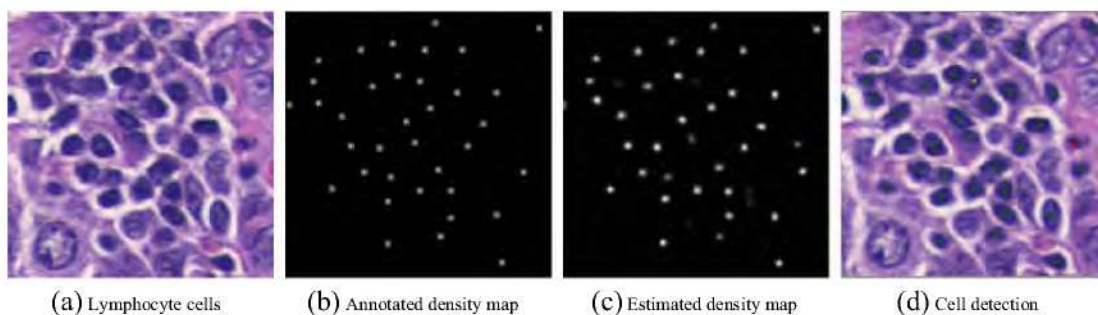
$$R = \frac{TP}{TP + FN} \quad (7)$$

$$F_1 = \frac{2 \times P \times R}{P + R} \quad (8)$$

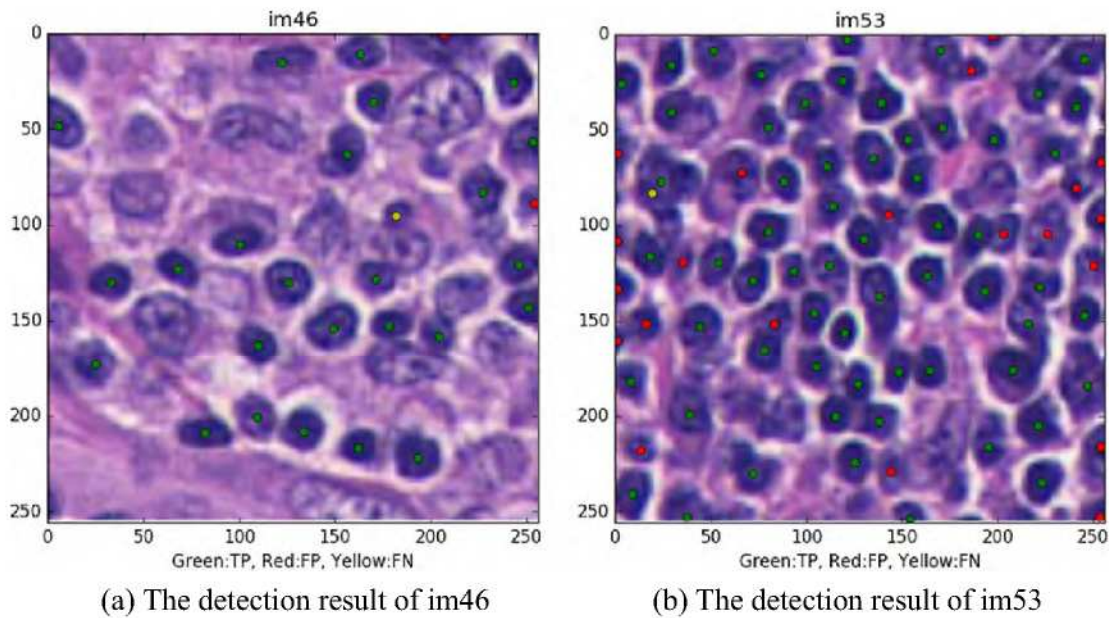
In the total number of nuclei detected from the proposed method, precision represents the ratio that exists in the ground truth image. It is also known as the positive predictive value. In the total number of nuclei present in the ground truth image, the percentage of nuclei detected by the method is recall. It is also called sensitivity. *F1-measure* is computed as the harmonic mean of precision and recall. *TP* represents true positive, *FP* stands for false positive, and *FN* stands for false negative [29].

