

# 级联稀疏卷积与决策树集成的病理图像细胞核分割方法

宋杰<sup>1,2</sup> 肖亮<sup>2,3</sup> 练智超<sup>2</sup>

**摘要** 数字病理图像分析对于乳腺癌、肾癌等良恶性分级诊断具有重要意义,其中细胞核的形态测量是病理量化分析的关键。然而,由于病理图像背景复杂,细胞核高密度分布、细胞粘连等,个体细胞核精准分割是一个挑战性问题。本文提出一个级联稀疏卷积与决策树集成学习的细胞核分割模型。该模型由稀疏可分离卷积模块和集成决策树学习的正则化回归模块堆叠级联组成,其中:前者采取秩-1张量分解学习机制,可分层抽取细胞核的多尺度方向分布式抽象特征;而后者采取随机采样、树剪枝以及正则化回归机制提升逐像素回归分类能力。相比于现有深度学习模型,该模型无需非线性激活和后向传播计算,参数规模较小,可实现端到端的学习。通过乳腺、前列腺、肾脏、胃和膀胱等多组病理图像的分割实验表明:该模型能够实现复杂数字病理图像中的高密度细胞核的快速个体目标检测和分割,在 Jaccard 相似性系数、F1 分数和平均边缘距离三个指标上均优于目前 CNN2、CNN3 和 U-Net 等深度学习方法,具有较好应用前景。

**关键词** 数字病理, 细胞核分割, 级联稀疏可分离卷积, 集成决策树, 正则化回归, 深层表征学习

**引用格式** 宋杰, 肖亮, 练智超. 级联稀疏卷积与决策树集成的病理图像细胞核分割方法. 自动化学报, 2021, 47(2): 378–390

**DOI** 10.16383/j.aas.c190672

## Cascade Sparse Convolution and Decision Tree Ensemble Model for Nuclear Segmentation in Pathology Images

SONG Jie<sup>1,2</sup> XIAO Liang<sup>2,3</sup> LIAN Zhi-Chao<sup>2</sup>

**Abstract** The quantitative analysis of digital pathology images plays a significant role in the diagnosis of benign and malignant diseases such as breast and prostate cancer, in which nuclear morphology measurement serve as a basis of quantitative analyses. However, due to the complex background of pathology images, dense distributions of nuclei, and nucleus adhesions, accurate segmentation of individual nuclei remains a challenging problem. In this paper, we propose a new method to automatically segment nuclei from digital pathology images with cascade sparse convolution and decision tree ensemble (CscDTE) model. In particular, the sparse separable convolution learning module and the decision tree ensemble learning module are stacked in a cascaded manner to form the CscDTE model. The former adopts rank-one tensor decomposition learning mechanism that can extract multiscale and multi-directional distributed abstract features; while the latter employs random sampling, pruning, and regularized regression mechanism to boost per-pixel regression and/or classification performance. Compared with the popular deep neural networks, the proposed CscDTE model does not require nonlinear activation and backpropagation computation, and depends on fewer parameters. Our CscDTE model is trained in a layer-wise manner that can achieve an end-to-end pixelwise learning and fast nuclear detection and segmentation in high-throughput imagery. We demonstrated the superiority of our method in terms of Jaccard index, F1 score, and average boundary distance by evaluating it on the multi-disease state and multi-organ dataset where consistently higher performance was obtained as compared to convolutional neural networks and fully convolutional networks.

**Key words** Digital pathology, nuclear segmentation, cascade sparse separable convolution, decision tree ensembles, regularized regression, deep representation learning

**Citation** Song Jie, Xiao Liang, Lian Zhi-Chao. Cascade sparse convolution and decision tree ensemble model for nuclear segmentation in pathology images. *Acta Automatica Sinica*, 2021, 47(2): 378–390

收稿日期 2019-09-23 录用日期 2020-03-12  
Manuscript received September 23, 2019; accepted March 12, 2020

国家自然科学基金 (61871226, 62001247), 国家重点研发计划 (2016YFF0103604), 中央高校基本科研专项资金 (30918011104), 江苏省社会发展重点研发计划 (BE2018727), 南京邮电大学引进人才科研启动基金 (NY219152), 江苏省高等学校自然科学研究面上项目 (20KJB520005) 资助

Supported by National Natural Science Foundation of China (61871226, 62001247), National Major Research Plan of China (2016YFF0103604), Fundamental Research Funds for the Central Universities (30918011104), Jiangsu Provincial Social Development Project (BE2018727), NUPTSF (NY219152), Natural Science

Foundation for Colleges and Universities in Jiangsu Province (20KJB520005)

本文责任编辑 张道强

Recommended by Associate Editor ZHANG Dao-Qiang

1. 南京邮电大学自动化学院、人工智能学院 南京 210023 2. 南京理工大学计算机科学与工程学院 南京 210094 3. 高维信息智能感知与系统教育部重点实验室 南京 210094

1. College of Automation & College of Artificial Intelligence, Nanjing University of Posts and Telecommunications, Nanjing 210023 2. School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094 3. Key Laboratory of Intelligent Perception and Systems for High-Dimensional Information of Ministry of Education, Nanjing 210094

复杂背景病理图像个体细胞分割是在成千上万个细胞汇集的图像中逐个分割出具有相对完整边界的细胞, 为后续细胞形态计算和病理特征提供定量分析. 传统的病理切片分割方法, 需要经过专门训练的病理医师在图像中逐个寻找感兴趣区域 (Regions of interest, RoI), 而后根据专业知识分析诊断. 一张病理切片通常包含成百上千个细胞或细胞核, 这给医师带来很大的工作负担, 疲劳阅片现象时有发生<sup>[1]</sup>. 虽然目前已有很多病理细胞图像处理方法, 但这些方法大多只关注于特定类型或单一器官的细胞或细胞核分割. 因此, 临床和医学研究迫切需要能够进行多个器官和疾病状态的细胞核病理图像高精度分割方法<sup>[2]</sup>. 然而, 如图 1 所示, 个体分割具有如下挑战: 一方面, 对于病理状态 (例如增生或某种癌症亚型) 图像而言, 由于细胞核增大, 并呈现染质浓集贴边, 即核内染质较浅, 而边缘附近染质较深; 而着色较深的核仁也大量出现在核内. 另一方面, 由于病理图像中往往细胞密度高、细胞间出现重叠和成团等突出问题, 加剧了个体细胞分割的难度. 目前方法体系主要分两类: 传统基于人工特征的图像分割方法和基于表示特征学习的图像分割方法.

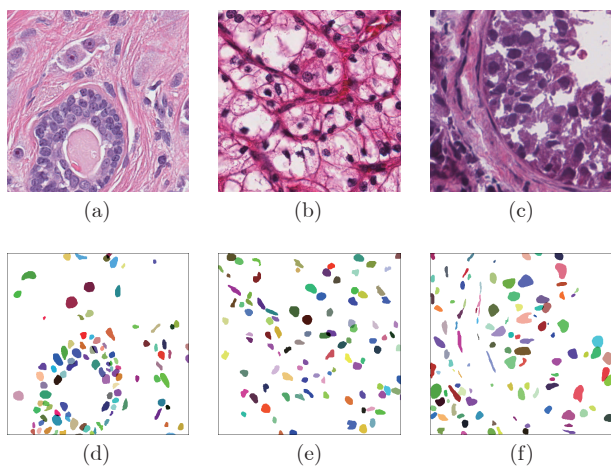


图 1 病理细胞核图像分割的挑战与人工分割结果

Fig.1 Challenges in nuclear segmentation and associated ground truth

传统图像分割方法包括: 水平集方法、图论方法和分水岭方法等. 水平集算法是目前比较流行的分割方法. 其基本模型包括两类: 基于边缘的水平集模型<sup>[3]</sup>和基于区域的水平集模型<sup>[4]</sup>. 近年来, 基于后者, 很多改进的方法<sup>[5]</sup>被提出来广泛应用于病理图像的细胞分割中. 在现有方法中, 基于图论的分割方法是病理图像分析中应用最为广泛的技术之一. 最常用算法是最小割算法<sup>[6]</sup>, 除了主流的图割方

法, 其他类型的图模型也已经应用到个体细胞核的分割, 例如属性关系图模型<sup>[7]</sup>. 另一类基于图像分析的代表性方法是分水岭算法, 分水岭算法通常期望目标细胞核内灰度分布均匀, 并且目标与背景具有明显灰度或颜色差异. 图像中的噪声、目标内部灰度变化, 都会产生过度分割的现象. 为此, 人们提出若干标记控制的分水岭及其变种算法<sup>[8-10]</sup>. 传统图像分割方法为了得到精确的个体分割结果, 通常需要分析 RoI 特有属性来设计额外的后处理步骤, 从而导致算法可迁移性较差.

虽然深度学习方法能较好地处理病理图像细胞核表现多样性的分割问题, 但由于网络架构、网络复杂性以及超参数的影响, 使其分割性能受到限制. 针对病理图像细胞核表现多样性的分割问题及深度神经网络的局限性, 国内外研究者提出了系列具有较小参数规模的快速浅层分割学习模型. 与现有深度学习模型相比, 浅层学习模型无需非线性激活和后向传播计算, 且学习模型参数规模较小. 传统浅层分割学习方法<sup>[11-12]</sup>为了得到密集预测结果, 通常使用像素周围的一个图像块作为模型的输入用于训练和预测. 由于只能提取一些局部的特征, 从而导致分割的性能受到限制. 而基于卷积操作的浅层学习模型<sup>[13-18]</sup>则是从抽象的特征中恢复出每个像素所属的类别, 从而可以实现比传统模型更加精确的分割结果. 然而, 这类浅层分割学习模型没有充分考虑像素间的局部依赖关系, 因此对图像细节特征快速捕获和紧致表达能力有待加强. 此外, 在传统集成决策树学习算法中最先学习得到的一两棵树对预测结果的影响最为显著, 这使得整个模块对这些树所做出的决策过于敏感, 容易产生过拟合现象.

为了解决这些问题, 本文的主要贡献是:

1) 提出了级联稀疏卷积与决策树集成 (Cascade sparse convolution and decision tree ensemble, CscDTE) 学习模型, 该模型没有非线性激活和后向传播计算, 且学习模型参数规模较小, 其特征学习过程具有一种替代深度神经网络的新型学习机制;

2) 采取多层稀疏可分离卷积特征学习捕获图像上下文特征; 采取秩-1 张量分解<sup>[19]</sup>的可分离卷积加速特征学习过程;

3) 建立集成决策树学习的正则化回归模型, 采取局部二阶近似逼近优化决策树, 提高分类回归泛化性能;

4) 在乳腺、前列腺、肾脏、胃和膀胱等多组病理图像的分割实验表明该模型优于目前 CNN2、CNN3 和 U-Net 等深度学习方法, 对于病理图像分割具有较好应用前景.

## 1 相关工作

### 1.1 浅层特征学习方法

浅层特征学习分割方法中代表性方法包括稀疏编码方法<sup>[11, 13]</sup>以及多层稀疏卷积特征学习方法<sup>[15-18]</sup>. 相比标准的卷积滤波器组来说, 稀疏可分离卷积滤波器组能够在不影响其性能的情况下大大降低计算复杂度. 例如 Sironi 等<sup>[13-15]</sup>通过学习一组可分离卷积滤波器有效提取图像中曲线状奇异边缘结构. 具体地, 稀疏卷积滤波器学习方法是利用一组样本, 通过利用稀疏编码方法<sup>[20-21]</sup>学习  $J$  个卷积滤波器  $\mathbf{f} = \{\mathbf{f}_j\}_{j=1}^J :=$

$$\arg \min_{\{\mathbf{f}_j\}, \{\mathbf{a}_j^i\}} \sum_i \left( \left\| I^i - \sum_{j=1}^J \mathbf{f}_j * \mathbf{a}_j^i \right\|_2^2 + \beta \cdot \sum_{j=1}^J \|\mathbf{a}_j^i\|_1 \right) \quad (1)$$

其中,  $\mathbf{f}_j \in \mathbf{R}^{d \times d}$ ,  $\mathbf{a}_j^i$  是输入图像  $I^i$  的稀疏编码,  $\beta$  是正则化参数. 模型 (1) 可通过交替方向法<sup>[22]</sup> 迭代优化更新  $\mathbf{a}_j^i$  和  $\mathbf{f}_j$  进行求解. 当得到一组卷积滤波器  $\{\mathbf{f}_j\}_{j=1}^J$  之后, 则可以采取  $\{\mathbf{f}_j\}_{j=1}^J$  对图像  $I$  进行卷积生成一组卷积特征图  $\mathbf{m} = \{\mathbf{m}_j\}_{j=1}^J$ , 其中,  $\mathbf{m}_j = \mathbf{f}_j * I$ , 符号  $*$  表示卷积运算. 受稀疏编码方法的启发, Song 等<sup>[16]</sup> 通过在线学习稀疏卷积特征并在模型更深层引入上下文信息实现了鲁棒的显微颗粒目标联立检测和分割. 进一步, Song 等<sup>[17]</sup> 通过结合分层稀疏特征学习和人工特征, 设计了 Boosting 判别

概率二叉决策树分类方法.

### 1.2 深层特征学习方法

随着深度学习的兴起, 研究者提出了系列深层特征表示学习的病理图像分割方法<sup>[23-27]</sup>. 代表性网络包括卷积神经网络 (Convolutional neural networks, CNN)<sup>[24-25]</sup> 和全卷积网络 (Fully convolutional network, FCN)<sup>[26-27]</sup>. 深度学习方法为解决细胞核表观多样性的图像分割问题提供了有效途径. 通常, CNN 网络在卷积层之后会接上若干个全连接层, 将卷积层产生的特征图映射成一个固定长度的特征向量, 以实现图像级的分割任务<sup>[28]</sup> (见图 2). 例如, Xing 等<sup>[24]</sup> 针对组织病理细胞核图像, 构建了两个卷积—池化层对和两个全连接层构成的 CNN2 模型, 实现了端到端的模型训练, 进而对测试图像的细胞核进行分割. 在此基础上, Kumar 等<sup>[25]</sup> 提出了更深的 CNN3 模型, 实现了更广泛病理图像的分割. 与 CNN 不同, FCN 可以接受任意尺寸的输入图像, 采用反卷积层对最后卷积层的特征图进行上采样, 使它恢复到输入图像相同的尺寸, 从而可实现逐像素预测, 并可保留原始输入图像的空间信息, 最后在上采样的特征图上进行逐像素分类<sup>[29]</sup>, 完成像素级的分割. 例如, Zhang 等<sup>[26]</sup> 通过结合 FCN 和基于图论的方法, 提出了一种宫颈癌细胞核分割新方法. 最近, Ronneberger 等训练了一种特殊的 U-Net 网络结构<sup>[27]</sup> (见图 3), 通过在收缩路径上捕获全局特征和在扩展路径上实现精确定位, 较好解决了

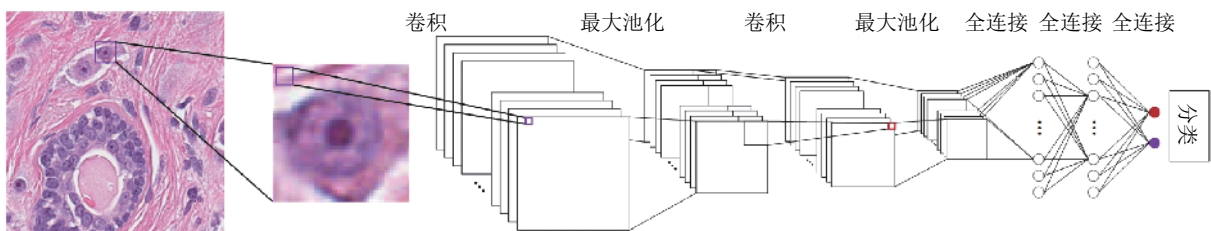


图 2 用于病理图像分割的 CNN 体系结构

Fig. 2 The CNN-style architecture for pathology image segmentation

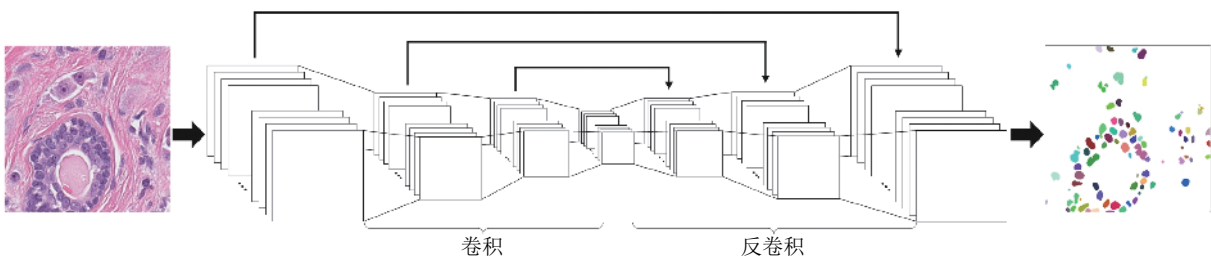


图 3 用于病理图像分割的 U-Net 体系结构

Fig. 3 The U-Net-style architecture for pathology image segmentation

复杂神经元结构的分割问题. 然而, 深度学习方法需要较大规模的训练数据集以经验拟合深层网络参数.

目前, Zhou 等提出了一种多粒度级联森林方法, 称为 gcForest<sup>[12]</sup>. 这是一种采取非神经网络, 而采取决策树集成的方法, 性能较之深度神经网络有很强的竞争力. 深度神经网络需要花大力气调参, 相比之下, gcForest 方法更容易训练得多. 实际上, 在几乎完全一样的超参数设置下, gcForest 在处理不同领域的不同数据时, 也能达到极佳的性能.

## 2 本文方法

针对病理图像细胞核表现多样性的分割问题以及深度神经网络的局限性, 本文提出了一种复杂病理细胞核分割的级联稀疏卷积与决策树集成 (CscDTE) 学习模型, 如图 4 所示. 假设  $(m, y)$  表示训练样本集, 其中,  $m = \{f_{1,j} * I\}_{j=1}^{J_1}$  是从输入图像  $I$  提取的特征图集,  $f_1$  是第一层学习得到的卷积滤波器组,  $y$  是尝试预测的目标. CscDTE 框架旨在渐进地提供一种非线性模型  $\varphi_h(m)$  拟合数据样本, 其中,  $h$  表示模型参数.