### 4.3 Evaluation on synthetic data

We randomly select 100 images of 200 images as training sets to optimize the detection algorithm, and other images as test sets, are used to evaluate the performance. Randomly repeated 5 times experiments to calculate the mean and variance are used to evaluate the performance. During the testing phase, each image is mapped to a density map, and then the local maximum gives the cell detection of the region (see Figure 10(d)). The study is



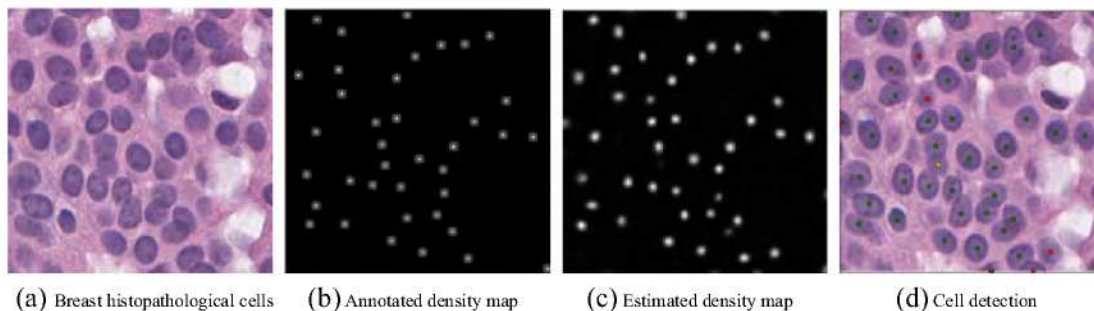
**Figure 11** Multiscale\_x2\_4layers applied on Lymphocyte cells. (a) Lymphocyte Cells. (b) Annotated density map of the label image. (c) Estimated density map. (d) Cell detection for the lymphocyte cell images



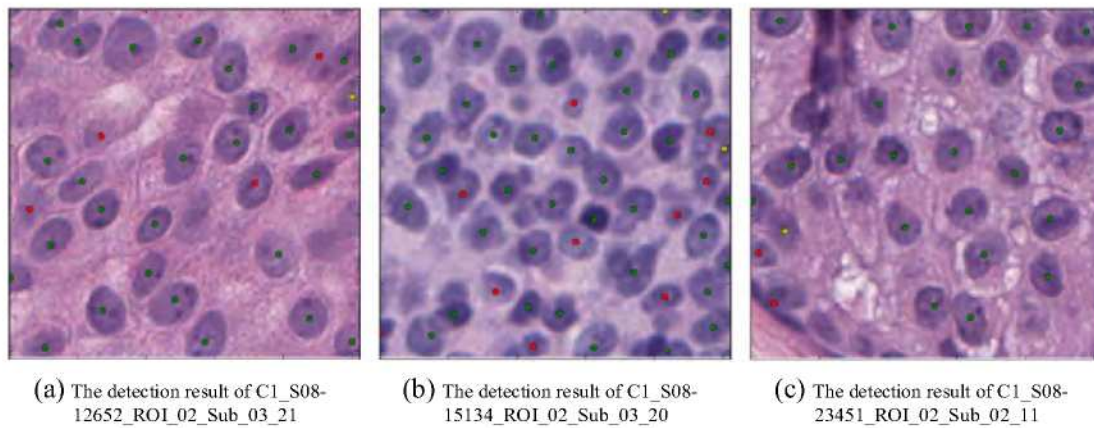
**Figure 12** Multiscale\_x2\_4layers applied on Lymphocyte cells. (a) The detection result of im46. (b) The detection result of im53

implemented with python on a workstation with Intel(R) Xeon(R) E5–2650 v2 CPU, 32 GB memory and the GPU is GTX1080. Comparison experiments are carried out between our proposed methods and the state-of-the-art approaches or models.

As Table 1 shows, FCRN-A performs slightly better than FCRN-B, Unet-like performs a little bit better than FCRN-A. Our proposed “Multiscale” series models obtain better performance than Unet-like, FCRN-A and FCRN-B. The size of the receptive field is more important than being able to provide more detailed information over the receptive field. Our hypothesis is that real difficulty of cell counting lies in the regression for large cell clusters, and a larger receptive field is needed to span them. From Table 1, we can see that the recall and F1-measure of our proposed method (e.g., Multiscale\_x2\_4layers) are 2.77% and 1.64% higher than FCRN-B [46], respectively. It is worth noting that for medical diagnosis, even if the diagnostic accuracy increased by 1% is also significant [25]. The performance of FCRN is much better than the state-of-the-art models [5, 14, 24], so our proposed methods are also better than [5, 14, 24].