### 2.1 稀疏可分离卷积特征学习

一般而言,  $f_j$  是满秩的, 当卷积较大图像时, 计算耗时. 如果将这些卷积滤波器  $\{f_j\}_{j=1}^J$  堆叠成一个三维张量  $Z \in \mathbf{R}^{d \times d \times J}$ , 其中,  $Z$  的第  $j$  个切片对应第  $j$  个卷积滤波器  $f_j$ , 可发现该张量存在大量冗余特征. 为此, 为了学习可分离核, 本文进一步引入低秩张量分解技术<sup>[19]</sup>, 利用一组秩为 1 (秩-1) 张量的线性组合来近似  $Z$ , 从而实现可分离卷积核, 加速特征计算. 该方法基于如下假设: 一个秩为  $R$  的滤波器  $f_j$  总能由一组秩-1 可分离核:  $\{k_r\}_{r=1}^R$  线

性表示:  $f_j = \sum_{r=1}^R c_j^r k_r$ , 其中,  $c_j^r$  为标量权重. 此时所有的  $f_j$  共享同一组秩-1 可分离核  $\{k_r\}_{r=1}^R$ , 只有权重系数  $c_j^r$  因不同的  $f_j$  而改变, 且  $R < J$ . 因此, 为了得到上式中的分解形式, 本文利用低秩张量分解求解如下优化问题来学习  $\{k_r\}_{r=1}^R$ :

$$\min_{\{a^r, b^r, c^r\}_r} \left\| Z - \sum_{r=1}^R a^r \circ b^r \circ c^r \right\|_2^2 \quad (2)$$

其中,  $a^r$  和  $b^r$  均是长度为  $d$  的向量,  $c^r$  是一个长度为  $J$  的向量, 符号  $\circ$  表示张量积. 式 (2) 的求解可通过使用共轭梯度下降法<sup>[30]</sup> 来实现. 如图 5 所示, 可分离核  $k_r$  可由典范多元分解 (Canonical polyadic decomposition, CPD) 成分  $a^r$  和  $b^r$  给出, 即  $k_r = a^r \circ b^r$ . 用于重构卷积滤波器  $f_j$  的权重  $c_j^r$  则由向量  $c^r$  的第  $j$  个成分给出. 为了捕获不同类型的图像特征, 本文相应学习多个可分离核组. 归于可分离卷积核的学习机制, 每个  $k_r$  都能分解成一个水平滤波器和一个垂直滤波器. 因此, 为了获得  $m$ , 首先使用  $\{k_r\}_{r=1}^R$  对图像  $I$  进行快速卷积, 生成一组可分离特征图. 然后, 针对每个像素位置, 将这些  $R$  维向量映射到一个  $J$  维向量. 这等价于在可分离特征图上应用  $J$  个空间维度为  $1 \times 1$  的滤波器<sup>[31]</sup>.

这样, 基于秩-1 张量分解学习机制, 本文实现了病理细胞核的多尺度方向分布式抽象特征的分层抽取. 模型的每层始于稀疏可分离卷积核组的训练及相应卷积特征图的计算. 然后以卷积特征作为样本输入, 通过集成决策树学习的正则化回归模块生成得分图预测结果, 如图 4 所示.

### 2.2 集成决策树学习的正则化回归模型

令  $f_i = f(x_i, I) \in \mathbf{R}^J$  表示输入图像  $I$  中像素点

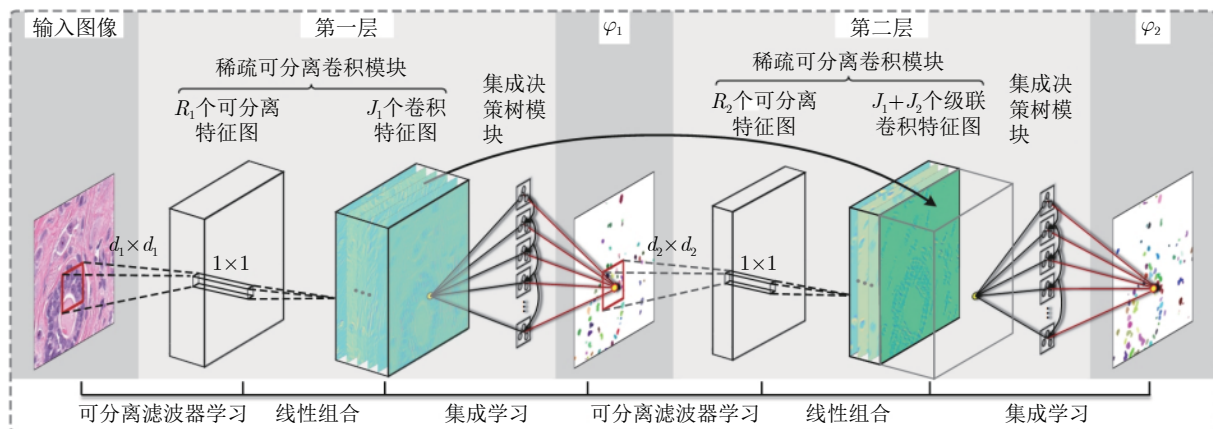


图 4 两层 CscDTE 学习模型示例

Fig. 4 Example of two-layer CscDTE architecture

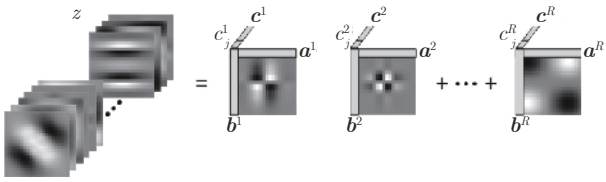


图 5 基于张量分解技术学习一组秩-1 可分离核  
Fig.5 Tensor decomposition for learning rank-1 separable kernels

$\mathbf{x}_i$  对应的特征向量, 则从给定的样本集  $(\mathbf{m}, \mathbf{y})$  中随机采样  $N$  对训练样本集  $\{(f_i, y_i)\}_{i=1}^N, y_i \in \mathbf{R}$ . CscDTE 框架中的集成决策树学习模块的目标是学习  $M$  棵既有预测能力又相对简单的决策树 (见图 6):  $\{h_l^m(\cdot)\}_{m=1}^M: \mathbf{R}^{J_l} \mapsto \mathbf{R}$ , 并使用这  $M$  棵决策树的组合来预测每层的输出:

$$\varphi_l(\boldsymbol{\alpha}_l, \mathbf{h}_l(f(\mathbf{x}, I))) = \sum_{m=1}^M \alpha_l^m h_l^m(f(\mathbf{x}, I)) \quad (3)$$

其中, 第  $l$  层第  $m$  棵决策树  $h_l^m(\cdot)$  定义如下:

$$h_l^m(f(\mathbf{x}, I)) = \mathbf{w}_{s_l^m}^m(f(\mathbf{x}, I)) \quad (4)$$

其中,  $s_l^m(\cdot): \mathbf{R}^{J_l} \mapsto \{1, 2, \dots, C_l^m\}$  为决策树结构, 表示将一个样本映射到相应的叶节点索引.  $C_l^m = |h_l^m(\cdot)|$  为当前决策树中叶节点的数目.  $\mathbf{w}_{s_l^m}^m(\cdot) = [w_1, \dots, w_{C_l^m}]^T \in \mathbf{R}^{C_l^m}$  为所有叶节点的响应值向量. 这样, 每棵决策树对应一个独立的结构和一组叶节点响应值. 那么, 给定一个输入样本, 就可以利用不同决策树的结构将其划分到叶节点中, 然后对相应叶节点的响应值进行累加求和来计算最终输出. 为

了学习模块中的  $M$  棵决策树, 本文以贪婪加性方式<sup>[32-33]</sup> 定义如下代价函数:

$$\mathcal{L}(\mathbf{y}, \varphi_l^m) = \left\| \mathbf{y} - \left( \varphi_l^{m-1} + \mathbf{w}_{s_l^m(f)} \right) \right\|_2^2 + \frac{1}{2} \lambda \left\| \mathbf{w}_{s_l^m(f)} \right\|_2^2 + \gamma \left| \mathbf{w}_{s_l^m(f)} \right| \quad (5)$$

其中, 第一项为数据拟合项, 该项中  $\mathbf{y} = [y_1, \dots, y_N]^T$ ; 第二项对应  $\ell_2$  正则化, 第三项  $\left| \mathbf{w}_{s_l^m(f)} \right|$  表示决策树  $h_l^m(\cdot)$  上叶节点的个数, 第二项和第三项将起到决策树剪枝作用<sup>[32]</sup>, 用于控制决策树的复杂度来避免过拟合, 参数  $\lambda$  为正则化系数, 参数  $\gamma$  表示引入额外叶节点复杂度成本. 直观地, 最小化上述代价函数有利于算法选择既有预测能力又相对简单的决策树.

### 2.3 局部二阶近似优化

为降低计算复杂度, 本文采用代价函数的一个局部二阶近似来对式 (5) 进行优化. 集成决策树学习的正则化回归算法通过计算损失函数的二阶导数, 有利于梯度下降的更快更准<sup>[34]</sup>. 这样, 对于一个固定的决策树结构  $s_l^m(\cdot)$ , 则可通过  $\mathcal{L}$  的局部二阶近似计算第  $c$  个叶节点的最优响应值, 即  $\hat{w}_c = G_{l,c}^m / \left( H_{l,c}^m + \frac{\lambda}{2} \right)$ , 其中,  $G_{l,c}^m$  和  $H_{l,c}^m$  分别对应一阶梯度和海森 (Hessian) 值. 一旦得到  $\hat{w}_c$ , 就可以通过计算最优  $\hat{\mathcal{L}}$  来选择最优决策树结构  $s_l^m(\cdot)$ . 算法终止条件为: 可预先设置决策树的最大深度值, 使得当决策树的生长到达该深度时, 算法停止继续分裂. 重复上述加性训练过程直至最后一棵决策树.