**Figure 13** Multiscale\_x2\_4layers applied on breast histopathological cells. (a) Breast histopathological cells. (b) Annotated density map of the label image. (c) Estimated density map. (d) Cell detection for the breast cancer histopathological image



**Figure 14** Multiscale\_x2\_4layers applied on breast histopathological cells. (a) The detection result of C1\_S08-12652\_ROI\_02\_Sub\_03\_21. (b) The detection result of C1\_S08-15134\_ROI\_02\_Sub\_03\_20. (c) The detection result of C1\_S08-23451\_ROI\_02\_Sub\_02\_11

#### 4.4 Evaluation on real data

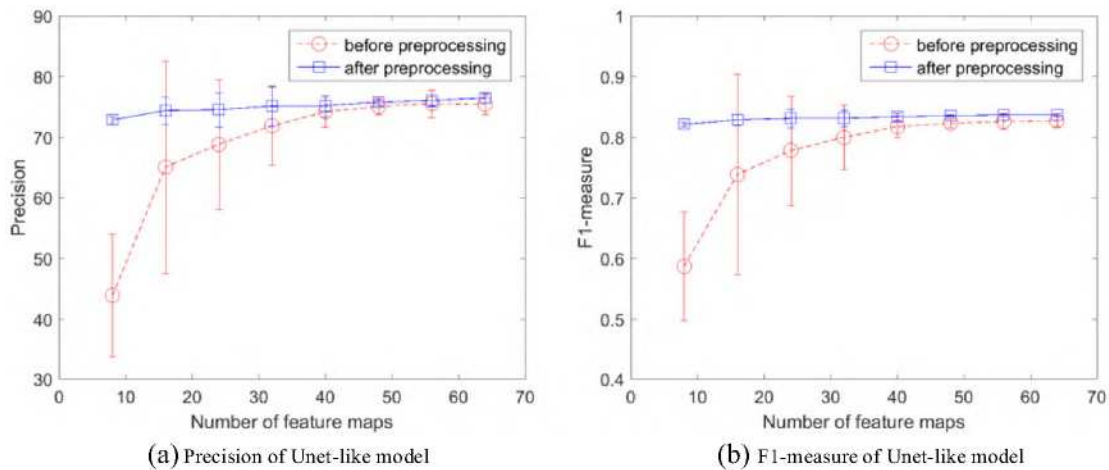
We also test the proposed methods on real datasets for detection. The proposed models are tested with two datasets: 1) 99 Lymphocyte cell images, the number of cells per image is about 30. 2) 285 breast cancer histopathological images. Each image contains roughly 35 cells. For each dataset, randomly repeated 5 experiments to calculate the mean and variance are used to evaluate the performance. All the cell centers are manually labeled by pathologists.

##### 4.4.1 Lymphocyte cell images

The metrics use to evaluate lymphocyte detection include: 1) precision; 2) recall; and 3) f1-measure (see Table 2). We test our proposed networks on Lymphocyte cells image for detection. Here, we only show figures for results from Multiscale\_x2\_4layers model in Figures 11 and 12. The detection results from the state-of-the-art models (EMaGACOR [13], LIPSyM [22], MSER-detector [4], Janowczyk et al. [2] and FCRN [46]) and our proposed models are compared to the manual detection results obtained from an expert clinician. The performances of our proposed models demonstrate a favorable comparison to the state-of-the-art which show (a) a precision of 86.00% and recall of 64.00% [13]; (b) precision of 70.21%, recall of 70.08% and F1-measure of 0.6984 [22] and (c) precision of 86.99%, recall of 90.03% and F1-measure of 0.8848 [4].