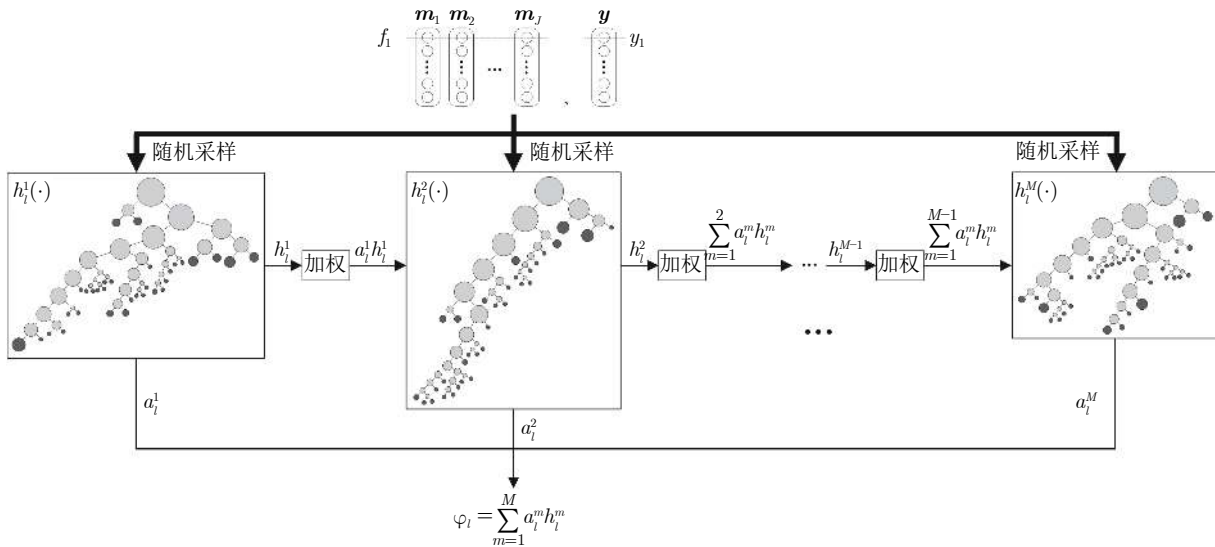


图 6 CscDTE 框架中的集成决策树学习模块的训练过程

Fig.6 Flowchart of the training procedure for the decision tree ensemble learning module of CscDTE framework

需要注意的是, 每次迭代学习得到最优决策树  $\hat{h}_l^m(\cdot)$  之后, 可利用线搜索算法<sup>[16]</sup> 计算其相应的贡献权重  $\alpha_l^m$ . 然后, 将当前模块更新为  $\varphi_l^m = \varphi_l^{m-1} + \alpha_l^m \cdot \hat{h}_l^m$ . 当然, 我们也可以取平均来赋值每个权重  $\alpha_l^m$  的值, 即  $\alpha_l^m = 1/M$ .

在测试阶段, 基于所学的决策树结构  $s_l^m(\cdot)$ , 每个新的特征样本被独立地推送到各棵训练好的决策树中. 当该特征样本到达叶节点时, 通过累加求和不同决策树最优  $\hat{w}_c$  值来确定此测试样本的预测响应.

## 2.4 CscDTE 学习模型

所构建的 CscDTE 框架由  $L-1$  个隐层和 1 个最终输出层构成, 且每层包含一个稀疏可分离卷积模块和一个集成决策树学习的正则化回归模块. 不妨将第  $l$  层的输出定义为  $\varphi_l = [\varphi_{l,1}, \varphi_{l,2}, \dots, \varphi_{l,n}]^T$ , 其中,  $n$  为图像的像素个数. CscDTE 模型将模块对进行堆叠, 这样上层集成决策树学习模块的得分图输出为下层稀疏可分离卷积模块的输入. 对于给定参数集合  $\mathbf{h}_l$ , CscDTE 模型的前向生成过程为:

$$\begin{aligned} \varphi_1 \left( \alpha_1, \mathbf{h}_1 \left( \{f_{1,j} * I\}_{j=1}^{J_1} \right) \right) &= \\ \alpha_1^T \mathbf{h}_1(\mathbf{m}) &= \sum_{m=1}^M \alpha_1^m h_1^m(\mathbf{m}) \\ \varphi_2 \left( \alpha_2, \mathbf{h}_2 \left( \mathbf{m}, \{f_{2,j} * \varphi_1\}_{j=1}^{J_2} \right) \right) &= \\ \alpha_2^T \mathbf{h}_2(\mathbf{m}, \mathbf{n}_1) &= \sum_{m=1}^M \alpha_2^m h_2^m(\mathbf{m}, \mathbf{n}_1) \\ &\dots \\ \varphi_L \left( \alpha_L, \mathbf{h}_L \left( \mathbf{m}, \{f_{L,j} * \varphi_{L-1}\}_{j=1}^{J_L} \right) \right) &= \\ \alpha_L^T \mathbf{h}_L(\mathbf{m}, \mathbf{n}_{L-1}) &= \sum_{m=1}^M \alpha_L^m h_L^m(\mathbf{m}, \mathbf{n}_{L-1}) \end{aligned} \quad (6)$$

其中,  $\mathbf{n}_l$  表示从第  $l$  层输出结果  $\varphi_l$  提取的特征图集, 且作为第  $l+1$  层集成决策树学习模块  $\varphi_{l+1}(\cdot)$  的输入.

如图 7 所示, 由于低级图像表现特征占主导作用, 较浅层对背景杂斑和细胞核内不均匀的染质较为敏感. 因此, 本文进一步引入高级上下文特征<sup>[35]</sup>. 具体地, 从得分图中训练一组新的稀疏可分离卷积核并获取新的上下文卷积特征图集  $\mathbf{n}_l$ . 为了实现更好的预测, 采用级联卷积特征  $[\mathbf{m}, \mathbf{n}_l]$  代替  $\mathbf{n}_l$  作为新层集成决策树学习模块的样本输入. 重复这一过程  $L$  次, 最终获得一个最大程度近似目标区域的输出 (见图 7(c)). 同时, 由于卷积特征的级联将产生分布式冗余特征, 为此, 本文采取了如下两个策略: 1) 在每层训练之前, 从输入样本中随机采样新的像素

位置以构建新的训练集; 2) 随机采样这些像素对应的卷积特征.

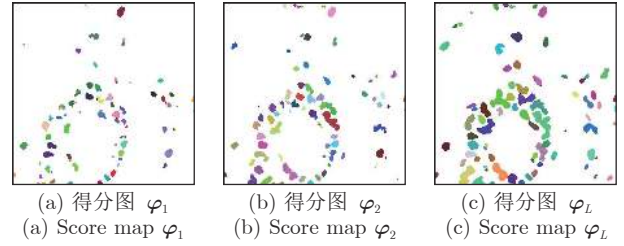


图 7 基于本文提出的 CscDTE 模型的分割改进  
Fig. 7 Improvement obtained by our CscDTE model

## 3 实验结果及分析

本节首先介绍用于测试分割方法的病理图像数据集, 并在此基础上, 给出所提出的 CscDTE 模型的最优超参数设置. 然后, 我们描述评估指标, 并对比标记控制的分水岭分割方法<sup>[10]</sup>、CNN2<sup>[24]</sup>、CNN3<sup>[25]</sup>、U-Net 技术<sup>[27]</sup> 以及 CscDTE 模型的分割性能及参数规模.

### 3.1 数据集与最优超参数设置

癌症基因组图谱 (The cancer genome atlas, TCGA) 是被广泛用于细胞核分割的病理数据集, 因为它覆盖了多家医院、多位病人、不同器官以及疾病状态信息. 为了最大化细胞核表现多样性, 本文使用 Kumar 等采集并标注的 TCGA 全切片图像 (Whole slide images, WSIs) 集<sup>[25]</sup> 作为算法的验证和比较. 其中, 每张图像大小为  $1000 \times 1000$ . 为了便于算法的验证和比较, 采用 Kumar 等的方法将整个数据集拆分成三个部分. 第一部分用于训练和验证, 对应 12 位病人, 3 个器官, 包括 4 张乳腺病理细胞核图像、4 张前列腺病理细胞核以及 4 张肾脏病理细胞核图像, 总共有 11 460 个细胞核. 因此, 不管是基于图像块级分类的分割方法还是基于像素级分类的分割方法, 即使不对数据进行扩充, 这个数目也足以使它们训练出相应的学习系统, 并产生预测的分割结果. 第二部分用于相同器官的测试, 对应不同病人, 相同三个器官, 图像总数为 6 幅. 大多数已有的细胞核分割学习算法只局限于同一器官的训练与测试, 因此, 这一部分能够对模型的泛化能力进行有效的评估. 第三部分用于更具挑战性的不同器官的测试, 对应 2 个器官, 包括 2 张胃病理细胞核图像和 2 张膀胱病理细胞核图像.

为了进一步验证本文方法的有效性, 本文也采用肾细胞癌 (Kidney renal cell carcinoma, KIRC)

病理数据集<sup>[36]</sup>来训练相应 CscDTE 模型, 并分析分割性能. 来自 TCGA 数据门户网站的 KIRC 病理数据集包含了不同类型的 WSIs, 涵盖了一定范围的 KIRC 病理分级. 实验中, 18 幅图像用于训练, 20 幅图像用于测试, 且平均每幅图像具有 96 个细胞核, 大小为  $400 \times 400$ .

基于验证集中的样本, 采用留一交叉验证方法<sup>[37]</sup>来确定 CscDTE 模型关键超参数的最优值, 包括训练样本数量、集成决策树学习模块尺寸、CscDTE 模型尺寸以及稀疏可分离卷积滤波器尺寸和数目. 表 1 列出了在 TCGA WSIs 病理数据集上所提出 CscDTE 模型的最优超参数值. 在训练过程中, 随机采样四分之一的像素位置用于集成决策树学习模块的训练, 在计算每棵树的响应之后, 进而使用所有的位置样本来学习式 (3) 中的权重  $\alpha_i^m$ .

### 3.2 分割比较

为了与深度学习技术进行对比, 本文研究了三种不同的网络体系: CNN2<sup>[24]</sup>、CNN3<sup>[25]</sup> 和 U-Net<sup>[27]</sup>, 并在 NVIDIA GeForce GTX 1080 Ti<sup>®</sup> 图形处理单元上使用 Tensorflow 框架训练所有的网络模型.