**Table 3** The comparison results of the cell detection of breast cancer histopathological images

Methods	Precision (%)	Recall (%)	F1-measure	RF
FCRN-A [46]	75.34 + -1.08	94.56 + -0.93	0.8336 + -0.0031	38
FCRN-B [46]	73.06 + -0.62	95.78 + -0.26	0.8241 + -0.0045	32
Unet-like	76.04 + -0.9	94.31 + -0.89	0.8366 + -0.0025	68
Multiscale_4layers	77.59 + -1.05	93.10 + -1.08	0.8411 + -0.0030	82
Multiscale_x2_4layers	78.41 + -1.26	92.61 + -1.37	0.8440 + -0.0021	108
Multiscale_3layers	76.39 + -0.82	93.50 + -0.66	0.8350 + -0.0028	46
Multiscale_x2_3layers	77.00 + -1.11	93.54 + -0.99	0.8392 + -0.0036	72
Multiscale_11_9_3layers	76.00 + -0.87	94.08 + -0.39	0.8351 + -0.0041	46



**Figure 15** The performance of Unet-like model with different input images. (a) Precision of Unet-like model as a function of the number of feature maps. (b) F1-measure of Unet-like model as a function of the number of feature maps

The performance of lymphocyte detection in [2] is almost the same level with our proposed models, but the literature obtained the results through a five steps pipeline (①patch extraction, ②cross-validation creation, ③database creation, ④training of DL classifier and ⑤generating output on test images), which is much more complex than our models. To increase the predictive power of the system, Janowczyk and Madabhushi [2] also artificially resize the lymphocytic images to be  $\times 4$  as large, so that the entire input space, when centered around a lymphocyte, contains lymphocyte pixels, allowing more of the weights in the network to be useful.

As the dataset shown, These images contain breast tissue with lymphocytic infiltration stained with Hematoxylin and Eosin (H and E) and digitized at  $20\times$ , with the spatial resolution of each image being  $100 \times 100$ . The average size of a lymphocyte is approximately 10 pixels in diameter, that's why the FCRN-A and FCRN-B [46] can obtain almost the same good performance as our proposed models with relatively small receptive field.

In lymphocyte detection result, green dots are the true positive (TP), red dots are the false positive (FP), and yellow dots are the false negative (FN) (see Figure 12). The image on the

**Table 4** The performance of Unet-like model with different input images. (before processing and after processing of the label images)

Nb of filter	Before processing of the label images			After processing of the label images		
	Precision (%)	Recall (%)	F1-measure	Precision (%)	Recall (%)	F1-measure
64	75.40 + -1.68	93.3 + -0.66	0.8269 + -0.0095	76.45 + -0.79	93.52 + -1.21	0.8366 + -0.0021
56	75.54 + -2.26	92.91 + -0.89	0.8257 + -0.0125	76.04 + -0.90	94.31 + -0.89	0.8366 + -0.0025
48	75.06 + -1.41	92.75 + -0.75	0.8228 + -0.0075	75.76 + -0.45	94.20 + -0.81	0.8352 + -0.0017
40	74.15 + -2.49	93.26 + -0.4	0.8173 + -0.0179	75.18 + -1.72	94.85 + -1.23	0.8335 + -0.0063
32	71.86 + -6.47	93.58 + -0.49	0.7996 + -0.0532	75.12 + -3.30	94.57 + -2.11	0.8311 + -0.0135
24	68.81 + -10.76	94.08 + -1.28	0.7777 + -0.0906	74.52 + -2.85	95.30 + -1.02	0.8309 + -0.0153
16	65.06 + -17.53	94.04 + -1.64	0.7385 + -0.1652	74.38 + -2.24	95.00 + -1.32	0.8285 + -0.0094
8	43.92 + -10.19	98.08 + -1.22	0.5875 + -0.0894	72.83 + -0.57	95.56 + -0.21	0.8210 + -0.0034

left (Figure 12(a)) has 25 TP/2 FP/1 FN. The false positives are on the edges, about 1 o'clock and 3 o'clock. We can see the false negative is quite small and not very clear, making it hard to detect it without encountering many false positives. The image on the right (Figure 12(b)) one has 60 TP/21 FP/1 FN. We can see the false negative is merge with the neighbor cell and almost become a whole cell that make it hard to detect. The false positives are quite a few, there are mainly two reasons: the first is that almost 10 false positives are on the edges in this image; the second is that the others 11 cells are false positives though their structures do look "lymphocyte-like".

#### 4.4.2 Breast cancer histopathological images

We randomly select 160 images of the entire 285 images as training sets to optimize the detection algorithm, and other images as test sets, are used to evaluate the performance. Here, we only show figures for results from Multiscale\_x2\_4layers model in Figures 13 and 14. The comparison performances of the cell detection of breast cancer histopathological images are shown in Table 3. From Table 3, we can see that the precision and F1-measure of our proposed method (e.g., Multiscale\_x2\_4layers) are 5.35% and 1.99% higher than FCRN-B [46], respectively. It is worth noting that for medical diagnosis, even if the diagnostic accuracy increased by 1% is also significant [25].

The performance of our proposed four layer "Multiscale" series models is slightly better but almost the same level with three layer "Multiscale" series models. The size of the receptive field is more important than being able to provide more detailed information over the receptive field. Our hypothesis is that real difficulty of cell counting lies in the regression for large cell clusters, and a larger receptive field is needed to span them.

In breast cancer histopathological cells detection result, green dots are the true positive (TP), red dots are the false positive (FP), and yellow dots are the false negative (FN). The image on the left (Figure 14(a)) has 27 TP/4 FP/1 FN. We can see the false negative is on the edge of the image and the boundary of a cell is not very clear making it hard to detect it. The structures of the false positives are look "cell-like", especially the one near the center of the image. The image in the middle (Figure 14(b)) one has 41 TP/9 FP/2 FN. False negatives are the same thing with the left image (Figure 14(a)) and almost all the false positives are look "cell-like". The image on the right (Figure 14(c)) one has 23 TP/2 FP/1 FN. We can see the boundary of a false negative is not very clear and the staining is junior making it hard to detect. The false positives are also look "cell-like".

From an overall perspective, we can see Figure 14 (a-c) vary largely with different stains, size of the cell, inhomogeneous intensities, background clutters and overlap issue. We obtain a promising result thanks to the multi-scale features and some tricks (e.g., skip, data augmentation, data preprocessing) in the multi-scale fully convolutional neural networks.

We also explore the interaction between the detection performance and the number of feature maps in fully convolutional neural networks. The result is shown in Table 4 and Figure 15. It can be seen that with the increase of the number of feature maps, the performance will increase and remain stable after the number is 56. In fact, when the number of feature maps is equal to 40, it begins to get a relatively good F1-measure. That is to say, the network is robust when the number of feature maps is in a wide range. At the same time, we can reach a promising and stable detection performance with a small number of feature maps when feeding the label images after preprocessing as inputs, that will reduce the inference time by a large margin and be good for embedded implementation.

## 5 Conclusions

In this paper, we propose several multi-scale fully convolutional neural networks models that are able to robustly and accurately detect the nuclei in pathology and microscopy images. The multi-scale feature maps will enlarge the receptive field and not only can detect the single, small size cells, but also benefit to detecting cells with big size and overlapping state. Compared with previous methods which need carefully designed features and combination strategies, the proposed feature learning method is easy to implement and reproduce. Since the networks are completely convolution, they can predict density maps for the input images with arbitrary sizes. To improve the detection performance of the networks, some data augmentation techniques are employed and regression of the proposed structured proximity space for patches is performed. The proposed algorithm is a general method and can be used for many detection applications. Both qualitative and quantitative experimental results show that our method is superior to several state-of-the-art approaches.

**Acknowledgments** The authors would like to thank Lehmassola et al. [23], Dr. Andrew Janowczyk et al. [2], and Dr. Zhang et al. [52] for publishing the datasets. We are grateful for helpful comments from the anonymous reviewers and the associate editor. This research was supported in part by the National Natural Science Foundation of China (Grant Nos. 21365008 and 61562013), and Natural Science Foundation of Guangxi Province (No. 2017GXNSFDA198025).