表 2 和表 3 分别给出了用于复杂 TCGA WSIs 病理图像分割的 CNN2 和 CNN3 体系结构, 包括连续的卷积—池化层对以及若干个全连接层. 最后一层是 Softmax 层, 旨在基于学习到的特征预测输入图像块属于每一类的概率. Xing 等<sup>[24]</sup> 采取两个卷积层和两个全连接层, 而 Kumar 等<sup>[25]</sup> 则采取三个卷积层和两个全连接层, 因为加入更多的层并不会显著提高网络的分割精确度, 反而会会增加计算时间. 为了获得最好的分割性能, 经验设置其他的超参数值, 包括卷积层中的滤波器尺寸和数量、隐层节点数目、以及输入输出尺寸等. 图 3 给出了用于复杂 TCGA WSIs 病理图像分割的 U-Net 体系结构<sup>[27]</sup>. 该体系结构由编码器和解码器组成. 编码器包含由修正线性单元 (Rectified linear unit, ReLU) 实现的填充卷积操作, 共两层. 每两个卷积层之后, 有一个步长为 2 的最大池化操作. 每次最大池化层下采样之后, 特征通道数加倍. 在解码器中, 每两个卷积层之前, 有一个  $2 \times 2$  的上采样操作, 且其输出与来自相应编码器部分的特征相结合. 最后两层分别是  $1 \times 1$  卷积操作和像素级 Softmax 输出, 其中,  $1 \times 1$  卷积层用于将特征图的通道数降至

所需的类别数. 在本文的 U-Net 网络实现中, 代替数据扩充, 只是随机提取了图像块, 不过仍然实现了令人满意的分割性能. 本文使用 Tensorflow 深度学习框架在 TCGA WSIs 病理数据集的所有样本上对 CNN2、CNN3 和 U-Net 网络均训练了 95 次 (Epochs). 针对每个网络, 分别从 12 幅训练图像和 3 幅验证图像中提取了 158 400 个图像块和 32 000 个图像块用于训练和验证, 其中包括相等数目的以细胞核像素为中心的样本块和以非细胞核像素为中心的负样本块. 另外, 基于 TCGA WSIs 病理数据集, 比较实验分析了标记控制的分水岭分割方法<sup>[10]</sup>.

图 8 和图 9 显示了在 5 个器官和疾病状态下, 病理细胞核图像的分割算法对比. 图 8 和图 9 中第 (a) ~ (g) 列分别是原始图像、人工分割结果、CscDTE 模型分割结果、U-Net 网络分割结果、CNN3 网络分割结果、CNN2 网络分割结果和标记控制的分水岭分割结果. 需要注意的是, 考虑到不同类型病理图像颜色的异质性, 在分割模型训练之前, 将原始图像从 RGB 颜色空间转换到 CIE LUV 均匀颜色空间并作线性归一化. 可以看出, 所提出 CscDTE 模型能有效应对细胞核大小、形状和方向的变化, 并最大程度地检测正确细胞核的数目. 另外, 针对存在染色质稀疏、重叠和复杂背景杂斑的情形, 例如, 部分前列腺病理图像和膀胱病理图像, CscDTE 模型也表现出了良好的鲁棒性 (见图 10). 相比之下, 虽然 U-Net 通过融入上下文信息能较好地避免复杂背景杂斑的干扰, 但由于网络深度的影响, 仍然对染色质稀疏的细胞核内细节不够敏感. CNN2 利用各向同性区域生长从网络输出中获取细胞核种子标记, 因而易受细胞核形状的影响, 尤其是带分割目标存在重叠和成团的情况, 产生欠分割的现象. CNN3 使用各向异性区域生长代替 CNN2 中的各向同性区域生长, 虽然增加了对重叠、成团分割的鲁棒性, 但其受限于图像块级的分类. 由于只能提取局部特征, 导致其分割精度相比于 CscDTE 模型较低. 作为传统图像分析的方法, 分水岭在分割灰度均匀的孤立细胞核时, 表现得相当优异, 例如图 8 中的肾脏细胞核分割. 然而, 对于 Kumar 病理数据中复杂背景图像, 其分割性能遭受大幅度的下降. 提出的 CscDTE 模型在训练和预测时间上也优于基于 CNN 和 FCN 的体系结构, 训练所有的细胞核仅花费了两个多小时左右.

表 1 提出的 CscDTE 模型的最优参数值. 像素位置样本总数为 800 000

Table 1 The optimal hyper-parameter values of our CscDTE model. The total number of pixel samples is 800 000

数据集	$N$	$L$	$d^2(R)$ ( $1^{\text{st}}$ layer; $2^{\text{nd}} \sim L^{\text{th}}$ layers)	$M$
TCGA WSIs	200 000	5	$11^2(25), 17^2(8), 21^2(72), 29^2(8), 43^2(44); 11^2(25), 21^2(36), 43^2(36)$	50

表 2 用于病理图像分割的 CNN2 体系结构  
Table 2 The CNN2 architecture for pathology image segmentation

网络层	滤波器尺寸	激活	输出尺寸	随机丢弃率
输入	-	-	55 × 55 × 3	-
卷积 1	6 × 6	ReLU	50 × 50 × 48	0.1
池化 1	2 × 2	Max	25 × 25 × 48	-
卷积 2	4 × 4	ReLU	22 × 22 × 48	0.2
池化 2	2 × 2	Max	11 × 11 × 48	-
全连接 1	-	ReLU	1024 × 1	0.5
全连接 2	-	ReLU	1024 × 1	0.5
输出	-	SoftMax	2 × 1	-

表 3 用于病理图像分割的 CNN3 体系结构  
Table 3 The CNN3 architecture for pathology image segmentation

网络层	滤波器尺寸	激活	输出尺寸	随机丢弃率
输入	-	-	51 × 51 × 3	-
卷积 1	4 × 4	ReLU	48 × 48 × 25	0.1
池化 1	2 × 2	Max	24 × 24 × 25	-
卷积 2	5 × 5	ReLU	20 × 20 × 50	0.2
池化 2	2 × 2	Max	10 × 10 × 50	-
卷积 3	6 × 6	ReLU	5 × 5 × 80	0.25
池化 3	2 × 2	Max	3 × 3 × 80	-
全连接 1	-	ReLU	1024 × 1	0.5
全连接 2	-	ReLU	1024 × 1	0.5
输出	-	SoftMax	3 × 1	-

图 11 显示了通过使用 KIRC 数据集对比细胞核分割算法的可视化结果. 图中第 1, 5、2, 6、3, 7 和 4, 8 行分别是原始图像、人工分割结果、本文方

法分割结果以及 U-Net 网络分割结果. 提出的

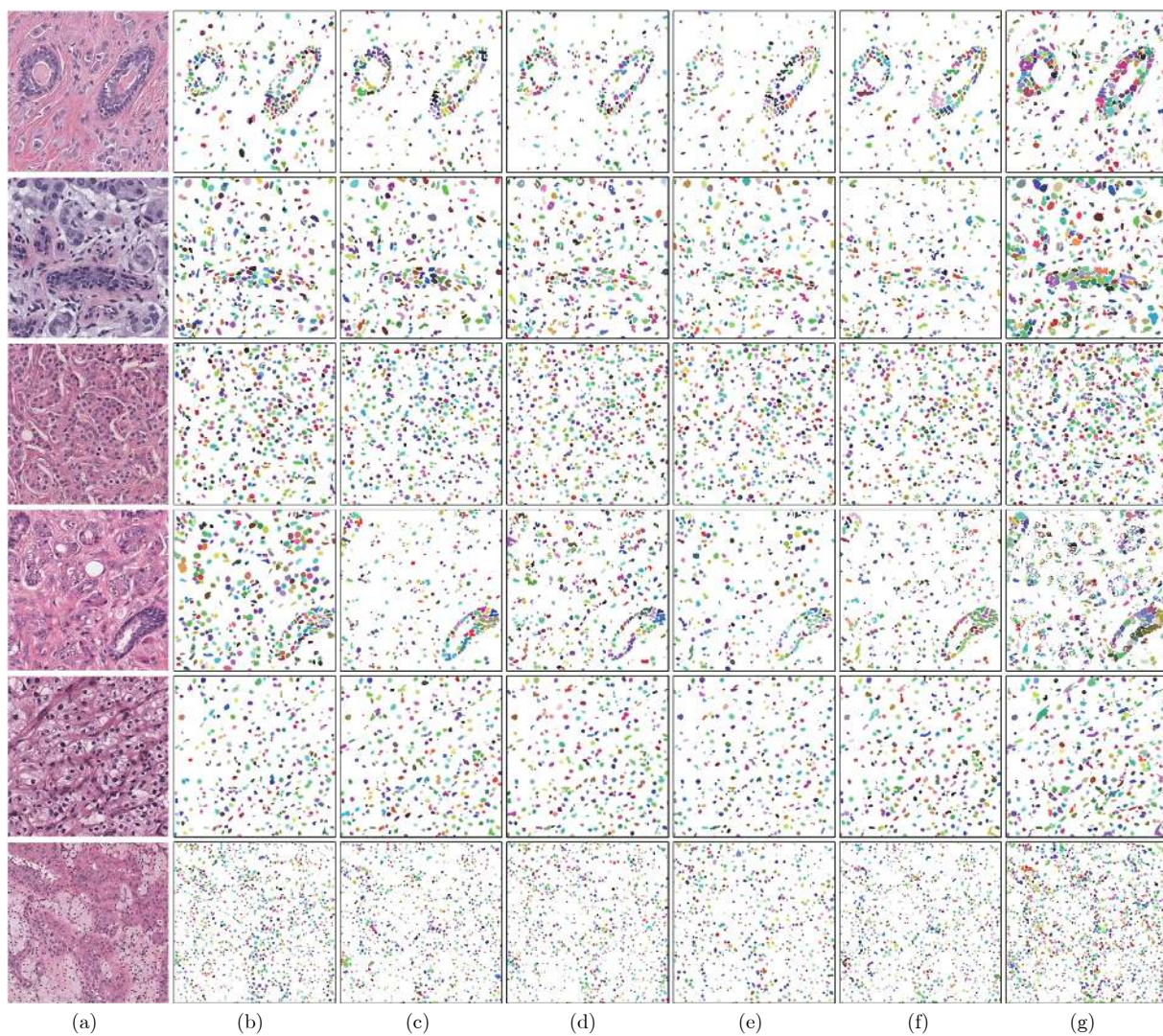


图 8 Kumar 病理细胞核相同器官图像的分割比较

Fig.8 Comparative segmentation results on the Kumar same-organs testing pathology images