## References

1. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süsstrunk, S.: Slic superpixels compared to state-of-the-art superpixel methods. *IEEE Trans Pattern Anal Mach Intell.* **34**(11), 2274 (2012)
2. Andrew, J., Anant, M.: Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. *Journal of Pathology Informatics.* **7**(1), 29 (2016)
3. Arteta, C., Lempitsky, V., Noble, J. A., Zisserman, A.: Learning to detect cells using non-overlapping extremal regions. In: *Medical Image Computing & Computer-assisted Intervention: Miccai International Conference on Medical Image Computing & Computer-assisted Intervention*, 15, 348–356 (2012)
4. Arteta, C., Lempitsky, V., Noble, J.A., Zisserman, A.: Learning to detect cells using non-overlapping extremal regions. In: *International Conference on Medical Image Computing and Computer Assisted Intervention (Lecture Notes in Computer Science). MICCAI*, 15, pp. 348–356 (2012)
5. Arteta, C., Lempitsky, V., Noble, J. A., & Zisserman, A.: Interactive object counting. In: *Proceedings of the European Conference on Computer Vision (ECCV)*, 8691, 504–518 (2014)
6. Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Trans Pattern Anal Mach Intell.* **26**(9), 1124 (2004)
7. Cai, Z., Fan, Q., Feris, R. S., Vasconcelos, N.: A unified multi-scale deep convolutional neural network for fast object detection. In: Leibe B., Matas J., Sebe N., Welling M. (eds.) *European Conference on Computer Vision*. pp.354–370. Springer, Cham (2016)
8. Cireşan, D. C., Giusti, A., Gambardella, L. M., Schmidhuber, J.: Mitosis detection in breast cancer histology images with deep neural networks. In: *International Conference on Medical Image Computing & Computer-assisted Intervention*, 16, 411–418 (2013)
9. Cruz-Roa A.A., Arevalo Ovalle J.E., Madabhushi A., González Osorio F.A.: A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection. In: Mori K., Sakuma I., Sato Y., Barillot C., Navab N. (eds.) *International Conference on Medical Image Computing and Computer-assisted Intervention (MICCAI)*, pp. 403–410. Springer, Berlin (2013)
10. Dan, C.C., Giusti, A., Gambardella, L.M., Schmidhuber, J.: Deep neural networks segment neuronal membranes in electron microscopy images. *Adv. Neural Inf. Proces. Syst.* **25**, 2852–2860 (2012)
11. Dong, B., Shao, L., Costa, M. D., Bandmann, O., Frangi, A. F.: Deep learning for automatic cell detection in wide-field microscopy zebrafish images. In: *Proceedings of IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pp. 772–776. IEEE New York (2015)