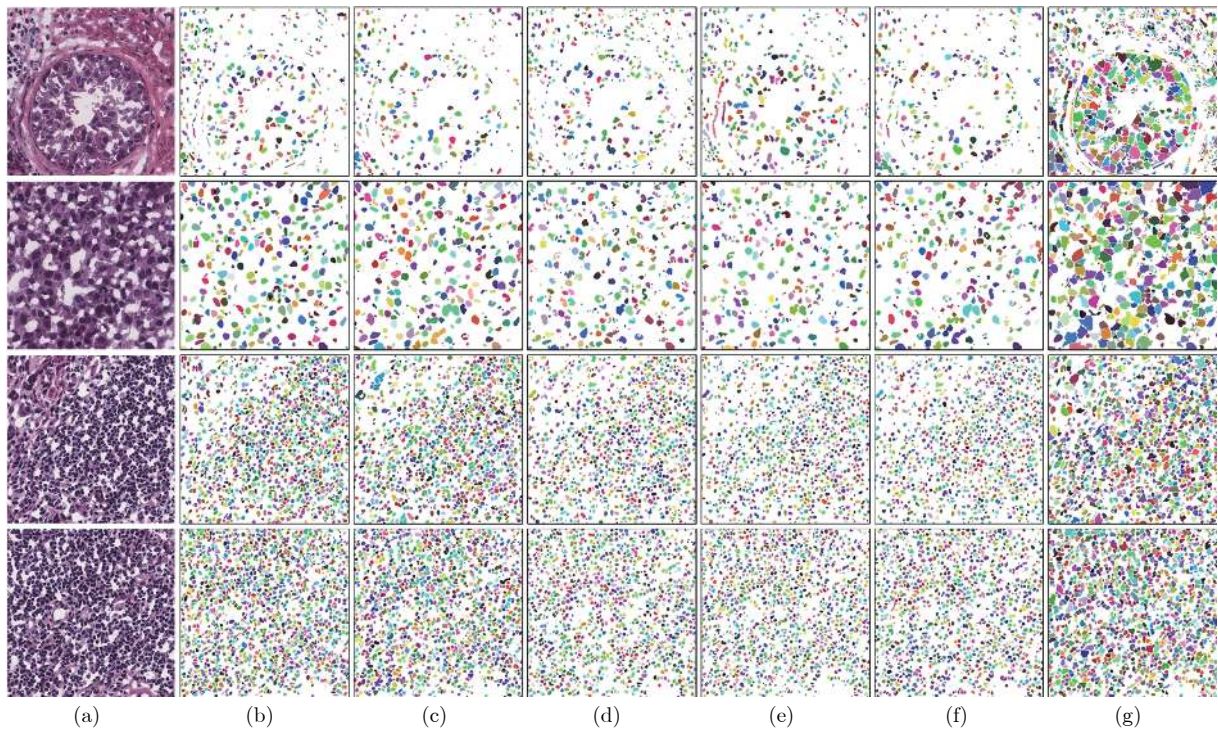


图 9 Kumar 病理细胞核不同器官图像的分割比较

Fig.9 Comparative segmentation results on the Kumar different-organs testing pathology images

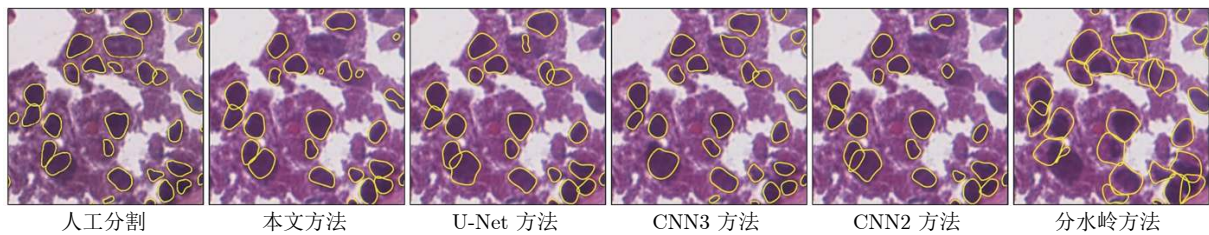


图 10 不同算法在具有重叠、形状及大小不规则的病理图像上的轮廓可视化分割结果图

Fig.10 Visualizing segmented contours of different algorithms on the pathology image with shape and color variations and overlapping regions

CscDTE 模型提供了一种定义复杂形式假设  $\varphi(\cdot)$  的级联方式, 并使用其最优参数拟合病理数据. 本文方法使得模型在训练过程中能够获取足够多的上下文信息. 如图 11 所示, 相比 U-Net, CscDTE 模型驱动的拟合重构能够起到信号正则化的作用. 因此, 即使当细胞核出现密集重叠以及形状、大小等变化时, 提出的分割方法也仍然有效.

### 3.3 定量分析

为了定量评估分割算法的性能, 本文分别使用 Jaccard 相似性系数、精确率、F1 分数和平均边缘距离 (ABD<sup>1</sup>) 三个指标来比较算法输出的掩模  $P$

<sup>1</sup>ABD 是基于像素级边缘的评价方法, 表示自动分割与人工分割目标边缘间的平均距离, 且漏分割和误分割会造成 ABD 值过大.

与 Ground Truth 掩模  $R$ , 计算如下

$$\text{Jaccard} = \frac{|P \cap R|}{|P \cup R|}, \text{Precision} = \frac{|P \cap R|}{|P|}, \text{F1} = \frac{2|P \cap R|}{|P| + |R|},$$

$$\text{ABD} = \frac{1}{2} \left( \frac{\sum_{x \in PB} d(x, RB)}{|PB|} + \frac{\sum_{u \in RB} d(u, PB)}{|RB|} \right) \quad (7)$$

其中,  $PB$  和  $RB$  分别对应  $P$  和  $R$  的边缘. 表 4 和表 5 给出了在 TCGA WSIs 病理数据集上不同算法实现的定量分割结果. 可以看出, 相比其他方法, 提出的 CscDTE 模型在不同器官图像和整个数据集上对三个指标值的改进明显. 由于本文方法更加有效地级联低级图像表观特征和高级上下文特征,

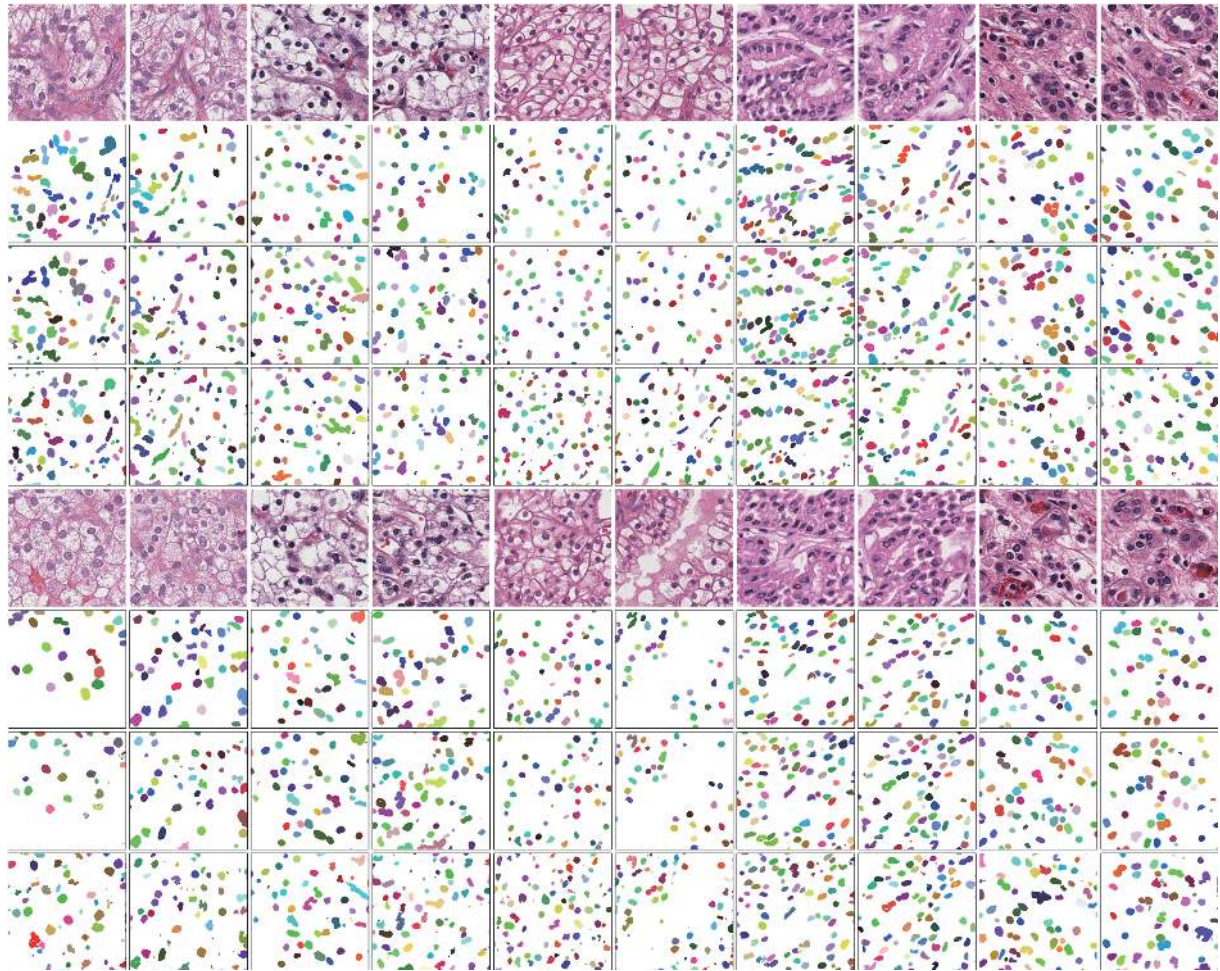


图 11 本文方法与 U-Net 方法在 KIRC 数据集上的分割比较  
 Fig.11 Comparative segmentation using ours and U-Net on the KIRC dataset

表 4 本文方法与其他方法在 Kumar 病理数据集上的分割性能 (均值) 比较  
 Table 4 Performance comparison of other methods with our CscDTE (For each metric, the mean are listed)

器官	图像	Jaccard相似性系数					精确率					F1分数					平均边缘距离				
		本文	U-Net	CNN3	CNN2	分水岭	本文	U-Net	CNN3	CNN2	分水岭	本文	U-Net	CNN3	CNN2	分水岭	本文	U-Net	CNN3	CNN2	分水岭
乳腺	1	<b>0.594</b>	0.539	0.570	0.569	0.545	0.810	<b>0.832</b>	0.756	0.764	0.591	<b>0.745</b>	0.700	0.726	0.725	0.706	<b>14.90</b>	42.78	29.68	64.46	42.36
	2	<b>0.675</b>	0.613	0.603	0.448	0.592	0.832	0.803	0.844	<b>0.853</b>	0.625	<b>0.806</b>	0.760	0.752	0.619	0.744	<b>12.99</b>	13.12	14.34	62.27	18.49
前列腺	1	0.646	0.634	<b>0.656</b>	0.625	0.585	0.867	0.858	0.858	<b>0.871</b>	0.679	0.785	0.776	<b>0.792</b>	0.769	0.739	12.56	15.87	11.41	<b>10.22</b>	24.40
	2	0.342	<b>0.485</b>	0.459	0.387	0.409	0.834	0.827	<b>0.906</b>	0.832	0.603	0.510	<b>0.653</b>	0.629	0.558	0.581	88.69	<b>34.51</b>	40.34	56.01	37.81
肾脏	1	<b>0.692</b>	0.673	0.641	0.691	0.570	0.752	0.782	<b>0.848</b>	0.755	0.606	<b>0.818</b>	0.804	0.781	0.817	0.726	<b>12.14</b>	15.25	19.47	12.24	30.03
	2	0.513	<b>0.539</b>	0.524	0.463	0.512	0.765	<b>0.789</b>	0.763	0.681	0.570	0.678	<b>0.701</b>	0.687	0.633	0.677	16.35	21.48	12.73	21.38	<b>9.48</b>
膀胱	1	<b>0.603</b>	0.551	0.581	0.550	0.325	<b>0.846</b>	0.756	0.720	0.795	0.330	<b>0.752</b>	0.711	0.735	0.709	0.490	41.97	86.13	<b>22.91</b>	52.33	79.61
	2	<b>0.759</b>	0.629	0.580	0.656	0.508	0.867	0.898	<b>0.926</b>	0.909	0.516	<b>0.863</b>	0.772	0.734	0.792	0.674	<b>6.58</b>	14.32	17.09	19.03	22.04
胃	1	<b>0.781</b>	0.645	0.622	0.577	0.647	0.881	<b>0.920</b>	0.824	0.913	0.692	<b>0.877</b>	0.784	0.767	0.732	0.785	<b>1.99</b>	7.44	8.49	9.16	3.67
	2	<b>0.785</b>	0.641	0.714	0.623	0.679	0.891	0.931	<b>0.948</b>	0.910	0.736	<b>0.880</b>	0.781	0.833	0.767	0.809	<b>1.09</b>	6.82	5.41	6.75	1.76
相同器官		0.577	<b>0.581</b>	0.576	0.531	0.536	0.810	0.815	<b>0.829</b>	0.793	0.612	0.724	<b>0.732</b>	0.728	0.687	0.696	26.27	23.84	<b>21.33</b>	37.76	27.10
不同器官		<b>0.732</b>	0.617	0.624	0.602	0.540	0.871	0.876	0.855	<b>0.882</b>	0.569	<b>0.843</b>	0.762	0.767	0.750	0.690	<b>12.91</b>	28.68	13.48	21.82	26.77
整体		<b>0.639</b>	0.595	0.595	0.559	0.537	0.835	<b>0.840</b>	0.839	0.828	0.595	<b>0.771</b>	0.744	0.744	0.712	0.693	20.93	25.77	<b>18.19</b>	31.39	26.97

表 5 本文方法与其他方法在 Kumar 病理数据集上的分割性能 (标准差) 比较

Table 5 Performance comparison of other methods with our CscDTE (For each metric, the standard deviation are listed)

器官	Jaccard相似性系数					精确率					F1分数					平均边缘距离				
	本文	U-Net	CNN3	CNN2	分水岭	本文	U-Net	CNN3	CNN2	分水岭	本文	U-Net	CNN3	CNN2	分水岭	本文	U-Net	CNN3	CNN2	分水岭
乳腺	0.057	0.052	0.023	0.086	0.033	0.016	0.021	0.062	0.063	0.024	0.043	0.042	0.018	0.075	0.027	1.35	20.97	10.85	1.55	16.88
前列腺	0.215	0.105	0.139	0.168	0.125	0.023	0.022	0.034	0.028	0.054	0.195	0.087	0.115	0.149	0.112	53.83	13.18	20.46	32.38	9.48
肾脏	0.127	0.095	0.083	0.161	0.041	0.009	0.005	0.060	0.052	0.026	0.099	0.073	0.067	0.130	0.035	2.98	4.41	4.77	6.46	14.53
膀胱	0.110	0.055	0.001	0.075	0.129	0.015	0.100	0.146	0.081	0.132	0.079	0.043	0.001	0.059	0.130	25.02	50.78	4.12	23.55	40.71
胃	0.003	0.003	0.065	0.033	0.023	0.007	0.008	0.088	0.002	0.031	0.002	0.002	0.047	0.025	0.017	0.64	0.44	2.18	1.70	1.35
相同器官	0.132	0.071	0.075	0.117	0.069	0.044	0.029	0.058	0.072	0.037	0.116	0.057	0.062	0.099	0.061	30.62	12.09	11.45	25.79	12.23
不同器官	0.087	0.044	0.063	0.047	0.161	0.020	0.081	0.105	0.058	0.185	0.061	0.034	0.046	0.037	0.145	19.52	38.45	8.00	21.02	36.39
整体	0.136	0.062	0.071	0.098	0.106	0.047	0.061	0.076	0.078	0.113	0.112	0.049	0.057	0.084	0.096	26.37	24.09	10.52	24.18	22.91

并结合集成决策树学习模块的使用, 提出方法在 Jaccard 相似性系数和 F1 分数整体值上分别超出表现最好的 U-Net 和 CNN3 方法 4.4% 和 2.7%, 同时在平均边缘距离值上相比 U-Net 减少了大约 4.8 个像素. 另外, 定量分割结果也反映出 U-Net 和 CNN3 在相同器官和不同器官测试方面均强于 CNN2, 而在前列腺病理细胞核图像上, 所提出方法略逊于 U-Net 和 CNN3. 需要注意的是: 1) 虽然标记控制的分水岭分割算法在相同器官测试上, 超过 CNN2 的性能, 但它们的分割精度较差, 造成了大量的误分割结果; 2) 本文算法无需非线性激活和后向传播计算, 虽然性能较之深度神经网络有很强的竞争力, 但由于端到端训练的简易性, 因而当染质稀疏较为严重时, 分割结果会有较大差异. 为了便于模型的显著性差异分析, 本文进一步采用双样本  $t$ -检验统计指标<sup>[38]</sup>来对不同方法进行比较. 基于 F1 分数, 虽然提出的 CscDTE 模型在整个 Kumar 数据集上与 U-Net 模型、CNN3 模型、CNN2 模型、分水岭模型之间没有显著性差异发生的概率 ( $p$  值) 分别为 0.491、0.493、0.197、0.110, 但在不同器官图像上却有本质区别, 概率分别为 0.001、0.141、0.018、0.021. 表 6 给出了本文方法和 U-Net 方法在 KIRC 测试图像上的各指标均值和标准差. 可以看出, CscDTE 模型的分割性能明显优于表现最好的 U-Net 模型, 且具有更好的可靠性.

对于逐像素分割任务而言, 学习模型的参数规模是另一个重要的性能评估指标. 学习模型中带参

数的层主要有卷积层、批归一化层以及全连接层等, 而激活函数、池化层和上采样层旨在提供一种非线性变换, 无需学习. 本文采取秩-1 可分离卷积核代替卷积滤波器, 且无需全连接层特征映射, 参数规模较之传统深度学习方法较小, 如表 7 所示.

#### 4 结论及后续工作

本文提出了一个级联稀疏卷积与决策树集成学习 (CscDTE) 的细胞核分割模型. 常用的分水岭分割方法和深度学习网络对算法初始化和病理图像数据的多样性较为敏感. 与之相比, CscDTE 模型具有四个方面的优势:

- 1) 为了实现鲁棒特征的提取和 CscDTE 模型算法的加速, 基于秩-1 张量分解学习机制, CscDTE 模型在各层引入稀疏可分离卷积模块, 分层抽取细胞核的多尺度方向分布式抽象特征;
- 2) 为了提升模型的分割性能, 充分考虑像素间的关系, CscDTE 模型使用级联卷积特征作为样本输入, 迭代有监督地训练集成决策树模块直到收敛;
- 3) 为了增强集成决策树模块对数据复杂性和可变性的自适应能力, CscDTE 模型采取随机采样、树剪枝以及正则化回归机制提升逐像素分布式特征集成回归分类能力;
- 4) 不同器官和疾病状态的复杂病理数据集上的大量试验表明了所提出的 CscDTE 方法相比于已有最先进技术的优越性.