12. Dundar, M.M., Badve, S., Bilgin, G., Raykar, V., Jain, R., Sertel, O., et al.: Computerized classification of intraductal breast lesions using histopathological images. *IEEE Trans. Biomed. Eng.* **58**(7), 1977–1984 (2011)
13. Fatakdawala, H., Xu, J., Basavanahally, A., Bhanot, G., Ganesan, S., Feldman, M., et al.: Expectation-maximization-driven geodesic active contour with overlap resolution (emagacor): application to lymphocyte segmentation on breast cancer histopathology. *IEEE Trans. Biomed. Eng.* **57**(7), 1676–1689 (2010)
14. Fiaschi, L., Koethe, U., Nair, R., Hamprecht, F. A.: Learning to count with regression forest and structured labels. In: *Proceedings of 21st International Conference on Pattern Recognition (ICPR)*, pp. 2685–2688. IEEE (2012)
15. Foran, D.J., Yang, L., Chen, W., Hu, J., Goodell, L.A., Reiss, M., et al.: Imageminer: a software system for comparative analysis of tissue microarrays using content-based image retrieval, high-performance computing, and grid technology. *J. Am. Med. Inform. Assoc.* **18**(4), 403–415 (2011)
16. García-Gojo, M.: State of the art and trends for digital pathology. *Stud Health Technol Inform.* **179**, 15–28 (2012)
17. Giusti, A., Dan, C.C., Masci, J., Gambardella, L.M., Schmidhuber, J.: Fast image scanning with deep max-pooling convolutional. *Neural Netw.* 4034–4038 (2013)
18. Guan, B.: Cell segmentation: 50 years down the road. *IEEE Signal Process. Mag.* **29**(5), 140–145 (2012)
19. Hu, R., Zhu, X., Cheng, D., et al.: Graph self-representation method for unsupervised feature selection. *Neurocomputing.* **220**, 130–137 (2017)
20. Ioffe, S., Szegedy, C.: Batch normalization: accelerating deep network training by reducing internal covariate shift. *International Conference on International Conference on Machine Learning.* 448–456 (2015)
21. Khoshdeli, M., Cong R., Parvin, B.: Detection of nuclei in H&E stained sections using convolutional neural networks. In: *2017 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, pp. 105–108. IEEE (2017). <https://doi.org/10.1109/BHI.2017.7897216>
22. Kuse, M., Wang, Y.F., Kalasannavar, V., Khan, M., Rajpoot, N.: Local isotropic phase symmetry measure for detection of beta cells and lymphocytes. *Journal of Pathology Informatics.* **2**(2), S2 (2011)
23. Lehmussola, A., Ruusuvoori, P., Selinummi, J., Huttunen, H., Yli-Harja, O.: Computational framework for simulating fluorescence microscope images with cell populations. *IEEE Trans. Med. Imaging.* **26**(7), 1010–1016 (2007)
24. Lempitsky, V., Zisserman, A.: Learning to count objects in images. In: *Advances in Neural Information Processing Systems (NIPS)*. **43**, 1324–1332 (2010)
25. Li, X., & Plataniotis, K. N.: Color model comparative analysis for breast cancer diagnosis using h and e stained images. In: *SPIE Medical Imaging International Society for Optics and Photonics*, 9420, 94200L–94200L-6 (2015)
26. Liu, F., Yang, L.: A novel cell detection method using deep convolutional neural network and maximum-weight independent set. In: Lu L., Zheng Y., Carneiro G., Yang L. (eds.) *Deep Learning and Convolutional Neural Networks for Medical Image Computing, Advances in Computer Vision and Pattern Recognition*, pp. 349–357. Springer, Cham (2015)
27. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis & Machine Intelligence.* **39**(4), 640–651 (2014)
28. López, C., Lejeune, M., Bosch, R., Korzynska, A., Garcíarojo, M., Salvadó, M.T., et al.: Digital image analysis in breast cancer: an example of an automated methodology and the effects of image compression. *Stud Health Technol Inform.* **179**, 155–171 (2012)
29. Pan, X., Li, L., Yang, H., et al.: Accurate segmentation of nuclei in pathological images via sparse reconstruction and deep convolutional networks. *Neurocomputing.* **229**, 88–99 (2017)
30. Parvin, B., Yang, Q., Han, J., Chang, H., Rydberg, B., Barcelloshoff, M.H.: Iterative voting for inference of structural saliency and characterization of subcellular events. *IEEE transactions on image processing a publication of the IEEE signal processing. Society.* **16**(3), 615–623 (2007)
31. Ren, W., Liu, S., Zhang, H., Pan, J., Cao, X., Yang, M. H.: Single Image Dehazing via Multi-scale Convolutional Neural Networks. In: *European Conference on Computer Vision*. pp. 154–169. Springer (2016)
32. Ronneberger, O., Fischer, P., & Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 9351, 234–241 (2015)
33. Saxe, A. M., McClelland, J. L., Ganguli, S.: Exact solutions to the nonlinear dynamics of learning in deep linear neural networks. *arXiv preprint arXiv:1312.6120.* (2014)
34. Sirinukunwattana, K., Shan, E. A. R., Tsang, Y. W., Snead, D., Cree, I., Rajpoot, N.: A Spatially Constrained Deep Learning Framework for Detection of Epithelial Tumor Nuclei in Cancer Histology Images. In: Wu G., Coupé P., Zhan Y., Munsell B., Rueckert D. (eds.) *Patch-Based Techniques in Medical Imaging. Lecture Notes in Computer Science*, vol. 9467, pp. 154–162. Springer, Cham (2015)