表 6 本文方法与 U-Net 在 KIRC 数据集上的分割性能 (均值, 标准差) 比较

Table 6 Performance comparison of U-Net with our CscDTE (For each metric, the mean and standard deviation are listed)

KIRC	Jaccard系数		精确率		F1分数		平均边缘距离	
	均值	标准差	均值	标准差	均值	标准差	均值	标准差
本文方法	<b>0.735</b>	<b>0.054</b>	<b>0.832</b>	<b>0.050</b>	<b>0.846</b>	<b>0.037</b>	<b>22.66</b>	<b>9.51</b>
U-Net	0.642	0.064	0.742	0.093	0.780	0.049	57.78	23.11

表 7 不同算法在病理数据集上的参数规模  
Table 7 Number of parameters of different algorithms on the pathology dataset

方法	本文算法	U-Net	CNN3	CNN2
参数规模	<b>184KB</b>	1.75MB	1.97MB	7.04MB

在现有框架的基础上, 将利用更先进的预测模型对集成决策树学习模块进行替换, 以实现更好的分割性能; 将结合语义分割, 在应用层面对 CscDTE 其进行拓展; 将探索管状显微结构的提取与重构, 例如血管、神经突等, 并拓展至三维医学图像应用。

## References

- Webster J D, Dunstan R W. Whole-slide imaging and automated image analysis: Considerations and opportunities in the practice of pathology. *Veterinary Pathology*, 2014, **51**(1): 211–223
- Yu K H, Zhang C, Berry G J, Altman R B, Re C, Rubin D L, et al. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nature Communications*, 2016, **7**(12474): 1–10
- Zheng Qiang, Dong En-Qing. Narrow band active contour model for local segmentation of medical and texture images. *Acta Automatica Sinica*, 2013, **39**(1): 21–30  
(郑强, 董恩清. 窄带主动轮廓模型及在医学和纹理图像局部分割中的应用. *自动化学报*, 2013, **39**(1): 21–30)
- Zhang Fan, Zhang Xin-Hong. Distance regularized level set image segmentation algorithm by means of dislocation theory. *Acta Automatica Sinica*, 2018, **44**(5): 943–952  
(张帆, 张新红. 基于位错理论的距离正则化水平集图像分割算法. *自动化学报*, 2018, **44**(5): 943–952)
- Yan P, Zhou X, Shah M, Wong S T C. Automatic segmentation of high-throughput RNAi fluorescent cellular images. *IEEE Transactions on Information Technology in Biomedicine*, 2008, **12**(1): 109–117
- Al-Kofahi Y, Lassoued W, Lee W, Roysam B. Improved automatic detection and segmentation of cell nuclei in histopathology images. *IEEE Transactions on Biomedical Engineering*, 2010, **57**(4): 841–852
- Arslan S, Ersahin T, Cetin-Atalay R, Gunduz-Demir C. Attributed relational graphs for cell nucleus segmentation in fluorescence microscopy images. *IEEE Transactions on Medical Imaging*, 2013, **32**(6): 1121–1131
- Koyuncu C F, Arslan S, Durmaz I, Cetin-Atalay R, Gunduz-Demir C. Smart markers for watershed-based cell segmentation. *PLoS One*, 2012, **7**(11): e48664
- Veta M, Huisman A, Viergever M A, van Diest P J, Pluim J P W. Automatic nuclei segmentation in H&E stained breast cancer histopathology images. *PLoS One*, 2013, **8**(7): e70221
- Dong F, Irshad H, Oh E Y, Lerwill M F, Brachtel E F, Jones N C, et al. Computational pathology to discriminate benign from malignant intraductal proliferations of the breast. *PLoS One*, 2014, **9**(12): e114885
- Zhou Y, Chang H, Barner K E, Parvin B. Nuclei segmentation via sparsity constrained convolutional regression. In: Proceedings of the 2015 IEEE International Symposium on Biomedical Imaging. New York, NY, USA: IEEE, 2015. 1284–1287
- Zhou Z-H, Feng J. Deep forest: Towards an alternative to deep neural networks. In: Proceedings of the 2017 International Joint Conference on Artificial Intelligence. Melbourne, Australia: Morgan Kaufmann, 2017. 3553–3559
- Sironi A, Tekin B, Rigamonti R, Lepetit V, Fua P. Learning separable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, **37**(1): 94–106
- Sironi A, Lepetit V, Fua P. Multiscale centerline detection by learning a scale-space distance transform. In: Proceedings of the 2014 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE, 2014. 2697–2704
- Sironi A, Türetken E, Lepetit V, Fua P. Multiscale centerline detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, **38**(7): 1327–1341
- Song J, Xiao L, Lian Z C. Contour-seed pairs learning-based framework for simultaneously detecting and segmenting various overlapping cells/nuclei in microscopy images. *IEEE Transactions on Image Processing*, 2018, **27**(12): 5759–5774
- Song J, Xiao L, Molaei M, Lian Z C. Multi-layer boosting sparse convolutional model for generalized nuclear segmentation from histopathology images. *Knowledge-Based Systems*, 2019, **176**: 40–53
- Sulam J, Pappas V, Romano Y, Elad M. Multilayer convolutional sparse modeling: pursuit and dictionary learning. *IEEE Transactions on Signal Processing*, 2018, **66**(15): 4090–4104
- Kolda T G, Bader B W. Tensor decompositions and applications. *SIAM Review*, 2009, **51**(3): 455–500
- Wright J, Ma Y, Mairal J, Sapiro G, Huang T S, Yan S. Sparse representation for computer vision and pattern recognition. *Proceedings of the IEEE*, 2010, **98**(6): 1031–1044
- Rigamonti R, Brown M, Lepetit V. Are sparse representations really relevant for image classification. In: Proceedings of the 2011 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Providence, RI, USA: IEEE, 2011. 1545–1552
- Bach F, Jenatton R, Mairal J, Obozienski G. Optimization with sparsity-inducing penalties. *Foundations and Trends in Machine Learning*, 2012, **4**(1): 1–106
- 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 Transactions on Medical Imaging*, 2016, **35**(1): 119–130
- Xing F, Xie Y, Yang L. An automatic learning-based framework for robust nucleus segmentation. *IEEE Transactions on Medical Imaging*, 2016, **5**(2): 550–566
- 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 Transactions on Medical Imaging*, 2017, **36**(7): 1550–1560
- Zhang L, Sonka M, Lu L, Summers R M, Yao J. Combining fully convolutional networks and graph-based approach for automated segmentation of cervical cell nuclei. In: Proceedings of the 2017 IEEE International Symposium on Biomedical Imaging. Melbourne, VIC, Australia: IEEE, 2017. 406–409
- Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: Proceedings of

- Medical Image Computing and Computer-assisted Intervention: MICCAI International Conference on Medical Image Computing and Computer-assisted Intervention. Berlin Heidelberg: Springer, 2015. 234–241
- 28 Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks. In: Proceedings of Advances in Neural Information Processing Systems 25: Proceedings of the 2012 Conference. Cambridge: MIT Press, 2012. 1097–1105
- 29 Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, **39**(4): 640–651
- 30 Acar E, Dunlavy D M, Kolda T G. A scalable optimization approach for fitting canonical tensor decompositions. *Journal of Chemometrics*, 2011, **25**(2): 67–86
- 31 Lin M, Chen Q, Yan S. Network In Network. In: Proceedings of the 2014 International Conference on Learning Representations. Banff, AB, Canada: OpenReview, 2014. 1–10
- 32 Hastie T, Tibshirani R, Friedman J. *The elements of statistical learning*. New York, NY, USA, Berlin Heidelberg: Springer, 2001
- 33 Rokach L. Ensemble-based classifiers. *Artificial Intelligence Review*, 2010, **33**(1-2): 1–39
- 34 Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, CA, USA: ACM, 2016. 785–794
- 35 Tu Z, Bai X. Auto-context and its application to high-level vision tasks and 3D brain image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(10): 1744–1757
- 36 Irshad H, Montaser-Kouhsari L, Waltz G, Bucur O, Nowak J A, Dong F, Knoblauch N W, Beck A H. Crowdsourcing image annotation for nucleus detection and segmentation in computational pathology: evaluating experts, automated methods, and the crowd. In: Proceedings of Pacific Symposium on Biocomputing, Fairmont Orchid, Hawaii, USA: PSB, 2014. 294–305
- 37 Mairal J, Bach F, Ponce J. Task-driven dictionary learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, **34**(4): 791–804
- 38 McDonald J H. *Handbook of biological statistics*. Baltimore, MD: Sparky House, 2014.



宋杰 博士, 南京邮电大学自动化学院、人工智能学院讲师. 主要研究方向为生物医学图像处理, 机器学习与模式识别.

E-mail: j.song0041@gmail.com

(**SONG Jie** Ph.D., lecturer at the College of Automation & College of Artificial Intelligence, Nanjing University of Posts and Telecommunications. His research interest covers biomedical image processing, machine learning, and pattern recognition.)



肖亮 博士, 南京理工大学计算机科学与工程学院教授. 主要研究方向为信号处理, 计算机视觉, 生物医学图像处理, 机器学习与模式识别. 本文通信作者.

E-mail: xiaoliang@mail.njust.edu.cn

(**XIAO Liang** Ph.D., professor at the School of Computer Science and Engineering, Nanjing University of Science and Technology. His research interest covers signal processing, computer vision, biomedical image processing, machine learning, and pattern recognition. Corresponding author of this paper.)



练智超 博士, 南京理工大学计算机科学与工程学院副教授. 主要研究方向为医学图像处理, 神经影像, 机器学习与模式识别.

E-mail: newlzcts@gmail.com

(**LIAN Zhi-Chao** Ph.D., associate professor at the School of Computer Science and Engineering, Nanjing University of Science and Technology. His research interest covers medical image processing, neuro-imaging, machine learning, and pattern recognition.)