35. Sirinukunwattana, K., Raza, S., Tsang, Y.W., Snead, D., Cree, I., Rajpoot, N.: Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. *IEEE Trans. Med. Imaging.* **35**(5), 1196–1206 (2016)
36. Sommer, C., Hoefler, R., Samwer, M., et al.: A deep learning and novelty detection framework for rapid phenotyping in high-content screening. *Mol. Biol. Cell.* (2017). <https://doi.org/10.1101/134627>
37. 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.* **62**(10), 2421–2433 (2015)
38. Song TH., Sanchez V., EIDaly H., Rajpoot N.: Simultaneous cell detection and classification with an asymmetric deep autoencoder in bone marrow histology images. In: Valdés Hernández M. and González-Castro V. (eds.) *Medical Image Understanding and Analysis. MIUA 2017. Communications in Computer and Information Science*, vol 723. pp. 829–838. Springer, Cham (2017)
39. Song, T., Sanchez, V., Eidaly, H., Rajpoot, N.: Hybrid deep autoencoder with Curvature Gaussian for detection of various types of cells in bone marrow trephine biopsy images. In: 2017 IEEE International Symposium on Biomedical Imaging (ISBI), pp. 1040–1043. IEEE (2017)
40. Su, H., Yin, Z., Kanade, T., & Huh, S.: Phase contrast image restoration via dictionary representation of diffraction patterns. In: *International Conference on Medical Image Computing and Computer-assisted Intervention*, Springer, 615–622 (2012)
41. Su, H., Xing, F., Kong, X., Xie, Y., Zhang, S., Yang, L.: Robust cell detection and segmentation in histopathological images using sparse reconstruction and stacked denoising autoencoders. In: Navab N., Hornegger J., Wells W., Frangi A. (eds.) *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Lecture Notes in Computer Science, vol. 9351, pp. 383–390. Springer, Cham (2015)
42. Wang, H., Cruz-Roa, A., Basavanahally, A., et al.: Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features. *Journal of Medical Imaging*, 1(3), 034003–1–8 (2014)
43. Wei, S., Zhou, M., Yang, F., Yang, C., Tian, J.: Multi-scale Convolutional Neural Networks for Lung Nodule Classification. In: Ourselin S., Alexander D., Westin CF., Cardoso M. (eds.) *Information Processing in Medical Imaging (IPMI)*, vol. 9123, pp. 588–599. Springer, Cham (2015)
44. Xie, Y., Kong, X., Xing, F., Liu, F., Su, H., Yang, L.: Deep voting: a robust approach toward nucleus localization in microscopy images. In: Navab N., Hornegger J., Wells W., Frangi A. (eds.) *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Lecture Notes in Computer Science, vol 9351. pp. 374–382. Springer, Cham (2015)
45. Xie, Y., Xing, F., Kong, X., Su, H., Yang, L.: Beyond Classification: Structured Regression for Robust Cell Detection Using Convolutional Neural Network. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, vol 9351. pp. 358–365. Springer, (2015)
46. Xie, W., Noble, J. A., Zisserman, A.: Microscopy cell counting and detection with fully convolutional regression networks. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, pp. 1–10. Taylor & Francis, Oxfordshire (2016)
47. Xing, F., Yang, L.: Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: a comprehensive review. *IEEE Rev. Biomed. Eng.* **9**, 234 (2016)
48. Xing, F., Xie, Y., Yang, L.: An automatic learning-based framework for robust nucleus segmentation. *IEEE Trans. Med. Imaging.* **35**(2), 550–566 (2015)
49. Xu, J., Xiang, L., Liu, Q., Gilmore, H., Wu, J., Tang, J., et al.: Stacked sparse autoencoder (ssae) for nuclei detection on breast cancer histopathology images. *IEEE Trans. Med. Imaging.* **35**(1), 119 (2016)
50. Yellin, F., Haeffele, B. D., Vidal, R.: Blood cell detection and counting in holographic lens-free imaging by convolutional sparse dictionary learning and coding. *IEEE International Symposium on Biomedical Imaging*, 650–653 (2017)
51. Zhang, S., Metaxas, D.: Large-scale medical image analytics: recent methodologies, applications and future directions. *Med. Image Anal.* **33**, 98–101 (2016)
52. Zhang, X., Liu, W., Dundar, M., Badve, S., Zhang, S.: Towards large-scale histopathological image analysis: hashing-based image retrieval. *IEEE Trans. Med. Imaging.* **34**(2), 496–506 (2015)
53. Zhang, X., Xing, F., Su, H., Yang, L., Zhang, S.: High-throughput histopathological image analysis via robust cell segmentation and hashing. *Med. Image Anal.* **26**(1), 306–315 (2015)
54. Zhu, X., Zhang, L., Huang, Z.: A sparse embedding and least variance encoding approach to hashing. *IEEE Trans. Image Process.* **23**(9), 3737–3750 (2014)
55. Zhu, X., Suk, H.I., Wang, L., et al.: A novel relational regularization feature selection method for joint regression and classification in AD diagnosis. *Med. Image Anal.* **75**(6), 570–577 (2015)
56. Zhu, X., Li, X., Zhang, S.: Block-row sparse multiview multilabel learning for image classification. *IEEE Transactions on Cybernetics.* **46**(2), 450–461 (2016)

57. Zhu, X., Suk, H.I., Lee, S.W., et al.: Subspace regularized sparse multitask learning for multiclass neurodegenerative disease identification. *IEEE Trans. Biomed. Eng.* **63**(3), 607–618 (2016)
58. Zhu, X., Li, X., Zhang, S., et al.: Robust joint graph sparse coding for unsupervised spectral feature selection. *IEEE Transactions on Neural Networks & Learning Systems.* **28**(6), 1263–1275 (2017)